# Sparse Spanners with Small Distance *and* Congestion Stretches

Costas Busch
School of Computer & Cyber Sciences
Augusta University
Augusta, Georgia, USA
kbusch@augusta.edu

Dariusz R. Kowalski
School of Computer & Cyber Sciences
Augusta University
Augusta, Georgia, USA
dkowalski@augusta.edu

Peter Robinson
School of Computer & Cyber Sciences
Augusta University
Augusta, Georgia, USA
perobinson@augusta.edu

## ABSTRACT

Given a graph $G$, a classical problem in graph theory is the construction of a spanner $H$ – a sparse subgraph of $G$ that closely approximates the distances between nodes in $G$. The distance stretch $\alpha$ of $H$ is the factor of how much the distances in $H$ increase versus $G$. Here, we consider sparse spanner constructions that can also preserve the node congestion of routing problems in $G$. The congestion stretch $\beta$ of $H$ is the factor of how much the (smallest) congestion of a routing problem increases in $H$ versus $G$. We introduce the notion of $(\alpha, \beta)$-DC-spanner (i.e., a Distance-Congestion-spanner) that simultaneously controls the stretches for distance and congestion. We show that for expander graphs with $n$ nodes, there is a $(3, O(\log n))$-DC-spanner with $O(n^{5/3})$ edges. We also examine $\Delta$-regular graphs with $\Delta \geq n^{2/3}$, where we show how to obtain a $(3, O(\sqrt{\Delta} \cdot \log n))$-DC-spanner with $O(n^{5/3} \log^2 n)$ edges. Finally, we show that there is a graph such that any optimal size 3-distance spanner has $\Omega(n^{7/6})$ edges and is a $(3, \Omega(n^{1/6}))$-DC-spanner.

## CCS CONCEPTS

• **Theory of computation** → **Sparsification and spanners**; **Routing and network design problems**.

## KEYWORDS

Sparse Spanners, Congestion, Stretch, Expanders

**ACM Reference Format:**
Costas Busch, Dariusz R. Kowalski, and Peter Robinson. 2024. Sparse Spanners with Small Distance *and* Congestion Stretches. In *Proceedings of the 36th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA '24), June 17–21, 2024, Nantes, France.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3626183.3659954

## 1 INTRODUCTION

A classic problem in graph theory is given a graph $G$ to construct a sparse spanner $H$ with distance stretch $\alpha$, where every path $p$ in $G$ has a respective path $p'$ in $H$, with the same source and destination, which is at most $\alpha$ times longer in $H$. Here, we consider the additional property where $H$ has $\beta$ congestion stretch, such that every routing (set of paths) in $G$ with node congestion $C$ has a respective routing in $H$ with node congestion at most $\beta C$. By combining the two stretch properties, we say that $H$ is an $(\alpha, \beta)$-DC-spanner of $G$, where DC-spanner is a shorthand for Distance-Congestion-spanner.

An $(\alpha, \beta)$-DC-spanner controls simultaneously the distance and congestion stretch. We are interested in finding sparse $(\alpha, \beta)$-DC-spanners, i.e., spanners with relatively small number of edges, for small parameters $\alpha$ and $\beta$. This can be particularly useful in network design problems which require a reduced number of edges without sacrificing the quality of the routing with respect to the original graph $G$. It also allows to reduce the total/average size of routing tables (due to sparsity of the used spanner $H$), while maintaining similar quality of considered routing requests (with respective overheads $\alpha$ and $\beta$). These routing requests may use different paths, within a subgraph $H$ of $G$, but they are still valid, at most $\alpha$ times longer and $\beta$ times more congested. See also Section 1.1 for other application examples.

We give the following results as summarized in Table 1:

- *Expander graphs:* In Theorem 2 we show that for expander graphs with $n$ nodes, there is a $(3, O(\log n))$-DC-spanner with $O(n^{5/3})$ edges, where the congestion stretch holds in expectation. For $\Delta$-regular expanders with large degree $\Delta = \Omega(n)$, it is possible to obtain an $(O(\log n), O(\log^3 n))$-DC-spanner with $O(n)$ edges. We also show that expander graphs (of any degree) have an $(O(\log n), O(\log^4 n))$-DC-spanner with $O(n \log n)$ edges.
- *Regular graphs:* In Theorem 3 we prove that any $\Delta$-regular graph, with $\Delta \geq n^{2/3}$, has a $(3, O(\sqrt{\Delta} \cdot \log n))$-DC-spanner with $O(n^{5/3} \log^2 n)$ edges. The result holds with high probability.
- *Lower bound:* Finally, in Theorem 4 we show the existence of a graph with node degrees $\Theta(n^{1/6})$ such that any optimal size 3-distance spanner has $\Omega(n^{7/6})$ edges and is a $(3, \Omega(n^{1/6}))$-DC-spanner.

A basic technique that we use to analyze congestion is that we replace a routing (set of paths) on $G$ with a set of matchings. Since each matching has node congestion 1 it is simpler to analyze congestion on the spanner. If each matching is replaced with a routing of congestion at most $x$ on the spanner $H$, the final congestion will be $O(x \log n)$ on $H$ (Lemma 22). A routing needs $O(n^3)$ distinct matchings, where some may repeat (Lemma 23).

The $\Delta$-regular graph result in Theorem 3 is based on random sampling the edges of $G$ with probability $1/\sqrt{\Delta}$. This allows the spanner to have congestion for a matching equal to the sampled graph degree $\Theta(\sqrt{\Delta})$. On the other hand, classic sparsification methods for distance spanners are not guaranteed to give low degree spanners, and the congestion for a matching may be $\Omega(n)$. Our random sampling method also allows efficient distributed implementation in the LOCAL model.



*SPAA '24, June 17–21, 2024, Nantes, France*

ACM ISBN 979-8-4007-0416-1/24/06.
https://doi.org/10.1145/3626183.3659954

| Result | Number of Edges | Distance Stretch | Congestion Stretch | Assumptions on Δ-Regular Input Graph |
|---|---|---|---|---|
| Theorem 2 | $O\left(n^{5/3}\right)$ | 3 | $O(\log^2 n)$ | Expander |
| [5] | $O(n)$ | $O(\log n)$ | $O(\log^3 n)$ | Expander with node degree $\Delta = \Omega(n)$ |
| [16] | $O(n \log n)$ | $O(\log n)$ | $O(\log^4 n)$ | Expander |
| Theorem 3 | $O(n^{5/3} \log^2 n)$ | 3 | $O(\sqrt{\Delta} \cdot \log n)$ | Δ-regular with $\Delta \geq n^{2/3}$ |
| Theorem 4 | $\Omega(n^{7/6})$ | 3 | $\Omega(n^{1/6})$ | $\Theta(n^{1/6})$ node degree |

**Table 1: Summary of our results, including bounds that follow from prior work.**

We obtain the entries in Table 1 referring to [16] and [5] due to the fact that bounded-degree expander graphs are highly suitable for routing, and allow solving *permutation routing* efficiently, where each node is the source and destination of exactly one message. More specifically, it is shown in Corollary 7.7.3 in [25] that there exist routing paths of length $O(\log n)$ where the edge congestion is limited to $O\left(\frac{\log n \cdot (\log \log n)^2}{\log \log \log n}\right) = O\left(\log^2 n\right)$, which we can translate to good bounds on the node congestion if the node degrees are small. Thus, given a dense regular expander and applying the sparsification algorithm of [16], we obtain an expander with logarithmic node degree, on which we can solve any matching routing problem with $O(\log^3 n)$ node congestion via permutation routing. On the other hand, for regular expanders with very high node degree, i.e., $\Delta = \Omega(n)$, the sparsification procedure of [5] yields an even sparser expander with just $O(n)$ edges and $O(\log^2 n)$ node congestion. For general routing problems, the congestion stretch for these results is multiplied by a factor of $\log n$, thus resulting in the bounds shown in Table 1.

## 1.1 Related Work

Graph spanners that obtain a small distance stretch were introduced in [23], and have since found numerous applications in distributed computing, ranging from routing [18, 24] to achieving more efficient information dissemination in networks [3, 7]. Simple and efficient algorithms are known (e.g., see [4]) that construct a $(2k-1)$-distance stretch spanner with $O(k \cdot n^{1+1/k} \log n)$ edges. This is close to optimal in achieving the best possible tradeoff between their size (i.e., the number of edges) and the provided distance stretch, due to a widely believed conjecture by Erdős [13], Bollobás [2], Bondy and Simonovits [6], which states that, for any $k \geq 1$, there are graphs with $\Omega\left(n^{1+1/k}\right)$ edges and girth at least $2k+2$, which implies that these graphs do not permit any $(2k-1)$-distance spanners as a proper subgraph.

Sparse graphs with good connectivity properties have been widely used for designing distributed fault-tolerant message-efficient protocols, e.g., consensus, in static [10, 14] and dynamic systems [17]. They also are useful is shared memory to schedule access to shared objects by distributed processes, in problems such as re-naming, store-and-collect, write-all and many others [9, 11]. Therefore, finding efficient spanners with small congestion could improve these applications in case the original graph was not sparse or connected enough.

Node congestion plays a significant role in communication in wireless networks, for instance, when routing packets through

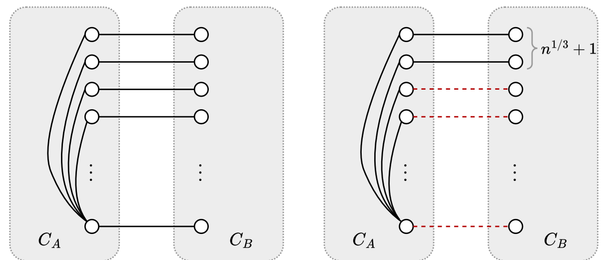



**Figure 1: An example of a vertex fault-tolerant spanner that does not necessarily guarantee low congestion.**

such networks, typically at most one packet can be received and forwarded by a node at a time, see e.g., [12]. Therefore, routing paths with smaller congestion result in lower packet latency and queue sizes of packets forwarded along these paths.

A related construction are fault-tolerant spanners, introduced for general graphs in [8], which extend the standard distance stretch spanners with some additional robustness guarantees. More specifically, an *$f$-vertex fault-tolerant ($f$-VFT) $(2k-1)$-spanner* continues to provide a distance stretch of $(2k-1)$ even after any set $F$ of up to $f$ vertices fail, where the stretch is measured with respect to the residual graph $G \setminus F$. [22] gives efficient algorithms for computing an $f$-VFT spanner of size $\tilde{O}\left(f^{1-1/k} n^{1+1/k}\right)$, which is known to be existentially optimal. Thus, to obtain an $f$-VFT 3-spanner that has an asymptotically equivalent size of $O(n^{5/3})$ as we obtain for DC-spanners with stretch 3, it must hold that $f \leq n^{1/3}$. This upper bound, however, turns out to be too restrictive to provide useful bounds on the node congestion. For instance, consider the graph on the left-hand side of Figure 1. We have two cliques $C_A$ and $C_B$ of size $n/2$ each that are inter-connected via a perfect matching. One possible way of constructing a $f$-VFT spanner, for $f = \lceil n^{1/3} \rceil$, is to include only a subset $M$ of the $\lceil n^{1/3} \rceil + 1$ of the matching edges, and also sparsify the cliques accordingly. However, for the routing problem that corresponds to the perfect matching, some node that is an endpoint of an edge in $M$ must have a congestion of at least $\Omega\left(n^{2/3}\right)$.

## Paper Outline

We continue as follows. In Section 2, we give basic definitions and preliminary results. We present the results on expanders in Section 3. Then, Section 4 presents the result for regular graphs. Next, we show a construction of graphs with small distance stretch but large congestion in Section 5. Section 6 presents and analyzes generic partition of arbitrary routing paths into matchings, which we used in Section 4. Distributed implementation of our spanner constructions are given in Section 7, while the conclusions and open problems are stated in Section 8.

## 2 DEFINITIONS AND PRELIMINARIES

Consider a graph $G = (V, E)$. We will also denote $V(G) = V$ and $E(G) = E$. Let $d_G(u, v)$ denote the distance between a pair of nodes $u, v \in V$. For a path $p$, let $l(p)$ denote the length (number of edges) of $p$. Let $N_G(v)$ denote the 1-neighborhood of $v \in V(G)$, namely, $N_G(v) = \{u : (u, v) \in E(G)\}$. The degree of $v$ is $\delta_G(v) = |N_G(u)|$.

A subgraph $H$ is a spanner graph of $G$ that has the same set of nodes, $V(H) = V(G)$, and it uses a subset of the edges, namely, $E(H) \subseteq E(G)$. A spanner $H$ of $G$ is sparse if the size of edge set $E(H)$ is significantly smaller than the size edge set $E(G)$.

Definition 1 ($\alpha$-distance-spanner). *For $a \geq 1$, a $\alpha$-distance-spanner of $G$ is a spanner graph $H$ such that for every pair of nodes $u, v \in V$,*

$$d_H(u, v) \leq \alpha \cdot d_G(u, v) .$$

A *routing problem* $R$ on $G$ is a set of pairs $R = \{(u_1, v_1), (u_2, v_2), \ldots, (u_k, v_k)\}$, where $u_i, v_i \in V$, and $u_i \neq v_i$, for all $1 \leq i \leq k$. For a pair $(u_i, v_i) \in R$, the node $u_i$ is the source and $v_i$ is the destination. A *routing* $P$ for $R$ is a set of paths $P = \{p_1, p_2, \ldots, p_k\}$, such that path $p_i$ has first node $u_i$ and last node $v_i$. Let $C(P, v) = |\{p_i : p_i \in P \wedge v \in p_i\}|$, denote the number of paths that use node $v$. The (node) congestion of routing $P$, denoted $C(P)$, is the maximum number of paths that use any node of $G$, namely, $C(P) = \max_{u \in V} C(P, u)$. Let $C(R)$ denote the smallest congestion achieved by any routing of $R$. We will use the notation $C_G(R)$ when we explicitly refer to routings of $R$ on graph $G$.

Similar to Definition 1, we can define a spanner related to congestion.

Definition 2 ($\beta$-congestion-spanner). *For $\beta \geq 1$, a $\beta$-congestion-spanner of $G$ is a spanner graph $H$ such that for every routing problem $R$ in $G$,*

$$C_H(R) \leq \beta \cdot C_G(R) .$$

To give some intuition why, for the distance stretch, we consider each routing path individually, whereas, for congestion stretch, we only consider the maximum node congestion and not each node individually. The reason for that is that, for the distance stretch, the substitute routing may replace each original path with a potentially longer path, according to the distance stretch factor. However, a new path may use a node $v$ which has not been used in the original routing, which means that node $v$ had congestion zero in the original routing. In that case, the congestion stretch with respect to $v$ is unbounded. For this reason, we consider the maximum node congestion for the congestion stretch.

For a routing $P = \{p_1, p_2, \ldots, p_k\}$ of some routing problem $R$, we say that a routing $P' = \{p'_1, p'_2, \ldots, p'_k\}$ is an $(\alpha, \beta)$*-stretch substitute of* $P$, if $P'$ is a routing for $R$ and $l(p'_i) \leq \alpha \cdot l(p_i)$ and $C(P') \leq \beta \cdot C(P)$. We can now combine Definitions 1 and 2 into the following spanner definition.

Definition 3 ($(\alpha, \beta)$-DC-spanner). *For $\alpha, \beta \geq 1$, a $(\alpha, \beta)$-DC-spanner of $G$ is a spanner graph $H$ such that for any routing $P$ in $G$ there is a respective $(\alpha, \beta)$-stretch substitute routing $P'$ in $H$.*

We say that an $(\alpha, \beta)$-DC-spanner graph construction is *explicit* if given $G$ and routing $P$, there is a polynomial-time algorithm to compute the spanner $H$ and the respective routing $P'$ in $H$. For an $(\alpha, \beta)$-DC-spanner $H$, we refer to $\alpha$ as the *distance stretch* and to $\beta$ as the *congestion stretch*. Analogously, if $H$ is an $\alpha$-distance spanner ($\beta$-congestion spanner), we refer to parameter $\alpha$ ($\beta$) as the distance stretch (congestion stretch).

Lemma 1. *A $(\alpha, \beta)$-DC-spanner of $G$ is also a $\alpha$-distance-spanner and a $\beta$-congestion-spanner of $G$.*

Proof. Let $H$ be an $(\alpha, \beta)$-DC-spanner of $G$.

We first show that $H$ has distance stretch $\alpha$. Consider the routing problem $R$ involving all edges in $G$, where for each edge $(u, v) \in E$, $u$ is a source and $v$ is a destination. Clearly, the respective routing for $R$ is the set of edges $E$. Since $H$ is an $(\alpha, \beta)$-DC-spanner of $G$, then with respect to $R$, for each edge $(u, v) \in E$ there is an alternative path from $u$ to $v$ in $H$ with length at most $\alpha$. Thus, for any path $p$ in $G$ there is a respective path $p'$ in $H$ which is obtained by replacing each edge of $p$ by its detour in $H$, such that $l(p') \leq \alpha \cdot l(p)$. Hence, $H$ is a $\alpha$-distance-spanner of $G$.

Now, we show that $H$ has congestion stretch $\beta$. Consider an arbitrary routing problem $R$ in $G$ with congestion $C_G(R)$. Let $P$ be the routing of $R$ in $G$ that has congestion $C(P) = C_G(R)$. Since $H$ is an $(\alpha, \beta)$-DC-spanner of $G$, there is a routing $P'$ of $R$ in $H$ that has congestion $C(P') \leq \beta \cdot C(P)$. Consequently, $C_H(R) \leq \beta \cdot C_G(R)$. Hence, $H$ is a $\beta$-congestion-spanner of $G$. □

Next, we show that the $(\alpha, \beta)$-DC-spanner property is not immediately implied from independently proving the distance and congestion stretch properties.

Lemma 2. *There is an infinite family of graphs and parameters $\alpha$, $\beta$, such that each instance $G$ has a spanner $H \subseteq G$ which is an $\alpha$-distance-spanner as well as a $\beta$-congestion-spanner of $G$, but $H$ is not an $(\alpha, \beta)$-DC-spanner of $G$.*

Proof. For any sufficiently large $n$, we define $G$ to be a graph of $2(\alpha - 1)n$ nodes, consisting of sets $A = \{a_1, \ldots, a_n\}$, $B = \{b_1, \ldots, b_n\}$, and $n$ sets $D_i = \{d_{i,1}, \ldots, d_{i,\alpha-1}\}$ $(1 \leq i \leq n)$. There is a perfect matching $M = \{(a_1, b_1), \ldots, (a_n, b_n)\}$ between $A$ and $B$, and the subgraph induced by $A$ (respectively $B$) forms a clique. Moreover, we connect the nodes $a_i, d_{i,1}, \ldots, d_{i,\alpha-1}, b_i$ via a simple path for each set $D_i$. We consider a spanning subgraph $H$, which we obtain by removing all edges in $M$ from $G$ except for the edge $(a_1, b_1)$.

We first argue that $H$ satisfies Definition 1: Clearly, any two nodes in $A$ still have distance 1 in $H$ and the same is true for any two nodes in $B$. For any two nodes $a_i \in A$ and $b_j \in B$, observe that there exists a path of length at most 3 in $H$ via the edge $(a_1, b_1)$, and thus $H$ is a 3-distance spanner of $G$.

Next, we will argue that $H$ is also a 2-congestion spanner of $G$, i.e., Definition 2 holds for $\beta = 2$: Consider any routing problem $R$ and let $P$ be any routing on $G$ that achieves the smallest possible congestion for $R$. We show how to obtain a routing $P'$ on $H$ that has the same congestion as follows: Consider any path $p \in P$. If $p$ does not use an edge that is in $G \setminus H$, we simply add $p$ to $P'$. Otherwise, suppose that $p$ uses some removed edge $(a_i, b_i)$. We obtain a modified path $p'$ from $p$ by simply replacing $(a_i, b_i)$ with the $(\alpha+1)$-length detour along the set $D_i$. As we do not change the parts of the paths that use edges between nodes in $A$ or between nodes in $B$, it is clear that $C_G(P, u) = C_H(P', u)$, for every $u \in A \cup B$. Next, we bound the increase in congestion for any node $d \in D_i$ $(1 \le i \le [n])$. According to the edges of $G$, every additional routing path in $P'$ that used the edge $(a_i, b_i)$ in $P$ instead must go across both $a_i$ and $b_i$ In $P'$, node $d$ may be on up to $\min\{C_G(P, a_i), C_G(P, b_i)\}$ additional paths compared to $P$. Since $C_G(P, a_i)$, $C_G(P, b_i)$, and $C_G(P, d)$ are bounded from above by $C_G(R)$, it follows that $C_G(P', d) \le 2C_G(R)$.

Finally, to see that $H$ is not an $(\alpha, \beta)$-DC-spanner, for any $\beta < \frac{|V(G)|}{2(\alpha-1)}$, consider the routing problem $R = \{(a_1, b_1), \dots, (a_n, b_n)\}$. The optimal routing for $R$ in $G$ has congestion 1, whereas any valid routing $P'$ for $R$ in $H$ must use the edge $(a_1, b_1)$ for each of the $n$ paths. It follows that $C_H(R) \ge n \cdot C_G(R) = \frac{|V(G)|}{2(\alpha-1)} C_G(R)$. □

Now we give a version of Definition 3 that involves probabilities. We modify the definition of the $(\alpha, \beta)$-DC-spanner so that for a routing $P$ in $G$ has a respective routing $P'$ in spanner $H$ with probability $\rho$. Our algorithms provide such probabilistic spanners.

Definition 4 (Probabilistic $(\alpha, \beta, \rho)$-DC-spanner). *For $\alpha, \beta \ge 1$ and $0 < \rho \le 1$, a $(\alpha, \beta, \rho)$-DC-spanner of $G$ is a spanner graph $H$ such that any routing $P$ in $G$ has with probability at least $\rho$ a respective $(\alpha, \beta)$-stretch substitute routing $P'$ in $H$. If the probability $\rho$ is clear from the context, we omit it and simply write $(\alpha, \beta)$-DC-spanner.*

In Section 6, we prove the following theorem, which enables us to obtain a DC-spanner for the general routing problem by leveraging a solution for the special case where the routing problem instance is a matching.

Theorem 1 (Decomposition into Matchings). *Consider a graph $G$ and a subgraph $H \subseteq G$ such that a matching routing problem $M$ has an $(\alpha', \beta')$-substitute routing on $H$ with probability at least $1 - \frac{1}{n^4}$. Then, $H$ is a probabilistic $(\alpha', O(\beta' \log n), 1/n)$-DC-spanner of $G$. Moreover, in the case that there (deterministically) exists such an $(\alpha', \beta')$-substitute routing for every matching routing problem, $H$ is an $(\alpha', O(\beta' \log n))$-DC-spanner of $G$.*

## 3 OBTAINING SPANNERS IN EXPANDERS

In this section, we focus on expander graphs and show how to obtain a spanner that achieves optimal distance stretch *and* also a congestion stretch that is almost optimal in expectation. Formally, we say that an $n$-node graph $G$ is a *(spectral) expander with expansion* $\lambda$, if $\max(|\lambda_2|, |\lambda_n|) \le \lambda$, where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of the adjacency matrix of $G$, ordered by decreasing magnitude. For instance, for Ramanujan graphs [19, 20], which attain near optimal expansion, we have $\lambda \le 2\sqrt{\Delta - 1}$.

Theorem 2. *Consider an $n^{2/3+\epsilon}$-regular expander graph $G$ with spectral expansion $\lambda$, where $\epsilon < \frac{1}{3} - \frac{3\log\log n}{\log n}$ and $\lambda \le o(n^{1/3+2\epsilon})$. There exists a 3-distance stretch spanner with $O\left(n^{5/3}\right)$ edges (w.h.p.), an expected node congestion of $O(\log n)$, and an overall congestion of $O(\log^2 n)$ (w.h.p.). Moreover, if the routing problem is a matching, then the expected node congestion is $1+o(1)$ and the overall congestion is $O(\log n)$.*

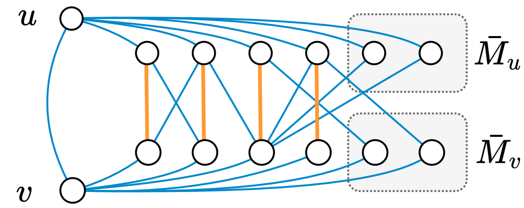



**Figure 2: The construction used in the proof of Lemma 4. The thick orange edges are a maximum matching between the neighborhoods of $u$ and $v$.**

For now, we assume that the routing problem is a matching, i.e., every node occurs at most once as either a source or a destination.

We prove the following technical result by making use of the expander mixing lemma [1], which we restate for completeness.

Lemma 3 (see [1, 15]). *Let $G$ be a $\Delta$-regular graph with spectral expansion $\lambda$. Then, for all subset of nodes $S, T \subseteq V(G)$:*

$$\left| e(S,T) - \frac{\Delta}{n} \cdot |S| \cdot |T| \right| \le \lambda \sqrt{|S| \cdot |T}.$$

Lemma 4. *Consider any two vertices $u$ and $v$ and with neighbors $N_u$ and $N_v$, respectively. Then, there exists a matching of size $\Delta\left(1 - \frac{\lambda n}{\Delta^2}\right)$ between $N_u$ and $N_v$.*

Proof. Consider any matching $M$ between $N_u$ and $N_v$ that is of maximum size, and let $m_0 = \Delta - |M|$. Let $\bar{M}_u \subseteq N_u$ be the subset of $u$'s neighbors that do not have an endpoint in $M$, and define $\bar{M}_v$ similarly. Clearly, we have that $|\bar{M}_u| = |\bar{M}_v| = m_0$. Moreover, by Lemma 3, it must hold that

$$\left| e(\bar{M}_u, \bar{M}_v) - \frac{\Delta}{n} m_0^2 \right| \le \lambda\, m_0 \,, \tag{1}$$

where $e(\bar{M}_u, \bar{M}_v)$ denotes the number of edges between $\bar{M}_u$ and $\bar{M}_v$. Since $M$ was chosen to be of maximum size, it follows that $e(\bar{M}_u, \bar{M}_v) = 0$, and (1) implies that $m_0 \le \frac{\lambda\, n}{\Delta}$. This tells us that $|M| \ge \Delta\left(1 - \frac{\lambda\, n}{\Delta^2}\right)$, as required. □

To construct the spanner $\mathcal{S}$, we sample every edge of the graph independently with probability $\frac{1}{n^\epsilon}$. Consider any edge $\{u, v\}$ that is not in the spanner. Let $M_{u,v}$ be the matching between the neighbors of $u$ and $v$ guaranteed by Lemma 4, and let $M_{u,v}^{\mathcal{S}} \subseteq M_{u,v}$ be the subset of these edges that are part of the spanner.

Lemma 5. *With high probability, for every edge $\{u, v\}$, we have*

$$|M_{u,v}^{\mathcal{S}}| \ge n^{2/3}\,(1 - o(1)) \;. \tag{2}$$

Proof. Let $X = |M_{u,v}^{\mathcal{S}}|$. From Lemma 4 and the upper bound on $\lambda$ in the premise of Theorem 2, we know that $\mathbf{E}[X] \ge n^{2/3}\left(1 - \frac{\lambda}{n^{1/3+2\epsilon}}\right) \ge n^{2/3}(1 - o(1))$. Since each edge is sampled independently, we can apply a standard Chernoff bound (e.g., Theorem 4.5 in [21]) to show that $\mathbf{Pr}\left[X \le (1-\delta) n^{2/3}(1 - o(1))\right] \le \exp\left(-\frac{\delta^2 \cdot n^{2/3}(1-o(1))}{2}\right)$, for any $\delta \in (0, 1)$. In particular, for

$\delta = \frac{c\sqrt{\log n}}{n^{1/3}}$ where $c > 0$ is a suitable constant, we get that, with high probability,

$$X > (1-\delta)n^{2/3}(1-o(1)) \geq n^{2/3}(1-o(1)) ,$$

where the last inequality follows because $\delta = o(1)$. □

Lemma 6. *With high probability, the distance stretch of the spanner is at most* 3.

Proof. It will be sufficient to argue that, for every edge $\{u,v\} \notin \mathcal{S}$, nodes $u$ and $v$ both have an edge in $\mathcal{S}$ to nodes in $M^{\mathcal{S}}_{u,v}$. Let Bad be the event that this does not happen. Consider any edge $\{x,y\} \in M^{\mathcal{S}}_{u,v}$. The probability that the edges $\{u,x\}$ and $\{v,y\}$ are both part of the spanner is $\frac{1}{n^{2\epsilon}}$. Due to Lemma 5, it follows that

$$\mathbf{Pr}[\mathrm{Bad}] \leq \left(1-\frac{1}{n^{2\epsilon}}\right)^{|M^{\mathcal{S}}_{u,v}|} \leq \exp\left(-n^{\frac{2}{3}-2\epsilon}(1-o(1))\right) \leq \frac{1}{n^3} ,$$

where the last inequality follows from the upper bound on $\epsilon$ assumed in the premise of Theorem 2). □

*Choosing the Replacement Paths.* For every edge $\{u,v\} \notin \mathcal{S}$ that is part of the (matching) routing problem, we choose a 3-hop replacement path in $\mathcal{S}$ by uniformly at random by picking one of the available 3-hop paths across $M^{\mathcal{S}}_{u,v}$ to route between $u$ and $v$.

Lemma 7. *The spanner contains at most $O(n^{5/3})$ edges (w.h.p.). Moreover, the expected node congestion for any matching routing problem is at most $1 + o(1)$, and (w.h.p.) the overall congestion is $O(\log n)$.*

Proof. Since we sample each edge independently with probability $\frac{1}{n^{\epsilon}}$, it follows by a standard Chernoff bound that the number of neighbors of any node in the spanner is at most $(1+o(1))\frac{\Delta}{n^{\epsilon}} = (1+o(1))n^{2/3}$ with high probability. Let $\mathcal{E}$ denote the event that this bound holds for all nodes, as well as the bound given in (2), for all every edge $\{u,v\}$. By a union bound, we can assume that $\mathcal{E}$ occurs with high probability, which shows the claimed bound on the size of the spanner.

Consider any edge $\{u,v\} \notin \mathcal{S}$. Recall that the replacement path for $\{u,v\}$ is chosen uniformly from all 3-hop paths connecting $u$ to $v$, where the middle edge is in the matching $M^{\mathcal{S}}_{u,v}$. Thus, the expected amount contributed to the congestion at nodes $x$ and $y$, where $\{x,y\} \in M^{\mathcal{S}}_{u,v}$, is at most $\frac{1}{|M^{\mathcal{S}}_{u,v}|}$, due to being used as a replacement edge for $\{u,v\}$. Each node may be part of such a matching for each one of its neighbors in $\mathcal{S}$. Let $T_w$ be the total congestion at a node $w$. We have

$$\mathbf{E}[T_w \mid \mathcal{E}] \leq \frac{(1+o(1))n^{2/3}}{|M^{\mathcal{S}}_{u,v}|} \leq \frac{1+o(1)}{1-o(1)} \leq 1+o(1) ,$$

where, in the second inequality, we have used the lower bound on $M^{\mathcal{S}}_{u,v}$ provided by (2). To complete the proof, observe that

$$\mathbf{E}[T_w] \leq \mathbf{E}[T_w \mid \mathcal{E}] + \frac{1}{n^2} \cdot \mathbf{E}[T_w \mid \neg\mathcal{E}] \leq 1+o(1) , \qquad (3)$$

since the maximum possible congestion at any node in the case where $\neg\mathcal{E}$ holds is $\Delta$.

To obtain a high probability bound, observe that $T_w$ is a sum of independent random variables, since each replacement path is chosen independently. Thus, it follows by a standard Chernoff bound that $T_w = O(\log n)$, and a union bound shows that the same holds for every node with high probability. □

**Algorithm 1:** Spanner Construction for Regular Graphs

**Input** : $\Delta$-regular graph $G = (V, E)$ with $|V| = n$ and $\Delta \geq n^{2/3}$
**Output**: Probabilistic $(3, O(\log^3 n))$-DC-spanner $H$

```
// Keep random edges
Δ' ← √Δ;
E' ← ∅;
foreach e ∈ E with probability ρ = Δ'/Δ do
    E' ← E' ∪ {e};
G' ← (V, E');
// Reinsert removed edges
c_1 ← constant such that 0 < c_1 < 1 − 1/Δ;
λ ← 2^7 ln^2 n / c_1;
Ê ← edges in G which are (λΔ', c_1Δ)-supported in at least one direction;
E'' ← E \ Ê;
H ← (V, E' ∪ E'');
```

So far, we have assumed that the routing problem is a matching. To obtain a bound on the congestion for general matching problems, we leverage Theorem 1, which shows that the overall congestion stretch is $O(\log^2 n)$ with high probability. This completes the proof of Theorem 2.

## 4 DC-SPANNER FOR Δ-REGULAR GRAPHS

Consider a $\Delta$-regular graph $G = (V,E)$, with $|V| = n$, $\Delta \geq n^{2/3}$.[1] We present an algorithm that constructs a probabilistic $(3, O(\log n))$-DC-spanner with $\widetilde{O}(n^{5/3})$ edges. We continue with describing the basic idea behind the algorithm. The details are in Algorithm 1. We first focus on independent paths of length 1 (matchings), and then in the final theorem we use the result from Section 6 about partitioning arbitrary routing paths into matchings.

*Basic Idea:* Let $\Delta' = \sqrt{\Delta}$. The first step to sparsify $G$ is that each edge is chosen to remain with probability $\Delta'/\Delta$. This gives a graph $G'$ with expected average degree $\Theta(\Delta')$ and $\Theta(n\Delta') = O(n\sqrt{n})$ edges. However, the graph $G'$ may be disconnected, or it may not be able to preserve the congestion to be close to that of the routing problem of $G$. In order to remedy these issues we may need to reinsert some of the removed edges back into $G'$.

We show that most of the removed edges can be replaced with alternative paths of length at most 3 (3-detours) in $G'$. Thus, $G'$ is a 3-distance-spanner for $G$. In fact, in the analysis we show that a removed edge $e$ has multiple alternative 3-detours. A routing on $G$ that uses $e$ can use instead one of the 3-detours picked at random. This allows the congestion to be controlled in $G'$.

However, there could be removed edges that do not have enough 3-detour alternatives to control the congestion. Such removed edges need to be reinserted. The analysis shows that the number of reinserted edges is $\widetilde{O}(n^{5/3})$. This allows the total number of edges to be $O(n\sqrt{n}) + \widetilde{O}(n^{5/3}) = \widetilde{O}(n^{5/3})$. This gives the final spanner $H$.

*Detours:* Consider graph $G$. A *2-detour* with base $\{u,z\}$ and router $v$ is a pair of edges $\{(u,v),(v,z)\}$ from $E(G)$. We say that the base $\{u,z\}$ is *supported* if there is a 2-detour that has it as a base. For

[1] Our result could be easily modified for non-regular graphs with all degrees in $\Theta(\Delta)$.

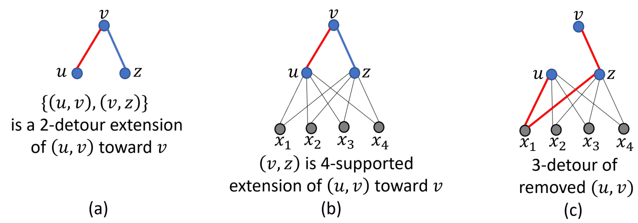



**Figure 3: (a) 2-detour extension; (b) 4-supported extension; (c) alternative 3-detour path for removed edge**

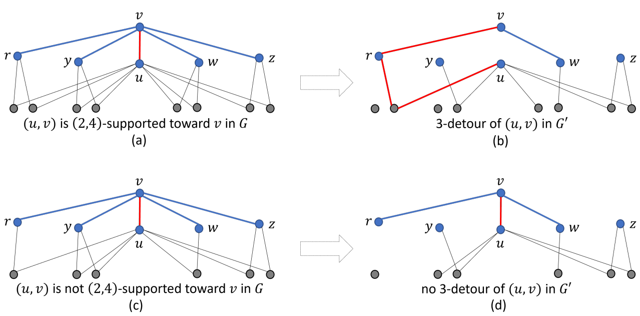



**Figure 4: (a) $(2,4)$-supported edge; (b) alternative 3-detour path; (c) not $(2,4)$-supported edge; (d) no alternative 3-detour**

example, the base $\{u,z\}$ in Figure 3.a is supported. Moreover, we say that a base $\{u,z\}$ is $a$-supported if there is a set $X$ of $a$ distinct nodes such that the 2-detour $(u,x),(x,z)$ exists in $G$ for each $x \in X$. For example, Figure 3.b depicts a 5-supported base $\{u,z\}$ with $X = \{v, x_1, x_2, x_3, x_4\}$.

An *extension* of edge $(u,v)$ toward node/router $v$ is an edge $(v,z)$, where $z \neq u$. The edge $(u,v)$ together with the extension are forming a 2-detour $\{(u,v)(v,z)\}$ with base $\{u,z\}$ (see Figure 3.a). Moreover, we say that the extension $(v,z)$ is $a$-supported if base $\{u,z\}$ is $(a+1)$-supported (where one of the 2-detours is $\{(u,v)(v,z)\}$). Figure 3.b depicts a 4-supported extension $(v,z)$ of $(u,v)$ toward $v$. In our algorithm, the $a$-supported extensions are important for finding multiple alternatives to replace a removed edge with a path of length 3 (a 3-detour). For example, Figure 3.c depicts how to replace a removed edge $(u,v)$ with one of 4 possible 3-detours, namely, the removed edge is replaced by the path $u, x_1, z, v$.

We say that edge $e = (u,v)$ is $(a,b)$-*supported* toward $v$ if the number of $a$-supported extensions of $e$ toward $v$ is $b$. Figure 4.a depicts an example of a $(2,4)$-supported edge $e = (u,v)$ toward $v$. In the example, node $r$ defines base $\{r,u\}$ which is 3-supported. Therefore, the extension $(v,r)$ of $e$ toward $v$ is 2-supported. Similar to $r$, each of the nodes $y,w,z$ defines 3 additional 2-supported extensions for $e$ toward $v$, making $e$ a $(2,4)$-supported edge toward $v$. The $(a,b)$-supported edge has in total $a \cdot b$ 3-detours through its $b$ $a$-supported extensions. Figure 4.b depicts one of such 3-detours of $e$ and how it can be used to replace $e$.

*Reinserted Edges:* An edge $e = (u,v)$ that is in $G$ but has not been selected for $G'$ may need to be reinserted. If $e$ is $(\lambda\Delta', c_1\Delta)$-supported toward $v$ (or $u$) in $G$ and one of its 3-detours toward $v$ (or $u$) still remains in $G'$, then $e$ does not need to be reinserted, since an alternative 3-detour is provided for $e$ in $G'$. Otherwise, $e$ is reinserted to become an edge of $H$. All the reinserted edges together with the edges of $G'$ form the edges of $H$.

As an example, Figures 4.c and 4.d depict an edge $(u,v)$ which is not $(2,4)$-supported in $G$ which may result it having no 3-detour at all in $G'$, and hence $e$ has to be reinserted (if it was removed from $G$).

## 4.1 Analysis

We continue with an analysis of the properties of the graphs in Algorithm 1. Unless otherwise stated, logarithms are base 2. We also use the natural logarithm whenever necessary. From the extension of Taylor sequences for exponential functions, we can obtain the following fact of inequalities.

Fact 1. *For any $x$, $1-x \leq e^{-x}$. For any $x \in [0,1/2]$, $1-x \geq e^{-2x}$. For any $x \geq 1$, $(1-1/x)^x \leq e^{-1}$.*

We will also use the following versions of the Chernoff bounds.

Lemma 8 (Chernoff bounds). *Let $X_1, X_2, \cdots, X_m$ be independent Poisson trials such that, for $1 \leq i \leq m$, $X_i \in \{0,1\}$. Then, for $X = \sum_{i=1}^m X_i$, $\mu = E[X]$, and any $0 \leq \delta \leq 1$, $\mathbf{Pr}[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2\mu}{3}}$, and $\mathbf{Pr}[X \leq (1-\delta)\mu] \leq e^{-\frac{\delta^2\mu}{2}}$, and for any $1 \leq \delta$, $\mathbf{Pr}[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta\mu}{3}}$.*

4.1.1 *Sparsity Analysis of H.* Next, we bound the number of edges $E'$ of $G'$.

Lemma 9. *$|E'| < n\Delta'$ with probability at least $1 - n^{-1}$.*

Proof. For edge $e \in E$, let $X_e$ be the random variable such that $X_e = 1$ if $e$ remains in $G'$ and otherwise $X_e = 0$. From Algorithm 1, we have that $P[X_e = 1] = \rho = \Delta'/\Delta$. We also have that $E' = \{e \in E \mid X_e = 1\}$. Note that since $G$ is $\Delta$-regular, $|E| = n\Delta/2$. Let $\mu = E[|E'|] = \sum_{e\in E} P[X_e = 1] = |E|\Delta'/\Delta = n\Delta'/2$.

Since for sufficiently large $n$, $\mu/3 \geq \ln n$, from Lemma 8 with $\delta = 1$ we have

$$\mathbf{Pr}[|E'| \geq 2\mu] = \mathbf{Pr}[|E'| \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2\mu}{3}} = e^{-\frac{\mu}{3}} \leq e^{-\ln n} = n^{-1}.$$

Hence, $|E'| < 2\mu = n\Delta'$ with probability at least $1 - n^{-1}$. □

The next result gives an upper bound on the size of $E''$, which is the set of edges of $G$ that are not $(\lambda\Delta', c_1\Delta)$-supported (in both directions) in $G$.

Lemma 10. *$|E''| = O(\lambda n^2 \Delta'/\Delta)$.*

Proof. Let $E'' = E \setminus \widehat{E}$. Consider an edge $e = (u,v) \in E''$, that is, $e$ is not $(\lambda\Delta', c_1\Delta)$-supported in either directions. Let $Y$ be the neighbors of $v$ different that $u$.Let $Y = Y_1 \cup Y_2$, where $Y_1$ are the nodes where for each $x \in Y_1$, the extension $(v,x)$ is $\lambda\Delta'$-supported, and $Y_2 = Y \setminus Y_1$. Thus, for each $x \in Y_2$, the extension $(v,x)$ is not $\lambda\Delta'$-supported. Since the degree of $v$ is $\Delta$, $|Y| = \Delta - 1$. Moreover, $|Y_1| < c_1\Delta$. Hence, $|Y_2| \geq \Delta - 1 - c_1\Delta = c_2\Delta$, for constant $c_2 = 1 - c_1 - 1/\Delta$. Note that $0 < c_2 < 1 - 1/\Delta$, since $0 < c_1 < 1 - 1/\Delta$.

Thus, at least $c_2\Delta$ of the extensions of $e$ towards $v$ (or $u$), are not $\lambda\Delta'$-supported. Each such extension $(v, x)$ (or $(u, x)$) corresponds to a "special" base $\{u, x\}$ (or $\{v, x\}$) of at most $\lambda\Delta'$ 2-detours, including the detour with router $v$ (resp. $u$). Hence, each $e \in Y_2$ defines $s = c_2\Delta$ special bases.

Let $B$ be the 2 detours of a special base $b$. Clearly, $|B| \le \lambda\Delta'$. Let $b = (u, x) \in B$. Let $\{(u, y), (y, x)\}$ be a 2-detour of $b$. Edge $(y, x)$ can be an extension of $(u, y)$ that also has special base $b$. Thus, special base $b$ can be shared by at most $2|B| \le 2\lambda\Delta'$ extensions (since each base has 2 edges each of which can be an extension). Potentially, all the up to $2\lambda\Delta'$ edges in $B$ could belong in $E''$. Therefore, the total number of unique special bases in $G$ defined from such extensions is at least $|E''|s/(2\lambda\Delta')$.

Note that the each special base is a pair of nodes. Hence, there are at most $\binom{n}{2} = O(n^2)$ special bases. Therefore, $|E''|s/(2\lambda\Delta') = |E''|c_2\Delta/(2\lambda\Delta') = O(n^2)$. Thus, $|E''| = O(\lambda n^2\Delta'/\Delta)$. □

Lemma 11. *$|E(H)| = O(n\Delta' + \lambda n^2\Delta'/\Delta)$ with probability at least $1 - n^{-1}$.*

Proof. From Algorithm 1 we have $E(H) = E' \cup E''$. From Lemma 9, $|E'| = O(n\Delta')$ with probability at least $1 - n^{-1}$. From Lemma 10, $|E''| = O(\lambda n^2\Delta'/\Delta)$. Hence, with probability at least $1 - n^{-1}$, $|E(H)| = O(n\Delta' + \lambda n^2\Delta'/\Delta)$. □

For $\Delta' = \sqrt{\Delta}$ and $n \ge \Delta \ge n^{2/3}$, from Lemma 11 we get, $|E(H)| = O(n\sqrt{\Delta} + \lambda n^2/\sqrt{\Delta}) = O(n\sqrt{n} + \lambda n^2/\sqrt{n^{2/3}}) = O(\lambda n^{5/3})$. Therefore, we have the following corollary.

Corollary 1. *For $\Delta' = \sqrt{\Delta}$ and $n \ge \Delta \ge n^{2/3}$, $|E(H)| = O(\lambda n^{5/3})$.*

#### 4.1.2 3-*Distance-Spanner Analysis for $H$.*

We will show that for each $e \in \widehat{E}$ there is a path of length 3 between its endpoints in $H$.

Consider an edge $e = (u, v) \in \widehat{E}$. Without loss of generality, $e$ is $(\lambda\Delta', c_1\Delta)$-supported toward $v$ in $G$. There are sets of nodes $A$ and $B$ in $G$ defined as follows. Let $A = \{r_1, \ldots, r_{c_1\Delta}\}$ be the $c_1\Delta$ nodes adjacent to $v$ such each $(v, r_i)$ is a $\lambda\Delta'$-supported extension of $e$ toward $v$ in $G$. (if there are more than $c_1\Delta$ extensions of $e$ toward $v$ that are $\lambda\Delta'$-supported, then for $A$ we pick and fix exactly $c_1\Delta$ nodes, and we will work with these choices for $A$ for the remaining of the analysis.) Let $B = \bigcup_{r_i} B(r_i)$, where $B(r_i)$ is the set that contains the $\lambda\Delta'$ routers of the 2-detours with base $\{u, r_i\}$ in $G$. To illustrate, in Figure 4.a, if $c_1\Delta = 4$ and $\lambda\Delta' = 2$, the set $A$ for edge $(u, v)$ would be $A = \{r, y, w, z\}$, and the set $B$ would be the bottom eight (gray colored) nodes that the nodes in $A$ connect to.

Let $A' \subseteq A$, be the set of nodes of $A$ which remain adjacent to $v$ in $G'$, that is, for each $r_i \in A'$, $(v, r_i) \in E'$. Let

$$Event_1 : \frac{c_1\Delta'}{2} \le |A'| \ .$$

Lemma 12. $\mathbf{Pr}[Event_1] \ge 1 - n^{-3}$.

Proof. For $r_i \in A$, let $X_i$ be the random variable such that $X_i = 1$ if $r_i$ remains in $A'$, while $X_i = 0$ otherwise. The events $X_i = 1$ and $X_j = 1$ are pairwise independent, for $i \ne j$. We have that $Pr[X_i = 1] = \rho$ (recall, $\rho = \Delta'/\Delta$). We have $|A'| = \sum_{r_i \in A} X_i$. Let $\mu = E[|A'|] = \sum_{r_i \in A} Pr[X_i = 1] = |A| \cdot \rho = c_1\Delta \cdot \Delta'/\Delta = c_1\Delta'$. For sufficiently large $n$, $\mu = c_1\Delta' \ge 24 \ln n$. Therefore, from Lemma 8 for $\delta = 1/2$ we have

$$\begin{aligned} \mathbf{Pr}\left[|A'| \le \frac{c_1\Delta'}{2}\right] &= \mathbf{Pr}\left[|A'| \le \frac{\mu}{2}\right] = \mathbf{Pr}[|A'| \le (1-\delta)\mu] \\ &\le e^{-\frac{\delta^2\mu}{2}} = e^{-\frac{\mu}{8}} \le e^{-3\ln n} = n^{-3} \ . \end{aligned}$$

□

Let $B(A') = \bigcup_{r_i \in A'} B(r_i)$ be the set of routers connected to set $A'$ in $G$. For each router $x \in B(A')$, denote by $h(x)$ the number of nodes in $A'$ which are adjacent to $x$ with respect to $G$.

We partition $B(A')$ into sets $B_1, \ldots, B_\zeta$, such that for each $x \in B_i$, $h(x) \in [2^{i-1}, 2^i)$, where $\zeta = 1 + \lceil \log |A'| \rceil \le 2 + \log |A'| \le 2 + \log n$.

Let $S_i = \sum_{x \in B_i} h(x)$. Let $m$ be such that $S_m = \max_i S_i$.

Lemma 13. $|B_m| \ge \frac{\lambda\Delta'|A'|}{2^{m+1}\ln n}$.

Proof. Since each $r_i \in A'$ is adjacent to $\lambda\Delta'$ nodes of $B(A')$, $\sum_{x \in B(A')} h(x) = \lambda\Delta'|A'|$. Thus, $\sum_i S_i = \sum_{x \in B(A')} h(x) = \lambda\Delta'|A'|$. We have for sufficiently large $n$,

$$S_m \ge \frac{\sum_i S_i}{\zeta} \ge \frac{\lambda\Delta'|A'|}{\zeta} \ge \frac{\lambda\Delta'|A'|}{2 + \log n} \ge \frac{\lambda\Delta'|A'|}{2\ln n} \ .$$

Since for each $x \in B_m$, $h(x) < 2^m$, we get

$$|B_m| \ge \frac{S_m}{2^m} \ge \frac{\lambda\Delta'|A'|}{2^m \cdot 2\ln n} = \frac{\lambda\Delta'|A'|}{2^{m+1}\ln n} \ .$$

□

Denote $B_i' \subseteq B_i$ the nodes of $B_i$ that in $G'$ remain connected to a node in $A'$, that is, for each $x \in B_i'$ there is a $r_j \in A'$ such that $(x, r_j) \in E'$. Let

$$Event_2 : |B_m'| \le \frac{c_1\lambda\Delta'}{2^5\ln n} \ .$$

Lemma 14. $\mathbf{Pr}[Event_2 \mid Event_1] \le n^{-3}$.

Proof. Since we assume $Event_1$ is true, $|A'| \ge c_1\Delta'/2$. Consider now a $B_i$, where $1 \le i \le \zeta$. For $x \in B_i$, let $Y_x^i$ be the random variable such that $Y_x^i = 1$ if $x \in B_i'$, and $Y_x^i = 0$, otherwise. Note that the variables $Y_x^i$ are mutually independent for all $x \in B_i$, due to the fact that the edges $h(x)$ that determine these events are disjoint for different $x$. The probability that none of the $h(x)$ edges of $x$ appear in $G'$ is $q = (1-\rho)^{h(x)}$. Hence, with probability $1 - q$, node $x$ stays connected to some node of $A'$ in $G'$. Since $h(x) \ge 2^{i-1}$, from Fact 1 we get

$$\begin{aligned} Pr[Y_x^i = 1] &= 1 - q = 1 - (1-\rho)^{h(x)} \ge 1 - e^{-h(x)\rho} \\ &\ge \min\left(\frac{h(x)\rho}{2}, \frac{1}{2}\right) \ge \min\left(2^{i-2}\rho, \frac{1}{2}\right) \ . \end{aligned} \tag{4}$$

Since $|B_i'| = \sum_{x \in B_i} Y_x^i$, from Equation (4) we get

$$\begin{aligned} E[|B_i'|] &= \sum_{x \in B_i} (Pr[Y_x^i = 1]) \ge \sum_{x \in B_i} \min\left(2^{i-2}\rho, \frac{1}{2}\right) \\ &= |B_i| \cdot \min\left(2^{i-2}\rho, \frac{1}{2}\right) \ . \end{aligned} \tag{5}$$

From Lemma 13 and Equation (5), for $i = m$ we get

$$E[|B_m'|] \ge |B_m| \cdot \min\left(2^{m-2}\rho, \frac{1}{2}\right) \ge \frac{\lambda\Delta'|A'|}{2^{m+1}\ln n} \cdot \min\left(2^{m-2}\rho, \frac{1}{2}\right) \ .$$

Since also $\rho\Delta' = (\Delta')^2/\Delta = 1$, we have $c_1\rho\Delta' = c_1$. Thus, if $2^{m-2}\rho < 1/2$ we get

$$\begin{aligned}\mathbf{E}[|B'_m|] &\geq \frac{\lambda\Delta'|A'|}{2^{m+1}\ln n}\cdot 2^{m-2}\rho = \frac{\lambda\Delta'\rho|A'|}{2^3\log n}\\ &\geq \frac{\lambda\Delta'\rho c_1\Delta'}{2^4\ln n} \geq \frac{c_1\lambda\Delta'}{2^4\ln n}\,. \end{aligned}\tag{6}$$

If $2^{m-2}\rho \geq 1/2$, since $m \leq \zeta \leq 2 + \log|A'|$,

$$\begin{aligned}\mathbf{E}[|B'_m|] &\geq \frac{\lambda\Delta'|A'|}{2^{m+1}\ln n}\cdot\frac{1}{2} = \frac{\lambda\Delta'|A'|}{2^{m+2}\ln n} \geq \frac{\lambda\Delta'|A'|}{2^{4+\log|A'|}\ln n}\\ &= \frac{\lambda\Delta'|A'|}{|A'|2^4\ln n} = \frac{\lambda\Delta'}{2^4\ln n}\,. \end{aligned}\tag{7}$$

From Equations (6) and (7), since $0 \leq c_1 < 1$, we get that $\mathbf{E}[|B'_m|] \geq \frac{c_1\lambda\Delta'}{2^4\ln n}$.

Since $\Delta \geq n^{2/3}$, $\Delta' = \sqrt{\Delta} \geq n^{1/3}$, for sufficiently large $n$ we get $\mu = \mathbf{E}[|B'_m|] \geq 24\ln n$. Hence, from Lemma 8 for $\delta = 1/2$ we have

$$\begin{aligned}\mathbf{Pr}\left[|B'_m| \leq \frac{c_1\lambda\Delta'}{2^5\ln n}\right] &\leq \mathbf{Pr}\left[|B'_m| \leq \frac{\mu}{2}\right] = \mathbf{Pr}[|B'_m| \leq (1-\delta)\mu]\\ &\leq e^{-\frac{\delta^2\mu}{2}} = e^{-\frac{\mu}{8}} \leq e^{-3\ln n} = n^{-3}\,. \end{aligned}$$

□

Lemma 15. *There is a 3-detour for edge $e = (u, v) \in \widehat{E}$ in $H$ with probability at least $1 - 3n^{-3}$.*

Proof. From Lemma 14, $|B'_m| > \frac{c_1\lambda\Delta'}{2^5\ln n}$ with probability at least $1 - n^{-3}$ (given a lower bound on the size of $|A'|$). Since $\lambda = 2^7\ln^2 n/c_1$, we get $|B'_m| > 4\Delta'\ln n$. Moreover, since $\Delta' = \sqrt{\Delta}$, $\rho = \Delta'/\Delta = 1/\Delta'$.

Given the above lower bound on $|B'_m|$, for sufficiently large $n$, the probability that none of the nodes in $B'_m$ remain connected to $u$ in $G'$ is

$$(1-\rho)^{|B'_m|} < \left(1 - \frac{1}{\Delta'}\right)^{4\Delta'\ln n} \leq e^{-4\ln n} \leq n^{-3}\,.$$

From Lemmas 12 and 14, the lower bound for $|B'_m|$ does not hold with probability at most $2n^{-3}$. Thus, the probability that none of the 3-detours of $e$ remain in $G'$ is at most $3n^{-3}$. Hence, with probability at least $1 - 3n^{-3}$, a 3-detour remains in $G'$ and thus in $H$. □

From Lemma 15, the probability that some edge in $\widehat{E}$ does not have a 3-detour in $H$, is at most $|\widehat{E}|\cdot 3n^{-3} \leq n^2\cdot 3n^{-3} \leq 3n^{-1}$. Hence, every edge in $\widehat{E}$ has a 3-detour in $H$ with probability at least $1 - 3n^{-1}$. Thus, we obtain the following corollary.

Corollary 2. *$H$ is a 3-distance-spanner of $G$, with probability at least $1 - 3n^{-1}$.*

*4.1.3 Congestion Analysis for $H$.* We start with a basic result for the node degree in $G'$, which can easily be shown using Lemma 8 (omitted proofs are included in the full paper).

Lemma 16. *Every node in $V(G')$ has degree at most $2\Delta'$, with probability at least $1 - n^{-4}$.*

Lemma 17. *For any matching $M$ in $G$, there is a routing $P$ in $H$ with congestion $C(P) \leq 1 + 2\sqrt{\Delta}$ with high probability.*

Proof. We can write $M = M_1 \cup M_2$, where $M_1 \subseteq E(G')$ and $M_2 \subseteq E(G) \setminus E(G')$. Suppose the respective routings in $H$ are $P_1$ and $P_2$, and $P = P_1 \cup P_2$.

For routing problem $M_1$, any $e \in M_1$ also appears in $H$. Hence, $P_1$ consists of the edges in matching $M_1$, which immediately gives $C(P_1) \leq 1$.

Next, consider routing problem $M_2$. Let $x$ be an arbitrary node in $G'$. For each $e \in M_2$, we find a 3-detour in $G'$ as described in Section 4.1.2. Thus, the congestion on $x$ due to $M_2$, can only be caused by neighbor nodes of $x$ which participate in an edge of $M_2$, and their 3-detours use $x$ in $G'$. Therefore, the maximum congestion on $x$ is at most its degree which by Lemma 16 is bounded by $2\Delta' = 2\sqrt{\Delta}$. Therefore, with high probability, $C(P_2) \leq 2\sqrt{\Delta}$. Consequently, $C(P) \leq C(P_1) + C(P_2) \leq 1 + 2\sqrt{\Delta}$. □

Theorem 3. *For $\Delta$-regular graph $G$, where $\Delta \geq n^{2/3}$, Algorithm 1 returns a probabilistic $(3, O(\sqrt{\Delta}\cdot\log n))$-DC-spanner (w.h.p.) that has $O(n^{5/3}\log^2 n)$ edges (w.h.p).*

Proof. The bound for the number of edges is from Corollary 1 and $\lambda = O(\log^2 n)$.

We continue with the stretches. According to Section 6 and Lemma 23, a routing $P$ can be decomposed into at most $O(n^3)$ matching routing problems in $G$.

From Corollary 2, every matching problem $M$ has a respective substitute-routing $P_M$ in $H$, such that each path in $P_M$ has a length of at most 3. From Lemma 17, the congestion of any specific $P_M$ is at most $O(\sqrt{\Delta})$. Hence, matching $M$ has a respective $(3, O(\sqrt{\Delta}))$-substitute routing $P_M$ on $H$. Therefore, by Theorem 1, the problem $P$ has a $(3, O(\sqrt{\Delta}\cdot\log n))$-substitute routing on $H$.

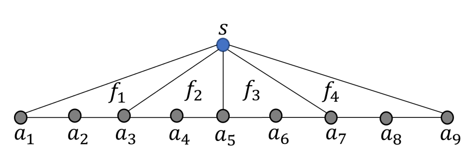
Note that we implicitly take the union bound of the remaining probabilities of the other lemmas and corollaries that we used above, all of which hold with high probability. □

## 5 DISTANCE SPANNERS WITH INHERENT LARGE CONGESTION STRETCH

We show that there is a graph such that if we try to sparsify it and produce a 3-distance spanner, the resulting spanner must have high congestion stretch. We start with the following result that will be used as a building block in the construction of the final graph.

Lemma 18. *There is a graph $G = (V, E)$ with $|V| = 2k + 2$ and $|E| = 3k + 1$, where $k \geq 1$, such that any 3-distance stretch spanner with $|E| - (k + x + 1)/3$ edges, where $0 \leq x \leq 2k - 1$, is also a $(3, \beta)$-congestion spanner with $\beta \geq x/4$.*

Proof. The graph $G = (V, E)$ has $2k$ nodes $a_1, \ldots, a_{2k+1}$ which are connected in a line arrangement with edges $(a_i, a_{i+1})$ for $1 \leq i \leq 2k$; this is the "line" subgraph of $G$ (see figure below for $k = 4$). There is a special node $s$, different than all the $a_i$, that connects to the odd indexed $a_i$ with "ray" edges. Namely, ray edge $r_i$ is $r_i = (s, a_{2i+1})$, where $0 \leq i \leq k$. Overall, $|V| = 2k + 2$ and $|E| = 3k + 1$.



Let "face" $f_i$, where $1 \leq i \leq k$, be the induced subgraph of $G$ with nodes $\{s, a_{2i-1}, a_{2i}, a_{2i+1}\}$. Hence, face $f_i$ consists of rays $r_{i-1}$ and $r_i$ and two consecutive line edges between the rays. Note that the $k$ faces $f_1, \ldots, f_k$ all have $s$ as a common node, and each face $f_i$ contains rays $r_{i-1}$ and $r_i$. Note also that for $2 \leq i \leq k-1$, face $f_i$ shares ray $r_{i-1}$ with $f_{i-1}$ and ray $r_i$ with $f_{i+1}$; we say that $f_i$ is adjacent to $f_{i-1}$ and $f_{i+1}$.

Any 3-distance spanner $H = (V, E_H)$ of $G$ cannot have resulted from three consecutive rays $r_i, r_{i+1}, r_{i+2}$ being removed from $G$, since there would be no substitute 3-distance path for $r_{i+1}$. Thus, $H'$ can have up to $k/3$ rays removed. Moreover, in any face $f_j$ we cannot remove at the same time a line edge and a ray edge, or two line edges, since in either case one of the line edges of $f_j$ will not have a 3-distance substitute path in $H'$. On the other hand, a single line edge of $f_j$ can safely be removed from $G$ if no other edge of $f_j$ is removed at the same time. Therefore, at most $k$ edges can be removed from $G$ to obtain $H$.

Let $E_H = E \setminus E'$. From the above discussion, $0 \leq |E'| \leq k$. We can write $E' = E_1 \cup E_2$, where $E_1$ are the removed line edges from $G$, while $E_2$ are the removed ray edges from $G$. For each $e \in E_1$, the respective face $f_i$ that contains $e$ cannot have any of its other edges removed from $G$, since $H'$ would not be a 3-distance spanner. Hence, $k' = k - |E_1|$ faces have their line edges intact in $H'$. The configuration that allows the maximum number of ray edges to be removed out of the $k'$ faces is when the $k'$ faces are adjacent in a sequence in $G$. In this configuration, the total number of ray edges is at most $k' + 1$, where at most $k' + 1/3$ can be removed. which gives $|E_2| \leq (k'+1)/3 = (k - |E_1| + 1)/3$. Therefore,

$$|E'| = |E_1| + |E_2| \leq |E_1| + (k - |E_1| + 1)/3 = (2|E_1| + k + 1)/3,$$

which gives

$$|E_1| \geq (3|E'| - k - 1)/2.$$

For $|E'| \geq (k + x + 1)/3$, where $0 \leq x \leq 2k-1$, we get $|E_1| \geq x/2$.

Consider a routing problem in $G$ consisting of the edges $E_1$. namely, the end-points of each edge in $E_1$ are a pair of source and respective destination nodes (in any direction). The congestion of such a routing problem in $G$ is at most 2, since two edges in $E_1$ may share a node in the common ray tip. The congestion in $H$ is at least $|E_1|$, since all the alternative 3-distance paths of the $|E_1|$ edges go through node $s$. Therefore, $H$ is a $(3, \beta)$-congestion spanner where $\beta \geq |E_1|/2 \geq x/4$. □

We can prove the following result using the probabilistic method (see the proof in the full paper).

Lemma 19. *For a set $N$ of size $n$, where $n$ is sufficiently large, there are $n$ subsets each of size $(n/17)^{1/6}$ such that:*

(i) *each element of $N$ is in $\Theta(n^{1/6})$ subsets, and*
(ii) *any pair of subsets has at most one common element.*

Theorem 4. *There is a graph $G$ such that any optimal size 3-distance spanner has $\Omega(n^{7/6})$ edges and is a $(3, \Omega(n^{1/6}))$-DC-spanner.*

Proof. To construct the graph $G = (V, E)$, we use $n$ instances of the graph in proof of Lemma 18, which we denote as $I_1, \ldots, I_n$. Each instance $I_i$ has its own separate special node $s_i$, and a set $L_i$ of $2k$ line nodes, where $2k = (n/17)^{1/6}$. We use $n$ nodes for all the lines. Each instance $I_i$ picks a random set of $2k$ out of the $n$ nodes. Then, $I_i$ uses an arbitrary order of the $L_i$ elements to build its line.

From Lemma 19, the instances are edge-disjoint. The total number of nodes is $|V| = 2n$, where $n$ are the special nodes and $n$ are the total lines nodes. The number of edges is $|E| = n \cdot (3k+1) = 3nk + n$.

Consider now a spanner graph $H$ of $G$ with $|E| - nk = n(3k+1) - nk = 2nk + n = \Omega(n^{7/6})$ edges. To preserve the 3-distance property, from Lemma 18, the only possibility is that each instance of $H$ reduces its edges by $k$, that is, $x = 2k-1$. Therefore, Lemma 18 implies that there is a routing instance on $H$ which can cause congestion stretch $\beta \geq x/4 = (2k-1)/4 = \Omega(n^{1/6})$. Note that from Lemma 19, we cannot decrease asymptotically the number of edges further without violating the 3-distance stretch property. □

We would like to note that from Lemma 19, the graph $G$ in Theorem 4 has node degrees $\Theta(n^{1/6})$. While $G$ is not exactly regular, the node degrees are within a constant factor of each other. Similarly, it easy to see that our upper bounds also hold for graphs with node degrees within a constant factor of each other.

## 6 PROOF OF THEOREM 1: DECOMPOSITION OF ROUTING INTO MATCHINGS

This section contains the proof of a technical result, used in Sections 3 and 4, about partitioning arbitrary routing paths into matchings. Given a matching $M$ on $G$, we define a respective routing problem $R_M$ such that each edge on $M$ is a source-destination pair of $R_M$, where we pick arbitrarily one of the incident nodes of the edge to be the source and the other the destination.

Consider now an arbitrary routing problem $R$ on $G$ with a respective routing $P$. We show how to find a substitute routing $P'$ on a spanner $H$ by using substitute routings for matchings on $G$. The benefit of this approach is that each matching $M$ is a routing (i.e. $P_M = M$) for $R_M$ with congestion 1, and it is easier to find a respective substitute routing for $M$ on $H$.

Assume for now that each matching $M$ in $G$ (routing $P_M$) has a respective $(\alpha', \beta')$-substitute routing $P'_M$ in spanner $H$. We first demonstrate the basic approach to convert $P$ to $P'$ for the case $C(P) = 1$ and then generalize for $C(P) \geq 1$.

*The case $C(P) = 1$.* Suppose that $C(P) = 1$. Let $G_P = (V, E_P)$ be the subgraph that consists only of the edges used in $P$. The degree in $G_P$ is at most 2, since each node is used by at most one path in $P$, and if the node is not the source or destination of the path then the path uses two incident edges, otherwise, it uses one incident edge.

We can then perform a coloring of $G_P$ with $m_P \leq 2$ colors. Each color $i$, $1 \leq i \leq m_P$, defines a matching $M_i \subseteq E_P$ on $G_P$. Clearly, $E_P = \bigcup_{1 \leq i \leq m_P} M_i$. Each matching $M_i$ corresponds to a routing problem $R_i$, where each edge in $M_i$ defines a source-destination pair of $R_i$, where we arbitrarily pick one of the edge's nodes to be the source and the other the destination.

Suppose that for each routing problem $R_i$ we have a corresponding routing $P_i$ in $H$ such that for each $p \in P_i$, $l(p) \leq \alpha'$, and also $C(P_i) \leq \beta'$, for appropriate parameters $\alpha'$ and $\beta'$.

We transform $P$ to a routing $P'$ in $H$. For each $p \in P$ we construct a respective path $p' \in P'$. Suppose that $p = (v_1, v_2, \ldots, v_l)$, where $v_k \in V$. If $e_i = (v_i, v_{i+1}) \in M_j$, we replace $e_i$ with the respective path in $P_j$ (by orienting the path from $v_i$ to $v_{i+1}$).

Lemma 20. *If $P$ has congestion $C(P) = 1$, then $P'$ is a $(\alpha', 2\beta')$-stretch substitute of $P$.*

Proof. Let $P = \{p_1, p_2, \ldots, p_k\}$ and $P' = \{p'_1, p'_2, \ldots, p'_k\}$. We first consider the distance stretch of $P'$. Since each edge of a path $p_i \in P$ is replaced by the respective substitute path of length at most $\alpha'$, we get that $l(p'_i) \leq \alpha' \cdot z = \alpha' \cdot l(p_i)$.

We now consider the congestion stretch of $P'$. For a path $p \in P$, let $E_{i,p}$ be the edges in $M_i$ used by $p$. Since $C(P) = 1$, each edge of $M_i$, $1 \leq i \leq m_P$ is used by exactly one path in $P$. Thus, $E_{i,p} \cap E_{i,p'} = \emptyset$, for $p \neq p', p, p' \in P$. Hence, each path $q_e \in P_i$ along edge $e \in M_i$, is used in exactly one path $p'_j \in P'$ which corresponds to $p_j \in P$ that uses $e$.

Since, $C(P_i) \leq \beta'$, we have that the congestion on any node due to $P_i$ does not exceed $\beta'$. Therefore, the congestion on any node due to all the routings $P_1, \ldots, P_{m_P}$ does not exceed $m_P \beta' \leq 2\beta'$. This is also the congestion of routing $P'$ since each path in each $P_i$ is used as a subpath in exactly one path in $P'$. □

*The general case $C(P) \geq 1$.* If we attempt to generalize the concept above for $C(P) = 1$, we run into the problem that an edge of a matching $M_i$ might be used by multiple paths in $P$. This edge-use multiplicity affects adversely the congestion stretch analysis of Lemma 20. To remedy this issue, we create additional matchings such that each matching edge is used in only one path in $P$.

The details are in Algorithm 2. Given routing $P$, the algorithm first creates a sequence of $r$ subgraphs $G_1, \ldots, G_r$, where $G_i = (V, Y_i)$, $Y_i \subseteq E$. Moreover, $Y_{i+1} \subseteq Y_i$, for $1 \leq i < r$. We refer to $G_i$ as the subgraph at level $i$. For any path $p \in P$, each edge $e \in p$ is assigned to one of the subgraphs $G_i$, and we say that the level of pair $(p, e)$ is $i$. Each edge $e \in Y_i$ is assigned to exactly one pair $(p, e)$.

Suppose $G_k$ has degree $d_k$. We perform an edge coloring of $G_k$ with $m_k \leq d_k + 1$ colors. Each color $i$ corresponds to a matching $M_{k,i}$, $1 \leq i \leq m_k$, with a respective routing problem $R_{k,i}$ defined on the edges of the matching. Suppose that for each routing problem $R_{k,i}$ we have a corresponding routing $P_{k,i}$ in $H$ such that for each $p \in P_{k,i}$, $l(p) \leq \alpha'$, and also $C(P_{k,i}) \leq \beta'$, for appropriate parameters $\alpha'$ and $\beta'$.

We transform $P$ to a routing $P'$ in $H$. Suppose that $p = (v_1, v_2, \ldots, v_l)$, where $v_k \in V$. If for $e_j = (v_i, v_{i+1})$ the pair $(p, e_j)$ is at level $i$, we replace $e_j$ with the respective path $q_e \in P_{k,i}$ (by orienting the path from $v_i$ to $v_{i+1}$).

Unless explicitly stated, logarithms are base 2. We continue to show a relation of the degrees of the subgraphs with the congestion.

Lemma 21. $\sum_{k=1}^{r}(d_k + 1) \leq 12 \cdot C(P) \log n$.

Proof. We have that $d_k \leq n - 1$. Since, $Y_{i+1} \subseteq Y_i$, for $1 \leq i < r$, we get that $d_{i+1} \leq d_i$. Divide the $r$ subgraph levels into $\chi = 1 + \lceil \log_2 n \rceil$ ranges $R_1, \ldots, R_\chi$, such that $R_j$ consists of the levels $k$ with $d_k \in [2^{j-1}, 2^j)$. Let $R_\xi$ be the range that maximizes the product $|R_\xi| 2^\xi$. Therefore,

$$\chi \cdot |R_\xi| 2^\xi \geq \sum_{k=1}^{r} d_k \geq \frac{1}{2} \sum_{k=1}^{r} (d_k + 1) .$$

Each edge of $Y_i$ (level $i$) is used by exactly one path in $P$. The congestion due to paths of $P$ at level $i$ is at least $d_k$. Since at least $\xi$ levels in range $R_\xi$ have degree at least $2^{\xi-1}$, and each node with degree $2^{\xi-1}$ at level $i + 1$ must have degree at least $2^{\xi-1}$ at level $i$ also (due to the fact that $Y_{i+1} \subseteq Y_i$), we get that $C(P) \geq |R_\xi| 2^{\xi-1}$. Consequently, for $n \geq 2$,

$$\begin{aligned} \sum_{k=1}^{r}(d_k + 1) &\leq 2\chi \cdot |R_\xi| 2^\xi \leq 4\chi \cdot C(P) \\ &\leq 4 \cdot C(P)(1 + \lceil \log_2 n \rceil) \leq 12 \cdot C(P) \log_2 n . \end{aligned}$$

□

**Algorithm 2:** Substitute Routings via Matchings

**Input:** Routing $P$ in graph $G = (V, E)$ for routing problem $R$; spanner $H = (V, E')$; each matching $M$ in $G$ has a respective $(\alpha', \beta')$-substitute routing in spanner $H$
**Output:** Routing $P'$ for $R$ in connected spanner graph $H = (V, E')$ built of substitute routings of matchings of $G$

```
foreach p ∈ P do
    A_p ← {e : e ∈ p};
r ← 0;
while there is an A_p which is not empty do
    r ← r + 1;
    Y_r ← ⋃_{p∈P} A_p;
    foreach e ∈ Y_r do
        Pick an A_p such that e ∈ A_p;
        Remove e from A_p;
        The level of (p, e) is r;
for k = 1 to r do
    G_k ← (V, Y_r); // subgraph of G induced by edges Y_k
    d_k ← degree of G_k;
    Color the edges of G_k with at most m_k ≤ d_k + 1 colors;
    for i = 1 to m_k do
        M_{k,i} ← matching corresponding to edges of color i;
        R_{k,i} ← routing such that each e = (u, v) ∈ M_{k,i} corresponds to
            a source u and destination v pair;
        P_{k,i} ← routing in H for R_{k,i} such that for each p ∈ P_{k,i},
            l(p) ≤ α' and C(P) ≤ β';
foreach p ∈ P do
    p' ← p; // initialize substitute path p' ∈ P'
    Let p = (v_1, v_2, ..., v_l);
    for j = 1 to l − 1 do
        e ← (v_j, v_{j+1});
        k ← level of (p, e);
        Suppose that e ∈ M_{k,i};
        Let q_e be the respective path of e in P_{k,i} (starting at v_1);
        Replace edge e in p' with q_e;
```

We continue with the main result of this section, which proves Theorem 1 for the case of (deterministic) DC-spanners:

Lemma 22. *$P'$ is a $(\alpha', 12 \cdot \beta' \log n)$-stretch substitute of $P$.*

Proof. Let $P = \{p_1, p_2, \ldots, p_k\}$ and $P' = \{p'_1, p'_2, \ldots, p'_k\}$. Since each edge of a path $p_i \in P$ is replaced by the respective substitute path of length at most $\alpha'$, we get that $l(p'_i) \leq \alpha' \cdot z = \alpha' \cdot l(p_i)$.

We now consider the congestion stretch of $P'$. Let $C_k$ be the congestion due the $k$-level routings $P_{k,1}, \ldots, P_{k,m_k}$. We have

$$C_k \leq \sum_{i=1}^{m_k} C(P_{k,i}) \leq \beta' m_k \leq \beta'(d_k + 1) .$$

Hence, since each $p_e \in P_{k,i}$ is used as a subpath in exactly one path of $P'$, from Lemma 21 we get

$$C(P') \leq \sum_{k=1}^{r} C_k \leq \beta' \sum_{k=1}^{r} (d_k + 1) \leq 12 \cdot \beta' C(P) \log n \ .$$

□

To complete the proof of Theorem 1 for probabilistic DC-spanners, it suffices that we show that we need to consider at most $O(n^3)$ distinct matchings. To see why this is the case, note that we have assumed in the premise of Theorem 1 that, for any given matching routing problem, there exists a suitable $(\alpha', \beta')$-substitute with probability at least $1 - \frac{1}{n^4}$, and hence we can simply take a union bound over all matchings for obtaining a result that holds with high probability.

Lemma 23. *The number of distinct matchings used to construct $P'$ is $O(n^3)$.*

Proof. Since $Y_{i+1} \subseteq Y_i$, $1 \leq i < r$, the number of distinct subgraphs $G_i$ is bounded by $|Y_1| \leq n^2$. Moreover, $d_i \leq d_1$. Thus, each level $i$ gives $m_i \leq d_i + 1 \leq d_1 + 1 \leq n + 1$ distinct matchings. Therefore, the total number of distinct matchings used to build $P'$ is at most $n^2 \cdot (n + 1) = O(n^3)$. □

## 7 DISTRIBUTED SPARSE SPANNER

We now show that Algorithm 1 described in Section 4 lends itself to a distributed implementation in the LOCAL model. First, every node $u$ samples each of its incident edges with probability $\rho$ and then informs each neighbor $v$ if it did sample the edge $(u, v)$. This ensures that every node knows its local neighborhood in the sampled subgraph $G'$. We can implement the part reinserting the removed edges by instructing the nodes to forward all information about $G$ and $G'$ that they learn for the next 3 rounds. Since the decision of whether to reinsert an edge only involves the 3-hop neighborhood of a node in graphs $G'$ and $G$, it follows that each node will obtain the sufficient knowledge for determining locally which of its incident edges are $(\lambda\Delta', c_1\Delta)$-supported. Finally, if node $u$ determines that an edge $(u, v)$ needs to be reinserted, it simply informs its neighbor $v$ using one more round.

Corollary 3. *There exists a $O(1)$-round distributed algorithm in the* LOCAL *model that computes a $(3, O(\log n))$-DC-spanner with high probability, on any $\Delta$-regular graph with $\Delta \geq n^{2/3}$.*

## 8 CONCLUSION

We introduced the problem of simultaneously controlling the distance and congestion stretches in spanner graphs. We presented algorithms for expanders and regular graphs. We also presented a lower bound that relates sparsity and congestion. Several open problems remain. One is to improve the gap between the lower and upper bounds for the 3-distance spanner in $\Delta$-regular graphs. Another is to increase the distance stretches for the spectral expanders and regular graphs; this may give better congestion bounds. Finally, it will be interesting to generalize the results from regular-degree graphs to arbitrary-degree graphs.

## ACKNOWLEDGMENTS

This paper is partly supported by NSF grant CNS-2131538.

## REFERENCES

[1] Noga Alon and Fan RK Chung. 1988. Explicit construction of linear sized tolerant networks. *Discrete Mathematics* 72, 1-3 (1988), 15–19.
[2] I Anderson. 1980. Béla Bollobás, Extremal Graph Theory (Academic Press, 1978), 488 pp.,£ 19· 50. *Proceedings of the Edinburgh Mathematical Society* 23, 1 (1980), 141–142.
[3] Baruch Awerbuch, Oded Goldreich, Ronen Vainish, and David Peleg. 1990. A trade-off between information and communication in broadcast protocols. *Journal of the ACM (JACM)* 37, 2 (1990), 238–256.
[4] Surender Baswana and Sandeep Sen. 2007. A simple and linear time randomized algorithm for computing sparse spanners in weighted graphs. *Random Structures & Algorithms* 30, 4 (2007), 532–563.
[5] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, and Luca Trevisan. 2020. Finding a bounded-degree expander inside a dense one. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 1320–1336.
[6] John A Bondy and Miklós Simonovits. 1974. Cycles of even length in graphs. *Journal of Combinatorial Theory, Series B* 16, 2 (1974), 97–105.
[7] Keren Censor-Hillel, Bernhard Haeupler, Jonathan Kelner, and Petar Maymounkov. 2012. Global computation in a poorly connected world: fast rumor spreading with no dependence on conductance. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*. 961–970.
[8] Shiri Chechik, Michael Langberg, David Peleg, and Liam Roditty. 2009. Fault-tolerant spanners for general graphs. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 435–444.
[9] Bogdan S. Chlebus and Dariusz R. Kowalski. 2005. Cooperative asynchronous update of shared memory. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, Harold N. Gabow and Ronald Fagin (Eds.). ACM, 733–739. https://doi.org/10.1145/1060590.1060698
[10] Bogdan S. Chlebus, Dariusz Rafal Kowalski, and Jan Olkowski. 2023. Deterministic Fault-Tolerant Distributed Computing in Linear Time and Communication. In *Proceedings of the 2023 ACM Symposium on Principles of Distributed Computing, PODC 2023, Orlando, FL, USA, June 19-23, 2023*. ACM, 344–354.
[11] Bogdan S. Chlebus, Dariusz R. Kowalski, and Alexander A. Shvartsman. 2004. Collective asynchronous reading with polylogarithmic worst-case overhead. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, László Babai (Ed.). ACM, 321–330.
[12] Vicent Cholvi, Pawel Garncarek, Tomasz Jurdzinski, and Dariusz R. Kowalski. 2022. Stable routing scheduling algorithms in multi-hop wireless networks. *Theor. Comput. Sci.* 921 (2022), 20–35. https://doi.org/10.1016/J.TCS.2022.03.038
[13] Paul Erdös. 1964. Extremal problems in graph theory. *Publ. House Cszechoslovak Acad. Sci., Prague* (1964), 29–36.
[14] Seth Gilbert, Peter Robinson, and Suman Sourav. 2023. Leader Election in Well-Connected Graphs. *Algorithmica* 85, 4 (2023), 1029–1066.
[15] Shlomo Hoory, Nathan Linial, and Avi Wigderson. 2006. Expander graphs and their applications. *Bull. Amer. Math. Soc.* 43, 4 (2006), 439–561.
[16] Ioannis Koutis and Shen Chen Xu. 2016. Simple parallel and distributed algorithms for spectral graph sparsification. *ACM Transactions on Parallel Computing (TOPC)* 3, 2 (2016), 1–14.
[17] Dariusz R. Kowalski and Miguel A. Mosteiro. 2021. Supervised Average Consensus in Anonymous Dynamic Networks. In *SPAA '21: 33rd ACM Symposium on Parallelism in Algorithms and Architectures, Virtual Event, USA, 6-8 July, 2021*, Kunal Agrawal and Yossi Azar (Eds.). ACM, 307–317.
[18] Christoph Lenzen, Boaz Patt-Shamir, and David Peleg. 2019. Distributed distance computation and routing with small messages. *Distributed Computing* 32 (2019), 133–157.
[19] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. 1988. Ramanujan graphs. *Combinatorica* 8, 3 (1988), 261–277.
[20] Grigorii Aleksandrovich Margulis. 1988. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problemy peredachi informatsii* 24, 1 (1988), 51–60.
[21] Michael Mitzenmacher and Eli Upfal. 2017. *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press.
[22] Merav Parter. 2022. Nearly optimal vertex fault-tolerant spanners in optimal time: sequential, distributed, and parallel. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. 1080–1092.
[23] David Peleg and Alejandro A Schäffer. 1989. Graph spanners. *Journal of graph theory* 13, 1 (1989), 99–116.
[24] David Peleg and Eli Upfal. 1989. A trade-off between space and efficiency for routing tables. *Journal of the ACM (JACM)* 36, 3 (1989), 510–530.
[25] Christian Scheideler. 2006. *Universal routing strategies for interconnection networks*. Vol. 1390. Springer.