Multiprocessor Architectures and Physical Law

Paul Vitányi*
CWI & Universiteit van Amsterdam
Address: CWI, Kruislaan 413
1098 SJ Amsterdam
The Netherlands

Abstract

We show that all highly symmetrical interconnection topologies for multiprocessors with low diameter require very long interconnect lengths. Therefore, such multicomputers do not scale well in the physical world with 3 dimensions. On the other hand, highly irregular (random) interconnection topologies have a very large subgraph of diameter two and therefore also require very long interconnect lengths. Hence the only scaling topologies for future massively parallel computers are high diameter regular ones, like mesh networks. The techniques used are symmetry properties in terms of orbits of automorphism groups of graphs, and a modern notion of randomness of individual objects, Kolmogorov complexity.

1 Introduction

In many areas of the theory of parallel computation we meet graph structured computational models which encourage the design of parallel algorithms where the cost of communication is largely ignored. Yet it is well known that the cost of computation - in both time and space - vanishes with respect to the cost of communication latency in parallel or distributed computing. We show that symmetric low diameter networks do not scale well; and present new results that random networks (and hence almost all networks) do not scale at all. This confirms that meshes are the way to go.

*CWI and Universiteit van Amsterdam. Email: paulvcwi.nl. Partially supported by the European Union through NeuroCOLT ESPRIT Working Group Nr. 8556, and by NWO through NFI Project ALADDIN under Contract number NF 62-376.

1.1 Multiprocessor Interconnect Architectures

Models of parallel computation that allow processors to randomly access a large shared memory, such as P-RAMs, or rapidly access a member of a large number of other processors, will necessarily have large latency. If we use 2^n processing elements of, say, unit size each, then the tightest they can be packed is in a 3-dimensional sphere of volume 2^n. Assuming that the units have no "funny" shapes, e.g., are spherical themselves, no unit in the enveloping sphere can be closer to all other units than a distance of radius R,

$$R = \left(\frac{3 \cdot 2^n}{4\pi} \right)^{1/3} \tag{1}$$

Because of the bounded speed of light, it is impossible to transport signals over $2^{\alpha n}$ ($\alpha > 0$) distance in polynomial $p(n)$ time. In fact, the assumption of the bounded speed of light says that the lower time bound on *any* computation using 2^n processing elements is $\Omega(2^{n/3})$ outright. Or, for the case of computations on networks which use n^α processors, $\alpha > 0$, the lower bound on the computation time is $\Omega(n^{\alpha/3})$.

Formerly, a wire had magical properties of transmitting data 'instantly' from one place to another (or better, to many other places). A wire did not take room, did not dissipate heat, and did not cost anything - at least, not enough to worry about. This was the situation when the number of wires was low, somewhere in the hundreds. Current designs use many millions of wires (on chip), or possibly billions of wires (on wafers). In a computation of parallel nature, most of the time seems to be spent on communication - transporting signals over wires. The present analysis allows us to see that any reasonable model for multicomputer computation must charge for communication. The communication cost will impact on both physical time and physical space costs.

2 The Problem with Symmetric Networks

At present, many popular multicomputer architectures are based on highly symmetric communication networks with small diameter. Like all networks with small diameter, such networks will suffer from the communication bottleneck above, i.e., necessarily contain *some* long interconnects (embedded edges). However, we can demonstrate that the desirable fast permutation properties of symmetric networks don't come free, since they require that the average of *all* interconnects is long. (Note that 'embedded edge,' 'wire,' and 'interconnect' are used synonymously.) To preclude objections that results like below hold only asymptotically (and therefore can be safely ignored for practical numbers of processors), or that processors are huge and wires thin (idem), we calculate precisely without hidden constants and assume that wires have length but no volume and can pass through everything. It is consistent with the results that wires have *zero* volume, and that *infinitely* many wires pass through a unit area. (Ω is used sometimes to simplify notation. The constant of proportionality can be reconstructed easily in all cases, and is never very small. The ideas presented in this section are more extensively elaborated in [5, 6].) Theorem 1 expresses a lower bound on the *average* edge length for *any* graph, in terms of certain symmetries and diameter. The argument is based on graph automorphism, graph topology, and Euclidean metric. For each graph topology we have examined, the resulting lower bound turned out to be sharp. Concretely, the problem is posed as follows. Let $G = (V, E)$ be a finite undirected graph, without loops or multiple edges, *embedded* in 3-dimensional Euclidean space. Let each embedded node have unit *volume*. For convenience of the argument, each node is embedded as a sphere, and is *represented* by the single point in the center. The *distance* between a pair of nodes is the Euclidean distance between the points representing them. The *length* of the embedding of an edge between two nodes is the distance between the nodes. How large does the *average* edge length need to be?

We illustrate the approach with a popular architecture, say the *binary n-cube*. Recall, that this is the network with $N = 2^n$ nodes, each of which is identified by an n-bit name. There is a two-way communication link between two nodes if their identifiers differ by a single bit. The network is represented by an undirected graph $C = (V, E)$, with V the set of nodes and $E \subseteq V \times V$ the set of edges, each edge corresponding with a communication link. There are $n2^{n-1}$ edges in C. Let C be embedded in 3-dimensional Euclidean space, each node as a sphere with unit volume. The distance between two nodes is the Euclidean distance between their centers. Let x be any node of C. There are at most $2^n/8$ nodes within Euclidean distance $R/2$ of x, with R as in Equation 1. Then, there are $\geq 7 \cdot 2^n/8$ nodes at Euclidean distance $\geq R/2$ from x. Construct a spanning tree T_x in C of depth n with node x as the root. Since the binary n-cube has diameter n, such a shallow tree exists. There are N nodes in T_x, and $N-1$ paths from root x to another node in T_x. Let P be such a path, and let $|P|$ be the *number of edges* in P. Then $|P| \leq n$. Let $length(P)$ denote the Euclidean length of the embedding of P. Since 7/8th of all nodes are at Euclidean distance at least $R/2$ of root x, the average of $length(P)$ satisfies

$$(N-1)^{-1} \sum_{P \in T_x} length(P) \geq \frac{7R}{16}$$

The average Euclidean length of an embedded edge *in a path P* is bounded below as follows:

$$(N-1)^{-1} \sum_{P \in T_x} \left(|P|^{-1} \sum_{e \in P} length(e) \right) \geq \frac{7R}{16n}. \quad (2)$$

This does *not yet* give a lower bound on the average Euclidean length of an edge, the average taken *over all edges* in T_x. To see this, note that if the edges incident with x have Euclidean length $7R/16$, then the average edge length *in each path* from the root x to a node in T_x is $\geq 7R/16n$, even if all edges not incident with x have length 0. However, the average edge length *in the tree* is dominated by the many short edges near the leaves, rather than the few long edges near the root. In contrast, in the case of the binary n-cube, because of its symmetry, if we squeeze a subset of nodes together to decrease local edge length, then other nodes are pushed farther apart increasing edge length again. We can make this intuition precise.

LEMMA 1 *The average Euclidean length of the edges in the 3-space embedding of C is at least* $7R/(16n)$.

PROOF. Denote a node a in C by an n-bit string $a_1 a_2 ... a_n$, and an edge (a, b) between nodes a and b differing in the kth bit by:

$$(a_1 ... a_{k-1} a_k a_{k+1} ... a_n, a_1 ... a_{k-1} (a_k \oplus 1) a_{k+1} ... a_n)$$

where $\oplus$ denotes modulo 2 addition. Since C is an undirected graph, an edge $e = (a, b)$ has two representations, namely (a, b) and (b, a). Consider the set A of automorphisms $\alpha_{v,j}$ of C consisting of

1. modulo 2 addition of a binary n-vector v to the node representation, followed by
2. a cyclic rotation over distance j.

Formally, let $v = v_1 v_2 ... v_n$, with $v_i = 0, 1$ $(1 \leq i \leq n)$, and let j be an integer $1 \leq j \leq n$. Then $\alpha_{v,j} : V \rightarrow V$ is defined by

$$\alpha_{v,j}(a) = b_{j+1} ... b_n b_1 ... b_j$$

with $b_i = a_i \oplus v_i$ for all i, $1 \leq i \leq n$.

Consider the spanning trees $\alpha(T_x)$ isomorphic to T_x, $\alpha \in A$. The argument used to obtain Equation 2 implies that for *each* α in A separately, in each path $\alpha(P)$ from root $\alpha(x)$ to a node in $\alpha(T_x)$, the average of $length(\alpha(e))$ over all edges $\alpha(e)$ in $\alpha(P)$ is at least $7R/16n$. Averaging Equation 2 additionally over all α in A, the same lower bound applies:

$$\frac{\sum_{\alpha \in A} \left[(N-1)^{-1} \sum_{P \in T_x} \left(|P|^{-1} \sum_{e \in P} length(\alpha(e)) \right) \right]}{N \log N} \geq \frac{7R}{16n} \quad (3)$$

Now fix a particular edge e in T_x. We average $length(\alpha(e))$ over all α in A, and show that this average equals twice the average edge length. Together with Equation 3 this will yield the desired result. For each edge f in C there are $\alpha_1, \alpha_2 \in A$, $\alpha_1 \neq \alpha_2$, such that $\alpha_1(e) = \alpha_2(e) = f$, and for all $\alpha \in A - \{\alpha_1, \alpha_2\}$, $\alpha(e) \neq f$. (For $e = (a, b)$ and $f = (c, d)$ we have $\alpha_1(a) = c$, $\alpha_1(b) = d$, and $\alpha_2(a) = d$, $\alpha_2(b) = c$.) Therefore, for each $e \in E$,

$$\sum_{\alpha \in A} length(\alpha(e)) = 2 \sum_{f \in E} length(f)$$

Then, for *any* path P in C,

$$\sum_{e \in P} \sum_{\alpha \in A} length(\alpha(e)) = 2|P| \sum_{f \in E} length(f) \quad (4)$$

Rearranging the summation order of Equation 3, and substituting Equation 4, yields the lemma. □

The symmetry property yielding such huge edge length is 'edge-symmetry.' To formulate the generalization of Lemma 1 for arbitrary graphs, we need some mathematical machinery. We recall the definitions from [1]. Let $G = (V, E)$ be a simple undirected graph, and let Γ be the automorphism group of G. Two edges $e_1 = (u_1, v_1)$ and $e_2 = (u_2, v_2)$ of G are *similar* if there is an automorphism γ of G such that $\gamma(\{u_1, v_1\}) = \{u_2, v_2\}$. We consider only connected graphs. The relation 'similar' is an equivalence relation, and partitions E into nonempty equivalence classes, called *orbits*, $E_1, ..., E_m$. We say that Γ *acts transitively* on each E_i, $i = 1, ..., m$. A graph is *edge-symmetric* if every pair of edges are similar $(m = 1)$.

Additionally, we need the following notions. If x and y are nodes, then $d(x, y)$ denotes the number of edges in a *shortest path* between them. Let D denote the *diameter* of G defined by D is the maximum over all node pairs x, y of $d(x, y)$. For $i = 1, ..., m$, define $d_i(x, y)$ as follows. For edges $\{x, y\} \in E$, if $\{x, y\} \in E_i$ then $d_i(x, y) = 1$, else $d_i(x, y) = 0$. Let Π be the set of shortest paths between x and y along edges in E. If x and y are not incident on the same edge $(\{x, y\} \notin E)$, then $d_i(x, y) = |\Pi|^{-1} \sum_{P \in \Pi} \sum_{e \in P} d_i(e)$. Clearly,

$$d_1(x, y) + \cdots + d_m(x, y) = d(x, y) \leq D$$

Denote $|V|$ by N. The ith *orbit frequency* is

$$\delta_i = N^{-2} \sum_{x, y \in V} \frac{d_i(x, y)}{d(x, y)},$$

$i = 1, ..., m$. Finally, define the *orbit skew coefficient* of G as $M = \min\{|E_i|/|E| : 1 \leq i \leq m\}$. Consider a d-space embedding of G, with embedded nodes, distance between nodes, and edge length as above. Let R be the *radius* of a d-space sphere with volume N, Equation 1 for $d = 3$. We are now ready to state the main result.

THEOREM 1 *Let graph G be embedded in d-space with the parameters above, and let $C = (2^d - 1)/2^{d+1}$.*

(i) Let $l_i = |E_i|^{-1} \sum_{e \in E_i} l(e)$ be the average length of the edges in orbit E_i, $i = 1, ..., m$. Then, $\sum_{1 \leq i \leq m} l_i \geq \sum_{1 \leq i \leq m} \delta_i l_i \geq CRD^{-1}$.

(ii) Let $l = |E|^{-1} \sum_{e \in E} l(e)$ be the average length of an edge in E. Then, $l \geq CRMD^{-1}$. [1]

The proof is a generalization of the argument for the binary n-cube, see [6]. Let us apply the theorem to a few examples.

EXAMPLE 1 (BINARY N-CUBE) Let Γ be an automorphism group of the binary n-cube, e.g., A in the proof of Lemma 1. Let $N = 2^n$. The orbit of each edge under Γ is E. Substituting R, D, $m = 1$, and $d = 3$ in Theorem 2 (ii) proves Lemma 1. Denote by L the *total* edge length $\sum_{f \in E} l(f)$ in the 3-space embedding of C. Then

$$L \geq \frac{7RN}{32} \quad (5)$$

[1] This constant C can be improved. For $d = 3$, from 7/16 to 3/4. Similarly, in 2 dimensions we can improve C from 3/8 to 2/3.

Recapitulating, the sum *total* of the lengths of the edges is $\Omega(N^{4/3})$, and the *average* length of an edge is $\Omega(N^{1/3}\log^{-1} N)$. (In 2 dimensions we obtain in a similar way $\Omega(N^{3/2})$ and $\Omega(N^{1/2}\log^{-1} N)$, respectively.) $\diamond$

EXAMPLE 2 (CUBE-CONNECTED CYCLES)
The binary n-cube has the drawback of unbounded node degree. Therefore, in the fixed degree version of it, each node is replaced by a *cycle* of n trivalent nodes, [4] whence the name *cube-connected cycles* or CCC. If $N = n2^n$, then the CCC version, say $CCC = (V, E)$, of the binary n-cube has N nodes, $3N/2$ edges, and diameter $D < 2.5n$. Theorem 1 shows: The average Euclidean length of edges in a 3-space embedding of CCC is at least $7R/(120n)$.

The *total* edge length is $\Omega(N^{4/3}\log^{-1} N)$ and the *average* edge length is $\Omega(N^{1/3}\log^{-1} N)$. (In 2 dimensions $\Omega(N^{3/2}\log^{-1} N)$ and $\Omega(N^{1/2}\log^{-1} N)$, respectively.) I expect that similar lower bounds hold for other fast permutation networks like the *butterfly-*, *shuffle-exchange-* and *de Bruijn* graphs. $\diamond$

EXAMPLE 3 (EDGE-SYMMETRIC GRAPHS) Recall that a graph $G = (V, E)$ is *edge-symmetric* if each edge is mapped to every other edge by an automorphism in Γ. We set off this case especially, since it covers an important class of graphs. (It includes the binary n-cube but excludes CCC.) Let $|V| = N$ and D be the diameter of G. Substituting $R, m = 1$, and $d = 3$ in Theorem 2 (i) we obtain: The average Euclidean length of edges in a 3-space embedding of an edge-symmetric graph is at least $7R/(16D)$.

For the complete graph K_N, this results in an average wire length of $\geq 7R/16$. I.e., the average wire length is $\Omega(N^{1/3})$, and the total wire length is $\Omega(N^{7/3})$.

For the complete bigraph $K_{1,N-1}$ (the star graph on N nodes) we obtain an average wire length of $\geq 7R/32$. I.e., the average wire length is $\Omega(N^{1/3})$, and the total wire length is $\Omega(N^{4/3})$.

For a N-node δ-dimensional mesh with wraparound (e.g., a ring for $\delta = 1$, and a torus for $\delta = 2$; for a formal definition see Appendix), this results in an average wire length of $\geq 7R/(8N^{1/\delta})$. I.e., the average wire length is $\Omega(N^{(\delta-3)/3\delta})$, and the total wire length is $\Omega(\delta N^{(4\delta-3)/3\delta})$. $\diamond$

EXAMPLE 4 (COMPLETE BINARY TREE) The complete binary tree T_n on $N - 1$ nodes ($N = 2^n$) has $n - 1$ orbits $E_1, ..., E_{n-1}$. Here E_i is the set of edges at level i of the tree, with E_1 is the set of edges incident with the leaves, and E_{n-1} is the set of edges incident with the root. Let l_i and l be as in Theorem 2 with $m = n - 1$. Then $|E_i| = 2^{n-i}$, $i = 1, ..., n - 1$, the orbit skew coefficient $M = 2/(2^n - 2)$, and we conclude from Theorem 2 (ii) that l is $\Omega(N^{-2/3}\log^{-1} N)$ for $d = 3$. This is consistent with the known fact l is $O(1)$. However, we obtain significantly stronger bounds using the more general part (i) of Theorem 2. In fact, we can show that 1-space embeddings of complete binary trees with $o(\log N)$ average edge length are impossible. Theorem 1 shows: The average Euclidean length of edges in a d-space embedding of a complete binary tree is $\Omega(1)$ for $d = 2, 3$, and $\Omega(\log N)$ for $d = 1$. $\diamond$

There is evidence that the lower bound of Theorem 1 is optimal in the sense of being within a constant multiplicative factor of an upper bound for several example graphs of various diameters, [6].

3 The Problem with Random Networks

Since low-diameter symmetric network topologies lead to high average interconnect length, it is natural to ask what happens with irregular topologies. In fact, it is sometimes proposed that since symmetric networks of low diameter lead to high interconnect length, one should use random networks where the presense or absence of a connection is determined by a coin flip. We shall show here that such networks will also have impossibly high average interconnect length.

One way to express irregularity or *randomness* of an individual network topology is by a modern notion of randomness called Kolmogorov complexity. We refer the reader to the textbook [2] for definitions and basic facts of Kolmogorov complexity. For the purpose of reading this article, it is sufficient to know that the *Kolmogorov complexity* $K(x)$ of a finite string x is simply the length of the shortest program, say in FORTRAN[2] encoded in binary, which prints x without any input. A similar definition holds conditionally, in the sense that $K(x|y)$ is the length of the shortest binary program which computes x given y as input. It can be shown that the Kolmogorov complexity is absolute in the sense of being independent of the programming language, up to a fixed additional constant term which depends on the programming language but not on x. We now fix one canonical programming language once and for all as reference and thereby $K()$.

A simple counting argument shows that for each y in the condition and each length n there exists at

[2]Or in Turing machine codes.

least one x of length n which is *incompressible* in the sense of $K(x|y) \geq n$, 50% of all x's of length n is incompressible but for 1 bit ($K(x|y) \geq n-1$), 75%th of all x's is incompressible but for 2 bits ($K(x|y) \geq n-2$) and in general a fraction of $1-2^{-c}$ of all strings cannot be compressed by more than c bits, [2].

Each graph $G=(V,E)$ on n nodes $V=\{0,\ldots,n-1\}$ can be coded (up to isomorphism) by a binary string of length $n(n-1)/2$. We enumerate the $n(n-1)/2$ possible edges in a graph on n nodes in standard order and set the ith bit in the string to 1 if the edge is present and to 0 otherwise. Conversely, each binary string of length $n(n-1)/2$ encodes a graph on n nodes. Hence we can identify each such graph with its corresponding binary string.

We shall call a graph G on n nodes *random* if it satisfies

$$K(G|n) \geq n(n-1)/2 - o(n). \tag{6}$$

Recall that $f(n)=o(n)$ iff $\lim_{n\to\infty} f(n)/n = 0$. Elementary counting shows that *a fraction* of at least

$$1 - o(n/2^n)$$

of all graphs on n nodes has that high complexity.

LEMMA 2 *The degree d of each node of a random graph satisfies $|d-(n-1)/2| = o(n)$.*

PROOF. Assume that the deviation of the degree d of a node v in G from $(n-1)/2$ is at least k. From the lower bound on $K(G|n)$ corresponding to the assumption that G is random, we can estimate an upper bound on k, as follows.

Describe G given n as follows. We can indicate which edges are incident on node v by giving the index of the connection pattern in the ensemble of

$$m = \sum_{|d-(n-1)/2| \geq k} \binom{n}{d} \leq 2^n e^{-k^2/(n-1)} \tag{7}$$

possibilities. The last inequality follows from a general estimate of the tail probability of the binomial distribution, with s_n the number of successful outcomes in n experiments with probability of success $0<p<1$ and $q=1-p$. Namely, Chernoff's bounds, [2], pp. 127-130, give

$$\Pr(|s_n - np| \geq k) \leq 2e^{-k^2/4npq}. \tag{8}$$

To describe G it then suffices to modify the old code of G by prefixing it with

- the identity of the node concerned in $\lceil \log n \rceil$ bits,
- the value of d in $\lceil \log n \rceil$ bits, possibly adding non-significant 0's to pad up to this amount,
- the index of the interconnection pattern in $\log m + 2\log\log m$ bits in self-delimiting form (this form requirement allows the concatenated binary sub-descriptions to be parsed and unpacked into the individual items: it encodes a separation delimiter, at the cost of adding the second term, [2]),

followed by the old code for G with the bits in the code denoting the presence or absence of the possible edges which are incident on the node v deleted.

Clearly, given n we can reconstruct the graph G from the new description. The total description we have achieved is an effective program of

$$\log m + 2\log\log m \pm O(\log n) + n(n-1)/2 - (n-1)$$

bits. This must be at least the length of the shortest effective binary program, which is $K(G|n)$ satisfying Equation 6. Therefore,

$$\log m + 2\log\log m \geq n - o(n).$$

Since we have estimated in Equation 7 that

$$\log m \leq n - (k^2/(n-1))\log e,$$

it follows that

$$k = o(n).$$

□

Note that we can, for concreteness sake, replace $o(n)$ everywhere by say $5n/\log n$. The lemma shows that each node is within two edges of about 50% of all nodes in G. Hence G contains a subgraph on about 50 % of its nodes of diameter 2. [3] Therefore, we have the following Theorem.

THEOREM 2 *A fraction of at least $1-o(n/2^n)$ of all graphs on n nodes (the incompressible, random, graphs) have average interconnect length of $\Omega(n^{1/3})$ in each 3-dimensional Euclidean space embedding (or $\Omega(n^{1/2})$ in each 2-dimensional Euclidean space embedding).*

Since both the very regular symmetric low diameter graphs and the random graphs have high average interconnect length which sharply rises with n, the only graphs which will scale feasibly up are symmetric fairly high diameter topologies like the mesh—which therefore will most likely be the interconnection pattern of the future massive multiprocessor systems.

[3] A similar extended argument shows that high complexity (random) graphs have a subgraph containing almost all nodes which has diameter 3.

4 Interpretation of the Results

An effect that becomes increasingly important at the present time is that most space in the device executing the computation is taken up by the wires. Under very conservative estimates that the unit length of a wire has a volume which is a constant fraction of that of a component it connects, we can see above that in 3-dimensional layouts for binary n-cubes, the volume of the $N = 2^n$ components performing the actual computation operations is an asymptotic fastly vanishing fraction of the volume of the wires needed for communication:

$$\frac{\text{volume computing components}}{\text{volume communication wires}} \in o(N^{-1/3})$$

If we charge a constant fraction of the unit volume for a unit wire length, and add the volume of the wires to the volume of the nodes, then the volume necessary to embed the binary n-cube is $\Omega(N^{4/3})$. However, this lower bound ignores the fact that the added volume of the wires pushes the nodes further apart, thus necessitating longer wires again. How far does this go? A rigorous analysis is complicated, and not important here. The following intuitive argument indicates what we can expect well enough. Denote the volume taken by the nodes as V_n, and the volume taken by the wires as V_w. The total volume taken by the embedding of the cube is $V_t = V_n + V_w$. The total wire length required to lay out a binary n-cube as a function of the volume taken by the embedding is, substituting $V_t = 4\pi R^3/3$ in Equation 5,

$$L(V_t) \geq \frac{7N}{32} \left(\frac{3V_t}{4\pi} \right)^{1/3}$$

Since $\lim_{n\to\infty} V_n/V_w \to 0$, assuming unit wire length of unit volume, we set $L(V_t) \approx V_t$. This results in a better estimate of $\Omega(N^{3/2})$ for the volume needed to embed the binary n-cube. When we want to investigate an upper bound to embed the binary n-cube under the current assumption, we have a problem with the unbounded degree of unit volume nodes. There is no room for the wires to come together at a node. For comparison, therefore, consider the fixed degree version of the binary n-cube, the CCC (see above), with $N = n2^n$ trivalent nodes and $3N/2$ edges. The same argument yields $\Omega(N^{3/2} \log^{-3/2} N)$ for the volume required to embed CCC with unit volume per unit length wire. It is known, that every small degree N-vertex graph, e.g., CCC, can be laid out in a 3-dimensional grid with volume $O(N^{3/2})$ using a unit volume per unit wire length assumption, [3, 4]. This neatly matches the lower bound.

Because of current limitations to layered VLSI technology, previous investigations have focussed on embeddings of graphs in 2-space (with unit length wires of unit volume). We observe that the above analysis for 2 dimensions leads to $\Omega(N^2)$ and $\Omega(N^2 \log^{-2} N)$ volumes for the binary n-cube and the cube-connected cycles, respectively. These lower bounds have been obtained before using *bisection width* arguments, and are known to be optimal, [4]. In [3] it is shown that we cannot always assume that a unit length of wire has $O(1)$ volume. (For instance, if we want to drive the signals to very high speed on chip.)

The lower bounds on the wire length above are *independent* of the ratio between the volume of a unit length wire and the volume of a processing element. This ratio changes with different technologies and granularity of computing components. Previous results may not hold for optical communication networks, intraconnected by optical wave guides such as glass fibre or guideless by photonic transmission in free space by lasers, while ours do. The arguments we have developed are purely geometrical, apply to any graph, and give optimal lower bounds in all cases we have examined. Our observations are mathematical consequences from the noncontroversial assumptions on 3 dimensional space and the Laws of Physics.

References

[1] F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.

[2] M. Li and P.M.B. Vitányi, *An Introduction to Kolmogorov Complexity and its Applications*, Springer-Verlag, New York, 1993.

[3] C. Mead and L. Conway, *Introduction to VLSI Systems*, Addison-Wesley, Reading, MA, 1980.

[4] J. Ullman, *Computational Aspects of VLSI*, Computer Science Press, Rockville, MD, 1984.

[5] P.M.B. Vitányi, Non-sequential computation and Laws of Nature, In: VLSI Algorithms and Architectures (Proceedings Aegean Workshop on Computing, 2nd International Workshop on Parallel Processing and VLSI), *Lecture Notes In Computer Science 227,* Springer Verlag, 1986, 108-120.

[6] P.M.B. Vitányi, Locality, communication and interconnect length in multicomputers, *SIAM J. Computing,* 17 (1988), 659-672.