

Augmented Ring Networks

William Aiello* Sandeep N. Bhatt† Fan R. K. Chung‡ Arnold L. Rosenberg§
Ramesh K. Sitaraman¶

June 27, 1997

Abstract

We study three augmentations of ring networks that are intended to decrease a ring's diameter significantly while increasing its structural complexity only modestly. **Chordal rings** enhance a ring network by adding noncrossing "shortcut" edges, which can be viewed as chords of the ring. **Express rings** are chordal rings whose chords are oriented either clockwise or counterclockwise, allowing them to be viewed as (noncrossing) arcs of the ring. **Multi-rings** append subsidiary rings to edges of a ring and, recursively, to edges of appended subrings. Important measures of structural complexity are: the **cutwidth** of an express ring, viz., the maximum number of arcs that cross "above" any ring edge (counting the edge itself); the **depth** of a multi-ring, viz., the level of recursive appending of subsidiary subrings. Our first result demonstrates the topological equivalence of these three modes of augmentation: for each augmented ring of one type, there are (graph-theoretically) isomorphic augmented rings of each of the other types; moreover, the cutwidth of an express ring is the depth of its isomorphic multi-ring, and vice versa. Our second focus is on the question of how much decrease in diameter is achievable for a given increase in structural complexity. We establish a tight diameter-cutwidth tradeoff for express rings: for each N and c , we exhibit a cutwidth- c , N -node express ring whose diameter is at most $2^{-1/c}cN^{1/c}$; and, we prove that no such express ring can have diameter smaller than $(4e)^{-1}cN^{1/c} - c/2$. Finally, we prove that our (nontraditional) insistence that the arcs in an express ring of given cutwidth be *noncrossing* at most doubles the diameter of the augmented ring.

*Bell Communications Research, 435 South St., Morristown, NJ 07960 (aiello@bellcore.com)

†Bell Communications Research, 435 South St., Morristown, NJ 07960 (bhatt@bellcore.com)

‡Dept. of Mathematics, Univ. of Pennsylvania, Philadelphia, PA 19104 (chung@math.upenn.edu)

§Dept. of Computer Science, Univ. of Massachusetts, Amherst, MA 01003 (rsnbrg@cs.umass.edu)

¶Dept. of Computer Science, Univ. of Massachusetts, Amherst, MA 01003 (ramesh@cs.umass.edu)

1 Introduction

Since the earliest uses of networks for communication and computation, there has been serious interest in ring networks because of their structural simplicity and (modest) fault tolerance. Of course, this interest has been moderated by the large diameter and small (bisection) bandwidth of rings. It was natural for people to seek ways to augment a ring network in a manner that decreases the network's diameter significantly while increasing its structural complexity only modestly. In this paper we study three such avenues for augmentation. After proving that the three avenues produce networks that are equivalent in a strong sense, we concentrate on determining how much decrease in diameter these avenues can yield for a given increase in structural complexity. We obtain a tight tradeoff between diameter and complexity from express networks (and, via inheritance, for the other two classes of augmented rings).

Our three equivalent augmentations of ring networks are as follows.

- A chordal ring is obtained from a ring network by adding noncrossing “shortcut” edges, which can be viewed as chords of the ring. Chordal rings were first proposed as cost-efficient interconnection networks for parallel architectures [1].
- An express ring is obtained from a chordal ring by orienting its “shortcut” chords in either the clockwise or counterclockwise sense, thereby turning the chords into arcs of the ring. An important measure of the structural complexity of an express ring is its cutwidth, i.e., the maximum number of arcs that cross “above” any ring edge (counting the edge itself). Express rings represent the simplest instance of the express networks of [7]; when generalized to allow crossing “shortcuts,” they also implement the shortcuts (or, hops) advocated in [10]-[12], [14] for optical communication networks.
- A multi-ring is obtained from a ring network by appending subsidiary rings to edges of the ring and, recursively, to edges of subsidiary subrings. An important measure of the structural complexity of a multi-ring is its depth, i.e., the depth of the recursive appending of subsidiary rings. Multi-rings are a variation on the theme of multi-level ring interconnection networks such as one finds in the KSR1 multiprocessor¹ [4]; they are also identical with the SONET (Synchronous Optical NETWORK) multi-rings [6] that are important in the realm of communication networks.

These families are defined in more detail in Section 2.1.

Our first, qualitative, results (Section 2.2) establish the equivalence of these three families of augmented ring networks in a strong sense. Most basically, for each augmented ring of one type, there are (graph-theoretically) isomorphic augmented rings of each of the other types. More importantly, any cutwidth- c express ring can be replaced by an isomorphic depth- c multi-ring, and vice versa.

¹Our multi-rings differ from those of the KSR1 in that our subsidiary rings hang off edges of the parent ring, while those of the KSR1 hang off nodes.

Our second set of results (Section 3) focuses on quantitative aspects of ring augmentation, improving and/or building on work discussed in Section 3.1. We establish a tight diameter-cutwidth tradeoff for express rings: for each c and N , we exhibit a cutwidth- c , N -node express ring whose diameter is at most $2^{-1/c}cN^{1/c}$ (Section 3.2); and, we prove that no cutwidth- c , N -node express ring can have diameter smaller than $(4e)^{-1}cN^{1/c} - c/2$ (Section 3.3). Our bounds are tight to within a small constant factor when c is “small;” for instance, when $c \leq \frac{1}{6} \log_2 N$, the bounds are within a factor of 12 of each other. Finally, we prove that our (nontraditional) insistence that the arcs in an express ring of given cutwidth be *noncrossing* at most doubles the diameter of the augmented ring (Section 4.1).

We close the paper by discussing possible avenues for extending our study. In Section 4.2, we somewhat simplify a proof from [14] which shows that the nontrivial diameter-cutwidth tradeoff we expose for express rings degenerates for express meshes. In Section 4.3, we present a partial extension of our tradeoff for augmented rings in which one charges for traversing long edges.

2 Three Augmented Ring Networks

In this section, we formally define our three avenues for augmenting rings, and we prove that all three avenues lead to the same family of enhanced ring networks.

2.1 The Three Augmentations

The N -node ring network \mathcal{R}_N has node-set $Z_N \stackrel{\text{def}}{=} \{0, 1, \dots, N-1\}$ and edge-set $\{(i, i+1 \bmod N) \mid i \in Z_N\}$. Fig. 1 depicts the ring \mathcal{R}_{18} .

Chordal rings. One obtains a chordal ring by adding noncrossing “shortcut” edges (which can be viewed as chords) to a ring network. Fig. 2 depicts a chordal augmentation of the ring \mathcal{R}_{18} . Chordal rings seem to have originated in [1], and their diameters were studied in [3].² Neither of the preceding references demands (as we do) that “shortcuts” be noncrossing, but this added requirement is essential for the equivalence we demonstrate in Section 2.2; and, as we show in Section 4.1, the requirement does not impact our quantitative results materially.

Express rings. One obtains an express ring by supplying (independently) either a clockwise or counterclockwise orientation to each “shortcut” of a chordal ring, thereby allowing the “shortcuts” to be viewed as (noncrossing) arcs of the ring. Fig. 3 depicts an express ring that is topologically equivalent to the chordal ring of Fig. 2, with the edge-orientations indicated in Table 1 (which does not list the underlying ring edges). Using the convention of the third column of the table, we shall henceforth avoid the need to label edges of express rings explicitly, by viewing the edges as *ordered* pairs of nodes. Henceforth, let us use the word hop to refer ambiguously to an arc of an express ring or an edge of the underlying ring.

²The converse problem, of determining how *deleting* edges can *increase* diameter is studied in [5].

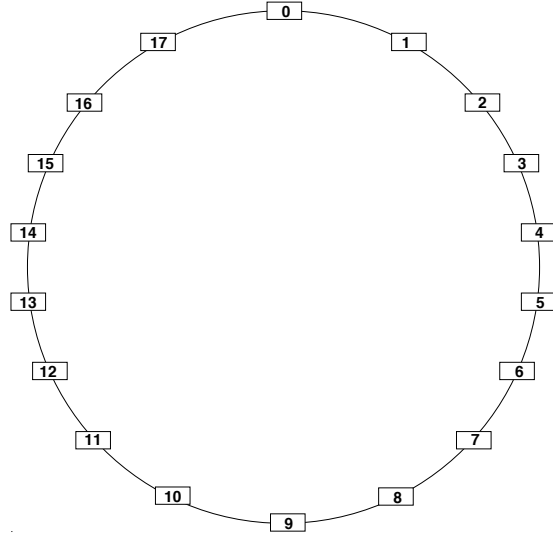


Figure 1: *The 18-node ring network \mathcal{R}_{18} .*

A natural measure of the complexity of an express network \mathcal{R} is its *cutwidth*, i.e., the maximum number of hops that “cross over” an adjacent pair of ring nodes. More precisely, if \mathcal{R} has N nodes, then its cutwidth is the maximum cardinality, over all $i \in Z_N$, of the set

$$\{(j, k) \in \text{Hops}(\mathcal{R}) \mid j \leq i \text{ and } k \geq i + 1 \bmod N\}.$$

Multi-rings. One obtains a multi-ring by appending subsidiary rings to edges of a ring network and, recursively, to edges of subsidiary subrings. Fig. 4 depicts a multi-ring that is topologically equivalent to the express ring of Fig. 3, as we shall verify in Section 2.2. Since the equivalence of multi-rings and our other augmentations is a bit subtler than the equivalence of chordal rings and express rings, we now define our notion of multi-ring formally.

The node-set of a c -level multi-ring \mathcal{M} consists of c pairwise disjoint sets, V_1, V_2, \dots, V_c , each V_i comprising the *level- i nodes* of \mathcal{M} . The edges of \mathcal{M} are specified implicitly as follows.

- The induced subgraph on the level-1 node-set V_1 is the (unique) *level-1 ring* of \mathcal{M} .
In Fig. 4, $V_1 = \{0, 4, 8, 12, 13\}$.
- The level-2 node-set V_2 is a disjoint union of some number $k_2 \leq |V_1|$ disjoint sets, $V_{2,1}, V_{2,2}, \dots, V_{2,k_2}$.
 - Each set $V_{2,j}$ is associated with a distinct pair of nodes $u_{2,j}, v_{2,j}$ that are adjacent in the level-1 ring of \mathcal{M} ;
 - the induced subgraph of \mathcal{M} on each node-set $V_{2,j} \cup \{u_{2,j}, v_{2,j}\}$ is a *level-2 ring* of \mathcal{M} .

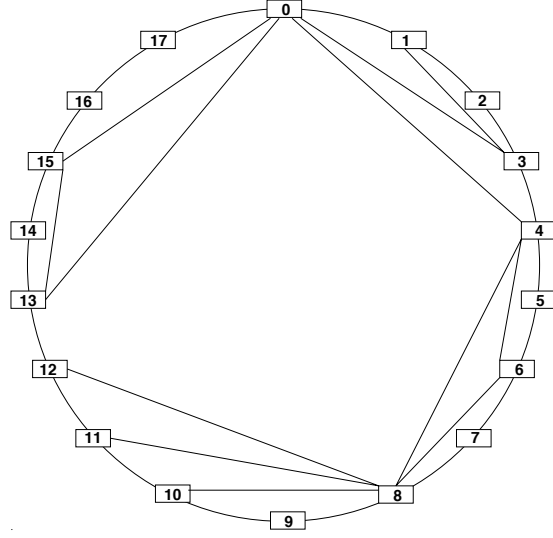


Figure 2: A chordal augmentation of \mathcal{R}_{18} .

In Fig. 4: $V_{2,1} = \{3\}$; $V_{2,2} = \{6\}$; $V_{2,3} = \{11\}$; $V_{2,4} = \{15\}$.

- For each level $i > 2$, the level- i node-set V_i is a disjoint union of some number $k_i \leq |V_{i-1}| + k_{i-1}$ disjoint sets, $V_{i,1}, V_{i,2}, \dots, V_{i,k_i}$.
 - Each set $V_{i,j}$ is associated with a distinct pair of adjacent nodes $u_{i,j}, v_{i,j}$ from a level- $(i-1)$ ring of \mathcal{M} . At least one of $u_{i,j}, v_{i,j}$ must be an element of V_{i-1} ; the other could belong to a lower-index V_k .
 - The induced subgraph of \mathcal{M} on each node-set $V_{i,j} \cup \{u_{i,j}, v_{i,j}\}$ is a level- i ring of \mathcal{M} .

In Fig. 4: $V_{3,1} = \{1\}$; $V_{3,2} = \{5\}$; $V_{3,3} = \{7\}$; $V_{3,4} = \{10\}$; $V_{3,5} = \{14\}$; $V_{3,6} = \{16, 17\}$; while $V_{4,1} = \{2\}$; $V_{4,2} = \{9\}$.

The reader should recognize how this definition formalizes our earlier description of the model.

2.2 The Equivalence of the Models

We now show that the three models of Section 2.1 are just different views of the same family of graphs.

Theorem 2.1 (a) [9] *In linear time, one can transform any chordal ring into a (graph-theoretically) isomorphic express ring of minimum cutwidth, and vice versa.*

(b) *In linear time, one can transform any cutwidth- c express ring into a (graph-theoretically) isomorphic c -level multi-ring, and vice versa.*

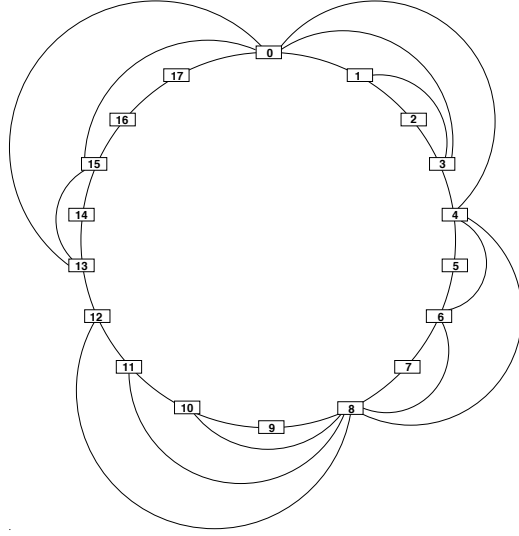


Figure 3: *An express ring realization of the chordal ring of Fig. 2.*

Proof. We prove only part (b) of the theorem, since the two claims of part (a) are either simple or known: One can transform an express ring to an equivalent chordal ring simply by ignoring the (clockwise, counterclockwise) arc orientations; one finds in [9] an efficient algorithm for transforming a chordal ring to an equivalent express ring having optimal cutwidth (a slightly simpler but less efficient algorithm appears in [15]).

Transforming an express ring to a multi-ring. Let \mathcal{R} be any cutwidth- c express ring. We transform \mathcal{R} into an equivalent c -level multi-ring by decomposing it into c shells: shell 1 will be a subring of \mathcal{R} , and each other shell will be a set of subrings. The subrings within each shell k will turn out to be the level- k subrings of \mathcal{R} 's equivalent multi-ring. The reader should note that each subring within shell $i > 1$ contains precisely one edge from a subring within shell $i - 1$; and distinct subrings within shell i may intersect in at most one node. The decomposition proceeds as follows.

Shell 1 of \mathcal{R} is the subring formed by the hops of \mathcal{R} that are “exposed,” in the sense of not being contained in any other hop of \mathcal{R} . The level-1 node-set of \mathcal{R} comprises the nodes of the shell-1 ring. In Fig. 3, shell 1 comprises the ring

$$0 \leftrightarrow 4 \leftrightarrow 8 \leftrightarrow 12 \leftrightarrow 13 \leftrightarrow 0,$$

so $V_1 = \{0, 4, 8, 12, 13\}$.

Shell 2 of \mathcal{R} is obtained by focusing, in turn, on each hop of shell 1. Say that (u, v) is such a hop, but that nodes u and v are not adjacent in the ring underlying \mathcal{R} . Then there is a path in \mathcal{R} that connects nodes u and v and that uses hops that become “exposed” when the hops of shell 1 are removed. The intermediate nodes of that path become one of the level-2 node-sets $V_{2,j}$ of \mathcal{R} ; and, that path, plus hop (u, v) becomes one of the subrings of shell 2. In Fig. 3, shell

Chord	Orientation	Arc notation
(0, 3)	clockwise	(0, 3)
(0, 4)	clockwise	(0, 4)
(0, 13)	counterclockwise	(13, 0)
(0, 15)	counterclockwise	(15, 0)
(1, 3)	clockwise	(1, 3)
(4, 6)	clockwise	(4, 6)
(4, 8)	clockwise	(4, 8)
(6, 8)	clockwise	(6, 8)
(8, 10)	clockwise	(8, 10)
(8, 11)	clockwise	(8, 11)
(8, 12)	clockwise	(8, 12)
(13, 15)	clockwise	(13, 15)

Table 1: *The arc-orientations that transform the chordal ring of Fig. 2 to the express ring of Fig. 3.*

2 comprises the subrings

$$\begin{aligned}
&0 \leftrightarrow 3 \leftrightarrow 4 \leftrightarrow 0 \\
&4 \leftrightarrow 6 \leftrightarrow 8 \leftrightarrow 4 \\
&8 \leftrightarrow 11 \leftrightarrow 12 \leftrightarrow 8 \\
&13 \leftrightarrow 15 \leftrightarrow 0 \leftrightarrow 13,
\end{aligned}$$

so $V_2 = \{3, 6, 11, 15\}$.

In general, shell $i + 1$ of \mathcal{R} is obtained by focusing, in turn, on each hop of shell i that contains at least one node from node-set $V_i \stackrel{\text{def}}{=} \bigcup_j V_{i,j}$. Say that (u, v) is such a hop, but that nodes u and v are not adjacent in the ring underlying \mathcal{R} . Then there is a path in \mathcal{R} that connects nodes u and v and that uses hops that become “exposed” when the hops of all shells $k \leq i$ are removed. The intermediate nodes of that path become one of the level- $(i + 1)$ node-sets $V_{i+1,j}$ of \mathcal{R} ; and, that path, plus hop (u, v) becomes one of the subrings of shell $i + 1$. In Fig. 3, shell 3 comprises the subrings

$$\begin{aligned}
&0 \leftrightarrow 1 \leftrightarrow 3 \leftrightarrow 0 \\
&4 \leftrightarrow 5 \leftrightarrow 6 \leftrightarrow 4 \\
&6 \leftrightarrow 7 \leftrightarrow 8 \leftrightarrow 6 \\
&8 \leftrightarrow 10 \leftrightarrow 11 \leftrightarrow 8 \\
&13 \leftrightarrow 14 \leftrightarrow 15 \leftrightarrow 13 \\
&15 \leftrightarrow 16 \leftrightarrow 17 \leftrightarrow 0 \leftrightarrow 15,
\end{aligned}$$

so $V_3 = \{1, 5, 7, 10, 14, 16, 17\}$; and shell 4 comprises the subrings

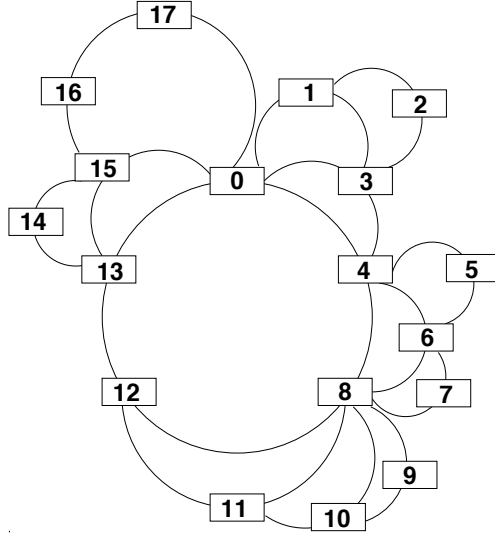


Figure 4: A multi-ring realization of the express ring of Fig. 3.

$$\begin{aligned}
 1 &\leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 1 \\
 8 &\leftrightarrow 9 \leftrightarrow 10 \leftrightarrow 8,
 \end{aligned}$$

so $V_4 = \{2, 9\}$.

We have, thus, constructed a multi-ring whose nodes and edges are, respectively, the nodes and hops of the express ring \mathcal{R} . The reader can verify that our running example has constructed the multi-ring of Fig. 4 from the express ring of Fig. 3.

Transforming a multi-ring to an express ring. The easiest way to perform this transformation proceeds in two steps. First, trace the obvious hamiltonian cycle in a given multi-ring, making all unused edges into chords of the ring obtained from the trace. Then, invoke part (a) to transform the resulting chordal ring into an express ring. The problem with this approach is that one must proceed carefully if one wants a c -level multi-ring to produce a cutwidth- c express ring. We choose, therefore to use an iterative transformation that allows us to keep track of the parameter c in both the multi-ring and the express ring in a perspicuous way.

We begin with a c -level multi-ring \mathcal{M} . We transform \mathcal{M} into an equivalent express ring by successively “collapsing” its subsidiary rings. We begin with the level-1 ring of \mathcal{M} . Make all of its edges *level-1 tentative edges* of the express ring $\mathcal{R}_{\mathcal{M}}$ that we are constructing. In general, we look at each tentative edge in turn. When we are looking at a level- i tentative edge (u, v) : if there is no level- $(i + 1)$ ring that contains edge (u, v) , then we make this a *final edge* of $\mathcal{R}_{\mathcal{M}}$; if there is a level- $(i + 1)$ ring $u \leftrightarrow w_1 \leftrightarrow w_2 \leftrightarrow \dots \leftrightarrow w_k \leftrightarrow v \leftrightarrow u$ that contains edge (u, v) , then we make this edge a *final arc* between nodes u and v , and we incorporate under this new arc the path $w_1 \leftrightarrow w_2 \leftrightarrow \dots \leftrightarrow w_k$, with each edge of the path being a *level- $(i + 1)$ tentative edge* of $\mathcal{R}_{\mathcal{M}}$. When no tentative edges remain—which must happen since each tentative edge is scanned just once before being converted to a final entity (edge or arc), we shall have produced

an express ring $\mathcal{R}_{\mathcal{M}}$ whose nodes and hops are, respectively, the nodes and edges of \mathcal{M} . The reader should be able to use this recipe to construct the the express ring of Fig. 3 from the multi-ring of Fig. 4. ■

3 The Diameter-Structure Tradeoff

As usual, the diameter, $\text{Diameter}(\mathcal{G})$, of a graph \mathcal{G} is the maximum distance between two nodes of \mathcal{G} . For positive integers c, N , we denote by $\text{Diameter}_c(N)$ the minimum diameter of any cutwidth- c , N -node express ring. This section is devoted to proving the following tradeoff between the cutwidth of an express ring (the parameter c), and the optimal diameter achievable for that cutwidth (the quantity $\text{Diameter}_c(N)$).

Theorem 3.1 *For all positive integers c, N ,*

$$\frac{1}{4e}cN^{1/c} - \frac{c}{2} \leq \text{Diameter}_c(N) \leq \left(\frac{1}{2}\right)^{1/c} cN^{1/c}.$$

Remarks.

- The upper and lower bounds in Theorem 3.1 differ by only a constant factor when $c \leq (1/\alpha) \log_2 N$, where α is any constant larger than $(1 + \log_2 e)$. For instance, when $c \leq \frac{1}{6} \log_2 N$, the bounds are within a factor of 12 of each other.
- By invoking the transformation of Theorem 2.1(b), one can infer from Theorem 3.1 a tradeoff between the number of levels of a multi-ring and the optimal diameter achievable for that number of levels.

We establish the upper bound of Theorem 3.1 in Section 3.2 and the lower bound in Section 3.3. Before we do so, however, we should review prior related work, both for the sake of completeness and for basic ideas that we shall use in our constructions and analyses.

3.1 Related Work on Diameter Bounds

The most relevant work we know of, for the study in this section, is the work of [12, 14]. Both of those sources study what can be described briefly as **unidirectional express paths**. In short, both sources start with a path and add hops in order to decrease the maximum distance from the left end of the path to any other node. We find in [12] the following exact, albeit implicit, determination of the quantity³ $\vec{D}_c(N)$, which is the minimum such distance for any cutwidth- c N -node unidirectional express path.

³The superposed arrow in “ \vec{D} ” is intended to stress the unidirectionality of the express-path problem.

Theorem 3.2 [12] *If*

$$\binom{c+d-1}{d} < N \leq \binom{c+d}{d},$$

then $\vec{D}_c(N) = d$.

This result is proved by embedding a carefully crafted unbalanced c -ary tree into the N -node path, with the root of the tree at the left end of the path.

Although the bounds of [14] are less definitive than those of [12], the methodology of [14] is of interest, especially as we contemplate generalizing our work to express rings whose arcs are weighted with a delay function which increases with the length of the arc; cf. Section 4.3. The upper bound in [14] is established via a very uniform recursive construction that is based on the well-known game:

I am thinking of an integer between 1 and N . To win the game, you must determine my integer with the fewest guesses, while never guessing a number that exceeds mine more than c times.

The natural recursive solution has you make the guesses $1, N^{1/c}, 2N^{1/c}, 3N^{1/c}, \dots$ until I tell you that you have exceeded my integer, at which point you have bracketed my integer in an interval of size $N^{1/c}$; you then recurse within this interval with parameter $c-1$ (since you have used one of your allowed big guesses). This technique leads to the recurrence

$$\vec{D}_c(N) \leq \min_x \left(\frac{N}{x} + \vec{D}_{c-1}(x) \right), \quad (1)$$

which yields the upper bound, $\vec{D}_c(N) \leq cN^{1/c}$, of [14]. The lower bound in that paper (which is a factor of c smaller) emerges from noting the following facts, where x denotes the size of the biggest ‘‘hop’’ in the guessing game.

- It takes at least N/x hops to get from the left end of the path to the right end.
- Once having landed at the left end of an interval of length x , using the biggest possible hops, it takes at least $\vec{D}_{c-1}(x)$ hops to get around within the interval. (This requires a bit of argumentation, but it remains valid even if one allows the hops of express paths to cross one another.)

This chain of reasoning leads to the recurrence

$$\vec{D}_c(N) \geq \min_x \left(\max \left\{ \frac{N}{x}, \vec{D}_{c-1}(x) \right\} \right), \quad (2)$$

which yields the lower bound, $\vec{D}_c(N) \geq \frac{1}{2}N^{1/c}$ of [14]. The reader should note that, were we able to replace the maximization in recurrence (2) by a summation, as in recurrence (1), then one would recover the factor of c that separates the upper and lower bounds of [14]. We shall return to this point in Section 4.3.

3.2 The Upper Bound in the Tradeoff

This section is devoted to proving the upper bound of Theorem 3.1, i.e., the inequality

$$\text{Diameter}_c(N) \leq \left(\frac{1}{2}\right)^{1/c} cN^{1/c}.$$

We choose a construction that follows the strategy of [14] rather than that of [12], for two reasons. Most importantly, the former strategy is more easily converted to a *bidirectional* one—and this bidirectionality buys us most of the improvement over prior bounds. Secondly, one would expect that our highly regular construction (and that of [14]) would be easier to implement than would be the construction of [12], at least within an electronic medium.

The basic recurrence. Let us be given an N -node ring \mathcal{R} which we want to augment with hops—up to cutwidth c —in order to obtain the claimed bound. Our augmentation proceeds in two stages.

- We augment \mathcal{R} with uniform upper-level hops, of common length ℓ (to be chosen later).⁴
- We augment the paths that connect the endpoints of each long hop via a cutwidth- $(c-1)$ bidirectional analog of the path-augmentation of [14]. Specifically, we view each inter-hop length- ℓ path as two length- $(\ell/2)$ paths—one from each endpoint—and we recursively augment each path.

We then route each path between node u of \mathcal{R} and node v , via the following 3-segment path.

1. Find a shortest path from node u to the upper-level hop “highway;” the path traversed crosses no more than $\ell/2$ nodes of \mathcal{R} .
2. Find the analogous path for node v .
3. Connect the two low-level paths just constructed via a shortest path of $\leq N/\ell$ upper-level hops.

The first two segments of this path involve a bidirectional solution of the path-augmentation problem, which, recall, seeks shortest paths from the endpoint of the path to an arbitrary other node w . Our bidirectional solution to this problem involves adding upper-level hops of uniform length ℓ *along the half of the path nearest the endpoint*, and recursing within these upper-level hops, down to the appropriate cutwidth. Routings in this solution involve 2-segment paths such as the following.

1. Find a shortest path from node w to the upper-level hop “highway;” the path traversed crosses no more than $\ell/2$ nodes of \mathcal{R} .

⁴As is customary in such constructions, we assume that all necessary divisibilities hold—such as N being divisible by ℓ here. Correcting such assumptions via rounding affects only low-order terms.

2. Connect the path just constructed via a shortest path of $\leq N/\ell$ upper-level hops to the endpoint of the path.

The construction just described gives rise to the following simultaneous recurrences, where $D_c(N)$ denotes the bidirectional analog of the unidirectional diameter-from-endpoint quantity $\bar{D}_c(N)$.

First, for the diameter-from-endpoint problem, our construction yields the recurrence

$$D_c(N) \leq \min \left(\frac{N}{\ell} + D_{c-1}(\ell/2) \right) \quad (3)$$

We already know from [14] that diameter $D_c(N)$ satisfies an inequality of the form

$$D_c(N) \leq \kappa_c N^{1/c},$$

for some parameter κ_c that depends only on c . Using recurrence (3), we determine that we can use a parameter κ_c that satisfies the recurrence

$$\kappa_c = \left(\frac{1}{2} \right)^{1/c} \left(\frac{1}{c-1} \right)^{1-1/c} \kappa_{c-1}^{1-1/c}. \quad (4)$$

Solving this recurrence (via substitution), we find that

$$\kappa_c = 2^{1/c-1} c$$

so that

$$D_c(N) \leq 2^{1/c-1} c N^{1/c}. \quad (5)$$

(Not surprisingly, this improves the unidirectional bounds of [14] by roughly a factor of 2, at least for large c .)

For the express-ring-diameter problem, our construction yields the following bound for $\text{Diameter}_c(N)$.

$$\text{Diameter}_c(N) \leq \min \left(\frac{N}{2\ell} + 2D_{c-1}(\ell/2) \right). \quad (6)$$

Using our bound (5) on $D_c(N)$, we convert inequality (6) to the explicit form

$$\text{Diameter}_c(N) \leq \min \left(\frac{N}{2\ell} + (c-1)\ell^{1/(c-1)} \right). \quad (7)$$

Standard techniques show that the optimal upper-level hops have common length $\ell = (N/2)^{1-1/c}$, so that

$$\text{Diameter}_c(N) \leq 2^{-1/c} c N^{1/c},$$

as claimed. ■

3.3 The Lower Bound in the Tradeoff

This section is devoted to proving the lower bound of Theorem 3.1, i.e., the inequality

$$\text{Diameter}_c(N) \geq \frac{1}{4e}cN^{1/c} - \frac{c}{2}.$$

We begin with the following lemma, which indicates where the elusive factor of c in the lower bound comes from.

Lemma 3.1 *Every cutwidth- c express ring is embeddable in a c -dimensional mesh with dilation 2.*

Proof. We begin by invoking Theorem 2.1(b) to replace the given cutwidth- c express ring \mathcal{R} by a topologically equivalent c -level multi-ring \mathcal{M} . We embed \mathcal{M} into the c -dimensional lattice⁵ with dilation 2, thereafter finding a mesh that encloses the image of \mathcal{M} . Our embedding employs the following strategy which minimizes the interference among the many subrings of \mathcal{M} .

- We number the dimensions of the c -dimensional lattice starting from 1, from left to right. For $i = 1, 2, \dots, c$, we allocate dimension i of the lattice to the level- i subrings of \mathcal{M} .
- We embed \mathcal{M} 's level- i subrings in a sinuous pattern, so that for $i > 1$, the (at most) two level- i rings incident to each level- $(i - 1)$ node of \mathcal{M} proceed in opposite directions (one proceeds to the “left” and the other to the “right”).

We now supply some details of the node-mapping component of the embedding, the edge-routing being specified implicitly.

Embedding \mathcal{M} 's level-1 subring. We use the familiar dilation-2 embedding of a ring in a path. To wit, letting this subring have nodes (in clockwise order) v_0, v_1, \dots, v_{n-1} , we take an n -step “walk” along dimension 1 of the lattice, starting at the origin $\langle 0, \vec{0} \rangle$. During the i th step of the walk, where $0 \leq i < n$, we visit node $\langle i, \vec{0} \rangle$ of the lattice. When i is even, we deposit node $v_{i/2}$ of the ring at this node; when i is odd, we deposit node $v_{n-\lceil i/2 \rceil}$ of the ring at this node; see Fig. 5.

Embedding \mathcal{M} 's level- k subrings, $k > 1$. Let us focus on an arbitrary level $k \geq 2$. Each level- k subring of \mathcal{M} contains two adjacent nodes of one of \mathcal{M} 's level- $(k - 1)$ subrings. We call a level- k subring *tight* if its two level- $(k - 1)$ nodes are adjacent in the c -dimensional lattice and *loose* if these nodes are distance 2 apart in the lattice. (The claimed dilation implies that these two categories of subrings exhaust the possibilities.) We formulate “standard” embeddings of tight and loose level- k subrings, which will then be oriented to create a sinuous pattern, and

⁵As usual, the *c-dimensional lattice* is the infinite analogue of the c -dimensional mesh, with nodes (lattice points) that are c -tuples of (positive and negative) integers and adjacencies between nodes that differ by precisely ± 1 in precisely one coordinate.

stretched when necessary to resolve collisions. In the details that follow, we repress subscripts to enhance clarity.

The “standard” embedding of a tight subring. If a tight subring has (even) length 2ℓ , then its standard embedding has the form

$$\begin{array}{ccccccc} \langle \vec{x}, i, 0, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i, 1, \vec{0} \rangle & \leftrightarrow & \cdots & \leftrightarrow & \langle \vec{x}, i, \ell - 1, \vec{0} \rangle \\ & & & & & & \updownarrow \\ \langle \vec{x}, i + 1, 0, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i + 1, 1, \vec{0} \rangle & \leftrightarrow & \cdots & \leftrightarrow & \langle \vec{x}, i + 1, \ell - 1, \vec{0} \rangle. \end{array}$$

If a tight subring has (odd) length $2\ell + 1$, then its standard embedding has the form

$$\begin{array}{ccccccc} \langle \vec{x}, i, 0, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i, 1, \vec{0} \rangle & \leftrightarrow & \cdots & \leftrightarrow & \langle \vec{x}, i, \ell - 1, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i, \ell, \vec{0} \rangle \\ & & & & & & & & \updownarrow \\ \langle \vec{x}, i + 1, 0, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i + 1, 1, \vec{0} \rangle & \leftrightarrow & \cdots & \leftrightarrow & \langle \vec{x}, i + 1, \ell - 1, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i + 1, \ell, \vec{0} \rangle \end{array}$$

The “standard” embedding of a loose subring. If a loose subring has (even) length 2ℓ , then its standard embedding has the form

$$\begin{array}{ccccccc} \langle \vec{x}, i, 0, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i, 1, \vec{0} \rangle & \leftrightarrow & \cdots & \leftrightarrow & \langle \vec{x}, i, \ell - 1, \vec{0} \rangle \\ & & & & & & \updownarrow \\ \langle \vec{x}, i + 2, 0, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i + 2, 1, \vec{0} \rangle & \leftrightarrow & \cdots & \leftrightarrow & \langle \vec{x}, i + 2, \ell - 1, \vec{0} \rangle. \end{array}$$

If a loose subring has (odd) length $2\ell + 1$, then its standard embedding has the form

$$\begin{array}{ccccccc} \langle \vec{x}, i, 0, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i, 1, \vec{0} \rangle & \leftrightarrow & \cdots & \leftrightarrow & \langle \vec{x}, i, \ell - 1, \vec{0} \rangle \\ & & & & & & \updownarrow \\ & & & & & & \langle \vec{x}, i + 1, \ell - 1, \vec{0} \rangle \\ & & & & & & \updownarrow \\ \langle \vec{x}, i + 2, 0, \vec{0} \rangle & \leftrightarrow & \langle \vec{x}, i + 2, 1, \vec{0} \rangle & \leftrightarrow & \cdots & \leftrightarrow & \langle \vec{x}, i + 2, \ell - 1, \vec{0} \rangle. \end{array}$$

Additionally, each standard embedding has a *positive aspect*, which we have just specified, and a *negative aspect*, which is obtained from the positive aspect by negating the dimension- k component of a lattice point, as in

$$\langle \vec{x}, i, \ell, \vec{0} \rangle \implies \langle \vec{x}, i, -\ell, \vec{0} \rangle.$$

The embedding. We proceed along the image-path of \mathcal{M} 's level- $(k - 1)$ subring. As we encounter each node v along this path, we use a positive-aspect standard embedding for one of node v 's level- k subrings and a negative-aspect standard embedding for node v 's other level- k subring. This alternation of positive and negative aspects creates our sinuous pattern and precludes any collisions between “adjacent” level- k subrings, i.e., subrings that share a node. This completes the embedding for levels $k > 2$. For level 2, however, there may be collisions

between “complementary” level- k subrings, i.e., subrings that are made adjacent in the lattice by our embedding of \mathcal{M} ’s level-1 ring; indeed, one level-2 subring may have collisions with both of its “complementary” ones. It is easy to verify that such collisions can be removed as follows. One starts at the origin of the lattice and proceeds along the path that contains \mathcal{M} ’s level-1 subring. One resolves each encountered collision between “complementary” level-2 subrings, by stretching a subpath of the colliding subring that is farther from the origin: one stretches the offending subpath of this subring by moving a path of the form

$$\langle i, \pm j, \vec{0} \rangle \leftrightarrow \langle i, \pm(j+1), \vec{0} \rangle \leftrightarrow \cdots \leftrightarrow \langle i, \pm \ell, \vec{0} \rangle$$

“one step along,” to become the path

$$\langle i, \pm(j+1), \vec{0} \rangle \leftrightarrow \langle i, \pm(j+2), \vec{0} \rangle \leftrightarrow \cdots \leftrightarrow \langle i, \pm(\ell+1), \vec{0} \rangle.$$

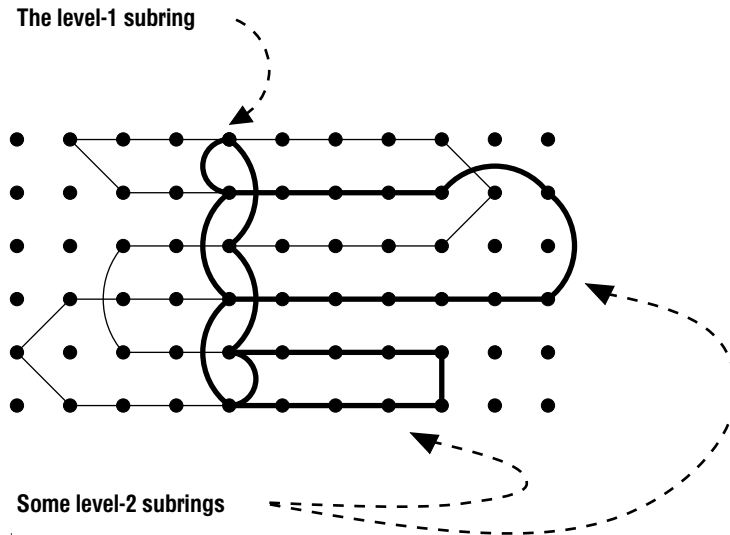


Figure 5: *Exemplifying the dilation-2 embedding of a “typical” 2-level multi-ring. The level-1 subring and two of the level-2 subrings are highlighted.*

Fig. 5 depicts the complete embedding of a “typical” 2-level multi-ring \mathcal{M} into the 2-dimensional lattice.

Summarizing, we embed \mathcal{M} ’s level-1 ring into the c -dimensional lattice with dilation 2. We then embed \mathcal{M} ’s higher-level subrings, level by level, using initial standard embeddings that incur dilation 2, “refined” by collision-resolving stretches that create length-2 subpaths from unit-length ones, hence never increase the dilation beyond 2. We thus have the advertised embedding of \mathcal{M} . ■

Returning to the proof of our lower bound: Lemma 3.1 shows each N -node cutwidth- c express ring to be an N -node connected subgraph of the (graph-theoretic) *square* of the c -dimensional lattice, i.e., the lattice augmented with all length-2 shortcuts. It follows that

the diameter of such an express ring can be no smaller than the smallest diameter of any such subgraph, which is, easily, one-half the smallest diameter, call it $d_c(N)$, of any N -node connected subgraph of the c -dimensional mesh. We can bound $d_c(N)$ from below by noting that it is no smaller than the smallest integer ρ such that

$$N \leq 2^c \binom{\rho + c}{c}. \quad (8)$$

This inequality follows from the fact that, for any r ,

$$\binom{r + c}{c}$$

is the number of lattice points within distance r of the origin in the nonnegative orthant of c -dimensional space—so that 2^c times this quantity is strictly larger than the number of lattice points in the radius- r c -dimensional ball. It follows that the integer ρ is no larger than the radius of the N -node c -dimensional ball. Straightforward reasoning and estimates now convert inequality (8) to the inequality $d_c(N) \geq (1/2e)cN^{1/c} - c$. It follows directly that $\text{Diameter}_c(N) \geq (1/4e)cN^{1/c} - c/2$ as was claimed. ■

4 Extensions

4.1 The Impact of the Noncrossing Requirement

The chordal rings and express rings that we have studied here differ from those often encountered, in our insistence that chords and hops, respectively, not cross. We show now that this structural simplification cannot cost us more than a factor of 2 in diameter. This bound cannot be improved in general, specifically, for rings having $N \geq 4$ nodes, and for infinitely many cutwidth parameters c . To wit, any *unit-diameter* N -node express (resp., chordal) ring is topologically equivalent to an N -node clique (completely connected graph). Since the N -node clique is not outerplanar for any $N \geq 4$ [13], such an augmented ring must have crossing hops (resp., chords).

Theorem 4.1 *Any cutwidth- c N -node express ring \mathcal{R} in which arcs are allowed to cross can be transformed to a crossing-free cutwidth- c N -node express ring \mathcal{R}' such that*

$$\text{Diameter}(\mathcal{R}') \leq 2 \cdot \text{Diameter}(\mathcal{R}).$$

Proof. We produce ring \mathcal{R}' from ring \mathcal{R} in two stages.

First, we replace \mathcal{R} by a cutwidth- c N -node express ring \mathcal{R}'' whose hops form a tree (whose edges may cross). Specifically, we obtain \mathcal{R}'' by choosing an arbitrary node v of \mathcal{R} and growing

a breadth-first spanning tree of hops outward from v . All paths within \mathcal{R}'' constitute a walk up the tree to its root v , then down the tree to the desired destination. Since \mathcal{R}'' is a *shortest-path* spanning tree of \mathcal{R} , it follows that

$$\text{Diameter}(\mathcal{R}'') \leq 2 \cdot \text{Diameter}(\mathcal{R}).$$

Next, we perform a series of transformations of \mathcal{R}'' to eliminate all crossings of hops—without increasing either diameter or cutwidth. We eliminate crossings starting with the edges/hops emanating from the root of the tree and proceeding steadily down toward the leaves. Since each transformation replaces (at least) one offending hop with a shorter hop, no “uncrossing” step can create a new crossing; this fact guarantees that our algorithm terminates. Our “uncrossing” algorithm is the analog for bidirectional trees of the similarly motivated algorithm in [12]—which has many fewer cases since its trees have each nonroot node lying to the right of its parent.

We describe our algorithm by specifying how it “uncrosses” two arbitrary crossing hops (a, c) and (b, d) while never increasing the level of any node, i.e., the node’s distance from the root. With no loss in generality (since the hops cross), assume that nodes $a, b, c,$ and d appear in that order in a clockwise traversal of the ring. Our primary breakdown into cases considers how hops (a, c) and (b, d) are oriented relative to the root of the spanning tree.

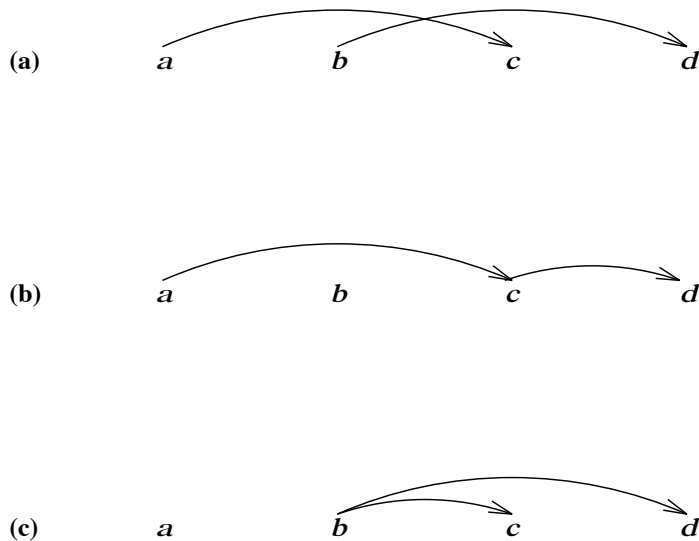


Figure 6: *Resolving a Case-1 crossing.*

Case 1. $\text{level}(a) < \text{level}(c)$; $\text{level}(b) < \text{level}(d)$; see Fig. 6(a). We branch on the relative levels of nodes a and b .

Case 1.a. $\text{level}(a) < \text{level}(b)$. Eliminate hop (b, d) ; add hop (c, d) ; see Fig. 6(b).

Case 1.b. $\text{level}(a) \geq \text{level}(b)$. Eliminate hop (a, c) ; add hop (b, c) ; see Fig. 6(c).

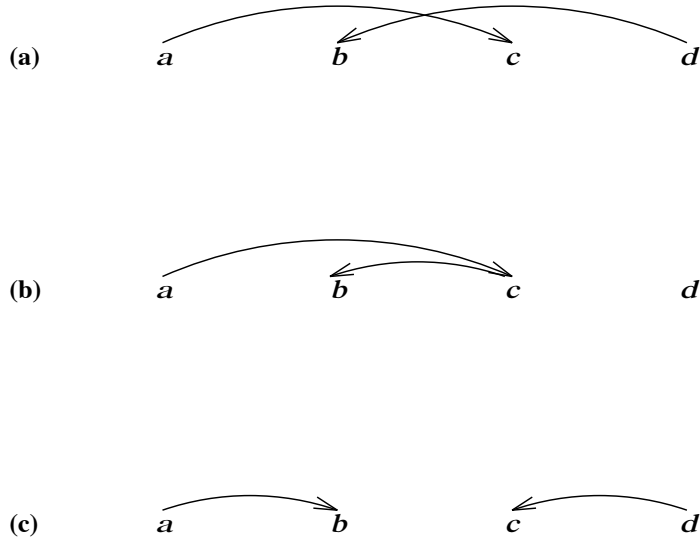


Figure 7: *Resolving a Case-3 crossing.*

Case 2. $\text{level}(a) > \text{level}(c)$; $\text{level}(b) > \text{level}(d)$. This case is clearly symmetric to Case 1.

Case 3. $\text{level}(a) < \text{level}(c)$; $\text{level}(b) > \text{level}(d)$; see Fig. 7(a). We branch on the relative levels of nodes a and d .

Case 3.a. $\text{level}(a) < \text{level}(d)$. Eliminate hop (d, b) ; add hop (c, b) ; see Fig. 7(b).

Case 3.b. $\text{level}(a) > \text{level}(d)$. This case is symmetric to Case 3.a.

Case 3.c. $\text{level}(a) = \text{level}(d)$. Eliminate hops (a, c) and (b, d) ; add hops (a, b) and (d, c) ; see Fig. 7(c).

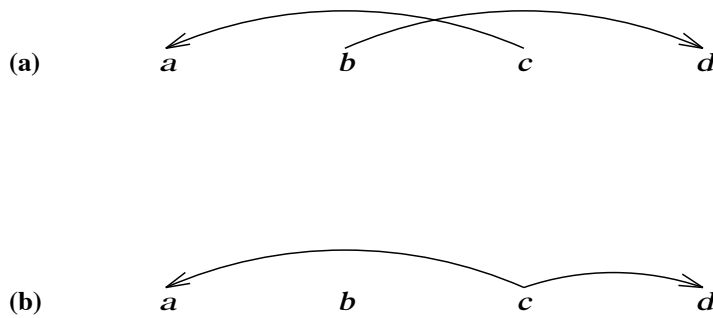


Figure 8: *Resolving a Case-4 crossing.*

Case 4. $\text{level}(a) > \text{level}(c)$; $\text{level}(b) < \text{level}(d)$; see Fig. 8(a). We branch on the relative levels of nodes b and c .

Case 4.a. $\text{level}(b) \leq \text{level}(c)$. Eliminate hop (b, d) ; add hop (c, d) ; see Fig. 8(b).

Case 4.b. $\text{level}(b) \geq \text{level}(c)$. This case is symmetric to Case 4.a.

By the time we reach the leaves of the spanning tree, we shall have transformed the express ring \mathcal{R}'' to the express ring \mathcal{R}' which has no crossing hops and which has diameter no greater than that of \mathcal{R}'' . ■

4.2 Express Meshes

We wish to consider the mesh-based analogue of the problem of Section 3. First, let us choose a reasonable notion of cutwidth for an augmented mesh. We assume that every augmenting edge is routed using a sequence of edges of the underlying mesh. The cutwidth of an augmented mesh is the number of edges that are routed along an edge of the underlying mesh; it is, therefore, just the congestion of the routing that underlies the augmentation. As with express rings, we say by convention that the unaugmented mesh has unit cutwidth.

Let us denote by $\text{Diameter}_c^{(\text{mesh})}(N)$ the analogue of $\text{Diameter}_c(N)$ for N -node square meshes. As the following theorem indicates, the diameter-cutwidth problem degenerates for meshes. The $c = 2$ case of the following theorem appears in [14].

Theorem 4.2 *For all positive integers N and all c ,*

$$\text{Diameter}_c^{(\text{mesh})}(N) = \Theta(\log_c N) = \Theta\left(\frac{\log N}{\log c}\right).$$

Proof Sketch. We refer the reader to [14] for the degree-based argument that yields the lower bound. For the upper bound, we sketch a more direct, hence slightly simplified alternative to the proof in that paper.

Our construction will add “one layer” of augmenting edges to the mesh (so that $c = 2$) in a way that connects all mesh-nodes via a balanced binary tree, thus achieving logarithmic diameter. The reader can easily generalize our construction to general cutwidth c , by constructing a c -ary spanning tree (instead of a binary tree), to obtain diameter $O(\log_c N)$.

Our construction begins by partitioning the N -node mesh \mathcal{M}_N into $N/\log_2^2 N$ submeshes, each of dimensions $\log_2 N \times \log_2 N$.

Next, for each submesh \mathcal{S} , use paths of mesh-edges to connect every node of \mathcal{S} to the node in the “bottom-left corner” of \mathcal{S} (using the standard drawing of \mathcal{M}_N). Each tree so created has depth $2\log_2 N$, hence diameter $\leq 4\log_2 N$.

Finally, use augmenting edges to connect all of the “bottom-left corner” nodes together via a binary tree. First, connect all nodes in the same row using row-trees. Then hook all the row-trees together using a single column-tree. Using the tree-layout of [8], we can route both the row trees and the column tree so that no two augmenting edges share a mesh-edge, ensuring that $c = 2$. ■

4.3 Honoring Physical Limitations

As is pointed out in [2, 17], physical limitations on fan-out and on signal transmission prevent the model that we have studied here from scaling indefinitely to larger architectures. Somewhat mitigating the fan-out problem is the fact that the upper bound of Theorem 3.1 indicates that near optimal diameter can be achieved using cutwidth that is only logarithmic in N . Somewhat mitigating the problem of long edges is the fact that, via Theorem 2.1(b), we can physically realize express rings as (almost-)subgraphs of c -dimensional meshes, which allows one always to achieve hop-lengths that are sublinear in N . Having said this, though, large fan-outs and long hops remain a problem that merits further study.

One possible avenue for increasing the range of scalability of express rings is to use the technique of repeated drivers, described in [16], to obtain logarithmic signal delay along long hops, at least in wafer-scale implementations, in which delays are largely capacitive. Such a delay model will honor physical constraints for larger values of N than will the isometric delay model of Section 3. Within this model, we have the following extension of the upper bound of Theorem 3.1, which is established by techniques similar to those of the Theorem.

Proposition 4.1 *For all c and N ,*

$$\text{Diameter}_c^{(\log)}(N) \leq kcN^{1/c}(\log N)^{1-1/c}.$$

Extending the lower bound of Theorem 3.1 seems to be significantly more challenging. The techniques that yield the optimal bounds of [12] and the strong lower bound of the current paper depend quite strongly on the isometry of all hops in an express ring. Hence, they cease to apply when we weight hops by some function of their lengths. The somewhat weaker techniques of [14], however, are quite robust in this respect. They can, therefore, be used to derive the following lower bound to complement the upper bound of Proposition 4.1.

Proposition 4.2 *For all c and N ,*

$$\text{Diameter}_c^{(\log)}(N) \geq k'N^{1/c}(\log N)^{1-1/c}.$$

While the dominant behavior of N is visible in this lower bound, the bound is off by a factor of c from our upper bound. We believe that the best road to bridging this gap would be to figure out how to replace the maximization in recurrence (2) by a summation (which would match the one in recurrence (1)). This alternative route to the lower bound in Theorem 3.1 should extend to non-unit delay models, such as the logarithmic model.

The avenue for achieving this extendible argument has eluded us thus far.

Acknowledgments. The research of A. Rosenberg was supported in part by NSF Grant CCR-92-21785 and in part by US-Israel BSF Grant 94-00266. The research of R. Sitaraman

was supported in part by NSF Grants CCR-94-10077 and CCR-9703017. A portion of this research was done while F. Chung and A. Rosenberg were visiting Bellcore, and a portion was done while A. Rosenberg was visiting the Dept. of Computer Science at the Technion (Israel Inst. of Technology).

References

- [1] B.W. Arden and H. Lee (1981): Analysis of chordal ring networks. *IEEE Trans. Comput.*, C-30, 291-295.
- [2] G. Bilardi and F.P. Preparata (1995): Horizons of parallel computation. *J. Parallel Distr. Comput.*
- [3] B. Bollobás and F.R.K. Chung (1988): The diameter of a cycle plus a random matching. *SIAM J. Discr. Math.* 1, 328-333.
- [4] H. Burkhardt et al. (1992): Overview of the KSR1 computer system. Tech. Rpt. KSR-TR 9202001, Kendall Square Research.
- [5] F.R.K. Chung and M.R. Garey (1984): Diameter bounds for altered graphs. *J. Graph Th.* 8, 511-534.
- [6] S. Cosares, I. Saniee, O. Wasem (1992): Network planning with the SONET toolkit. *Bellcore Exchange* (Sept./Oct., 1992) 8-15.
- [7] W.J. Dally (1991): Express cubes: improving the performance of k -ary n -cube interconnection networks. *IEEE Trans. Comp.* 40, 1016-1023.
- [8] M.J. Fischer and M.S. Paterson (1980): Optimal tree layout. *12th ACM Symp. on Theory of Computing*, 177-189.
- [9] A. Frank, T. Nishizeki, N. Saito, H. Suzuki, É. Tardos (1992): Algorithms for routing around a rectangle. *Discr. Appl. Math.* 40, 363-378.
- [10] O. Gerstel and S. Zaks (1994): The virtual path layout problem in fast networks. *13th ACM Symp. on Principles of Distributed Computing*, 235-243.
- [11] O. Gerstel and S. Zaks (1995): The virtual path layout problem in ATM ring and mesh networks. *1st Conf. on Structure, Information and Communication Complexity*, Carleton Univ. Press, Ottawa.
- [12] O. Gerstel, A. Wool, S. Zaks (1997): Optimal layouts on a chain ATM network. *Discr. Appl. Math.*, to appear.
- [13] F. Harary (1969): *Graph Theory*. Addison-Wesley, Reading, Mass.

- [14] E. Kranakis, D. Krizanc, A. Pelc (1995): Hop-congestion tradeoffs for high-speed networks. *7th IEEE Symp. on Parallel and Distr. Processing*, 662-668.
- [15] A. Litman and A.L. Rosenberg (1993): Balancing communication in ring-structured networks. Tech. Rpt. 93-80, Univ. Massachusetts.
- [16] C. Mead and L. Conway (1980): *Introduction to VLSI Systems*. Addison-Wesley, Reading, Mass.
- [17] P.M.B. Vitanyi (1988): A modest proposal for communication costs in multicomputers. In *Concurrent Computations: Algorithms, Architecture, and Technology* (S.K. Tewksbury, B.W. Dickinson, S.C. Schwartz, eds.) Plenum Press, N.Y., 203-216.