

Periodically Regular Chordal Rings

Behrooz Parhami, *Fellow, IEEE*, and Ding-Ming Kwai

Abstract—Chordal rings have been proposed in the past as networks that combine the simple routing framework of rings with the lower diameter, wider bisection, and higher resilience of other architectures. Virtually all proposed chordal ring networks are node-symmetric, i.e., all nodes have the same in/out degree and interconnection pattern. Unfortunately, such regular chordal rings are not scalable. In this paper, periodically regular chordal (PRC) ring networks are proposed as a compromise for combining low node degree with small diameter. By varying the PRC ring parameters, one can obtain architectures with significantly different characteristics (e.g., from linear to logarithmic diameter), while maintaining an elegant framework for computation and communication. In particular, a very simple and efficient routing algorithm works for the entire spectrum of PRC rings thus obtained. This flexibility has important implications for key system attributes such as architectural scalability, software portability, and fault tolerance. Our discussion is centered on unidirectional PRC rings with in/out-degree of 2. We explore the basic structure, topological properties, optimization of parameters, VLSI layout, and scalability of such networks, develop packet and wormhole routing algorithms for them, and briefly compare them to competing fixed-degree architectures such as symmetric chordal rings, meshes, tori, and cube-connected cycles.

Index Terms—Chordal rings, fault tolerance, greedy routing, hierarchical parallel architectures, interconnection networks, routing algorithms, skip links.

1 INTRODUCTION

NUMEROUS interconnection networks have been used or proposed for parallel computers and there are widely varying views on the relative merits of different network topologies [23], [25]. The ring interconnection scheme has proven quite effective in certain distributed and small-scale parallel systems [5] in view of its low node degree and simple routing algorithm. However, the diameter of a simple ring would be too large for effective utilization in a massively parallel system. As a result, multilevel and hybrid architectures utilizing ring connections at various levels of a hierarchically structured network or as a basis for synthesizing richer interconnection schemes have been proposed [2], [3], [4], [5], [10], [12], [14], [18], [19], [28], [29], [30], [34].

The multilevel ring structure of KSRI's interconnection network [14] and the QuickRing network [28] are good examples of the hierarchical approach. The chordal ring [2], generalized chordal ring [3], [10], and distributed loop network [5] architectures, in which each node is also connected to one or more distant nodes through *skip links* or *chords*, provide examples of the second approach. Such skip links reduce the network diameter at the expense of increased node degree and wiring complexity. Because the basic ring structure is preserved, many nice features of a simple ring, including ease of routing and deadlock avoidance, carry over to these enhanced ring architectures.

Fig. 1 shows a simple unidirectional ring with eight nodes and a node-symmetric chordal ring with the same number of nodes in which chords, or forward skip links, of length 3 have been added to each node. This chordal ring

network is node-symmetric because each node has the same number of skip links (1) and the same set of *skip distances* (3). More generally, the degree of each node may be greater than two and multiple chords or forward skip links may originate from each node.

Low-diameter node-symmetric chordal rings require high node degrees. If the node degree is fixed at k , then a node-symmetric chordal ring is quite similar to a k -dimensional mesh in terms of both diameter and bisection width [24], [26], [33].

In this paper, we show that by relaxing the symmetry requirement and opting instead for periodic regularity, low node degree and small diameter can be achieved simultaneously. We analyze the resulting networks and show them to possess advantages over meshes, tori, and cube-connected cycles with regard to topological parameters and ease of routing. Even though we derive our networks by combining a ring structure with a mesh-like architecture, the combining process bears no relation to those previously examined [1].

The remainder of this paper is organized as follows. We begin by reviewing node-symmetric chordal rings in Section 2. Periodically regular chordal rings are introduced and analyzed in Section 3, where a semigreedy routing algorithm is also presented and shown to be quite efficient. We discuss the problem of optimally selecting the PRC ring network parameters in Section 4. Section 5 deals with VLSI layout and scalability issues. Sections 6 and 7 discuss fault-tolerant and wormhole routing algorithms, respectively. Section 8 contains our conclusions.

2 NODE-SYMMETRIC CHORDAL RINGS

The discussion of node-symmetric chordal rings in this section draws heavily from the notation and results of [12].

Definition 1. (chordal ring) Consider an N -node ring with node labels $0, 1, \dots, N - 1$, and ring connection going from

• The authors are with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560. E-mail: parhami@ece.ucsb.edu.

Manuscript received 15 Nov. 1996; revised 1 Sept. 1998.
For information on obtaining reprints of this article, please send e-mail to: tpd@computer.org, and reference IEEECS Log Number 100341.

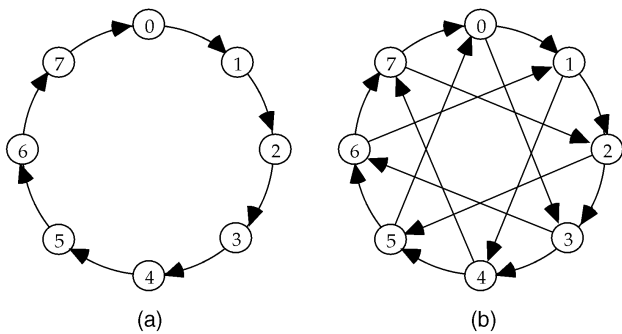


Fig. 1. Two types of previously proposed ring networks: (a) Simple 8-node unidirectional ring, and (b) example of 8-node chordal ring with chords, or forward skip links, of length 3 (skip distance = 3).

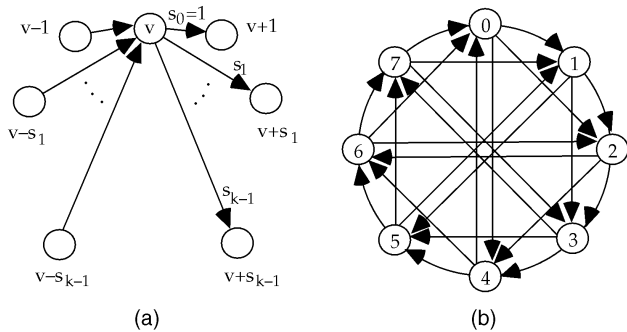


Fig. 2. Defining and exemplifying node-symmetric chordal rings. (a) Input/output connections of node v in a degree- k chordal ring, and (b) 8-node degree-3 chordal ring with skip distances $s_1 = 2$ and $s_2 = 4$.

each node v to node $v + 1(\text{mod } N)$. Let there also be unidirectional skip links or chords from each node v to nodes $v + s_1, v + s_2, \dots, v + s_{k-1}(\text{mod } N)$, with

$$1 < s_1 < s_2 < \dots < s_{k-1} < N.$$

Such an augmented ring is known as a degree- k (symmetric, unidirectional) chordal ring with the skip set $\{s_1, s_2, \dots, s_{k-1}\}$. In the remainder of this paper, it will be understood that all arithmetic in node-index expressions is modulo N . For convenience, we define $s_0 = 1$ and $s_k = N$. Hence, node v is connected to nodes $v + s_j$ for $0 \leq j \leq k$, where $1 = s_0 < s_1 < s_2 < \dots < s_{k-1} < s_k = N$. Fig. 2 shows the neighbors of a given node v in the general case along with an example for $k = 3$.

A shortest path leading from node u to node w consists of a number of skips of each type. Because of node-symmetry, the required skips of each type can be traversed in any order, leading to many distinct paths. Let $d(u, w)$ be the distance from node u to node w along a shortest path and $n_h(u, w)$, $0 \leq h \leq k - 1$, be the number of skip links of type s_h included in the shortest path. Then:

$$d(u, w) = n_0(u, w) + n_1(u, w) + \dots + n_{k-1}(u, w).$$

Given the set of skip distances $\{s_h | 0 \leq h \leq k - 1\}$, the problem of finding a shortest-path from node u to node w requires the precomputation of a size- N table in each node specifying the skip link to be taken for each possible

destination node or, equivalently, for each ring distance from the current node.

Example 1. With $k = 3$ and skips $s_1 = 10$ and $s_2 = 16$, shortest paths for $w - u = 32, 33$, or 34 start with s_2 , whereas, for $w - u = 24, 25, 30$, or 31 , s_1 should be taken first and, for $w - u = 26, 27, 28$, or 29 , either s_1 or s_2 will do as the first step.

In most practical cases, however, a greedy algorithm (that selects the largest skip not overshooting the destination node) performs quite well and leads to near-optimal and, under some conditions, to optimal paths. Examples where simple greedy routing is optimal include the special case where each skip distance s_{h+1} is a multiple of the next lower skip distance s_h , $0 \leq h \leq k - 1$, which in turn covers power-of-2 network size and skip distances that are most likely to be used in practice.

The greedy routing algorithm can be easily described from the viewpoint of an arbitrary intermediate node v , initially the source u and eventually the destination w , and executed in a distributed manner (see Algorithm 1 below). When the node degree k is large, the routing algorithm would become more efficient if the variable h is made part of the packet header and decremented by nodes each time the index offset $w - v$ is smaller than the skip s_h .

Algorithm 1: Greedy packet routing on a node-symmetric chordal ring.

```

Node  $v = ig + j (j < g)$ , upon producing or receiving
a packet destined for node  $w$ , does the following
if  $w = v$ 
then remove the packet; stop
else
    find the unique  $h$  such that  $s_h \leq w - v < s_{h+1}$ ;
    send the packet to node  $v + s_h$ 
endif
    
```

The inequality $n_h(u, w) \leq \lceil s_{h+1}/s_h \rceil - 1$ is clearly satisfied when routing is done by Algorithm 1. Hence, the routing distance $d_{Alg1}(u, w)$ from node u to node w satisfies:

$$d_{Alg1}(u, w) \leq s_1/s_0 - 1 + \lceil s_2/s_1 \rceil - 1 + \dots + \lceil s_k/s_{k-1} \rceil - 1 < \left(\sum_{h=0}^{k-1} s_{h+1}/s_h \right) - 1 = E.$$

To minimize the worst-case bound for $d_{Alg1}(u, w)$, the right-hand-side expression E must be minimized. Equating

$$\partial E / \partial s_h = 1/s_{h-1} - s_{h+1}/s_h^2$$

with 0, we obtain the optimal skips

$$s_h^{opt} = N^{h/k}$$

and the optimal worst-case routing distance or routing diameter

$$D_{Alg1}^{opt} \approx kN^{1/k},$$

which is basically that of a k -dimensional mesh with unidirectional and wrap-around links. Such meshes are sometimes referred to as Manhattan street networks [7], [20]. This is not surprising in view of the fact that an s_k -node

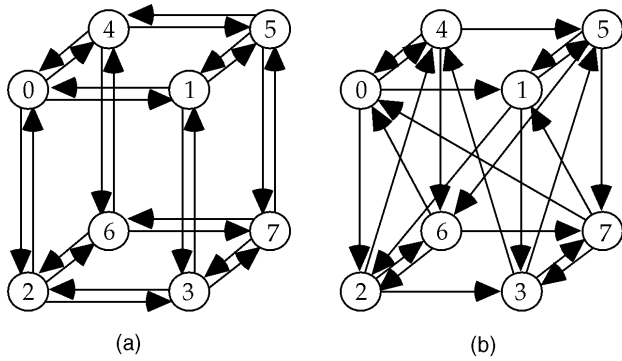


Fig. 3. (a) Three-dimensional $2 \times 2 \times 2$ mesh/torus, and (b) the chordal ring network of Fig. 2b redrawn to expose its mesh-like structure.

chordal ring with skips s_1, s_2, \dots, s_{k-1} can be redrawn to resemble an $s_1 \times (s_2/s_1) \times \dots \times (s_k/s_{k-1})$ mesh.

Example 2. A 64-node symmetric chordal ring with skip distance $s_1 = 8$, when laid out on an 8×8 grid with nodes appearing in row-major order, yields a torus-like structure with column wrap-around links to the same column and row wrap-around links to the next row (the ILLIAC IV topology [25]).

Example 3. The 8-node symmetric chordal ring of Fig. 2b has been redrawn in Fig. 3b to show its relationship with the $2 \times 2 \times 2$ mesh/torus of Fig. 3a. Note that all eight links along one of the dimensions, plus four links in each of the other two dimensions, are identical in Figs. 3a and 3b. The remaining links in Fig. 3b have been redirected.

The case where each $s_h + 1$ is divisible by s_h , $0 \leq h \leq k - 1$, merits special attention. In this case, greedy routing does in fact lead to a shortest path. From the bound previously derived for $d_{Alg1}(u, w)$, we see that in this special case, the diameter of the chordal ring network is upper bounded by $\sum_{h=0}^{k-1} s_{h+1}/s_h - k$. Theorem 1 shows that this bound is tight.

Theorem 1. The diameter D of an N -node symmetric chordal ring network with skip distances s_1, s_2, \dots, s_{k-1} , such that s_{h+1} is divisible by s_h , $0 \leq h \leq k - 1$, is exactly equal to the bound $\sum_{h=0}^{k-1} s_{h+1}/s_h - k$.

Proof. The bound for $d_{Alg1}(u, w)$ shows that, when each s_{h+1} is divisible by s_h , any node can be reached in at most $\sum_{h=0}^{k-1} s_{h+1}/s_h - k$ steps using the greedy routing algorithm. Hence, $D \leq \sum_{h=0}^{k-1} s_{h+1}/s_h - k$. The proof is complete upon noting that the distance from node 0 to node $N - 1$ is exactly $\sum_{h=0}^{k-1} s_{h+1}/s_h - k$. \square

When $N^{1/k}$ is an integer and s_h , $1 \leq h \leq k - 1$, is optimally chosen (as discussed earlier) to be $N^{h/k}$, the exact diameter of the node-symmetric chordal ring becomes:

$$D = k(N^{1/k} - 1)$$

Example 4. With $N = 125$ and $k = 3$, the optimal skip distances are $\{5, 25\}$ and the chordal ring diameter is easily verified to be $D = 3(125^{1/3} - 1) = 12$.

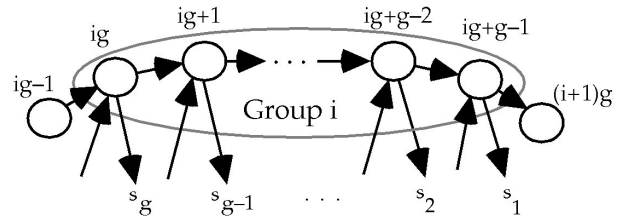


Fig. 4. Nodes within the i th g -node group in a PRC ring and their associated input and output skip links.

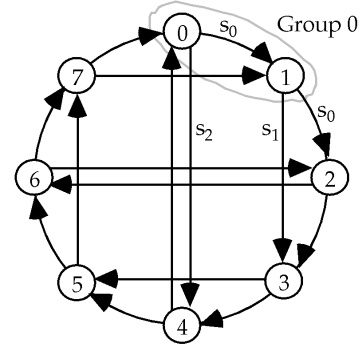


Fig. 5. PRC ring with $N = 8$, group length $g = 2$, and skip distances $s_1 = 2$ and $s_2 = 4$.

Example 5. With $N = 1,024$ and $k = 5$, the set of skip distances is optimally chosen to be $\{4, 16, 64, 256\}$, leading to the network diameter

$$D = 5(1,024^{1/5} - 1) = 15.$$

3 PERIODICALLY REGULAR SKIPS

Node-symmetric chordal ring networks seem wasteful in that long-distance, medium-distance, and short-distance links are provided for every node. In a manner similar to deriving the cube-connected cycles (CCC) architecture [28] from the hypercube, one can distribute the various skips among a sequence of nodes, each having at most one skip link. The N nodes are thus split into N/g groups of g consecutive nodes, where g divides N .

Definition 2. (periodically regular chordal ring) In a PRC ring, the N nodes are divided into N/g groups of g nodes each, where the group length, or period, g divides N , and there are g skip distances $s_1 < s_2 < \dots < s_g$. For notational convenience, we define $s_0 = 1$ and $s_{g+1} = N$. Each node v is connected to node $v + 1$ via the basic unidirectional ring connection. The node $v = ig + j$ ($j < g$), or the j th node in the i th group, is also connected via a skip link with skip distance s_{g-j} to node $v + s_{g-j}$. In order to assure that node in-degree/out-degree is uniformly equal to 2, we require that all skip distances s_h , $1 \leq h \leq g$, be multiples of g so that each node is guaranteed to be the destination of one, and only one, skip link. The detailed structure of the i th g -node group in the resulting PRC ring is depicted in Fig. 4. Fig. 5 shows an 8-node PRC ring with $g = 2$, $s_1 = 2$, and $s_2 = 4$.

A PRC ring can be viewed as a subgraph of a node-symmetric chordal ring. For example, Fig. 5 depicts an 8-node PRC ring that can be derived by pruning some of the links in the symmetric chordal ring network of Fig. 2b.

As in Section 2, let $n_h(u, w)$, $0 \leq h \leq g$, be the number of skip links of type s_h included in a shortest path of length $d(u, w)$ from node u to node w . Whereas in the case of node-symmetric chordal rings $d(u, w)$ depends only on $w - u$, here it truly depends on both u and w . In other words, the relative positions of source and destination nodes in their respective groups also affect the length of the shortest path between them.

Example 6. In the PRC ring of Fig. 5, the shortest path from node 0 to node 2 is of length 2 but node 1 is directly connected to node 3. In both cases $w - u = 2$.

Before deriving the diameter of such networks, it is helpful to discuss a semigreedy routing algorithm that performs quite well in cases where nodes u and w are not very close to each other. The routing algorithm is based on taking the skip links in the order $s_g, s_{g-1}, \dots, s_2, s_1$. This is similar to *dimension-order* routing in certain other networks. Since node u in general does not have the first (longest) required skip link, the packet is first routed to a node that does. This is the only deviation from the “greedy” strategy.

Algorithm 2: *Semigreedy routing on a PRC ring, with skips used in descending order.*

Node $v = ig + j(j < g)$, upon producing or receiving a packet that is headed for the destination node w , does the following

```

if  $w = v$ 
then remove the packet; stop
else
  if  $s_{g-j} \leq w - v < g - 1 + s_{g-j+1}$ 
  then send the packet to node  $v + s_{g-j}$ 
  else send the packet to node  $v + 1$ 
  endif
endif
```

Note that Algorithm 2 is not a pure greedy algorithm in that it does not choose among the two locally available links the one that would take the packet closer to the destination. The $s_{g-j} \leq w - v$ part of the condition for the second “if” statement is the greedy part which eventually prevails. Initially, however, the $w - v < g - 1 + s_{g-j+1}$ part may force routing of the packet to node $v + 1$ even when node $v + s_{g-j}$ is closer to the destination. In effect, the algorithm compares the current skip s_{g-j} against the next longer skip s_{g-j+1} , which is $g - 1$ steps downstream, to determine which one would take the packet closer to its destination.

It is easy to see that Algorithm 2 frequently routes packets via nonoptimal paths.

Example 7. Consider a PRC ring with group size $g = 2$ and skips $s_1 = 10$ (for odd-numbered nodes) and $s_2 = 16$ (for even-numbered nodes). To route from node $u = 2i$ to node $w = 2i + 21$, Algorithm 2 uses the path

$$2i \xrightarrow{16} 2i + 16 \xrightarrow{1} 2i + 17 \xrightarrow{1} 2i + 18 \xrightarrow{1} 2i + 19 \xrightarrow{1} 2i + 20 \xrightarrow{1} 2i + 21,$$

whereas the optimal (shortest) path is:

$$2i \xrightarrow{1} 2i + 1 \xrightarrow{10} 2i + 11 \xrightarrow{10} 2i + 21.$$

The routing path may be nonoptimal even if each skip distance s_{h+1} is a multiple of the next shorter one s_h .

The worst-case number of routing steps required by Algorithm 2 is

$$D_{Alg1} = 2g - 1 + \sum_{h=0}^g x_h,$$

where x_h is the number of times that an s_h -type skip is taken to route the packet to the destination node w from the node $g[u/g]$; viz the closest node to the source node u that has the longest, or type- s_g , skip link. The added term $2g - 1$ results from the worst case where $g - 1$ steps are needed to route from node u to node $g[u/g]$, plus g steps to move from one skip distance to the next lower one; viz, s_g to s_{g-1}, \dots, s_2 to s_1, s_1 to s_0 .

As in the analysis of Algorithm 1, the inequality

$$x_h \leq \lceil s_{h+1}/s_h \rceil - 1$$

holds for all h ; actually, here we can prove the slightly improved bound

$$x_h \leq \lceil (s_{h+1} - 1)/s_h \rceil - 1$$

in view of the extra s_0 step taken between successive skips of different lengths (e.g., if we could not take the skip $s_{h+1} = 25$, we should not be able to take four skips of length $s_h = 6$ after we have stepped forward on an s_0 link), but we will use the first bound for simplicity. For x_0 , we can derive a tighter bound. The preceding argument suggests that:

$$x_0 \leq s_1/s_0 - 1 = s_1 - 1.$$

We observe that $x_0 \leq s_1 - 2$, since if $x_0 = s_1 - 1$, the s_0 steps and the transition step from s_1 just before the s_0 steps could be combined into a single s_1 step. That is, the sequence of forwarding steps

$$\begin{array}{ccccccccccc} s_1 & s_1 & \dots & s_1 & s_0 & s_0 & s_0 & \dots & s_0 & & \\ \hline & & & & & & & & & x_0 = s_1 - 1 & \text{steps} \end{array}$$

can be replaced with $x_1 + 1$ steps of s_1 . Using the tighter bound for x_0 given above, the worst case routing distance of Algorithm 2 is upper-bounded by

$$D_{Alg2} = s_1/s_0 + \lceil s_2/s_1 \rceil + \dots + \lceil s_{g+1}/s_g \rceil + g - 3 < 2g - 3 + \sum_{h=0}^g s_{h+1}/s_h,$$

where the inequality is obtained by observing that each of the g terms $\lceil s_{h+1}/s_h \rceil$ is strictly less than $s_{h+1}/s_h + 1$.

It is worth noting that the condition used in Algorithm 2 to decide between taking the skip or the local outgoing link essentially involves comparing the destination node index

TABLE 1
Minimum Diameter D_{opt} and the Associated Average Internode Distance $\Delta_{\text{opt},D}$ for PRC Rings with Different Power-of-2 Sizes N and Group Lengths g

N	g	D_{opt}	$\Delta_{\text{opt},D}$	$s_h (1 \leq h \leq g)$
8	2	3	2.0	2, 4
16	2	4	2.7	4, 6
32	2	6	3.6	6, 14
	4	6	3.6	4, 8, 12, 16
64	2	8	4.9	6, 20
	4	7	4.4	8, 20, 24, 28
128	2	10	6.4	40, 60
	4	8	5.4	16, 28, 36, 60
	8	10	6.1	8, 16, 24, 32, 40, 48, 56, 64
256	2	14	8.2	106, 116
	4	10	6.5	32, 68, 76, 116
	8	11	6.9	24, 32, 56, 72, 80, 104, 112, 120
512	4	12	7.7	36, 76, 168, 200
	8	12	7.8	8, 56, 88, 152, 160, 184, 200, 224
	2	22	13.5	252, 458
1024	4	14	9.0	212, 320, 344, 436
	8	13	8.7	208, 216, 264, 344, 376, 400, 464, 504

ease of algorithm development will likely make it desirable to “round” the preceding skips to, say, 16 and 4.

A corollary of Theorem 2 is that if the skip ratio s_{h+1}/s_h is the same for $0 \leq h \leq g$ and there are a logarithmic number of skips, i.e., $g = O(\log N)$, then the PRC ring will have logarithmic diameter. This is further clarified in the next section.

4 OPTIMAL PERIOD AND SKIPS

The diameter $D = \sum_{h=0}^g s_{h+1}/s_h - 3$ of a PRC ring given by Theorem 2 is minimized for

$$\partial D / \partial s_h = 1/s_{h-1} - s_{h+1}/s_h^2 = 0.$$

This leads to the optimal skips

$$s_h^{\text{opt}}(g) = N^{h/(g+1)}$$

for any given group size g . Since, in general, $N^{1/(g+1)}$ is not an integer, the rest of this analysis is approximate and is only meant to provide insight into how logarithmic diameter can be attained. Substituting $s_{h+1}/s_h = N^{1/(g+1)}$ into the expression for D leads us to:

$$D_{\text{opt}}(g) = (g + 1)N^{1/(g+1)} - 3.$$

We next equate

$$\partial D_{\text{opt}}(g) / \partial g = N^{1/(g+1)} [1 - (\ln N) / (g + 1)]$$

with 0 to find:

$$g_{\text{opt},D} = \ln N - 1.$$

Finally, substituting $g_{\text{opt},D}$ in the expression for $D_{\text{opt}}(g)$, we get:

$$D_{\text{opt}} = e \ln N - 3.$$

The preceding was an approximate analysis beginning with the assumption that each skip distance s_{h+1} is divisible by the next lower skip s_h . Given the number N of nodes and the group length g , the optimal skip distances s_h^{opt} yielding minimal diameter for the PRC ring do not in general satisfy the preceding property. This is also true for node-symmetric chordal rings for which the determination of minimal diameter, along with the corresponding skip distances, remains an open problem [5].

The results of an exhaustive search through all possible group sizes and skip distances (Table 1) indicate that both the diameter and average distance can be minimized by choosing a particular group length whose value depends on N . Furthermore, we see from Table 1 that the diameter and average distance do not change significantly when the group length g is varied around the optimal value. Some of the skip distances in Table 1 are quite counterintuitive and show the limitations of the preceding approximate analysis. For example, the optimal skip distances $s_1 = 252$ and $s_2 = 458$ with $N = 1,024$ and $g = 2$ are nowhere close to $1,024^{1/3} \approx 10$ and $1,024^{2/3} \approx 102$.

Fig. 6 shows the effect of choosing various skip distances s_1 and s_2 on the PRC ring diameter when $N = 64$ and $g = 2$. Node-symmetric chordal rings are represented by the horizontally placed diagonal ($s_1 = s_2$), while the left half of the same diagonal ($s_1 = s_2 < 0$) represents the class of

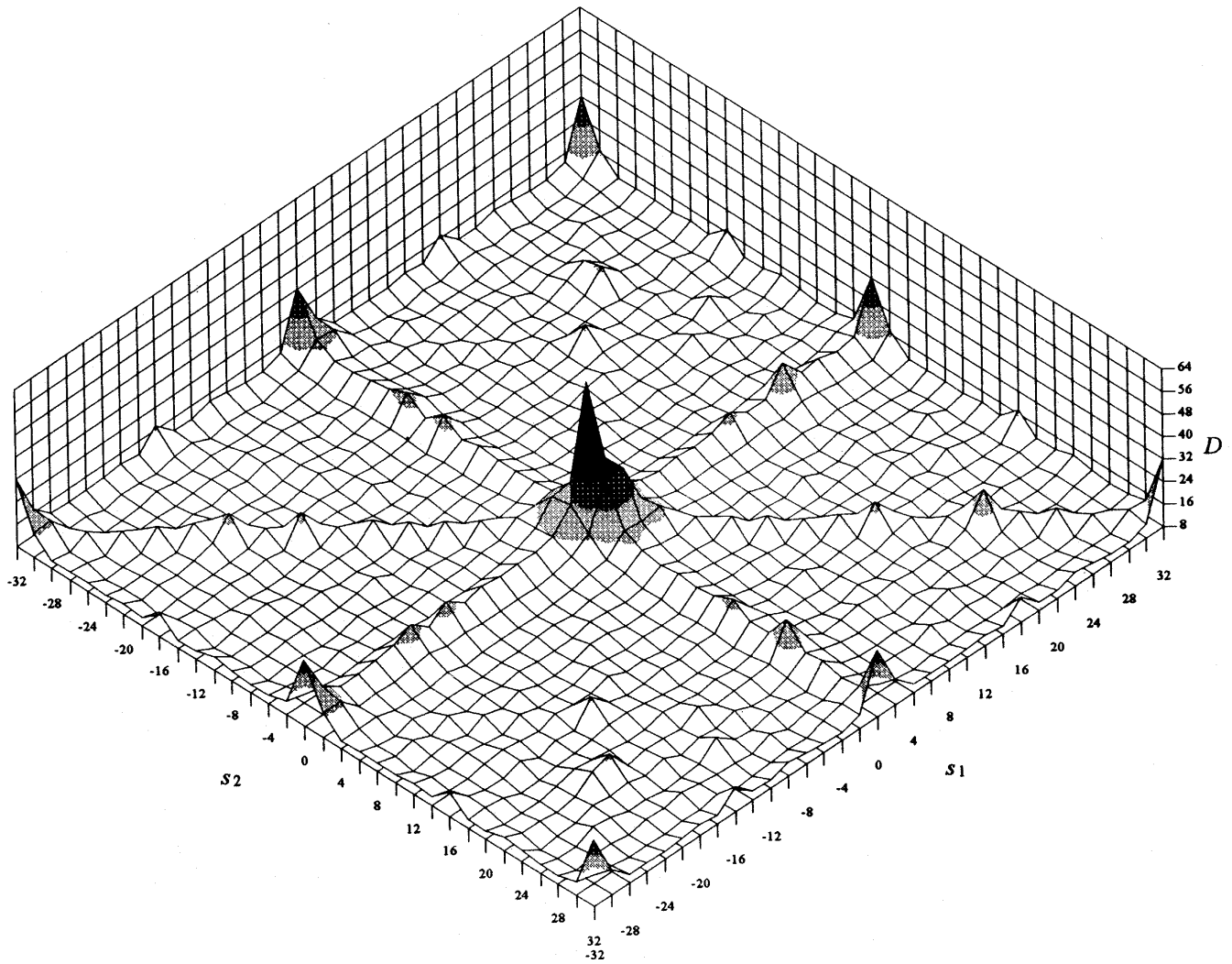


Fig. 6. Variations in the diameter of a 64-node PRC ring with group length $g = 2$ and various choices of skip distances s_1 and s_2 . Negative, or backward, skips are equivalent to forward skips with distances greater than $N/2$.

forward-loop backward-hop networks [13], [29]. Our selection of skip distances satisfying $s_1 < s_2$, with both skips nonzero, falls below the diagonal, where the diameter does not vary significantly. Hence, other criteria, such as weighted average internode distance for commonly used communication patterns and effects of scaling to expand the network size, may be more important in selecting the best values for the group length and skip distances.

Let us continue by selecting the network size and skip distances to be powers of 2 and consider running semigroup (also known as global reduction, fan-in, or tree) computations on such PRC rings. A semigroup computation on $N = 2^n$ data items is defined as the application of an associative binary operator \otimes to all the data items; viz computing the expression $x_0 \otimes x_1 \otimes \dots \otimes x_{N-1}$. In what follows, we assume that the binary operator \otimes is also commutative, so the data items and intermediate results can be paired or combined in any order. This will make our algorithm more efficient (as explained later in Example 9) and is not a significant restriction in practice. Addition, OR,

AND, maximum, minimum, and many other binary operators of practical interest are in fact commutative.

Algorithm 4: *Semigroup computation, for a commutative operator, on a PRC ring with $N = 2^n$ nodes and power-of-2 group length and skip distances.*

Conceptually, operations are performed in n iterations or rounds where, at the start of iteration i , $0 \leq i \leq n-1$, there are 2^{n-i} data items left to be combined. Thus, we begin with 2^n data items, distributed one per node, in iteration 0 and terminate by combining 2 data items in iteration $n-1$.

To implement iterations 0 through $\log_2 g - 1$ that reduce the number of data items by a factor of g , we let each group shift $g-1$ data items to the first node of the next group which combines them with its own data into a single value. At this point, we have $2^n/g$ data items all of which are stored at nodes having skip links of length s_g .

The s_g skip links are then used to send the data items forward by successive distances $2^{n-1}, 2^{n-2}, \dots, s_g$, where

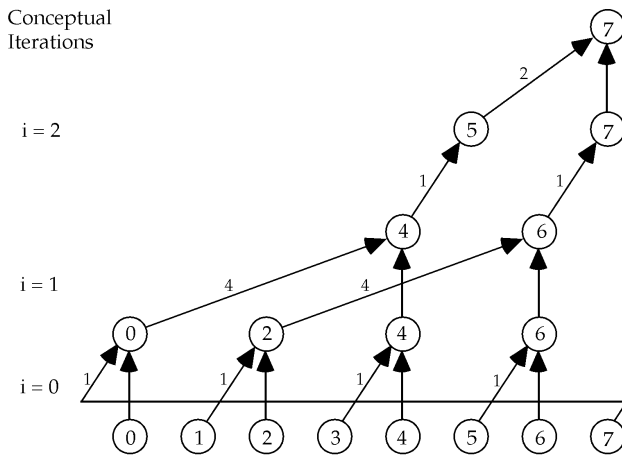


Fig. 7. Semigroup computation performed on 8-node PRC ring with group length $g = 2$ and skip distances $s_1 = 2, s_2 = 4$. Arrows are labeled with skip distances.

upon arrival at each destination, they are combined with the existing values. These $n - \log_2 s_g$ iterations, which successively compress the data into the last $1/2, 1/4, \dots, 1/(2^n/s_g)$ of the ring, require

$$(2^{n-1} + 2^{n-2} + \dots + s_g)/s_g = 2^n/s_g - 1$$

routing steps.

Finally, we need a scheme to combine the remaining s_g/g data items into a single value. Since skips of length s_g can no longer be used, we shift the remaining data items forward by one node to allow the use of the next shorter skip distance s_{g-1} . The process continues until the iteration variable i reaches $n - \log_2 s_1 + \log_2 g$, at which point the data items reside at the last nodes of s_1/g consecutive groups. The computation can then be completed via $s_1 - g$ routing steps using the skips $s_0 = 1$ to shift all remaining values to the last of these s_1/g nodes.

Based on the analysis presented within Algorithm 4, the total number of routing steps needed for performing the semigroup computation is

$$\begin{aligned} T_{\text{semigroup}} &= g - 1 + [(s_{g+1}/s_g - 1) + 1] + \dots \\ &\quad + [(s_3/s_2 - 1) + 1] + (s_2/s_1 - 1) + s_1 - g \\ &= \sum_{h=0}^g s_{h+1}/s_h - 2, \end{aligned}$$

which is one step more than the diameter of a PRC ring, as given by Theorem 2.

Example 9. Fig. 7 shows the semigroup computation performed according to the preceding algorithm on the 8-node PRC ring of Fig. 5. The algorithm executes in three conceptual iterations or rounds, requiring 1, 1, and 2 routing steps, respectively. The interleaving of values and branches of the reduction tree shown in Fig. 7 is the reason why our semigroup computation algorithm needs commutativity in addition to the standard associativity requirement for the semigroup binary operator.

Theorem 3. The minimum diameter of a 2^n -node PRC ring with group length g and power-of-2 skip distances s_1, s_2, \dots, s_g , is $2n + g - 2 \log_2 g - 3$, where $g + \log_2 g \leq n \leq 2g + \log_2 g$.

Proof. Given the group length g , there are $n - \log_2 g$ permissible powers of 2 from which the g skip distances $s_h, 1 \leq h \leq g$, can be selected. These are $2^{n-1}, 2^{n-2}, \dots, 2^{\log_2 g} = g$. Thus, we need to have:

$$n - \log_2 g \geq g.$$

It is then easy to see that the group length g satisfies:

$$g + \log_2 g = n - \varepsilon \text{ where } 0 \leq \varepsilon \leq g.$$

Let $2^{m_h} = s_{h+1}/s_h, 0 \leq h \leq g$, be the h th skip ratio. Recall that $s_{g+1} = 2^n$ and $s_0 = 1$. We then have:

$$m_0 = \log_2 g + \Delta \text{ where } \Delta \geq 0$$

$$\sum_{h=0}^g m_h = n$$

$$D = \sum_{h=0}^g 2^{m_h} - 3 \text{ (by Theorem 2).}$$

Since $m_0 \leq n - g = \log_2 g + \varepsilon$, we have $0 \leq \delta \leq \varepsilon$. Select the remaining m_h values, $1 \leq h \leq g$, such that $m_h = 2$ in $\varepsilon - \delta$ cases and $m_h = 1$ in the remaining $g - \varepsilon + \delta$ cases. It is easily verified that:

$$n = \sum_{h=0}^g m_h = g + \log_2 g + \varepsilon$$

$$D = \sum_{h=0}^g 2^{m_h} - 3 = (2^\delta g - 2\delta) + 2g + 2\varepsilon - 3.$$

The diameter is minimized for $\delta = 0$, leading to:

$$D = 3g + 2\varepsilon - 3 = 2n + g - 2 \log_2 g - 3.$$

The proof is complete upon showing that reduction of the group length to $g/2$ does not improve the diameter relative to the above. As shown earlier, a $2^{n-g/2-1}$ -node PRC ring with group length $g/2$ has the minimized diameter:

$$2(n - g/2 - 1) + g/2 - 2(\log_2 g - 1) - 3.$$

The factor of $2^{g/2+1}$ increase in the number of nodes leads to all $g/2 + 1$ skip ratios 2^{m_h} being multiplied by 2 (the m_h values being incremented by 1). The diameter then becomes

$$\begin{aligned} D' &= 2[2(n - g/2 - 1) + g/2 - 2(\log_2 g - 1)] - 3 \\ &= 4n - g - 4 \log_2 g - 3 = D + 2\varepsilon, \end{aligned}$$

which is no less than that obtained for group length g . \square

Table 2 illustrates the result of Theorem 3 by listing the minimum diameters of 2^n -node PRC rings, for n in the range $3 \leq n \leq 14$, together with the worst-case routing distance of Algorithm 2. The set of skip distances thus obtained is not unique, as elements of the skip-ratio power set $\{m_h | 0 \leq h \leq g\}$, given in the rightmost column of Table 2, can be permuted as long as $m_0 > \log_2 g$.

TABLE 2
Minimum Diameter D_{opt} and Worst-Case Routing Distance D_{Alg2} for PRC Rings with Optimal Group Length g ,
Along with the Skip-Ratio Power Set $\{m_h | 0 \leq h \leq g\}$ to Achieve D_{opt} , Where $s_{h+1}/s_h = 2^{m_h}$

$N = 2^n$	g	D_{opt}	D_{Alg2}	$\{m_h 0 \leq h \leq g\}$
8	2	3	4	{1, 1, 1}
16	2	5	6	{2, 1, 1}
32	2	7	8	{2, 2, 1}
64	2	9	10	{2, 2, 2}
	4	9	11	{2, 1, 1, 1, 1}
128	4	11	13	{2, 2, 1, 1, 1}
256	4	13	15	{2, 2, 2, 1, 1}
512	4	15	17	{2, 2, 2, 2, 1}
1024	4	17	19	{2, 2, 2, 2, 2}
2048	4	21	23	{3, 2, 2, 2, 2}
	8	21	27	{3, 1, 1, 1, 1, 1, 1, 1}
4096	8	23	29	{3, 2, 1, 1, 1, 1, 1, 1}
8192	8	25	31	{3, 2, 2, 1, 1, 1, 1, 1}
16384	8	27	33	{3, 2, 2, 2, 1, 1, 1, 1}

Example 10. For $n = 5$ and $g = 2$, the skip sets $\{2, 8\}$, $\{4, 8\}$, and $\{4, 16\}$, corresponding to the skip-ratio power sets $\{1, 2, 2\}$, $\{2, 1, 2\}$, and $\{2, 2, 1\}$, respectively, all lead to the same diameter $D = 7$. We will exploit this flexibility in choosing the optimal skip distances to obtain efficient VLSI layouts for PRC rings.

The proof of Theorem 3 implies that when $n = g + \log_2 g$, the group lengths g and $g/2$ are equally good as they both lead to the diameter

$$D = 3g - 3.$$

Since the diameter is the same in both cases, the smaller group length $g/2$ should be selected in view of the fact that it reduces the worst-case routing distance of Algorithm 2 which, as shown in Section 3, equals the diameter plus the group length.

It is instructive to compare the diameter $3g - 3$, derived above, with the diameter $2.5g - 2$ of a CCC network of the same size [21]. The slightly larger diameter of the PRC ring is due to the unidirectional links and lower node degree (CCC has node in/out-degree of 3). The PRC ring diameter can be reduced to $2g$ by using bidirectional ring connections, while still maintaining the unidirectional skip links (node in/out-degree becomes 3).

5 VLSI LAYOUT AND SCALABILITY

In this section, we present a VLSI layout scheme for PRC rings using the standard grid model in which nodes are placed at grid points and are connected by links that are routed through evenly spaced horizontal and vertical grid lines on two separate wiring planes. The layout area is obtained as the product of the numbers of horizontal and vertical grid lines which contain a node or link segment of the network.

The complexity of a realizable network is limited by its wire density. The bisection width of a network, defined as the minimum number of links that must be cut in order to divide it into two equal halves, is a good predictor of wire density.

Theorem 4. *The bisection width B of a PRC ring with group length g and skip distances s_1, s_2, \dots, s_g is at most $2 + 2 \sum_{h=1}^g s_h/g$.*

Proof. Consider a cut as in the example depicted in Fig. 8 and focus on the top part of the PRC ring. There are at most s_h/g links of type s_h from one side of the cut to the other side. Summing the resulting terms, adding 1 for the s_0 link, and doubling to account for the opposite side of the ring lead to the claimed upper bound for B . \square

Corollary. *With group length g and power-of-2 skip distances s_1, s_2, \dots, s_g , the bisection width B of the PRC ring cannot exceed $4s_g/g$.*

Proof. From $s_g \geq 2s_{g-1} \geq \dots \geq 2^{g-1}s_1$, we conclude that $s_h \leq s_g/2^{g-h}$. Substituting this upper bound for s_h in the upper bound for B given by Theorem 4, we get:

$$B \leq 2 + 2(2 - 1/2^{g-1})s_g/g.$$

The preceding inequality, combined with $s_g/g \leq 2^{g-1}$, yields $B \leq 4s_g/g$. \square

Example 11. Fig. 8 depicts a PRC ring with $N = 2^n = 16$, $g = 2$, and the skip set $\{2, 4\}$. By Theorem 4, the bisection width of this network is no larger than $2 + 2(2 + 1) = 8$ corresponding to the number of links cut by the double dotted line. The corollary to Theorem 4 also yields the same bound $4 \times 4/2 = 8$.

To lay out a PRC ring, we arrange the nodes in a snake-like fashion, with nodes in each group aligned vertically and nodes in each subring connected by s_h skip links

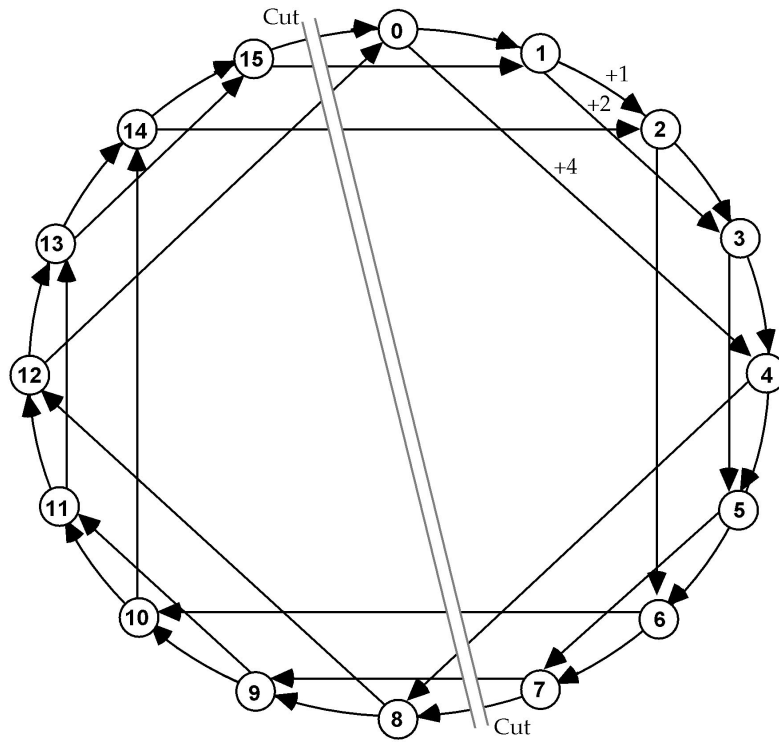


Fig. 8. PRC ring with $N = 16$, $g = 2$, $s_1 = 2$, $s_2 = 4$.

aligned in one row. Thus, the N nodes are placed in N/g columns and $2 \sum_{h=1}^g s_h/g$ rows. The vertical grid lines are assigned to the basic ring connections and the horizontal grid lines are used for skip links. The long wrap-around connections can be avoided, and the wires in each subring balanced, by applying the standard technique of folding [17], which is applicable in both the vertical and horizontal directions. Fig. 9 shows the above procedure applied to the layout of the 8-node PRC ring of Fig. 5.

Based on the preceding layout scheme, we show that the required layout area can be made to be a linear function of N and that the maximum wire length depends on the group length g and the largest skip distance s_g .

Theorem 5. *An N -node PRC ring with group length g and skip distances s_1, s_2, \dots, s_g can be laid out in $O(Ns_g/g^2)$ area with the longest wire being of length $O(s_g/g)$.*

Proof. The preceding layout scheme places the nodes in N/g columns and $2 \sum_{h=1}^g s_h/g$ rows. The layout area for the PRC ring is thus:

$$A = (N/g + 2)(4 + 2 \sum_{h=1}^g s_h/g).$$

Using the result of the corollary to Theorem 4, we find:

$$A \leq (N/g + 2)(4s_g/g + 2) = O(Ns_g/g^2).$$

The longest wire for our layout is in the smallest subring, where N/s_g nodes are connected by wires of total length $2(N/g + 2)$. The length of the longest wire is thus $O(s_g/g)$. \square

One can construct a large PRC ring by concatenating smaller ones with the same group length and skips.

Example 12. The layout of the 16-node PRC ring of Fig. 8 is depicted in Fig. 10, which consists of two copies of Fig. 9, placed side by side and merged.

Note that the construction demonstrated in Example 12 (Fig. 10) can be repeated three, four, five, or more times, each time increasing the layout area by the same amount. A consequence of this observation is that PRC rings are readily scalable, with the network size growing gradually and capable of assuming values that are not powers of 2. The price paid for this modularity and scalability properties is that the network diameter will in general not be optimal for its size.

Whereas an optimized network with small diameter and large bisection width is desirable, its cost/performance ratio suffers from dramatically increased area and wire length [6], [32]. For optimal PRC rings, this leads to the layout area $O(N^2/\log^2 N)$ and the length of the longest wire $O(N/\log N)$, which are of the same orders of magnitude as the respective requirements of CCC networks [28]. As we fix the period and the skips, the diameter will grow, while the layout area decreases and the longest wire length remains constant. The exact result depends on how the parameters are selected to strike a balance between cost and performance.

Example 13. Table 3 lists several parameters of 2^n -node PRC rings for $6 \leq n \leq 14$ when the group length and skip set are fixed at $g = 4$ and $\{4, 16, 64, 256\}$, respectively. The actual diameter D and average internode distance Δ for the resulting PRC rings are compared against the optimal diameter D_{opt} and worst-case and average-case routing

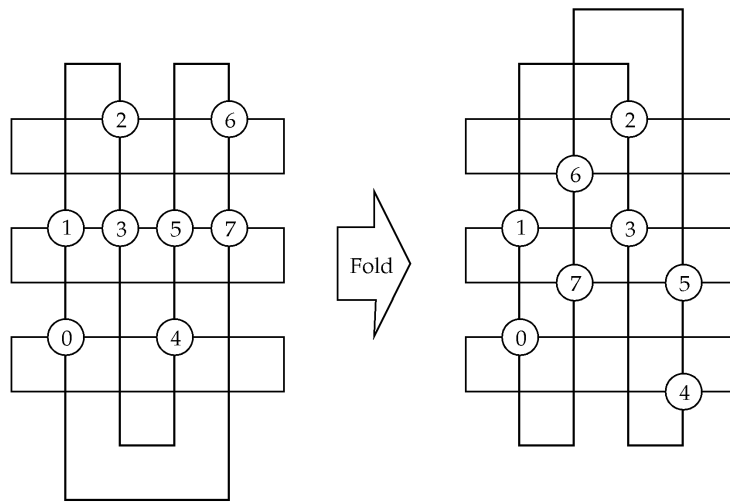


Fig. 9. VLSI layouts for 8-node PRC ring with $g = 2$, $s_1 = 2$, and $s_2 = 4$. Link directions are not shown.

distances when routing is done by Algorithm 2 (D_{Alg2} and Δ_{Alg2}). The D and Δ parameters for 2D torus, 2D mesh, and hypercube of the same sizes are also included for comparison. The group length $g = 4$ and skip set $\{4, 16, 64, 256\}$ have been chosen to allow expandability within a limited range while keeping the diameter close to the minimum possible. We note that with this strategy, both the diameter and average internode distance seem to fall between those of 2D torus networks and optimal PRC rings. Table 3 also indicates that Algorithm 2 performs quite well and is very close to optimal in terms of both the worst-case and average routing distance.

The performance of a parallel architecture depends not only on the internal network structure but also on its input/output bandwidth. Referring to Fig. 10, we note that I/O ports can be conveniently inserted on the ring, at the boundaries between adjacent groups, through the provision of two-way multiplexers at the upper and lower parts of the

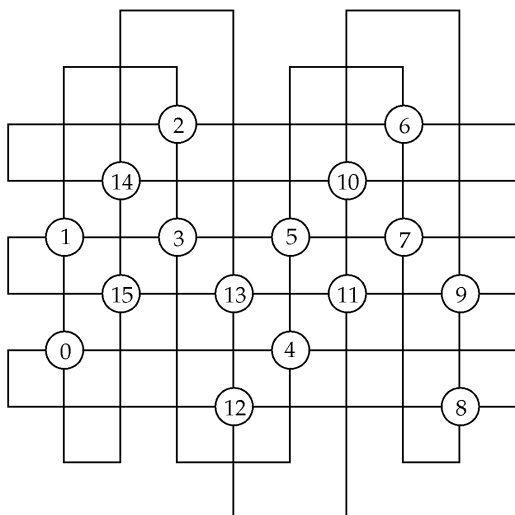


Fig. 10. VLSI layout of a folded 16-node PRC ring with $g = 2$, $s_1 = 2$, and $s_2 = 4$.

diagram. The resulting $O(N/g)$ external bandwidth is adequate to make the I/O time comparable to, or lower than, the running time of any algorithm. A similar scheme has been suggested for the CCC network [28].

6 FAULT-TOLERANT ROUTING

Like symmetric chordal rings which can be viewed as fault-tolerant loops, PRC rings can tolerate certain node and link failures without jeopardizing the connectivity of the remaining fault-free nodes. Link failures are tolerable due to the existence of multiple paths from a source to a destination. Node failures are tolerable due to the bypassing capability resulting from the presence of skip links; albeit, some healthy nodes may have to be bypassed along with the faulty ones, thus reducing the size of the surviving ring. We say that a set of node and link faults is tolerable if a working ring, that is not too small, can be salvaged from the faulty PRC ring. In this section, we present some results that bear on the extent of fault tolerance offered by PRC rings.

Consider an interval $[v + 1, v + f]$ of faulty or otherwise unusable nodes, preceded and followed by the healthy nodes v and $v + f + 1$. To maintain the ring connectivity by bypassing these f nodes, there must exist a skip link from node u to node w such that $u \leq v$ and $w \geq v + f + 1$. Since the longest skip distance is s_g , the number of tolerable consecutive node faults is $s_g - g$ in the worst case (if the first fault is in node $v = ig$, then $g - 1$ healthy nodes preceding it must also be bypassed if an s_g skip is to be used). Hence, $s_g - g$ is an upper bound on the number of faults that can be tolerated.

Example 14. The PRC ring of Fig. 8 can tolerate up to two consecutive node failures, since $s_g - g = 4 - 2 = 2$. Such faulty pairs reduce the ring size by $s_g - 1 = 4 - 1 = 3$ nodes. Any sequence of three faulty nodes that begins at an even-numbered node is intolerable.

Random patterns of faults will in general require multiple intervals of nodes to be bypassed in order to establish the ring connectivity.

TABLE 3
Comparing the Diameter D and Average Internode Distance Δ for Different Network Topologies

$N=2^n$	Periodically Regular Chordal Ring					2D Torus		2D Mesh		Hypercube	
	D_{opt}	D	Δ	D_{Alg2}	Δ_{Alg2}	D	Δ	D	Δ	D	Δ
64	9	11	6.1	12	6.3	8	4.1	14	5.3	6	3.0
128	11	12	6.9	13	7.3	12	6.1	22	8.0	7	3.5
256	13	14	8.1	15	8.6	16	8.0	30	10.7	8	4.0
512	15	15	8.8	17	9.7	24	12.0	46	16.0	9	4.5
1024	17	17	10.0	19	11.0	32	16.0	62	21.3	10	5.0
2048	21	21	12.0	23	13.2	48	24.0	94	32.0	11	5.5
4096	23	29	16.1	31	17.3	64	32.0	126	42.7	12	6.0
8192	25	45	24.1	47	25.3	96	48.0	190	64.0	13	6.5
16384	27	77	40.1	79	41.3	128	64.0	254	85.3	14	7.0

Theorem 6. A PRC ring with group length g and skip distances s_1, s_2, \dots, s_g is guaranteed to maintain the ring connectivity for a (possibly small) subset of its nodes under f random faults only if the number of faults is strictly less than $\sum_{h=1}^g s_h/g$.

Proof. Since $s_g \leq N/2$, a surviving ring must contain nodes at both sides of any bisection and should thus have two links crossing any bisection, one in each direction. From the proof of Theorem 4, we know that there exist bisections that cut $1 + \sum_{h=1}^g s_h/g$ links in each direction. Thus, given that two of these links emanate from the same node, in the worst case, we cannot lose more than $\sum_{h=1}^g s_h/g - 1$ nodes and still maintain the ring connectivity. Note that a skip link can become unusable by the failure of its source or destination node. \square

Example 15. The 8-node PRC ring of Fig. 5 can recover from any $4/2 + 2/2 - 1 = 2$ node failures. With three node failures, however, the existence of an embedded ring, even a very small one, cannot be guaranteed. If, for instance, nodes 0, 1, and 2 fail, the ring connectivity is lost.

Careful examination of Example 15 reveals that node 3 cannot receive, and node 7 cannot send, a message without using nodes 0-2. These nodes will thus become dangling once nodes 0-2 are removed. Removal of these unusable nodes [35] creates other unusable nodes, until all healthy nodes for this example become unusable. Since every usable node has at least one forward and one backward neighbor, an embedded ring can be constructed beginning with any usable node u . In each step, the ring is extended by using the ring link if possible and the skip link if not. This strategy ensures that we obtain the largest possible embedded ring. The size of such a surviving ring can be determined by placing an upper bound on the number of nodes that may become unusable in the worst case.

Theorem 7. An N -node PRC ring with group length g and power-of-2 skip distances s_1, s_2, \dots, s_g can maintain the ring connectivity for at least $Ng/(2s_g - 2g)$ nodes in the presence of any tolerable fault pattern.

Proof. Using a skip link to connect a node v to its successor in the embedded ring implies that v has no usable ring

neighbor. For each such skip link used, at most $\sum_{h=1}^g s_h/g - 2$ nodes can be removed. Denoting the number of nodes in the embedded ring by l , and the number of s_h -type links included in it by x_h , we have:

$$l = \sum_{h=0}^g x_h \geq N - \left(\sum_{h=1}^g s_h/g - 2 \right) \sum_{h=0}^g x_h.$$

The above leads to

$$l \geq N / \left(\sum_{h=1}^g s_h/g - 1 \right) = Ng / (2s_g - 2g),$$

where, in the last step of the derivation, we have made use of the corollary to Theorem 4; viz, $2 + 2 \sum_{h=1}^g s_h/g \leq 4s_g/g$. \square

The bound given by Theorem 7 is quite pessimistic and in most cases we can embed larger rings into a faulty PRC ring. However, this is the best result we have been able to derive for arbitrary worst-case fault patterns.

Example 16. The 8-node PRC ring of Fig. 5 can embed a ring with at least $Ng/(2s_g - 2g) = 8 \times 2/(8 - 4) = 4$ nodes in the presence of two faulty nodes. One can easily verify that when nodes 0 and 3 are faulty, making nodes 4 and 5 unusable, the remaining nodes 1, 2, 6, and 7 can form a 4-node ring.

Algorithm 2 can be used for routing on faulty PRC rings after a slight modification. Suppose each node has a status bit for the availability of its outgoing skip link. Then, the availability of s_{g-j} is added to the condition for sending the packet to node $v + s_{g-j}$. This automatically forces the packet to move forward to the next node that has an s_{g-j} skip or to use a number of s_{g-j-1} hops instead. The simplicity of fault-tolerant routing by a modified version of Algorithm 2 is a direct result of its distributed decision-making structure. Algorithm 3 does not have a correspondingly simple and efficient fault-tolerant version.

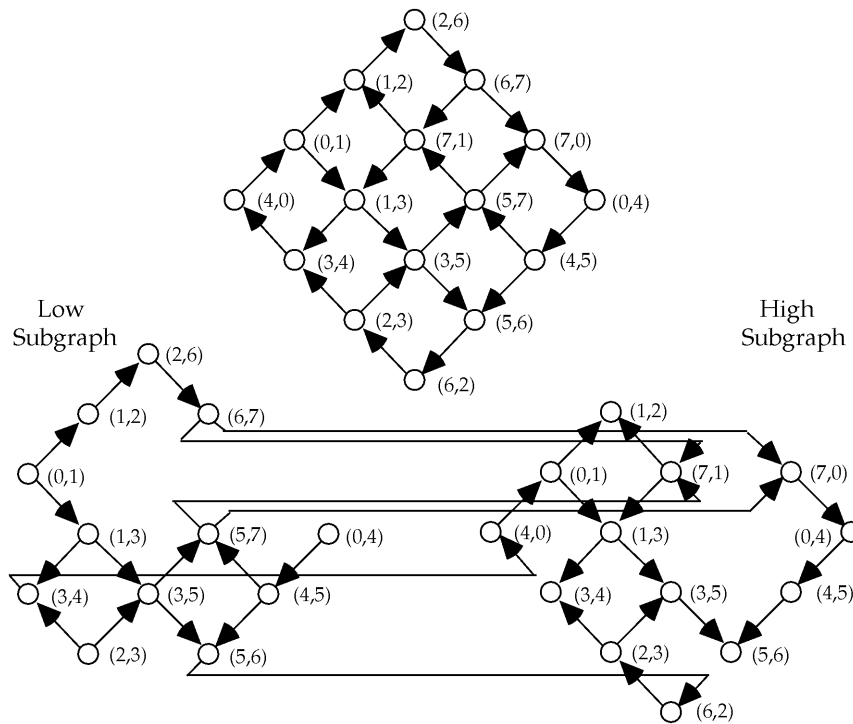


Fig. 11. Dependency graph for an 8-node PRC ring with $g = 2$, $s_1 = 2$, and $s_2 = 4$ (top) along with its cycle-free version with two logical channels (bottom).

7 WORMHOLE ROUTING

With wormhole routing, the message routing path is established by advancing the head of a packet, sending the rest of the packet in a pipelined fashion, and releasing the links after the tail of the packet has passed through [22]. A packet is decomposed into flow control digits, or flits, that are spread out over contiguous links on the routing path and prevent other packets from using those links. When the header flit is blocked, all of the flits of the packet stop advancing. In such a case, the packets involved must wait for the required links to be released and deadlock may occur.

The blockages are due to the fact that the routing paths of multiple packets may overlap. One way to deal with this problem is to share the physical links through multiplexing at different time slots to separate the routing paths temporally; this leads to the use of multiple logical links (virtual channels), each with its own flit buffer, for each physical link. Such a scheme is viable provided that the bandwidth of a physical link can support the average data rate needed by these logical links. For efficient utilization, the number of logical links sharing the same physical link should be minimized.

Consider wormhole routing on a simple ring. Ring networks are prone to deadlocks in view of possible circular waiting among links. The dependency graph of a simple ring, i.e., a graph whose nodes represent the edges $(v, v + 1)$ of the original ring and whose edges go from node $(v, v + 1)$ to node $(v + 1, v + 2)$ to indicate that the latter edge can be used after the former one by the routing algorithm, is also a ring. Since lack of a cycle in the dependency graph is sufficient for deadlock-free routing [9], [11], using two

logical channels per link can solve the deadlock problem. A packet is routed on the "low" logical channels until it needs to go from a node v to a node with a smaller number, at which time it switches to the "high" logical channels and stays there until it gets to the destination. Since the transition from a node v to a node with a smaller number can occur at most once, no circular waiting is possible.

Deadlock-free wormhole routing on PRC rings essentially follows the same strategy, using two logical channels per links. This low complexity is one of the main advantages of PRC rings relative to competing networks and is due to the underlying ring structure. Each link $C(u, v)$, from node u to node v , in the PRC ring is split into a low logical link $C_0(u, v)$ and a high logical link $C_1(u, v)$. By following the strategy outlined above for simple rings, i.e., by starting with the low logical channels and moving to the high channels upon transition to a successor node with a smaller number than the previous node, the dependency graph is transformed into a cycle-free graph that guarantees deadlock freedom.

Example 17. The dependency graph for the PRC ring of Fig. 5, when route selection is done according to Algorithm 2, is shown at the top of Fig. 11. Note that even though it is possible to use the link $C(2, 3)$ after $C(1, 2)$, Algorithm 2 never uses this combination; thus, there is no arrow from node $(1, 2)$ to node $(2, 3)$ in the dependency graph. Use of low and high logical channels, as outlined above, transforms the dependency graph into a cycle-free version also depicted in Fig. 11.

Algorithm 5 is the deadlock-free wormhole routing counterpart to Algorithm 2. A header flit received over

the logical channel p ($p = 0$ for low, $p = 1$ for high), is sent out over the logical channel p if the next node ($v + s_{g-j}$ or $v + 1$, depending on the destination) has a larger number and over the logical channel 1 otherwise. Thus, the output logical channel is determined by ORing p with the result of comparing v with $v + s_{g-j}$ or $v + 1$.

Algorithm 5: *Deadlock-free wormhole routing on a PRC ring, with skips used in descending order.*

Node $v = ig + j$ ($j < g$), upon producing or receiving a header flit from the logical link $C_p(u, v)$ for a packet that is headed for the destination node w , does the following if $w = v$

then remove the header and successive flits; stop else

```

if  $s_{g-j} \leq w - v < g - 1 + s_{g-j+1}$ 
then  $q := p \vee (v > v + s_{g-j})$ ; route along  $C_q(v, v + s_{g-j})$ 
else  $q := p \vee (v > v + 1)$ ; route along  $C_q(v, v + 1)$ 
endif

```

endif

We have limited our discussion of packet and wormhole routing on PRC rings to point-to-point or unicast communication. Multicasting, including its special case of broadcasting, is easily handled with packet switching. Wormhole-routed multicasting algorithms for PRC rings can also be developed with moderate effort using a variety of schemes for representing the destination node set and for performing the message distribution (tree-based or path-based algorithms). Details are omitted here for brevity [27].

8 CONCLUSION

We have introduced periodically regular chordal (PRC) rings as an attractive alternative for realizing scalable parallel architectures. These networks combine the benefits of low node degree, small diameter, and a very simple routing framework. Our discussion centered on basic network structure, topological properties, routing algorithms, optimization of parameters, and comparison to meshes, tori, and cube-connected cycles.

We showed that PRC rings have regular and modular VLSI layouts. They also provide smaller diameter and average distance than similar-sized 2D meshes, 2D tori, and CCC networks, support simpler routing algorithms, and are more easily adapted to fault tolerance in routing and parallel computations. The latter property is due to the fact that the basic greedy routing algorithm continues to work as long as link and node faults are detectable, whereas there is no correspondingly simple fault-tolerant routing scheme for the other networks listed above.

In our discussions thus far, we have assumed that each node has exactly one skip link. This assumption can be relaxed, leading to both richer and sparser interconnections without unduly complicating the routing algorithm. For example in Figs. 9 and 10, removing the skip links corresponding to $s_2 = 4$ will cause significant vertical compression of the layout, leading to lower cost and correspondingly larger diameter and lower performance. This flexibility offers a mechanism for cost/performance trade-offs [24] similar to what can be accomplished for CCC networks [28].

The preceding, and other extensions to the basic architecture, form fruitful areas for further research. Work is also required on the development of parallel algorithms. With suitable parametrization, a single algorithm can run on a large class of PRC rings, spanning a wide range in terms of cost and performance. This software portability among systems of varying capabilities is of utmost importance.

For ease of reference, we conclude by listing key open problems, some of which were alluded to earlier in the paper:

- Evaluating the performance of PRC rings in realistic application contexts; in particular, quantifying the difference between Algorithms 2 and 3.
- Deriving general closed-form expressions for the diameter and average inter-node distance of node-symmetric chordal rings and PRC rings.
- Determining the implications of changing the constraints $s_1 < s_2 < \dots < s_g$ on the skip set to $s_1 \leq s_2 \leq \dots \leq s_g$, possibly also mixing real and null skips.
- Studying PRC rings with bidirectional links and/or multiple skips per node.
- Obtaining tighter bounds for the capacity of PRC rings to tolerate faults and looking at the related issues of fault detection and diagnosis.

We are currently studying these problems and will report some results in the near future.

REFERENCES

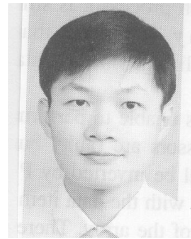
- [1] D.P. Agrawal, C. Chen, and J.R. Burke, "A Comparison of Hybrid Graph Based Topologies," *IEEE Computer Architecture Technical Committee Newsletter*, pp. 5-8, Winter 1994-1995.
- [2] B.W. Arden and H. Lee, "Analysis of Chordal Ring Networks," *IEEE Trans. Computers*, vol. 30, no. 4, pp. 291-295, Apr. 1981.
- [3] B.W. Arden and K.-T. Tang, "Routing for Generalized Chordal Rings," *Proc. ACM 18th Computer Science Conf.*, pp. 271-275, Feb. 1990.
- [4] P. Banerjee, "The Cubical Ring Connected Cycles: A Fault-Tolerant Parallel Computation Network," *IEEE Trans. Computers*, vol. 37, no. 5, pp. 632-636, May 1990.
- [5] J.-C. Bermond, F. Comellas, and D.F. Du, "Distributed Loop Computer Networks: A Survey," *J. Parallel and Distributed Computing*, vol. 24, no. 1, pp. 2-10, Jan. 1995.
- [6] G. Bilardi and F.P. Preparata, "Horizons of Parallel Computation," *J. Parallel and Distributed Computing*, vol. 27, no. 2, pp. 172-182, June 1995.
- [7] T.Y. Chung, N. Sharma, and D.P. Agrawal, "Cost-Performance Tradeoffs in Manhattan Street Networks Versus 2-D Torus," *IEEE Trans. Computers*, vol. 43, no. 2, pp. 240-243, Feb. 1994.
- [8] W.J. Dally, "Virtual-Channel Flow Control," *IEEE Trans. Parallel and Distributed Systems*, vol. 3, no. 2, pp. 194-205, Mar. 1992.
- [9] W.J. Dally and C.L. Seitz, "Deadlock-Free Message Routing in Multiprocessor Interconnection Networks," *IEEE Trans. Computers*, vol. 36, no. 5, pp. 547-553, May 1987.
- [10] K.W. Doty, "New Designs for Dense Processor Interconnection Networks," *IEEE Trans. Computers*, vol. 33, no. 5, pp. 447-450, May 1984.
- [11] J. Duato, "A Necessary and Sufficient Condition for Deadlock-Free Adaptive Routing in Wormhole Networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 6, no. 10, pp. 1,055-1,067, Oct. 1995.
- [12] L.M. Huisman and S. Kundu, "Highly Reliable Symmetric Networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 5, no. 1, pp. 94-97, Jan. 1994.
- [13] F.K. Hwang, "Comments on Reliable Loop Topologies for Large Local Computer Networks," *IEEE Trans. Computers*, vol. 36, no. 3, pp. 383-384, Mar. 1987.

- [14] Kendall Square Research, "KSR1 Technology Background," 1992.
- [15] D.-M. Kwai and B. Parhami, "Periodically Regular Chordal Rings as Fault-Tolerant Loops," *Proc. Pacific Rim Int'l Symp. Fault-Tolerant Systems*, pp. 210-215, Dec. 1995.
- [16] D.-M. Kwai and B. Parhami, "Comparing the Performance Parameters of Two Network Structures for Scalable Massively Parallel Processors," *Proc. Int'l Workshop Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, pp. 273-277, Feb. 1996.
- [17] D.-M. Kwai and B. Parhami, "Periodically Regular Chordal Rings: Generality, Scalability, and VLSI Layout," *Proc. Eighth IEEE Symp. Parallel and Distributed Processing*, pp. 148-151, Oct. 1996.
- [18] S. Latifi, M.M. Azevedo, and N. Bagherzadeh, "The Star Connected Cycles: A Fixed-Degree Network for Parallel Processing," *Proc. Int'l Conf. Parallel Processing*, vol. I, pp. 91-95, 1993.
- [19] I.Y. Lee and S.-D. Wang, "Ring-Connected Networks and Their Relationship to Cubical Ring Connected Cycles and Dynamic Redundancy Networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 6, no. 9, pp. 988-996, Sept. 1995.
- [20] N.F. Maxemchuk, "Routing in the Manhattan Street Network," *IEEE Trans. Comm.*, vol. 35, no. 5, pp. 503-512, May 1985.
- [21] D.S. Meliksetian and C.Y.R. Chen, "Optimal Routing Algorithm and the Diameter of the Cube-Connected Cycles," *IEEE Trans. Parallel and Distributed Systems*, vol. 4, no. 10, pp. 1,172-1,178, Oct. 1993.
- [22] L.M. Ni and P.K. McKinley, "A Survey of Wormhole Routing Techniques in Direct Networks," *Computer*, vol. 26, no. 2, pp. 62-76, Feb. 1993.
- [23] L.M. Ni and D.K. Panda, "A Report of the ICPP'94 Panel on Sea of Interconnection Networks: What's Your Choice," *IEEE Computer Architecture Technical Committee Newsletter*, pp. 31-44, Winter 1994-1995.
- [24] B. Parhami, "Periodically Regular Chordal Ring Networks for Massively Parallel Architectures," *Proc. Fifth Symp. Frontiers Massively Parallel Computation*, pp. 315-322, Feb. 1995.
- [25] B. Parhami, *Introduction to Parallel Processing: Algorithms and Architectures*. Plenum Press, 1999.
- [26] B. Parhami and D.-M. Kwai, "A Characterization of Symmetric Chordal Rings Using Redundant Number Representations," *Proc. Int'l Conf. Systems Eng.*, pp. 467-472, July 1996.
- [27] B. Parhami and D.-M. Kwai, "Wormhole Routing on a Class of High-Performance Fixed-Degree Parallel Processor Networks," *Proc. 11th Int'l Conf. Parallel and Distributed Computing and Systems*, pp. 376-378, Oct. 1998.
- [28] F.P. Preparata and J. Vuillemin, "The Cube-Connected Cycles: A Versatile Network for Parallel Computation," *Comm. ACM*, vol. 4, no. 5, pp. 300-309, May 1981.
- [29] C.S. Raghavendra, M. Gerla, and A. Avizienis, "Reliable Loop Topologies for Large Local Computer Networks," *IEEE Trans. Computers*, vol. 34, no. 1, pp. 46-55, Jan. 1985.
- [30] M. Valerio, L.E. Moser, P.M. Melliar-Smith, and P. Sweazey, "The QuickRing Network," *Proc. ACM 22nd Computer Science Conf.*, pp. 233-240, Mar. 1994.
- [31] J. van Leeuwen and R.B. Tan, "Interval Routing," *The Computer J.*, vol. 30, no. 4, pp. 298-307, Aug. 1987.
- [32] P.M. Vitanyi, "Locality, Communication, and Interconnect Length in Multicomputers," *SIAM J. Computing*, vol. 17, no. 4, pp. 659-672, Apr. 1988.
- [33] C.K. Wong and D. Coppersmith, "A Combinatorial Problem Related to Multimodule Memory Organization," *J. ACM*, vol. 21, no. 3, pp. 392-402, July 1974.
- [34] C.-H. Yeh and B. Parhami, "Cyclic Petersen Networks—Efficient Fixed-Degree Interconnection Networks for Large-Scale Multi-computer Systems," *Proc. Int'l Conf. Parallel and Distributed Processing: Techniques and Applications*, pp. 549-560, Aug. 1996.
- [35] G.W. Zimmerman and A.H. Esfahanian, "Chordal Rings as Fault-Tolerant Loops," *Discrete Applied Math.*, vol. 37/38, pp. 563-573, July 1992.



Behrooz Parhami received his PhD in computer science from University of California, Los Angeles, in 1973. Presently, he is a professor in the Department of Electrical and Computer Engineering, University of California, Santa Barbara. His research deals with parallel architectures and algorithms, computer arithmetic, and reliable computing. In his previous position with Sharif University of Technology in Tehran, Iran (1974-1988), he was also involved in the areas of educational planning, curriculum development, standardization efforts, technology transfer, and various editorial responsibilities, including a five-year term as editor of *Computer Report*, a Farsi-language computing periodical. His technical publications include more than 170 papers in journals and international conferences, a Farsi-language textbook, and an English/Farsi glossary of computing terms. Two textbooks on computer arithmetic (Oxford) and parallel processing (Plenum) have recently been published.

Dr. Parhami is a fellow of both the IEEE and the British Computer Society, a member of the ACM, and a Distinguished Member of the Informatics Society of Iran for which he served as a founding member and president during 1979-1984. He also served as a chairman of IEEE Iran Section (1977-1986) and received the IEEE Centennial Medal in 1984.



Ding-Ming Kwai received the BS and MS degrees in Taiwan from the National Cheng Kung University, Tainan, and the National Chiao Tung University, Hsinchu, in 1987 and 1989, respectively, and the PhD degree from the University of California, Santa Barbara, in 1997. He was with the Chung Cheng Institute of Technology, Taoyuan, Taiwan, as a reserve officer during 1989-1991, and with the Hualon Microelectronics Corporation, Hsinchu, Taiwan,

as a design engineer during 1991-1993. His research interests include parallel processing, VLSI architectures, and fault-tolerant computing.