SIAM J. COMPUT.
Vol. 30, No. 5, pp. 1427–1442

A NEAR-TIGHT LOWER BOUND ON THE TIME COMPLEXITY OF DISTRIBUTED MINIMUM-WEIGHT SPANNING TREE CONSTRUCTION*

DAVID PELEG† AND VITALY RUBINOVICH‡

Abstract. This paper presents a lower bound of $\Omega(D + \sqrt{n}/\log n)$ on the time required for the distributed construction of a minimum-weight spanning tree (MST) in weighted n-vertex networks of diameter $D = \Omega(\log n)$, in the bounded message model. This establishes the asymptotic near-optimality of existing time-efficient distributed algorithms for the problem, whose complexity is $O(D + \sqrt{n}\log^* n)$.

Key words. distributed algorithm, minimum weight spanning tree, lower bound

AMS subject classifications. 05C85, 68Q22, 68Q25

PII. S0097539700369740

1. Introduction.

The study of distributed algorithms for minimum-weight spanning tree (MST) construction was initiated by the pioneering work of Gallager, Humblet, and Spira [GHS83], which introduced a basic distributed technique for the problem and presented a message-optimal algorithm with time complexity $O(n \log n)$ on an n-vertex network. This result was later improved to a message-optimal algorithm with time complexity $O(n)$ by Awerbuch [A87].

However, for many natural distributed network problems, the parameter controlling the time complexity is not the number of vertices but rather the network's diameter D, namely, the maximum distance between any two vertices (measured in hops). This holds, for example, for leader election and related problems [P90].

It is easy to verify that $\Omega(D)$ time is required for distributed MST construction in the worst case. More formally, for every two integers $n \geq 2$ and $1 \leq D \leq \lfloor n/2 \rfloor$ there exist weighted n-vertex networks of diameter D (say, based on a $2D$-vertex ring with $n - 2D$ vertices attached to it as leaves) on which any distributed MST algorithm will require at least D time.

Hence, a natural question is whether $O(D)$-time algorithms exist for distributed MST construction as well. More generally, the problem of devising $o(n)$ (though possibly not message-optimal) distributed algorithms for MST construction was introduced in [GKP98].

Clearly, in the extreme model allowing the transmission of an unbounded-size message on a link in a single time unit (cf. [L92]), the problem can be trivially solved in time $O(D)$ by collecting the entire graph's topology and all the edge weights into a central vertex, computing an MST locally and broadcasting the result throughout the network. The problem thus becomes interesting in the more realistic, and rather

*Received by the editors March 27, 2000; accepted for publication (in revised form) August 1, 2000; published electronically November 8, 2000. An extended abstract has appeared in *A near-tight lower bound on the time complexity of distributed MST construction*, in Proceedings of the 40th IEEE Symposium on Foundations of Comoputer Science, IEEE Computer Society Press, Los Alamitos, CA, 1999, pp. 253–261.

http://www.siam.org/journals/sicomp/30-5/36974.html

†Department of Computer Science and Applied Mathematics, The Weizmann Institute of Science, Rehovot, 76100 Israel (peleg@wisdom.weizmann.ac.il). The work of this author was supported in part by a grant from the Israel Ministry of Science and Art.

‡Department of Mathematics and Computer Science, Bar Ilan University, Ramat Gan, Israel (rabinov@macs.biu.ac.il).

common, B-*bounded-message* model (henceforth referred to simply as the B model), in which message size is bounded by some value B (usually taken to be either constant or $O(\log n)$), and a vertex may send at most one message on each edge at each time unit.

The algorithm presented in [GKP98] for distributed MST construction in this model (with $B = O(\log n)$-bit messages) has time complexity $O(D + n^{\epsilon} \log^* n)$ for $\epsilon = \ln 6 / \ln 3 \approx 0.613$. This was later improved to $O(D + \sqrt{n} \log^* n)$ in [KP98]. Similar bounds were recently obtained by us using other algorithmic methods, but none of those methods were able to break the $\sqrt{n}$-time barrier, indicating that distributed MST might be harder than other distributed network problems such as leader election or breadth-first search (BFS) tree construction.

It is important to mention that the algorithms of [GHS83, A87, GKP98, KP98] discussed above were analyzed under the (natural) assumption that the weight of each edge can be represented as an integer small enough to be included in a single message. This assumption is adopted in the current paper.

The current paper concerns establishing the asymptotic near-optimality of the algorithm of [KP98], by showing that $\tilde{\Omega}(\sqrt{n})$ is a lower bound[1] as well, even on low diameter networks. Specifically, for any integers $K, m \geq 2$, we construct a family of $O(m^{2K})$-vertex networks of diameter $D = O(Km)$ for which $\Omega(m^K/(BK))$ time is required for constructing a minimum spanning tree in the B model. Fixing some positive integer $m \geq 2$, we get that for every integer $n \geq 1$ there exists a family of n-vertex networks of diameter $\Theta(\log n)$ for which MST construction requires $\Omega(\sqrt{n}/(B \log n))$ time in the B model.

While it is not clear that the $\Omega(\log n)$ limitation on the diameters for which the lower bound holds is essential, *some* limitation must apparently exist. This follows from the observation that the n-vertex complete graph ($D = 1$) admits a simple $O(\log n)$ time distributed MST construction algorithm.

Towards proving the lower bound on distributed MST construction, we first establish a lower bound on the time complexity of a problem referred to as the *mailing problem*, which can be informally stated as follows. Given a particular type of graph named F_m^K, for integers $m, K \geq 2$, and two vertices s and r in it, it is required to deliver an m^K-bit string $\mathcal{X}$ generated in s to r. The graph F_m^K has $n = O(m^{2K})$ vertices and diameter $O(Km)$, yet we show that the time required for mailing from s to r on F_m^K in the B model is considerably larger than the diameter, namely, $\Omega(m^K/(BK)) = \Omega(\sqrt{n}/(BK))$.

The rest of the paper is organized as follows. First, a definition of the model and the mailing problem is given in section 2. Section 3 handles the mailing problem for the case of $K = 2$. It defines the graphs F_m^2, having diameter $D = O(m) = O(n^{1/4})$ and shows a lower bound of $O(m^2) = O(\sqrt{n})$ on the time complexity of the mailing problem for m^2-bit strings. This result is then used in section 4 to prove that the same lower bound applies also to the time complexity of the MST problem on weighted versions of the graphs F_m^2. The next two sections extend these two results, respectively, to graphs F_m^K, $K \geq 3$ with diameters down to $O(\log n)$. Finally, section 7 discusses some open problems.

2. Preliminaries.

2.1. The model. A point-to-point communication network is modeled as an undirected graph $G(V, E)$, where the vertices of V represent the network processors

[1] $\tilde{\Omega}$ is a relaxed variant of the Ω notation that ignores polylog factors.

and the edges of E represent the communication links connecting them. Vertices are allowed to have unique identifiers. The vertices do not know the topology or the edge weights of the entire network, but they may know the IDs of their neighbors and the weights of the corresponding edges.

A weight function $\omega : E \to \mathbb{R}^+$ associated with the graph assigns a nonnegative integer weight $\omega(e)$ to each edge $e = (u, v) \in E$. The weight $\omega(e)$ is known to the adjacent vertices, u and v. The vertices can communicate only by sending and receiving messages over the communication links. Communication is carried out in a synchronous manner; i.e., all the vertices are driven by a global clock. Messages are sent at the beginning of each round and are received at the end of the round. (Clearly, our lower bounds hold for asynchronous networks as well.) At most one B-bit message can be sent on each link in one direction on every round. It is assumed that B is large enough to allow the transmission of an edge weight in a single message. The model also allows vertices to detect the absence of a message on a link at a given round, which can be used to convey information. Hence at each communication round, a link can be at one of $2^B + 1$ possible states, i.e., it can either transmit any of 2^B possible messages or remain silent.

The length of a path p in the network is the number of edges it contains. The *distance* between two vertices u and v is defined as the length of the shortest path connecting them in G. The *diameter* of G, denoted D, is the maximum distance between any two vertices of G.

2.2. The mailing problem. The mailing problem is defined in the following situation. We are given a graph G with two distinguished vertices s and r, referred to as the *sender* and the *receiver*, respectively. Both the sender s and the receiver r store b boolean variables each, $X_1^s, \ldots, X_b^s$ and $X_1^r, \ldots, X_b^r$, respectively, for some integer $b \geq 1$. An instance of the problem consists of an initial assignment $\mathcal{X} = \{x_i \mid 1 \leq i \leq b\}$, where $x_i \in \{0, 1\}$, to the variables of s, such that $X_i^s = x_i$. Given such an instance, the mailing problem requires s to deliver the string $\mathcal{X}$ to the receiver r, i.e., upon termination, the variables of r should contain the output $X_i^r = x_i$ for every $1 \leq i \leq b$. Henceforth, we refer to this problem as $\texttt{Mail}(G, s, r, b)$. Throughout sections 3 and 4, we consider this problem on graphs F_m^2 with $b = m^2$ for some integer $m \geq 2$. In sections 5 and 6, we deal with the problem on graphs F_m^K with $b = m^K$ for $K \geq 3$.

2.3. The distributed MST problem. Formally, the *minimum spanning tree (MST)* problem can be stated as follows. Given a graph $G(V, E)$ and a weight function ω on the edges, it is required to find a spanning tree $MST(G) \subseteq E$ whose total weight, $\omega(MST(G)) = \sum_{e \in MST(G)} \omega(e)$, is minimal. In the distributed model, the input and output of the MST problem are represented as follows. Each vertex knows the ID's of its closest neighbors and the weights of the corresponding edges. A degree-d vertex $v \in V$ with neighbors $u_1, \ldots, u_d$ has d *weight variables* $W_1^v, \ldots, W_d^v$, with W_i^v containing the weight of the edge connecting v to u_i, i.e., $W_i^v = \omega(v, u_i)$. The output of the MST problem at each vertex v is an assignment to the (boolean) output variables $Y_1^v, \ldots, Y_d^v$, assigning

$$Y_i^v = \begin{cases} 1, & (u_i, v) \in MST(G), \\ 0 & \text{otherwise.} \end{cases}$$

3. A lower bound for the mailing problem on F_m^2.

3.1. The graphs F_m^2. Let us now define the collection of graphs denoted F_m^2 for $m \geq 2$. The two basic units in the construction are the *ordinary path* $\mathcal{P}$ on m^2+1 vertices,

$$V(\mathcal{P}) = \{v_0, \ldots, v_{m^2}\} \quad \text{and} \quad E(\mathcal{P}) = \{(v_i, v_{i+1}) \mid 0 \leq i \leq m^2 - 1\},$$

and the *highway* $\mathcal{H}$ on $m+1$ vertices,

$$V(\mathcal{H}) = \{h_{im} \mid 0 \leq i \leq m\} \quad \text{and} \quad E(\mathcal{H}) = \{(h_{im}, h_{(i+1)m}) \mid 0 \leq i \leq m-1\}.$$

Each highway vertex h_{im} is connected to the corresponding path vertex v_{im} by a *spoke* edge (h_{im}, v_{im}), as in Figure 1.

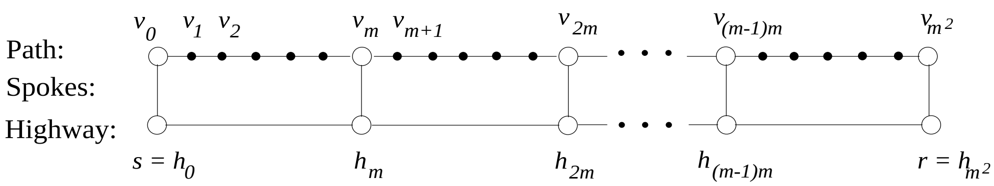

FIG. 1. *The connections between the path and the highway.*

The graph F_m^2 is constructed by taking m^2 copies of the ordinary path $\mathcal{P}$, denoted $\mathcal{P}^1, \ldots, \mathcal{P}^{m^2}$, and connecting all of them to the *same* highway $\mathcal{H}$. The vertex h_0 is the intended sender s, and the vertex h_{m^2} is the intended receiver r. (See Figure 2.)

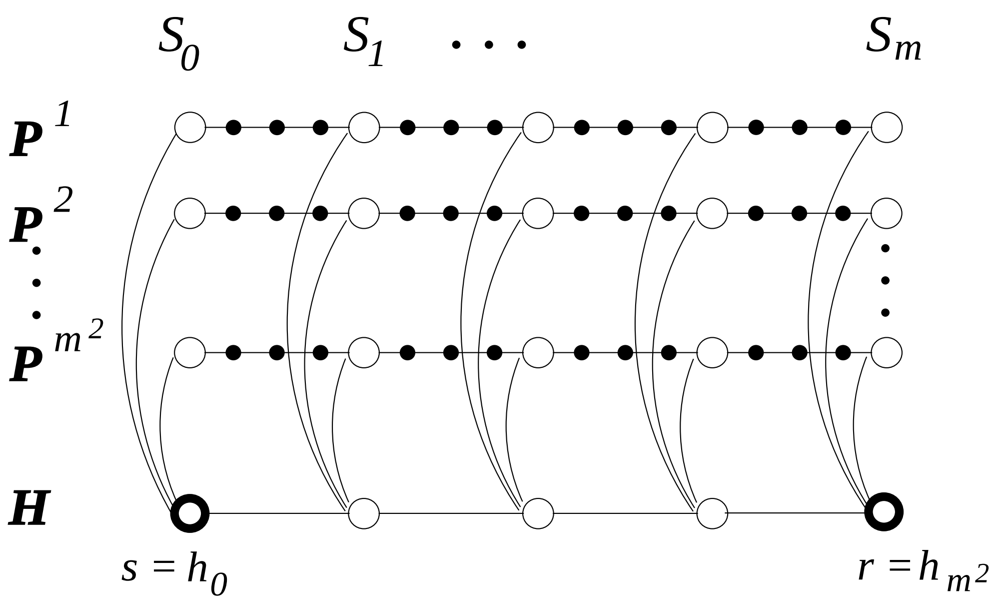

FIG. 2. *The graph F_m^2.*

Visualizing the graph F_m^2 as organized in a cylindrical shape, the spoke edges can be grouped into $m+1$ *stars* S_i, $0 \leq i \leq m$, where each star S_i consists of the highway vertex h_{im} and the m^2 vertices $v_{im}^1, \ldots, v_{im}^{m^2}$ connected to it by spoke edges. Hence

$$V(S_i) = \{h_{im}\} \cup \{v_{im}^1, \ldots, v_{im}^{m^2}\} \quad \text{and} \quad E(S_i) = \{(v_{im}^j, h_{im}) \mid 1 \leq j \leq m^2\}.$$

The vertex and edge sets of the graph F_m^2 are thus

$$V(F_m^2) = V(\mathcal{H}) \cup \bigcup_{j=1}^{m^2} V(\mathcal{P}^j) \quad \text{and} \quad E(F_m^2) = \bigcup_{i=0}^{m} E(S_i) \cup \bigcup_{j=1}^{m^2} E(\mathcal{P}^j) \cup E(\mathcal{H}).$$

FACT 3.1. *The graph F_m^2 consists of $n = \Theta(m^4)$ vertices, and its diameter is $\Theta(m)$.*

3.2. The lower bound. We would now like to prove that solving the mailing problem on the graph F_m^2 with a $b = m^2$-bit string $\mathcal{X}$ requires $\Omega(m^2/B)$ time in the B model. Intuitively, this happens because routing the string $\mathcal{X}$ from s to r along ordinary paths would be too slow; hence our only hope is to route the string along the highway, or at least use interleaved paths, mixing highway segments with segments of ordinary paths. However, F_m^2 does not have sufficient capacity for routing all m^2 bits from s to r along such short (or "relatively short") paths.

This intuition yields a rather simple proof of the claim if we limit ourselves to a restricted class of algorithms, referred to as *explicit delivery* algorithms. These are algorithms in which the input bits are required to be delivered in an explicit way, namely, each bit x_i must be shipped from s to t along some path p_i. (Naturally, the paths of different bits may be identical or partly overlap.) However, we would like the lower bound to apply also to *arbitrary* algorithms, in which the information can be conveyed from s to r in arbitrary ways. This may include applying arbitrary functions to the bits at s and sending the resulting values, possibly modifying and "recombining" these values in intermediate nodes along the way, in a way that will allow r to extract the original bits from the messages it receives. For handling such a general class of algorithms, the proof must be formalized in a more careful way.

Let us start with an outline of the proof. Consider the set of possible states a vertex v may be in at any given stage t of the execution of a mailing algorithm on some m^2-bit input $\mathcal{X}$. (The state of a vertex consists of all its local data; hence it is affected by its input, topological knowledge, and history, namely, all incoming messages.) As the computation progresses, the tree of possible executions grows, and thus the set of possible states of v becomes larger. In particular, when the execution starts at round 0, each of the vertices is in one specific initial local state, except for the sender s, which may be in any one of 2^{m^2} states, determined by the value of the input string $\mathcal{X}$. Upon termination, the string $\mathcal{X}$ should be known to the receiver r, meaning that r should be in one of 2^{m^2} states. Our argument is based on analyzing the growth process of the sets of possible states, and showing that this process is slow, forcing the algorithm to spend at least $\Omega(m^2/B)$ time until the set of possible states of r is of size 2^{m^2}.

We now continue with a more detailed and formal proof. Consider some arbitrary algorithm $\mathcal{A}_{\mathrm{mail}}$, and let $\varphi_{\mathcal{X}}$ denote the execution of $\mathcal{A}_{\mathrm{mail}}$ on an m^2-bit input $\mathcal{X}$ in the graph F_m^2. For $1 \le i \le m$, define the *tail set* of the graph F_m^2, denoted T_i, as follows. For every $1 \le j \le m^2$, define the *tail* of the path $\mathcal{P}^j$ as

$$T_i(\mathcal{P}^j) = \{v_l^j \mid i \le l \le m^2\}.$$

Let $\beta(i)$ denote the least integer δ such that $\delta m \ge i$, and define the tail of $\mathcal{H}$ as

$$T_i(\mathcal{H}) = \{h_{jm} \mid \beta(i) \le j \le m\}.$$

Now, the tail set of F_m^2 is the union of those tails,

$$T_i = T_i(\mathcal{H}) \cup \bigcup_j T_i(\mathcal{P}^j).$$

(See Figure 3.) For $i = 0$, the definition is slightly different, letting

$$T_0 = V \setminus \{h_0\}.$$

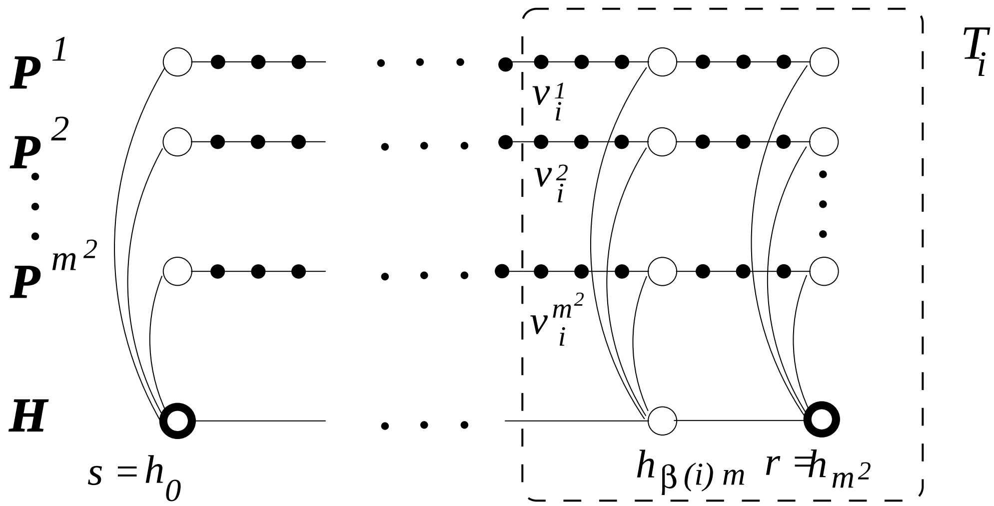

FIG. 3. *The tail set T_i in the graph F_m^2.*

Denote the *state* of the vertex v at the beginning of round t during the execution $\varphi_{\mathcal{X}}$ on the input $\mathcal{X}$ by $\sigma(v, t, \mathcal{X})$. In two different executions $\varphi_{\mathcal{X}}$ and $\varphi_{\mathcal{X}'}$, a vertex reaches the same state at time t, i.e., $\sigma(v, t, \mathcal{X}) = \sigma(v, t, \mathcal{X}')$, iff it receives the same sequence of messages on each of its incoming links; for different sequences, the states are distinguishable.

For a given set of vertices $U = \{v_1, \ldots, v_l\} \subseteq V$, a *configuration*

$$C(U, t, \mathcal{X}) = \langle \sigma(v_1, t, \mathcal{X}), \ldots, \sigma(v_l, t, \mathcal{X}) \rangle$$

is a vector of the states of the vertices of U at the beginning of round t of the execution $\varphi_{\mathcal{X}}$. Denote by $\mathcal{C}[U, t]$ the collection of all possible configurations of the subset $U \subseteq V$ at time t over all executions $\varphi_{\mathcal{X}}$ of algorithm $\mathcal{A}_{\text{mail}}$ (i.e., on all legal inputs $\mathcal{X}$), and let $\rho[U, t] = |\mathcal{C}[U, t]|$.

Prior to the beginning of the execution (i.e., at the beginning of round $t = 0$), the input string $\mathcal{X}$ is known only to the sender s. The rest of the vertices are found in some initial state, described by the configuration $C_{init} = C(T_0, 0, \mathcal{X})$, which is independent of $\mathcal{X}$. Thus, in particular, $\rho[T_0, 0] = 1$. (Note, however, that $\rho[V, 0] = 2^{m^2}$.)

Our main lemma is the following.

LEMMA 3.2. *For every $0 \le t < m^2$,*

$$\rho[T_{t+1}, t + 1] \le (2^B + 1) \cdot \rho[T_t, t].$$

Proof. The lemma is proved by showing that in round $t+1$ of the algorithm, each configuration in $\mathcal{C}[T_t, t]$ branches off into at most $2^B + 1$ different configurations of $\mathcal{C}[T_{t+1}, t+1]$.

Fix a configuration $\hat{C} \in \mathcal{C}[T_t, t]$, and let $\delta = \beta(t+1)$. The tail set T_{t+1} is connected to the rest of the graph by the highway edge $f_{\delta-1} = (h_{(\delta-1)m}, h_{\delta m})$ and by the m^2 path edges $e_t^j = (v_t^j, v_{t+1}^j)$, $1 \le j \le m^2$. (See Figure 4.)

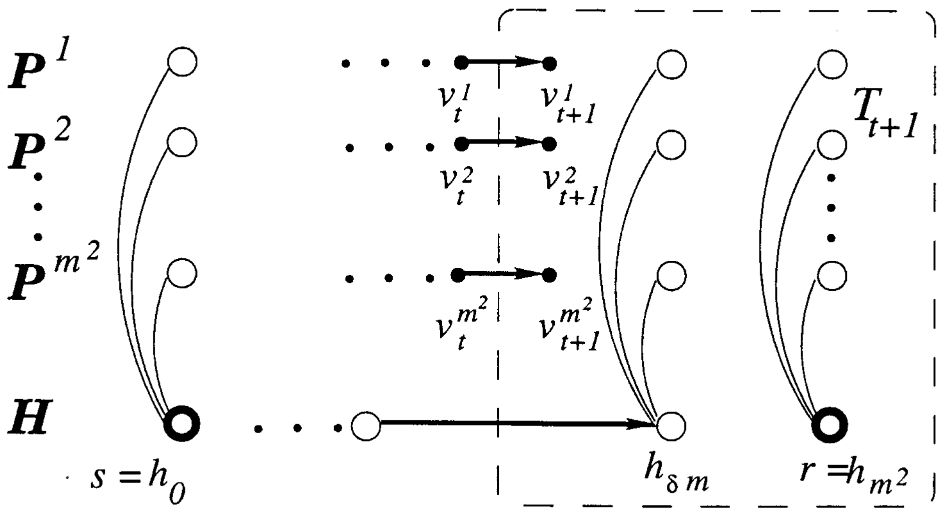

FIG. 4. *The edges entering the tail set T_{t+1}.*

Let us count the number of different configurations in $\mathcal{C}[T_{t+1}, t+1]$ that may result from the configuration $\hat{C}$. Starting from the configuration $\hat{C}$, each vertex v_t^j is restricted to a single state, and hence it sends a single (well determined) message over the edge e_t^j to v_{t+1}^j, thus not introducing any divergence in the execution. The same applies to all the edges internal to T_{t+1}. As for the highway edge $f_{\delta-1}$, the vertex $h_{(\delta-1)m}$ is not in the set T_t; hence it may be in any one of many possible states, and the value passed over this edge into the set T_{t+1} is not determined by the configuration $\hat{C}$. However, due to the restriction of the B-bounded-message model, at most $2^B + 1$ different behaviors of $f_{\delta-1}$ can be observed by $h_{\delta m}$. Thus altogether, the configuration $\hat{C}$ branches off into at most $2^B + 1$ possible configurations $\hat{C}_1, \ldots, \hat{C}_{2^B+1} \in \mathcal{C}[T_{t+1}, t+1]$, differing only by the state $\sigma(h_{\delta m}, t+1, \mathcal{X})$. The lemma follows. $\square$

Applying Lemma 3.2 and the fact that $\rho[T_0, 0] = 1$, we get the following result.

COROLLARY 3.3. *For every $0 \le t < m^2$,*

$$\rho[T_t, t] \le (2^B + 1)^t.$$

Let t_{end} denote the time it takes algorithm $\mathcal{A}_{\text{mail}}$ to complete the mailing. As argued earlier, at that time, the receiver r may be in at least 2^{m^2} different states, hence necessarily $\rho[T_{t_{end}}, t_{end}] \ge 2^{m^2}$. Applying Corollary 3.3, we get that $(2^B + 1)^{t_{end}} \ge 2^{m^2}$, implying the following.

LEMMA 3.4. *For every $m \ge 1$, solving the mailing problem* $\texttt{Mail}(F_m^2, h_0, h_{m^2}, m^2)$ *in the B model requires $\Omega(m^2/B)$ time.*

4. A lower bound for the MST problem on $\mathcal{J}_m^2$. In this section, we use the results achieved in the previous sections and show that in the B model for $B \geq 3$, the distributed MST problem cannot be solved faster than $\Omega(m^2/B)$ on weighted versions of the graphs F_m^2. In order to prove this lower bound, we define in subsection 4.1 a family of weighted graphs $\mathcal{J}_m^2$, based on F_m^2 but differing in their weight assignments. Then in subsection 4.2, we show that any algorithm solving the MST problem on the graphs of $\mathcal{J}_m^2$ can also be used to solve the mailing problem on F_m^2 with the same time complexity. Subsequently, the lower bound for the distributed MST problem follows from the lower bound given in the previous section for the mailing problem in F_m^2.

4.1. The graph family $\mathcal{J}_m^2$. The graphs F_m^2 defined earlier were unweighted. In this subsection, we define for every graph F_m^2 a family of weighted graphs

$$\mathcal{J}_m^2 = \{J_{m,\gamma}^2 = (F_m^2, \omega_\gamma) \mid 1 \leq \gamma \leq 2^{m^2}\},$$

where ω_γ is an edge weight function.

Recall that in the graph F_m^2 there are three types of edges, namely, highway edges, edges of paths $\mathcal{P}^j$, and star spokes. In all the weight functions ω_γ, all the edges of the highway $\mathcal{H}$ and the paths $\mathcal{P}^j$ are assigned the weight 0. The spokes of all stars except S_0 and S_m are assigned the weight 4. The spokes of the star S_m are assigned the weight 2.

The only differences between different weight functions ω_γ occur on the m^2 spokes of the star S_0. Specifically, each of these m^2 spokes is assigned a weight of either 1 or 3; thus there are 2^{m^2} possible combinations of weight assignments. (See Figure 5.)

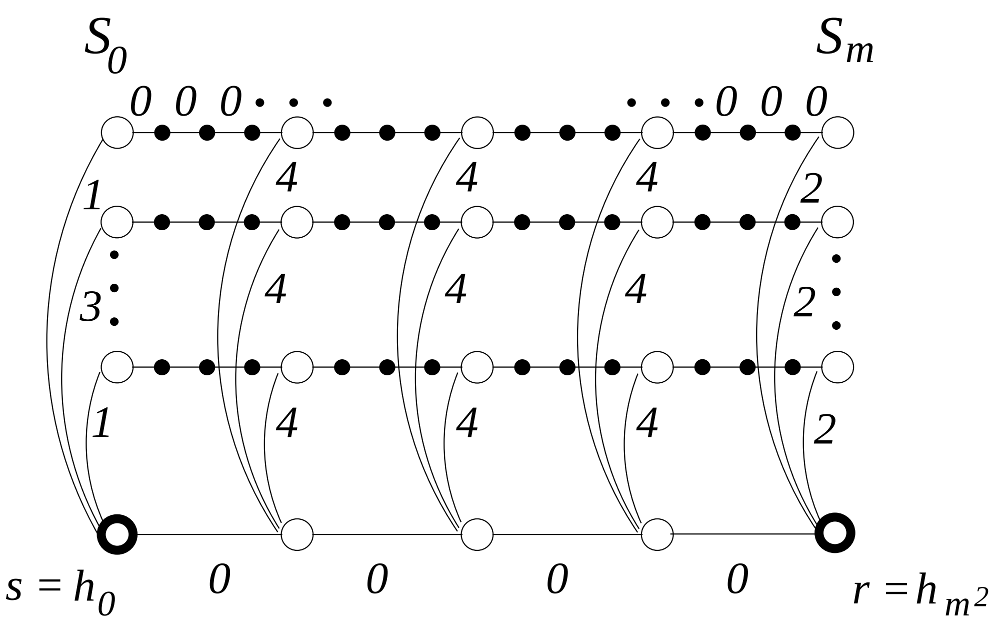

FIG. 5. *The edge weights assigned to $J_{m,\gamma}^2$.*

Since discarding all spoke edges of weight 4 from the graph $J_{m,\gamma}^2$ leaves it connected, and since every spoke edges of weight 4 is the heaviest edge on some cycle in the graph, the following is clear.

LEMMA 4.1. *No spoke edge of weight* 4 *belongs to the MST of* $J^2_{m,\gamma}$ *for every* $1 \le \gamma \le 2^{m^2}$.

Let us pair the spoke edges of S_0 and S_m, denoting the jth pair (for $1 \le j \le m^2$) by

$$PE^j = \{(s, v^j_0), (r, v^j_{m^2})\}.$$

LEMMA 4.2. *For every* $1 \le \gamma \le 2^{m^2}$ *and* $1 \le j \le m^2$, *exactly one of the two edges of* PE^j *belongs to the MST of* $J^2_{m,\gamma}$, *namely, the lighter one.*

Proof. Since the MST must be connected, at least one of the two edges of PE^j must belong to it, as otherwise the path $\mathcal{P}^j$ is completely disconnected from the rest of the graph, by Lemma 4.1. It remains to show that the MST cannot contain both edges of PE^j.

The proof is by contradiction. Consider the cycle in $J^2_{m,\gamma}$ consisting of the edges of $\mathcal{H}$, PE^j, and $\mathcal{P}^j$, and suppose that *both* edges of PE^j are in the MST. In order for the MST to be cycle-free, at least one edge e of either the highway $\mathcal{H}$ or the path $\mathcal{P}^j$ must not belong to the MST. Since the edges of $\mathcal{H}$ and $\mathcal{P}^j$ have zero weight, $\omega_\gamma(e) = 0$. Hence deleting the heavier edge of the pair PE^j and adding the edge e instead leaves us with a lighter tree than the original one, leading us to contradiction. ∎

LEMMA 4.3. *For every* $m \ge 2$ *and* $1 \le \gamma \le 2^{m^2}$, *all the edges of the highway* $\mathcal{H}$ *and the paths* $\mathcal{P}^j$, *for* $1 \le j \le m^2$, *belong to the MST of* $J^2_{m,\gamma}$.

Figure 6 illustrates the remaining candidate edges to join the MST. Bold edges belong to the MST under any edge weight function. Of the remaining edges, exactly one of each pair will join the MST, depending on the particular weight assignment to the spoke edges of the star S_0.

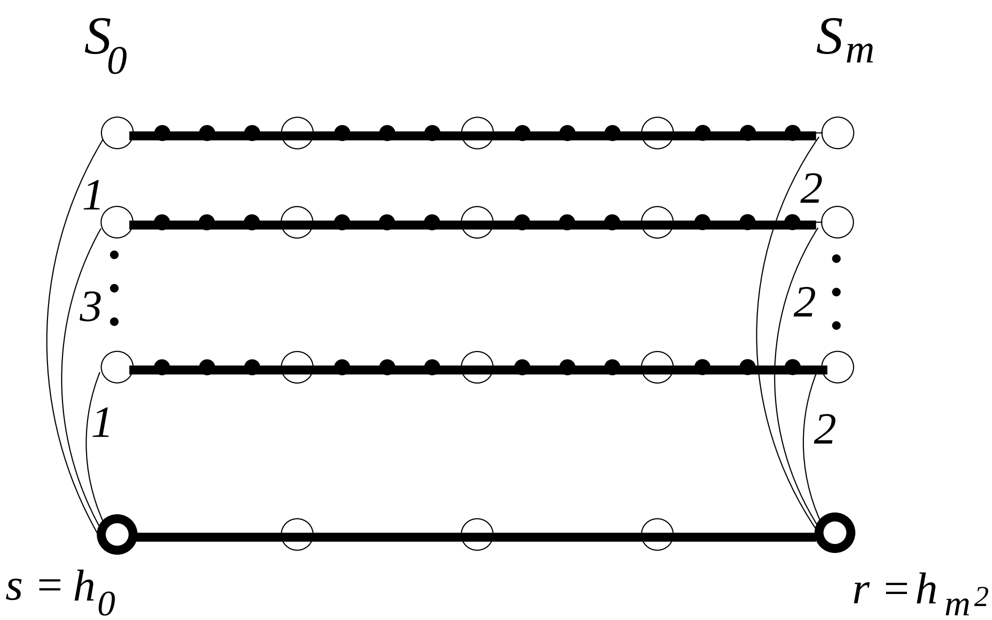

FIG. 6. *The remaining candidate edges to join the MST of* $J^2_{m,\gamma}$. *Bold edges belong to the MST under any edge weight function.*

4.2. The lower bound.

LEMMA 4.4. *Any distributed algorithm for constructing an MST on the graphs of the class $\mathcal{J}_m^2$ can be used to solve the* $\texttt{Mail}(F_m^2, h_0, h_{m^2}, m^2)$ *problem on F_m^2 with the same time complexity.*

Proof. Consider an algorithm $\mathcal{A}_{\rm mst}$ for the MST problem, and suppose that we are given an instance of the $\texttt{Mail}(F_m^2, h_0, h_{m^2}, m^2)$ problem with input string $\mathcal{X}$. We use the algorithm $\mathcal{A}_{\rm mst}$ to solve this instance of the mailing problem as follows. The sender $s = h_0$ initiates the construction of an instance of the MST by turning F_m^2 into a weighted graph from $\mathcal{J}_m^2$, setting the edge weights as follows: for each $x_i \in \mathcal{X}$, $1 \le i \le m^2$, it sets the weight variable W_i^s corresponding to the spoke edge $e_i \in E(S_0)$ to be

$$W_i^s = \begin{cases} 3, & x_i = 1, \\ 1, & x_i = 0. \end{cases}$$

The rest of the graph edges are assigned fixed weights as specified in subsection 4.1. Note that the weights for all the edges except those connecting s to its immediate neighbors in S_0 do not depend on the particular input instance at hand. Hence a single round of communication between s and its S_0 neighbors suffices for performing this assignment; s assigns its edges weights according to its input string $\mathcal{X}$, and needs to send at most one message to each of its neighbors on S_0, to notify it about the weight of the spoke connecting them.

Every vertex v in the network, upon receiving the first message of algorithm $\mathcal{A}_{\rm mst}$, assigns the values defined by the edge weight function ω_γ to its variables W_i^v. (As discussed earlier, this does not require v to know γ, as its assignment is identical under all weight functions ω_γ, $1 \le \gamma \le 2^{m^2}$.) From this point on, v may proceed with executing algorithm $\mathcal{A}_{\rm mst}$ for the MST problem.

Once algorithm $\mathcal{A}_{\rm mst}$ terminates, the receiver vertex r determines its output for the mailing problem, by setting $X_i^r \leftarrow Y_i^r$ for $1 \le i \le m^2$.

By Lemma 4.2, the lighter edge of each pair PE^j, for $1 \le j \le m^2$, belongs to the MST; thus in the set of variables $Y_1^r, \ldots, Y_{m^2}^r$ obtained by the vertex r as a result of solving the MST problem, $Y_j^r = 1$ corresponds to the assignment of $\omega(h_0, v_0^j) = 3$ to the jth edge of S_0, while $Y_j^r = 0$ corresponds to the assignment of 1 to that edge; hence for every j, Y_j^r equals x_j, the jth bit of $\mathcal{X}$. Hence the resulting algorithm has correctly solved the given instance of the mailing problem. $\square$

Combined with Lemma 3.4, we now have the following theorem.

THEOREM 4.5. *For every $m \ge 1$, any distributed algorithm for constructing an MST on the graphs of the family $\mathcal{J}_m^2$ in the B model for $B \ge 3$ requires $\Omega(m^2/B)$ time.* $\square$

COROLLARY 4.6. *Any distributed algorithm for the MST problem in the B model for $B \ge 3$ requires $\Omega(\sqrt{n}/B)$ time on some n-vertex graphs of diameter $O(n^{1/4})$.*

5. A lower bound on the mailing problem on F_m^K. This section generalizes the results of the previous section to the graphs F_m^K for $K \ge 3$, thus establishing the desired lower bound.

5.1. The graphs F_m^K. Given two integer parameters $m, K \ge 2$, construct the graph F_m^K as follows. The two basic units are still the path and the highway, with the following changes. The basic path $\mathcal{P}$ now has $m^K + 1$ vertices, i.e.,

$$V(\mathcal{P}) = \{v_0, \ldots, v_{m^K}\} \quad \text{and} \quad E(\mathcal{P}) = \{(v_i, v_{i+1}) \mid 0 \le i \le m^K - 1\}.$$

There are $K-1$ highways, denoted $\mathcal{H}^1, \ldots, \mathcal{H}^{K-1}$. The level-$\ell$ highway $\mathcal{H}^\ell$ consists of $m^\ell + 1$ vertices, i.e.,

$$V(\mathcal{H}^\ell) = \{h^\ell_{im^{K-\ell}} \mid 0 \le i \le m^\ell\} \quad \text{and}$$
$$E(\mathcal{H}^\ell) = \{(h^\ell_{im^{K-\ell}}, h^\ell_{(i+1)m^{K-\ell}}) \mid 0 \le i \le m^\ell - 1\}.$$

Each highway vertex $h^\ell_{im^{K-\ell}}$ is connected to the corresponding path vertex $v^j_{im^{K-\ell}}$ by a spoke edge. Figure 7 depicts these connections for the case of $m = K = 3$.

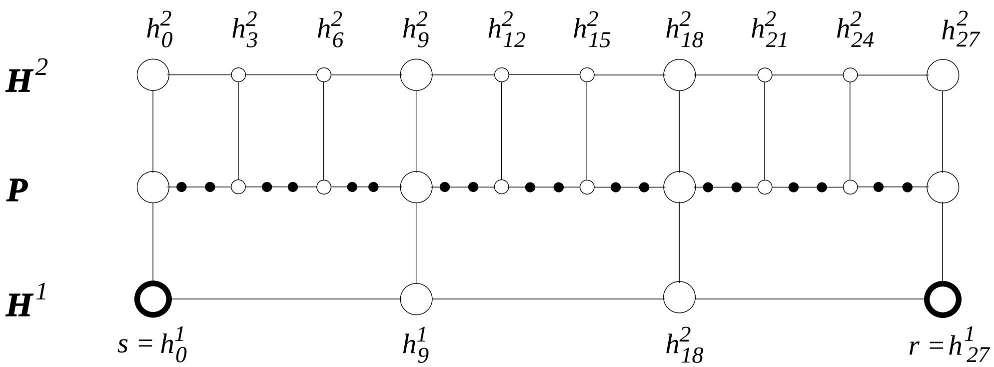

FIG. 7. *The connections between the path $\mathcal{P}$ and the highways $\mathcal{H}^1$ and $\mathcal{H}^2$ for $m = K = 3$.*

The graph F^K_m is constructed by taking m^K copies of the path $\mathcal{P}$, denoted $\mathcal{P}^1, \ldots, \mathcal{P}^{m^K}$, and connecting them all to the same level-ℓ highway $\mathcal{H}^\ell$, for each $1 \le \ell \le K-1$. The vertex h^1_0 is the intended sender s, and the vertex $h^1_{m^K}$ is the intended receiver r. (See Figure 8, showing the graph F^3_3.)

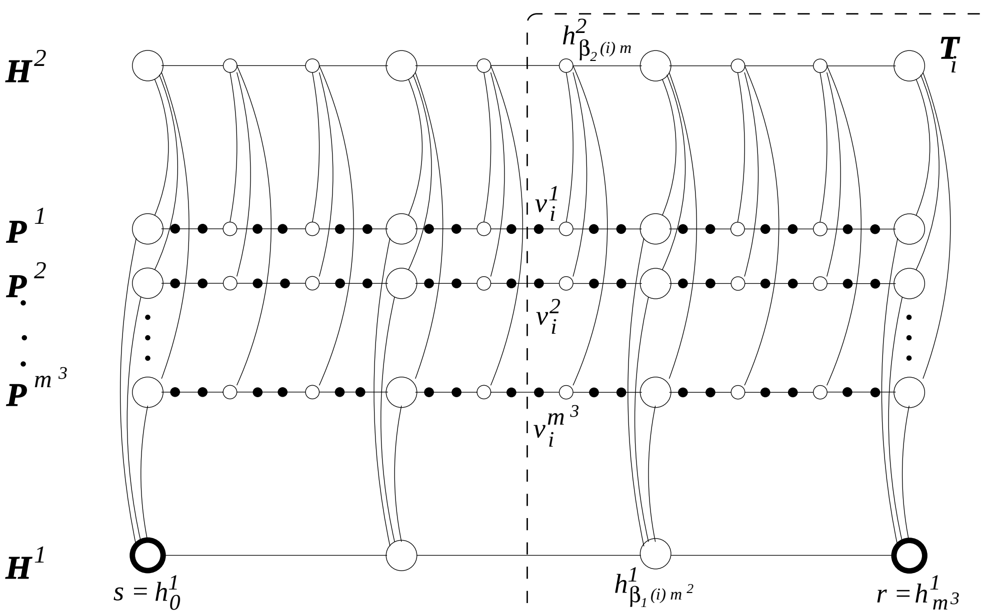

FIG. 8. *The graph F^3_m (here also $m = 3$).*

The following two facts are easily verified.

LEMMA 5.1. *The cardinality of F_m^K is $n = \Theta(m^{2K})$ and its diameter is $\Theta(Km)$.*

5.2. The lower bound. The lower bound for the mailing problem can be extended from F_m^2 to F_m^K for $K \geq 3$ in a natural way. Consider some arbitrary algorithm $\mathcal{A}_{\mathrm{mail}}$, and let $\varphi_{\mathcal{X}}$ denote the execution of $\mathcal{A}_{\mathrm{mail}}$ on the input $\mathcal{X}$ in the graph F_m^K. The notion of a tail set is generalized to F_m^K for $K \geq 3$ as follows. For every $1 \leq j \leq m$, define the *tail* of the path $\mathcal{P}^j$ as before, i.e.,

$$T_i(\mathcal{P}^j) = \{v_l^j \mid i \leq l \leq m^K\}.$$

Let $\beta^\ell(i)$ denote the least integer δ such that $\delta m^{K-\ell} \geq i$, and define the tail of $\mathcal{H}^\ell$ as

$$T_i(\mathcal{H}^\ell) = \{h^\ell_{jm^{K-\ell}} \mid \beta^\ell(i) \leq j \leq m^\ell\}.$$

The tail set of F_m^K is the union of those tails,

$$T_i = T_i(\mathcal{H}) \cup \bigcup_j T_i(\mathcal{P}^j).$$

(See Figure 8.) Again, for $i = 0$ the definition is $T_0 = V \setminus \{h_0^1\}$.

The main lemma becomes the natural extension of Lemma 3.2, and its proof is similar. The notions of configuration, collection of configurations, and absolute size ρ of collections of possible configurations are defined in the same way as in section 3.2.

LEMMA 5.2. *For any $0 \leq t < m^K$,*

$$\rho[T_{t+1}, t+1] \leq (2^B + 1)^{K-1} \cdot \rho[T_t, t].$$

Proof. The lemma is proved by showing that in round $t+1$ of the algorithm, each configuration in $\mathcal{C}[T_t, t]$ branches off into at most $(2^B+1)^{K-1}$ different configurations of $\mathcal{C}[T_{t+1}, t+1]$.

Fix a configuration $\hat{C} \in \mathcal{C}[T_t, t]$. The tail set T_{t+1} is connected to the rest of the graph by the highway edges $f^\ell_{\beta^\ell(t+1)-1} = (h^\ell_{(\beta^\ell(t+1)-1)m^{K-\ell}}, h^\ell_{\beta^\ell(t+1)m^{K-\ell}})$, for every $1 \leq \ell \leq K-1$, and by the m^K path edges $e_t^j = (v_t^j, v_{t+1}^j)$, $1 \leq j \leq m^K$.

Consider the number of different configurations in $\mathcal{C}[T_{t+1}, t+1]$ that may result from $\hat{C}$. Starting from the configuration $\hat{C}$, each vertex v_t^j is restricted to a single state, and hence it sends a single (well determined) message over the edge e_t^j to v_{t+1}^j, thus not introducing any divergence in the execution. The same applies to all the edges internal to T_{t+1}.

The situation with highway edges $f^\ell_{\beta^\ell(t+1)-1}$ is different as there are K possible cases. When $\beta^\ell(t+1) = \beta^\ell(t)$ for $1 \leq \ell \leq K-1$, the vertices $h^\ell_{(\beta^\ell(t)-1)m^{K-\ell}}$ are not in the set T_t; hence their state is not defined by the choice of $\hat{C}$. The value passed over the edge $f^\ell_{\beta^\ell(t)-1}$ into the set T_{t+1} is thus unknown. However, due to the restriction of the B-bounded-message model, at most 2^B+1 different behaviors of can be observed by each vertex $h^\ell_{\beta^\ell(t)m^{K-\ell}}$, resulting in 2^B+1 possible states for each such node. Considering the entire set $\{h^\ell_{\beta^\ell(t)m^{K-\ell}} \mid 1 \leq \ell \leq K-1\}$, the single state $\hat{C}$ results in $(2^B+1)^{K-1}$ states of the tail set T_{t+1} at time $t+1$.

In the rest of the cases, $\beta^\ell(t+1) = \beta^\ell(t) + 1$ for at least one ℓ, when passing to the next tail set causes the exclusion of the highway point $h^\ell_{\beta^\ell(t)m^{K-\ell}}$ from the tail set. Here, a well defined message is sent over $f^\ell_{\beta^\ell(t+1)-1}$ since the state of $h^\ell_{\beta^\ell(t)m^{K-\ell}}$

is defined by the configuration $\hat{C}$. It follows that in these cases the number of possible states of T_{t+1} is less than $(2^B+1)^{K-1}$.

Altogether, the configuration $\hat{C}$ branches off into at most $(2^B+1)^{K-1}$ possible configurations $\hat{C}_1,\ldots,\hat{C}_{(2^B+1)^{K-1}} \in \mathcal{C}[T_{t+1},t+1]$, differing by the states $\sigma(h^\ell_{\beta^\ell(t+1)m^{K-\ell}},t+1,\mathcal{X})$. The lemma follows. □

COROLLARY 5.3. *For any* $0 \le t < m^K$,

$$\rho[T_t,t] \le (2^B+1)^{(K-1)t}.$$

Letting t_{end} denote the time it takes algorithm $\mathcal{A}_{\rm mail}$ to complete the mailing, we derive, similar to the proof for $K=2$, that necessarily

$$(2^B+1)^{(K-1)t_{end}} \ge \rho[T_{t_{end}},t_{end}] \ge 2^{m^K},$$

implying the following.

LEMMA 5.4. *For every* $K,m\ge 2$, *solving the mailing problem* $\mathtt{Mail}(F^K_m,h^1_0,h^1_{m^K},m^K)$ *in the B model requires* $\Omega(m^K/(BK))$ *time.*

6. A generalized lower bound on the MST on $\mathcal{J}^K_m$. Finally, we show the lower bound for the MST problem on the weighted versions of the graphs F^K_m.

6.1. The graph families $\mathcal{J}^K_m$. Let us define the families of weighted graphs $\mathcal{J}^K_m$. For every two integers $m,K\ge 2$, let

$$\mathcal{J}^K_m = \{J^K_{m,\gamma} = (F^K_m,\omega^K_\gamma) \mid 1\le\gamma\le 2^{m^K}\},$$

where ω^K_γ is the weight function defined as follows. All the edges of the highways $\mathcal{H}^\ell$ for $1\le\ell\le K-1$ and the paths $\mathcal{P}^j$ for $1\le j\le m^K$ are assigned zero weight. It remains to assign the weights to the spoke edges.

Consider a subgraph of F^K_m, consisting of all the available paths $\mathcal{P}_j$, $1\le j\le m^K$, a single highway $\mathcal{H}^\ell$ and all the connections of this highway to the paths. Consider a single node $h^\ell_{im^{K-\ell}}$ of the highway $\mathcal{H}^\ell$ and the set of all its connections to $\mathcal{P}^1,\ldots,\mathcal{P}^{m^K}$. Following the terminology of the case $K=2$, this is termed the *level-ℓ star* $S^{K,\ell}_{m,i}$. There are m^ℓ such stars at level ℓ.

Consider the collection of the stars $S^{K,\ell}_{m,i}$ for $1\le i\le m^\ell$, $2\le\ell\le K-1$ (excluding the first star of each level ℓ). The spokes of these stars are assigned the weight 4. The spokes of the first star of each level, $S^{K,\ell}_{m,0}$, are assigned as follows. The spokes connecting the star centers h^ℓ_0, $1\le\ell\le K-1$ to the extreme vertex v^1_0 of the path $\mathcal{P}^1$ are assigned zero weight. The rest of the spokes are assigned the weight 4.

For the collection of the level-1 stars, $S^{K,1}_{m,i}$, the assignment is as follows. The spokes of all the stars except the two extreme ones, $S^{K,1}_{m,0}$ and $S^{K,1}_{m,m}$, are assigned the weight 4. The spokes of the last star $S^{K,1}_{m,m}$ are assigned weight 2. The weight assignment to the m^K spokes of the star $S^{K,1}_{m,0}$ depends on the particular function ω^K_γ, with each spoke assigned a value of either 1 or 3, as in section 4.2.

LEMMA 6.1. *No spoke edge of weight* 4 *belongs to the MST of* $J^K_{m,\gamma}$ *for every* $1\le\gamma\le m^K$.

Proof. Following the proof of Lemma 4.1, it can be shown that the elimination of all spoke edges of weight 4 from the graph $J^K_{m,\gamma}$ leaves the graph connected. Since all the edges of all the highways and basic paths have zero weight, none of their edges is eliminated. Consider the connectivity of the highway $\mathcal{H}^1$ and the basic paths

$\mathcal{P}^1, \dots, \mathcal{P}^{m^K}$. By construction, the spokes of two stars, namely, $S^{K,1}_{m,0}$ and $S^{K,1}_{m,m}$, have edges of weight at most 3, which guarantees that every basic path is connected to $\mathcal{H}_1$. The rest of the highways $\mathcal{H}^\ell$, for $2 \le \ell \le K-1$, are connected to the node v^1_0 of the basic path $\mathcal{P}^1$ by their nodes h^ℓ_0 via a zero-weight edge. Thus all the highways are connected to the path $\mathcal{P}^1$ and through it to the highway $\mathcal{H}^1$ and all the other basic paths.

Hence every spoke edge of weight 4 occurs as the heaviest edge on some cycle in the graph, implying the lemma. ∎

LEMMA 6.2. *For every $2 \le \ell \le K-1$, the edge (h^ℓ_0, v^1_0) belongs to the MST.*

Proof. By Lemma 6.1, no spoke edge of weight 4 belongs to the MST. By construction, each of the highways $\mathcal{H}^\ell$ for $2 \le \ell \le K-1$ is connected to the rest of the graph by spokes of weight 4 and by a single zero-weight spoke of the star $S^{K,\ell}_{m,0}$ connecting it to $\mathcal{P}^1$. Thus in order for the MST to be connected, the zero-weight edge must belong to the MST. ∎

Let us pair the spoke edges of $S^{K,1}_{m,0}$ and $S^{K,1}_{m,m}$ connecting $\mathcal{H}^1$ to $\mathcal{P}^j$ for $1 \le j \le m^K$, denoting the jth pair (for $1 \le j \le m^K$) by

$$PE^j = \{(s, v^j_0), (r, v^j_{m^K})\}.$$

By a proof similar to that of Lemma 4.2, we get the following lemma.

LEMMA 6.3. *For every $1 \le j \le m^K$, exactly one of the two edges of PE^j belongs to the MST of $J^K_{m,\gamma}$, namely, the lighter one.*

6.2. The generalized lower bound on distributed MST. We obtained an instance of the MST problem, in which the membership of edges in the MST is predetermined for all but the m^K edge pairs PE^j. Following the proof method of Lemma 4.4, we show that any algorithm solving the distributed MST problem on $\mathcal{J}^K_m$ can be used for solving the mailing problem in the same time complexity, implying the following.

THEOREM 6.4. *For every $m, K \ge 2$, any distributed algorithm for constructing an MST on the graphs of the family $\mathcal{J}^K_m$ in the B model for $B \ge 3$ requires $\Omega(m^K/(BK))$ time.*

Proof. Consider an algorithm $\mathcal{A}_{\rm mst}$ for the MST problem, and suppose that we are given an instance of the $\texttt{Mail}(F^K_m, h^1_0, h^1_{m^K}, m^K)$ problem with input string $\mathcal{X}$. We use the algorithm $\mathcal{A}_{\rm mst}$ to solve this instance of the mailing problem as follows. The sender $s = h^1_0$ initiates the construction of an instance of the MST by turning F^K_m into a weighted graph from $\mathcal{J}^K_m$, setting the edge weights as follows: for each $x_i \in \mathcal{X}$, $1 \le i \le m^K$, it sets the weight variable W^s_i corresponding to the spoke edge $e_i \in E(S^{K,1}_{m,0})$ (the first 1-level star), to be as in the proof of Lemma 4.4. The rest of the graph edges are assigned fixed weights as specified in section 6.1. Note that again, the weights for all the vertices except s and its immediate neighbors in $S^{K,1}_{m,0}$ do not depend on the particular input instance at hand; hence a single round of communication between s and its $S^{K,1}_{m,0}$ neighbors suffices for performing this assignment.

From this point on, we may proceed with executing algorithm $\mathcal{A}_{\rm mst}$ for the MST problem. Once algorithm $\mathcal{A}_{\rm mst}$ terminates, the receiver r determines its output for the mailing problem, by setting $X^r_i \leftarrow Y^r_i$ for $1 \le i \le m^K$.

The fact that the resulting algorithm has correctly solved the given instance of the mailing problem is established as in the proof of Lemma 4.4, relying on Lemma 6.3. The theorem now follows from Lemma 5.4. ∎

COROLLARY 6.5. *For every $K \geq 2$, there exists a family of n-vertex graphs of diameter $O(Kn^{1/(2K)},)$ such that any distributed algorithm for the MST problem in the B model for $B \geq 3$ requires $\Omega(\sqrt{n}/(BK))$ time on some of those graphs.*

COROLLARY 6.6. *For every $n \geq 2$, there exists a family of n-vertex graphs of diameter $O(\log n)$ such that any distributed algorithm for the MST problem in the B model for $B \geq 3$ requires $\Omega(\sqrt{n}/(B \log n))$ time on some of those graphs.*

Finally, let us comment that it has recently been shown that using Yao's method [Yao77] it is possible to extend the lower bound of Lemma 5.4 on the mailing problem into a lower bound on the expected time complexity of any *randomized* (Las Vegas) distributed algorithm for the mailing problem (see [P00, Chapter 24, Exercise 9]). This, in turn, yields the following lower bound on the time complexity of randomized algorithms for distributed construction: For every $n \geq 2$, there exists a family of n-vertex graphs of diameter $O(\log n)$ such that any randomized Las Vegas distributed algorithm for the MST problem in the B model requires $\Omega(\sqrt{n}/(B \log n))$ expected time on some of those graphs.

7. Open problems. Several interesting problems can be considered for future research. The first direction concerns the limitations of the presented lower bound. To begin with, the lower bound does not seem to extend to diameters lower than $O(\log n)$. As graphs with $D = 1$ admit an $O(\log n)$ distributed algorithm for MST construction, one may expect an interesting interdependence between the time to construct an MST and the network's diameter.

Second, one may consider a model allowing L-bit edge weights for $L > B$. While our lower bound still holds, stronger bounds may apply. Note that the transmission of an edge weight can be carried out in this model by sending $\Theta(L/B)$ separate messages. Hence each of the existing algorithms for distributed MST can be adapted to this model with a multiplicative slowdown of L/B. The algorithm of [KP98], for instance, will have time complexity $O((D + \sqrt{n} \log^* n)L/B)$. However, it is less clear whether this slowdown is necessary or if it can be avoided. It seems easy to verify (say, by considering a ring with two diametrically opposing edges having the extreme weights) that $\Omega(L/B)$ is indeed a lower bound on the time complexity of the problem in this model. However, it is plausible that the algorithm of [KP98] can be modified using pipelining ideas to yield a time complexity close to $O(D + \sqrt{n} \log^* n + L/B)$.

Another research direction is to try to reduce the *communication* complexity of the nearly time optimal algorithm of [KP98] from $O(|E| + n^{3/2})$ towards the lower bound of $O(|E| + n \log n)$.

Finally, one may consider the possibility of devising faster algorithms that construct an *approximation* to the MST, namely, a spanning tree whose total weight is near-minimum. To the best of our knowledge, nothing nontrivial is currently known about this problem, and little is known about distributed approximation algorithms in general, but this direction may well deserve further study.

REFERENCES

[A87] B. AWERBUCH, *Optimal distributed algorithms for minimum-weight spanning tree, counting, leader election and related problems*, in Proceedings of the 19th Symposium on Theory of Computing, ACM, New York, 1987, pp. 230–240.

[GHS83] R. GALLAGER, P. HUMBLET, AND P. SPIRA, *A distributed algorithm for minimum-weight spanning trees*, ACM Trans. Programming Languages Systems, 5 (1983), pp. 66–77.

[GKP98] J. A. GARAY, S. KUTTEN, AND D. PELEG, *A sublinear time distributed algorithm for minimum-weight spanning trees*, SIAM J. Comput., 27 (1998), pp. 302–316.

[KP98] S. Kutten and D. Peleg, *Fast distributed construction of small k-dominating sets and applications*, J. Algorithms, 28 (1998), pp. 40–66.

[L92] N. Linial, *Locality in distributed graph algorithms*, SIAM J. Comput., 21 (1992), pp. 193–201.

[P90] D. Peleg, *Time-optimal leader election in general networks*, J. Parallel Distrib. Comput., 8 (1990), pp. 96–99.

[P00] D. Peleg, *Distributed Computing: A Locality-Sensitive Approach*, SIAM, Philadelphia, PA, 2000.

[PR99] D. Peleg and V. Rubinovich, *A near-tight lower bound on the time complexity of distributed MST construction*, in Proceedings of the 40th IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 1999, pp. 253–261.

[Yao77] A. Yao, *Probabilistic computations: Towards a unified measure of complexity*, in Proceedings of the 17th IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 1977, pp. 222–227.