Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

J. Parallel Distrib. Comput. 64 (2004) 183–190

Journal of Parallel and Distributed Computing

http://www.elsevier.com/locate/jpdc

Incomplete k-ary n-cube and its derivatives

Behrooz Parhami* and Ding-Ming Kwai

Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560, USA

Received 12 January 1997; revised 12 December 1999

Abstract

Incomplete or pruned k-ary n-cube, $n \geqslant 3$, is derived as follows. All links of dimension $n-1$ are left in place and links of the remaining $n-1$ dimensions are removed, except for one, which is chosen periodically from the remaining dimensions along the intact dimension $n-1$. This leads to a node degree of 4 instead of the original $2n$ and results in regular networks that are Cayley graphs, provided that $n-1$ divides k. For $n=3$ $(n=5)$, the preceding restriction is not problematic, as it only requires that k be even (a multiple of 4). In other cases, changes to the basis network to be pruned, or to the pruning algorithm, can mitigate the problem. Incomplete k-ary n-cube maintains a number of desirable topological properties of its unpruned counterpart despite having fewer links. It is maximally connected, has diameter and fault diameter very close to those of k-ary n-cube, and an average internode distance that is only slightly greater. Hence, the cost/performance tradeoffs offered by our pruning scheme can in fact lead to useful, and practically realizable, parallel architectures. We study pruned k-ary n-cubes in general and offer some additional results for the special case $n=3$.

Keywords: Cayley graph; Fault diameter; Fault tolerance; Fixed-degree network; Interconnection network; k-Ary n-cube; Pruning; Routing algorithm, VLSI layout

1. Introduction

Pruning is the process of removing links from a basis network with the goal of simplifying its implementation (scalability, modularity, VLSI layout, etc.) and overcoming the bandwidth limitations at various levels of the packaging hierarchy, while maintaining the smaller diameter and average internode distance associated with higher degrees of connectivity. A number of useful interconnection networks were derived by pruning suitably chosen networks or were subsequently shown to be pruned versions of other networks. Examples include the cube-connected cycles with power-of-2 size [16], packed exponential connections [9], periodically regular chordal rings [15], and binary Gaussian cubes [8]. In the preceding networks, each node is provided with a subset of its original connections and a group of nodes collectively provide the same communication capabilities as one node of the original network.

One advantage of a pruned network is that it may be capable of emulating the original basis network efficiently, given their structural similarities. However, if pruning is not done carefully, the routing algorithm may become so complicated and/or message traffic so unbalanced as to hurt performance. We are thus motivated to focus on pruning schemes that produce symmetric networks. Our pruned k-ary n-cubes $(n \geqslant 3)$ are derived as follows. All links of dimension $n-1$ are left in place and all but one of the links for the remaining $n-1$ dimensions are removed in a periodic fashion along the intact dimension $n-1$. This leads to a node degree of 4 instead of the original $2n$ and results in regular networks that are Cayley graphs, provided that $n-1$ divides k. For $n=3$ $(n=5)$, the preceding restriction is not problematic, as it only requires that k be even (a multiple of 4). In other cases, changes to the basis network to be pruned, or to the pruning scheme, can mitigate the problem.

Whereas pruning leads to reduced node degree, Cartesian or cross-product networks work in the opposite direction [17]. The k-ary n-cube [6] is simply the cross-product of n rings of size k. Pruning a cross-product network may be viewed as an attempt to offset the complexity introduced by using Cartesian products, while maintaining some of the benefits gained. Pruned

*Corresponding author. Fax: +1-805-893-3262.
E-mail address: parhami@ece.ucsb.edu (B. Parhami).

0743-7315/$ - see front matter
doi:10.1016/j.jpdc.2003.11.009

versions of k-ary n-cubes have been shown to be quite effective [2,13], but the studies thus far have been restricted to 3D torus as the basis network. Advantages of more general incomplete k-ary n-cubes over their unpruned counterparts may be viewed in two complementary ways:

- Simplified packaging due to an effective reduction in dimensionality.
- Accommodation of wider communication channels with a fixed pin limit.

Together, these advantages more than offset the negative effects of less favorable topological parameters such as slightly increased diameter and lower bisection bandwidth.

2. Network structure and symmetries

Consider a $k \times k \times k$ torus, where k is even. In the pruned version of this torus, links of dimensions x and y are alternately removed [2]. Each node (x, y, z), $0 \leqslant x, y, z \leqslant k-1$, is connected to two neighbors $(x, y, z \pm 1)$ and also has two other neighbors: $(x \pm 1, y, z)$ if z is even, or $(x, y \pm 1, z)$ if z is odd. Throughout the paper, it will be understood that node-index expressions are evaluated modulo k. Since dimensions x, y, and z can be permuted without changing the connectivity, the preceding yields a node-transitive (symmetric) graph of degree 4. An example, with $k = 4$, is depicted in Fig. 1.

An nD k-torus (k-ary n-cube) with k^n nodes, $n \geqslant 3$, can be similarly pruned to a constant node degree of 4. In the resulting network, node $(a_0, a_1, \ldots, a_{n-1})$ is connected to two dimension $n-1$ neighbors $(a_0, a_1, \ldots, a_{n-1} \pm 1)$ and also has two other neighbors: $(a_0, a_1, \ldots, a_i \pm 1, \ldots, a_{n-1})$, where $a_{n-1} \bmod (n-1) = i$. To assure that an equal number of links are provided along each pruned dimension (a necessary condition for symmetry), we require that k be a multiple of $n-1$. Then, every group of $n-1$ consecutive nodes around each k-node ring in dimension $n-1$ possess a complete set of dimensional links.

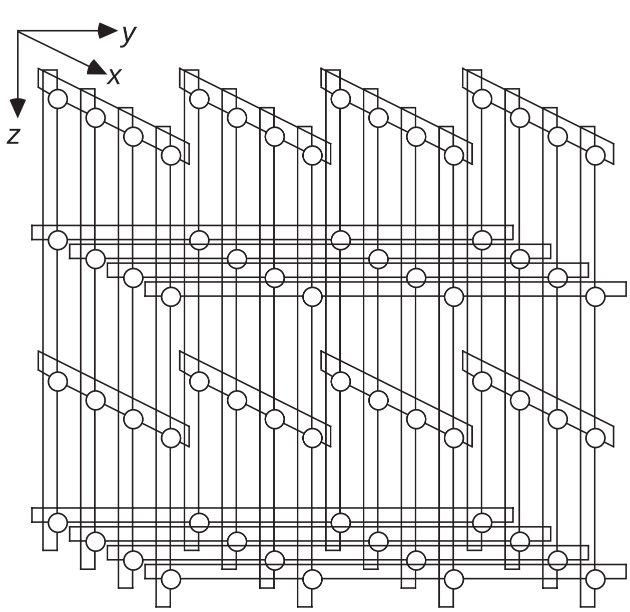

Fig. 1. Three-dimensional pruned 4-torus (4-ary 3-cube).

The k-ary n-cube is known to belong to the class of Cayley graphs and is thus node-transitive. Our proof that incomplete k-ary n-cube shares the same property is similar to that for n-CCC [5]. Let $*$ be an associative binary operator and Ω be a subset (generator set) from a finite group Γ such that $\imath \notin \Omega$ (where $\imath$ is the identity element) and for each $\omega \in \Omega$, its inverse ω^{-1} (satisfying $\omega * \omega^{-1} = \imath$) is also in Ω. A Cayley graph [1] is a graph whose nodes x and $x * \omega$, both in Γ, are connected iff $\omega \in \Omega$. The proof that a network is a Cayley graph is constructed by specifying the structure of the group (the node set Γ and the operator $*$) and the associated generator set Ω. For more details, the reader is referred to [12].

Theorem 1. *Incomplete k-ary n-cube is a Cayley graph.*

Proof. The set Γ consists of node indices $(a_0, a_1, \ldots, a_{n-1}), 0 \leqslant a_i \leqslant k-1$. To simplify our presentation, we write the node index $(a_0, a_1, \ldots, a_{n-1})$ as (α, β), where $\alpha = [a_0, a_1, \ldots, a_{n-2}]^T$ is an $(n-1)$-vector and $\beta = a_{n-1}$. We next define the group operator $*$ as

$$(\alpha, \beta) * (\alpha', \beta') = (\alpha + \Phi^{\beta}\alpha', \beta + \beta'),$$

where Φ^{β} is the βth power of the $(n-1) \times (n-1)$ matrix:

$$\Phi = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ : & : & : & & : & : \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{bmatrix}.$$

The first addition in the expression defining $*$ is component-wise and both additions are modulo k. The matrix Φ has the periodic property $\Phi^i = \Phi^{i+j(n-1)}$, where $0 \leqslant i \leqslant n-2$. The proof is complete upon specifying the identity element $\imath = ([0, 0, \ldots, 0]^T, 0)$ and the generator set $\Omega = \{([0, 0, \ldots, 0]^T, 1), ([0, 0, \ldots, 0]^T, k-1), ([1, 0, \ldots, 0]^T, 0), ([k-1, 0, \ldots, 0]^T, 0)\}$, whose closure under inverse makes all links bidirectional. Note that the operator $*$ corresponds precisely to our definition of incomplete k-ary n-cube; i.e., it connects $([a_0, \ldots, a_{n-2}]^T, a_{n-1})$ to $([a_0, \ldots, a_{n-2}]^T, a_{n-1} \pm 1)$ and $([a_0, \ldots, a_i \pm 1, \ldots, a_{n-2}]^T, a_{n-1})$ if $a_{n-1} \bmod (n-1) = i$. □

Note that because dimension $n-1$ is treated differently from other dimensions in the proof of Theorem 1, the proof remains valid for a class of networks that are derived from pruning a $k \times k \times \cdots \times k \times l$ torus, where $l > k$. This generalization leads, among other things, to an alternate derivation of the cube-connected cycles network as a Cayley graph. It also reduces the burden imposed by the requirement that k be divisible by $n-1$, because only l, and not k, needs this property. Thus, k can be chosen to be a power of 2, which is often most convenient in practice.

In addition to node-symmetry, which is a direct consequence of the Cayley graph result of Theorem 1, incomplete k-ary n-cubes are also edge-symmetric; meaning that the network looks exactly the same when viewed from different edges. Edge-transitivity has important practical implications for balanced utilization of network resources and for fault tolerance.

Theorem 2. *Incomplete k-ary n-cube is edge-transitive.*

Proof. We first note that the k^n nodes of the incomplete k-ary n-cube constitute a set of k^{n-1} node-disjoint rings of k nodes, in two different ways; one of the sets is formed by edges of dimension $n-1$ and the other set by edges of all other dimensions (see Fig. 1). Consider two edges from different dimensions (the case of same-dimension edges is trivial). If both edges are in k-rings of dimensions other than $n-1$, then their roles can be exchanged by simply renumbering the dimensions. If one edge belongs to dimension $\delta, 0 \leqslant \delta \leqslant n-2$, and the other is from dimension $n-1$, then relabeling links of dimensions $\{0, 1, \ldots, n-2\}$ as dimension $n-1$, and vice versa, will do the trick (again, refer to Fig. 1). □

3. Diameter and shortest paths

Given two nodes $A = (a_0, a_1, \ldots, a_{n-1})$ and $B = (b_0, b_1, \ldots, b_{n-1})$, the Lee distance [4] between them is defined as $\sum_{i=0}^{n-1} |\Delta d_i|$, where the offset $|\Delta d_i| = \min(|b_i - a_i|, k - |b_i - a_i|)$ is the minimum possible number of routing steps along dimension i in moving from A to B. In the complete k-ary n-cube, the Lee distance is also the length of the shortest path between the nodes. To incorporate the direction (forward/positive or backward/negative) of the shortest dimensional distance from A to B, we define:

$$\Delta d_i = \text{if } b_i - a_i \pmod{k} \leqslant a_i - b_i \pmod{k} \text{ then } |\Delta d_i| \text{ else } -|\Delta d_i|.$$

Theorem 3 shows that the diameter (maximum internode distance) of the incomplete k-ary n-cube is equal to or only slightly larger than that of its unpruned counterpart.

Theorem 3. *The diameter of the incomplete k-ary n-cube is*

$$n\lfloor k/2 \rfloor + \max(2n - 4 - \lfloor k/2 \rfloor, 0) \quad \text{if } k \geqslant 2n-2,$$
$$n\lfloor k/2 \rfloor + \max(n - 3, \lceil k/2 \rceil) \quad \text{if } k = n-1.$$

Proof. We can establish the network diameter by constructing a shortest path from source node $(0, 0, \ldots, 0)$ to destination node $(\Delta d_0, \Delta d_1, \ldots, \Delta d_{n-1})$, where $|\Delta d_i| \leqslant k/2$, and maximizing its length. Links of dimension $n-1$ are accessible at every node, so their traversal involves no added overhead. Links of dimension 0 are directly accessible from the source node and can be traversed at the outset, if needed. As we take the required $|\Delta d_{n-1}|$ hops along dimension $n-1$, links from the other $n-2$ dimensions become accessible for traversal. Thus, if $|\Delta d_{n-1}| \geqslant n-2$, links for every other dimension become accessible while routing along dimension $n-1$ and there will be no extra hop compared to routing on a complete k-ary n-cube. We need not consider this case further. For $|\Delta d_{n-1}| < n-2$, we must consider two cases, as routing in the direction indicated by the sign of Δd_{n-1} may not be the best choice when the extra hops needed to accommodate the unencountered dimensional links are taken into account.

Case 1 (*Route according to the sign of* Δd_{n-1}): In this case, $|\Delta d_{n-1}|$ of the $n-2$ dimensions can be traversed with no added overhead. We may need up to $2(n-2-|\Delta d_{n-1}|)$ extra hops for the remaining dimensions; going past the destination along dimension $n-1$ and returning to cancel the redundant moves. Thus, the worst-case routing distance is $\sum_{i=0}^{n-1} |\Delta d_i| + 2n - 4 - 2|\Delta d_{n-1}|$.

Case 2 (*Route in the direction opposite to the sign of* Δd_{n-1}): In this case, $k - |\Delta d_{n-1}|$ nodes are visited.

Subcase 2a: If $k - |\Delta d_{n-1}| \geqslant n-2$, then nothing more is needed and the total routing distance is $\sum_{i=0}^{n-1} |\Delta d_i| + k - 2|\Delta d_{n-1}|$.

Subcase 2b: If, however, $k - |\Delta d_{n-1}| < n-2$, we may need $2(n-2-k+|\Delta d_{n-1}|)$ additional hops, making the worst-case distance $\sum_{i=0}^{n-1} |\Delta d_i| + 2n - 4 - k$.

To proceed with the proof, it is more convenient to consider the case $k = n-1$ separately (recall that $n-1$ divides k). For $k \geqslant 2(n-1)$, we have $k - |\Delta d_{n-1}| \geqslant k/2 > n-2$ and $k > 2n-4$. So, case 1, which always has a smaller distance than the applicable subcase 2a, leads to the overall worst-case distance $(n-1)\lfloor k/2 \rfloor + 2n - 4$ for $|\Delta d_{n-1}| = 0$, which, combined with the fact that the diameter cannot be less than $n\lfloor k/2 \rfloor$, proves the first statement of the theorem (for $k \geqslant 2n-2$).

For $k = n-1$, we have $k - |\Delta d_{n-1}| \geqslant n-2$ if $|\Delta d_{n-1}| \leqslant 1$. So, subcase 2a, which has a smaller distance than case 1, leads to a worst-case total distance $(n-1)\lfloor k/2 \rfloor + k = n\lfloor k/2 \rfloor + \lceil k/2 \rceil$. The remaining case where $|\Delta d_{n-1}| \geqslant 2$ makes the total routing distance $\sum_{i=0}^{n-1} |\Delta d_i| + 2n - 4 - 2|\Delta d_{n-1}|$ for case 1 and $\sum_{i=0}^{n-1} |\Delta d_i| + 2n - 4 - k$ for subcase 2b. Since the latter two values vary in opposite directions as $|\Delta d_{n-1}|$ changes, the smaller of the two is maximized if they are (approximately) equal; this leads to $|\Delta d_{n-1}| = \lceil k/2 \rceil$ (or $\lfloor k/2 \rfloor$) and the worst-case shortest distance $n\lfloor k/2 \rfloor + n - 3$. The latter two cases prove the second statement of the theorem (for $k = n - 1$). □

Corollary 1. *The diameter of the incomplete k-ary n-cube is at most $k-2$ larger than that of its unpruned counterpart. This worst case occurs only when $k = n-1$ and $n \geqslant 7$. For $k \geqslant 4n - 4$, which almost always holds in practice (because usually n is relatively small and k is rather large), the diameter is equal to that of the unpruned network.*

Proof. In Theorem 2, the additive *max* terms represent the increase in diameter due to pruning. In the second statement of Theorem 2, $n - 3 = k - 2$ is larger than $\lceil k/2 \rceil$ if $k \geqslant 6$ $(n \geqslant 7)$. The additive term in the first statement never exceeds $n - 3$, which is in turn less than $k/2$. □

Corollary 2. *The shortest path between nodes A and B in the incomplete k-ary n-cube is at most $2n - 4$ hops longer than that of its unpruned counterpart.*

Proof. The path lengths derived for cases 1 and 2 in the proof of Theorem 2 readily lead to this result. For $|\Delta d_{n-1}| \leqslant k - n + 2$, subcase 2a applies and the length of the shortest path is derived, from the smaller of case 1 and subcase 2a, to be $\sum_{i=0}^{n-1} |\Delta d_i| + \min(2n-4, k) - 2|\Delta d_{n-1}|$. This path length exceeds $\sum_{i=0}^{n-1} |\Delta d_i|$ by no more than $2n - 4$. For $|\Delta d_{n-1}| > k - n + 2$, subcase 2b applies and the length of the shortest path is derived, from the smaller of case 1 and subcase 2b, to be $\sum_{i=0}^{n-1} |\Delta d_i| + 2n - 4 - \max(k, 2|\Delta d_{n-1}|) = \sum_{i=0}^{n-1} |\Delta d_i| + 2n - 4 - k$, leading to a considerably smaller increase, if any, over the shortest path in the unpruned torus. As an example, for $n = 3$, the shortest path between two nodes in an incomplete k-ary 3-cube is no more than 2 hops longer than that in an unpruned 3D k-torus. □

Intuitively, the fact that shortest paths are only slightly longer in the worst case suggests that the average internode distance for an incomplete k-ary n-cube should be very close to that of its unpruned counterpart. We do not have closed-form results for the average internode distance, except in the case of $n = 3$ (see Section 6). However, extensive numerical experimentation has shown that the increase is indeed negligible in practice [11].

4. Other topological properties

The bisection width of a network is the minimum number of links that must be removed to divide the network into two equal parts. Bisection width is an indicator of communication capacity and also sets a lower bound on interconnect length for a given diameter [3]. Theorem 4 shows that the bisection width of the incomplete k-ary n-cube is smaller than that of its unpruned counterpart by a factor of $n - 1$.

Theorem 4. *The bisection width of the incomplete k-ary n-cube is $2k^{n-1}/(n-1)$.*

Proof. It is easily seen that the network can be bisected by removing $2k^{n-1}/(n-1)$ links, e.g., by a hyperplane, near $a_0 = k/2$, that cuts only links of dimension 0. To complete the proof, we must show that the network cannot be bisected by removing fewer than $2k^{n-1}/(n-1)$ links. Assume the opposite, i.e., the existence of a bisection cut with $2k^{n-1}/(n-1) - 1$ or fewer links. The incomplete k-ary n-cube contains at least $k^{n-1}/(n-1)$ rings of length k in each dimension. Let us call the nodes that form any such k-ring a "string". Each node belongs to two strings. Consider the worst-case number of nodes that have at least one of their strings disconnected by the hypothesized bisection. Given that cutting one pair of links can create no more than k such nodes, their number is no greater than $k^n/(n-1) - k/2$. The remaining nodes, that number at least $k^n(n-2)/(n-1) + k/2$, remain fully connected, given that there is at least one connected string along every dimension. Because $(n-2)/(n-1) \geqslant 1/2$, the assumed bisection cut cannot exist. □

Embedding of an unpruned k-ary n-cube into a pruned network allows us to run algorithms of the k-ary n-cube on the incomplete network with reasonable efficiency, provided that the dilation and congestion of the embedding are not too large. The dilation is defined as the maximum distance in the host graph between nodes onto which a pair of neighboring nodes of the guest graph are mapped. The congestion is defined as the maximum number of times that a link is used to route messages between such pairs of nodes. The aforementioned embedding also has fault tolerance implications in that it allows us to emulate algorithms for an intact network on a faulty one that is missing certain links. In deriving the dilation and congestion of the embedding in Theorem 5, we assume that each bidirectional link is composed of two unidirectional links.

Theorem 5. *The incomplete k-ary n-cube can embed its unpruned counterpart with dilation $2\lceil n/2 \rceil - 1$ and congestion $2\lfloor n/2 \rfloor \lceil n/2 \rceil - 2\lfloor n/2 \rfloor + 1$.*

Proof. Recall that every set of $n-1$ consecutive nodes along dimension $n-1$ have a complete set of dimensional links. Thus, a node $(a_0, a_1, \ldots, a_{n-1})$ can gain access to dimension i link by taking at most $\lceil (n-2)/2 \rceil = \lceil n/2 \rceil - 1$ steps along dimension $n-1$. The dilation is obtained by doubling the above to account for routing forward/backward along dimension $n-1$ and adding 1 for the hop along dimension i itself.

To perform communication along dimension $i, 0 \leqslant i \leqslant n-2$, we cluster every $n-1$ nodes along dimension $n-1$, centered at a node $(b_0, b_1, \ldots, b_{n-1})$ with $b_{n-1} \bmod (n-1) = i$. There are $\lfloor n/2 \rfloor - 1$ nodes on one side and $\lceil n/2 \rceil - 1$ nodes on the other side of the aforementioned center that share the dimension i link with node $(b_0, b_1, \ldots, b_{n-1})$. The congestion is obtained by noting that the dimension $n-1$ link to $(b_0, b_1, \ldots, b_{n-1})$ is used $2\lfloor n/2 \rfloor (\lceil n/2 \rceil - 1)$ times to route forward and backward along the pruned dimensions and adding 1 for routing along dimension $n-1$. □

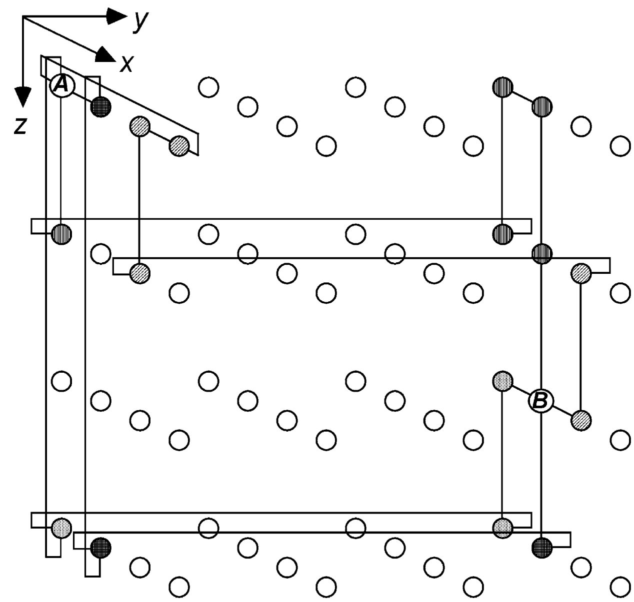

Fig. 2. Example of four parallel paths between two nodes (A and B).

5. Parallel paths and fault tolerance

One important consequence of edge-transitivity (Theorem 2) is that the number of node-disjoint parallel paths between any two nodes is the largest possible, i.e., equal to the node degree [12]. Such parallel paths allow us to select alternate routes and thus lead to maximal fault tolerance for the given node degree. An optimal (tight) upper bound for the lengths of these parallel paths yields the network's fault diameter [10]; that is, the maximum length of the shortest fault-free path in the presence of $d-1$ or fewer faults in a network with node degree d. The fault diameter of the k-ary n-cube is known to be $n\lfloor k/2 \rfloor + 1$ [7].

In Theorem 6, we build four parallel paths between two nodes in an incomplete k-ary n-cube with $k \geqslant 4n-4$ and derive their maximum length. In such a case, the network diameter is $n\lfloor k/2 \rfloor$ (see Corollary 1), making it easier to compare the derived fault diameter with that of the corresponding unpruned k-ary n-cube. This also leads to a worst-case result in view of the fact that the length k of our k-cycles is much greater than n. Even though the existence of four parallel paths in the case of $n=3$ is guaranteed only for $k \geqslant 8$, examining the parallel paths depicted in Fig. 2 (with $k=4$) might be helpful in understanding the proof.

Theorem 6. *For $k \geqslant 4n-4$, the fault diameter of the incomplete k-ary n-cube is at most $n\lfloor k/2 \rfloor + \lceil k/2 \rceil - n + 2$.*

Proof. We prove the result by constructing four parallel paths between node $(0, 0, \ldots, 0)$ and the arbitrary node $(a_0, a_1, \ldots, a_{n-1})$, where $a_{n-1} \bmod (n-1) = i$. Table 1 contains the specifications of the four paths, together with a record of the maximum path length to each intermediate point in the leftmost column. Node $(0, 0, \ldots, 0)$ has links of dimensions 0 and $n-1$. To ensure that the paths are node-disjoint, they start by visiting nodes $(1, 0, \ldots, 0)$, $(k-1, 0, \ldots, 0)$, $(0, 0, \ldots, 1)$, $(0, 0, \ldots, k-1)$ and from there proceed respectively to $(a_0, \ldots, a_i + 1, \ldots, a_{n-1})$, $(a_0, \ldots, a_i - 1, \ldots, a_{n-1})$, $(a_0, a_1, \ldots, a_{n-1} + 1)$, $(a_0, a_1, \ldots, a_{n-1} - 1)$ before arriving at node $(a_0, a_1, \ldots, a_{n-1})$. That the maximum path length does not exceed $n\lfloor k/2 \rfloor + \lceil k/2 \rceil - n + 2$ is evident from Table 1. To demonstrate why the four paths are node-disjoint, we indicate how they are constructed. For path 1, after the initial hop, dimensions are adjusted in order, from $n-2$ down to 1, with traversals of different dimensions separated by a single downward step along dimension $n-1$; this is followed by adjustment of dimensions 0 and $n-1$. Path 4 is quite similar to path 1, but is made to pass through different nodes (node labels starting mostly with 0 as opposed to 1). A key difference is that dimensions i and $n-1$ are adjusted last to ensure separation near the end. Paths 2 and 3 are similarly related, this time taking upward steps in dimensions $n-1$. Again, it is easily verified that the node labels encountered on these two paths are distinct from each other and from those of paths 1 and 4. □

6. Incomplete 3D k-torus

The case of $n=3$ is of special interest to us because lower dimensional networks are easier to build and

Table 1
Four node-disjoint paths from node 0^n to node $a_0 a_1 \ldots a_{n-1}$ in an incomplete k-ary n-cube[a]

Max hop distance	Node on path 1	Node on path 2	Node on path 3	Node on path 4
0	0^n	0^n	0^n	0^n
1	10^{n-1}	$(k-1)0^{n-1}$	$0^{n-1}1$	$0^{n-1}(k-1)$
2	$10^{n-2}(k-1)$	$(k-1)0^{n-2}1$	…	…
$\lfloor k/2 \rfloor + 1$	…	…	$0a_1 0^{n-3}1$	$0^{n-2}a_{n-2}(k-1)$
$\lfloor k/2 \rfloor + 2$	$10^{n-3}a_{n-2}(k-1)$	$(k-1)a_1 0^{n-3}1$	…	…
…	…	…	…	$0^{i+1}a_{i+1} \ldots a_{n-2}$ $(k-n+i+1)$
…	…	…	$0a_1 \ldots a_{i-1}$ $0^{n-i-1}i$	…
…	$10^{n-i}a_{n-i} \ldots a_{n-2}$ $(k-i)$	$(k-1)a_1 \ldots a_{i-1}$ $0^{n-i-1}i$	…	$0^i(a_i+1)a_{i+1} \ldots$ $a_{n-2}(k-n+i+1)$
$i\lfloor k/2 \rfloor + i$	…	…	$0a_1 \ldots (a_{i-1})$ $0^{n-i-2}i$	…
$i\lfloor k/2 \rfloor + i + 1$	$10^{n-i-1}a_{n-i-1} \ldots$ $a_{n-2}(k-i)$	$(k-1)a_1 \ldots a_i$ $0^{n-i-2}i$	…	…
$(n-1)\lfloor k/2 \rfloor + n - 1$	…	…	$a_0 \ldots a_{i-1}(a_i-1)$ $a_{i+1} \ldots a_{n-2}(n-1)$	$a_0 \ldots a_{i-1}(a_i+1)$ $a_{i+1} \ldots a_{n-2}(k-n-1)$
$(n-1)\lfloor k/2 \rfloor + n$	$a_0 a_1 \ldots a_{n-2}$ $(k-n+1)$	$a_0 a_1 \ldots a_{n-2}$ $(n-1)$	…	…
$(n-1)\lfloor k/2 \rfloor + k - n + 2$	$a_0 a_1 \ldots a_{n-1}$	$a_0 a_1 \ldots a_{n-1}$	$a_0 a_1 \ldots a_{n-1}$	$a_0 a_1 \ldots a_{n-1}$

[a] For a symbol $\sigma \in \{0, 1\}$, the notation σ^j stands for j repetitions of σ.

potentially offer greater performance when the cost and delay of interconnects, packaging limitations such as pin count, and the complexity of routing decisions are considered. In this special case, both the dilation and congestion of embedding a complete k-ary n-cube into the incomplete one (Theorem 5) are equal to 3. We can also derive a closed-form exact expression for the average internode distance in this case that serves to confirm our intuition regarding the negligible effect of pruning on this parameter.

Theorem 7. *The average internode distance of incomplete* 3D *k-torus is* $3k/4 + 2k^{-1} - 2k^{-2}$.

Proof. We need to find the sum of shortest distances from node (0,0,0) to all other nodes, dividing the result by the number k^3 of nodes (one could divide by $k^3 - 1$ to account for the fact that a node does not send messages to itself, but dividing by k^3 leads to simpler expressions and is adopted here). In the case of a complete 3D torus, the sum of distances would be $3k^4/4$ and the average internode distance is $3k/4$. For the incomplete torus, the proof of Theorem 3 suggests that path lengths do not change except when the destination node (x, y, z) satisfies $y \neq 0$ and $z = 0$. In the latter case, two additional hops are added. As the number of such nodes is $k^2 - k$, we must add $2k^2 - 2k$ to the original total distance of $3k^4/4$, proving the desired result. □

To keep the interconnect length balanced, long wraparound links should be avoided. The usual solution to this problem is folding, i.e., shuffling the odd- and even-indexed nodes in each cycle and placing them on different grid lines of the standard VLSI grid model. For a 3D k-torus, the volume occupied will be $8k^3$, as each node and its links occupy a $2 \times 2 \times 2$ subgrid after folding. The pruned links of the incomplete 3D k-torus allow a much more compact layout, requiring a volume of only $3k^3$. Fig. 3 depicts an example. To begin with, we can allocate a $1 \times 2 \times 2$ subgrid to each node. Furthermore, in stacking the layers, folding along one or the other dimension is required, not both of them. This reduces the volume by an additional factor of 4/3.

Like the unpruned 3D k-torus, the incomplete version is also Hamiltonian, meaning that it contains a ring encompassing all the nodes as a subgraph. Here is an explicit construction of the Hamiltonian cycle for the incomplete 3D k-torus with $k \geqslant 4$. Beginning from node (0,0,0), we proceed along dimension z, from $z = 0$ to $z = k - 1$, and from there to node $(0, 1, k-1)$. We next go backward along dimension z, from $k - 1$ down to 0, and from there to node $(k-1, 1, 0)$. These sweeps are repeated $2k - 1$ times, each time visiting k nodes, until we reach node $(1, 1, k-1)$. We next move one step forward in dimension y, $k - 1$ steps back in dimension z, and one step back in dimension x to node $(0, 2, 0)$. From node $(0, 2, 0)$, the entire process can be repeated, leading to node $(0, 4, 0)$. It is easily seen that such a path will eventually return to $(0, 0, 0)$, after visiting all other nodes, via a wraparound link from $(1, k-1, k-1)$; see Fig. 4 for an example.

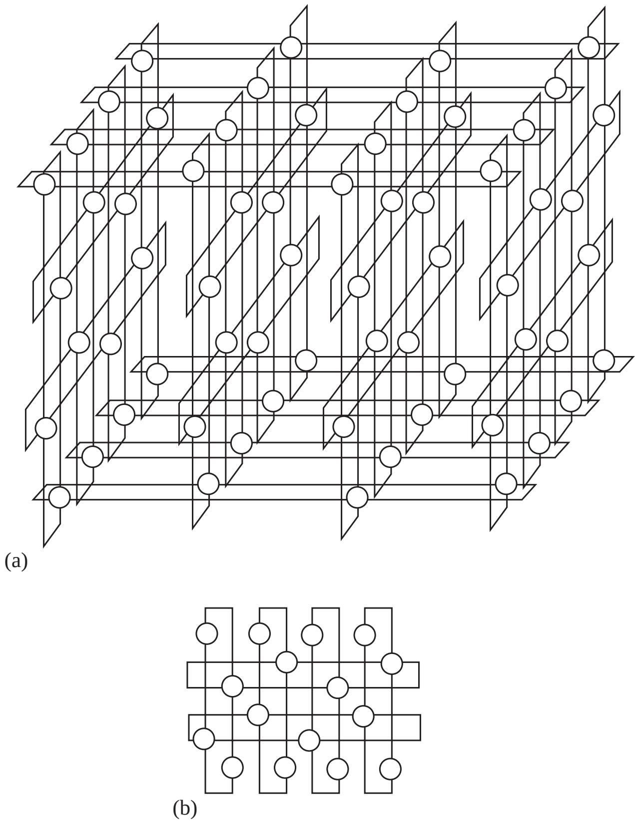

Fig. 3. Folded layout of a pruned 3D 4-torus network. (a) 3D view, (b) Side view.

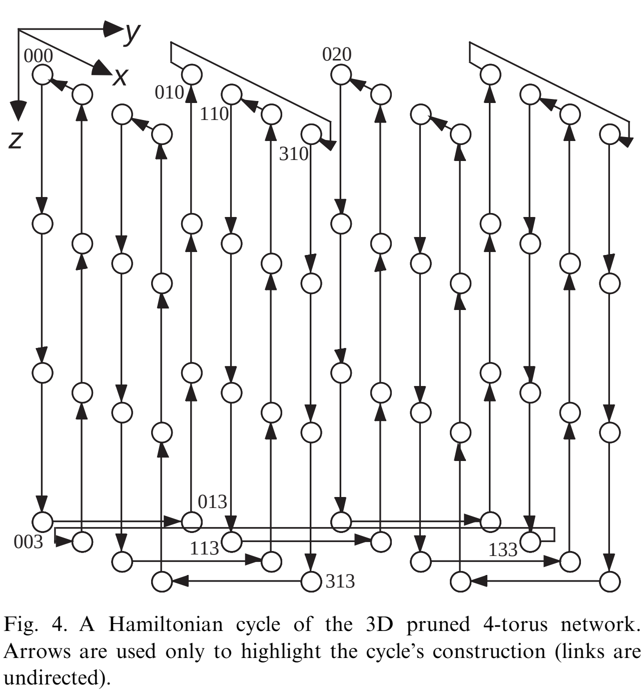

Fig. 4. A Hamiltonian cycle of the 3D pruned 4-torus network. Arrows are used only to highlight the cycle's construction (links are undirected).

7. Conclusion

In this paper, we applied a pruning scheme to the k-ary n-cube to reduce its node degree from $2n$ to 4. We showed that by removing links from the k-ary n-cube in a regular (periodic) fashion, many desirable properties of the basis network can be preserved. The pruned network obtained in this manner remains in the class of Cayley graphs and is additionally edge-transitive. The diameter and fault diameter of the resulting pruned network are close to those of the original network. These results can be viewed as additional evidence that k-ary n-cube is highly resilient, given that many of its desirable properties are maintained after removing a large number of links. A complementary conclusion is that dense connectivity is unnecessary for ensuring these properties. Given the same node capacity and complexity, we can use a smaller node degree, with correspondingly wider channels, to achieve improved communication performance [11].

Even though we focused on a particular type of pruning, with a degree-4 network as the end result, other pruning schemes can be easily envisaged. For instance, pruning to a node degree of 6, rather than 4, can be done by keeping two common dimensions and removing all but one of the remaining dimensions from each node in a periodic fashion. The resulting improvement in network diameter and fault diameter will be minimal, given the near-optimal results already achieved with node degree 4. However, improvements in other aspects, such as embeddings and communication performance, may make such networks worthy of additional investigation. Pruning is also applicable to other topologies [14]; for example, it can be applied to chordal rings [15] and to general torus networks with unequal side lengths. These latter derivative networks, an example of which was discussed immediately after Theorem 1, may prove important because in most practically built mesh- or torus-based parallel computers, unequal side lengths have been employed for various reasons, including packageability and incremental expandability.

References

[1] S.B. Akers, B. Krishnamurthy, A group-theoretic model for symmetric interconnection networks, IEEE Trans. Comput. 38 (4) (April 1989) 555–566.
[2] R. Alverson, D. Callahan, D. Cummings, B. Koblenz, A. Porterfield, B. Smith, The tera computer system, Proceedings of the ACM International Conference on Supercomputing, Amsterdam, June 1990, pp. 1–6.
[3] G. Bilardi, F.P. Preparata, Horizons of parallel computation, J. Parallel Distrib. Comput. 27 (June 1995) 172–182.
[4] B. Bose, B. Broeg, Y. Kwon, Y. Ashir, Lee distance and topological properties of k-ary n-cubes, IEEE Trans. Comput. 44 (August 1995) 1021–1030.
[5] G.E. Carlson, J. Cruthirds, H. Section, C. Wright, Interconnection networks based on generalization of cube-connected cycles, IEEE Trans. Comput. 34 (August 1985) 769–772.
[6] W.J. Dally, Performance analysis of k-ary n-cube interconnection networks, IEEE Trans. Comput. 39 (6) (June 1990) 775–785.

[7] Q.-P. Gu, S. Peng, Fault tolerant routing in Toroidal networks, IEICE Trans. Inform. Systems E79-D (August 1996) 1153–1159.
[8] W.-J. Hsu, M.J. Chung, Z. Hu, Gaussian networks for scalable distributed systems, Computer J. 39 (5) (1996) 417–426.
[9] W.W. Kirkman, D. Quammen, Packed exponential connections—a hierarchy of 2-D meshes, Proceedings of the International Parallel Processing Symposium, April 1994, pp. 464–470.
[10] M.S. Krishnamoorthy, B. Krishnamurthy, Fault diameter of interconnection networks, Comput. Math. Appl. 13 (5/6) (1987) 577–582.
[11] D.-M. Kwai, B. Parhami, Comparing torus, pruned torus, and Manhattan Street networks as interconnection architectures for highly parallel computers, Proceedings of the International Conference on Parallel and Distributed Computing and Systems, Cambridge, MA, November 1999, pp. 19–22.
[12] S. Lakshmivarahan, J.-S. Jwo, S.K. Dahl, Symmetry in interconnection networks based on Cayley graphs of permutation group: a survey, Parallel Comput. 19 (1993) 361–401.
[13] J. Nguyen, J. Pezaris, G. Pratt, S. Ward, Three-dimensional network topologies, Proceedings of the International Workshop Parallel Computer Routing and Communication, Seattle, May 1994, pp. 101–115.
[14] B. Parhami, Introduction to Parallel Processing: Algorithms and Architectures, Plenum Press, New York, 1999.
[15] B. Parhami, D.-M. Kwai, Periodically regular chordal rings, IEEE Trans. Parallel Distrib. Systems 10 (6) (June 1999) 658–672.
[16] F.P. Preparata, J. Vuillemin, The cube-connected cycles: a versatile network for parallel computation, Comm. ACM 24 (5) (May 1981) 300–309.
[17] A. Youssef, Design and analysis of product networks, Proceedings of the Symposium Frontiers of Massively Parallel Computation, February 1995, pp. 521–528.