DIAMETER MINIMIZATION IN NETWORKS FOR **SIMD** MACHINES

J. Ramon BEIVIDE
José L. BALCÁZAR

RR85/03

ABSTRACT

We propose an algebraic constructive method which allows to find a certain kind of networks having optimal diameter. These interconnection networks have been proposed for SIMD machines. We compare the results with other designs, pointing out the improvement achieved with this method.

RESUM

Es proposa un mètode algebraic constructiu que permet trobar un cert tipus de xarxes òptimes amb respecte al seu diàmetre. Aquestes xarxes d'interconnexió han estat propostes per a màquines SIMD. Es comparen els resultats amb altres dissenys, posant en evidència la ganància obtinguda amb aquest mètode.

I. INTRODUCTION.

A SIMD machine, as described by Siegel [8], is a system which consists of a control unit and N processing elements, connected by a interconnection network. The control unit is able to broadcast instructions to every processing element, so that those elements which are active at each time execute simultaneously the same instruction on its own data. A memory module is available to each of them to hold the data. Communication among the processing elements is provided by the interconnection network.

The two basic configurations for SIMD machines [3] are shown in the figure 1. An example of topology (a) is the ILLIAC IV computer [1], and an example of topology (b), with two alignment networks, is the BSP computer [4]. SIMD machines have been designed to be programmed with parallel algorithms for matrix operations.

An important problem in this type of systems is the design of an appropriate storage scheme for the data access in the memory modules. These schemes should provide fast access to the most frequently used sections of data. They should provide conflict-free access to rows, columns, diagonals, and subarrays, which are sections of data typically used in matricial computations. Several storage schemes have been proposed [2], [7], in which the so called p-ordered vectors arise. We deal with this kind of vectors.

An important task for the interconnection networks when working with this kind of schemes is to unscramble p-ordered vectors. The necessity of this organization is due to the fact that vectors of data, obtained from memory modules in a single access, are not usually in the required order.

Swanson [10] suggests a type of interconnection networks known as k-apart networks to unscramble the p-ordered vectors. These networks require

$$\left\lceil \frac{(8N+1)^{1/2}-3}{2} \right\rceil$$

steps through an interconnection network consisting of two k-apart networks. However, Lang and Stone [5] proved that perfect shuffle networks perform the same task in $\log_2 N$ steps. For this reason, Siegel [8], [9] does not consider the networks proposed by Swanson.

In this paper we propose a modification of the k-apart networks, by establishing bidirectional links among the registers of the network. In this way, unscrambling of p-ordered vectors can be achieved within

$$\left\lceil \frac{(2N-1)^{1/2}-1}{2} \right\rceil$$

steps through the interconnection network. Both Lang and Stone upper bounds and ours are shown in figure 2, in continuous and discrete forms. It can be seen from the figure that our performances are in general significantly better for values less than N=86, with three values in which both networks require 6 steps. We consider that these values are practical for particular implementations.

The structure of this paper is as follows: in section II we define some concepts based in the paper of Swanson [10], as well as the extension we propose of the k-apart networks; in section III we give an algebraic solution for the problem of finding two values of k giving an optimal network, for any number of memory modules. This algebraic solution has also practical interest for interconnection networks connecting each processing element with four neighbour processing elements, as in the case of the ILLIAC IV computer [1];

we consider this issue in section IV. Finally, we present in section V some ideas regarding the control of the optimal network.

II. DEFINITIONS.

In this section we present a set of definitions that will be used throughout the paper. As an initial remark, we point out that we assume a prime number of memory modules; this fact yields a number of practical advantages, as observed by Budnik and Kuck [2] and Lawrie [6].

<u>Definition</u> 1. An N-vector A is p-ordered, with $1 \leq p \leq N-1$, if its contents are described by

$$A\ (\ pi\ \underline{\mathrm{mod}}\ N\ ) = i$$

where $0 \leq i \leq N-1$. If N is a prime, then it is possible to define any p-ordered vector.

<u>Definition</u> 2. N registers are interconnected with a k-apart network, $1 \leq k \leq N-1$, if the contents of register (ki <u>mod</u> N) can be transferred directly to register i, with $0 \leq i \leq N-1$. If N is a prime, then it is possible to define any k-apart network.

<u>Definition</u> 3. The group $A_N = \langle A, +_N \rangle$ includes the set A of the nonnegative integers less than N, with the addition modulo N as operation.

<u>Definition</u> 4. The group $M_N = \langle M, \cdot_N \rangle$ includes the set M of the nonnegative integers less than N and relatively prime to N, with the product modulo N as operation. If N is a prime, then the group M_N is cyclic.

Proposition. A k-apart network, with $k \in M_N$, becomes a k'-apart network, with $k' \in M_N$, when the links are inverted. Further, k' is the inverse element of k in the group M_N.

Proof. Let A be a k-apart network. Then for $0 \leqslant i \leqslant N-1$ and $1 \leqslant k \leqslant N-1$,

$$\text{reg} (ki \underline{\text{mod}} N) \text{-->} \text{reg} (i)$$

following the notation of Swanson [10]. When the links are inverted,

$$\text{reg} (i) \text{-->} \text{reg} (ki \underline{\text{mod}} N).$$

Taking as a value of i the quantity $k^{-1}j$ mod N, for $0 \leqslant j \leqslant N-1$, we obtain

$$\text{reg} (k^{-1}j \underline{\text{mod}} N) \text{-->} \text{reg} (k(k^{-1}j \underline{\text{mod}} N) \underline{\text{mod}} N).$$

A straightforward manipulation shows that

$$k(k^{-1}j \underline{\text{mod}} N) \underline{\text{mod}} N = j$$

and hence the inverted network is

$$\text{reg} (k^{-1}j \underline{\text{mod}} N) \text{-->} \text{reg} (j)$$

which is the k^{-1}-apart network by definition.

∎

From now on, we consider k-apart bidirectional networks, in which the transfer of data may be performed in both directions.

III. AN OPTIMAL INTERCONNECTION NETWORK.

In this section we propose an algebraic method for finding an optimal interconnection network consisting of two k-apart bidirectional networks. By using two k-apart networks with $k=k1$, $k=k2$, $1 \leqslant k1, k2 \leqslant N-1$, any p-ordered vector, $1 \leqslant p \leqslant N-1$, can be unscrambled if there are two integers i, j such that:

$$(k1^{i} \,.\, k2^{j}) \ \underline{\text{mod}}\ N = p \qquad (1)$$

Positive values of i or j indicate transfers through the k1- or k2-apart networks respectively; and negative values indicate transfers through the inverse networks.

As we have a prime number of memory modules, the group M_N is cyclic and hence isomorphic to the group A_{N-1}. Therefore we state our results for A_{N-1}, using any generator g of M_N in the following way:

$$k1 = g^{x}, \qquad k2 = g^{y}, \qquad p = g^{z}$$

where $0 \leqslant x, y, z \leqslant N-2$, and we can rewrite (1) as

$$(ix + jy) \ \underline{\text{mod}}\ N-1 = z \qquad (2)$$

In order to achieve the unscrambling of p-ordered vectors as fast as possible, one must find integers i, j fulfilling (2) and such that $|i|+|j|$ be minimal. As z may take N-1 different values, the maximal values of $|i|$ and $|j|$ must be minimized by choosing in an inteligent manner the fixed values of x and y.

<u>Lemma</u> 1. For every $m \geqslant 0$ there are exactly $2m^2+2m+1$ different ways of choosing couples i, j with $|i|+|j| \leqslant m$.

Proof. The i may take 2m+1 different values, ranging from -m to m. For every i there are $2(m-|i|)+1$ possible values for j. Hence there are

$$\sum_{i=-m}^{m} 2(m-|i|)+1$$

different ways of choosing the couples i, j. Evaluating the series,

$$\sum_{i=-m}^{m} 2(m-|i|)+1 = 2m^2+2m+1$$

□

Theorem. Let m be an integer, $m > 2$, let $N_m = 2m^2+2m+1$, and let N be such that $N_{m-1} < N \leqslant N_m$. Then for every z, $0 \leqslant z \leqslant N-1$, there are two integers i, j such that

$$(im + j(m+1)) \ \underline{\text{mod}}\ N = z \qquad (3)$$

with $|i|+|j| \leqslant m$.

Remark. For N as in the theorem, no couple can behave better than (m, m+1), because if $|i|+|j| < m$ then the number of elements which can be generated is at most N_{m-1} according to lemma 1.

Before proving the theorem we state a set of auxiliar lemmas.

<u>Lemma</u> 2. If $|i|+|j| \leq m$ and $|p|+|q| \leq m$, then

$$(im + j(m+1)) = (pm + q(m+1)) \qquad (4)$$

implies $i=p$ and $j=q$.

<u>Proof</u>. By straightforward manipulation we obtain

$$(p-i)m = (j-q)(m+1) \qquad (5)$$

As GCD (m, m+1) = 1, m must divide $(j-q)$. Without loss of generality, we assume that $j \geq q$. By hypothesis, $|j| \leq m$ and $|q| \leq m$; hence,

$$|j-q| \leq |j|+|q| \leq 2m.$$

There are three cases:

(i) $j-q = 0$. By substituting in (4), we have $j=q$ and $i=p$.

(ii) $j-q = m$. By substituting in (4), we have $p-i = m+1$. On the other hand, from the hypothesis we have

$$|(p-i)+(j-q)| \leq |p|+|i|+|j|+|q| \leq 2m \qquad (5).$$

Substituting m for $j-q$ and $m+1$ for $p-i$ we get a contradiction.

(iii) $j-q = 2m$. By substituting in (4) we have $p-i = 2(m+1)$. Substituting again in (5) we get a contradiction.

∎

Lemma 3. Let m be an integer, $m > 2$, let $N_m = 2m^2+2m+1$, and let N be such that $N_{m-1} < N \leqslant N_m$. Let $c = ((i+j)m + j) \ \underline{\text{mod}}\ N$, with $|i|+|j| \leqslant m$. Then exactly one of the following cases holds:

(i) $(i+j)m + j \geqslant 0$, and we have $c = (i+j)m + j < N$;

(ii) $(i+j)m + j < 0$, and we have $c = (i+j)m + j + N$.

Proof. For $m > 2$ it holds that $m < m^2-2m+2$. We have:

$$| (i+j)m + j | \leqslant (|i|+|j|)m + |j| \leqslant m^2+m <$$
$$< m^2 + m^2 - 2m +2 = 2m^2 - 2m + 2 = N_{m-1} + 1 \leqslant N$$

Thus,

$$-N < (i+j)m + j < N$$

and the lemma follows.

□

Remark. For the lemma it is enough that $m^2+m < N$. For $m = 2$ and $N > 6$ it is also true. On the other hand, it is easily seen that the main theorem is still true for $N = 6$. Hence all the cases for $m > 1$ are covered.

Lemma 4. Let m be an integer, $m > 2$, N_m and N as in lemma 3, and let $t = N_m-N$. Let i, j be such that $|i|+|j| \leqslant m$. If $((i+j)m + j) \ \underline{\text{mod}}\ N < m^2+m+1-t$ then only case (i) of lemma 3 holds.

Proof. We show that case (ii) leads to a contradiction.

$$((i+j)m + j) \ \underline{\text{mod}}\ N < m^2+m+1-t \implies$$
$$((i+j)m + j) + N < m^2+m+1-t \implies$$
$$((i+j)m + j) < m^2+m+1-t-N \implies$$

$$((i+j)m + j) < m^2+m+1-t-2m^2-2m-1+t \implies$$
$$((i+j)m + j) < -m^2-m$$

As $((i+j)m + j) < 0$,

$$| (i+j)m + j | > m^2 + m$$

but as $|i|+|j| \leq m$ we have a contradiction.

$\square$

Lemma 5. Let m be an integer, $m > 2$. Let i, j be such that $|i|+|j| \leq m$. If $((i+j)m + j) \bmod N > m^2+m$ then only case (ii) of lemma 3 holds.

Proof. We show that case (i) leads to a contradiction.

$$((i+j)m + j) \bmod N > m^2+m$$

As case (i) implies that $((i+j)m + j) > 0$,

$$| (i+j)m + j | > m^2 + m$$

which is the same contradiction as before.

$\square$

Now we are ready to prove the main theorem.

Proof of theorem. Let i, j and p, q be two different couples with $|i|+|j| \leq m$ and $|p|+|q| \leq m$. Assume that

$$((i+j)m + j) \bmod N = ((p+q)m + q) \bmod N$$

Then for each of the two terms of the equality a different case of lemma 3 holds; if it were the same case, we would have

$$((i+j)m + j) = ((p+q)m + q)$$

contradicting lemma 2.

Let t be N_m-N as before. Now, the elements less than $m^2+m+1-t$ can be only generated by the case (i) of lemma 3 (because of lemma 4) and the elements greater than m^2+m can be only generated by the case (ii) of lemma 3 (because of lemma 5). Hence, the only couples generating repeated elements are the 2t couples generating the t elements between $m^2+m+1-t$ and m^2+m, each element being generated by at most two couples corresponding to the two cases of lemma 3. Now we have proved that at least $2m^2+2m+1-t = N$ different elements are generated. As there are exactly N elements, all of them have been generated. This completes the proof.

□

IV. A COMPARISON WITH THE ILLIAC IV NETWORK.

As mentioned before, the ILLIAC IV is an array processor having 64 processing elements. Each processing element can send data to one of four neighbour processing elements with just one transfer. We may characterize the network of the ILLIAC IV by four routing functions [3]:

$$\begin{aligned} P_{+1}(i) &= (i+1) \text{ \underline{mod} } N \\ P_{-1}(i) &= (i-1) \text{ \underline{mod} } N \\ P_{+r}(i) &= (i+r) \text{ \underline{mod} } N \qquad (6) \\ P_{-r}(i) &= (i-r) \text{ \underline{mod} } N \end{aligned}$$

for $0 \leqslant i \leqslant N-1$. In this particular case N = 64 and $r = \sqrt{N} = 8$. With this network, sending data from processing element i to processing element j requires at most $\sqrt{N}-1$ transfers through the network.

It is apparent now that we have here the same problem we have discussed in the last section. From our solution we see that, as

$$N_{m-1} < 64 \leq N_m \qquad \text{for } m = 6,$$

we can use the following routing functions instead of (6):

$$P_{+6}(i) = (\ i+6\ )\ \underline{\text{mod}}\ N$$
$$P_{-6}(i) = (\ i-6\ )\ \underline{\text{mod}}\ N$$
$$P_{+7}(i) = (\ i+7\ )\ \underline{\text{mod}}\ N$$
$$P_{-7}(i) = (\ i-7\ )\ \underline{\text{mod}}\ N$$

With this network sending data from any processing element to another requires at most $\sqrt{N}-2$ transfers, which is one less than the upper bound indicated above.

Anyway, for this particular case N = 64 we can find this optimal value of $\sqrt{N}-2$ changing only one of the two networks, because the couples (1,10) and (3,8) are also optimal. Observe also that an interconnection of 256 processing elements with a network similar to the one of the ILLIAC IV will require in the worst case 15 transfers, while 11 transfers will suffice when using a network obtained from our theorem.

V. THE CONTROL OF THE NETWORK.

We focus in this section on the problem of finding the fastest way for a communication through the optimal network proposed in the section III. We solve the problem for the additive group; a solution for the unscrambling of p-ordered vectors is obtained easily from the isomorphism.

Our problem is as follows: we have a bidirectional network with at most N_m elements, in which each node is connected to four neighbours nodes at distances m and m+1; given two nodes x, y of the network, we want numbers i, j, such that:

(i) $|i|+|j| \leqslant m$

(ii) $im + j(m+1) = (y-x) \underline{\bmod} N$

Observe that $im + j(m+1) = (i+j)m + j$. Hence, an integer division of y-x mod N by m will give a remainder j and a quotient i+j from which a couple (i,j) is obtained. This couple allows to reach node y from node x. Unfortunately, there is no warranty that $|i|+|j|$ is less than or equal to m. We will show how to find an optimal couple (i',j') from this one if it is not optimal.

First, we will proceed to this computation only if (y-x) mod N is strictly less than m^2+m. If y-x mod $N= m^2+m$ then an optimal solution is given by i = 0, j = m. If y-x mod N is greater than m^2+m, then the inverse of y-x mod N in A_N is smaller, and we can solve the problem for this inverse and then change the sign of the solutions.

Now, if y-x mod N is smaller than m^2+m, then the integer division by m will give a remainder j with $0 \leqslant j < m$ and a quotient (i+j) with $0 \leqslant (i+j) \leqslant m$. If $|i|+|j| \leqslant m$ then a solution has been found. Assume that $|i|+|j| > m$. We prove that $i < 0$.

If i is nonnegative, then $|i| = i$; also, $|j| = j$. Hence, $i+j = |i|+|j| > m$. But the quotient of the division is i+j and cannot be greater than m, because the dividend was smaller than m^2+m. Hence i is negative. From $0 \leqslant (i+j)$ we deduce $|i| = -i \leqslant j$.

Now define $i' = i+m+1$ and $j' = j-m$. (It is easily seen that these are the new quotient and remainder if the integer division is performed with negative remainder.) Then

$$i'm + j'(m+1) = (i+m+1)m + (j-m)(m+1) = im + j(m+1)$$

and the couple (i',j') generates also $y-x \bmod N$. On the other hand, as $|i| \leqslant j < m+1$,

$$|i'| = |i+m+1| = |\ m+1-|i|\ | = m+1 - |i|$$

and as $0 \leqslant j < m$,

$$|j'| = |j-m| = m-j = m-|j|.$$

Hence,

$$|i'|+|j'| = m+1-|i|+m-|j| = 2m+1-(|i|+|j|) < 2m+1-m = m+1$$

because we assumed $|i|+|j| > m$. Thus we have

$$|i'|+|j'| \leqslant m$$

as desired.

Using this method, the values of i and j may be obtained with PLA's or from a ROM memory. In the second case the size of the needed ROM would be

$$(N-1) \times (2\lceil \log_2 ((N-1)^{1/2}-1)/2 \rceil + 1))$$

where the "+1" is due to the sign bit.

VI. CONCLUSIONS.

We consider that one of the contributions of our work is to provide an example of a situation in which number-theoretic methods give solutions to nontrivial problems on computer architecture. In particular, a uniform algebraic solution has been found giving for any number of nodes an optimal interconnection network, consisting of two bidirectional networks; our optimality criteria has been the minimization of the number of steps through the network in the worst case.

These networks seem to admit several applications. We have studied its use for unscrambling p-ordered vectors, obtaining better results than perfect shuffle networks for values of N less than 86. Swanson's solution is significatively improved for every value of N.

Comparing this solution with the network in the ILLIAC IV computer, we show that the upper bound on the number of steps through the network is improved in 1 by the network we propose, with no extra hardware; this improvement could be greater for higher values of N.

Finally, we have shown how the control of the network can be performed, giving a simple algorithm of routing that can be easily implemented.

We will continue this line of research looking for similar characterizations of the optimal configurations for three or more networks with directed or not directed links, thus allowing a comparison with other existing networks, like n-cube and PM2I.

REFERENCES

[1] G. H. Barnes et al. "The ILLIAC IV computer", IEEE Trans. Comput., vol. C-17, pp. 746-757, Aug 1968.

[2] P. Budnik and D. J. Kuck, "The organization and use of parallel memories", IEEE Trans. Comput., vol. C-20, pp. 1566-1569, Dec 1971.

[3] K. Hwang and F. A. Briggs, "Computer architecture and parallel processing", McGraw-Hill Book Company 1984, pp. 325-342.

[4] D. J. Kuck and R. A. Stokes, "The Burroughs Scientific Processor", IEEE Trans. Comput., vol. C-31, pp. 363-376, May 1982.

[5] T. Lang and H. S. Stone, "A shuffle-exchange network with simplified control", IEEE Trans. Comput., vol. C-25, pp 55-65, Jan. 1976.

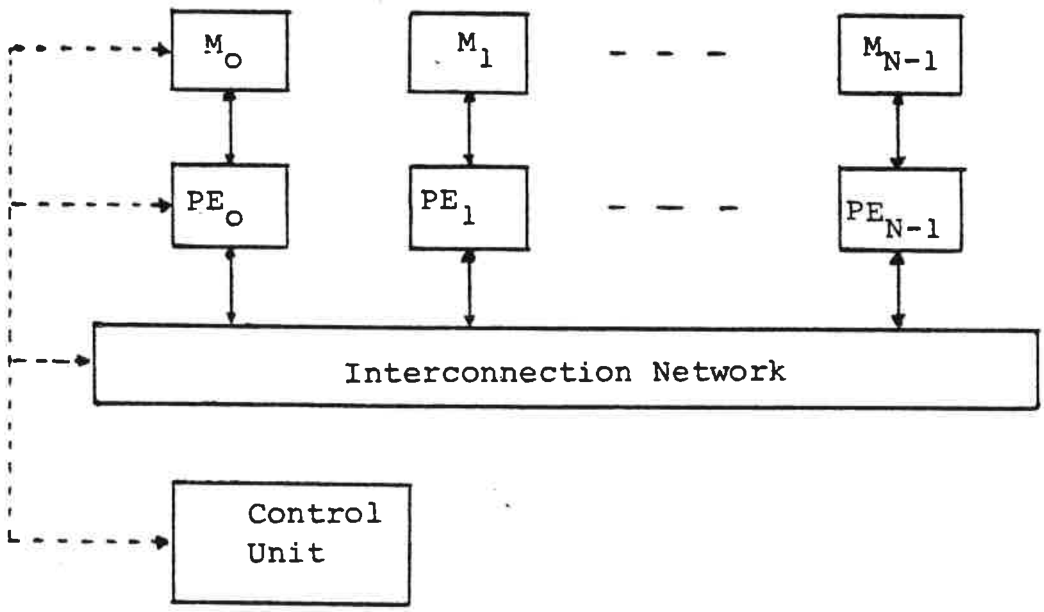
[6] D. H. Lawrie and C. R. Vora, "The prime memory system for array access", IEEE Trans. Comput., vol. C-31, pp. 435-442, May 1982.

[7] H. D. Shapiro, "Theoretical limitations on the efficient use of parallel memories", IEEE Trans. Comput., vol. C-27, pp. 421-428, May 1978.

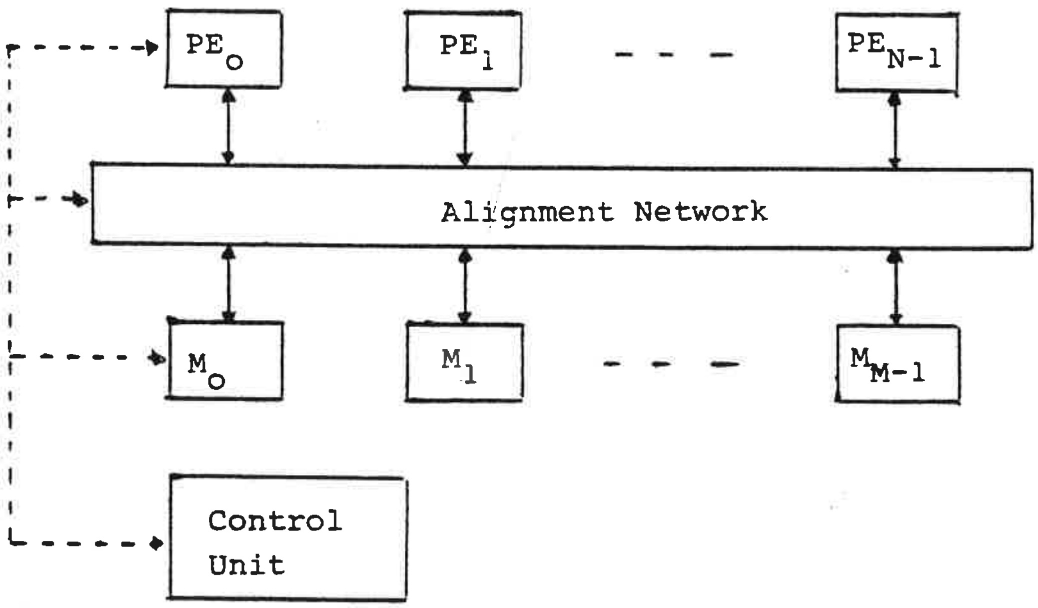
[8] H. J. Siegel, "Analysis techniques for SIMD machine interconnection networks and the effects of processor address masks", IEEE Trans. Comput., vol. C-26, pp. 153-161, Feb. 1977.

[9] H. J. Siegel, "A model of SIMD machines and a comparison of various interconnection networks", IEEE Trans. Comput., vol. C-28, pp. 907-917, Dec. 1979.

[10] R C. Swanson, "Interconnections for parallel memories to unscramble p-ordered vectors", IEEE Trans. Comput., vol. C-23, pp. 1105-1115, Nov. 1974.



Topology a



Topology b

Figure

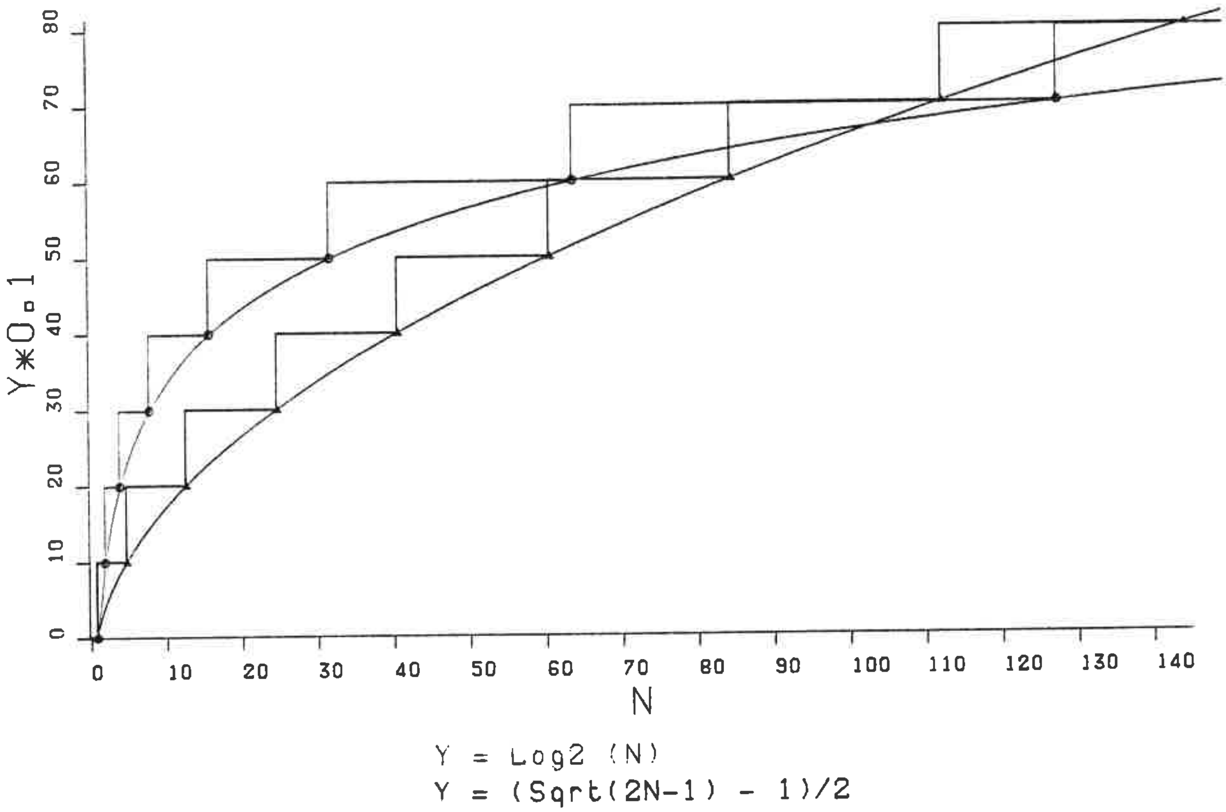

Y*0.1
0
10
20
30
40
50
60
70
80
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
N
Y = Log2 (N)
Y = (Sqrt(2N-1) - 1)/2

Figure 2