2014 IEEE Annual Symposium on Foundations of Computer Science

Topology Matters in Communication

Arkadev Chattopadhyay and Jaikumar Radhakrishnan
School of Technology and Computer Science
Tata Institute of Fundamental Research
Mumbai, India
Email: arkadev.c@tifr.res.in, jaikumar@tifr.res.in

Atri Rudra
Department of Computer Science and Engineering
University at Buffalo, SUNY
Buffalo, USA
Email: atri@buffalo.edu

***Abstract*—We consider the communication cost of computing functions when inputs are distributed among the vertices of an undirected graph. The communication is assumed to be point-to-point: a processor sends messages only to its neighbors. The processors in the graph act according to a pre-determined protocol, which can be randomized and may err with some small probability. The communication cost of the protocol is the total number of bits exchanged in the worst case.**

Extending recent work that assumed that the graph was the complete graph (with unit edge lengths), we develop a methodology for showing lower bounds that are sensitive to the graph topology. In particular, for a broad class of graphs, we obtain a lower bound of the form $\Omega(k^2 n)$, for computing a function of k inputs, each of which is n-bits long and located at a different vertex. Previous works obtained lower bounds of the form $\Omega(kn)$. This methodology yields a variety of other results including the following: A tight lower bound (ignoring poly-log factors) for Element Distinctness, settling a question of Phillips, Verbin and Zhang (SODA '12, [1]); a distributed XOR lemma; a lower bound for composed functions, settling a question of Phillips et al. [1]; new topology-dependent bounds for several natural graph problems considered by Woodruff and Zhang (DISC '13, [2]).

To obtain these results we use tools from the theory of metric embeddings and represent the topological constraints imposed by the graph as a collection of cuts, each cut providing a setting where our understanding of two-party communication complexity can be effectively deployed.

***Keywords*-Communication Complexity; Metric Embeddings; Distributed Computing.**

I. Introduction

We consider the cost of communication associated with computation in a distributed model of computing. The processors are located at the vertices of an undirected graph and communicate with each other by exchanging messages. Some of the processors, called terminals, have inputs; the others just participate in the computation. All processors collaborate to compute a boolean function of the inputs. In this paper, we study the communication cost of this task for various functions. Such a study has been carried out in the past in various guises and lower and upper bounds have been derived. These results typically assume complete connectivity, or restricted families of graph (e.g. the line graph, the ring, etc.). We develop a methodology for deriving lower bounds that relates the communication cost to well-known parameters of the graph, making them sensitive to the topology of the graph.

In order to present the background to our work and our results in a common setting, it will be helpful to use the following terminology. As usual, the graph will consist of vertices and edges. Every vertex of the graph has a processor. Some of the processors, known as terminals, receive inputs. The processors operate synchronously with a common clock, and communicate by sending bits across the edges of the graph according to a fixed protocol. We will assume that before a bit is sent, the receiver knows that it is going to receive one. We will assume unit cost for transmitting a bit across an edge. The goal is to compute a common function of the inputs; we will assume that at the end a predetermined terminal node has the correct answer. Thus, a computational problem in this model is of the form $p = (f, G, K, \Sigma)$, where G is the underlying graph, K is the set of terminals, Σ is the alphabet from which each terminal receives its input, and $f : \Sigma^K \to \{0,1\}$ is the function that the processors must compute at the end. In our discussions, whenever the underlying graph becomes clear from the context, we drop it from our notation. If Σ is not explicitly mentioned, then it is assumed to be $\{0,1\}^n$. Throughout the paper we will use k for the size of the set K of terminals.

When the computation is randomized (the main focus of this paper) we will assume that the processors share a random string. Let $R_\epsilon(p)$ be the expected (over the protocol's randomness) communication cost for the worst-case input of a protocol for p that on all inputs errs with probability at most ϵ. This definition differs from the usual definition, where $R_\epsilon(p)$ is defined in terms of worst case communication. It will be easier to state our proofs using this definition; the two notions are closely related, and the lower bounds using our definition are stronger. When stating our upper bounds we will indicate if they hold even for worst-case communication. In particular, $R_0(p)$ will be the worst-case expected communication cost of a zero-error communication protocol for p. The worst case deterministic complexity of the problem p is defined similarly, and is denoted by $D(p)$.

The two-party communication complexity model is

0272-5428/14 $31.00 © 2014 IEEE
DOI 10.1109/FOCS.2014.73

clearly a special case of the above model: the graph is just an edge. Multi-party generalizations of this model, where processors communicate using broadcasts has been considered in several earlier works. The point-to-point communication assumed above is closer to the multi-party communication model of Dolev and Feder [3] who showed that in general that in this model the deterministic complexity is polynomially bounded by the non-deterministic complexity. Subsequently, Duris and Rolim [4] obtained lower bounds in this model on the deterministic and non-deterministic communication complexity. These works assume that each processor may send a bit to any other processor at the same cost, that is, the underlying graph is completely connected.

The point to point model has seen a recent surge of interest [1], [5], [2], [6]. This is because the point-to-point model (as opposed to the broadcast model) arguably better captures many of the modern day networks and has been studied in the many distributed models: e.g. the BSP model of Valiant [7], models for MapReduce [8], [9], massively parallel models to compute conjunctive queries [10], [11], distributed models for learning [12], [13] and in the core distributed computing literature [14].

The recent surge in interest in this model is also in part motivated by proving lower bounds for the distributed functional monitoring framework (see e.g. the recent survey [15]), which generalizes the popular data streaming model [16]. However, all of the recent work assumes that the underlying topology is fully connected. In our opinion this is a strict restriction since in many situations assuming full connectivity would be too strong an assumption. Indeed in areas like sensor networks, researchers have considered the effects of network topology with some success [17] for simple topologies like trees.

The following is the motivating question for our work (which was also mentioned as an interesting direction to pursue in [6]):

> *Can we prove lower bounds on multi-party communication complexity for point to point communication that are sensitive to the topology of the connections between the processors?*

To see how the network topology can make a big difference (for graphs G that have the same number of vertices), consider the trivial protocol that computes the function f by collecting the n-bit inputs of all processors at one designated processor. If the underlying graph is the complete graph (or has constant diameter), then this trivial protocol can be implemented with a total communication of only about kn bits. This is tight, because the designated processor needed to receive kn bits in total. If, however, we were forced to implement this protocol when the underlying graph is the line graph, the total communication would be $\Omega(k^2 n)$. Focusing only on the information each processor needs to obtain would lead us to neglect such factors of k. Our interest is in identifying situations where we can recover this extra $\Theta(k)$ factor (up to a poly-log loss) in our lower bounds and in general match the bound of the trivial algorithm for any topology.

Not surprisingly, the role of topology in computation has been studied extensively in distributed computing. There are three main differences between works in this literature and ours. First, the main objective in distributed computation is to minimize the end to end delay of the computation, which in communication complexity terminology corresponds to the number of rounds need to compute a given function. By contrast, we consider the related but different measure of the total amount of communication. Second, the effect of network topology on the cost of communication has been analyzed to quite an extent when the networks are *dynamic* (see for example the recent survey of Kuhn and Oshman [18]). By contrast, in this work we are mainly concerned with static networks of arbitrary topology. Finally, there has also been work on proving lower bounds for distributed computing on static networks, see e.g. the recent work of Das Sarma et al. [19]. This line of work differs from ours in at least two ways. First, their aim is to prove lower bounds on the number of rounds needed to compute, especially when the edges of the graph are capacitated. Our results, on the other hand, focus on the total communication needed without placing any restriction on the capacities of the edges or the number of rounds involved. Second, the sense in which their bounds are topology dependent is different than ours. For instance, when they prove a lower bound involving the diameter of a graph, they do so by carefully constructing a particular family of graphs for which the claimed bound holds. To the best of our knowledge, their bounds are not known to extend to graphs outside this family. By contrast, our results hold asymptotically for *every* graph.

Our work, in this regard, is perhaps closer to the earlier work of Tiwari [20], who considered computing a function when the inputs are distributed over graphs with simple topologies. In our terminology, Tiwari considered the deterministic complexity $D(f, P_k, \{1, k+1\}, \Sigma)$, where P_k denotes the path of length k.[1] He also considered $D(f, G, V(G), \Sigma)$ for specific functions f (including the element distinctness problem, which we also consider in this paper) and G being a path, grid or ring graph. In all these cases, Tiwari proved lower bounds that match the trivial protocol where all the terminals send their inputs to a

[1] In particular, Tiwari made the following remarkable *linear array conjecture*, which in our terminology can be stated as: $D(f, P_k, \{1, k+1\}, \Sigma) = kD(f, P_2, \{1, 2\}, \Sigma)$ for *all* functions f. That is, the communication costs do not change much if the non-terminal vertices perform no real computation, beyond relaying the messages between the terminals. He further developed techniques for showing lower bounds in this setting, and confirmed that the conjecture was true for several functions. In general, however, the conjecture was found to be false, and it got resolved as follows: Kushilevitz, Linial and Ostrovsky [21] exhibited a function for which $D(f, P_k, \{1, k+1\}, \Sigma) \leq (\frac{3}{4} + o(1))kD(f, P_2, \{1, 2\}, \Sigma)$; Dietzfelbinger [22] showed that $D(f, P_k, \{1, k+1\}, \Sigma) \geq \gamma k D(f, P_2, \{1, 2\}, \Sigma)$, for a constant $\gamma \approx 0.275$.

designated terminal. To the best of our knowledge, Tiwari's results were not generalized to arbitrary topologies.

Our Contributions: We present a general framework to prove lower bounds for general topologies. It is standard to view the computation across a cut as a two-party problem and extract lower bounds using two-party lower bounds: in particular, Tiwari's results on mesh and ring topologies utilized cuts.[2] We refine this method by combining the effect of considering communication across several cuts. The choice of cuts will be guided by the the theory of metric embeddings as in the celebrated works of Bourgain [23] and London, Linial and Rabinovich [24]. None of our proofs are technically difficult in themselves, probably because the tools we use are quite non-trivial. We believe our main contribution is more conceptual: we identify certain key components and show how to combine them to obtain topology dependent lower bounds. We also believe that our framework is fairly general and should be widely applicable. In particular, this framework allows us to generalize many of the existing lower bounds proved assuming the complete graph topology to general graphs. The following are our results.

Element distinctness: In the *element distinctness* problem $(\mathrm{ED}, G, K, \{0,1\}^n)$, each terminal node in $K \subseteq V(G)$ is given an n-bit binary string as input. The goal is to determine if all the inputs are distinct. We express our bounds using the following graph theoretic parameters (see Goddard and Oellerman [25]). The length of the shortest path between u and v will be denoted by $d(u,v)$. The diameter $\mathrm{diam}_G(K)$ of K is the maximum value of $d(u,v)$ as u and v range over K. For a vertex $v \in V(G)$, its *status w.r.t.* K, denoted by $\sigma_K(v)$, is the sum of the distances from u to all other vertices in K. A vertex v whose status w.r.t. K is minimum is said to be a *median* for K, and its status is denoted by $\sigma_G(K)$.

Theorem I.1. *Suppose $n \geq 2\log k$. We have the following upper bounds.*

(i) For all $\frac{1}{2} \geq \epsilon \geq 0$, $R_\epsilon(\mathrm{ED}, K) = O(\sigma_G(K)\log(k/\epsilon))$.

(ii) $R_0(\mathrm{ED}, K) = O(\mathrm{diam}_G(K)n + \sigma_G(K)\log k)$.

The following almost matching lower bounds are true.

(i) For all constant $1/2 > \epsilon > 0$, $R_\epsilon(\mathrm{ED}, K) = \Omega(\sigma_G(K)/\log k)$.

(ii) $R_0(\mathrm{ED}, K) = \Omega(\mathrm{diam}_G(K)n + \sigma_G(K)/\log k)$.

In Section II, we present the proof of this theorem, with a discussion of the tools we employ.

An XOR lemma: Suppose the underlying graph G is on vertex set $[N]$ and $\mathcal{M} = \{(i_\ell, j_\ell)\}_{\ell=1}^{k/2}$ is a partition of K (the set of terminals) into disjoint pairs of distinct elements. For a function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, let $f^{\oplus,\mathcal{M}} \equiv \oplus_{(i,j)\in\mathcal{M}} f(X_i, X_j)$. A natural way to compute $f^{\oplus,\mathcal{M}}$ is to independently compute each $f(X_i, X_j)$ and combine the answers. The cost of this naive protocol will be about $d(\mathcal{M}) \cdot R_{\epsilon/k}(f)$, where $d(\mathcal{M}) = \sum_{(u,v)\in\mathcal{M}} d(u,v)$. Is this optimal? We show the following.

Theorem I.2. *For every constant $1/20 \geq \epsilon > 0$ and every binary function f and every set of pairs $\mathcal{M}$ as described above, the following hold.*

(i)

$$R_\epsilon\left(f^{\oplus,\mathcal{M}}, K\right) \geq \Omega\left((R_{2\epsilon}(f) - 2) \cdot \frac{d(\mathcal{M})}{\sqrt{k} \cdot \log k \cdot \log(nk)}\right).$$

(ii)

$$R_\epsilon\left(f^{\oplus,\mathcal{M}}, G, K, \Sigma\right) \geq \Omega\left((R'_{2\epsilon}(f) - 2) \cdot \frac{d(\mathcal{M})}{\log k \cdot \log(nk)}\right),$$

where $R'_\delta(f)$ is the optimal two-party communication complexity of f with randomized protocols that err with probability δ that can be proved via distributional complexity on product distributions.

Distributed OR/AND: It is natural to consider what happens if one replaces the XOR function in Theorem I.2 with OR or AND. More precisely, define: $f^{\vee,\mathcal{M}} \equiv \bigvee_{(u,v)\in\mathcal{M}} f(X_u, X_v)$ and $f^{\wedge,\mathcal{M}} \equiv \bigwedge_{(u,v)\in\mathcal{M}} f(X_u, X_v)$, where $\mathcal{M}$ is a disjoint pairing of the terminals in K. Unlike in the XOR case, it turns out that the bounds depend on f in a more delicate manner. For instance, it can be shown that $R\big(\mathrm{NEQ}^{\vee,\mathcal{M}}\big)$ and $R\big(\mathrm{EQ}^{\wedge,\mathcal{M}}\big)$ are both roughly the cost of the minimal steiner tree connecting the terminal nodes of K. We leave the details of this result for the full version. On the other hand, we also show the following topology sensitive bounds:

Theorem I.3. *Let $\mathcal{M}$ be a disjoint pairing of the terminal nodes of G that each hold an n-bit string with $n \geq 2$. Then, $R_\epsilon\big(EQ^{\vee,\mathcal{M}}\big)$ is $\Omega\big(d(\mathcal{M})/\log k\big)$ and there exists a randomized protocol with worst case-cost $O\big(d(M)\log k\big)$.*

Of particular interest for applications to graph problems are $\mathrm{DISJ}^{\vee,\mathcal{M}}$ and $\mathrm{DISJ}^{\wedge,\mathcal{M}}$. We establish the following:

Theorem I.4. *Let each of the terminal nodes in K hold an n-bit string and $\mathcal{M}$ be a disjoint pairing of these nodes. Then,*

(i)

$$R_\epsilon\big(DISJ^{\vee,\mathcal{M}}\big) = \Omega\left(\frac{d(M)}{\log(k)} \cdot n\right).$$

(ii)

$$R_\epsilon\big(DISJ^{\wedge,\mathcal{M}}\big) = \Omega\left(\frac{d(M)}{\log(k)} \cdot n\right).$$

[2] We remark that for the randomized version of the linear array conjecture stated earlier, using cuts would immediately yield $R_\epsilon(f, P_k, \{1, k+1\}, \Sigma) = \Omega(kR_\epsilon(f, P_2, \{1,2\}, \Sigma))$.

While we provide the argument of Theorem I.2 later, the proofs of Theorems I.3 and I.4 can be found in the full version.

Graph problems: Using some of the above results, we extend the lower bounds on graph problems considered in [2] to the general topology case. In particular, in these problems the k terminals get k subgraphs of some graph H (edges can be duplicated) and the terminals want to solve one of the following five problems: determining (i) the degree of a vertex in H, (ii) if H is acyclic; (iii) if H is triangle-free; (iv) if H is bipartite and (v) if H is connected. In all these cases we show that the trivial algorithm of all terminals sending their subgraphs to a designated terminal is the best possible up to an $O(\log k)$ factor. Our reductions for (ii)-(v) are different from those in [2] as the hard problem in [2] can be solved with $O(kn)$ communication for any topology.

A. The organization of the rest of the paper

In Section II we present the full proof of Theorem I.1. We give an overview of our method to prove lower bounds in Section II-A and then compare our techniques with those of existing work in Section II-B. We present the proof of Theorem I.2 in Section III. Section IV presents some of our reductions that prove our lower bounds for graph problems. All omitted proofs can be found in the full version of the paper [26].

II. An example: element distinctness

In this section we analyze the Element Distinctness problem and prove Theorem I.1. The proof will point to a general methodology for showing lower bounds on the communication cost of distributed protocols. We summarize the main features of this methodology at the end of the section and compare our techniques with those of existing related work.

Upper bounds: The upper bounds follow using standard randomized protocols to test equality of strings.

(i) Let $v \in V(G)$ be a median for K. Each terminal node in K computes a random ℓ-bit hash of its input treating the shared random string as a hash function. These fingerprints are sent to v, where it is verified that they are all distinct. By the union bound, the probability of error is at most $\binom{|K|}{2}2^{-\ell}$. Thus, if ℓ is chosen to be around $2\log|K| + \log(\frac{1}{\epsilon})$, this quantity is at most ϵ.

(ii) It suffices to show that there is a randomized protocol that (i) always terminates declaring success or failure after communicating $O(\mathrm{diam}_G(K)n + \sigma_G(K)\log|K|)$ bits, (ii) on all inputs, with probability at least $\frac{1}{2}$ declares success (iii) when it declares success, gives the correct answer to the problem, but when it declares failure, commits to no answer to the problem. (Repeating this protocol until success is obtained, yields a protocol that terminates with probability 1, and always gives the correct answer; the expected communication of the new protocol is at most twice the original protocol's worst case communication.) We start with the protocol for part (i) with $\epsilon = \frac{1}{2}$. If all hashes are different we declare that the elements are distinct. If some two fingerprints match, the corresponding n-bit inputs at the two terminals are compared (at one of the terminals, say); if the inputs match, we declare that the elements are not distinct; if the inputs don't match, we declare failure but give no answer. On all inputs, this protocol terminates successfully with probability at least $\frac{1}{2}$, but its answer is never wrong.

Lower bounds: Many of the key ideas of the paper will already be encountered in this section. Let us first gather our tools.

Metric embeddings: Consider an unweighted network G. A cut in G is a pair (V_A, V_B) of disjoint non-empty sets of vertices such that $V_A \cup V_B = V(G)$. We say that a pair of vertices $\{u, v\}$ is cut by C, and write $\{u, v\} \in C$, if $(u, v) \in (V_A \times V_B) \cup (V_B \times V_A)$. We need a collection of cuts $\mathcal{C}$ such that each pair of nodes of G is cut by roughly as many $C \in \mathcal{C}$ as the distance between them.

Definition II.1. *Let K be a subset of vertices of a graph G. A collection (multiset) $\mathcal{C}$ of cuts is said to be an (α, β)-family for K, if the following two conditions hold.*

(i) For $u, v \in K$, $|\{C \in \mathcal{C} : \{u, v\} \in C\}| \geq \alpha \cdot d(u, v)$;

(ii) For all $\{u, v\} \in E(G)$, $|\{C \in \mathcal{C} : \{u, v\} \in C\}| \leq \beta$.

The stretch of $\mathcal{C}$ is the minimum (taken over all choices of α and β satisfying (i) and (ii)) of the quantity $\frac{\beta}{\alpha}$.

Approximating distances in graphs using cuts in this manner is a central tool in the theory of metric embeddings [24]; it has numerous applications in network design and optimization. For us its value stems from the following observation.

Proposition II.2. *Let G be an undirected graph, and let $g : E(G) \to \mathbb{R}^+$. Suppose $\mathcal{C}$ is an (α, β)-family of cuts for G. Let $h : \mathcal{C} \to \mathbb{R}^+$ be such that for all $C \in \mathcal{C}$, we have $\sum_{e \in C} g(e) \geq h(C)$. Then, $\sum_e g(e) \geq \frac{1}{\beta}\sum_C h(C)$.*

In our applications, $g(e)$ will be the expected number of bits exchanged along the edge e; and $h(C)$ will be the complexity of an appropriate two-party communication problem induced by C. This allows us to derive lower bounds for general distributed protocols from two-party communication lower bounds. The loss in translation from the graph protocol to the two-party problem will be governed by the quantity $\frac{\beta}{\alpha}$. It is, therefore, desirable to obtain families of cuts with small stretch. The celebrated results of Bourgain [23] and Linial, London and Rabinovich [24] provides us a general tool for obtaining such families.

Lemma II.3. *In every graph G for every non-empty $K \subseteq V(G)$, there is a family $\mathcal{C}$ of cuts for K with stretch $O(\log|K|)$.*

Two-party communication complexity: To apply Proposition II.2 we need $h(C)$ to be a lower bound on the expected communication across C. We suggested earlier that $h(C)$ could be derived from an appropriate two-party communication complexity problem induced by C. We now make this association explicit. Let $C = (V_A, V_B)$. Fix a protocol Π for $(\mathrm{ED}, G, K, \{0,1\}^n)$ and a distribution on its inputs. We will imagine that the inputs corresponding to vertices in V_A are with Alice and the inputs corresponding to the vertices in V_B are with Bob. They wish to determine if the inputs are all distinct. One strategy available to Alice and Bob is to simulate the protocol Π and return the answer it gives. To simulate the protocol Π, Alice performs computations that take place at vertices in V_A and Bob performs all computations that take place at vertices in V_B. They share the messages that travel across the edges cut by C. In particular, the resulting two-party protocol will have expected communication exactly equal to the expected communication across the cut C in Π. (This can be formally justified; we omit the details.) Thus, one candidate for $h(C)$ is the expected communication cost of the best two-party communication protocol under the given distribution. Before we proceed, we must do the following: (i) identify a good distribution to use; (ii) show that the corresponding two-party problem is hard to solve for all cuts under this distribution. The two lower bounds for $(\mathrm{ED}, G, K, \{0,1\}^n)$ claimed in Theorem I.1 will use different distributions. For the first lower bound the following natural distribution μ_0 will be effective: all input nodes receive distinct strings, each of the $\binom{2^n}{k}$ possibilities being equally likely. For the second lower bound, we will use μ_1 defined as follows. Let $u, v \in K$ such that $d(u,v) = \mathrm{diam}_G(K)$. Now define μ_1 to be the uniform distribution on those inputs where all terminals other than u and v are assigned distinct values, and u and v are assigned a common value distinct from the others. We are thus led to a two-party communication problem of the following form. Alice receives a sequence S_A of a distinct elements of $\{0,1\}^n$ and Bob a sequence S_B of b distinct elements of $\{0,1\}^n$. They need to determine if $S_A \cap S_B = \emptyset$.

Lemma II.4. *(i) Suppose S_A and S_B are chosen uniformly at random and disjoint from each other; we call this distribution also μ_0. Let $a \leq b$ and $2^n \geq 4(a+b)+1$. Let Π be a two-party ϵ-error randomized communication protocol for the above two-party problem (thus, it computes the answer correctly with probability at least $1-\epsilon$ on all inputs). Then,*

$$\mathop{\mathbb{E}}_{(S_A,S_B)\sim\mu_0}[|\Pi(S_A,S_B)|] = \Omega(\min\{a,b\}).$$

(ii) Suppose S_A and S_B are as above, but chosen so that their first elements are equal but they are otherwise disjoint; call this distribution μ_1. Let $a, b \geq 1$, $2^n \geq a+b$, and let Π be a zero-error protocol for the above two-party problem. Then,

$$\mathop{\mathbb{E}}_{(S_A,S_B)\sim\mu_0}[|\Pi(S_A,S_B)|] \geq \log(2^n - (a+b) + 2).$$

We will justify this later. With this we are now ready to establish the lower bounds claimed in Theorem I.1.

Proof: (of Theorem I.1)

(i) Considered an ϵ-error randomized protocol for $(\mathrm{ED}, G, K, \{0,1\}^n)$. Let $\mathcal{C}$ be an (α,β)-family of cuts with stretch $O(\log|K|)$ promised by Lemma II.3. Assume the inputs are chosen according to the distribution μ_0 defined above. We apply Proposition II.2 taking $g(e)$ as the expected number of bits exchanged across edge e and taking $h(C)$ to be the lower bound on the two-party communication problem given by Lemma II.4. (We use linearity of expectation, where the expectations are over the choice of inputs from μ_0 and the random choices implicit in Π.) For a cut $C = (V_A, V_B)$, let $K_A = V_A \cap K$ and $V_B = V_B \cap K$. We conclude that

$$\begin{aligned}
\mathop{\mathbb{E}}_{X\sim\mu_0}[|\Pi(X)|] &= \sum_{e\in E(G)} g(e) \\
&\geq \frac{1}{\beta}\sum_{C\in\mathcal{C}} h(C) \\
&\geq \frac{1}{\beta}\sum_{(V_A,V_B)\in\mathcal{C}} \Omega(\min(|K_A|,|K_B|)) \\
&\geq \frac{1}{\beta k}\sum_{(V_A,V_B)\in\mathcal{C}} \Omega(|K_A||K_B|).
\end{aligned}$$

Note that

$$\begin{aligned}
\sum_{(V_A,V_B)\in\mathcal{C}} |K_A||K_B| &= \sum_{\{u,v\}\subseteq K} |\{C\in\mathcal{C} : \{u,v\}\in C\}| \\
&\geq \sum_{\{u,v\}\subseteq K} \alpha\cdot d(u,v) \\
&= \frac{\alpha}{2}\sum_{u\in K}\sigma_K(u).
\end{aligned}$$

Therefore, $\mathbb{E}[|\Pi(X)|] \geq (\frac{\alpha}{2\beta k})\sum_{u\in K}\Omega(\sigma_K(u)) \geq (\frac{\alpha}{2\beta})\Omega(\sigma_G(K))$, and $\mathbb{E}[|\Pi(X)|] = \Omega(\sigma_G(K)/\log k)$.

(ii) Since $R_0(\mathrm{ED}, G, K, \{0,1\}^n) \geq R_\epsilon(\mathrm{ED}, G, K, \{0,1\}^n)$ for all $\epsilon \geq 0$, we conclude from part (i) that $R_0(\mathrm{ED}, G, K, \{0,1\}^n) = \Omega(\sigma_G(K)/\log k)$. It thus suffices to show that $R_0(\mathrm{ED}, G, K, \{0,1\}^n) \geq \mathrm{diam}_G(K)n$. For this, we will use a different family of cuts and a different distribution. Let $u, v \in K$ such that $d(u,v) = \mathrm{diam}_G(K)$. For $i = 0, 1, \ldots, \mathrm{diam}_G(K)-1$, let $V_i = \{w \in V : d(u,w) \leq i\}$, and $C_i = (V_i, V \setminus V_i)$. Consider the family $\mathcal{C} = \{C_0, C_1, \ldots, C_{d(u,v)-1}\}$ of cuts. Now we use μ_1 defined above, under which all terminals other than u and v are assigned distinct values, and u and v are assigned a common value

distinct from the others. Since no edge appears in more than one cut, arguing as above, we obtain from Proposition II.2 and Lemma II.4 that

$$\mathop{\mathbb{E}}_{\mu_1}[|\Pi|] \geq \sum_{i}^{d(u,v)} \log(2^n - k + 2) = \mathrm{diam}_G(K) \log(2^n - k + 2).$$

■

It remains to establish Lemma II.4.

Proof: (of Lemma II.4)

(i) Now, suppose $a \leq b$ and $N = 2^n \geq 4(a+b)$. Identify $[N]$ with $\{0,1\}^n$. We may assume that S_A and S_B are sets instead of sequences. Thus S_A and S_B are subsets of $[N]$ chosen uniformly at random such that $|S_A| = a$, $|S_B| = b$ and $S_A \cap S_B = \emptyset$. Suppose there is an ϵ-error two-party randomized protocol Π for the set disjointness problem (its error probability is bounded by ϵ for all inputs (S_A, S_B) even outside the support of μ_0). We may assume (by repeating the protocol) that $\epsilon \leq \frac{1}{8}$. Let

$$\mathop{\mathbb{E}}_{(S_A,S_B)\sim\mu_0}[|\Pi(S_A, S_B)|] = \ell. \tag{1}$$

We wish to show that $\ell = \Omega(a)$.

We will appeal to the following result of Razborov [27]. Let $N' = 4a + 1$. Define distributions ρ_0 and ρ_1 on pairs of subsets (T_A, T_B) of $[N']$ as follows. In ρ_0, the the sets T_A and T_B are disjoint a-element subsets of $[N']$ chosen at random; in ρ_1, the sets T_A and T_B are a-element subsets of $[N']$ chosen at random with the condition that they intersect at exactly one place. Let $\rho = \frac{3}{4}\rho_0 + \frac{1}{4}\rho_1$. Then, there is a constant $\alpha > 0$ such that for all randomized protocols Γ for the set disjointness problem with worst-case communication less than αa,

$$\Pr_{(T_A,T_B)\sim\rho}[\Gamma \text{ errs on } (T_A, T_B)] \geq \frac{1}{4}.$$

Now consider the following protocol Γ for sets generated according to the distribution ρ.

Input: Alice receives T_A and Bob receives T_B.

Step 1: Let π be a random permutation of $[N]$ shared by Alice and Bob. Let $S_A = \pi(T_A)$ and $S_B = \pi(T_B \cup \{N'+1, \ldots, N'+(b-a)\})$ (note $N \geq N' + b - a$). Note that $T_A \cap T_B = \emptyset$ iff $S_A \cap S_B = \emptyset$; furthermore, if T_A and T_B are disjoint, (S_A, S_B) have distribution precisely μ_0.

Step 2: Alice and Bob simulate $\Pi(S_A, S_B)$ but limit the communication to 6ℓ bits. If Π is terminated before it finishes naturally, we say 'Yes, the sets intersect'; otherwise, we return the answer Π returns.

One can verify that Γ errs only when Π errs or when Π is terminated prematurely and $S_A \cap S_B = \emptyset$. Thus, our assumption (1) and Markov's inequality imply

$$\begin{aligned}\Pr[\Gamma \text{ errs }] &\leq \frac{1}{8} + \Pr[S_A \cap S_B = \emptyset] \\ &\quad \cdot \Pr[|\Pi| > 6\ell \mid S_A \cap S_B = \emptyset] \\ &< \frac{1}{8} + \frac{3}{4} \times \frac{1}{6} = \frac{1}{4}.\end{aligned}$$

It follows that $6\ell \geq \alpha a$, that is, $\ell = \Omega(a)$.

(ii) Let X be the first element of S_A and S_B, when $(S_A, S_B) \sim \mu_1$. Let Y be values at the other locations of S_A and S_B. Let Π have expected cost c. We first fix Y and the public randomness of Π so that the resultant deterministic protocol, denoted by Π_D, has expected cost c. Using the rectangle property of transcripts we conclude that for no two different values of X the transcripts of Π_D are the same. Using a well-known property of such prefix-free transcripts, we conclude that $c \geq \log(2^n - a - b + 2)$.

■

A. The outline of the method

The above proof for element distinctness exemplifies the following natural methodology. Let (F, G, K, Σ) be a distributed computing problem. Each cut $C = (V_A, V_B)$ induces a two-party communication complexity problem F_C as follows: Alice's input corresponds to the input $x_A \in \Sigma^{K_A}$ and Bob receives the input $x_B \in \Sigma^{K_B}$, where $K_A = K \cap V_A$ and $K_B = K \cap V_B$. The function they wish to compute is $F_C(x_A, x_B) = F(x)$, where x is the input for the vertices in K obtained by combining x_A and x_B.

Cuts: Intuitively, the main contributors to the complexity of the problem are the distances between the inputs. As in the above example, we will define a suitable family of cuts that captures this. In most cases, we will use the cuts promised by Lemma II.3.

Distribution: As stated in the introduction, the difficulty of the problem across any one cut is often not hard to establish based on results already available. However, in order to use of Proposition II.2, we need to somehow show inputs that are hard simultaneously across all cuts. This will be achieved by arguing not in terms of worst-case bounds but expected bounds, for which it will be crucial to obtain the right distribution.

Two-party communication complexity: Suppose we have defined the distribution μ on inputs x for F. For each cut C, this naturally defines a distribution on the inputs to F_C; we refer to this distribution as μ_C. Standard results from two-party communication complexity will be deployed here. These results will also suggest the right global distribution that we can impose on the inputs to the problem.

Putting it together: This will be straightforward, and we will just use Proposition II.2 in the following form.

Proposition II.5. $R_\epsilon(F,K) \geq \frac{1}{\beta}\sum_{C\in\mathcal{C}} \mathrm{ED}_{\mu_C,\epsilon}(F_C)$*, where* $\mathcal{C}$ *is the* (α,β)*-family of cuts with stretch* $\gamma = \beta/\alpha = O(\log k)$ *obtained from Lemma II.3, and* $\mathrm{ED}_{\mu_C,\epsilon}(F_C)$ *is the minimum expected communication (averaged over inputs drawn from* μ *and the coin tosses of the protocol) of a protocol* Γ_C *for* F_C *that errs with probability at most* ϵ *for* all *inputs. Note that* Γ_C *is required to err with probability at most* ϵ *even for inputs not in the support of* μ_C*.*

In subsequent sections, we will use this method to derive lower bounds for a variety of problems.

B. Connections to Related Work

Finally, we put our techniques in the context of existing techniques used to argue lower bounds (most of which are for the case when G is fully connected or almost equivalently when G is a star graph). In particular, we will argue that our techniques essentially generalize many of the existing techniques.

The first lower bounds for the message-passing Number in Hand (NIH) model seems to be due to Duris and Rolim [4]. They also considered the complete graph topology (co-ordinator model) and their bounds were for deterministic and non-deterministic complexity. In particular it uses a generalization of 2-party fooling set argument that does not seem to apply to bounded error randomized protocols. Very recently, the symmetrization technique was introduced by Phillips et al. [1] and was further developed by Woodruff and Zhang [5], [2], [28]. At a very high level, the core idea in symmetrization is as follows. First we consider the case when G is a star graph of diameter 2 with a co-ordinator node at the center. Prototypical hard problems to consider are functions of the form $\vee_{i=1}^{k} f(X_i, Y)$, where the center gets Y and the k leaves of G get $X_1, \ldots, X_k$ (i.e. in this case $K = V(G)$). If ν is a hard distribution for (the 2-party) function f, then the trick is to define a hard distribution μ on $X_1, \ldots, X_k, Y$ such that for every $i \in [k]$ the effective distribution on (X_i, Y) is ν. Then the argument, slightly re-phrased in our language, proceeds as follows: pick a random cut among the k cuts corresponding to the k edges. Then by definition of μ the induced 2-party problem across each cut is f and hence, the communication complexity is $\Omega(R(f))$, where $R(f)$ is the randomized two-party communication complexity of f. Then we note that since the cut was picked completely at random, and the distribution μ is symmetric with respect to the leaf-nodes, the communication across such a random cut in expectation is $\Theta(1/k)$ of the total communication, which leads to an overall $\Omega(k \cdot R(f))$ lower bound on the total communication. By contrast, our technique does not need this symmetric property though our use of linearity of expectation seems similar. Indeed, we show how to recover the lower bound on the OR of f from [2] using our techniques in the full version of the paper. Note that the cuts in a star-graph are all similar, as all leaves are symmetric with respect to the prototypical example. As identified by the authors [1] themselves, this property seems to be lost even for star graphs when the inputs held by leaf-nodes are not symmetric with respect to the function that players want to compute. For general graph topology, there might be very little symmetry left. In particular, in our technique, the cuts obtained are arbitrary with no guarantee of symmetry. Nevertheless, our technique seems flexible enough to handle such cases.

One technique that we cannot (yet) handle with ours is the result of Braverman et al. [6] that proves a lower bound of $\Omega(kn)$ on the set disjointness problem on the star graph. It is an interesting open question to see if we can port the techniques of Braverman et al. to our setting.

As mentioned in the introduction, Tiwari's work [20] did consider the case when G is not well connected. His lower bounds for graphs other than the line graph were obtained by embedding the graph G on to a line graph. In our terminology, the set of cuts considered by Tiwari were the layers of a BFS. For the simple cases of G being a grid or a cycle, these set of cuts are a $(1,1)$ family of cuts for G. In our work we use the full power of metric embedding to handle general topologies in G. One can think of our use of metric embeddings as embedding the general graph G into suitable subgraphs of a grid graph. The other difference is that Tiwari considered deterministic communication complexity. This seems to present some difficulties when applying to cuts in the following sense: only 2-party lower bounds proved using certain techniques like the fooling set or rank arguments could be employed. The randomized protocol setting of our work appears to be more flexible. When working across a cut, our method seems oblivious of the particular technique used to derive the relevant 2-party bound on the expected communication.

We conclude this discussion by pointing out another area where cuts have been used to prove lower bounds on communication. In the network coding literature one of the outstanding open questions is the so called Li and Li conjecture [29], [30], [31], which asks whether the gap between the trivial upper bound and the cut based lower bound is necessary. In fact, like our work this gap of $\Theta(\log k)$ is due to the use of metric embedding results. However, the cut-based lower bound in network coding seems more relevant to proving a lower bound for the version of our problem where we are interested in minimizing the total number of rounds. Further, in network coding, the main goal seems to be transmitting a set of bits from a source to a sink. By contrast, we deal with function computation in the spirit of classical communication complexity.

III. Distributed XOR Lemma

As observed in the introduction, the naive protocol gives

$$R(f^{\oplus,\mathcal{M}}) = d(\mathcal{M})R_{\epsilon/|\mathcal{M}|}(f) = O(d(\mathcal{M})R_\epsilon(f)\log k).$$

In this section, we show the lower bound claimed in Theorem I.2. Our proof will be a straightforward consequence of a result of Barak et al. [32], which we now present.

Recall that $f : \Sigma \times \Sigma \to \{0,1\}$. For $m \geq 1$, let $f^{\oplus m} : \Sigma^m \times \Sigma^m \to \{0,1\}$, where $f^{\oplus m}((X_1,\ldots,X_m),(Y_1,\ldots,Y_m)) = \bigoplus_i f(X_i,Y_i)$. Fix a distribution distribution ν on $\Sigma \times \Sigma$, and let

$$\mathrm{ED}_{\nu^m,\frac{1}{20}}(f^{\oplus m}) = \mathop{\mathbb{E}}_{\nu^m}[R_{\frac{1}{20}}(f^{\oplus m})],$$

and let $D_{\nu,\frac{1}{5}}(f)$ be the $(\frac{1}{5})$-error distributional complexity of f under ν (that is, the minimum worst-case communication over all deterministic protocols for f whose error probability is at most $\frac{1}{5}$ when inputs are drawn from ν).

Theorem III.1 (Barak et al. [32])**.** *(i)* $\mathrm{ED}_{\nu^m,\frac{1}{20}}(f^{\oplus m})\log(20ER(f^{\oplus m})) = \Omega(\left((D_{\nu,\frac{1}{5}(f)}-2)\sqrt{m}\right)$.

(ii) If ν itself is a product distribution on $\Sigma \times \Sigma$, then $\mathrm{ED}_{\nu^m,\frac{1}{20}}(f^{\oplus m})\log(20\mathrm{ED}(f^{\oplus m})) = \Omega(\left((D_{\nu,\frac{1}{5}(f)}(f)-2)\cdot m\right)$.

Remark 1. *The LHS of both (i) and (ii) differ from those in Theorem 2.8 and Theorem 2.9 of [32], because we define the complexity of a randomized protocol to be its expected communication cost instead of its worst-case communication cost. Our claims follow from theirs by a straightforward application of Markov's inequality.*

With this, we can justify Theorem I.2.

Proof: We will use the methodology outlined in the Section II-A. Let $F = f^{\oplus,\mathcal{M}}$. Let $\mathcal{C}$ be a (α,β)-family of cuts for K with stretch $\gamma = \beta/\alpha = O(\log k)$ obtained from Lemma II.3.

The next step requires defining a distribution on Σ^K. Let ν be the distribution on $\Sigma\times\Sigma$ such that $D_{\nu,\frac{1}{5}}(f)$ is maximized. Our distribution μ on Σ^K will be obtained by placing one copy of ν on each pair in $\mathcal{M}$. That is, if $K=[k]$ and $\mathcal{M} = \{(i_1,j_1),(i_2,j_2),\ldots,(i_{k/2},j_{k/2})\}$, then for each $X \in \Sigma^k$, we have

$$\mu(X) = \nu(X_{i_1},X_{j_1})\nu(X_{i_2},X_{j_2})\cdots\nu(X_{i_{k/2}},X_{j_{k/2}}).$$

The third step requires finding a lower bound on $\mathrm{ED}(F_C)$ for each cut $C \in \mathcal{C}$.

Claim III.2. *Suppose the cut C separates ℓ pairs in $\mathcal{M}$.*

(i) $\mathrm{ED}_{\mu_C,\frac{1}{20}}[F_C]\log(20\mathrm{ED}_{\mu_C,\frac{1}{20}}[F_C]) = \Omega\left((D_{\nu,\frac{1}{5}}(f)-2)\sqrt{\ell}\right)$.

(ii) If ν is itself a product distribution, then $\mathrm{ED}_{\mu_C,\frac{1}{20}}[F_C]\log(20\mathrm{ED}_{\mu_C,\frac{1}{20}}[F_C]) = \Omega\left((D_{\nu,\frac{1}{5}}(f)-2)\cdot\ell\right)$.

We will justify this claim later. Let us first see how this implies our theorem. Consider part (i) of our theorem. We have from Proposition II.5 that

$$R_{\frac{1}{20}}(F,K) \geq \frac{1}{\beta}\sum_{C\in\mathcal{C}}\mathrm{ED}_{\mu_C,\frac{1}{20}}[F_C].$$

Multiplying both sides by $\log(20R_{\frac{1}{20}(F,K)}) \geq \log(20\mathrm{ED}_{\mu_C,\frac{1}{20}}[F_C])$ gives

$$\begin{aligned} R_{\frac{1}{20}}(F,K)&\log(20R_{\frac{1}{20}(F,K)}) \\ &\geq \frac{1}{\beta}\sum_{C\in\mathcal{C}}\mathrm{ED}_{\mu_C,\frac{1}{20}}[F_C](\log 20\mathrm{ED}_{\mu_C,\frac{1}{20}}[F_C]) \\ &\geq \frac{1}{\beta}\sum_{C\in\mathcal{C}}\tilde{\Omega}\left((D_{\nu,\frac{1}{5}}(f)-2)\sqrt{\ell_C}\right), \end{aligned}$$

where ℓ_C is the number of pairs of $\mathcal{M}$ separated by C and we use the first part of the above claim to justify the last inequality. Note that $\sum_{C\in\mathcal{C}}\sqrt{\ell_C} \geq \sum_{C\in\mathcal{C}}\ell_C/\sqrt{k} \geq \alpha\cdot d(\mathcal{M})/\sqrt{k}$. Thus,

$$R_{\frac{1}{20}}(F,K)\log(10R_{\frac{1}{20}}(F,K)) = \Omega\left((D_{\nu,\frac{1}{5}}(f)-2)d(\mathcal{M})/(\gamma\sqrt{k})\right).$$

Yao's lemma (strong duality) implies that for our choice of ν, $D_{\nu,\frac{1}{5}}(f) \geq R_{\frac{1}{5}}(f)$. Thus,

$$R_{\frac{1}{20}}(F,K)\log(10R_{\frac{1}{20}(F,K)}) = \Omega\left((R_{\frac{1}{5}}(f)-2)d(\mathcal{M})/(\gamma\sqrt{k})\right).$$

The proof of the second part is similar, so we omit it; but we still need to verify the claim above. Again, the two-parts are similar and we present our justification only for the first part. Fix C and a randomized protocol Γ_C for F_C, and write ℓ for ℓ_C. Let $\mathcal{M} = \{(i_1,j_1),(i_2,j_2),\ldots,(i_{k/2},j_{k/2})\}$, and let the ℓ pairs of $\mathcal{M}$ cut by C be $(i_1,j_1),\ldots,(i_\ell,j_\ell)$. We will now derive a protocol Γ for the function $f^{\oplus\ell}$.

Step 1: Alice receives $X = (X_1,X_2,\ldots,X_\ell)$ and Bob receives inputs $Y = (Y_1,Y_2,\ldots,Y_\ell)$.

Step 2: Alice and Bob construct inputs $X' \in \Sigma^K$ for Γ_C as follows. First we deal with the pairs cut by C: we set $(X'_{i_r},X'_{j_r}) = (X_r,Y_r)$ for $r = 1,2,\ldots,\ell$. For $r = \ell+1,\ldots,k/2$, we choose (X'_{i_r},X'_{j_r}) according to ν, independently for different j.

Step 3: Alice and Bob simulate $\Gamma_C(X')$; let the answer it returns be z. Then Γ returns $z \oplus \bigoplus_{r=\ell+1}^{k/2} f(X'_{i_r},X'_{j_r})$ ($f(X'_{i_r},X'_{j_r})$ is available in their shared randomness).

Now, Γ errs in computing $f^{\oplus\ell}(X,Y)$ only if Γ_C errs in computing $F_C(X')$. Also, if the inputs to Γ_C are drawn from ν^ℓ, then the inputs Alice and Bob present to Γ_C are drawn according to $\nu^{k/2}$. Thus, the expected communication between Alice and Bob is precisely $\mathbb{E}_\mu[|\Gamma_C|]$, and the

probability of error is at most $\frac{1}{20}$. Thus, our claim follows from Theorem III.1 (i). ■

IV. Lower Bounds for Graph Problems

In this section, we will consider the case when the k terminals are trying to compute a function about a graph $H = (V, E)$ that is distributed among the k terminals. In particular, for terminal $i \in K$, we will denote its subgraph by H_i. We study the graph based problems considered in [2] and show that in all of them the trivial algorithm where all terminals send their input to one terminal is the best possible algorithm (up to an $O(\log k)$ factor). In particular, these give topology dependent extensions to the corresponding results in [2]. In this section, we will not explicitly differentiate whether the edge sets of H_i are disjoint for every i or not. In what follows we will use $|H_i|$ to denote the size of H_i ($i \in K$).

We begin with a technical result that we will use:

Lemma IV.1. *For any graph G and subset of terminals K with even $k = |K|$, let $\mathcal{M}(K)$ denote the set of all disjoint pairings in K. Then for any $k \geq 2$,*

$$\frac{1}{2} \cdot \sigma_G(K) \leq \max_{M \in \mathcal{M}(K)} d(M) \leq \sigma_G(K).$$

In this section, we consider the following three problems.

Acyclicity: Given H_i to terminal $i \in K$, the terminals need to decide if H is acyclic or not.

Triangle-Freeness: Given H_i to terminal $i \in K$, the terminals have to decide if H has a triangle or not.

Bipartitiness: Given H_i to terminal $i \in K$, the terminals have to decide if H is bipartite or not.

We will show that for all of the problems above, the trivial algorithm is the best.

Theorem IV.2. *Each of the problems of acyclicity, triangle-freeness and bipartitenes on graph G needs $\Omega\left(\sigma_G(K) \cdot \frac{\max_i |H_i|}{\log k}\right)$ communication.*

Proof: We will prove the claimed lower bounds by showing the claimed lower bound on the problem where the terminals have to decide if (i) H is a forest vs (ii) H has a triangle. (Note that solving any of acyclicity, triangle-freeness or bipartitenes will determine which of the two cases H falls in.)

Consider the DISJ$^{\vee M}$, where we pick the pairing M so that it maximizes $d(M)$ (and hence, we can apply Lemma IV.1). For notational convenience let us assume that $K = [k]$ and that

$$M = \{(i, i') | i \in [k/2] \text{ and } i' = i + k/2\}.$$

Assume that for the DISJ$^{\vee M}$ problem, terminal $i \in [k]$ gets a set/vector $X_i \in \{0,1\}^n$. We now define the subgraphs H_i. To begin with we have

$$V = \cup_{j=1}^{k/2} U_j \cup \{w^1, \ldots, w^k\},$$

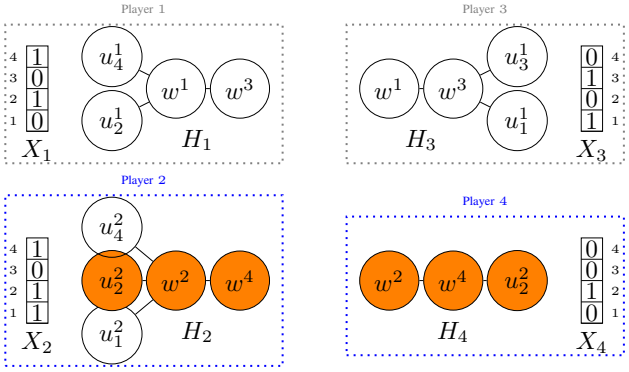

Figure 1. Illustration of the reduction in proof of Theorem IV.2 for $n = k = 4$. The pairing $M = (1,3), (2,4)$ is denoted by the paired terminals having the same colored boxes. In this example the overall graph $H = \cup_{i=1}^{4} H_i$ has a triangle and the three participating nodes are colored in orange.

where

$$U_j = \{u_1^j, \ldots, u_n^j\}.$$

For every $p \in [k]$, H_p has the following edges: $(w^p, w^{p'})$ and $(w^p, u_j^{p'})$ for every $j \in [n]$ such that $X_p(j) = 1$, where $p' = p$ if $p \leq k/2$ and $p' = p - k/2$ otherwise. See Figure 1 for an illustration of this reduction.

Note that each terminal p can construct its subgraph just from its X_p and that H is in case (i) if DISJ$^{\vee M}(X_1, \ldots, X_k) = 0$ and is in case (ii) otherwise. Theorem I.4 along with the lower bound in Lemma IV.1 and the fact that for every $i \in [k]$, $|H_i| = \Theta(n)$ completes the proof. ■

In the full version of the paper [26] we present analogous lower bounds for determining the degree of a vertex in H and checking if H is connected or not.

Acknowledgment

We thank the anonymous reviewers for their comments, which helped improve the presentation of the paper. Special thanks to the reviewer who pointed out [20] to us. We thank the organizers of the 2012 seminar on Algebraic and Combinatorial Methods in Computational Complexity at Dagstuhl for providing the venue for the initial discussion of the problems considered in this paper. We thank Anupam Gupta for answering questions on metric embeddings. Thanks also to David Woodruff and Qin Zhang for answering our questions on their paper [2].

AC's research is supported in part by a Ramanujan Fellowship of the DST. AR's research is supported in part by NSF grant CCF-0844796.

References

[1] J. M. Phillips, E. Verbin, and Q. Zhang, "Lower bounds for number-in-hand multiparty communication complexity, made easy," in *SODA*, 2012, pp. 486–501.

[2] D. Woodruff and Q. Zhang, "When distributed computation is communication expensive," in *DISC*, 2013, pp. 16–30.

[3] D. Dolev and T. Feder, "Multiparty communication complexity," in *FOCS*, 1989, pp. 428–433.

[4] P. Duris and J. Rolim, "Lower bounds on the multiparty communication complexity," *J.Comput.Syst.Sci.*, vol. 56, no. 1, pp. 90–95, 1998.

[5] D. Woodruff and Q. Zhang, "Tight bounds for distributed functional monitoring," in *STOC*, 2012, pp. 941–960.

[6] M. Braverman, F. Ellen, R. Oshman, T. Pitassi, and V. Vaikuntanathan, "A tight bound for set disjointness in the message-passing model," in *FOCS*, 2013, pp. 668–677.

[7] L. G. Valiant, "A bridging model for parallel computation," *Commun. ACM*, vol. 33, no. 8, pp. 103–111, 1990.

[8] H. J. Karloff, S. Suri, and S. Vassilvitskii, "A model of computation for mapreduce," in *SODA*, 2010, pp. 938–948.

[9] M. T. Goodrich, N. Sitchinava, and Q. Zhang, "Sorting, searching, and simulation in the mapreduce framework," in *ISAAC*, 2011, pp. 374–383.

[10] P. Beame, P. Koutris, and D. Suciu, "Communication steps for parallel query processing," in *PODS*, 2013, pp. 273–284.

[11] P. Koutris and D. Suciu, "Parallel evaluation of conjunctive queries," in *PODS*, 2011, pp. 223–234.

[12] M.-F. Balcan, A. Blum, S. Fine, and Y. Mansour, "Distributed learning, communication complexity and privacy," in *COLT*, 2012, pp. 26.1–26.22.

[13] H. Daumé III, J. M. Phillips, A. Saha, and S. Venkatasubramanian, "Protocols for learning classifiers on distributed data," in *AISTATS*, 2012, pp. 282–290.

[14] A. Drucker, F. Kuhn, and R. Oshman, "The communication complexity of distributed task allocation," in *PODC*, 2012, pp. 67–76.

[15] G. Cormode, "The continuous distributed monitoring model," *SIGMOD Rec.*, vol. 42, no. 1, pp. 5–14, May 2013. [Online]. Available: http://doi.acm.org/10.1145/2481528.2481530

[16] S. Muthukrishnan, "Data streams: Algorithms and applications," *Foundations and Trends in Theoretical Computer Science*, vol. 1, no. 2, 2005.

[17] H. Kowshik and P. Kumar, "Optimal function computation in directed and undirected graphs," *IEEE Transcations on Information Theory*, vol. 58, no. 6, pp. 3407–3418, 2012.

[18] F. Kuhn and R. Oshman, "Dynamic networks: models and algorithms," *SIGACT News*, vol. 42, no. 1, pp. 82–96, 2011.

[19] A. Das Sarma, S. Holzer, L. Kor, A. Korman, D. Nanongkai, G. Pandurangan, D. Peleg, and R. Wattenhofer, "Distributed verification and hardness of distributed approximation," *SIAM Journal on Computing*, vol. 41, no. 5, pp. 1235–1265, 2012. [Online]. Available: http://dx.doi.org/10.1137/11085178X

[20] P. Tiwari, "Lower bounds on communication complexity in distributed computer networks," *J. ACM*, vol. 34, no. 4, pp. 921–938, 1987.

[21] E. Kushilevitz, N. Linial, and R. Ostrovsky, "The linear-array conjecture in communication complexity is false," *Combinatorica*, vol. 19, no. 2, pp. 241–254, 1999.

[22] M. Dietzfelbinger, "The linear-array problem in communication complexity resolved," in *STOC*, F. T. Leighton and P. W. Shor, Eds. ACM, 1997, pp. 373–382.

[23] J. Bourgain, "On lipschitz embedding of finite metric spaces in hilbert space," *Israel J. Math.*, vol. 52, no. 1-2, pp. 46–52, 1995.

[24] N. Linial, E. London, and Y. Rabinovich, "The geometry of graphs and some of its algorithmic applications," *Combinatorica*, vol. 15, no. 2, pp. 215–245, 1995.

[25] W. Goddard and O. R. Oellermann, "Distance in graphs," in *Structural Analysis of Complex Networks*, M. Dehmer, Ed. Springer Science + Business Media, 2011.

[26] A. Chattopadhyay, J. Radhakrishnan, and A. Rudra, "Topology matters in communication," *ECCC Tech report TR14-074*, 2014.

[27] A. Razborov, "On the distributional complexity of Disjointness," *Theor.Comput.Sci.*, vol. 106, no. 2, pp. 385–390, 1992.

[28] D. Woodruff and Q. Zhang, "An optimal lower bound for distinct elements in the message passing model," in *SODA*, 2014, pp. 718–733.

[29] Z. Li and B. Li, "Network coding in undirected networks," in *CISS*, 2004.

[30] ——, "Network coding: The case of multiple unicast sessions," in *Allerton Conference on Communications*, vol. 16, 2004.

[31] N. J. Harvey, R. D. Kleinberg, and A. R. Lehman, "Comparing network coding with multicommodity flow for the k-pairs communication problem," *MIT LCS Tech report MIT-LCS-TR-964*, 2004.

[32] B. Barak, M. Braverman, X. Chen, and A. Rao, "How to compress interactive communication," *SIAM J. Computing*, vol. 42, no. 3, pp. 1327–1363, 2013.