# Low Distortion Spanners

SETH PETTIE
University of Michigan

A *spanner* of an undirected unweighted graph is a subgraph that approximates the distance metric of the original graph with some specified accuracy. Specifically, we say $H \subseteq G$ is an $f$-spanner of $G$ if any two vertices $u, v$ at distance $d$ in $G$ are at distance at most $f(d)$ in $H$. There is clearly some tradeoff between the sparsity of $H$ and the *distortion* function $f$, though the nature of the optimal tradeoff is still poorly understood.

In this paper we present a simple, modular framework for constructing sparse spanners that is based on interchangable components called *connection schemes*. By assembling connection schemes in different ways we can recreate the additive 2- and 6-spanners of Aingworth et al. and Baswana et al., and give spanners whose multiplicative distortion quickly tends toward 1. Our results rival the simplicity of all previous algorithms and provide substantial improvements (up to a doubly exponential reduction in edge density) over the comparable spanners of Elkin & Peleg and Thorup & Zwick.

Categories and Subject Descriptors: G.2.2 [**Discrete Mathematics**]: Graph Theory—*Graph algorithms*

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Spanner, metric embedding

## 1. INTRODUCTION

An $f$-spanner of an undirected, unweighted graph $G$ is a subgraph $H$ such that

$$\delta_H(u, v) \leq f(\delta_G(u, v))$$

holds for every pair of vertices $u, v$, where $\delta_H$ is the distance metric w.r.t. $H$. The premier open problem in this area is to understand the necessary tradeoffs between the sparsity of $H$ and the distortion function $f$.[1] The problem of finding a sparse spanner is one in the wider area of *metric embeddings*, where *distortion* is almost universally defined to be multiplicative, of the form $f(d) = t \cdot d$ for some $t \geq 1$. Spanners, however, can possess substantially stronger properties. The recent work of Elkin and Peleg [2004] and Thorup and Zwick [2006] shows that the multiplicative distortion $f(d)/d$ can tend toward 1 as $d$ increases; in this situation the nature of

[1] We are interested in absolute guarantees on the distortion that hold regardless of $G$. A relative guarantee would be of the form: for a given $G$ and $f$, the size of $H$ is within some factor of the smallest $f$-spanner of $G$.

Supported by NSF CAREER grant no. CCF-0746673. Author's address: Seth Pettie, Department of Electrical Engineering and Computer Science, University of Michigan, 2260 Hayward Street, Ann Arbor, MI 48109. Author's email: pettie@umich.edu.

the tradeoff is between the sparsity of the spanner and the *rate* of convergence. It is unknown whether this type of tradeoff is the best possible or whether there exist arbitrarily sparse *additive* spanners, where $f(d) = d + O(1)$ and the tradeoff is between sparsity and the constant hidden in the $O(1)$ term.

### 1.1 Applications.

The original application of spanners was in the efficient simulation of synchronized protocols in unsynchronized networks [Awerbuch 1985; Peleg and Ullman 1989]. Thereafter spanners were used in the design of low-stretch routing schemes using small routing tables [Peleg and Upfal 1989; Awerbuch et al. 1990; Awerbuch and Peleg 1992; Cowen 2001; Cowen and Wagner 2004; Roditty et al. 2008; Thorup and Zwick 2001], computing almost shortest paths in distributed networks [Elkin and Zhang 2006], and in approximation algorithms for geometric spaces[2] [Narasimhan and Smid 2007]. A recent application of spanners is in the design of approximate distance oracles and labeling schemes [Thorup and Zwick 2005; Baswana and Sen 2007; Roditty et al. 2005; Baswana and Kavitha 2006] for arbitrary metrics. Tree spanners have found a number of uses in recent years, such as solving diagonally dominant linear systems [Spielman and Teng 2004] and various approximation algorithms [Fakcharoenphol et al. 2004]. (Tree spanners cannot have any non-trivial distortion in the worst case so weaker notions are used, such as average distortion and expected distortion over a distribution of spanning trees.) In all the applications cited above the quality of the solution is directly related to the quality of the underlying spanners.

### 1.2 Sparseness-Distortion Tradeoffs.

It was observed early on [Peleg and Schaffer 1989; Althöfer et al. 1993] that a spanner has multiplicative distortion $t$ if and only if $f(1) = t$, that is, if the distance between adjacent vertices in $G$ is at most $t$ in the spanner $H$. Althöfer et al. [1993] proved that the sparsest multiplicative $t$-spanner has precisely $m_{t+2}(n)$ edges, where $m_g(n)$ is the maximum number of edges in a graph with $n$ vertices and girth at least $g$.[3] The upper bound follows from a trivial greedy algorithm (similar to Kruskal's minimum spanning tree algorithm) and the lower bound is also simple. In any graph with girth $t+2$, removing any edge shifts the distance of its endpoints from 1 to at least t+1. Thus, the only multiplicative $t$-spanner is the graph itself. It is easy to show that $m_{2k+1}(n)$ and $m_{2k+2}(n)$ are $O(n^{1+1/k})$ and it has been conjectured by Erdős and others (see [Erdős 1963; Thorup and Zwick 2005]) that this bound is asymptotically tight. However, it has only been proved for $k = 1, 2, 3$, and 5; see [Wenger 1991; Thorup and Zwick 2005] for a longer discussion on the girth conjecture. The tradeoff between sparseness and $f(1)$ is fully understood inasmuch as it amounts to proving the girth conjecture. The only other situation that is understood to a similar degree is the threshold $D$ beyond which $f$ is isometric, i.e., where $f(d) = d$, for all $d \geq D$. Bollobas et al. [2006] showed that these so called distance preservers have $\Theta(n^2/D)$ edges. The only known lower bound for

[2]The term spanner is often used to refer to any type of graph that approximates an underlying metric. However, in this paper *spanner* always refers to a subgraph of an undirected graph.
[3]Girth is the length of the shortest cycle.

an intermediate distance was given recently by Woodruff [2006], who showed that $f(d) < d + 2k$ holds only if the spanner has $\Omega(k^{-1}n^{1+1/k})$ edges.

It is perfectly consistent with the girth conjecture and Woodruff's lower bound that there are spanners with size $O(n^{1+1/k})$ and constant additive distortion $f(d) = d+2k-2$, though little progress has been made in proving or disproving their existence. Aingworth et al. [1999] (see also [Dor et al. 2000; Elkin and Peleg 2004; Thorup and Zwick 2006]) showed that there are additive 2-spanners with size $O(n^{3/2})$, which is optimal, and Baswana et al. [2009] gave an additive 6-spanner with size $O(n^{4/3})$. Below the $O(n^{4/3})$ threshold the best known tradeoff is quite weak; it is shown in [Baswana et al. 2009] that there is an $O(n^{1+\epsilon})$-sized spanner with $f(d) = d + O(n^{1-3\epsilon})$, for any $\epsilon \in (0, 1/3)$.

One nice property of additive spanners is that $f(d)/d$ quickly tends toward 1 as $d$ increases. Elkin and Peleg [2004] and Thorup and Zwick [2006] have shown that this property can be achieved without directly addressing the problem of guaranteeing a constant additive distortion. Elkin and Peleg [2004] define an $(\alpha, \beta)$-spanner to be one with distortion $f(d) = \alpha d + \beta$. They show the existence of $(1 + \epsilon, \beta)$-spanners with size $O(\beta n^{1+1/k})$, where $\beta$ is roughly $(\epsilon^{-1}\log k)^{\log k}$. Thorup and Zwick [2006] gave a remarkably simple spanner construction with similar but incomparable properties. They showed that there is an $O(kn^{1+1/k})$-size $(1+\epsilon, O(\lceil 1 + \frac{2}{\epsilon}\rceil^{k-2}))$-spanner, which holds for *all* $\epsilon$ simultaneously. When $\epsilon^{-1}$ is chosen to be $\Theta(d^{1/(k-1)})$ the distortion function is $f(d) = d + O(d^{1-1/(k-1)} + 2^k)$. Notice that the $\beta$ of the Thorup-Zwick spanner is exponentially larger than that of Elkin and Peleg.

### 1.3 Our Results.

In this paper we present a simple, modular framework for constructing low distortion spanners that generalizes much of the recent work on additive and $(\alpha, \beta)$-spanners. In our framework a spanner is expressed as a list of *connection schemes*, which are essentially interchangeable components that can be combined in various ways. This framework simplifies the construction of spanners and greatly simplifies their analysis. Once the list of connection schemes is fixed the size and distortion of the spanner follow from some straightforward linear recurrences. In our framework it is possible to succinctly express the additive 2-spanners of [Aingworth et al. 1999; Elkin and Peleg 2004; Thorup and Zwick 2006] and the additive 2- and 6-spanners of Baswana et al. [2009], as well as the additive 4-spanner suggested in Coppersmith and Elkin [2006]. By properly combining connection schemes we can simultaneously improve the sparseness and distortion of both the Elkin-Peleg and Thorup-Zwick spanners.

One nice feature of our framework is that it is possible to obtain *linear* size spanners with relatively good distortion. Previous to this work the only linear size spanners [Althöfer et al. 1993; Halperin and Zwick 1996] had $O(\log n)$ multiplicative distortion. (The Elkin-Peleg spanners always have $\Omega(n(\epsilon^{-1}\log\log n)^{\log\log n})$ edges. The size of the Thorup-Zwick spanners is $\Omega(n \log n)$, though at this sparsity the guaranteed distortion is quite weak.) We can construct an $O(n)$-size $(5 + \epsilon, \beta)$-spanner, where $\epsilon > 0$ is constant and $\beta = \text{polylog}(n)$, as well as an additive $\tilde{O}(n^{9/16})$-spanner. Under relatively mild assumptions we can actually push the density *and* multiplicative distortion arbitrarily close to 1. For graphs with *quadratic* *expansion* there are $(1+\epsilon,\beta)$-spanners with $(1+\epsilon)n$ edges, for any $\epsilon > 0$. By quadratic expansion we mean that the number of vertices within distance $D$ of any vertex is at least $D^2$.

### 1.4 Organization.

Section 2 introduces some notation and explains how spanners are constructed from layers of connection schemes. Section 3 presents a general framework for analyzing the distortion of spanners based on their underlying connection schemes and in Section 4 we present the algorithms behind the connection schemes. In Section 5 we discuss some open problems.

## 2. NOTATION AND OVERVIEW

Throughout the paper $G = (V, E)$ denotes the input graph. We denote by $\delta_H(u, v)$ and $P_H(u, v)$ the distance from $u$ to $v$ in $H$ and the associated shortest path, respectively. In general there are many shortest paths between two vertices. We insist that if $x, y \in P_H(u, v)$ then $P_H(x, y) \subseteq P_H(u, v)$. Whenever $H$ is omitted it is assumed to be $G$. Our spanner constructions all refer to vertex sets $V_0, V_1, \ldots, V_o$, where $V_0 = V$ and $V_j$ is derived by sampling $V_{j-1}$ with probability $q_j/q_{j-1}$, where $1 = q_0 > q_1 > \cdots > q_o$. Thus, the expected size of $V_j$ is $nq_j$. Let $p_j(v)$ be the closest vertex in $V_j$ to $v$, breaking ties arbitrarily, and let $\text{rad}_j(v) = \delta(v, p_j(v))$. If $j = o+1$ then $p_{o+1}(v)$ is non-existent and $\text{rad}_{o+1}(v) = \infty$ by definition. Let $\text{Ball}(v, r) = \{u : \delta(v, u) < r\}$. We define $\mathcal{B}_j^{\epsilon}(v) = \text{Ball}(v, \epsilon \cdot \text{rad}_{j+1}(v))$, where $\epsilon$ is taken to be 1 if omitted, and $\mathcal{B}_j^{-}(v) = \text{Ball}(v, \text{rad}_{j+1}(v) - 1)$. Let $\bar{\mathcal{B}}_j^{x}(v) = \mathcal{B}_j^{x}(v) \cup \{p_{j+1}(v)\}$, where $x$ is '$-$' or some $\epsilon$. (Note that $\mathcal{B}_j$ is defined w.r.t. the distance to the closest $V_{j+1}$ vertex.)

In Section 4 we describe five connection schemes called $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$, and $\mathbf{x}$. In our framework a spanner can be expressed by choosing an *order* $o$ and a list of the connection schemes employed at each level. For instance, in our compact notation the spanner $\mathbf{ABB}$ employs scheme $\mathbf{A}$ at level zero and scheme $\mathbf{B}$ at levels 1 and 2, where in this case $o = 2$. When a connection scheme is employed at level $j$ it returns a subgraph that connects each $v \in V_j$ to some subset of the vertices in $\bar{\mathcal{B}}_j(v)$; the particulars depend on the scheme used. The overall properties of the spanner are determined by the sequence of connection schemes and, in general, a larger order $o$ leads to a sparser spanner with higher distortion. Figure 1 lists the specifications for the different schemes and Figure 2 lists some of the interesting spanners that can be generated from $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{x}\}^*$.

The connection schemes $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ and $\mathbf{x}$ all produce subgraphs that connect certain pairs of vertices by shortest or almost shortest paths. The three features of a connection scheme we care about are the pairs of vertices to be connected, the guaranteed distortion, and the expected size of the subgraph as a function of the sampling probabilities. The properties of each of the connection schemes are given in Figure 1. (Notice that some of the connection schemes have slightly stronger properties when used at the highest level $o$.) Let us decipher a few of the lines in Figure 1. When $\mathbf{A}$ is used at level $j$ it returns a subgraph $H_j$ such that for $v \in V_j$ and $u \in \bar{\mathcal{B}}_j(v)$, $\delta_{H_j}(v, u) = \delta(v, u)$, and furthermore, the expected size of

| Connection Scheme | Connected Pairs | Distortion | Expected Size |
|---|---|---|---|
| **A** | $V_j \times \bar{\mathcal{B}}_j(\cdot)$ | exact | $nq_j/q_{j+1}$ |
| **B** | $V_j \times \bar{\mathcal{B}}_j^-(\cdot)$ | $d + 2(\log d + 1)$ | $n\sqrt{q_j/q_{j+1}}$ |
| | $V_j \times \bar{\mathcal{B}}_j^{1/2}(\cdot) \cap V_j$ | $d+2$ | $n\sqrt{q_j/q_{j+1}}$ |
| | $V_o \times V_o$ | $d+2$ | $n\sqrt{nq_o}$ |
| **C** | $V_j \times \bar{\mathcal{B}}_j^{1/3}(\cdot) \cap V_j$ | exact | $n + nq_j^2/q_{j+1}^{3/2}$ |
| | $V_o \times V_o$ | exact | $n + n^{5/2}q_o^2$ |
| $\mathbf{D}(r)$ | $V_j \times \bar{\mathcal{B}}_j(\cdot) \cap \mathrm{Ball}(\cdot, r) \cap V_j$ | exact | $nrq_j^2/q_{j+1}$ |
| **x** | $V_j \times \{p_{j+1}(\cdot)\}$ | exact | $n$ |

Fig. 1. The connection schemes. Here $0 \le j \le o$. Schemes **B** and **C** have slightly stronger guarantees at $j = o$.

$H_j$ is on the order of $nq_j/q_{j+1}$. (The notation $V_j \times \bar{\mathcal{B}}_j(\cdot)$ is short for the set of pairs $\{(v,u) : v \in V_j,\ u \in \bar{\mathcal{B}}_j(v)\}$.) Like **A**, schemes **C** and **D** have no distortion but connect fewer pairs of vertices. For $v \in V_j$, scheme **C** only connects the pair $(v,u)$ if $u$ is in *both* $V_j$ and $\bar{\mathcal{B}}_j^{1/3}(v)$. Scheme $\mathbf{D}(r)$ requires $u$ to be in $V_j$, $\bar{\mathcal{B}}_j(v)$, and $\mathrm{Ball}(v,r)$, where $r$ is a given parameter that influences the size of the subgraph. Scheme **B** guarantees two grades of distortion. If $u$ is in both $\bar{\mathcal{B}}_j^{1/2}(v)$ and $V_j$ the additive distortion is 2 and if $u$ is in $\bar{\mathcal{B}}_j^-(v)$ the additive distortion is $2(\log d + 1)$, where $d = \delta(v,u)$.[4] Scheme **x** simply connects every $v \in V_j$ to the nearest vertex $p_{j+1}(v) \in V_{j+1}$. In every case, applying a scheme to level $j$ creates a subgraph that depends solely on $V_j$ and $V_{j+1}$ (since $\mathcal{B}_j(v)$ is defined w.r.t. $V_{j+1}$) and the expected size of this subgraph depends solely on $n, q_j$, and $q_{j+1}$.

In Section 3 we show how connection schemes can be composed in various ways to yield spanners with different sparseness-distortion tradeoffs. The construction and analysis of these spanners is inspired by the *distance emulators* of Thorup and Zwick [2006]. In Section 4 we present the construction algorithms for schemes $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$, and **x**. Schemes $\mathbf{A}, \mathbf{D}$, and **x** are trivial but surprisingly powerful. Scheme **B** uses the generic *path buying* algorithm of Baswana et al. [2009] and scheme **C** is based on the pairwise distance preservers of Coppersmith and Elkin [2006].

## 3. MODULAR SPANNER CONSTRUCTION

Before describing our construction in its full generality let us walk through a relatively small example that illustrates all the major concepts. The spanner construction corresponding to the encoding **ABB** begins by sampling vertex sets $V = V_0 \supset V_1 \supset V_2$, where $\mathbb{E}[|V_1|] = q_1 n$ and $\mathbb{E}[|V_2|] = q_2 n$. It returns the spanner $H = H_0 \cup H_1 \cup H_2$, where $H_0$ is the subgraph returned by connection scheme **A** applied to the zeroth level, and $H_1$ and $H_2$ are the subgraphs returned by **B** applied to levels 1 and 2. By the properties of schemes **A** and **B** (refer to Figure 1), $\mathbb{E}[|H|]$ is on the order of $n/q_1 + n\sqrt{q_1/q_2} + n\sqrt{nq_2}$. Thus, regardless of how we analyze

[4] The constants "1/2" and "1/3" appearing in schemes **B** and **C** are chosen to satisfy the following properties. If $u \in \mathcal{B}_j^{1/2}(v)$ then $v \in \mathcal{B}_j(u)$, and if $u \notin \mathcal{B}_j(v)$ then $\mathcal{B}_j^{1/3}(v)$ and $\mathcal{B}_j^{1/3}(u)$ are disjoint.

| Encoding | Distortion $f(d)$ | Size | Notes |
|---|---|---|---|
| $\mathbf{A}^2$ or $\mathbf{B}$ | $d+2$ | $O(n^{3/2})$ | (1) |
| $\mathbf{AC}$ | $d+4$ | $O(n^{3/2}),\ \Omega(n^{4/3})$ | (2) |
| $\mathbf{AB}$ | $d+6$ | $O(n^{4/3})$ | (3) |
| $\mathbf{A}^{o+1}$ | $d+O(d^{1-1/o}+3^o)$ | $O(on^{1+1/o})$ | (4) |
| not appl. | $d+O(d^{1-1/o}+2^o)$ | $O(on^{1+1/o})$ | |
| $\mathbf{AB}^2$ | $d+O(\sqrt{d})$ | $O(n^{6/5})$ | **new** |
| $\mathbf{AB}^2\mathbf{C}$ | $d+O(d^{2/3})$ | $O(n^{25/22})$ | **new** |
| $\vdots$ | | | |
| $\mathbf{AB}^2\mathbf{C}^{o-2}$ | $d+O(od^{1-1/o}+o^o)$ | $O(on^{1+\frac{(3/4)^{o-2}}{7-2(3/4)^{o-2}}})$ | **new**, (5) |

Linear or Near-Linear Size Spanners:

| | | | |
|---|---|---|---|
| not appl. | $O(d\log n)$ | $O(n)$ | (6) |
| $\mathbf{AD}^{\log\log n}$ | $(5+\epsilon)d+\beta$ | $O(n)$ | **new**, (7) |
| $\mathbf{xCC}$ | $d+\tilde{O}(n^{9/16})$ | $O(n)$ | **new** |
| $\mathbf{xD}^{\log\log n}$ | $(1+\epsilon)d+\beta'$ | $(1+\epsilon)n$ | **new**, (8) |
| not appl. | $(1+\epsilon)d+\beta''$ | $O(n\beta'')$ | (9) |
| $\mathbf{AC}^{O(\log\log\epsilon^{-1})}\mathbf{D}^{\log\log n}$ | $(1+\epsilon)d+\beta'''$ | $O(n\log\log(\epsilon^{-1}\log\log n))$ | **new**, (10) |

(1) The additive 2-spanners of Aingworth et al. [1999], Dor et al. [2000], Elkin and Peleg [2004], and Thorup and Zwick [2006] differ only in the details; the encoding $\mathbf{A}^2$ captures Thorup and Zwick's construction. The additive 2-spanner $\mathbf{B}$ of Baswana et al. [2009] is quite different.

(2) The additive 4-spanner $\mathbf{AC}$, has, by Coppersmith and Elkin's analysis [2006], at most $O(n^{3/2})$ edges but could have as few as $\Theta(n^{4/3})$.

(3) The additive 6-spanner $\mathbf{AB}$ is from Baswana et al. [2009].

(4) Thorup and Zwick [2006] analyzed two spanners with size $O(on^{1+1/o})$ and additive distortion $O(d^{1-1/o})$. The one that fits within our notational framework ($\mathbf{A}^{o+1}$) is slightly weaker inasmuch as the sublinear additive distortion becomes apparent for distances greater than $3^o$ rather than $2^o$.

(5) The exponent $1+\frac{(3/4)^{o-2}}{7-2(3/4)^{o-2}}$ is always strictly less than $1+(3/4)^{o+3}$. For $o=\log_{4/3}\log n - O(1)$ the spanner size is $O(n\log\log n)$.

(6) The standard $O(n)$-size, $O(\log n)$-spanners for weighted [Althöfer et al. 1993] and unweighted graphs [Halperin and Zwick 1996; Peleg 2000] do not fit within our framework.

(7) Here $\beta = O(\epsilon^{-1})^{\log\log n}$.

(8) These bounds hold for graphs with *quadratic* expansion, meaning the number of vertices at distance $D$ from any vertex is at least $D^2$. Here $\beta' = O(\epsilon^{-1}\log\log n)^{\log\log n}$.

(9) In Elkin and Peleg's spanners [2004], $\beta'' = (\epsilon^{-1}\log\log n)^{\log\log n}$.

(10) Here $\beta''' = O(\epsilon^{-1}\log\log n)^{\log\log n}$.

Fig. 2. Some of the spanners generated by $\{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}(\cdot),\mathbf{x}\}^*$.

the distortion of $\mathbf{ABB}$, it is wisest to choose $q_1 = n^{-1/5}$ and $q_2 = n^{-3/5}$, making $\mathbb{E}[|H|] = O(n^{6/5})$.

To analyze the distortion of $\mathbf{ABB}$, let $v$ and $v'$ be vertices at distance $d = \delta(u,v)$, which, we assume for convenience is square: $d = \Delta^2$ for some integer $\Delta$. Let $v_\ell \in P(v,v')$ be the vertex for which $\delta(v,v_\ell) = \ell\Delta$, so $v = v_0$ and $v' = v_\Delta$. To travel from $v$ to $v'$ in $H$ we will first categorize all segments $(v_\ell,\ldots,v_{\ell+1})$ as being *successful* or *failed*. A successful segment is one for which $\delta_{H_0\cup H_1}(v_\ell,v_{\ell+1}) \le \Delta+6$. For unsuccessful/failed segments we have the alternative guarantee that

$\delta_{H_0 \cup H_1}(v_\ell, p_2(v_\ell)) \le 2\Delta + 5$. Let $(v_{\ell,0}, v_{\ell,1}, \ldots, v_{\ell,\Delta})$ be a segment, where $v_{\ell,0} = v_\ell$ and $v_{\ell,\Delta} = v_{\ell+1}$. Some prefix and suffix of the segment will be present in $H_0$. Let $z = v_{\ell,s}$ and $z' = v_{\ell,\Delta-s'}$ be the first and last vertices, respectively, for which the edges $(z, v_{\ell,s+1})$ and $(v_{\ell,\Delta-s'-1}, z')$ do not appear in $H_0$. By definition of the connection scheme $\mathbf{A}$, $H_0$ contains a shortest path from $z$ to every $u \in \mathcal{B}_0(z)$ and a shortest path from $z$ to $p_1(z)$. Since $(z, v_{\ell,s+1}) \notin H_0$, this implies that $\delta_{H_0}(z, p_1(z)) = 1$. The same reasoning shows $\delta_{H_0}(z', p_1(z')) = 1$. We may repeat the same argument, using the relationship between $p_1(z)$ and $p_1(z')$ within $H_1$ in the same way we reasoned about consecutive vertices in $P(v, v')$ w.r.t. $H_0$. By definition of the scheme $\mathbf{B}$, if $p_1(z') \in \mathcal{B}_1^{1/2}(p_1(z))$ or $p_1(z) \in \mathcal{B}_1^{1/2}(p_1(z'))$ then $\delta_{H_1}(p_1(z), p_1(z')) \le \delta(p_1(z), p_1(z')) + 2$. If this is the case we will call the segment *successful*. If the segment failed then $\mathrm{rad}_2(p_1(z))$ and $\mathrm{rad}_2(p_1(z'))$ must both be at most $2 \cdot \delta(p_1(z), p_1(z'))$, and, as a consequence, $\delta_H(v_\ell, p_2(v_\ell)) \le \delta_H(v_\ell, p_1(z)) + 2\delta_H(p_1(z), p_1(z'))$. Thus, for a successful segment the distance from $v_\ell$ to $v_{\ell+1}$ in $H$ is $\delta(v_\ell, p_1(z)) + \delta(p_1(z), p_1(z')) + 2 + \delta(p_1(z'), v_{\ell+1}) \le (s+1) + (\delta(v_\ell, v_{\ell+1}) - s - s' + 2) + 2 + (s'+1) = \delta(v_\ell, v_{\ell+1}) + 6$. For an unsuccessful segment $\delta_H(v_\ell, p_2(v_\ell)) \le \delta_H(v_\ell, p_1(z)) + 2\delta_H(p_1(z), p_1(z')) \le (s+1) + 2(\delta(v_\ell, v_{\ell+1}) - s - s' + 2) \le 2\delta(v_\ell, v_{\ell+1}) + 5$. Given these bounds on successful and unsuccessful segments we can bound $\delta_H(v, v')$ as follows. If there are no failed segments then $\delta_H(v, v') \le \Delta(\Delta + 6) = d + O(\sqrt{d})$. In general, let us redefine $z$ and $z'$ as, respectively, the first vertex of the first failed segment and the last vertex of the last failed segment. In a similar fashion we redefine $s$ and $s'$ to be the number of segments between $v$ and $z$ and $v'$ and $z'$, respectively. Then, from the bounds established above, $\delta_H(v, z) \le s(\Delta+6)$, $\delta_H(v', z') \le s'(\Delta+6)$, and both $\delta_H(z, p_2(z))$ and $\delta_H(z', p_2(z'))$ are at most $2\Delta + 5$. By the properties of scheme $\mathbf{B}$ applied to the second level, $\delta_{H_2}(p_2(z), p_2(z')) \le \delta(p_2(z), p_2(z')) + 2$, which is at most $2(2\Delta+5) + (\Delta - s - s')\Delta + 2$. By concatenating all the paths we see that $\delta_H(v, v') \le (s+s')(\Delta+6) + 4\Delta + 12 + (\Delta - s - s')\Delta$, which is maximized when $s + s' = \Delta - 1$. Thus $\delta_H(v, v') \le \Delta^2 + 10\Delta + 12 = d + O(\sqrt{d})$. This concludes the distortion analysis of $\mathbf{ABB}$.

*Remark* 3.1. The analysis above would go through in much the same way had we encoded the spanner by $\mathbf{ABC}$. The analysis of the distortion would be identical, except that $H_2$ would preserve the distance between $p_2(z)$ and $p_2(z')$ without an additive error of 2. However, the expected size of the spanner would now be on the order of $n/q_1 + n\sqrt{q_1/q_2} + n^{5/2}q_2^2$, which is minimized at $q_1 = n^{-3/14}$ and $q_2 = n^{-9/14}$. Thus, the size of the $\mathbf{ABC}$ spanner would be $O(n^{17/14})$, which is slightly worse than $\mathbf{ABB}$'s size of $O(n^{6/5})$.

We can generalize the distortion analysis above to any spanner defined by a finite string $\tau \in \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}(\cdot)\}^{o+1}$. Let $v, v'$ be two vertices at distance $\Delta^o$, for some integer $\Delta$. To travel from $v$ to $v'$ in the spanner $H = H_0 \cup \cdots \cup H_o$ we divide up $P(v, v')$ into segments of length $\Delta^{o-1}$ and categorize each segment as a success or failure. A successful segment $(v_\ell, \ldots, v_{\ell+1})$ is one for which $\delta_H(v_\ell, v_{\ell+1})$ is *short*; failed segments have the guarantee that $\delta_H(v_\ell, p_o(v_\ell)) = \delta(v_\ell, p_o(v_\ell))$ is *short*. The term "short" here reflects a function that depends on $\Delta$, the encoding $\tau$, and the length of the segment. Specifically, given some fixed $\tau$, $\mathsf{S}^j_\Delta$ and $\mathsf{F}^j_\Delta$ are selected such that if $\delta(v_\ell, v_{\ell+1}) = \Delta^j$, then either $\delta_H(v_\ell, v_{\ell+1}) \le \mathsf{S}^j_\Delta$ or $\delta_H(v_\ell, p_{j+1}(v_\ell)) \le \mathsf{F}^j_\Delta$. Using a generalized form of our analysis of $\mathbf{ABB}$ spanners, we show how

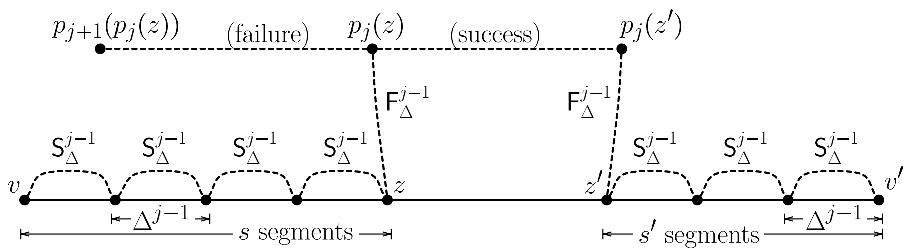


Fig. 3. The vertices $v$ and $v'$ are at distance at most $\Delta^j$. Either $\delta_H(v, v') \leq \mathsf{S}_\Delta^j$ or $\mathrm{rad}_{j+1}(v) = \delta(v, p_{j+1}(v)) \leq \mathsf{F}_\Delta^j$. The distance in $H$ from $p_j(z)$ to $p_j(z')$ (success) or $p_{j+1}(p_j(z))$ (failure) depends on $\tau(j)$.

$\mathsf{S}_\Delta^j$ and $\mathsf{F}_\Delta^j$ can be expressed in terms of $\mathsf{S}_\Delta^{j-1}$ and $\mathsf{F}_\Delta^{j-1}$. In Lemma 3.3 we derive recursive expressions for $\mathsf{S}_\Delta^j$ and $\mathsf{F}_\Delta^j$ for spanners based on schemes $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$. Lemmas 3.5 and 3.6 solve these recurrences for certain classes of spanners and Thereoms 3.7–3.11 illustrate the sparseness-distortion tradeoffs that can be achieved with different combinations of connection schemes.

*Definition* 3.2. **(Success and Failure)** Let $H$ be a spanner defined by some finite string $\tau \in \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{x}\}^*$. We define $\mathsf{S}_\Delta^j$ and $\mathsf{F}_\Delta^j$ to be minimal such that for any two vertices $v, v'$ at distance at most $\Delta^j$ , where $0 \leq j \leq o$, at least one of the following inequalities holds:

$$\delta_H(v, v') \leq \mathsf{S}_\Delta^j \qquad \text{or} \qquad \delta_H(v, p_{j+1}(v)) \leq \mathsf{F}_\Delta^j$$

It is assumed that if $\tau$ includes the connection scheme $\mathbf{D}(\cdot)$ then these bounds only hold if $\Delta$ is below some threshold.

Note that any two vertices at distance at most $\Delta^o$ must be connected in $H$ by a path of length at most $\mathsf{S}_\Delta^o$, that is, every such path must be a success. Such a path cannot fail because $V_{o+1}$ does not exist, and, therefore $\delta_H(v, p_{o+1}(v))$ is undefined. This simply reflects the fact that in our connection schemes, all vertices in $V_o$ are connected by (nearly) shortest paths in $H$.

Lemma 3.3 shows that $\mathsf{S}$ and $\mathsf{F}$ are bounded by some straightforward recurrences. It only considers spanners that employ scheme $\mathbf{A}$ at the zeroth level, which is generally the wisest choice.

Lemma 3.3. **(Recursive Expressions)** *Consider a spanner defined by* $\tau =$

$\mathbf{A} \cdot \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}(\cdot)\}^o$. *Then* $\mathsf{S}^0_\triangle = \mathsf{F}^0_\triangle = 1$ *holds for all* $\Delta$ *and:*

$$\mathsf{F}^j_\triangle \leq \begin{cases} 3\mathsf{F}^{j-1}_\triangle + \Delta^j & \text{for } \tau(j) \in \{\mathbf{A}, \mathbf{D}(r)\}, \text{ provided } r \geq 2\mathsf{F}^{j-1}_\triangle + \Delta^j \\ 5\mathsf{F}^{j-1}_\triangle + 2\Delta^j & \text{for } \tau(j) = \mathbf{B} \\ 7\mathsf{F}^{j-1}_\triangle + 3\Delta^j & \text{for } \tau(j) = \mathbf{C} \end{cases}$$

$$\mathsf{S}^j_\triangle \leq \text{max of } \Delta\mathsf{S}^{j-1}_\triangle \text{ and}$$

$$\begin{cases} (\Delta-1)\mathsf{S}^{j-1}_\triangle + 4\mathsf{F}^{j-1}_\triangle + \Delta^{j-1} + 2 & \text{for } \tau(j) = \mathbf{B} \\ (\Delta-1)\mathsf{S}^{j-1}_\triangle + 4\mathsf{F}^{j-1}_\triangle + \Delta^{j-1} & \text{for } \tau(j) \in \{\mathbf{A}, \mathbf{C}, \mathbf{D}(r)\} \\ & \text{(provided } r \geq 2\mathsf{F}^{j-1}_\triangle + \Delta^j) \end{cases}$$

PROOF. For the base case of $j = 0$, consider any adjacent $v, v'$ in $G$. If the edge $(v, v')$ is in $H_0$ (returned by $\mathbf{A}$ at level 0) then $\delta(v, v') = 1 = \mathsf{S}^0_\triangle$. If not then, by the definition of $\mathbf{A}$, $v' \notin \mathcal{B}_0(v)$ and $\delta_{H_0}(v, p_1(v)) = 1 = \mathsf{F}^0_\triangle$.

Define $v, v', z, z', s$, and $s'$ as in our earlier analysis of $\mathbf{ABB}$. That is, $\delta(v, v') = \Delta^j$, $P(v, v')$ is divided into segments with length $\Delta^{j-1}$, and $P(v, v')$ either consists solely of successful segments or contains a prefix of $s$ successful segments, ending at $z$, and a suffix of $s'$ successful segments beginning at $z'$. Figure 3 illustrates the case when $z$ and $z'$ exist.

If the spanner does not contain a short path from $p_j(z)$ to $p_j(z')$ (failure) then we can conclude that $p_j(z') \notin \mathcal{B}_j(p_j(z))$ if $\tau(j) = \mathbf{A}$ or $\mathbf{D}$, that $p_j(z') \notin \mathcal{B}^{1/2}_j(p_j(z))$ if $\tau(j) = \mathbf{B}$, and that $p_j(z') \notin \mathcal{B}^{1/3}_j(p_j(z))$ if $\tau(j) = \mathbf{C}$. It follows that (for $r$ sufficiently large):

$$\delta_H(p_j(z), p_{j+1}(p_j(z))) \leq \begin{cases} 2\mathsf{F}^{j-1}_\triangle + (\Delta - s - s')\Delta^{j-1} & \text{if } \tau(j) = \mathbf{A} \text{ or } \mathbf{D}(\cdot) \\ 2(2\mathsf{F}^{j-1}_\triangle + (\Delta - s - s')\Delta^{j-1}) & \text{if } \tau(j) = \mathbf{B} \\ 3(2\mathsf{F}^{j-1}_\triangle + (\Delta - s - s')\Delta^{j-1}) & \text{if } \tau(j) = \mathbf{C} \end{cases}$$

The distance from $v$ to $p_{j+1}(v)$ is at most $\delta(v, p_{j+1}(p_j(z)))$. We may bound $\delta_H(v, p_{j+1}(v))$ as follows:

$$\begin{aligned} \delta_H(v, p_{j+1}(v)) &\leq \delta(v, z) + \delta(z, p_j(z)) + \delta(p_j(z), p_{j+1}(p_j(z))) \\ &\leq s\Delta^{j-1} + \mathsf{F}^{j-1}_\triangle + t(2\mathsf{F}^{j-1}_\triangle + (\Delta - s - s')\Delta^{j-1}) \\ &\qquad\qquad \{t = 1, 2, 3 \text{ depending on } \tau(j)\} \\ &\leq (s + t(\Delta - s - s'))\Delta^{j-1} + (2t+1)\mathsf{F}^{j-1}_\triangle \\ &\leq (2t+1)\mathsf{F}^{j-1}_\triangle + t\Delta^j \qquad \{\text{worst case is } s = s' = 0\} \end{aligned}$$

We obtain the claimed bounds on $\mathsf{F}^j_\triangle$ by setting $t = 1, 2$, and $3$ when $\tau(j)$ is, respectively, either $\mathbf{A}$ or $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{C}$. This covers the case when the path $v \ldots v'$ is a failure. One way for it to be a success is if each of the $\Delta$ segments is a success, that is, if $z$ and $z'$ do not exist. In general there will be some unsuccessful segments and we can only declare the path successful if there is a *short* path from $p_j(z)$ to $p_j(z')$. We demand a shortest path if $\tau(j) \in \{\mathbf{A}, \mathbf{C}, \mathbf{D}\}$ and tolerate an additive error of 2 if $\tau(j) = \mathbf{B}$. We can now bound $\mathsf{S}^j_\triangle$ as follows:

$$\begin{aligned}\delta_H(v,v') &\le \max\{\Delta\mathsf{S}^{j-1}_\triangle,\ \ \delta_H(v,z) + \delta_H(z,p_j(z)) + \delta_H(p_j(z),p_j(z')) \\ &\qquad\qquad + \delta_H(p_j(z'),z') + \delta_H(z',v')\} \\ &\le \max\{\Delta\mathsf{S}^{j-1}_\triangle,\ \ (s+s')\mathsf{S}^{j-1}_\triangle + 4\mathsf{F}^{j-1}_\triangle + (\Delta - s - s')\Delta^{j-1}\ \ [+2]\} \\ &\le \max\{\Delta\mathsf{S}^{j-1}_\triangle,\ \ (\Delta-1)\mathsf{S}^{j-1}_\triangle + 4\mathsf{F}^{j-1}_\triangle + \Delta^{j-1}\ \ [+2]\}\end{aligned}$$

where the "[+2]" is only present if $\tau(j) = \mathbf{B}$. □

*Remark* 3.4. The bounds in Lemma 3.3 can be improved a bit if $\tau(j) = \mathbf{A}$. We ignored these improvements because they have no effect on our constructions. When $\tau(j) = \mathbf{A}$ one can easily show that $\mathsf{F}^j_\triangle \le 2\mathsf{F}^{j-1}_\triangle + \Delta^j$ and $\mathsf{S}^j_\triangle \le (\Delta-1)\mathsf{S}^{j-1}_\triangle + 2\mathsf{F}^{j-1}_\triangle + \Delta^{j-1}$.

Lemma 3.5 solves these recurrences for spanners that use only schemes $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$. These schemes are generally sufficient to obtain our best sparseness-distortion tradeoffs, so long as the resulting spanner has $\Omega(n \log\log n)$ edges. To obtain sparser spanners we require the use of schemes $\mathbf{D}(\cdot)$ and $\mathbf{x}$.

Lemma 3.5. **(ABC Spanners)** *Let $H$ be a spanner defined by an encoding $\tau \in \mathbf{A}\{\mathbf{A},\mathbf{B},\mathbf{C}\}^o$. If $\Delta \ge 8$ and $c = 3\Delta/(\Delta-7)$ then:*

$$\begin{aligned}\mathsf{F}^j_\triangle &\le c\Delta^j \\ \mathsf{S}^j_\triangle &\le \begin{cases} \Delta^j + 4cj\Delta^{j-1} & \text{for } j \le \Delta \\ (4c+1)\Delta^j & \text{for } j \ge \Delta \end{cases}\end{aligned}$$

*Furthermore, $\mathsf{F}^o_\triangle = 0$, that is, if $\delta(u,v) \le \Delta^o$ then $\delta_H(u,v) \le \mathsf{S}^o_\triangle$.*

Proof. Taking the worst cases from Lemma 3.3 we have $\mathsf{F}^j_\triangle \le 7\mathsf{F}^{j-1}_\triangle + 3\Delta^j$ and $\mathsf{S}^j_\triangle \le \max\{\Delta\mathsf{S}^{j-1}_\triangle,\ (\Delta-1)\mathsf{S}^{j-1}_\triangle + 4\mathsf{F}^{j-1}_\triangle + \Delta^{j-1} + 2\}$. One can easily verify by induction that $\mathsf{F}^j_\triangle \le c\Delta^j$. We now show that $\mathsf{S}^j_\triangle$ is at most $\Delta^j + 4cj\Delta^{j-1} - 1$, and for $j \ge \Delta$, that it is at most $(4c+1)\Delta^j - 1$; these inequalities clearly hold for $j = 1$. First consider the case $j \le \Delta$ and assume the claim holds for $j-1$.

$$\begin{aligned}\mathsf{S}^j_\triangle &\le \max\{\Delta\mathsf{S}^{j-1}_\triangle,\ \ (\Delta-1)\mathsf{S}^{j-1}_\triangle + 4\mathsf{F}^{j-1}_\triangle + \Delta^{j-1} + 2\} \\ &\le \max\{\Delta^j + 4c(j-1)\Delta^{j-1} - \Delta, \\ &\qquad (\Delta-1)(\Delta^{j-1} + 4c(j-1)\Delta^{j-2} - 1) + 4c\Delta^{j-1} + \Delta^{j-1} + 2\} \\ &\le \max\{\Delta^j + 4cj\Delta^{j-1} - 1, \\ &\qquad \Delta^j + 4c(j-1)\Delta^{j-1} + 4c\Delta^{j-1} - \big(4c(j-1)\Delta^{j-2} + \Delta + 1\big)\} \\ &\le \Delta^j + 4cj\Delta^{j-1} - 1\end{aligned}$$

Notice that for $j = \Delta$ this bound is precisely $(4c+1)\Delta^j - 1$, which serves as our base case for the bounds on $\mathsf{S}^j_\triangle$ for $j > \Delta$:

$$\begin{aligned}\mathsf{S}^j_\triangle &\le \max\big\{\Delta\mathsf{S}^{j-1}_\triangle,\ \ (\Delta-1)\mathsf{S}^{j-1}_\triangle + 4\mathsf{F}^{j-1}_\triangle + \Delta^{j-1} + 2\big\} \\ &\le \max\big\{(4c+1)\Delta^j - \Delta,\ \ (4c+1)(\Delta^j - \Delta^{j-1}) + 4c\Delta^{j-1} + \Delta^{j-1} - \Delta + 3\big\} \\ &\le (4c+1)\Delta^j - 1\end{aligned}$$

□

Lemma 3.5 states that in any spanner generated by some string in $\mathbf{A}\cdot\{\mathbf{A},\mathbf{B},\mathbf{C},\}^o$, the distortion is given by the function $f(d)=d+O(od^{1-1/o})$, *provided* that $d$ is at least $8^o$. As we will see later, $o$ can be as large as $\log_{4/3}\log n$ which means that these spanners have weak guarantees for $d<8^{\log_{4/3}\log n}<(\log n)^{7.23}$. However, by using just the $\mathbf{A}$ and $\mathbf{D}$ connection schemes we can approximate polylogarithmic distances much better. Theorem 3.8 shows that in these spanners the multiplicative distortion quickly improves as a function of distance: it goes from logarithmic to log-logarithmic, to constant, and ultimately tending towards 1.

LEMMA 3.6. **(AD Spanners)** *Consider a spanner defined by the encoding* $\tau=\mathbf{AD}(r)\mathbf{D}(r^2)\ldots\mathbf{D}(r^o)$, *where* $r\geq 4$. *Then* $\mathsf{F}_2^j<3^{j+1}$, $\mathsf{F}_3^j=(j+1)3^j$, $\mathsf{S}_2^j<6\cdot 3^j$, *and* $\mathsf{S}_3^j<4j3^j$. *For* $\Delta$ *in the range* $[4,r-2]$ *and* $c'=\frac{\Delta}{\Delta-3}$ *the following bounds hold:*

$$\begin{aligned}\mathsf{F}_\Delta^j &\leq c'\Delta^j\\ \mathsf{S}_\Delta^j &\leq \begin{cases}\Delta^j+4c'j\Delta^{j-1} & \textit{for } j\leq\Delta\\ (4c'+1)\Delta^j & \textit{for } j\geq\Delta\end{cases}\end{aligned}$$

PROOF. We first consider $\Delta=2$. Applying the bound from Lemma 3.3 we have $\mathsf{F}_2^0=1$ and $\mathsf{F}_2^j\leq 3\mathsf{F}_2^{j-1}+2^j$. One can easily check that $\mathsf{F}_2^j=3^{j+1}-2^{j+1}$ is the exact bound. Notice that we can only apply Lemma 3.3 if $r$ is sufficiently large. In particular we require that $2\mathsf{F}_2^{j-1}+2^j\leq r^j$, which holds since $r\geq 4$. Assuming the stated bound on $\mathsf{S}_2^{j-1}$ holds, we have

$$\begin{aligned}\mathsf{S}_2^j &\leq \max\left\{2\mathsf{S}_2^{j-1},\ \mathsf{S}_2^{j-1}+4\mathsf{F}_2^{j-1}+2^{j-1}\right\}\\ &\leq \max\left\{12\cdot 3^{j-1},\ 6\cdot 3^{j-1}+4(3^j-2^j)+2^{j-1}\right\} \qquad \{\text{Ind. ass.: } \mathsf{S}_2^{j-1}\leq 6\cdot 3^{j-1}\}\\ &\leq 6\cdot 3^j\end{aligned}$$

For $\Delta=3$, $\mathsf{F}_3^0=1$ and $\mathsf{F}_3^j\leq 3\mathsf{F}_3^{j-1}+3^j$. One can check that $\mathsf{F}_3^j=(j+1)3^j$ satisfies these recurrences. We assume the stated bound on $\mathsf{S}_3^{j-1}$ and bound $\mathsf{S}_3^j$ as:

$$\begin{aligned}\mathsf{S}_3^j &\leq \max\left\{3\mathsf{S}_3^{j-1},\ 2\mathsf{S}_3^{j-1}+4\mathsf{F}_3^{j-1}+3^{j-1}\right\}\\ &\leq \max\left\{4(j-1)3^j,\ 8(j-1)3^{j-1}+4j3^{j-1}+3^{j-1}\right\}\\ &\leq 4j3^j\end{aligned}$$

We now turn to the general case of $\Delta\geq 4$. Assume inductively that $\mathsf{F}_\Delta^j=c'\Delta^j-(c'-1)3^j$. (At the base case, $\mathsf{F}_\Delta^0=1=c'\Delta^0-(c'-1)3^0$.) Using the recurrence from Lemma 3.3 and the inductive assumption we have $\mathsf{F}_\Delta^j=3(c'\Delta^{j-1}-(c'-1)3^{j-1})+\Delta^j=c'\Delta^j-(c'-1)3^j<c'\Delta^j$. We can bound $\mathsf{S}_\Delta^j$ as $\min\{\Delta^j+4c'j\Delta^{j-1},\ (4c'+1)\Delta^j\}$ using the same proof from Lemma 3.5 simply by substituting $c'$ for $c$. □

Theorem 3.7 illustrates some nice sparseness-distortion tradeoffs for spanners composed of schemes $\mathbf{A},\mathbf{B}$, and $\mathbf{C}$. It only considers those generated by sequences $\mathbf{ABBC}^{o-2}$, which turns out to optimize sparseness without significantly affecting the distortion. (In other words, **ABCCC** would be denser than **ABBCC** and could only improve lower order terms in the distortion.)

THEOREM 3.7. *The spanner generated by* **ABB** *has* $O(n^{6/5})$ *edges and distortion function* $f(d) = d + O(\sqrt{d})$. *The spanner generated by* $\mathbf{ABBC}^{o-2}$ *has* $O(on^{1+\nu})$ *edges, where* $\nu = \left(\frac{3}{4}\right)^{o-2}/(7 - 2\left(\frac{3}{4}\right)^{o-2})$, *and distortion function* $d + O(od^{1-1/o} + 8^o)$. *For all* $o$, $1+\nu$ *is strictly less than* $1 + (3/4)^{o+3}$.

PROOF. Let $H$ be the spanner defined by **ABB**. $H$ has on the order of $n/q_1 + n\sqrt{q_1/q_2} + n\sqrt{nq_2}$ edges, which is $O(n^{6/5})$ for $q_1 = n^{-1/5}$ and $q_2 = n^{-3/5}$. By Lemma 3.5, if $\delta(v, v') \le \Delta^2$ then $\delta_H(v, v') \le \mathsf{S}^2_\Delta = \Delta^2 + O(\Delta)$. (Recall that such a path cannot fail because every pair of vertices in $V_2$ is connected by a nearly shortest path.) In the general case let $H$ be generated by $\mathbf{AB}^2\mathbf{C}^{o-2}$, for some $o \ge 3$. If $v$ and $v'$ are at distance at most $\Delta^o \ge 8^o$ then by Lemma 3.5 $\delta(v, v') \le \min\{\Delta^o + O(o\Delta^{o-1}),\ O(\Delta^o)\}$. Thus, for any distance $d$ (possibly less than $8^o$) the distortion is $f(d) = d + O(od^{1-1/o} + 8^o)$. We now choose the sampling probabilities so as to optimize the size of $H$. They will be selected so that each of the levels zero through $o$ contributes about the same number of edges, say $n^{1+\nu}$. Since the first three levels contribute $n/q_1 + n\sqrt{q_1/q_2} + n\sqrt{q_2/q_3}$ edges (scheme **A** at level 0, **B** at 1 and 2), it follows that $q_1 = n^{-\nu}, q_2 = n^{-3\nu}$, and $q_3 = n^{-5\nu}$. Starting from the other end, level $o$ (scheme **C**) contributes $n + n^{2.5}q_o^2$ implying $q_o = n^{-3/4+\nu/2}$. For $3 \le j < o$, level $j$ contributes on the order of $n + nq_j^2/q_{j+1}^{3/2}$ edges, implying $q_j = q_{j+1}^{3/4}n^{\nu/2}$. Assuming inductively that $q_{j+1} = n^{-\left(\frac{3}{4}\right)^{o-j} + \nu\left(2 - \frac{3}{2}\left(\frac{3}{4}\right)^{o-(j+1)}\right)}$ (which holds for the base case $j + 1 = o$), we have, for $3 \le j < o$:

$$
\begin{aligned}
q_j &= q_{j+1}^{3/4}n^{\nu/2} \\
&= n^{-\left(\frac{3}{4}\right)^{o-j+1} + \frac{3}{4}\nu(2 - \frac{3}{2}\left(\frac{3}{4}\right)^{o-(j+1)}) + \nu/2} \\
&= n^{-\left(\frac{3}{4}\right)^{o-(j-1)} + \nu(2 - \frac{3}{2}\left(\frac{3}{4}\right)^{o-j})}
\end{aligned}
$$

The only sampling probability under two constraints is $q_3$, which means that $\nu$ should be selected to satisfy:

$$n^{-5\nu} = n^{-\left(\frac{3}{4}\right)^{o-2} + \nu(2 - \frac{3}{2}\left(\frac{3}{4}\right)^{o-3})}$$

This equality holds for $\nu = \left(\frac{3}{4}\right)^{o-2}/(7 - \frac{3}{2}\left(\frac{3}{4}\right)^{o-3})$. The size of $H$ is, therefore, on the order of $on^{1+\nu}$. $\square$

Let us briefly compare the size bounds obtained above to the spanners of Thorup and Zwick [2006]. For distortions $d + O(\sqrt{d})$, $d + O(d^{2/3})$, and $d + O(d^{3/4})$ the spanners of Theorem 3.7 have sizes on the order of $n^{6/5}$, $n^{25/22}$, and $n^{103/94}$ in contrast to $n^{4/3}, n^{5/4}$, and $n^{6/5}$ obtained in [Thorup and Zwick 2006]. The separation in density becomes sharper as $o$ increases. For $o = \log_{4/3}\log n - C$ (any constant $C$), the size and distortion of our spanners is $O(n \log\log n)$ and $d + O(od^{1-1/o} + o^o)$, in contrast to [Thorup and Zwick 2006], where the size and distortion are $O(on^{1+1/o})$ and $d + O(d^{1-1/o} + 2^o)$. In this case Theorem 3.7 gives a doubly exponential improvement in density. In some ways Theorem 3.7 is our strongest result. However, when the order $o$ is large (close to $\log_{4/3}\log n$) and the distance being approximated is very short, the spanners of Theorem 3.7 cannot guarantee good distortion. Theorem 3.8 addresses some of these shortcomings.

THEOREM 3.8. *Let $H$ be the spanner encoded by $\mathbf{AD}(r)\mathbf{D}(r^2)\cdots\mathbf{D}(r^o)$, where $o = \log\log_r n$. Then $H$ has $O(r^2 n)$ edges and if $\delta(u,v) = d$ then:*

$$\delta_H(u,v) \le \begin{cases} d \cdot 6(\log n)^{\log(3/2)} & \text{for } d \ge \log n \\ d \cdot 4\log\log n & \text{for } d \ge (\log n)^{\log 3} \\ d \cdot \left(5 + \frac{12}{\Delta - 3}\right) & \text{for } d \ge (\log n)^{\log \Delta}, 4 \le \Delta \le r-2 \\ d + O(\epsilon d) & \text{for } d \ge (o/\epsilon)^o \text{ and } o/\epsilon \le r-2 \end{cases}$$

PROOF. We first show how the sampling probabilities can be selected so that $o = \log\log_r n$ and $|H| = O(r^2 n)$. The subgraph returned by the lowest level connection scheme $\mathbf{A}$ has size roughly $n/q_1$. We want to choose the sampling probabilities so that $\mathbf{D}(r^{j+1})$ contributes half as many as $\mathbf{D}(r^j)$ and $\mathbf{D}(r)$ half as many as $\mathbf{A}$. The number contributed by $\mathbf{D}(r^j)$ is $nr^j q_j^2/q_{j+1}$ which should be on the order of $n/(q_1 2^j)$. We let $q_1 = 1/r^2$ and let $q_j = 2^{h(j)}/r^{g(j)}$. Thus, the size contributed by $\mathbf{D}(r^j)$ is $nr^j 2^{2h(j)-h(j+1)} r^{g(j+1)-2g(j)}$ and should be roughly $nr^2/2^j$. It follows that $h(j+1) = 2h(j) + j$ and $g(j+1) = 2g(j) - j + 2$, where $f(1) = 0$ and $g(1) = 2$. One can verify that these constraints are satisfied for $h(j) = 2^j - (j+1)$ and $g(j) = 2^j + (j-1)$. We can stop at the earliest level $o$ such that $(nq_o)^2 r^o = O(n)$. For $o = \log\log_r n$ we have

$$(nq_o)^2 r^o = n^2 \left(\frac{2^{2^o-(o+1)}}{r^{2^o+(o-1)}}\right)^2 r^o < n^2 \left(\frac{2}{r}\right)^{2^{o+1}} = n^2 \left(\frac{2}{r}\right)^{2\log n/\log r} < n$$

When two vertices are at distance $d = 2^o < \log n$, Lemma 3.6 shows that in the spanner they are at distance at most $6 \cdot 3^o < 6(\log n)^{\log 3}$: thus, a multiplicative distortion of $6(\log n)^{\log(3/2)}$. For $d = 3^o < (\log n)^{\log 3}$, Lemma 3.6 says the distance in the spanner is at most $4(o+1)3^o < d \cdot 4\log\log n$. The other cases are treated in the same fashion, by appealing to the bounds on $\mathsf{S}^o_\Delta$ proved in Lemma 3.6, for $\Delta = 4, 5, \ldots$. □

Theorem 3.8 says that there is a linear size spanner whose multiplicative distortion can be driven arbitrarily close to 5 at the cost of an additive polylog($n$) term. The exponent in our polylog($n$) term is likely to be improvable though this additive term can not be eliminated entirely. If we want Theorem 3.8 to produce a $(1+\epsilon,\beta)$-spanner then $r$ must be at least $\epsilon^{-1}\log\log n$ and the size at least $n(\epsilon^{-1}\log\log n)^2$. The dependence on $\epsilon$ here is already a significant improvement over the comparable spanners of Elkin and Peleg [2004], which always have size $\Omega(n(\epsilon^{-1}\log\log n)^{\log\log n})$. However, Theorem 3.9 shows that our space bound can be improved doubly-exponentially.

THEOREM 3.9. *For a constant $c''$ let $r = c''\epsilon^{-1}\log\log n + 2$, $\gamma = 2\log_{4/3}\log r$ and $o = \log\log_r n$. The spanner encoded by $\mathbf{AC}^\gamma\mathbf{D}(r^{\gamma+1})\mathbf{D}(r^{\gamma+2})\cdots\mathbf{D}(r^o)$ is a $(1+\epsilon,\beta)$-spanner with size $O(n\log\log(\epsilon^{-1}\log\log n))$, where $\beta = r^o$.*

PROOF. We first analyze the distortion of the spanner. Let $u$ and $v$ be two vertices at distance $d = (r-2)^o$ in the original graph. Using the same analysis from Lemma 3.3 and Theorem 3.7 it follows that the distance in the spanner $H$ is:

$$
\begin{aligned}
\delta_H(u,v) &\le d + 4c(o+1)d^{1-1/o} && \text{From Thm. 3.7, } c = 3(r-2)/(r-9)\\
&\le d + \frac{12(r-2)}{r-9}\frac{(o+1)d}{r-2}\\
&\le d(1 + \frac{12(\log\log_r n + 1)}{c''\epsilon^{-1}\log\log n - 7}\\
&\le d(1+\epsilon)
\end{aligned}
$$

If two vertices are at a distance $d' > d$ we can simply chop up the shortest path into segments of length at most $d$ and consider each separately. It follows that the distance in the spanner is $(1+\epsilon)d' + \beta$.

We choose the sampling probabilities so that $\mathbf{A}$ contributes $O(\gamma n)$ edges (asymptotically the size of the spanner), $\mathbf{C}^\gamma$ contributes $O(\gamma n)$ edges ($O(n)$ per $\mathbf{C}$), and $\mathbf{D}(r^{\gamma+1})\cdots\mathbf{D}(r^o)$ contributes $o(n)$ in total. It follows that $q_1 = 1/\gamma$ and $nq_j^2/q_{j+1}^{3/2} = n$, for $1 \le i \le \gamma$. These are satisfied for $q_j = \gamma^{-\left(\frac{4}{3}\right)^{j-1}}$. For $j \ge \gamma+1$ the number of edges contributed is $nr^j q_j^2/q_{j+1}$, which should be on the order of $n/2^j$. Let $Q = q_{\gamma+1} = \gamma^{-\left(\frac{4}{3}\right)^{\gamma}}$. For $j > \gamma$ we write $q_j$ as $Q^{h(j)}(2r)^{g(j)}$ and select $h, g$ such that $\mathbf{D}(r^j)$ contributes around $n/2^j$ edges. That is, $nr^j q_j^2/q_{j+1} = nr^j Q^{2h(j)-h(j+1)}(2r)^{2g(j)-g(j+1)} = n/2^j$. It follows that $h, g$ obey the equalities $h(j+1) = 2h(j)$, $g(j+1) = 2g(j) + j$, with $h(\gamma+1) = 1$, and $g(\gamma+1) = 0$. One can verify that $h(j) = 2^{j-(\gamma+1)}$ and $g(j) = (\gamma+2)2^{j-(\gamma+1)} - (j+1)$ satisfy these constraints. What remains is to show that the number of edges contributed by $\mathbf{D}(r^o)$ (at most $(q_o n)^2 r^o$) is negligible.

$$
\begin{aligned}
q_o &= Q^{2^{o-(\gamma+1)}}(2r)^{(\gamma+2)2^{o-(\gamma+1)}-(o+1)}\\
&\le \left(\left(\gamma^{-(4/3)^\gamma}\right)(2r)^{\gamma+2}\right)^{2^{o-(\gamma+1)}}\\
&\le \left(r^{-\log r \log\gamma}(2r)^{\gamma+2}\right)^{2^{o-(\gamma+1)}} && \{\gamma = \left\lceil 2\log_{4/3}\log r\right\rceil\}\\
&\le r^{-\log r(\log\gamma - 1)2^{o-(\gamma+1)}} && \{(2r)^{\gamma+2} < r^{\log r}\}\\
&\le r^{-\log r(\log\gamma-1)\log_r n 2^{-\gamma-1}} && \{o = \lceil\log\log_r n\rceil\}\\
&< 1/n
\end{aligned}
$$

□

Previous to our work the only linear sized spanners for general graphs had $O(\log n)$ multiplicative distortion [Althöfer et al. 1993; Halperin and Zwick 1996]. Theorem 3.8 shows that a multiplicative distortion tending toward 5 can be achieved within this size bound. In the following theorem we show that there are linear size spanners whose additive distortion is $\tilde{O}(n^{9/16})$.

THEOREM 3.10. *Every graph contains an additive $\tilde{O}(n^{9/16})$-spanner with $O(n)$ edges.*

PROOF. Consider the spanner described by the sequence **xCC**. Notice that the size of the subgraph returned by **x** is $n$ regardless of the sampling probability $q_1$. In other words, the size and distortion of the spanner is not uniquely determined by the short encoding **xCC**. The spanner $H$ includes the shortest paths between all pairs in $V_2 \times V_2$, a shortest path from each $v \in V_1$ to each $u \in \bar{\mathcal{B}}_1^{1/3}(v) \cap V_1$, and for every $u$, the path $P(u, p_1(u))$. By Theorem 4.4 the expected size of this spanner is on the order of $n + nq_1^2/q_2^{3/2} + n^{2.5}q_2^2$, which is balanced when $q_1 = n^{3/4} \cdot q_2^{7/4}$. We require that for any two vertices $u$ and $v$ at distance $D = 2q_1^{-1}\log n$, some vertex of $V_1$ lies on $P(u,v)$. (This property holds w.h.p. and can easily be checked in polynomial time.) The desired additive distortion will be on the order of $D$ so we can restrict our attention to shortest paths between vertices $u, v \in V_1$. Let $u' \in P(u,v)$ be the closest vertex to $u$ in $V_1$. We have that $\delta(u,u') \leq D$ and if $u' \in \mathcal{B}_1^{1/3}(u)$ then $\delta_H(u,u') = \delta(u,u')$. If this is the case we take the shortest path from $u$ to $u'$ and then consider the shortest path from $u'$ to $v$. Suppose that $u' \notin \mathcal{B}_1^{1/3}(u)$ and, symmetrically, that $v' \notin \mathcal{B}_1^{1/3}(v)$, where $v' \in P(u,v)$ is the closest vertex to $v$ in $V_1$. Then $\delta_H(u, p_2(u)) \leq 3D$ and $\delta_H(v, p_2(v)) \leq 3D$. Since all pairs in $V_2 \times V_2$ are connected by shortest paths in $H$, $\delta_H(p_2(u), p_2(v)) \leq \delta(u,v) + 6D$. Thus, for *any* two vertices $u, v \in V$, $\delta_H(u,v) \leq \delta(u,v) + O(D)$. If our desired spanner size is $O(n^{1+\epsilon})$ we would set $q_2 = n^{-3/4+\epsilon/2}$ and $q_1 = n^{-9/16+7\epsilon/8}$. The additive distortion would be $O(n^{9/16-7\epsilon/8}\log n)$. Setting $\epsilon = 0$ gives the theorem. □

All of the spanners presented so far have an inherent tradeoff between sparseness and distortion. Theorem 3.11 shows that for a large class of graphs *both* the multiplicative distortion and density of the spanner can be driven arbitrarily close to 1. Theorem 3.11 applies to graphs with quadratic expansion, meaning the number of vertices within distance $D$ of any vertex is at least $D^2$.

THEOREM 3.11. *Every graph with quadratic expansion contains a* $(1+\epsilon, \beta)$*-spanner with* $(1+\epsilon)n$ *edges, for any* $\epsilon > 0$ *and* $\beta = O(\epsilon^{-1}\log\log n)^{\log\log n}$.

PROOF. Let $\Delta = 4\epsilon^{-1}\log\log n$. The spanner is generated by the sequence $\mathbf{xD}((\Delta+2)^5)\mathbf{D}((\Delta+2)^6)\cdots\mathbf{D}((\Delta+2)^{\log\log n})$, where $o = \log\log n - 4$ is the order. Unlike all previous spanners we select the set $V_1$ deterministically such that $|V_1| = O(n/\Delta^8)$ and $\delta(v, p_1(v)) \leq \Delta^4$ for every vertex $v$. Thereafter $V_2, \ldots, V_{\log\log n - 4}$ are selected by random sampling. To obtain $V_1$ we select a maximal set of vertices such that the distance between any two is at least $\Delta^4$. Since the $(\Delta^4/2)$-neighborhood of each is disjoint and contains at least $(\Delta^4/2)^2$ vertices, it follows that $|V_1| \leq 4n/\Delta^8$. By the maximality of this set it also follows that $\delta(v, p_1(v)) < \Delta^4$ for all $v$. We choose the sampling probabilities such that each level contributes roughly $n/(\Delta+2)$ edges. Thus, in total the size of the spanner is $n + on/(\Delta+2) < n(1+\epsilon)$. We let $q_j$ be of the form $(\Delta+2)^{-g(j)}$ (and assume for simplicity that $q_1$ is also in this form despite the fact that $V_1$ was generated deterministically.) The number of edges contributed by the $j$th level is $nq_j^2(\Delta+2)^{j+4}q_2^{-1} = n(\Delta+2)^{g(j+1)-2g(j)+j+4}$. Thus $g(j+1) = 2g(j) - (j+5)$ and $g(1) = 8$. One can easily verify that $g(j) = 2^{j-1} + j + 6$. Thus, the number of edges contributed at the highest level is $|V_o|^2(\Delta+2)^{o+4} = o(n)$ We'll analyze the distortion using the standard framework, except that all distances will be rescaled to be in units of $U = \Delta^4$. Thus $\mathsf{F}_\Delta^j$ and $\mathsf{S}_\Delta^j$ are w.r.t. vertices at distance $U\Delta^j$ rather than $\Delta^j$. Since our choice of $V_1$ ensures that $\delta(v, p_1(v)) \leq \Delta^4 = U$

for all $v$, we let $\mathsf{F}^o_\Delta = \mathsf{S}^o_\Delta = U$. (We could set $\mathsf{S}^o_\Delta$ to be anything here since all paths of length $U$ are guaranteed to fail.) Using the exact same analysis from Lemma 3.6 (but in units of $U$) it follows that $\mathsf{F}^j_\Delta \le c'U\Delta^j$ and $\mathsf{S}^j_\Delta \le U\Delta^j + 4c'jU\Delta^{j-1}$, where $c' = \Delta/(\Delta-3)$. If the distance between two vertices is at most $d = U\Delta^o$, their distance in the spanner is at most:

$$\begin{aligned}
\mathsf{S}^o_\Delta &= U\Delta^o(1 + 4c'o/\Delta) \\
&= d(1 + 4\frac{\Delta}{\Delta-3}\frac{\log\log n - 4}{\Delta}) \\
&= d(1 + \frac{4(\log\log n - 4)}{4\epsilon^{-1}\log\log n - 3}) \\
&\le d(1+\epsilon)
\end{aligned}$$

Again, long shortest paths should be analyzed by chopping them up into pieces of length at most $U\Delta^o$, which are then analyzed separately. It follows that this is a $(1+\epsilon, \beta)$-spanner with $\beta = O(\epsilon^{-1}\log\log n)^{\log\log n}$ $\square$

## 4. THE CONNECTION SCHEMES

In this section we make use of the assumption that shortest paths are closed under taking subpaths, i.e., if $x, y \in P(u,v)$ then $P(x,y) \subseteq P(u,v)$.

### 4.1 The Trivial Schemes $\mathbf{A}, \mathbf{D}$, and $\mathbf{x}$

Connection schemes $\mathbf{A}$ and $\mathbf{D}$ are trivial but surprisingly powerful. The subgraph returned by $\mathbf{A}$ at level $j$ is, by definition, $\bigcup_{v\in V_j, u\in\bar{\mathcal{B}}_j(v)\}} P(v,u)$, that is, a breadth first search tree from every $v \in V_j$ containing $p_{j+1}(v)$ and all vertices $u$ closer to $v$ than $p_{j+1}(v)$. The expected size of this subgraph is at most $\sum_{v\in V}\Pr[v \in V_j]\cdot\mathbb{E}[|\bar{\mathcal{B}}_j(v)| - 1] \le nq_j/q_{j+1}$. Scheme $\mathbf{A}$ was introduced by Thorup and Zwick [2006].

Scheme $\mathbf{D}(r)$ returns the subgraph $\bigcup_{v,u\in V_j\,:\,u\in\bar{\mathcal{B}}_j(v)\cap\mathrm{Ball}(v,r)} P(u,v)$. Notice that $\mathbf{D}$ only connects the pair $v, p_{j+1}(v)$ if they are at distance at most $r$. To bound the size of the subgraph returned by $\mathbf{D}$ we pessimistically assume that each path contributes exactly $r$ edges. The expected size is then $r\cdot\sum_{v\in V}\Pr[v\in V_j]\cdot\mathbb{E}[|\bar{\mathcal{B}}_j(v)\cap V_j| - 1] \le nrq_j^2/q_{j+1}$.

Scheme $\mathbf{x}$ returns the subgraph $\bigcup_{v\in V_j} P(v, p_{j+1}(v))$, which is a collection of disjoint trees, regardless of $j$ or the sampling probabilities $q_j, q_{j+1}$. Thus, the subgraph returned has fewer than $n$ edges.

### 4.2 Scheme $\mathbf{B}$

Our objective is to find a subgraph $H$ such that for every $v \in V_j$ and every $u \in \mathcal{B}^-_j(v)$:

$$\delta_H(v,u) \le \delta(v,u) + 2(\log\delta(v,u) + 1) \tag{1}$$

Furthermore, if $u$ is in $V_j$ as well as $\mathcal{B}^{1/2}_j(v)$ (or if $j = o$ and $u \in V_o$) then:

$$\delta_H(v,u) \le \delta(v,u) + 2 \tag{2}$$

Our algorithm is based on a generalized and iterative version of the path-buying algorithm introduced by Baswana et al. [2009]. In the $i$th iteration we receive a subgraph $H^{(i-1)}$ that distorts the distance (additively) from $v \in V_j$ to some subset of $\mathcal{B}_j^-(v)$ by at most $2(i-1)$. The output of the iteration is $H^{(i)} \supseteq H^{(i-1)}$, which connects $v$ to a larger subset of $\mathcal{B}_j^-(v)$, at the cost of a larger distortion. We prove that $H^{(\log n)}$ has the distortion claimed in Eqn. (1). We show Eqn. (2) is guaranteed to hold even in $H^{(1)}$.

Each iteration is an instantiation of the generic *path-buying* algorithm, which was introduced by Baswana et al. [2009] to obtain an additive 6-spanner. In iteration $i$ we consider each $v \in V_j$ and $u \in B_j^-(v)$ and a certain path $P^{(i)}(v,u)$ that may be longer than the shortest path $P(v,u)$ by $2(i-1)$ edges. (Note that $\left|P^{(1)}(v,u)\right| = |P(v,u)|$.) Based on certain evolving *cost* and *value* functions we either ignore $P^{(i)}(v,u)$ or *purchase* it, including all its edges in $H^{(i)}$. The pseudocode in Figure 4 will be meaningful only after we specify all the parameters of the algorithm. We need to choose cost and value functions, as well as the initial subgraph $H^{(0)}$ passed to the first iteration. We also have to explain how the paths $P^{(i)}(v,u)$ are chosen. Let us start with $H^{(0)}$. We randomly select each vertex to be a *center* with probability $q'$. Every vertex that is adjacent to a center is *covered* and if $v$ is covered, $c(v)$ refers to an arbitrary adjacent center. The initial subgraph $H^{(0)}$ consists of all edges that are incident to at least one uncovered vertex and all edges of the form $(v, c(v))$ connecting vertices to their centers. It is easy to show that $\mathbb{E}[\left|H^{(0)}\right|] \le n/q'$ with an analysis similar to that of scheme **A**.

If $P$ is a path let $C(P) = \{c(v) \mid v \in P \text{ is covered}\}$. The cost and value of a path change depending on a specified subgraph $H$ and the iteration $i$. In any iteration $\mathrm{cost}_H(P) = |P \backslash H|$. Let $P$ be a path (not necessarily shortest) from $v$ to $u$. The value of $P$ in the $i$th iteration is defined to be:

$$\mathrm{value}_{H,i}(P) = \left| \left\{ c \in C(P) \;\middle|\; \begin{array}{l} \delta_{H\cup P}(v,c) < \delta_H(v,c) \\ \quad \text{and} \\ \delta_H(v,c) > \delta(v,c) + 2(i-1) \end{array} \right\} \right|$$

That is, the cost of $P$ is the number of edges that we need to include in $H$ so that it contains $P$. The value of $P$ is the number of $c \in C(P)$ such that $H \cup P$ is more accurate than $H$ in approximating the distance from $v$ to $c$, provided that $H$ is not already sufficiently accurate, i.e., if $\delta_H(v,c) \le \delta(v,c) + 2(i-1)$.

In iteration $i$ we choose $P^{(i)}(v,u)$ to be the concatenation of $P_{H^{(i-1)}}(v, w_i)$ with $P(w_i, u)$, where $w_i$ is the farthest vertex from $v$ on $P(v,u)$ such that $\delta_{H^{(i-1)}}(v, w_i) \le \delta(v, w_i) + 2(i-1)$. Call $w_i$ the *intermediary* between $v$ and $u$.

Iteration $i$: $\quad (1 \le i < \log n)$
1. $H \leftarrow H^{(i-1)}$
2. For each $v \in V_j$ and $u \in \mathcal{B}_j^-(v)$
3. Let $w_i$ be the last vertex in $P(v,u)$ s.t. $\delta_{H^{(i-1)}}(v, w_i) \le \delta(v, w_i) + 2(i-1)$.
4. Let $P = P^{(i)}(v,u) = P_{H^{(i-1)}}(v, w_i) \cdot P(w_i, u)$
5. If $4 \cdot \mathrm{value}_{H,i}(P) \ge \mathrm{cost}_H(P)$
6. $H \leftarrow H \cup P$ $\quad$ {*$P$ is purchased*}
7. $H^{(i)} \leftarrow H$

Fig. 4. One iteration of scheme **B**'s path buying algorithm.

Lemma 4.1 basically says that the intermediary vertices $w_1, w_2, w_3, \ldots$ get geometrically closer to $u$ in each iteration of the algorithm of Figure 4.

LEMMA 4.1. *Let $w_i$ be the intermediary between $v \in V_j$ and $u \in \mathcal{B}_j^-(v)$ at iteration $i$. Then $|C(P(w_i, u))| \leq \lceil |C(P(v, u))| / 2^{i-1} \rceil$.*

PROOF. We assume that if $x, y$, and $z$ are consecutive vertices on $P(v, u)$ with $c(x) = c(z)$ then $y = c(x)$. Choosing $P(v, u)$ in this way helps to reduce the costs of paths since $(x, y)$ and $(y, z)$ have already been included in $H^{(0)}$. Recall that in iteration $i$ we choose $P^{(i)}(v, u)$ to be the concatenation of $P_{H^{(i-1)}}(v, w_i)$ with $P(w_i, u)$.

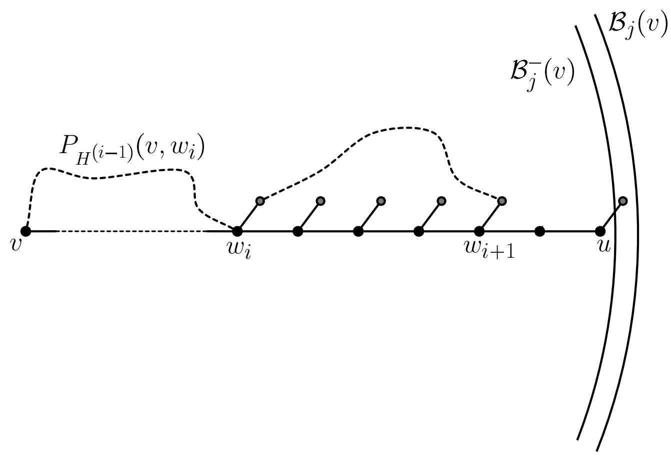


Fig. 5. The shaded vertices represent the centers of vertices in $P(v, u)$.

The proof is by induction on the number of iterations. Since $w_1$ is on $P(v, u)$ we have $|C(P(w_1, u))| \leq |C(P(v, u))|$, so the claim holds initially. Now consider when the path $P = P^{(i)}(v, u) = P_{H^{(i-1)}}(v, w_i) \cdot P(w_i, u)$ is examined in iteration $i$. See Figure 5. There are two cases to consider, depending on whether $P$ is purchased or not. If the algorithm purchases $P$ then $\delta_{H^{(i)}}(v, u) \leq \delta_{H^{(i-1)}}(v, w_i) + \delta(w_i, u) \leq \delta(v, u) + 2(i - 1)$. Therefore, $w_{i'} = u$ for all $i' > i$ and $|C(P(w_{i'}, u))| \leq 1$. If the algorithm refuses to purchase $P$ then $4 \cdot \text{value}_{H,i}(P) < \text{cost}_H(P)$. Since all edges in $P_{H^{(i-1)}}(v, w_i)$ are necessarily in $H^{(i-1)} \subseteq H$, $\text{cost}_H(P) = \text{cost}_H(P(w_i, u))$, which is at most $2\,|C(P(w_i, u))| - 1$. Combining inequalities we have $\text{value}_{H,i}(P) \leq (2\,|C(P(w_i, u))| - 2)/4$. In other words, for strictly more than half of the $c \in C(P(w_i, u))$, either $\delta_H(v, c) \leq \delta(v, c) + 2(i-1)$ or $\delta_{H \cup P}(v, c) = \delta_H(v, c)$. These two cases are actually the same since $\delta_{H \cup P}(v, c) \leq \delta_{H^{(i-1)}}(v, w_i) + 2(i - 1) + \delta(w_i, c) = \delta(v, c) + 2(i - 1)$. Let $w_{i+1}$ be the last vertex in $P(v, u)$ for which $c = c(w_{i+1})$ satisfies the above inequality. It follows that $\delta_{H^{(i)}}(v, w_{i+1}) \leq \delta(v, c) + 2(i-1) + 1 \leq$

$\delta(v, w_{i+1}) + 2i$ and that $|C(P(w_{i+1}, u))| \leq \lceil |C(P(w_i, u))| /2 \rceil$. From the induction hypothesis it follows that $|C(P(w_{i+1}, u))| \leq \lceil |C(P(v, u))| /2^i \rceil$. □

THEOREM 4.2. *Let $H = H^{(\log n)}$ be the subgraph produced by $\log n$ iterations of the path-buying algorithm. For $v \in V_j, u \in \mathcal{B}_j^-(v)$, and $d = \delta(v, u)$, it holds that $\delta_H(v, u) \leq d + 2(\log d + 1)$. If $u \in V_j \cap \mathcal{B}_j^{1/2}(v)$ or $u \in V_j$ and $j = o$, then $\delta_H(v, u) \leq d + 2$. When the center vertices are chosen with suitable probability the expected size of $H$ is $O(n\sqrt{q_j/q_{j+1}})$.*

PROOF. *Distortion Bounds.* One consequence of Lemma 4.1 is that if $d = \delta(v, u)$ then $w_{\log d+1}$ is either $u$ or the last covered vertex on $P(v, u)$, implying that $\delta_{H^{(\log d+1)}}(v, u) \leq \delta(v, u) + 2(\log d + 1)$. If $u \in V_j \cap \mathcal{B}_j^{1/2}(v)$ we can prove better bounds. Notice that $\text{rad}_{j+1}(u) > \text{rad}_{j+1}(v)/2$, which implies that $v \in \mathcal{B}_j(u)$. In the first iteration of the path-buying algorithm the paths $P^{(1)}(v, u) = P(v, u)$ and $P^{(1)}(u, v) = P(u, v)$ were considered separately. If both were not purchased then, by the definition of the value function, $\delta_{H^{(1)}}(v, c) \leq \delta(v, c)$ for strictly more than half the $c \in C(P(v, u))$. Similarly $\delta_{H^{(1)}}(u, c) \leq \delta(u, c)$ also holds for strictly more than half the $c$, meaning there exists some $c^* \in C(P(v, u))$ such that $\delta_{H^{(1)}}(v, u) \leq \delta_{H^{(1)}}(v, c^*) + \delta_{H^{(1)}}(c^*, u) \leq \delta(v, u) + 2$. If $j = o$ then $\text{rad}_{o+1}(v) = \infty$ for all vertices $v$. In this case both $P(v, u)$ and $P(u, v)$ would be considered in the first iteration of the path buying algorithm and the analysis above applies.

*Sparseness Bounds.* Let $P_1, P_2, \ldots$ be the sequence of paths purchased by the algorithm and let $\text{cost}(P_\ell)$ and $\text{value}(P_\ell)$ be w.r.t. the time of purchase. By the definition of the cost function the total size of $H^{(\log n)}$ is $\left|H^{(0)}\right| + \sum_\ell \text{cost}(P_\ell)$, which is at most $\left|H^{(0)}\right| + 4 \sum_\ell \text{value}(P_\ell)$ because of our criterion for purchasing paths. The sum $\sum_\ell \text{value}(P_\ell)$ counts pairs $(v, c)$ where $v \in V_j$ and for some $u \in \mathcal{B}_j^-(v)$, $c(u) = c$. This implies that $c \in \mathcal{B}_j(v)$. The expected number of such pairs $(v, c)$ is therefore at most $nq'q_j/q_{j+1}$ since $\mathbb{E}[|V_j|] = nq_j$ and for any $v$, $\mathbb{E}[|\mathcal{B}_j(v)|] = q_{j+1}^{-1}$. Recall that $q'$ is the probability for sampling centers. If we could show that each pair is counted in $\sum_\ell \text{value}(P_\ell)$ at most a constant number of times then the expected size of $H^{(\log n)}$ would be on the order of $n/q' + nq'q_j/q_{j+1}$. Setting $q' = \sqrt{q_{j+1}/q_j}$ gives a bound of $O(n\sqrt{q_j/q_{j+1}})$. Consider the *first* time the pair $(v, c)$ was counted. That is, we find the first purchased path $P_\ell$ where $c \in C(P)$, the first vertex of $P$ is $v \in V_j$, and $\delta_{H \cup P_\ell}(v, c) < \delta_H(v, c)$. If $P_\ell$ was purchased in iteration $i$ then $\delta_{H \cup P_\ell}(v, c) \leq \delta(v, c) + 2i$. Thus, after $P_\ell$ is purchased $(v, c)$ can be counted at most two more times. Every time $(v, c)$ is counted the distance between $v$ and $c$ is reduced, and after the distance is at most $\delta(v, c) + 2(j - 1)$, $(v, c)$ will never be counted again. □

### 4.3 Scheme **C**

To analyze the size of the subgraph returned by **C** we appeal to a lemma of Coppersmith of Elkin [2006]. Let $\mathcal{Q} \subseteq \mathcal{P}$ be a set of shortest paths. We say that $v$ is a *branching point* for two shortest paths $P, P' \in \mathcal{Q}$ if $P$ and $P'$ intersect and $v$ is an endpoint on the path $P \cap P'$. Notice that if $P$ and $P'$ have just one vertex in common it would be the unique endpoint on the edgeless path $P \cap P'$; see Figure 6. Let $\text{br}(v)$ be the number of pairs $P, P' \in \mathcal{Q}$ such that $v$ is a branching point for $P, P'$, and let $\text{br}(\mathcal{Q}) = \sum_{v \in V} \text{br}(v)$.

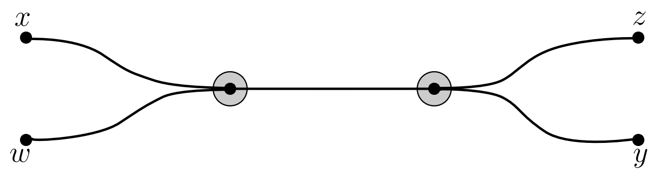


Fig. 6. Branching points on the intersecting paths $P(x,y)$ and $P(w,z)$.

THEOREM 4.3. *(Coppersmith and Elkin) Let $\mathcal{Q}$ be a set of shortest paths and $G(\mathcal{Q}) = \bigcup_{P\in\mathcal{Q}} P$. Then $|G(\mathcal{Q})| \le n + O(\sqrt{n\mathrm{br}(\mathcal{Q})})$.*

PROOF. Let $\deg(v)$ be the degree of $v$ in $G(\mathcal{Q})$. Notice that $\mathrm{br}(v) \ge \binom{\lceil \deg(v)/2\rceil}{2}$. There must be at least $\lceil \deg(v)/2\rceil$ paths in $\mathcal{Q}$ that intersect $v$, no two of which use the same edges incident to $v$. Each pair of these paths contributes to $\mathrm{br}(v)$. We bound the size of $G(\mathcal{Q})$ as:

$$\begin{aligned} |G(\mathcal{Q})| &= \frac{1}{2}\sum_v \deg(v) \\ &= n + \sum_{v\,:\,\deg(v)\ge 3} O(\sqrt{\mathrm{br}(v)}) && \{\text{from the above observation}\} \\ &= n + O(\sqrt{n\mathrm{br}(\mathcal{Q})}) && \{\text{from the concavity of sqrt}\} \end{aligned}$$

□

THEOREM 4.4. *Let $\mathcal{Q} = \{P(v,u) \,:\, v \in V_j,\, u \in \bar{\mathcal{B}}_j^{1/3}(v)\}$. Then $\mathbb{E}[|G(\mathcal{Q})|] = n + O(nq_j^2/q_{j+1}^{3/2})$. If $\mathcal{Q}' = \{P(v,u) \,:\, (v,u) \in V_o \times V_o\}$ then $\mathbb{E}[|G(\mathcal{Q}')|] = n + O(n^{2.5}q_o^2)$.*

PROOF. Let $v, w, v', w' \in V_j$, where $v' \in \mathcal{B}_j^{1/3}(v)$ and $w' \in \mathcal{B}_j^{1/3}(w)$. We first argue that if $P(v,v')$ and $P(w,w')$ intersect then $w, w' \in \mathcal{B}_j(v)$. The contrary scenario is depicted in Figure 7. For any vertex $w$, $\mathrm{rad}_{j+1}(w) \le \delta(w,v) + \mathrm{rad}_{j+1}(v)$. Thus, If $w$ lies outside $\mathcal{B}_j(v)$ then $\mathcal{B}_j^{1/3}(w) \cap \mathcal{B}_j^{1/3}(v)$ must be empty.

Let $v_a$ be the $a$th farthest vertex from $v = v_1$, breaking ties arbitrarily.

$$\begin{aligned} \mathbb{E}[\mathrm{br}(\mathcal{Q})] \;&\le\; 2 \sum_{v\in V, 1<a<b<c} \Pr[\{v, v_a, v_b, v_c\} \subseteq V_j \wedge \{v_a, v_b, v_c \in \mathcal{B}_j(v)] \\ &=\; 2 \sum_{v\in V, c\ge 4} \Pr[|\mathcal{B}_j(v)| \ge c] \cdot \binom{c-2}{2} \cdot q_j^4 \\ &=\; 2 \sum_{v\in V, c\ge 4} (1-q_{j+1})^c \cdot \binom{c-2}{2} \cdot q_j^4 \\ &=\; O(nq_j^4/q_{j+1}^3) \end{aligned}$$

The second line follows since $v_c \in \mathcal{B}_j(v)$ if and only if $|\mathcal{B}_j(v)| \ge c$, and once $v_c$ and $v = v_1$ are chosen there are $\binom{c-2}{2}$ ways to choose $v_a$ and $v_b$. The last line follows since $(1-q_{j+1})^c$ is bounded by a constant for $c < 1/q_{j+1}$ and geometrically

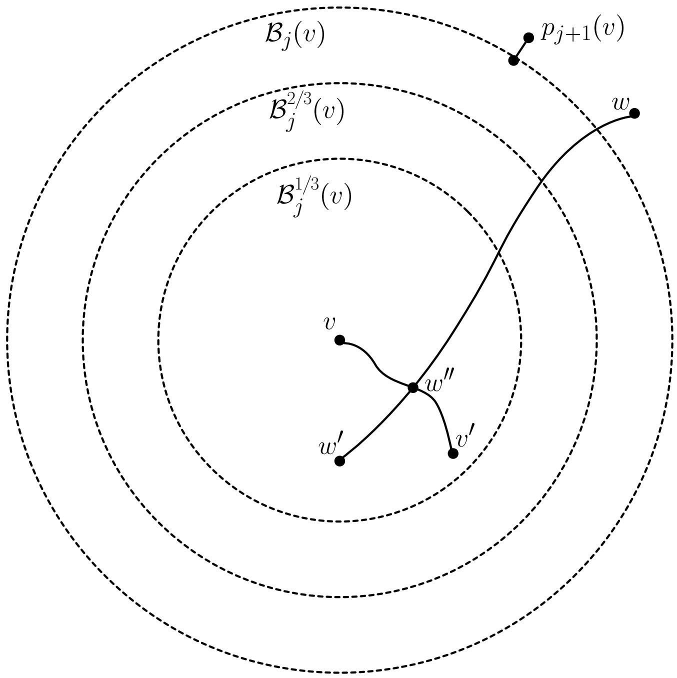


Fig. 7. An impossible situation depicted: $P(v, v')$ and $P(w, w')$ intersect, $v' \in \mathcal{B}_j^{1/3}(v)$, $w' \in \mathcal{B}_j^{1/3}(w)$, and $w \notin \mathcal{B}_j(v)$.

decaying thereafter. Thus, $\mathbb{E}[|G(\mathcal{Q})|] = n + O(\sqrt{n\text{br}(\mathcal{Q})}) = n + O(nq_j^2/q_{j+1}^{3/2})$. Similarly, $\text{br}(\mathcal{Q}')$ is sharply concentrated around its mean—at most $(q_{j+1}n)^4$—and $\mathbb{E}[|G(\mathcal{Q}')|] = n + O(\mathbb{E}[\sqrt{n\text{br}(\mathcal{Q}')}]) = n + O(n^{2.5}q_o^2)$. $\square$

## 5. CONCLUSION

In this paper we have shown that nearly all the recent work on additive and low-distortion spanners can be seen as merely instantiating a generic algorithm based on modular connection schemes. The contribution of this work is not only a simpler way to look at spanners. On purely quantitative terms our constructions provide substantially better distortion than [Elkin and Peleg 2004; Thorup and Zwick 2006] at any desired level of sparsity. Our constructions can also produce *linear* sized spanners, a feature that is conspicuously absent from recent spanner constructions.

One clear avenue for further research is to expand our repertoire of connection schemes. The schemes we use either have no distortion or constant additive distortion, which is actually much stronger than we need to guarantee the good overall distortion of the spanner. A natural idea is to apply the ideas presented here in a recursive manner, by composing connection schemes with strong guarantees to form sparser connection schemes with weaker guarantees.

Although the specific tradeoffs of our results could certainly be improved, the framework of this paper seems inherently incapable of generating purely additive spanners at any desired level of sparseness.[5] It is unclear whether a fundamentally new technique is required to find sparse additive spanners or whether the generic path-buying algorithm of Baswana et al. [2009] could be generalized for this purpose. In any case, proving or disproving the existence of additive spanners remains the chief open problem in this area.

## REFERENCES


Aingworth, D., Chekuri, C., Indyk, P., and Motwani, R. 1999. Fast estimation of diameter and shortest paths (without matrix multiplication). *SIAM J. Comput. 28,* 4, 1167–1181.

Althöfer, I., Das, G., Dobkin, D., Joseph, D., and Soares, J. 1993. On sparse spanners of weighted graphs. *Discrete and Computational Geometry 9*, 81–100.

Awerbuch, B. 1985. Complexity of network synchronization. *J. ACM 32,* 4, 804–823.

Awerbuch, B., Bar-Noy, A., Linial, N., and Peleg, D. 1990. Improved routing strategies with succinct tables. *J. Algorithms 11,* 3, 307–341.

Awerbuch, B. and Peleg, D. 1992. Routing with polynomial communication-space trade-off. *SIAM J. Discrete Math. 5,* 2, 151–162.

Baswana, S. and Kavitha, T. 2006. Faster algorithms for approximate distance oracles and all-pairs small stretch paths. In *Proc. 47th IEEE Symposium on Foundations of Computer Science (FOCS).* 591–602.

Baswana, S., Kavitha, T., Mehlhorn, K., and Pettie, S. 2009. Additive spanners and $(\alpha, \beta)$-spanners. *ACM Trans. on Algorithms*. to appear.

Baswana, S. and Sen, S. 2007. A simple and linear time randomized algorithm for computing sparse spanners in weighted graphs. *J. Random Structures and Algs. 30,* 4, 532–563.

Bollobás, B., Coppersmith, D., and Elkin, M. 2006. Sparse subgraphs that preserve long distances and additive spanners. *SIAM J. Discr. Math. 9,* 4, 1029–1055.

Coppersmith, D. and Elkin, M. 2006. Sparse source-wise and pair-wise preservers. *SIAM J. Discr. Math. 20,* 2, 463–501.

Cowen, L. J. 2001. Compact routing with minimum stretch. *J. Algor. 28*, 170–183.

Cowen, L. J. and Wagner, C. G. 2004. Compact roundtrip routing in directed networks. *J. Algor. 50,* 1, 79–95.

Dor, D., Halperin, S., and Zwick, U. 2000. All-pairs almost shortest paths. *SIAM J. Comput. 29,* 5, 1740–1759.

Elkin, M. and Peleg, D. 2004. $(1 + \epsilon, \beta)$-spanner constructions for general graphs. *SIAM J. Comput. 33,* 3, 608–631.

Elkin, M. and Zhang, J. 2006. Efficient algorithms for constructing $(1 + \epsilon, \beta)$-spanners in the distributed and streaming models. *Distributed Computing 18,* 5, 375–385.

Erdős, P. 1963. Extremal problems in graph theory. In *Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963), pages 29–36. Publ. House Czechoslovak Acad. Sci., Prague.*

Fakcharoenphol, J., Rao, S., and Talwar, K. 2004. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. Syst. Sci. 69,* 3, 485–497.

Halperin, S. and Zwick, U. 1996. Unpublished.

Narasimhan, G. and Smid, M. 2007. *Geometric Spanner Networks.* Cambridge University Press.

Peleg, D. 2000. *Distributed Computing: A Locality-Sensitive Approach.* SIAM.

Peleg, D. and Schaffer, A. A. 1989. Graph spanners. *Journal of Graph Theory 13*, 99–116.

Peleg, D. and Ullman, J. D. 1989. An optimal synchronizer for the hypercube. *SIAM J. Comput. 18*, 740–747.


[5] Of course, a new connection scheme could guarantee an additive distortion directly (as in our scheme **B**) but composing such a scheme with itself would—without major modifications to the framework—seem to lead directly to a distortion of the form $f(d) = d + O(\sqrt{d})$.

PELEG, D. AND UPFAL, E. 1989. A trade-off between space amd efficiency for routing tables. *J. ACM 36,* 3, 510–530.

RODITTY, L., THORUP, M., AND ZWICK, U. 2005. Deterministic constructions of approximate distance oracles and spanners. In *Proc. 32nd Int'l Colloq. on Automata, Lang., and Prog. (ICALP)*. 261–272.

RODITTY, L., THORUP, M., AND ZWICK, U. 2008. Roundtrip spanners and roundtrip routing in directed graphs. *ACM Trans. on Algorithms 4,* 3, Article 29.

SPIELMAN, D. A. AND TENG, S.-H. 2004. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proc. 36th Annual ACM Symposium on Theory of Computing (STOC)*. 81–90.

THORUP, M. AND ZWICK, U. 2001. Compact routing schemes. In *Proc. 13th ACM Symp. on Parallel Algorithms and Architectures (SPAA)*. 1–10.

THORUP, M. AND ZWICK, U. 2005. Approximate distance oracles. *J. ACM 52,* 1, 1–24.

THORUP, M. AND ZWICK, U. 2006. Spanners and emulators with sublinear distance errors. In *Proc. 17th ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 802–809.

WENGER, R. 1991. Extremal graphs with no $C^4$s, $C^6$s, or $C^{10}$s. *J. Comb. Theory Ser. B 52,* 1, 113–116.

WOODRUFF, D. 2006. Lower bounds for additive spanners, emulators, and more. In *Proc. 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 389–398.

ACKNOWLEDGMENTS

We would like to thank Uri Zwick, Mikkel Thorup, and Michael Elkin for many helpful comments.

Received Month Year; revised Month Year; accepted Month Year