Routing Schemes and Distance Oracles in the Hybrid Model

Fabian Kuhn
Universität Freiburg, Germany

Philipp Schneider
Universitä Freiburg, Germany

Abstract

The HYBRID model was introduced as a means for theoretical study of *distributed* networks that use various communication modes. Conceptually, it is a synchronous message passing model with a *local communication mode*, where in each round each node can send large messages to all its neighbors in a local network (a graph), and a *global communication mode*, where each node is allotted limited (polylogarithmic) bandwidth per round to communicate with *any* node in the network.

Prior work has often focused on shortest paths problems in the local network, as their global nature makes these an interesting case study how combining communication modes in the HYBRID model can overcome the individual lower bounds of either mode. In this work we consider a similar problem, namely computation of *distance oracles* and *routing schemes*. In the former, all nodes have to compute *local tables*, which allows them to look up the distance (estimates) to any target node in the local network when provided with the *label* of the target. In the latter, it suffices that nodes give the next node on an (approximately) shortest path to the target.

Our goal is to compute these local tables as fast as possible with labels as small as possible. We show that this can be done *exactly* in $\widetilde{O}(n^{1/3})$ communication rounds and labels of size $\Theta(n^{2/3})$ bits. For constant stretch approximations we achieve labels of size $O(\log n)$ in the same time. Further, as our main technical contribution, we provide computational lower bounds for a variety of problem parameters. For instance, we show that computing solutions with stretch below a certain constant takes $\widetilde{\Omega}(n^{1/3})$ rounds even for labels of size $O(n^{2/3})$.

2012 ACM Subject Classification Theory of computation → Distributed algorithms; Mathematics of computing → Graph algorithms

Keywords and phrases Distributed Computing, Graph Algorithms, Complexity Analysis

Digital Object Identifier 10.4230/LIPIcs.DISC.2022.28

Related Version *Full Version*: https://doi.org/10.48550/arXiv.2202.06624

1 Introduction

Real networks often employ multiple communication modes. For instance, mobile devices combine high-bandwidth, short-range wireless communication with relatively low-bandwidth cellular communication (c.f., 5G [4]). Other examples are software defined networking [22] or hybrid data centers, which combine wireless and wired communication [11] or optical circuit switching and electrical packet switching [23].

In this article we utilize the theoretical abstraction of such hybrid communication networks provided by [5] which became known as *hybrid model* and was designed to reflect a high-bandwidth local communication mode and a low-bandwidth global communication mode, capturing one of the main aspects of real hybrid networks. Fundamentally, the hybrid model builds on the concept of *synchronous message passing*, a classic model to investigate round complexity in distributed systems.

36th International Symposium on Distributed Computing (DISC 2022).
Editor: Christian Scheideler; Article No. 28; pp. 28:1–28:22
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

▶ **Definition 1** (Synchronous Message Passing, c.f., [17]). *Let V be a set of n nodes with unique identifiers $ID : V \to [n] \overset{\text{Def}}{=} \{1, \ldots, n\}$. Time is slotted into discrete rounds consisting of the following steps. First, all nodes receive the set of messages addressed to them in the last round. Second, nodes conduct computations based on their current state and the set of received messages to compute their new state (randomized algorithms also include the result of some random function). Third, based on the new state the next set of messages is sent.*

Synchronous message passing has a clear focus on investigating round complexity, i.e., the number of communication rounds required to solve a problem with an input distributed over all nodes. For this purpose, nodes are usually assumed to be computationally unbounded. Occasionally this model is "overexploited", e.g., nodes are supposed solve $\mathcal{NP}$-complete problems on their local data, however we will refrain from that. The hybrid model places additional restrictions on the size of message and which nodes can exchange them.

▶ **Definition 2** (Hybrid model [5]). *The* $\mathsf{HYBRID}(\lambda, \gamma)$ *model is a synchronous message passing model (Def. 1), subject to the following restrictions.* Local mode: *nodes may send a message per round of maximum size λ bits to each of their neighbors in a connected graph.* Global mode: *nodes can send and receive messages of total size at most γ bits per round to/from any other node(s) in the network. If the restrictions are not adhered to, then a strong adversary selects the messages that are delivered.*

The parameter spectrum of the $\mathsf{HYBRID}(\lambda, \gamma)$ model covers the standard models LOCAL, $\mathsf{CONGEST}$, CLIQUE (aka "Congested Clique") and NCC ("Node Capacitated Clique", [5]) as marginal cases: LOCAL: $\lambda = \infty, \gamma = 0$, $\mathsf{CONGEST}$: $\lambda = O(\log n), \gamma = 0$, CLIQUE (+ Lenzen's routing protocol [16]): $\lambda = 0, \gamma = n \log n$, NCC: $\lambda = 0, \gamma = O(\log^2 n)$. Given the ramifications of investigating $\mathsf{HYBRID}(\lambda, \gamma)$ in its entirety, we narrow our scope (for our upper bounds) to a particular parametrization that pushes both communication modes to one extreme end of the spectrum. Following the arguments of [5] we leave the size of local messages unrestricted (modeling high local bandwidth) and allow only $\text{polylog}\, n$ bits of global communication per node per round (modeling severely restricted global bandwidth). Formally, we define the "standard" hybrid model as combination of the standard LOCAL and NCC models: $\mathsf{HYBRID} := \mathsf{HYBRID}\big(\infty, O(\log^2 n)\big)$. Our lower bounds are given for the more general $\mathsf{HYBRID}(\infty, \gamma)$ model, which implies lower bounds for the weaker HYBRID model.

A fundamental aspect of the Internet Protocol is packet forwarding, where every node has to compute a routing function, which – when combined with target-specific information stored in the packet header – must indicate the neighbor which the packet has to be forwarded to such that it reaches its intended destination. A correct *routing scheme* consists of these routing functions and a unique *label* per node, such that the packet forwarding procedure induces a path in the network from *any* source node to *any* destination node specified by the corresponding label attached to the packet.

We distinguish *stateful* and *stateless* routing schemes. In the former, routing can be based on additional information accumulated in the packet header as the packet is forwarded, whereas in the latter routing decisions are completely oblivious to the previous routing path. A related problem is the computation of distance oracles, which is similar to the all pairs shortest paths (APSP) problem. Each node must compute an oracle function that provides the distance (or an estimate) to any other node when provided with the corresponding label. Formal definitions are given in Section 1.3 (Def. 3, 4 and 5).

The first goal is to gather the required data for labels, routing and oracle functions in as few rounds as possible. This is particularly important for dynamic or unreliable networks where changes in distances or topology necessitates (frequent) re-computation. In this work we allow to relabel nodes (a.k.a. *labeling scheme*), where the new labels may contain information that helps with distance estimation and routing decisions. This gives rise to our second goal; keeping node labels small.[1] The third goal is to speed up the actual packet forwarding process to minimize latency and alleviate congestion. Given a graph with edge weights corresponding to (e.g.) link-latencies, we want to minimize the largest detour any packet takes relative to the corresponding shortest path. This is also known as *stretch*. Analogously, for distance oracles we want to minimize the worst estimation error relative to the true distance.

In this work we are interested in solving the above problems in a distributed setting (c.f., Definition 1). This has particular importance given the distributed nature of many real networks where routing problems are relevant (most prominently, the Internet) and where providing a centralized view of the whole network is prohibitively expensive. Given that usually vast quantities of data must be routed, we interested in computing routing schemes and distance oracles for the local communication network, which offers the majority of the throughput, whereas the throughput of global network is considered negligible. The global mode is used to speed up the computation of routing schemes (significantly, as we will see in this work) and can also be used to send the (relatively small) labels to a source, which can then be stored at that node for the duration of a session.

From an algorithmic standpoint, computing routing schemes and distance oracles is an inherently *global* problem. That is, allowing only local communication (i.e., the LOCAL model) it takes $\Omega(n)$ rounds to accomplish this.[2] A similar observation can be made for the global communication mode (NCC model). If we are only allowed to use global communication and each node initially only knows its incident edges in the local network, it takes $\widetilde{\Omega}(n)$ [3] rounds to compute routing schemes and distance oracles. This article addresses the question whether the combination of the two communication modes in the HYBRID model can overcome the $\widetilde{\Omega}(n)$ lower bound of the individual modes LOCAL and NCC.

Our answer to this is two-pronged. First we show that indeed, we can compute routing schemes and distance oracles significantly faster, for instance, $\widetilde{O}(n^{1/3})$ rounds and labels of size $\Theta(n^{2/3})$ suffice for an exact solution (c.f., Theorem 27). This separates these problems from the APSP problem, which can be used to solve our problems at hand, but has a $\widetilde{\Omega}(\sqrt{n})$ lower bound even for extremely crude approximations [5]. Second, we show that the HYBRID model is not arbitrarily powerful by giving polynomial lower bounds for these problems (depending on the stretch) that hold even for relatively large labels and unbounded local memory. For instance, we show that it takes $\widetilde{\Omega}(n^{1/3})$ rounds to solve either problem exactly, even for unweighted graphs and labels of size $O(n^{2/3})$ (c.f., Theorem 14). We provide nuanced results, depending on stretch and the type of problem, summarized in the following.

1.1 Contributions and Overview

Table 1 gives a simplified overview of our complexity results for the various forms of routing scheme and distance oracle problems. Our main contributions revolve around computational lower bounds for computing distance oracles and stateless and stateful routing schemes in

[1] In other settings, the amount of information stored at nodes for routing and distance estimation is considered, as well. Since the nodes in our model are computationally unbounded we do not focus on that. Our lower bounds have also no restriction on the local information.

[2] Note that any graph problem can be solved in $O(n)$ rounds in LOCAL by collecting the graph and solving the problem locally at some node. This makes global problems uninteresting for the LOCAL model, unless communication restrictions are increased (c.f., CONGEST) or decreased (c.f., HYBRID).

[3] The $\widetilde{O}(\cdot)$ notation suppresses multiplicative terms that are polylogarithmic in n.

the HYBRID model. Lower bounds for approximations are summarized in the first three groups of Table 1. We also consider lower bounds on unweighted graphs, see the first row of the fourth group of Table 1. These lower bounds hold regardless of the allowed local memory and hold for randomized algorithms with constant success probability.

Furthermore, we give upper bounds for the HYBRID model, summarized in the last two groups of Table 1. The implied computational upper bounds for the exact and approximate problem variants are almost tight with some of our corresponding lower bounds (more on that further below). In the following paragraphs we aim to give some intuitive understanding into how our techniques work by shedding most of the proof-details and occasionally pointing out how the results are generalized in the main part. We also show how the techniques used in this work relate to – and are distinct from – prior work.

In the main part, the results on lower bounds are formulated for the more general HYBRID(∞, γ) model (which therefore hold for HYBRID(λ, γ)), and thus γ appears as parameter. For instance, the precise bound on unweighted graphs or for approximations on weighted graphs is $\Omega(n^{1/3}/\gamma^{1/3})$ rounds for labels of size up to $c \cdot n^{2/3} \cdot \gamma^{1/3}$ for some $c > 0$ (c.f., Theorem 14). That is, we get a polynomial lower bound for the HYBRID(∞, γ) model for any $\gamma = n^t$ for constant $t < 1$. For easier readability we plug in the "standard" HYBRID model with $\gamma = \widetilde{O}(1)$, which lets us hide γ using the $\widetilde{\Omega}$ notation.

Table 1 Selected and simplified contributions of this paper.

problem	stretch	complexity	label-size	reference
distance oracles	$3-\varepsilon$	$\widetilde{\Omega}(n^{1/3})$	$O(n^{2/3})^{\dagger}$	Theorem 20
	ℓ	$\widetilde{\Omega}\big(n^{1/f(\ell)}\big)^{\ddagger}$	$O\big(n^{2/f(\ell)}\big)^{\dagger\ddagger}$	Theorem 20, 21
stateless routing schemes	$\sqrt{3}-\varepsilon$	$\widetilde{\Omega}(n^{1/3})$	$O(n^{2/3})^{\dagger}$	Theorem 23
	$\sqrt{5}-\varepsilon$	$\widetilde{\Omega}(n^{1/5})$	$O(n^{2/5})^{\dagger}$	Theorem 23
	$1+\sqrt{2}-\varepsilon$	$\widetilde{\Omega}(n^{1/7})$	$O(n^{2/7})^{\dagger}$	Theorem 23
stateful routing schemes	$\sqrt{2}-\varepsilon$	$\widetilde{\Omega}(n^{1/3})$	$O(n^{2/3})^{\dagger}$	Theorem 25
	$\frac{5}{3}-\varepsilon$	$\widetilde{\Omega}(n^{1/5})$	$O(n^{2/5})^{\dagger}$	Theorem 25
	$\frac{7}{4}-\varepsilon$	$\widetilde{\Omega}(n^{1/7})$	$O(n^{2/7})^{\dagger}$	Theorem 25
	≈ 1.78	$\widetilde{\Omega}(n^{1/11})$	$O(n^{2/11})$	Theorem 25
all of the above on *unweighted graphs*	exact	$\widetilde{\Omega}(n^{1/3})$	$O(n^{2/3})^{\dagger}$	Theorem 14
	$1+\varepsilon$	$\widetilde{O}(n^{1/3}/\varepsilon)$	$\Theta(\log n)$	Theorem 28
all of the above on *weighted graphs*	exact	$\widetilde{O}(n^{1/3})$	$\Theta(n^{2/3})$	Theorem 27
	3	$\widetilde{O}(n^{1/3})$	$\Theta(\log n)$	Theorem 28

†) The lower bound on round complexity holds for node labels of at most that size.
‡) For some function $f(\ell)$ that is linear in ℓ.

Communication Lower Bound in HYBRID. The proofs of lower bounds are based on an intermediate problem which describes the complexity of communicating information between distinct node sets in the HYBRID model, which we can then translate into the realm of classic information theory. We sketch the idea bottom up, neglecting generalizations and most of the details. We start out with a two party communication problem, where Alice is given the state of some random variable X and needs to communicate it to Bob (c.f., Definition 8). Any communication protocol that achieves this needs to communicate $H(X)$ (Shannon entropy of X, see Def. 29) bits in expectation (c.f., Corollary 31), which is a consequence of the source coding theorem (replicated in Lemma 30).

In Section 2 we reduce this to the HYBRID setting into what we call the *node communication problem*. There, we have two sets of nodes A and B, where nodes in A "collectively know" the state of some random variable X and need to communicate it to B (for more precise information see Definition 7). We show a reduction where a HYBRID algorithm that solves the node communication problem on sets A and B, which are at sufficiently large distance in the local graph, can be used to derive a protocol for the two party communication problem (c.f., Lemma 9). We conclude that for sets A, B with distance at least h it takes $\widetilde{\Omega}\big(\min(H(X)/n, h)\big)$ rounds to solve this problem (Theorem 10).

The node communication problem extends on [13, 5] and can be seen as a natural intermediate problem that lies at the core of many interesting problems in the HYBRID(∞, γ) model and might be useful as black-box for other reductions. A major difference to [13] is that there, two party set disjointness is reduced to a distributed decision problem, whereas we show that learning super-linear information from distant parts of a graph is hard in the HYBRID model. The APSP lower bound of [5] uses the observation that a large number of distances has to be learned by a single node. However, labels (even as small as $O(\log n)$) prohibit this idea in our setting, as these must be treated as "free information" that can provide these distances. On a technical level we generalize the approach for the HYBRID(∞, γ) model and strengthen the proof for randomized protocols for *any* constant success probability (see Appendix A).

Lower Bounds on Unweighted Graphs. In Section 3 we construct a reduction from the node communication problem to distance oracle and routing scheme computation in the HYBRID model. The goal is to encode some random variable X with large entropy (super-linear in n) into some randomized part of our local communication graph such that some node set A knows X by vicinity. We construct such a graph Γ (see Figure 1a) from the complete bipartite graph $G_{k,k} = (A, E)$ and a (i.i.d.) random k^2-bit-string $X = (x_e)_{e \in E}$ with $H(X) = \Theta(k^2)$. Then an edge $e \in E$ of $G_{k,k}$ is present in Γ iff $x_e = 1$.

The nodes in A collectively know X since they are incident to the edges sampled from $G_{k,k}$. We designate k nodes of A (one side of the bipartition in $G_{k,k}$) as the "target nodes". We connect each target with a path of length h to one of k "source nodes" which takes the role of B (see Figure 1a). We show that if the nodes in B learn the distances to the target nodes they also learn about the (non-)existence of the edges sampled from $G_{k,k}$ and can conclude the state of X and thus have solved the node communication problem. Balancing the parameters k and h we conclude that $\widetilde{\Omega}(n^{1/3})$ rounds of communication must have taken place to solve the distance oracle problem exactly (Theorem 14).

For routing schemes we have to adapt the graph Γ. We add a slightly longer alternative route from sources to targets (Figure 1a, left side) and show that the state of X can be concluded from the first routing decision the sources have to make, i.e., whether this alternative path is used or not. One caveat is that the nodes are only supposed to give a distance or make a correct routing decision when also provided with the target-label. We choose the labels sufficiently small such that the "free information", given in form of the labels of all targets, is negligible. Still, labels up to size $O(n^{2/3})$ do not change the above narrative, cf. Theorem 14.

Lower Bounds for Approximations. In Section 4 we show how to use graph weights to obtain lower bounds for approximations. We replace $G_{k,k}$ with a balanced, bipartite graph $G = (A, E)$ with k nodes and girth ℓ (length of the shortest cycle in G). As before, the existence of an edge $e \in E$ in Γ is determined by a random bit string $X = (x_e)_{e \in E}$, c.f., Figure 1b. If some edge $e \in E$ is not in Γ, then the detour between endpoints of e using other edges sampled from G is at least $\ell-1$ edges, which translates into almost the same multiplicative detour, by assigning large weights to those edges. Any algorithm for distance oracles with stretch slightly smaller than $\ell-1$ can then be used to solve the node communication problem.

To maximize $H(X)$ we need to maximize the density of G. However, it is known that girth and density of a graph are opposing goals: a graph with girth $2g+1$ can have at most $O\big(n^{1+1/g}\big)$ edges (c.f., [2], simplified in Lemma 34). This inherently limits the amount of information we can encode in Γ and we show in Lemma 18 how graph density affects lower bounds for the node communication problem. The good news is, that for some girth values, graphs that achieve their theoretical density limit actually exist and have been constructed (c.f., [6, 19], simplified form given in Lemma 35). For higher girth values, graphs that come close to that limit are known (c.f., [15], simplified form in Lemma 36). Utilizing these graphs we prove *polynomial* lower bounds for the *distance oracle* problem for small stretch values (c.f., Theorem 20) but also for arbitrary constant stretch (c.f., Theorem 21). Theorem 21 is heavily parametrized, but to sum it up in a simpler way: for any constant stretch ℓ we attain a polynomial lower bound of $\widetilde{\Omega}(n^{1/f(\ell)})$, that is, $f(\ell)$ is constant as well (roughly $f(\ell) \approx \frac{3}{2}\ell$).

A major part is dedicated to lower bounds for approximate stateless and stateful *routing schemes* (c.f., Definitions 4, 5). Here, the introduction of weights, stretch and girth introduces considerable complexity when using our techniques. The main reason for this is that in routing problems a wrong routing decision at the source can often still be completed into a routing path of relatively good quality, even more so for stateful routing, where a packet may "backtrack" to try different paths (for those reasons the lower bounds are also more limited in terms of stretch). We carefully optimize our bad case graph such that it maximizes the stretch for certain round complexities. The results are given in the second and third group of Table 1 with details in Theorems 23 and 25.

Our reductions from the node communication problem to various bad graph instances for distance oracle and routing scheme problems are distinct from those in [13], which shows a $\widetilde{\Omega}(n^{1/3})$ lower bound for computing the diameter that reduces from the two party set-disjointness problem. Our bad case graphs are also distinct from those for APSP in [5], which creates a bottleneck for a single node, that must learn $\widetilde{\Omega}(n)$ bits (which does not work for labeling schemes). Lower bounds for the same problem in the CONGEST model are provided by [12]. Here, the goal is to construct a graph G that has a communication bottleneck (usually a small cut) and small diameter D_G as the $\Omega(D_G)$ lower bound is trivial. By contrast, as the HYBRID model contains LOCAL, small cuts do not help us and we have a trivial $O(D_G)$ upper bound, thus we need to look at graphs with relatively large diameter.

Upper Bounds. For our computational upper bounds (given in Section 5) we show how to reduce the computation of distance oracles and routing schemes for *general graphs* in the HYBRID model to shortest path problems. In particular we draw on fast solutions for the so called *random sources shortest paths problem* (RSSP) [7], where all nodes must learn their distance to a set of i.i.d. randomly sampled nodes, say S. After solving RSSP, our strategy is to use the distance between a node u and the nodes in S as its label $\lambda(u)$.

Roughly speaking, provided that u is sufficiently "far away", a node v can combine $\lambda(u)$ with its own distances to S to compute its distance (estimate) to u. If u is "close" then we can use the local network to compute the distance directly. While this gives us only distance oracles, routing schemes can also be derived. Simply speaking, we can always send a packet to a neighbor that has the best distance (estimate) to u, although some further care must be taken for approximations. Note that this process is oblivious to previous routing decisions so the obtained routing scheme is stateless (c.f., Definition 4).

A trade-off arises from the local exploration around nodes and the global computation depending on the size of S (since we solve RSSP on S), which balances out to a round complexity of $\widetilde{O}(n^{1/3})$ with $|S| \in \widetilde{O}(n^{2/3})$. We obtain exact algorithms (distance oracles and routing schemes) with labels of size $\Theta(n^{2/3})$ (however we can further decrease label size to $\Theta(n^{2/3-\zeta})$ at a cost of $\widetilde{O}(n^{1/3+\zeta})$ rounds, c.f., Theorem 27). Note that this is tight up to polylog n factors as is shown by the corresponding lower bound in Table 1 group 4 line 1 (which holds even on unweighted graphs).

For smaller labels we obtain a 3-approximation on weighted graphs and a $(1+\varepsilon)$ approximation on unweighted graphs in $\widetilde{O}(n^{1/3})$ rounds (for constant $\varepsilon > 0$) with labels of size $O(\log n)$ (c.f., Theorem 28), which is as small as labels can asymptotically be to be able to identify the destination. Compare this to our lower bounds: even much larger labels of size $\Theta(n^{2/3})$ do *not* help to improve the stretch by much, as this still takes $\widetilde{\Omega}(n^{1/3})$ rounds for stretch of $3-\varepsilon$ for distance oracles on weighted graphs, and stretch 1 on unweighted graphs (see Table 1).

Our upper bounds separate the distance oracle (and routing) problem from the related *all pairs shortest paths* (APSP) problem, where nodes must give their distance to all other nodes *without* using labels and where a $\widetilde{\Omega}(n^{1/2})$ lower bound is known even for stretch up to some $\alpha \in \widetilde{\Theta}(n^{1/2})$ [13]. This separation is not the case in the LOCAL and NCC models, where either problem has round complexity $\widetilde{\Theta}(n)$ rounds in general. Our results show that labels of limited size of $O(\log n)$ bits helps to significantly speed up computing (approximate) distances to all destinations in the HYBRID model.

1.2 Related Work

There was an early effort to approach hybrid networks from a theoretic angle [1], with a conceptually different model. Research on the current take of the HYBRID model was initiated by [5] in the context of shortest paths problems, which most of the research has focused on so far. As shortest paths problems problems are closely related, we give a brief account of the recent developments. An overview of distance oracles and routing schemes in other models is provided in the full version of this article [14].

Shortest Paths in the Hybrid Model. [5] introduced an information dissemination scheme to efficiently broadcast small messages to all nodes in the network. Using this protocol, they derive various solutions for shortest paths problems. For instance, for SSSP: a $(1+\varepsilon)$ stretch, $\widetilde{O}(n^{1/3})$-round algorithm and a $(1/\varepsilon)^{O(1/\varepsilon)}$-stretch, $\widetilde{O}(n^{\varepsilon})$-round algorithm. Further, an approximation of APSP with stretch 3 in $\widetilde{O}(n^{1/2})$ rounds, which closely matches their $\widetilde{\Omega}(n^{1/2})$ lower bound (which holds for much larger stretch). Subsequently, [13] introduced a protocol to efficiently uni-cast small messages between dedicated source-target pairs in the HYBRID model, which they use to solve APSP and SSSP exactly in $\widetilde{O}(n^{1/2})$ and $\widetilde{O}(n^{2/5})$ rounds, respectively. For computing the diameter they provide algorithms (e.g., a $3/2+\varepsilon$ approximation in $\widetilde{O}(n^{1/3})$ rounds) and a $\widetilde{\Omega}(n^{1/3})$ lower bound. [7] combines the techniques of [13] with a densitiy sensitive approach, to solve $n^{1/3}$-SSP (thus SSSP) exactly and compute a $(1+\varepsilon)$-approximation of the diameter in $\widetilde{O}(n^{1/3})$ rounds. [8] uses density awareness in a different way to improve SSSP to $\widetilde{O}(n^{5/17})$ rounds for a small stretch of $(1+\varepsilon)$. For classes of sparse graphs (e.g., cactus graphs) [10] demonstrates that exact solutions in $\widetilde{O}(1)$ rounds are possible even in the harsher hybrid combination CONGEST and NCC. The recent result of [24] implies an $(1+\varepsilon)$ approximation of SSSP in $\widetilde{O}(1)$ rounds in *general* graphs in the hybrid CONGEST and NCC regime, since their partwise aggregation model can be simulated efficiently in the HYBRID model as shown by [3]. The article by [3] also derandomized the dissemination protocol of [5] to obtain a deterministic APSP-algorithm with stretch $\frac{\log n}{\log \log n}$ in $\widetilde{O}(n^{1/2})$ rounds.

1.3 Preliminaries

General Definitions. The scope of this paper is solving graph problems, typically in the undirected local graph $G = (V, E)$. Edges have weights $w : E \to [W]$, where W is at most polynomial in n, thus the weight of an edge and of a simple path fits into a $O(\log n)$ bit message (whereas define $\log := \log_2$). Graph G is considered unweighted if $W = 1$. Let $w(P) = \sum_{e\in P} w(e)$ denote the length of a path $P \subseteq E$. Then the *distance* between two nodes $u, v \in V$ is $d_G(u, v) := \min_{u\text{-}v\text{-path } P} w(P)$. A path with smallest length between two nodes is called a *shortest path*. Let $|P|$ be the number of edges (or *hops*) of a path P. The *hop-distance* between two nodes u and v is defined as: $\text{hop}_G(u, v) := \min_{u\text{-}v\text{-path } P} |P|$. We generalize this to sets $U, W \subseteq V$ (whereas $\text{hop}_G(v, v) := 0$): $\text{hop}_G(U, W) := \min_{u\in U, w\in W} \text{hop}_G(u, w)$. The *diameter* of G is defined as: $D_G := \max_{u,v\in V} \text{hop}_G(u, v)$. Let the *$h$-hop distance* from u to v be: $d_{G,h}(u, v) := \min_{u\text{-}v\text{-path } P, |P|\le h} w(P)$. If there is no u-v path P with $|P| \le h$ we define $d_h(u, v) := \infty$. We drop the subscript G, if G is clear from the context. We consider the following problems:

▶ **Definition 3** (Distance Oracles). *Every node $v \in V$ of a graph $G = (V, E)$ needs to compute a label $\lambda(v)$ and an oracle function $o_v : \lambda(V) \to \mathbb{N}$, such that $o_v(\lambda(u)) \ge d(u, v)$ for all $u \in V$. An oracle function o_v is an (α, β)-approximation if $o_v(\lambda(u)) \le \alpha \cdot d(u, v) + \beta$ for all $u, v \in V$, that is, α, β are the multiplicative and additive approximation error, respectively. We speak of a stretch of α in case of an $(\alpha, 0)$-approximation. If the stretch is one, we call o_v exact.*

▶ **Definition 4** (Stateless Routing Scheme). *Every node $v \in V$ of a graph $G = (V, E)$ needs to learn a label $\lambda(v)$ and a routing function (sometimes called "table") $\rho_v : \lambda(V) \to N(v) \cup \{v\}$ where $N(v)$ are adjacent nodes of v in G (whereas we formally set $\rho_v(\lambda(v)) := v$). The functions ρ_v must fulfill the following correctness condition. Let $v_0 := v$ and recursively define $v_i := \rho_{v_{i-1}}(\lambda(u))$. Then the routing functions $\rho_v, v \in V$ must satisfy $v_h = u$ for some $h \in \mathbb{N}$. Let $P_\rho(u, v)$ be the path induced by the visited nodes $v_0, \dots, v_h$. We call ρ an (α, β)-approximation if $w(P_\rho(u, v)) \le \alpha d(u, v) + \beta$ for all $u, v \in V$.*

▶ **Definition 5** (Stateful Routing Scheme). *This is mostly defined as in the stateless case, with the difference that ρ_v can additionally depend on the information gathered along the path that has already been visited by a packet (which would be stored in its header). Note that the routing path defined by such a function ρ might have loops.*

▶ **Definition 6** (Randomized Graph Algorithms). *We say that an algorithm has success probability p, if it succeeds with probability at least p on every possible input graph (however, some of our results are restricted to unweighted input graphs). Specifically, for our upper bounds we aim for success with high probability (w.h.p.), which means with success probability at least $1 - \frac{1}{n^c}$ for arbitrary constant $c > 0$.*

2 Node Communication Problem

In this section we create an "information bottleneck" in the HYBRID model between two (distant) parts of the local communication graph. We do this for the more general $\text{HYBRID}(\infty, \gamma)$ model, where we have a global communication bandwidth of γ bits per node per round, which also has the advantage of avoiding logarithmic terms and O-notation as long as possible (recall that $\text{HYBRID} = \text{HYBRID}(\infty, O(\log^2 n))$).

We first introduce some definitions. We say that the nodes from some set $A \subseteq V$ *collectively know* the state of a random variable X, if its state can be derived from the information that the nodes A have. Or, in terms of information theory, given the state or input S_A of all nodes A (interpreted as a random variable), then the conditional entropy $H(X|S_A)$ (see Definition 29), also known as the amount of new information of X provided that S_A is already known, is zero. Similarly, we say that the state of X is *unknown* to $B \subseteq V$, if the initial information of the nodes B does not induce any knowledge on the outcome of X. Or formally, that for the state S_B of the nodes B we have that $H(X|S_B) = H(X)$, meaning that all information in X is new even if S_B is known.

▶ **Definition 7** (Node Communication Problem). *Let $G = (V, E)$ be some graph. Let $A, B \subset V$ be disjoint sets of nodes and $h := hop(A, B)$. Furthermore, let X be a random variable whose state is collectively known by the nodes A but unknown to any set of nodes disjoint from A. An algorithm $\mathcal{A}$ solves the* node communication problem *if the nodes in B collectively know the state of X after $\mathcal{A}$ terminates. We say $\mathcal{A}$ has success probability p if $\mathcal{A}$ solves the problem with probability at least p for* any *state X can take (in line with our Definition 6 of success probability for graph algorithms).*

The goal of Lemma 9 is to reduce a more basic communication problem, for which we can provide lower bounds using information theory (c.f., Appendix A) to the node communication problem. Analogously to node sets, we define that Alice knows some random variable X, which is unknown to Bob as follows. Given that S_{Alice} and S_{Bob} are their respective inputs then we have $H(X|S_{\text{Alice}}) = 0$ and $H(X|S_{\text{Bob}}) = H(X)$.

▶ **Definition 8** (Two Party Communication Problem). *Given two computationally unbounded parties, Alice and Bob, where initially Alice knows the state of some random variable X which is unknown to Bob. A communication protocol $\mathcal{P}$ is said to solve that problem if after its execution Bob can derive the state of X from the transcript of all exchanged messages. Performance is measured in the length of the transcript in bits. We say $\mathcal{P}$ has success probability p if $\mathcal{P}$ solves the problem with probability at least p for* any *state X can take.*

The reduction from the 2-party communication problem to the node communication problem uses the following simulation argument: Alice and Bob can together simulate a HYBRID(∞, γ) model algorithm for the node communication problem and use it solve the 2-party communication problem. The proof is deferred to Appendix C.

▶ **Lemma 9.** *Any algorithm $\mathcal{A}$ that solves the* node communication problem *(Def. 7) in the* HYBRID(∞, γ) *model on some local graph $G = (V, E)$ with $n = |V|$ and $A, B \subset V$ in $T < h = hop(A, B)$ rounds with success probability p can be used to obtain a protocol $\mathcal{P}$ that solves the two party communication problem (Def. 8) with the same success probability p and transcript length at most $T \cdot n \cdot \gamma$.*

We plug in the lower bound for the 2-party communication problem (c.f. Lemma 32 in Appendix A) to derive a lower bound for the node communication problem. Note that this theorem only depends on the hop distance h between A, B, the entropy of X and the number of nodes n and is otherwise agnostic to the local graph. Also note that a lower bound that holds in expectation is also a worst case lower bound.

▶ **Theorem 10.** *Any algorithm that solves the node communication problem (Def. 7) on some n-node graph in the* HYBRID(∞, γ) *model with success probability at least p, takes at least $\min\left(\frac{pH(X)-1}{n \cdot \gamma}, h\right)$ rounds in expectation, where $H(X)$ denotes the entropy of X.*

Proof. We have to show that a randomized, HYBRID(∞, γ) algorithm $\mathcal{A}$ that solves the node communication problem in *less* than h rounds with success probability p takes at least $\frac{pH(X)-1}{n \cdot \gamma}$ rounds. Presume, for a contradiction, that $\mathcal{A}$ has an expected running time $T < h$ *and* $T < \frac{pH(X)-1}{n \cdot \gamma}$. This implies $T \cdot n \cdot \gamma < p \cdot H(X) - 1$.

Invoking Lemma 9 gives us a protocol $\mathcal{P}$ with the same success probability p and with a transcript of length at most $T \cdot n \cdot \gamma$. With the inequality above, this means in the protocol $\mathcal{P}$, Alice sends *less* than $p \cdot H(X) - 1$ bits to Bob in expectation. This contradicts the fact that $p \cdot H(X) - 1$ is a lower bound for this due to Appendix A Lemma 32. ◀

We have to accommodate the fact that in the routing problem or distance oracle problem, the nodes have to give a distance estimation or next routing neighbor only when provided with the label of the target node. Therefore we have to slightly amend Theorem 10, which will later allow us to argue that even if we assume that nodes have advance knowledge of a selection of sufficiently small labels, the lower bound will not change asymptotically.

▶ **Corollary 11.** *If A is allowed to communicate y bits to B for free, then any algorithm that solves the node communication problem on some n-node graph (Def. 7) in the* $\mathsf{HYBRID}(\infty, \gamma)$ *model with success probability at least p, takes at least $\min\left(\frac{pH(X)-1-y}{n\cdot\gamma}, h\right)$ rounds in expectation (i.e., also in the worst case).*

3 Lower Bounds For Unweighted Graphs

In this and the following section we aim to reduce from the node communication problem in the $\mathsf{HYBRID}(\infty, \gamma)$ model given in Definition 7, to the problem of computing routing tables or distance oracles, which works as follows. We define a graph $\Gamma = (V_\Gamma, E_\Gamma)$ such that, first, the solution of the routing or distance oracle problems informs a subset $B \subset V_\Gamma$ about the exact state of some random variable X that is encoded by the subgraph induced by $A \subset V_\Gamma$. Second, X has a large entropy (we aim for super-linear in n). And third, the distance $hop(A, B)$ between both sets is sufficiently large.

▶ **Definition 12.** *Let $X = (x_{ij})_{i,j\in[k]} \in \{0,1\}^{k^2}$ be a bit sequence of length k^2. Let $\Gamma = (V_\Gamma, E_\Gamma)$ (shown in Figure 1a) be an unweighted graph with source nodes $s_1, \dots, s_k \in V_\Gamma$, transit nodes $u_1, \dots, u_k \in V_\Gamma$ and target nodes $t_1, \dots, t_k \in V_\Gamma$. Each source s_i has a path of length h hops to the transit nodes u_i. We have an edge between u_i and t_j if and only if $x_{ij} = 1$. Additionally, there are two nodes $v, v' \in V_\Gamma$ connected by a path of h hops. The nodes v and v' have an edge to each source s_i or target t_i, respectively (Figure 1a).*

This construction has the following properties.

(1) The distance from source s_i to t_j is larger for $x_{ij} = 0$ than for $x_{ij} = 1$ (as shown by the subsequent Lemma 13).

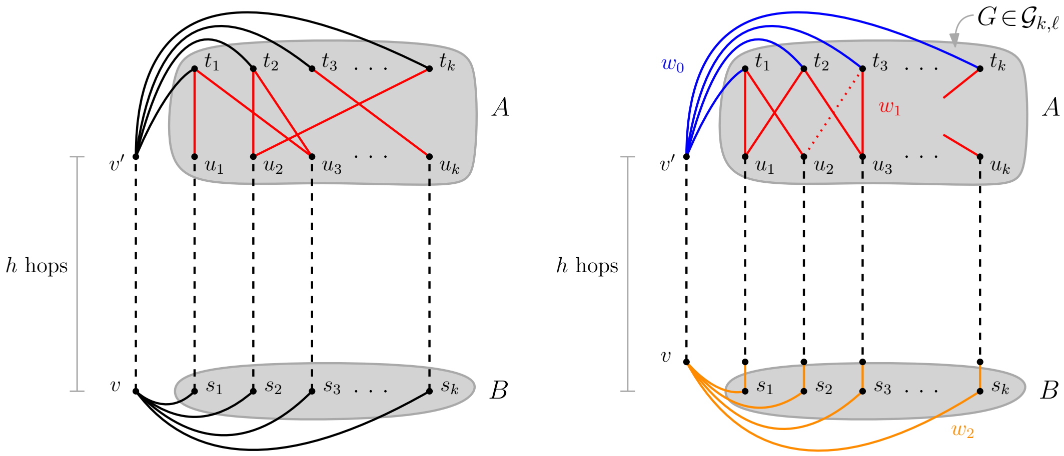
(2) For all $i, j \in [k]$, independently set $x_{ij} = 1$ with probability $\frac{1}{2}$, else $x_{ij} = 0$. This maximizes $H(X) = -k^2 \cdot \frac{\log(1/2)}{2} = \frac{k^2}{2}$.

(3) Let $A = \{u_1, \dots, u_k, t_1, \dots, t_k\}$, $B = \{s_1, \dots, s_k\}$, i.e., $hop(A, B) = h$.

▶ **Lemma 13.** *If $x_{ij} = 1$ then $d(s_i, t_j) = h+1$ and the shortest s_i-t_j-path contains v, else $d(s_i, t_j) = h+2$ and it does not contain v.*

Proof. Any path from s_i to t_j has to cross the vertex cut $U := \{u_1, \dots, u_k, v'\}$ (c.f., Figure 1a). Such a path has to include a path of length h to reach a node of U, as well as an additional edge connecting U to t_j and therefore $d(s_i, t_j) \geq h + 1$. However, we also have $d(s_i, t_j) \leq h + 2$, due to the path along the nodes $s_i, v, \dots, v', t_j$ (c.f., Figure 1a) that has length $h + 2$.

If $x_{ij} = 1$, i.e., $\{u_i, t_j\} \in E$, then the path along the nodes $s_i, \dots, u_i, t_j$ has length $h + 1$. Note that all nodes in $U \setminus \{u_i, v'\}$ are at distance at least $h + 2$ from s_i (c.f., Figure 1a), so every path via one of the nodes $U \setminus \{u_i, v'\}$ has distance at least $h + 3$. In the case $x_{ij} = 0$, i.e., $\{u_i, t_j\} \notin E_\Gamma$, this is also true for the path via u_i and the only path with distance $h + 2$ is the one via v'. ◀



(a) Graph $\Gamma = (V_\Gamma, E_\Gamma)$. Bit string $X = (x_{ij}) \in \{0,1\}^{k^2}$ determines red edges. E.g., $\{u_1, t_2\} \notin E_\Gamma$ and $\{u_1, t_3\} \in E_\Gamma$ means $x_{12} = 0$ and $x_{13} = 1$.

(b) Graph Γ constructed around $G = (A, E) \in \mathcal{G}_{k,\ell}$ with $m := |E|$ edges (red) and from $X = (x_e)_{e \in E} \in \{0,1\}^m$, where $e \in E$ is part of Γ iff $x_e = 1$.

Figure 1 Lower bound graphs. Unweighted (left) and weighted (right).

The idea to prove the next theorem is that if the nodes in B learn the distance to the nodes in $t_1, \dots, t_k$, then their combined knowledge can be used to infer the state of the random string X that is collectively known by the nodes in A. The proof the applies Theorem 10 and we maximize the lower bound by balancing k (which givers the size of $H(X)$) and h. Formally, we apply Corollary 11 to account for the labels $\lambda(t_1), \dots, \lambda(t_k)$ that are given "for free" to the nodes in B, which gives us an upper bound on their size for which the lower bound in round complexity still holds. The full proof is given in [14].

▶ **Theorem 14.** *Even on* unweighted graphs*, any randomized algorithm that computes exact (stateless or stateful) routing schemes or distance oracles in the* HYBRID(∞, γ) *model with constant success probability takes* $\Omega(n^{1/3}/\gamma^{1/3})$ *rounds. This holds for labels of size up to* $c \cdot n^{2/3} \cdot \gamma^{1/3}$ *(for a fixed constant* $c > 0$*).*

4 Lower Bounds for Approximations

Our next construction relies on the existence of families of graphs that have high girth and maintain relatively high density. We modify the basic construction above, essentially by replacing the upper part of Γ with a random selection of edges from a graph of that family (and also making Γ weighted). Besides high density we require the following.

▶ **Definition 15.** $\mathcal{G}_{k,\ell}$ *is a graph family, s.t. for all* $G = (A, E) \in \mathcal{G}_{k,\ell}$*: (i)* $|A| = 2k$*, (ii)* G *has (even) girth at least* ℓ*, (iii)* G *is balanced and bipartite.*

Removing an edge from $G \in \mathcal{G}_{k,\ell}$ incurs a large detour of at least $\ell - 1$ hops between the endpoints of that edge, since otherwise there would be a cycle shorter than ℓ in G. This observation is often used to prove certain bounds for low stretch subgraphs (one prominent example is the lower bound on the size of low stretch spanners). and can be exploited to introduce a stretch into our lower bound construction. We construct this formally as follows (however, first consulting Figure 1b will presumably be more helpful to the reader).

▶ **Definition 16.** *Let* $G = (A, E) \in \mathcal{G}_{k,\ell}$ *with* $m := |E|$ *edges and let* $\{u_1, \dots, u_k\} \cup \{t_1, \dots, t_k\} = A$ *be the bipartition of* G*. Graph* $\Gamma = (V_\Gamma, E_\Gamma)$ *(shown in Figure 1b) has a similar structure as the unweighted construction (Def. 12), where the main difference is the way how the nodes* $\{u_1, \dots, u_k\} \cup \{t_1, \dots, t_k\}$ *are connected by edges in* Γ*.*

Let $X = (x_e)_{e \in E} \in \{0,1\}^m$ be a bit string of length $m = |E|$, i.e., each bit x_e corresponds to an edge e of G. For each u_i, t_j we have $\{u_i, t_j\} \in E_\Gamma$, if and only if $\{u_i, t_j\} \in E$ and $x_{\{u_i,t_j\}} = 1$. In a slight change from the previous construction, we make the path from v to v' of hop length $h-1$. The weights of Γ are assigned as follows. Edges between the node v' and some t_j have weight w_0. Edges between nodes u_i, t_j have weight w_1. Edges incident to some s_i have weight w_2.

We have the following properties.

(1) Let $e=\{u_i,t_j\}\in E$. w_0, w_1, w_1 can be chosen s.t. $d(s_i, t_j)$ is much longer for $x_e = 0$ than for $x_e = 1$ (c.f., Lemma 17).

(2) For each edge $e \in E$ of G, set $x_e = 1$ i.i.d. with probability $\frac{1}{2}$, else $x_e = 0$. This maximizes the entropy $H(X) = \frac{m}{2}$.

(3) For nodes A of G and $B := \{s_1, \dots, s_k\}$ we have $\text{hop}(A, B) = h$.

We observe that the distances $d(s_i, t_j)$ between nodes s_i, t_j with $e = \{u_i, t_j\} \in E$ depend on whether e is in Γ ($x_e = 1$), or not ($x_e = 0$), the proof is deferred to Appendix D. Conceptually, we later choose weights $w_1 \ll w_0$, so that x_e induces a large difference in $d(s_i, t_j)$.

▶ **Lemma 17.** *Consider Γ (Def. 16), constructed from $G = (A, E) \in \mathcal{G}_{k,\ell}$ and X. Let $w_1 < w_0 < (\ell - 1)w_1$. Let $e = \{u_i, t_j\} \in E$. Then we have:*

(i) *The shortest s_i-t_j-path contains v if and only if $x_e = 0$.*

(ii) *If $x_e = 1$ then $d(s_i, t_j) = w_2 + w_1 + h - 1$.*
If $x_e = 0$, then $d(s_i, t_j) = w_2 + w_0 + h - 1$.

For the reduction from the node communication problem to our concrete routing and distance oracle problems, we start with a technical lemma that analyzes the running time of any algorithm $\mathcal{A}$ that solves the node communication problem in Γ for the dedicated node sets A, B and the random variable X from which Γ is constructed.

In particular, we express the lower bound from Theorem 10 as function of $n := |V_\Gamma|$, the density of $G \in \mathcal{G}_{k,\ell}$ given by a parameter δ and the global communication capacity γ. The lemma is the result of balancing a trade off between the distance $h = \text{hop}(A, B)$ and the number of nodes $\Theta(k)$ of G (which governs the entropy $H(X) = \Theta(k^{1+\delta})$ when the density of G is fixed).For the proof details see [14].

▶ **Lemma 18.** *Consider Γ constructed from random variable X and $G = (A, E) \in \mathcal{G}_{k,\ell}$ (Def. 16) with $|E| = \Theta(k^{1+\delta})$ edges (for $\delta > 0$ and k of our choosing). Let $\mathcal{A}$ be an algorithm that solves the node communication problem on Γ with X, node sets $A, B \subset V_\Gamma$ and $h = hop(A, B)$ in the* HYBRID(∞, γ) *model (all parameters as in Def. 16). We can choose $k = \Theta\big(\frac{n}{h}\big)$ such that $\mathcal{A}$ takes $\Omega\big((\frac{n^\delta}{\gamma})^{\frac{1}{2+\delta}}\big)$ rounds. There exists a constant $c > 0$ such that this holds even when we allow exchanging $c \cdot k^{1+\delta}$ bits from A to B for free.*

4.1 Distance Oracles

The first lower bound with stretch is for the distance oracle problem. The idea is as follows. In case there is a direct edge $e = \{u_i, t_j\}$ (i.e., $x_e = 1$), the distance from s_i to t_j is almost $\ell-1$ times shorter, than if that is not the case. Hence, by learning an approximation of $d(s_i, t_j)$ with a stretch slightly lower than $\ell-1$, the node s_i can conclude if e exists or not, i.e., if $x_e = 1$ or $x_e = 0$. Hence the nodes $B = \{s_1, \dots, s_k\}$ collectively learn the random variable X and thus solve the node communication problem.

This lemma is kept general such that we can plug in any graph with density parameter δ and girth ℓ. Note that the girth ℓ fundamentally limits the density parameter δ; the correspondence between the two is roughly $\delta \in O\big(\frac{1}{\ell}\big)$ as shown in Appendix B. For a more intuitive understanding we suggest plugging in the complete bipartite graph $G_{k,k}$ which has girth $\ell = 4$ and $\Theta(k^2)$ edges (i.e., density parameter $\delta = 1$). The proof appears in [14].

▶ **Lemma 19.** *Consider Γ constructed from $G = (A, E) \in \mathcal{G}_{k,\ell}$ with $|E| = \Theta\big(k^{1+\delta}\big)$ edges for some $\delta > 0$. Any algorithm that solves the* distance oracle problem *on Γ with stretch $\alpha_\ell = \ell-1-\varepsilon$ (for any const. $\varepsilon > 0$) and constant success probability in the* HYBRID(∞, γ) *model takes $\Omega\big((\frac{n^\delta}{\gamma})^{\frac{1}{2+\delta}}\big)$ rounds, for labels up to size $c \cdot n^{\frac{2\delta}{2+\delta}} \cdot \gamma^{\frac{\delta}{2+\delta}}$ (for a fixed const. $c > 0$).*

It remains to insert graphs $G \in \mathcal{G}_{k,\ell}$ into Lemma 19. We aim for graphs $G = (V, E) \in \mathcal{G}_{k,\ell}$ with $|E| \in \Theta\big(k^{1+\delta}\big)$ that maximize both girth ℓ and density parameter δ. As outlined in Appendix B, these are opposing objectives, and for even girth $\ell \geq 4$ we know that $\delta \in O\big(\frac{2}{\ell-2}\big)$ (from applying Lemma 34 on uneven girth $\ell - 1$). Bipartite graphs of girth ℓ that reach $\delta \in \Theta\big(\frac{2}{\ell-2}\big)$ can be constructed for small girth ℓ (summarized in Lemma 38) from which we obtain Theorem 20. But for higher girth we have to settle for δ below this threshold (see Lemma 39), this is reflected in Theorem 21.

▶ **Theorem 20.** *Any algorithm that solves the* distance oracle problem *in the* HYBRID(∞, γ) *model with constant success probability with*

- *stretch $3-\varepsilon$ takes $\Omega\big((\frac{n}{\gamma})^{\frac{1}{3}}\big)$ rounds for label size $\leq c \cdot (n^2\gamma)^{\frac{1}{3}}$*
- *stretch $5-\varepsilon$ takes $\Omega\big(\frac{n^{1/5}}{\gamma^{2/5}}\big)$ rounds for label size $\leq c \cdot (n^2\gamma)^{\frac{1}{5}}$*
- *stretch $7-\varepsilon$ takes $\Omega\big(\frac{n^{1/7}}{\gamma^{3/7}}\big)$ rounds for label size $\leq c \cdot (n^2\gamma)^{\frac{1}{7}}$*
- *stretch $11-\varepsilon$ takes $\Omega\big(\frac{n^{1/11}}{\gamma^{5/11}}\big)$ rounds for label size $\leq c \cdot (n^2\gamma)^{\frac{1}{11}}$*

for any const. $\varepsilon > 0$ and a fixed const. $c > 0$.

Proof. By Lemma 38 there are bipartite, balanced graphs with girth $\ell \in \{4, 6, 8, 12\}$ and $\Theta(n^{1+\frac{2}{\ell-2}})$ edges, thus $\delta(\ell) = \frac{2}{\ell-2}$. In particular, we have $\delta(4) = 1$, $\delta(6) = \frac{1}{2}$, $\delta(8) = \frac{1}{3}$, $\delta(12) = \frac{1}{5}$, which yield the desired results when plugged into Lemma 19. ◀

Applying Lemma 19 on the densest known graphs with larger girth (see Lemma 39), we obtain the subsequent theorem. The parametrization is complex due to a case distinction in Lemma 39, the upshot is that for any constant stretch and sufficiently small γ we still get polynomial (in n) lower bounds for labels up to some polynomial size (in n).

▶ **Theorem 21.** *Any algorithm that solves the* distance oracle problem *in the* HYBRID(∞, γ) *model with constant success probability for any const. $\varepsilon > 0$ and a fixed const. $c > 0$, with*

- *stretch $\ell-1-\varepsilon$ for $\ell \geq 14$ with $\ell \equiv 2 \bmod 4$ takes $\Omega\Big(n^{\frac{2}{3\ell-8}} / \gamma^{\frac{3\ell-10}{6\ell-6}}\Big)$ rounds for label size $\leq c \cdot n^{4/(3\ell-8)} \cdot \gamma^{2/(3\ell-8)}$*
- *stretch $\ell-1-\varepsilon$ for $\ell \geq 16$ with $\ell \equiv 0 \bmod 4$ takes $\Omega\Big(n^{\frac{2}{3\ell-10}} / \gamma^{\frac{3\ell-12}{6\ell-8}}\Big)$ rounds for label size $\leq c \cdot n^{4/(3\ell-10)} \cdot \gamma^{2/(3\ell-10)}$.*

Proof. By Lemma 39 there are bipartite, balanced graphs with even girth $\ell \geq 14$ that have (i) $\Theta(n^{1+\frac{4}{3\ell-10}})$ edges if $\ell \equiv 2 \bmod 4$, or (ii) $\Theta(n^{1+\frac{4}{3\ell-12}})$ edges if $\ell \equiv 0 \bmod 4$. Thus in case (i) we have $\delta(\ell) = \frac{4}{3\ell-10}$ and in case (ii) $\delta(\ell) = \frac{4}{3\ell-12}$. Plugging $\delta(\ell)$ into Lemma 19 gives the desired result. ◀

4.2 Stateless Routing Scheme

For lower bounds of routing schemes we exploit the observation that for an edge $e = \{s_i, t_j\} \in E$ the node s_i learns about the existence of e in Γ, i.e., whether $x_e = 0$ or $x_e = 1$, from the decision to send a packet with destination t_j first to v or not. More precisely, we show that $x_e = 0$ if and only if v is the first routing neighbor for the packet with destination t_j.

However, we have to decrease the stretch of our lower bound in order that this works. The main obstacle is that the decision of s_i to send a packet with target t_j directly towards u_i instead of node v (left path) does not impact the distance of the routing path that one can still obtain by that much.

In particular, in the case of *stateless* routing, a packet that travels from s_i to u_i and finds that the direct edge $\{u_i, t_j\}$ does *not* exist, could still use any other edge $\{u_i, t_p\}$ and then the two edges $\{t_p, v'\}, \{v', t_j\}$ to get to t_j (e.g., in Figure 1b from s_2 to t_3). This would mislead s_i as the first routing node was *not* v, yet $x_e = 0$.

The target is to prohibit this and some other troublesome routing options by making them exceed the stretch guarantee. However, this gives us additional restrictions that dominate the resulting system of inequalities for higher girth ℓ of G, in particular we gain no improvement in the stretch for $\ell \geq 8$. The proof is deferred to [14].

▶ **Lemma 22.** *Consider Γ constructed from $G = (A, E) \in \mathcal{G}_{k,\ell}$ with $|E| = \Theta\big(k^{1+\delta}\big)$ edges for some $\delta > 0$. For any constant $\varepsilon > 0$ let $\alpha_\ell = \sqrt{\ell - 1} - \varepsilon$ for $\ell \leq 6$ and $\alpha_\ell = 1 + \sqrt{2} - \varepsilon$ for $\ell \geq 8$. Any algorithm that computes a* stateless routing scheme *on Γ with stretch α_ℓ and constant success probability in the* HYBRID(∞, γ) *model takes $\Omega\big((\frac{n^\delta}{\gamma})^{\frac{1}{2+\delta}}\big)$ rounds. This holds for labels of size $c \cdot n^{\frac{2\delta}{2+\delta}} \gamma^{\frac{\delta}{2+\delta}}$ and fixed constant $c > 0$.*

We plug graphs $G \in \mathcal{G}_{k,\ell}$ into Lemma 22. Since in this case we get no improvements in the stretch for girth $\ell \geq 8$ it suffices to apply Lemma 38. Beside the changed values for the stretch, the proof is the same as that of Theorem 20, we just have to use the corresponding values of δ from Lemma 38 for $\ell = 4, 6, 8$.

▶ **Theorem 23.** *Any algorithm that solves the* stateless routing problem *in the* HYBRID(∞, γ) *model with constant success probability with*

- *stretch $\sqrt{3} - \varepsilon$ takes $\Omega\big((\frac{n}{\gamma})^{\frac{1}{3}}\big)$ rounds for label size $\leq c \cdot (n^2\gamma)^{\frac{1}{3}}$*
- *stretch $\sqrt{5} - \varepsilon$ takes $\Omega\big(\frac{n^{1/5}}{\gamma^{2/5}}\big)$ rounds for label size $\leq c \cdot (n^2\gamma)^{\frac{1}{5}}$*
- *stretch $1 + \sqrt{2} - \varepsilon$ takes $\Omega\big(\frac{n^{1/7}}{\gamma^{3/7}}\big)$ rounds for label size $\leq c \cdot (n^2\gamma)^{\frac{1}{7}}$*

for any const. $\varepsilon > 0$ and a fixed const. $c > 0$.

4.3 Stateful Routing Scheme

We obtain similar lower bound results for the approximate *stateful* routing problem, however with even smaller stretch. Recall that in the stateful version the problem is relaxed in the sense that a routing decision may also depend on the information a packet has gathered along the previous routing path.

Since this permits loops in the routing path, it opens up additional options for routing a packet from s_i to t_j that we need to prohibit. For instance, a packet could first travel to u_i, then check if the direct edge $\{u_i, t_j\}$ is present, and if not travel back to s_i to take the shorter route via v instead. Note that this path has the same number of red and blue edges as the shortest path directly to v and then to t_j (c.f. Figure 1b).

The trick is to make the weight w_2 (orange edges) of all incident edges of s_i more expensive, such that revisiting s_i breaks the approximation guarantee. This again forces the source s_i to make the correct decision with the first node it routes the packet to, which renders the ability to travel in loops and learn along the way useless. We carefully optimize the involved parameters to obtain the following lemma (the details are given in [14]).

▶ **Lemma 24.** *Consider Γ constructed from $G = (A, E) \in \mathcal{G}_{k,\ell}$ with $|E| = \Theta\big(k^{1+\delta}\big)$ edges for some $\delta > 0$. For any constant $\varepsilon > 0$ let $\alpha_4 = \sqrt{2} - \varepsilon$, $\alpha_6 = \frac{5}{3} - \varepsilon$, $\alpha_8 = \frac{7}{4} - \varepsilon$. For $\ell \geq 10$ let $\alpha_\ell = \frac{3+\sqrt{17}}{4} - \varepsilon \approx 1.78$. Any algorithm that computes a* stateful routing scheme *on Γ with stretch α_ℓ and constant success probability in the* HYBRID(∞, γ) *model takes $\Omega\big((\frac{n^\delta}{\gamma})^{\frac{1}{2+\delta}}\big)$ rounds. This holds for labels of size $c \cdot n^{\frac{2\delta}{2+\delta}} \cdot \gamma^{\frac{\delta}{2+\delta}}$ and fixed constant $c > 0$.*

Again, our actual lower bounds come from inserting graphs $G \in \mathcal{G}_{k,\ell}$ into Lemma 22. Our best stretch is obtained for $\ell = 10$, but unfortunately we have a gap for that value in Lemma 38. Therefore, for the largest stretch value we use a graph $G \in \mathcal{G}_{k,12} \subseteq \mathcal{G}_{k,10}$, which has the drawback of not being as dense. Aside from different stretch values, the proof follows that of Theorem 20, by inserting the values of δ from Lemma 38 for $\ell = 4, 6, 8, 12$.

▶ **Theorem 25.** *Any algorithm that solves the* stateful routing problem *in the* HYBRID(∞, γ) *model with constant success probability with*

- *stretch* $\sqrt{2} - \varepsilon$ *takes* $\Omega\big((\frac{n}{\gamma})^{\frac{1}{3}}\big)$ *rounds for label size* $\leq c \cdot (n^2\gamma)^{\frac{1}{3}}$
- *stretch* $\frac{5}{3} - \varepsilon$ *takes* $\Omega\big(\frac{n^{1/5}}{\gamma^{2/5}}\big)$ *rounds for label size* $\leq c \cdot (n^2\gamma)^{\frac{1}{5}}$
- *stretch* $\frac{7}{4} - \varepsilon$ *takes* $\Omega\big(\frac{n^{1/7}}{\gamma^{3/7}}\big)$ *rounds for label size* $\leq c \cdot (n^2\gamma)^{\frac{1}{7}}$
- *stretch* $\frac{3+\sqrt{17}}{4} - \varepsilon$ *takes* $\Omega\big(\frac{n^{1/11}}{\gamma^{5/11}}\big)$ *rounds for label size* $\leq c \cdot (n^2\gamma)^{\frac{1}{11}}$

for any const. $\varepsilon > 0$ *and a fixed const.* $c > 0$.

5 Upper Bounds

In this section, we sketch our algorithms for computing routing schemes and distance oracles in the HYBRID model. Due to space limitations, the details are given in the full version [14], where we will also show that distance oracles imply *stateless* routing schemes with the same label-size, stretch and asymptotic round complexity, thus we concentrate on the former. We combine two techniques to compute exact distance oracles, namely skeleton graphs and fast algorithms for the random sources shortest paths (RSSP) problem, where all nodes need to determine their distance to a set of nodes that was sampled i.i.d. from V.

Skeleton graphs were first used by [20] and became one of the main tools used in the context of shortest path algorithms in the HYBRID model (c.f., [5, 13, 7]). Simply speaking, a skeleton consists of a (usually relatively small) set of nodes sampled with some probability $\frac{1}{x}$ and virtual edges formed between sampled nodes at most $h \in \widetilde{O}(x)$ hops apart. The main property is that for any $u, v \in V$, there is a sampled node at least every h hops on a shortest path from u to v w.h.p. Here, we only require this sampling property, the explicit construction of the skeleton graph occurs under the hood when solving the RSSP problem. The following result is known.

▶ **Lemma 26** (c.f., [7]). *There is a* HYBRID *algorithm that solves the RSSP problem for sampling probability* $1/x$ *for* $x \geq 1$ *exactly and w.h.p. in* $\widetilde{O}(n^{1/3} + n/x^2)$ *rounds.*

Our algorithm to compute distance oracles roughly consists of the following steps. First, we sample the nodes for the skeleton graph S i.i.d. with probability $\frac{1}{x}$. Then we do a local exploration to depth $h \in \widetilde{O}(x)$. If our target node has a shortest path with less or equal h hops, distance decisions can be made based on the so acquired local knowledge.

For shortest paths from v to u with more than h hops we do the following. We first solve the RSSP problem using Lemma 26. Then each node v puts the resulting distance to all skeleton nodes in S within h hops into its label . Since there is a skeleton node on a shortest path from v to u w.h.p., v can reconstruct the distance to u given the label $\lambda(u)$ as follows.

$$o_v(\lambda(u)) = \min\Big(d_h(v,u), \min_{s \in S} d(v,s) + d(s,u)\Big)$$

It remains to balance the trade-off in round complexity of $h \in \widetilde{O}(x)$ between the local search of h rounds and the HYBRID computation of RSSP distances (see Lemma 26), which is optimal for $x \in \widetilde{O}(n^{1/3})$. Note that the size of labels scales in the number of sampled nodes, thus $O\big(\frac{n}{x}\big) = O(n^{2/3})$, however, by deviating from the optimal round complexity, we can decrease the label size. The following proof of the following theorem is given in [14].

▶ **Theorem 27.** *For any* $\zeta \geq 0$ exact *distance oracles and* stateless *routing schemes with labels of size* $O(n^{\frac{2}{3}-\zeta})$ *bits can be computed in* $\widetilde{O}(n^{\frac{1}{3}+\zeta})$ *rounds in the* HYBRID *model w.h.p.*

We can also trade higher stretch for smaller labels. Given that each node u only puts its distance to the *closest* skeleton node into its label $\lambda(u)$ (which must be within h hops due to the sampling property) we can still recover distance oracles with stretch 3 on weighted and $(1+\varepsilon)$ on unweighted graphs with label size only $O(\log n)$. For details of the exact procedure and the proof of the following theorem see [14].

▶ **Theorem 28.** *Distance oracles and* stateless *routing schemes with label-size* $O(\log n)$ *can be computed in* HYBRID *w.h.p. and*

- *stretch 3 in* $\widetilde{O}(n^{1/3})$ *rounds on weighted graphs,*
- *stretch* $1+\varepsilon$ *for* $0<\varepsilon\leq 1$ *in* $\widetilde{O}\big(\frac{n^{1/3}}{\varepsilon}\big)$ *rounds on unweighted graphs.*

References

1 Yehuda Afek, Gad M. Landau, Baruch Schieber, and Moti Yung. The power of multimedia: Combining point-to-point and multiaccess networks. *Information and Computation*, 84(1):97–118, January 1990. doi:10.1016/0890-5401(90)90035-G.

2 Noga Alon, Shlomo Hoory, and Nathan Linial. The moore bound for irregular graphs. *Graphs and Combinatorics*, 18, September 2001. doi:10.1007/s003730200002.

3 Ioannis Anagnostides and Themis Gouleakis. Deterministic Distributed Algorithms and Lower Bounds in the Hybrid Model. In Seth Gilbert, editor, *35th International Symposium on Distributed Computing (DISC 2021)*, volume 209 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:19, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.DISC.2021.5.

4 Arash Asadi, Vincenzo Mancuso, and Rohit Gupta. An sdr-based experimental study of outband d2d communications. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pages 1–9, 2016. doi:10.1109/INFOCOM.2016.7524372.

5 John Augustine, Kristian Hinnenthal, Fabian Kuhn, Christian Scheideler, and Philipp Schneider. Shortest paths in a hybrid network model. In *Proceedings of the Thirty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '20, pages 1280–1299, USA, 2020. Society for Industrial and Applied Mathematics.

6 Clark T. Benson. Minimal regular graphs of girths eight and twelve. *Canadian Journal of Mathematics*, 18:1091–1094, 1966. doi:10.4153/CJM-1966-109-8.

7 Keren Censor-Hillel, Dean Leitersdorf, and Volodymyr Polosukhin. Distance Computations in the Hybrid Network Model via Oracle Simulations. In Markus Bläser and Benjamin Monmege, editors, *38th International Symposium on Theoretical Aspects of Computer Science (STACS 2021)*, volume 187 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1–21:19, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.STACS.2021.21.

8 Keren Censor-Hillel, Dean Leitersdorf, and Volodymyr Polosukhin. On sparsity awareness in distributed computations. In *Proceedings of the 33rd ACM Symposium on Parallelism in Algorithms and Architectures*, SPAA '21, pages 151–161, New York, NY, USA, 2021. Association for Computing Machinery. doi:10.1145/3409964.3461798.

9 Paul Erdős and Miklós Simonovits. Compactness results in extremal graph theory. *Comb*, 2(3):275–288, 1982.

10 Michael Feldmann, Kristian Hinnenthal, and Christian Scheideler. Fast hybrid network algorithms for shortest paths in sparse graphs. In *Proc. of the 24th International Conference on Principles of Distributed Systems (OPODIS 2020)*, pages 31:1–31:16, 2020. doi:10.4230/LIPIcs.OPODIS.2020.31.

11 Kai Han, Zhiming Hu, Jun Luo, and Liu Xiang. Rush: Routing and scheduling for hybrid data center networks. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 415–423, 2015. doi:10.1109/INFOCOM.2015.7218407.

12 Taisuke Izumi and Roger Wattenhofer. Time lower bounds for distributed distance oracles. In Marcos K. Aguilera, Leonardo Querzoni, and Marc Shapiro, editors, *Principles of Distributed Systems*, pages 60–75, Cham, 2014. Springer International Publishing.

13 Fabian Kuhn and Philipp Schneider. Computing shortest paths and diameter in the hybrid network model. In *Proceedings of the 39th Symposium on Principles of Distributed Computing*, PODC '20, pages 109–118, New York, NY, USA, July 2020. Association for Computing Machinery. doi:10.1145/3382734.3405719.

14 Fabian Kuhn and Philipp Schneider. Routing schemes and distance oracles in the hybrid model. *arXiv preprint*, 2022. arXiv:2202.06624.

15 Felix Lazebnik, Vasiliy A. Ustimenko, and Andrew J. Woldar. Upper bounds on the order of cages. *the electronic journal of combinatorics*, pages R13–R13, 1997.

16 Christoph Lenzen. Optimal deterministic routing and sorting on the congested clique. In *Proc. 32nd Symp. on Principles of Distr. Comp. (PODC)*, pages 42–50, 2013.

17 Nancy A Lynch. *Distributed algorithms*. Elsevier, 1996.

18 Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

19 Robert Singleton. On minimal graphs of maximum even girth. *J. Comb. Theory*, 1:306–332, 1966.

20 Jeffrey D. Ullman and Mihalis Yannakakis. High-probability parallel transitive-closure algorithms. *Journal on Computing*, 20(1):100–125, 1991.

21 Jacques Verstraëte. Extremal problems for cycles in graphs. In *Recent trends in combinatorics*, pages 83–116. Springer, 2016.

22 Stefano Vissicchio, Laurent Vanbever, and Olivier Bonaventure. Opportunities and research challenges of hybrid software defined networks. *SIGCOMM Comput. Commun. Rev.*, 44(2):70–75, April 2014. doi:10.1145/2602204.2602216.

23 Guohui Wang, David G. Andersen, Michael Kaminsky, Konstantina Papagiannaki, T.S. Eugene Ng, Michael Kozuch, and Michael Ryan. C-through: Part-time optics in data centers. In *Proceedings of the ACM SIGCOMM 2010 Conference*, SIGCOMM '10, pages 327–338, New York, NY, USA, 2010. Association for Computing Machinery. doi:10.1145/1851182.1851222.

24 Goran Zuzic, Gramoz Goranci, Mingquan Ye, Bernhard Haeupler, and Xiaorui Sun. Universally-optimal distributed shortest paths and transshipment via graph-based ℓ1-oblivious routing. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2549–2579. SIAM, 2022.

A Information Theoretic Concepts

The Shannon entropy of a random variable X can be thought of as the average information conveyed by a realization of X and is defined as follows.

▶ **Definition 29** (Entropy, c.f., [18]). *The Shannon entropy of a random variable $X:\Omega \to S$ is defined as $H(X) := -\sum_{x\in S} \mathbb{P}(X=x) \log\big(\mathbb{P}(X=x)\big)$. For two random variables X,Y the joint entropy $H(X,Y)$ is defined as the entropy of (X,Y). The conditional entropy is $H(X|Y) = H(X,Y) - H(Y)$. The transinformation is defined as $I(X;Y) = H(X) - H(Y|X)$.*

The Entropy $H(X)$ gives a lower bound for expected number of bits required for encoding the state of a random variable. This is entailed by Shannon's [18] source coding theorem.

▶ **Lemma 30** (c.f., [18]). *Given a random variable X with outcomes from some set S and an uniquely decodable code $f : S \to \{0,1\}^*$ with expected code length $\mathbb{E}(|f(X)|)$. Then $\mathbb{E}(|f(X)|) \geq H(X)$.*

In particular, in a two party communication setting (see Definition 8) this implies that $H(X)$ constitutes a lower bound for the worst case number of bits that have to be transmitted from one party that knows the state of X to some party that needs to learn it.

▶ **Corollary 31.** *Bob must receive at least $H(X)$ bits from Alice in expectation, as part of any protocol solving the two party communication problem (Def. 8).*

Proof. Assume, for a contradiction, that we have a protocol $\mathcal{P}$ in which sending *less* than $H(X)$ bits from Alice to Bob always suffices to solve the two party communication problem. Clearly, for any two possible outcomes $x_1, x_2 \in S$ of X, the transcript of the communication occurring between Alice and Bob must be different as otherwise Bob would not be able to distinguish x_1 from x_2. But then we could use the transcript of $\mathcal{P}$ for any given outcome of $x \in S$ of X as uniquely decodable code for x of expected length less than $H(X)$, a contradiction to Lemma 30. ◀

Using information theoretic concepts, the above statement generalizes for a protocol that has a probability of at least p that Bob can successfully decode the state of X after it is terminates.

▶ **Lemma 32.** *Bob must receive at least $p \cdot H(X) - 1$ bits from Alice in expectation, as part of any protocol that solves the two party communication problem (c.f., Def. 8) with probability at least p.*

Proof. We assume that the random variable X has a finite number of outcomes (which is sufficient for our purposes), i.e., $X \in \{x_1, \dots, x_k\}$ for some $k \in \mathbb{N}$. Assuming the outcome $X = x_i$, let y_i be the output that Bob makes after the randomized communication protocol terminates. Then

$$y_i = \begin{cases} x_i, & \text{with probability } p_i \\ x_j \text{ and } j \neq i, & \text{with probability } (1 - p_i), \end{cases}$$

where $p \le p_i \le 1$. That means we have another random variable Y dependent on X, which describes Bob's guess about the state of X. It remains to prove that the information about X that is still contained in Y, is large. This is known as the *transinformation* $I(X;Y)$ and since Bob "learns" the state of Y, at least $I(X;Y)$ must have been transmitted from Alice to Bob. In particular, we want to show $I(X;Y) \ge p \cdot H(X) - 1$. The transinformation (see Def. 29) can also be written as follows

$$I(X;Y) = \sum_{i,j \in [k]} \mathbb{P}(X = x_i, Y = x_j) \cdot \log \frac{\mathbb{P}(X = x_i, Y = x_j)}{\mathbb{P}(X = x_i)\mathbb{P}(Y = x_j)}$$

Analyzing this directly is tricky since the output distribution of Y for the second case, where $Y \neq X$, is not specified (and can not be made such, without losing the generality of the claim). So we have to take a detour by defining a third random variable Z that tells us if the protocol was successful.

$$Z = \begin{cases} 1, & \text{if } y_i = x_i \\ 0, & \text{else.} \end{cases}$$

To simplify the analysis of the transinformation we assume that Bob gets to know Z "for free" and since $H(Z) \le 1$ the additional information about X from learning Y is not significantly reduced. Formally, we first show $I(X;Y) \ge I(X;Y,Z) - H(Z)$ which allows us to analyze $I(X;Y,Z)$ instead.

The conditional entropy $H(A|B)$ describes the amount of "new" information in some random variable A given that we already know random variable B. In the following steps we will use the fact that $H(Z|X,Y) = 0$ since Z is functionally dependent on X and Y and we will the chain rule of entropy $H(A,B) = H(A|B)+H(B)$. We plug this into the alternative characterization of transinformation

$$\begin{aligned}
I(X;Y,Z) &= H(X) - H(X|Y,Z) && \textit{def. of } I(X;Y,Z)\\
&= H(X) - H(X,Z|Y) + H(Z|Y) && \textit{chain rule}\\
&= H(X) - H(Z|X,Y) - H(X|Y) + H(Z|Y) && \textit{chain rule}\\
&= H(X) - H(X|Y) + H(Z|Y) && H(Z|X,Y) = 0\\
&= I(X;Y) + H(Z|Y) && \textit{def. of } I(X;Y)\\
&\leq I(X;Y) + H(Z).
\end{aligned}$$

This implies $I(X;Y) \geq I(X;Y,Z) - H(Z)$, and it remains to show that $I(X;Y,Z)$ is large. The random variable Z helps in the following way. For any x_i we have

$$\mathbb{P}(Y{=}x_i, Z{=}1) = p_i \cdot \mathbb{P}(X{=}x_i) = \mathbb{P}(X{=}x_i, Y{=}x_i, Z{=}1),$$

since $Z = 1$ means that $Y = x_i$ is only possible if $X = x_i$. We obtain

$$\begin{aligned}
I(X;Y,Z) &= \sum_{i,j\in[k],z\in\{0,1\}} \mathbb{P}(X{=}x_i, Y{=}x_j, Z{=}z) \cdot \log \frac{\mathbb{P}(X{=}x_i, Y{=}x_j, Z{=}z)}{\mathbb{P}(X{=}x_i) \cdot \mathbb{P}(Y{=}x_j, Z{=}z)}\\
&\geq \sum_{i\in[k]} \mathbb{P}(X{=}x_i, Y{=}x_i, Z{=}1) \cdot \log \frac{\mathbb{P}(X{=}x_i, Y{=}x_i, Z{=}1)}{\mathbb{P}(X{=}x_i) \cdot \mathbb{P}(Y{=}x_i, Z{=}1)}\\
&= \sum_{i\in[k]} p_i \cdot \mathbb{P}(X{=}x_i) \cdot \log \frac{1}{\mathbb{P}(X{=}x_i)}\\
&\geq p \cdot \sum_{i\in[k]} \mathbb{P}(X{=}x_i) \cdot \log \frac{1}{\mathbb{P}(X{=}x_i)} = p \cdot H(X)
\end{aligned}$$

Finally, we have $I(X;Y) \geq I(X;Y,Z) - H(Z) \geq p \cdot H(X) - 1$. ◀

B Density of Bounded Girth Graphs

We reproduce a few known and conjectured results from extremal graph theory, in particular that the number of edges in cycle-free graphs can be bounded from above and below. We are going to formulate these results in the context and granularity that we require in this article (neglecting constants, in particular). First, there is a long standing conjecture from Erdős and Simonovits [9].[4]

▶ **Conjecture 33** (by [9]). *For any $k \in \mathbb{N}$, there is an n-node graph with girth $\geq 2k+1$ and $\Theta(n^{1+\frac{1}{k}})$ edges.*

It is known that a graph with average degree d and girth $2k + 1$ has $n \in \Omega(d^k)$ nodes due to [2]. This translates into the following lemma:

[4] [9] states in Conjecture 5 that there are graphs without cycles of a fixed length with the claimed density, and conjectures that the same holds for excluding smaller cycles as well (below Theorem 2 of [9]).

▶ **Lemma 34** (c.f., [2]). *Any n-node graph with girth at least $2k+1, k \in \mathbb{N}$ has at most $O(n^{1+\frac{1}{k}})$ edges.*

Conjecture 33 is known to be true for some parameters of k due to [19] and [6].

▶ **Lemma 35** (c.f., [19, 6]). *For $k = 2, 3, 5$ there are n-node graphs with girth $2k+1$ and $\Theta(n^{1+\frac{1}{k}})$ edges.*

There are more general lower bounds for graphs for arbitrary girth by [15] which the survey [21] summarizes as follows:

▶ **Lemma 36** (c.f., [15, 21]). *For any $k \geq 2$ there is a n-node graph with girth $2k+1$ and $\Theta(n^{1+\frac{2}{3k-2}})$ edges if k is even, and $\Theta(n^{1+\frac{2}{3k-3}})$ if k is odd.*

Above we mention only uneven girth, whereas in this paper we are mostly interested in (balanced) bipartite graphs which naturally have even girth. Note that given a graph with girth $2k+1$, one easily obtains a balanced, bipartite graph of even girth $2k+2$ with the same asymptotic order and size by constructing the bipartite double cover.

▶ **Lemma 37.** *Let $G = (V, E)$ be a n-node graph with girth $2k+1$, then there is a balanced, bipartite graph $G' = (V', E')$ with girth $2k+2$, $|V'| = 2|V|$ and $|E'| = 2|E|$.*

Proof. Let $V' := \bigcup_{v \in V}\{v_1, v_2\}$, i.e., for each node $v \in V$ we create two copies. Further, let $E' = \bigcup_{\{u,v\} \in E}\{\{u_1, v_2\}, \{v_1, u_2\}\}$, i.e., for each edge $\{u, v\}$ in E we create two "crossing" edges between the node copies u_1, v_2 and u_2, v_1. Any cycle of G' must form a corresponding cycle in G, by taking the original edge $\{u, v\}$ for each edge $\{u_1, v_2\}$ in that cycle. Thus G' can not have a cycle shorter than $2k + 1$. Further, by construction, we have a (balanced) bipartition of G' given by the nodes with index 1 and 2, respectively. Since G' is bipartite, it can not contain an odd cycle, hence the girth is at least $2k + 2$. ◀

Combining Lemma 37 with Lemma 35 and the n-node clique which has girth 3 and $\Theta(n^2)$ edges, we obtain the following lemma.

▶ **Lemma 38.** *For $\ell = 4, 6, 8, 12$ there are balanced, bipartite n-node graphs with girth ℓ and $\Theta(n^{1+\frac{2}{\ell-2}})$ edges.*

Note that Lemma 38 this is tight, since for any even $\ell \geq 4$ we obtain the upper bounds $\Omega(n^{1+\frac{2}{\ell-2}})$ in Lemma 34 by plugging in the smaller uneven girth $\ell - 1$. For all other even girths we have to fall back on Lemma 36. Combining it with Lemma 37 gives us the lemma below. Note that we do not apply this lemma for girth 10 as we can get the same asymptotic number of edges for the higher (= better) girth 12 from Lemma 38.

▶ **Lemma 39.** *For any even $\ell \geq 14$ there is a balanced, bipartite n-node graph with girth ℓ and $\Theta(n^{1+\frac{4}{3\ell-10}})$ edges if $\ell \equiv 2 \bmod 4$, or $\Theta(n^{1+\frac{4}{3\ell-12}})$ edges if $\ell \equiv 0 \bmod 4$.*

C Proof of Lemma 9

Proof. We derive a protocol $\mathcal{P}$ that uses (i.e., simulates) algorithm $\mathcal{A}$ in order to solve the two-party communication problem. First, we make a few assumptions about the initial knowledge of both parties in particular about the graph G from the node communication problem, you can think of this information as hard coded into the instructions of $\mathcal{P}$. The important observation is that none of these assumptions give Bob any knowledge about X.

Specifically, assume that Alice is given complete knowledge of the topology G and inputs of all nodes in G (in particular the state of X and the source codes of all nodes specified by $\mathcal{A}$). Bob is given the same for the subgraph induced by $V \setminus A$, which means that the state of X remains unknown to Bob (c.f., Def. 7). To accommodate randomization of $\mathcal{A}$, both are given the same copy of a string of random bits (determined randomly and independently from X) that is sufficiently long to cover all "coin flips" used by any node in the execution of $\mathcal{A}$.

Alice and Bob simulate the following nodes during the simulated execution of algorithm $\mathcal{A}$. For $i \in [h-1]$ let $V_i := \{v \in V \mid \text{hop}(v, A) \leq i\}$ be the set of nodes at hop distance at most i from A. Note that $A \subseteq V_i$ for all i. In round 0 of algorithm $\mathcal{A}$, Alice simulates all nodes in A and Bob simulates all nodes in $V \setminus A$. However, in subsequent rounds $i > 0$, Alice simulates the larger set $A \cup V_i$ and Bob simulates the smaller set $B \cup V \setminus V_i$.

Figuratively speaking, in round i Bob will relinquish control of all nodes that are at hop distance i from set A, to Alice. This means, in each round, every node is simulated *either* by Alice *or* by Bob. We show that each party can simulate their nodes correctly with an induction on i. Initially ($i = 0$), this is true as each party gets the necessary inputs of the nodes they simulate. Say we are at the beginning of round $i > 0$ and the simulation was correct so far. It suffices to show that both parties obtain all messages that are sent (in the HYBRID(∞, γ) model) to the nodes they currently simulate.

The communication taking place during execution of $\mathcal{A}$ in the HYBRID(∞, γ) model is simulated as follows. If two nodes that are currently simulated by the *same* party, say Alice, want to communicate, then this can be taken care as part of the internal simulation by Alice. If a node that is currently simulated (w.l.o.g.) by Bob wants to send a message over the *global* network to some node that Alice simulates, then Bob sends that message directly to Alice as part of $\mathcal{P}$, and that message becomes part of the transcript.

Now consider the case where a *local* message is exchanged between some node u simulated by Alice and some node v simulated by Bob. Then in the subsequent round Alice will *always* take control of v, as part of our simulation regime. Thus Alice can continue simulating v correctly as she has all information to simulate all nodes all the time anyway (Alice is initially given all inputs of all nodes). Therefore it is *not* required to exchange any *local* messages across parties for the correct simulation.

After T simulated rounds, Bob, who simulates the set B until the very end (as $T < h$), can derive the state of X from the local information of B with success probability at least p (same as algorithm $\mathcal{A}$). Hence, using the global messages that were exchanged between Alice and Bob during the simulation of algorithm $\mathcal{A}$ we obtain a protocol $\mathcal{P}$ that solves the two party communication problem with probability p. Since total global communication is restricted by $n \cdot \gamma$ bits per round in the HYBRID(∞, γ) model, Alice sends Bob at most $T \cdot n \cdot \gamma$ bits during the whole simulation. ◀

D Proof of Lemma 17

Proof. Let $U := \{u_1, \ldots, u_k, v\}$ be the vertex cut that separates any s_i from any t_j. The shortest *simple* s_i-t_j-path that crosses U via v has length $w_2 + w_0 + h - 1$ *independently* from x_e (simple implies that a path can not "turn around" and go via u_i).

Consider the shortest s_i-t_j-path that does *not* contain v. In the case $x_e = 1$, i.e., $e = \{u_i, t_j\}$ exists in Γ, this s_i-t_j-path is forced to cross U via u_i and then goes directly to t_j via e, and thus has length $w_2 + w_1 + h - 1$.

Let us analyze the length of the s_i-t_j-path that does *not* contain v for the case $x_e = 0$ (i.e., $e \notin E_\Gamma$). Let G' be the subgraph that corresponds to G after removing each edge $e' \in E$ with $x_{e'} = 0$. Then that s_i-t_j-path has to traverse G' to reach t_j. The sub-path from u_i to t_j in G' has to use at least $(\ell{-}1)$ edges, because otherwise $e = \{u_i, t_j\}$ would close a loop of less than ℓ edges in G' (and thus also in G), contradicting the premise that G has girth ℓ. Thus, for $e \notin E_\Gamma$ any s_i-t_j-path that does not contain v has length *at least* $w_2 + (\ell{-}1)w_1 + h - 1$.

We sum up the cases. If $x_e = 1$, then the s_i-t_j-path *not* containing v of length $w_2{+}w_1{+}h{-}1$ is shorter than the one via v of length $w_2 + w_0 + h - 1$, since $w_1 < w_0$. If $x_e = 0$, then the s_i-t_j-path via v of length $w_2 + w_0 + h - 1$ is shorter than the one not containing v of length at least $w_2 + (\ell - 1)w_1 + h - 1$ due to $w_0 < (\ell{-}1)w_1$. ◀