A Hierarchy of Lower Bounds for Sublinear Additive Spanners*

Amir Abboud
Stanford University

Greg Bodwin
Stanford University

Seth Pettie
University of Michigan

Abstract

Spanners, emulators, and approximate distance oracles can be viewed as *lossy* compression schemes that represent an unweighted graph metric in small space, say $\tilde{O}(n^{1+\delta})$ bits. There is an inherent tradeoff between the sparsity parameter δ and the *stretch function* f of the compression scheme, but the qualitative nature of this tradeoff has remained a persistent open problem.

It has been known for some time that when $\delta \geq 1/3$ there are schemes with constant *additive* stretch (distance d is stretched to at most $f(d) = d + O(1)$), and recent results of Abboud and Bodwin show that when $\delta < 1/3$ there are no such schemes. Thus, to get practically efficient graph compression with $\delta \to 0$ we must pay super-constant additive stretch, but exactly how much do we have to pay?

In this paper we show that the lower bound of Abboud and Bodwin is just the first step in a *hierarchy* of lower bounds that characterize the asymptotic behavior of the optimal stretch function f for sparsity parameter $\delta \in (0, 1/3)$. Specifically, for any integer $k \geq 2$, any compression scheme with size $O(n^{1+\frac{1}{2^k-1}-\epsilon})$ has a *sublinear additive stretch* function f:

$$f(d) = d + \Omega(d^{1-\frac{1}{k}}).$$

This lower bound matches Thorup and Zwick's (2006) construction of sublinear additive *emulators*. It also shows that Elkin and Peleg's $(1+\epsilon, \beta)$-spanners have an essentially optimal tradeoff between δ, ϵ, and β, and that the sublinear additive spanners of Pettie (2009) and Chechik (2013) are not too far from optimal. To complement these lower bounds we present a new construction of $(1+\epsilon, O(k/\epsilon)^{k-1})$-spanners with size $O((k/\epsilon)^{h_k} k n^{1+\frac{1}{2^{k+1}-1}})$, where $h_k < 3/4$. This size bound improves on the spanners of Elkin and Peleg (2004), Thorup and Zwick (2006), and Pettie (2009). According to our lower bounds neither the size nor stretch function can be substantially improved.

Our lower bound technique exhibits several interesting degrees of freedom in the framework of Abboud and Bodwin. By carefully exploiting these freedoms, we are able to obtain lower bounds for several related combinatorial objects. We get lower bounds on the size of (β, ϵ)-*hopsets*, matching Elkin and Neiman's construction (2016), and lower bounds on *shortcutting sets* for digraphs that preserve the transitive closure. Our lower bound simplifies Hesse's (2003) refutation of Thorup's conjecture (1992), which stated that adding a linear number of shortcuts suffices to reduce the diameter to polylogarithmic. Finally, we show matching upper and lower bounds for graph compression schemes that work for graph metrics with girth at least $2\gamma + 1$. One consequence is that Baswana et al.'s (2010) additive $O(\gamma)$-spanners with size $O(n^{1+\frac{1}{2\gamma+1}})$ cannot be improved in the exponent.

1 Introduction

Spanners [46], *emulators* [27, 57], and *approximate distance oracles* [56] can be viewed as kinds of compression schemes that approximately encode the distance metric of a (dense) undirected input graph $G = (V, E)$ in small space, where the notion of *approximation* is captured by a non-decreasing *stretch function* $f : \mathbb{N} \to \mathbb{N}$.

Spanners. An $f(d)$-*spanner* $G' = (V, E')$ is a subgraph of G for which $\text{dist}_{G'}(u, v)$ is at most $f(\text{dist}_G(u, v))$. An (α, β)-spanner is one with stretch function $f(d) = \alpha d + \beta$. Notable special cases include *multiplicative* α-*spanners* [46, 8, 30, 56, 10, 9], when $\beta = 0$, and *additive* β-*spanners* [6, 27, 30, 57, 9, 62, 20, 38], when $\alpha = 1$. See [30, 9, 57, 47, 20, 43] for "mixed" spanners with $\alpha > 1, \beta > 0$.

Emulators. An $f(d)$-*emulator* (also called a *Steiner spanner* [8]) is a *weighted* graph $G' = (V' \supseteq$

*Full version: `http://arxiv.org/abs/1607.07497` . Supported by NSF grants CCF-1217338, CNS-1318294, CCF-1417238, CCF-1514339, CCF-1514383, and CCF-1637546, and BSF Grant 2012338. Email: `abboud@cs.stanford.edu`, `gbodwin@cs.stanford.edu`, `pettie@umich.edu`.

V, E', w') such that for each $u, v \in V$, $\text{dist}_{G'}(u,v) \in [\text{dist}_G(u,v), f(\text{dist}_G(u,v))]$. In other words, one is allowed to add Steiner points ($V' \backslash V$) and long-range (weighted) edges $(u,v) \in E' \backslash E$ such that distances are non-contracting.

(Unconstrained) Distance Oracles. For our purposes, an $f(d)$-approximate distance oracle using space s is a bit string in $\{0,1\}^s$ such that given $u, v \in V$, an estimate $\widetilde{\text{dist}}(u,v) \in [\text{dist}_G(u,v), f(\text{dist}_G(u,v))]$ can be computed by examining only the bit string. Note: the term "oracle" was used in [56] to indicate that $\widetilde{\text{dist}}(u,v)$ is computed in constant time [44, 3, 21]. Later work considered distance oracles with non-constant query time [48, 5, 4, 31]. In this paper we make no restrictions on the query time at all. Thus, for our purposes distance oracles generalize spanners, emulators, and related objects.

In this paper we establish essentially optimal tradeoffs between the size of the compressed graph representation and the asymptotic behavior of its stretch function f. In order to put our results in context we must recount the developments of the last 30 years that investigated multiplicative, additive, (α, β), and sublinear additive stretch functions.

1.1 Multiplicative Stretch

Historically, the first notion of stretch studied in the literature was purely multiplicative stretch. Althöfer et al. [8] quickly settled the problem by showing that any graph contains an α-spanner with at most $m_{\alpha+2}(n)$ edges, and that the claim is false for $m_{\alpha+2}(n)-1$. Here $m_g(n)$ is the maximum number of edges in a graph with n vertices and girth g. The upper bound of [8] follows directly from the observation that a natural greedy construction never closes a cycle with length at most $\alpha+1$; the lower bound follows from the fact that no strict subgraph of a graph with girth $\alpha+2$ is an α-spanner.[1] It has been conjectured [32, 17, 15] that the trivial upper bound $m_{2k+1}(n), m_{2k+2}(n) = O(n^{1+1/k})$ is sharp up to the leading constant, but this *Girth Conjecture* has only been proved for $k=1$ (trivial), and $k \in \{2,3,5\}$ [18, 33, 49, 60, 58, 12, 39]. See [40, 41, 61] for lower bounds on $m_g(n)$.

1.2 Additive Stretch

It is somewhat disingenuous to apply the girth argument to lower bound multiplicative stretch since it *only* applies to adjacent vertices, i.e., at distance $d=1$. For example, it does not imply that $f(d) = (2k-1)d$ is an optimal stretch *function* for spanners with size $O(n^{1+1/k})$. In general, the Girth Conjecture only implies that size $O(n^{1+1/k})$-spanners have $f(d) \geq 2k-d$; for example (α, β)-spanners must have $\alpha + \beta \geq 2k-1$.

Aingworth, Chekuri, Indyk, and Motwani [6] gave a construction of an additive 2-spanner with size $\tilde{O}(n^{3/2})$, which is optimal in the sense that neither the additive stretch 2 nor exponent 3/2 can be unilaterally improved.[2] This result raised the tantalizing possibility that there exist arbitrarily sparse additive spanners. Dor, Halperin, and Zwick [27] observed that additive 4-*emulators* exist with size $\tilde{O}(n^{4/3})$, i.e., the emulator introduces weighted edges connecting distant vertex pairs. Baswana, Kavitha, Mehlhorn, and Pettie [9] constructed additive 6-spanners with size $O(n^{4/3})$ and Chechik [20] constructed additive-4 spanners with size $\tilde{O}(n^{7/5})$. See [62, 38, 30, 57, 27, 9] for other constructions of additive 2- and 6-spanners.

The "4/3" exponent proved to be very resilient, for both emulators and spanners with additive stretch. This led to a line of work establishing additive spanners below the $n^{4/3}$ threshold with stretch polynomial in n [16, 9, 47, 20, 13]. The additive spanners of Bodwin and Williams [14] with stretch function $f(d) = d + n^{\epsilon}$ have size that is the minimum of $O(n^{\frac{4}{3} - \frac{7\epsilon}{9} + o(1)})$ and $O(n^{\frac{5}{4} - \frac{5\epsilon}{12} + o(1)})$.

1.3 Sublinear Additive Stretch

Elkin and Peleg [30] showed that the "4/3 barrier" could also be broken by tolerating $1+\epsilon$ multiplicative stretch. In particular, for any integer κ and real $\epsilon > 0$, there are $(1+\epsilon, \beta)$-spanners with size $O(\beta n^{1+1/\kappa})$, where $\beta = O(\epsilon^{-1} \log \kappa)^{\log \kappa}$. The construction algorithm and size-bound both depend on ϵ. Thorup and Zwick [57] gave a surprisingly simple construction of an $O(kn^{1+\frac{1}{2^{k+1}-1}})$-size *emulator* with $(1+\epsilon, O(k/\epsilon)^{k-1})$-type stretch.

Thorup and Zwick's emulator has the special property that its stretch holds for every $\epsilon > 0$ *simultaneously*, i.e., it can be selected as a function of d. Judiciously choosing $\epsilon = k/d^{\frac{1}{k}}$ leads to an emulator with a *sublinear additive stretch*

[1] Removing any edge stretches the distance between its endpoints from 1 to at least $\alpha+1$. Moreover, since every graph contains a bipartite subgraph with at least half the edges, $m_{2k+1} \leq 2m_{2k+2}(n)$ for every k. Thus, there are $(2k-1)$-spanners with size $O(m_{2k+2}(n))$.

[2] Moreover, later results of Bollobás et al. [16] show that for spanner size $O(n^{3/2})$, the stretch *function* $f(d) = d+2$ is optimal for $1 \leq d \leq \Theta(\sqrt{n})$. See [30, 57, 9, 38] for constructions of additive 2-spanners with size $O(n^{3/2})$.

function $f(d) = d + O(kd^{1-\frac{1}{k}} + 3^k)$.[3] Thorup and Zwick also showed that this same stretch function also applies to their earlier [56] construction of multiplicative $(2k+1)$-spanners with size $O(kn^{1+\frac{1}{k+1}})$. Pettie [47] gave a construction of sublinear additive *spanners* whose size-stretch tradeoff is closer to the Thorup-Zwick emulators. For stretch function $d + O(kd^{1-\frac{1}{k}} + 3^k)$ the size is $O(kn^{1+\frac{(3/4)^{k-2}}{7-2(3/4)^{k-2}}})$, which is always $o(n^{1+(3/4)^{k+3}})$ for any fixed k. At their sparsest, Thorup and Zwick's emulators [57] and Pettie's spanners [47] have size $O(n \log\log n)$ and stretch $f(d) = d + O(\log\log n) \cdot d^{1-\Theta(1/\log\log n)} + (\log n)^{\log_2 3}$. Pettie [47] gave an even sparser $(1+\epsilon, O(\epsilon^{-1}\log\log n)^{\log\log n})$-spanner with size $O(n \log\log(\epsilon^{-1}\log\log n))$.

1.4 Lower Bounds Woodruff proved that any $k^{-1}n^{1+1/k}$-size spanner with stretch function f must have $f(k) \geq 3k$. As a corollary, additive $(2k-2)$-spanners must have size $\Omega(k^{-1}n^{1+1/k})$, independent of the status of the Girth Conjecture. Bollobás, Coppersmith, and Elkin [16] showed that if the stretch f is such that $f(d) = d$ for $d \geq D$, then $\Omega(n^2/D)$-size is necessary and sufficient for spanners and emulators.

In a recent surprise, Abboud and Bodwin [1] proved that no additive β-spanners, emulators, nor distance oracles exist with $\beta = O(1)$ and exponent less than 4/3. More precisely, any construction of these three objects with additive $\beta = O(1)$ stretch has size $\Omega(n^{4/3}/2^{O(\sqrt{\log n})})$ and any construction with size $O(n^{4/3-\epsilon})$ has additive stretch $\beta = n^\delta$ for some $\delta = \delta(\epsilon)$. This result explained why all prior additive spanner constructions had a strange transition at 4/3 [27, 57, 9, 20, 14, 38, 62], but it did not suggest what the *optimal* stretch function should be for sparsity $n^{1+\delta}$ when $\delta \in [0, 1/3)$.

1.5 New Results

Distance Oracle Lower Bounds. Our main result is a hierarchy of lower bounds for spanners, emulators, and distance oracles, which shows that tradeoffs offered by Thorup and Zwick's [57] sublinear additive emulators [57] and Elkin and Peleg's $(1+\epsilon, \beta)$-spanners cannot be substantially improved. Building on Abboud and Bodwin's [1] $\Omega(n^{4/3-o(1)})$ lower bounds for additive spanners, we prove that for every integer $k \geq 2$ and $d < n^{o(1)}$, there is a graph $\mathcal{H}_k$ on n vertices and $n^{1+\frac{1}{2^k-1}-o(1)}$ edges such that any spanner with size $n^{1+\frac{1}{2^k-1}-\epsilon}$, $\epsilon > 0$, stretches vertices at distance d to at least $d + c_k d^{1-\frac{1}{k}}$ for a constant $c_k = \Theta(1/k)$. More generally, we exhibit graph families that cannot be *compressed* into distance oracles on $n^{1+\frac{1}{2^k-1}-\epsilon}$ bits such that distances can be recovered below this error threshold. The consequences of this construction are that the existing sublinear additive emulators [57], sublinear additive spanners [47, 20], and $(1+\epsilon, \beta)$-spanners [30, 57, 47] are, to varying degrees, close to optimal. Specifically,

- The $(d + O(kd^{1-\frac{1}{k}} + 3^k))$-emulator [57] with size $O(n^{1+\frac{1}{2^{k+1}-1}})$ cannot be improved by more than a constant factor in the stretch $O(kd^{1-\frac{1}{k}})$, or by a $o(1)$ in the exponent $1 + \frac{1}{2^{k+1}-1}$.
- The sublinear additive *spanners* of Pettie [47] and Chechik [20] probably have suboptimal exponents, but not by much. For example, the exponent of Chechik's [20] $\tilde{O}(n^{20/17})$-size $(d + O(\sqrt{d}))$-spanner is within 0.034 of optimal and the exponent of Pettie's [47] $O(n^{25/22})$-size $(d + O(d^{2/3}))$-spanner is within 0.07 of optimal.
- When $\epsilon \geq 1/n^{o(1)}$, the existing constructions of $(1+\epsilon, O(k/\epsilon)^{k-1})$-spanners [30, 57, 47] with size $O\Big((k/\epsilon)^{O(1)} n^{1+\frac{1}{2^{k+1}-1}}\Big)$ cannot be substantially improved in either the additive $O(k/\epsilon)^{k-1}$ term or the exponent $1+\frac{1}{2^{k+1}-1}$. This follows from the fact that any spanner with stretch of type $(1 + \hat{\epsilon}, O(k/\hat{\epsilon})^{k-1})$, for every $\hat{\epsilon} \geq \epsilon$ functions as a $(d + O(kd^{1-\frac{1}{k}}))$-spanner for distances $d \leq O(k/\epsilon)^k$. However, there is no reason to believe that the *size* of such $(1+\epsilon, \beta)$-spanners must depend on ϵ, as it does in the current constructions.

There is an interesting new hierarchy of *phase transitions* in the interplay between our lower bounds previous upper bounds [57]. Let C be a sufficiently large constant and c be a sufficiently small constant. If one wants a graph compression scheme with stretch $f(d) = d + C\sqrt{d}$, then one needs only $\widetilde{O}(n^{8/7})$ bits of space to store an emulator [57]. However, if we

[3] The Thorup-Zwick emulator can easily be converted to a $(1+\epsilon, \beta)$-spanner by replacing weighted edges with paths up to length β. A careful analysis shows the size of the resulting spanner can be made $O((k/\epsilon)^{O(1)} n^{1+\frac{1}{2^{k+1}-1}})$ (see the full version) which would slightly improve on [30]. Elkin [personal communication, 2013] has stated that with minor changes, the Elkin-Peleg [30] spanners can also be expressed as $(1+\epsilon, O(k/\epsilon)^{k-1})$-spanners with size $O((k/\epsilon)^{O(1)} n^{1+\frac{1}{2^{k+1}-1}})$. We state these bounds in Figure 1 rather than those of [30] in order to facilitate easier comparisons with subsequent constructions [57, 20, 47], and the new constructions in the full version of this paper.

Citation		**Stretch Function** $d + O\left(\sqrt{d}\right)$ or $\left(1+\epsilon, O\left(\frac{1}{\epsilon}\right)\right)$	$d + O\left(d^{\frac{2}{3}}\right)$ or $\left(1+\epsilon, O\left(\frac{1}{\epsilon}\right)^2\right)$	$d + O\left(d^{\frac{3}{4}}\right)$ or $\left(1+\epsilon, O\left(\frac{1}{\epsilon}\right)^3\right)$	$d + O\left(kd^{1-\frac{1}{k}}\right)$ or $\left(1+\epsilon, O\left(\frac{k}{\epsilon}\right)^{k-1}\right)$
Elkin & Peleg	SPAN.	$O\left(\epsilon^{-O(1)} n^{\frac{8}{7}}\right)$	$O\left(\epsilon^{-O(1)} n^{\frac{16}{15}}\right)$	$O\left(\epsilon^{-O(1)} n^{\frac{32}{31}}\right)$	$O\left(\left(\frac{k}{\epsilon}\right)^{O(1)} n^{1+\frac{1}{2^{k+1}-1}}\right)$
Thorup & Zwick	EMUL.	$O\left(n^{\frac{8}{7}}\right)$	$O\left(n^{\frac{16}{15}}\right)$	$O\left(n^{\frac{32}{31}}\right)$	$O\left(kn^{1+\frac{1}{2^{k+1}-1}}\right)$
	SPAN.	$O\left(n^{\frac{4}{3}}\right)$	$O\left(n^{\frac{5}{4}}\right)$	$O\left(n^{\frac{6}{5}}\right)$	$O\left(kn^{1+\frac{1}{k+1}}\right)$
Pettie	SPAN.	$O\left(n^{\frac{6}{5}}\right)$	$O\left(n^{\frac{25}{22}}\right)$	$O\left(n^{\frac{103}{94}}\right)$	$O\left(kn^{1+\frac{(3/4)^{k-2}}{7-2(3/4)^{k-2}}}\right)$
Chechik	SPAN.	$\tilde{O}\left(n^{\frac{20}{17}}\right)$			
New	SPAN.	$O\left(\epsilon^{-\frac{2}{7}} n^{\frac{8}{7}}\right)$	$O\left(\epsilon^{-\frac{7}{15}} n^{\frac{16}{15}}\right)$	$O\left(\epsilon^{-\frac{18}{31}} n^{\frac{32}{31}}\right)$	$O\left(\left(\frac{k}{\epsilon}\right)^{h} kn^{1+\frac{1}{2^{k+1}-1}}\right)$
New Lower Bounds	ALL	$\Omega\left(n^{\frac{4}{3}-o(1)}\right)$	$\Omega\left(n^{\frac{8}{7}-o(1)}\right)$	$\Omega\left(n^{\frac{16}{15}-o(1)}\right)$	$\Omega\left(n^{1+\frac{1}{2^k-1}-o(1)}\right)$

Figure 1: A summary of spanners and emulators with $(1+\epsilon, O(k/\epsilon)^{k-1})$-type stretch and sublinear additive stretch $d + O(kd^{1-\frac{1}{k}})$. Note: the new lower bounds do not contradict the upper bounds; the lower bounds are for stretch functions with smaller leading constants in the $O(k/\epsilon)^{k-1}$ and $O(kd^{1-\frac{1}{k}})$ terms. In the last cell of the table, $h = \frac{3\cdot 2^{k-1}-(k+2)}{2^{k+1}-1} < 3/4$, which improves the dependence on ϵ that can be obtained from modified versions of existing constructions [30, 57].

want a slightly improved stretch $f(d) = d + c\sqrt{d}$, then, by our lower bound, the space requirement leaps to $\Omega(n^{4/3-o(1)})$. In general, the optimal space for stretch function $f(d) = d + c'd^{1-1/k}$ takes a polynomial jump as we shift c' from some sufficiently large constant $O(k)$ to a sufficiently small constant $\Omega(1/k)$.

An important take-away message from our work is that the sublinear additive stretch functions of type $f(d) = d + O(d^{1-1/k})$ used by Thorup and Zwick [57] are exactly of the "right" form. For example, such plausible-looking stretch functions as $f(d) = d + O(d^{1/3})$ and $f(d) = d + O(d^{2/3}/\log d)$ could only exist in the narrow bands not covered by our lower bounds: between space $n^{4/3-o(1)}$ and $n^{4/3}$ and between space $n^{8/7-o(1)}$ and $n^{8/7}$.

Spanner Upper Bounds. To complement our lower bounds we provide new upper bounds on the sparsity of spanners with stretch of type $(1 + \hat{\epsilon}, O(k/\hat{\epsilon})^{k-1})$, which holds for every $\hat{\epsilon} \geq \epsilon$. Our new spanners have size $O((k/\epsilon)^h kn^{1+\frac{1}{2^{k+1}-1}})$, where $h = \frac{3\cdot 2^{k-1}-(k+2)}{2^{k+1}-1} < 3/4$. This construction improves on the bounds that can be derived from [57, 30, 47] in the dependence on ϵ.[4] For example, one consequence of this result is an $O(D^{1/7}n^{8/7})$-size spanner that functions as a $(d + O(\sqrt{d}))$-spanner for all $d \leq D$. This size bound is an improvement on Chechik's $(d+O(\sqrt{d}))$-spanner, as long as $D < n^{4/17}$.

Hopset Lower Bounds. Hopsets are fundamental objects that are morally similar to emulators. They were explicitly defined by Cohen [23] but used implicitly in many earlier works [59, 37, 22, 51]. Let $G = (V, E, w)$ be an arbitrary undirected *weighted* graph and $H \subset \binom{V}{2}$ be a set of edges called the *hopset*. In the united graph $G' = (V, E \cup H, w)$, the weight of an edge $(u, v) \in H$ is the length of the shortest path in G between u and v. Define the β-limited distance in G', denoted $\text{dist}^{(\beta)}_{G'}(u, v)$, to be the length of the shortest path from u to v that uses at most β edges in G'.[5] We call H a (β, ϵ)-*hopset*, where $\beta \geq 1, \epsilon > 0$, if, for any $u, v \in V$, we have

$$\text{dist}^{(\beta)}_{G'}(u, v) \leq (1 + \epsilon)\,\text{dist}_G(u, v).$$

There is clearly some three-way tradeoff between β, ϵ, and $|H|$. Elkin and Neiman [28] recently showed that any graph has a (β, ϵ)-hopset with size $\tilde{O}(n^{1+1/\kappa})$, where $\beta = O\left(\frac{\log\kappa}{\epsilon}\right)^{\log\kappa}$.[6]

In this work, we show that any construction of (β, ϵ)-hopsets with worst-case size $n^{1+\frac{1}{2^k-1}-\delta}$, where $k \geq 1$ is an integer and $\delta > 0$, must have $\beta = \Omega_k\left(\frac{1}{\epsilon}\right)^k$. For example, hopsets with $\beta = o(1/\epsilon)$ must have size $\Omega(n^{2-o(1)})$ and those with $\beta = o(1/\epsilon^2)$ must have size $\Omega(n^{4/3-o(1)})$. This essentially matches the Elkin-Neiman tradeoff, up to a constant in β that depends on k.

Lower Bounds on Shortcutting Digraphs. In 1992, Thorup [53] conjectured that the diameter of any directed graph $G = (V, E)$ could be drastically reduced with a small number of *shortcuts*. In particular, there exists another directed graph $G' = (V, E')$ with $|E'| = O(|E|)$ and the same transitive closure relation as G ($\rightsquigarrow$), such that if $u \rightsquigarrow v$, then there is a poly$(\log n)$-length path from u to v in G'. Thorup's conjecture was confirmed for trees [53, 55, 19] and planar graphs [54], but finally refuted by Hesse [34] for general graphs. In this paper we give a simpler 1-page proof of Hesse's refutation by modifying our spanner lower bound construction.

Spanners for High-Girth Graphs. Our lower bounds apply to the class of *all* undirected graph metrics. Baswana, Kavitha, Mehlhorn, and Pettie [9] gave sparser spanners for a *restricted* class of graph metrics. Specifically, graphs with girth at least $2\gamma+1$ contain additive 6γ-spanners with size $O(n^{1+\frac{1}{2\gamma+1}})$. We adapt our lower bound construction to prove that the exponent $1 + \frac{1}{2\gamma+1}$ is optimal, assuming the Girth Conjecture, and more generally we give lower bounds on compression schemes for the class of graphs with girth at least $2\gamma + 1$. Any scheme that uses $n^{1+\frac{1}{(\gamma+1)2^{k-1}-1}-\epsilon}$ bits must have stretch $f(d) \geq d + \Omega(d^{1-1/k})$, for any $d < n^{o(1)}$. We also give new constructions of emulators and spanners for girth-$(2\gamma + 1)$ graphs that shows that the exponent $1 + \frac{1}{(\gamma+1)2^{k-1}-1}$ is the best possible.

1.6 Related Work

Much of the recent work on spanners has focused on preserving or approximating distances between specified *pairs* of vertices. See [25, 2, 1] for lower bounds on pairwise spanners and [25, 47, 26, 36, 35, 2, 43, 50] for upper bounds.

[4] No bounds of this type are stated explicitly in [57] or [30]. In order to get a bound of this type—with the $1 + \frac{1}{2^{k+1}-1}$ exponent and some poly$(1/\epsilon)$ dependence on ϵ— one must only adjust the sampling probabilities of [57]; however, adapting [30] requires slightly more significant changes [Elkin, personal communication, 2013].

[5] Note that whereas $\text{dist} = \text{dist}^{(\infty)}$ is metric, $\text{dist}^{(\beta)}$ does not necessarily satisfy the triangle inequality for finite β.

[6] It is likely that Elkin and Neiman's tradeoff could be more precisely stated as follows: for any positive integer k and $\epsilon > 0$, there is an $\tilde{O}(n^{1+\frac{1}{2^{k+1}-1}})$ size (β, ϵ)-hopset with $\beta = O(k/\epsilon)^k$.

Pairwise spanners have proven to be useful tools for constructing (sublinear) additive spanners; see [47, 20, 13].

The space/stretch tradeoffs offered by the best distance oracles [21, 44, 45, 3, 5, 4, 31] are strictly worse than those of the best spanners and emulators, even though distance oracles are entirely *unconstrained* in how they encode the graph metric. This is primarily due to the requirement that distance oracles respond to queries quickly. There are both unconditional [52] and conditional [24, 44, 45] lower bounds suggesting that distance oracles with reasonable query time cannot match the best spanners or emulators.

2 Formal Statement of Our Theorems

In this section we give the formal statements of the results that we discussed in the introduction. All proofs can be found in the full version of the paper.

2.1 Lower Bounds for Spanners We start with our main result that exhibits a hierarchy of additive spanner lower bounds.

THEOREM 2.1. *For any integer $k \geq 2$ and a sufficiently small constant $c_k = O(1/k)$, any spanner construction with stretch function bounded by $f(d) \leq d + c_k d^{1-\frac{1}{k}} + \tilde{O}(1)$ has size $\Omega(n^{1+\frac{1}{2^{k+1}-1}-o(1)})$ in the worst case.*

The next theorem presents our lower bound for mixed spanners.

THEOREM 2.2. **($(1+\epsilon,\beta)$-Spanner Lower Bounds)** *Any $(1+\epsilon,\beta)$ spanner construction with worst-case size at most $n^{1+\frac{1}{2^{k+1}-1}-\delta}$, $\delta > 0$, has $\beta = \Omega\left(\frac{1}{\epsilon(k-1)}\right)^{k-1}$.*

Theorem 2.2 shows that the existing $(1+\epsilon, O(k/\epsilon)^{k-1})$-spanners with size $O((k/\epsilon)^{O(1)} n^{1+\frac{1}{2^{k+1}-1}})$ are optimal in the following sense. If k is constant then we cannot improve β by more than a constant factor $\approx (k^2)^{k-1}$ without increasing the exponent to $1 + \frac{1}{2^k-1} - o(1)$. Moreover, any constant reduction in the exponent increases β to $\Theta(1/(k\epsilon))^k$.

The next theorem strengthens the results above showing that the lower bounds hold for any kind of compression.

THEOREM 2.3. **(Distance Oracle Lower Bounds)** *Consider any data structure for the class of n-vertex undirected graphs that answers approximate distance queries. If its stretch function is:*

- *$f(d) \leq d + c_k d^{1-\frac{1}{k}} + \tilde{O}(1)$ for an integer k and a sufficiently small constant $c_k < 2/(k-1)^{1-1/k}$, or*
- *$f(d) \leq (1+\epsilon)d + \beta$ where $\beta = o\left(\left(\frac{1}{\epsilon(k-1)}\right)^{k-1}\right)$*

then on some graph, the data structure occupies at least $n^{1+\frac{1}{2^k-1}-o(1)}$ bits of space.

2.2 Lower Bounds for Hopsets Next, we show lower bounds on the tradeoffs between β and ϵ in (β,ϵ)-hopsets, subject to an upper bound on the number of edges in the hopset.

THEOREM 2.4. *Fix a positive integer k and parameter $\epsilon > 1/n^{o(1)}$. Any construction of (β,ϵ)-hopsets with size $n^{1+\frac{1}{2^k-1}-\delta}$, $\delta > 0$, has $\beta = \Omega_k\left(\frac{1}{\epsilon}\right)^k$.*

Observe that Theorem 2.4 implies several interesting corollaries: any (β,ϵ)-hopset with $\beta = o(1/\epsilon)$ must have size $\Omega(n^{2-o(1)})$ and any such hopset with $\beta = o(1/\epsilon^2)$ must have size $\Omega(n^{4/3-o(1)})$.

2.3 Lower Bounds on Compressing High Girth Graphs Baswana et al. [9] showed that the class of graphs with girth (length of the shortest cycle) larger than 4 contains additive spanners below the 4/3 threshold. For example, graphs with girth 5 contain additive 12-spanners with size $O(n^{6/5})$.

THEOREM 2.5. ([9]) *For any integer $\gamma \geq 1$, any graph with girth at least $2\gamma+1$ contains an additive 6γ-spanner on $O(n^{1+\frac{1}{2\gamma+1}})$ edges.*

We extend our lower bound technique to show that the exponent of Theorem 2.5 is optimal. More generally, we establish a hierarchy of tradeoffs for sublinear additive graph compression schemes that depend on k and γ.

Our construction uses a slightly stronger, but equivalent, statement of the Girth Conjecture [32, 17, 15] that asserts a lower bound on the degree rather than the total size: For any integer $\gamma \geq 1$, there exists a graph with n vertices, girth $2\gamma+2$, and minimum degree $\Omega(n^{1/\gamma})$.

We prove the following theorem. Observe that by setting $k = 2$, Theorem 2.6 implies that the exponent of Theorem 2.5 cannot be improved.

THEOREM 2.6. *Fix integers $\gamma \geq 1, k \geq 2$. Consider any data structure that answers approximate distance* *queries for the class of n-vertex undirected graphs with girth at least $2\gamma + 1$. Assuming the Girth Conjecture, if the stretch function of the data structure is $f(d) < d + c_{k,\gamma} d^{1-1/k}$, for $c_{k,\gamma} \approx \frac{2}{(\gamma(k-1))^{1-1/k}}$ and d sufficiently large then on some graph the data structure occupies at least $\Omega(n^{1+\frac{1}{(\gamma+1)2^{k-1}-1}-o(1)})$ bits.*

We also provide a matching upper bound extending the result of Baswana et al. [9] for larger $k > 2$.

THEOREM 2.7. *Let G be a graph with girth at least $2\gamma + 1$. There is an additive 4γ-emulator and additive 6γ-spanner for G with size $O(n^{1+\frac{1}{2\gamma+1}})$. For any integer $k \geq 2$, there is an $(d + O(\gamma k d^{1-1/k}))$-emulator for G with size $O(kn^{1+\frac{1}{(\gamma+1)2^k-1}})$ and a $(1 + \epsilon, O(\gamma k/\epsilon)^{k-1})$-spanner for G with size $O((\gamma k/\epsilon)^h kn^{1+\frac{1}{(\gamma+1)2^k-1}})$, where $h = \frac{3\cdot 2^{k-1}-(k+2)}{(\gamma+1)2^k-1} < \frac{3}{2(\gamma+1)}$.*

3 Conclusion

In this paper, we characterized the optimal asymptotic behavior of sublinear additive stretch functions f for spanners, emulators, or any graph compression scheme. Roughly speaking, any representation using $n^{1+\frac{1}{2^k-1}-\delta}$ bits (for any $\delta > 0$) must have stretch function $f(d) = d + \Omega(d^{1-\frac{1}{k}})$. Previous constructions of sublinear additive emulators [57] and $(1 + \epsilon, \beta)$-spanners ([30, 57] and the construction in the full version of this paper) show that neither the exponent $1 + \frac{1}{2^k-1}$ nor additive stretch $\Omega(d^{1-\frac{1}{k}})$ can be improved, for any k,

The main distinction between $(1 + \epsilon, \beta)$-spanners [30, 57, 47] and sublinear additive emulators/spanners [57, 47, 20] is that constructions of the former take ϵ as a parameter (which affects the size of the spanner) whereas the latter have $(1+\epsilon, \beta)$-stretch for all ϵ, that is, ϵ can be chosen in the analysis. An interesting open question is whether one can match the size-stretch tradeoff of Thorup and Zwick's optimal *emulators* [57] with a *spanner*. (Constructions in [47, 20] are off from [57] (and our lower bounds) by a polynomial factor.) It would be possible to construct such spanners given a *pairwise* spanner with a sublinear additive stretch function. For example, when $S \subset V$ with $|S| = \Omega(n^{4/7})$ and $P = S \times S$, does there exist a pairwise spanner for P with stretch $d + O(\sqrt{d})$ and size $O(|P|)$? If such an object existed, we would immediately have an optimal $(d + O(\sqrt{d}))$-spanner with size $O(n^{8/7})$; see [57, 47].

Our lower bounds match the existing upper bounds in the distance regime $2^{\Omega(k)} \ll d < n^{o(1)}$, while they say nothing when $d = 2^{O(k)}$ and they are weaker when $d = n^{\Omega(1)}$. An interesting open problem is to understand the sparseness-stretch tradeoffs available when $d = O(2^k)$ is tiny (see [29, 9, 43]) and when $d = n^{\Omega(1)}$ is very large [16, 14].

Acknowledgment. We are grateful to Virginia Vassilevska Williams for useful technical discussions and for advice about the directions taken by this paper.

References

[1] A. Abboud and G. Bodwin. The 4/3 additive spanner exponent is tight. In *Proceedings 48th Annual ACM Symposium on Theory of Computing (STOC)*, pages 351–361, 2016.

[2] A. Abboud and G. Bodwin. Error amplification for pairwise spanner lower bounds. In *Proceedings 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 841–854, 2016.

[3] I. Abraham and C. Gavoille. On approximate distance labels and routing schemes with affine stretch. In *Proceedings 25th International Symposium on Distributed Computing (DISC)*, pages 404–415, 2011.

[4] R. Agarwal. The space-stretch-time tradeoff in distance oracles. In *Proceedings 22nd Annual European Symposium Algorithms (ESA)*, pages 49–60, 2014.

[5] R. Agarwal and P. B. Godfrey. Distance oracles for stretch less than 2. In *Proceedings 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 526–538, 2013.

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

[7] N. Alon. Testing subgraphs in large graphs. In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 434–441, 2001.

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

[9] S. Baswana, T. Kavitha, K. Mehlhorn, and S. Pettie. Additive spanners and (α, β)-spanners. *ACM Trans. on Algorithms*, 2009.

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

[11] F. Behrend. On sets of integers which contain no three terms in arithmetic progression. *Proc. Nat. Acad. Sci.*, 32:331–332, 1946.

[12] C. Benson. Minimal regular graphs of girth eight

and twelve. *Canadian Journal of Mathematics*, 18:1091–1094, 1966.

[13] G. Bodwin and V. Vassilevska Williams. Very sparse additive spanners and emulators. In *Proceedings 2015 Conference on Innovations in Theoretical Computer Science (ITCS)*, pages 377–382, 2015.

[14] G. Bodwin and V. Vassilevska Williams. Better distance preservers and additive spanners. In *Proceedings 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 855–872, 2016.

[15] B. Bollobás. *Extremal graph theory*, volume 11 of *London Mathematical Society Monographs*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1978.

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

[17] J. Bondy and M. Simonovits. Cycles of even length in graphs. *J. Combinatorial Theory, Series B*, 16:97–105, 1974.

[18] W. G. Brown. On graphs that do not contain a Thomsen graph. *Canad. Math. Bull.*, 9:281–285, 1966.

[19] B. Chazelle. Computing on a free tree via complexity-preserving mappings. *Algorithmica*, 2(3):337–361, 1987.

[20] S. Chechik. New additive spanners. In *Proceedings 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 498–512, 2013.

[21] S. Chechik. Approximate distance oracles with improved bounds. In *Proceedings 47th Annual ACM on Symposium on Theory of Computing (STOC)*, pages 1–10, 2015.

[22] E. Cohen. Using selective path-doubling for parallel shortest-path computations. *Journal of Algorithms*, 22(1):30–56, 1997.

[23] E. Cohen. Polylog-time and near-linear work approximation scheme for undirected shortest-paths. *J. ACM*, 47:132–166, 2000.

[24] H. Cohen and E. Porat. On the hardness of distance oracle for sparse graph. *CoRR*, abs/1006.1117, 2010.

[25] D. Coppersmith and M. Elkin. Sparse sourcewise and pair-wise preservers. *SIAM J. Discrete Mathematics*, 20(2):463–501, 2006.

[26] M. Cygan, F. Grandoni, and T. Kavitha. On pairwise spanners. In *Proceedings 30th International Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 209–220, 2013.

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

[28] M. Elkin and O. Neiman. Hopsets with constant hopbound, and applications to approximate shortest paths. In *Proc. 57th IEEE Symposium on Foundations of Computer Science (FOCS), to appear*, 2016.

[29] M. Elkin and D. Peleg. $(1+\epsilon,\beta)$-spanner constructions for general graphs. In *Proc. 33rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 173–182, 2001.

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

[31] M. Elkin and S. Pettie. A linear-size logarithmic stretch path-reporting distance oracle for general graphs. *ACM Trans. Algorithms*, 12(4):50, 2016.

[32] P. Erdős. 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*, 1963.

[33] P. Erdős, A. Rényi, and V. T. Sós. On a problem of graph theory. *Studia Sci. Math. Hungar.*, 1:215–235, 1966.

[34] W. Hesse. Directed graphs requiring large numbers of shortcuts. In *Proceedings 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 665–669, 2003.

[35] T. Kavitha. New pairwise spanners. In *Proceedings 32nd International Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 513–526, 2015.

[36] T. Kavitha and N. M. Varma. Small stretch pairwise spanners and approximate d-preservers. *SIAM J. Discrete Mathematics*, 29(4):2239–2254, 2015.

[37] P. N. Klein and S. Subramanian. A randomized parallel algorithm for single-source shortest paths. *J. Algor.*, 25(2):205–220, 1997.

[38] M. B. T. Knudsen. Additive spanners: A simple construction. In *Proceedings 14th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT)*, pages 277–281, 2014.

[39] F. Lazebnik and V. A. Ustimenko. New examples of graphs without small cycles and of large size. *European J. of Combinatorics*, 14:445–460, 1993.

[40] F. Lazebnik, V. A. Ustimenko, and A. J. Woldar. A new series of dense graphs of high girth. *Bulletin of the AMS*, 32(1):73–79, 1995.

[41] F. Lazebnik, V. A. Ustimenko, and A. J. Woldar. A characterization of the components of the graphs $d(k,q)$. *Discrete Mathematics*, 157(1–3):271–283, 1996.

[42] J. Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel J. Math.*, 93:333–344, 1996.

[43] M. Parter. Bypassing Erdős' girth conjecture: Hybrid stretch and sourcewise spanners. In *Proceedings 41st International Colloquium Automata, Languages, and Programming (ICALP)*, pages 608–619, 2014.

[44] M. Pătraşcu and L. Roditty. Distance oracles beyond the Thorup-Zwick bound. *SIAM J. Comput.*, 43(1):300–311, 2014.

[45] M. Patrascu, L. Roditty, and M. Thorup. A new infinity of distance oracles for sparse graphs. In *Pro-*

ceedings 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS), pages 738–747, 2012.

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

[47] S. Pettie. Low distortion spanners. *ACM Transactions on Algorithms*, 6(1), 2009.

[48] E. Porat and L. Roditty. Preprocess, set, query! *Algorithmica*, 67(4):516–528, 2013.

[49] I. Reiman. Über ein Problem von K. Zarankiewicz. *Acta. Math. Acad. Sci. Hungary*, 9:269–273, 1958.

[50] L. Roditty, M. Thorup, and U. Zwick. Deterministic constructions of approximate distance oracles and spanners. In *Proc. 32nd Int'l Colloq. on Automata, Lang., and Prog. (ICALP)*, pages 261–272, 2005.

[51] H. Shi and T. H. Spencer. Time-work tradeoffs of the single-source shortest paths problem. *Journal of Algorithms*, 30(1):19–32, 1999.

[52] C. Sommer, E. Verbin, and W. Yu. Distance oracles for sparse graphs. In *Proceedings 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 703–712, 2009.

[53] M. Thorup. On shortcutting digraphs. In *Proceedings 18th International Workshop on Graph Theoretic Concepts in Computer Science (WG)*, pages 205–211, 1992.

[54] M. Thorup. Shortcutting planar digraphs. *Combinatorics, Probability & Computing*, 4:287–315, 1995.

[55] M. Thorup. Parallel shortcutting of rooted trees. *J. Algorithms*, 23(1):139–159, 1997.

[56] M. Thorup and U. Zwick. Approximate distance oracles. *J. ACM*, 52(1):1–24, 2005.

[57] M. Thorup and U. Zwick. Spanners and emulators with sublinear distance errors. In *Proc. 17th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 802–809, 2006.

[58] J. Tits. Sur la trialité et certains groupes qui s'en déduisent. *Publ. Math. I.H.E.S.*, 2:14–20, 1959.

[59] J. D. Ullman and M. Yannakakis. High-probability parallel transitive-closure algorithms. *SIAM J. Comput.*, 20(1):100–125, 1991.

[60] R. Wenger. Extremal graphs with no C^4s, C^6s, or C^{10}s. *J. Combin. Theory Ser. B*, 52(1):113–116, 1991.

[61] A. Woldar and V. Ustimenko. An application of group theory to extremal graph theory. In *Group theory, Proceedings of the Ohio State-Denison Conference*, pages 293–298, River Edge, NJ, 1993. World Sci. Publishing.

[62] D. P. Woodruff. Additive spanners in nearly quadratic time. In *Proceedings 37th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 463–474, 2010.