Reprinted from INFORMATION AND COMPUTATION
All Rights Reserved by Academic Press, New York and London

Vol. 72, No. 3, March 1987
Printed in Belgium

Area-Time Optimal Division for $T = \Omega((\log n)^{1+\epsilon})$*

K. MEHLHORN

Fachbereich 10, Informatik, Universität des Saarlandes, 6600, Saarbrücken, West Germany

AND

F. P. PREPARATA

Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801

Area-time optimal VLSI division circuits are described for all computation times in the range $[\Omega((\log n)^{1+\varepsilon}), O(\sqrt{n})]$ for arbitrary $\varepsilon > 0$. © 1987 Academic Press, Inc.

1. INTRODUCTION

A simple transformation of right-shift to integer division shows that the area-time (AT^2) complexity of any network for the computation of the inverse of an n-bit number (referred to here as "divider") is bounded from below by $\Omega(n^2)$. A trivial fan-in argument also gives $T = \Omega(\log n)$. A family of AT^2-optimal dividers has been proposed some time ago by Mehlhorn (1984). A network of this family can be constructed for each computation time T in the range $[\Omega(\log^2 n), O(\sqrt{n})]$. Since then considerable progress has been made in the design of faster dividers (Reif, 1983), culminating in the result of Beame, Cook, and Hoover (1984) illustrating an $O(\log n)$-time divider (i.e., a time-optimal network in the hypothesis of bounded-fan-in components). However, the Beame–Cook–Hoover network (referred to here as the BCH network) does not achieve area optimality. Thus, it is natural to ask the question of the existence of area-time optimal dividers for $T = o(\log^2 n)$. This paper provides an affirmative answer for $T \in [\Omega(\log n)^{1+\varepsilon}), O(\log^2 n)]$ for any positive constant $\varepsilon \leqslant 1$. It must be pointed out that the proposed networks are so complicated—notwithstanding their area-time optimality—that they are exclusively of theoretical interest.

* This work was supported by the DFG, SFB 124, TP B2, VLSI Entwurf und Parallelität, and by NSF Grant ECS–84–10902.

0890-5401/87 $3.00

Copyright © 1987 by Academic Press, Inc.
All rights of reproduction in any form reserved.

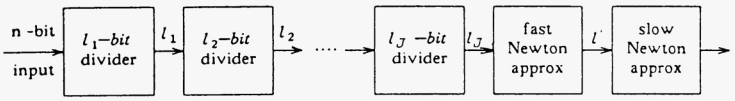

FIG. 1. Block structure of the divider.

The network (see Fig. 1) consists of $J+2$ cascaded modules, where $J \simeq 1/\varepsilon$. The first J modules are modified dividers of the BCH type, computing a sequence of approximations of the inverse with increasing numbers of bits $l_1 \leqslant l_2 \leqslant \cdots \leqslant l_J < n$.

The last two modules are designed to complete the buildup of the result size from l_J to n bits by implementing the Newton approximation method, which, at each iteration doubles the length of the result. This is carried out in two phases, respectively executed by the "fast" and "slow" approximators. The fast approximator basically consists of a single area-time optimal fastest multiplier, used to execute the initial iterations; the slow approximator is instead a cascade of affordably slow multipliers, each executing one of the final iterations. Note that the cascade of the two Newton approximators structurally coincides with Mehlhorn's (1984) divider.

The paper is organized as follows. In Section 2 we present a more efficient implementation of the BCH method leading to a circuit referred to as "modified BCH divider." In Section 3 we discuss an alternative method for the computation of the inverse, which uses the modified BCH method as a subroutine. Finally, in Section 4 we illustrate the combination of the previous techniques with the Newton approximation, to yield our proposed network, while Section 5 contains a few closing remarks.

In this paper we shall frequently refer to under- and overapproximations of the reciprocal of a number. For brevity, given an n-bit number x in the interval $[1/2, 1)$ (i.e., a normalized fraction with n bits to the right of the binary point), we say that for $l \leqslant n$, v is an l-bit **underinverse** or an l-bit **overinverse** of x depending upon whether $v = \lfloor 2^l/x \rfloor \cdot 2^{-l}$ or $v = \lceil 2^l/x \rceil \cdot 2^{-l}$. Equivalently, $v \cdot x = 1 \mp \delta$, with $\delta < 2^{-l}$ or v has two significant bits to the left and l significant bits to the right of the point.

2. AN EFFICIENT IMPLEMENTATION OF THE BCH METHOD

In this section we first describe (a variant of) the BCH method (Beame *et al.*, 1984) and then modify it so as to reduce its area requirement.

The original BCH method computes the n-bit underinverse of an n-bit number x by adding the first n powers of $u = 1 - x$ and truncating the n^2-bit result to its leading n-bits. Each power of u is computed individually and the n powers are subsequently added together. A power u^k is computed

by taking the "logarithm" of u, multiplying it by k, and then taking the "antilogarithm."

Since taking logarithms of large numbers is very hard, the method resorts to a modular representation and works as follows:

ALGORITHM INVERSE $1(x)$.

Input: an n-bit number x in the range $[1/2, 1)$. Given are m (small, possibly consecutive) primes $p_1, \ldots, p_m$ such that

$$\prod_{j=1}^{m} p_j \geqslant 2^{(n^2)} \qquad \text{(Note that } m \simeq n^2/\log n\text{)}$$

(n is assumed to be a power of two)

Output: an $(n+2)$-bit number v in the range $(1, 2)$, so that $v \times x = 1 + \delta$ with $\delta < 2^{-n}$

(1) **begin** $u := (1-x)\,2^n$; (*u is an integer *)

(2) **for** j, $1 \leqslant j \leqslant m$

(3) **pardo** $b_j := u \bmod p_j$;

(4) compute r_j so that $a_j^{r_j} = b_j$, where
a_j is a generator of the multiplicative group of $\mathbf{Z}^*_{p_j}$;

(5) **for** $l = 0$ **to** $\log n - 1$

(6) **do** $m_j^{(l)} := a_j^{r_j 2^l \bmod (p_j - 1)}$ (*$m_j^{(l)} = u^{2^l} \bmod p_j$*)

(7) **od**;

(8) $v_j := \prod_{l=0}^{\log n - 1} (m_j^{(l)} + 2^{n2^l}) \bmod p_j$
(*$v_j = 2^{n(n-1)} \cdot \sum_{t=0}^{n-1} (u/2^n)^t \bmod p_j$*)

(9) $V_j := v_j M_j \bmod (p_1 \cdots p_m)$
(* first step of Chinese remaindering *)

(10) **odpar**;

(11) $v := \sum_{j=1}^{m} V_j \bmod (p_1 \cdots p_m)$; (* second step of Chinese remaindering*)

(12) $v :=$ truncate v to the first $n+1$ bits, insert one 0 to the left, and set point after the second bit from the left

(13) **end**

Let us next describe the different steps of this algorithm in more detail. In this description we will make frequent use of the following two facts:

(1) One can multiply two k-bit integers in time T and area A, where $AT^2 = O(k^2)$ and $T \in [\Omega(\log k), O(\sqrt{k})]$. This is the result of (Mehlhorn and Preparata, 1983).

(2) One can add m k-bit integers in time $O(\log m + \log k)$ and area $O(km \cdot \log m)$. This can be achieved by expressing the m integers in redundant representation (see, e.g., [4–6]) and then adding them in a tree-like fashion. The tree has depth $O(\log m)$ and requires area $O(m \log m)$ for every bit position. Each level of the tree introduces a delay of just $O(1)$ thanks to the redundant number representation.

We are now ready to describe the circuit in more detail. We start with the parallel loop, lines 2–10.

Line 3. This line is easily executed in time $O(\log n)$ and area $O(n(\log n)^2)$ for each p_j by expressing u by its binary expansion $u = \sum_{t=0}^{n-1} u_t 2^t$, $u_t \in \{0, 1\}$, storing the numbers $2^t \bmod p_j$ in a table and performing the required additions in redundant number representation. We leave the details of this step to the reader.

Line 4. Step 4 is realized by a table loop-up, i.e., by a look-up in a table which gives the value of r_j for each possible value of b_j. Since p_j can certainly be expressed using $2\text{-}\log n$ bits this table has n^2 entries of $2 \log n$ bits each. We realize this table by $2 \log n$ H-trees each requiring area $O(n^2)$. Thus the total area is $O(n^2 \log n)$ for each p_j, and a table look-up takes time $O(\log n)$.

Note that the $2 \log n$ slices of the table are accessed in parallel. Also note that this circuit can be pipelined (its period is $O(1)$ in technical terms) and therefore $O(\log n)$ look-ups can also be performed in time $O(\log n)$ using the same area. This observation is important for step 6.

Lines 5, 6, 7. Consider a fixed l first. We first compute

$$R_j^{(l)} = r_j 2^l \bmod (p_j - 1)$$

as outlined in line 3. Note that the l-place shift does not have to be executed explicitly; it only determines which powers of two need to be looked-up. The computation of $R_j^{(l)}$ takes time $O(\log n)$ and area $O(n(\log n)^2)$. We perform this computation in parallel for all l, $0 \leqslant l \leqslant \log n - 1$.

The integer $m_j^{(l)}$ is computed from $R_j^{(l)}$ by look-up in a table of "antilogarithms." The $\log n$ look-ups are pipelined and take time $O(\log n)$ and area $O(n^2 \log n)$ for each p_j (refer to the description of line 4).

Finally note that $m_j^{(l)} = a_j^{r_j 2^l \bmod (p_j - 1)} = b_j^{2^l} \bmod p_j = u^{2^l} \bmod p_j$.

Line 8. We use a tree of multipliers. This tree has depth $O(\log \log n)$ and has $\log n$ nodes. Each node contains a circuit multiplying two $2 \log n$ bit numbers and reducing the result $\bmod p_j$ in time $O(\log \log n)$ and area $O((\log n)^2)$. This shows that step 8 takes time $O(\log n)$ and area $O(n)$. Both estimates are very generous.

Finally note that

$$\prod_{l=0}^{\log n-1}(2^{n2^l}+m_j^{(l)})=\prod_{l=0}^{\log n-1}(2^{n2^l}+u^{(2^l)})=2^{n(n-1)}\cdot\sum_{t=0}^{n-1}(u/2^n)^t.$$

Line 9. Let $M_j=[(p_1\cdots p_m)/p_j]^{p_j-1} \bmod (p_1\cdots p_m)$. Then M_j is the coefficient of v_j required for Chinese remaindering (Knuth, 1981). The number M_j is precomputed and stored in a register of length $O(n^2)$. We multiply v_j by M_j by dividing M_j into $n^2/\log n$ pieces of length $O(\log n)$, performing $n^2/\log n$ multiplications in parallel and then summing the results. This can certainly be done in time $O(\log n)$ and area $O(n^2\log n)$. Also the reduction $\bmod(p_1\cdots p_m)$ can be done in that area and time. Indeed, let q be an integer in $[0, 2^{n^2+\log n})$ and $M \triangleq p_1\cdots p_m$; then $q \bmod M = q-\lfloor q/M\rfloor\cdot M$. Thus we perform, in time $O(\log n)$ and area $O(n^2)$, a multiplication of q by an approximation of $1/M$ of precision $2^{-n^2-\log n}$ (having only $O(\log n)$ significant bits), followed by a multiplication of M by $\lfloor q/M\rfloor$.

Summary. Lines(3) to (9) take time $O(\log n)$ and area $O(n^2\log n)$ for each p_j. Since u^n has n^2 bits we have $m=\Theta(n^2/\log n)$ and each modulus is representable in $2\log n$ bits. We realize loop (2) to (10) by having a module for each modulus and hence the loop takes time $O(\log n)$ and area $O(n^4)$.

Line 11. In line 11 we add m numbers of n^2 bits each and reduce $\bmod(p_1\cdots p_m)$. This takes time $O(\log n)$ and area $O(m\log m\cdot n^2)=O(n^4)$.

LEMMA 1. *There exists a circuit which computes the n-bit inverse of an n-bit number in time* $O(\log n)$ *and area* $O(n^4)$.

Proof. Immediate from the discussion above. ∎

The enormous space requirement of the method sketched above is essentially due to the fact that the powers of u are computed with $\Theta(n^2)$ bits of precision. However, only the leading $n+\log n$ bits are truly needed for the computation of v. This observation is the key to the "modified" BCH method, to be described next. In the modified method we compute the powers of an l-bit integer u in m rounds (this m has nothing to do with the m in algorithm INVERSE 1), where m is a design parameter to be selected. In each round we compute the sum of $s=(l)^{1/m}$ consecutive powers using the method of Lemma 1. We call s the *depth* of the method. This takes time $O(\log l)$ and area $O((ls)^2)$ and yields a result of $O(ls)$ bits. The space requirement results from the fact that only $ls/\log(ls)$ different prime moduli, each of length $2\log(ls)$ bits, must be used. We truncate this result to $l+\lceil\log 12m\rceil$ bits and start the next round. The details are as follows.

ALGORITHM INVERSE $2(x)$

Input: an l-bit number $x \in [1/2, 1)$ and an integer $s = (l)^{1/m}$.
Output: an $(l+2)$-bit number $v \in (1, 2)$

begin $u_0 := 1 - x$;
for $i = 0$ **to** $m - 1$ **do**
begin
$\sigma_i := \sum_{j=0}^{s-1} u_i^j$
$u_{i+1} :=$ truncate u_i^s to $q = l + \lceil \log 12m \rceil$ bits right of point;
end;
$v :=$ truncate $\sigma_0 \sigma_1 \cdots \sigma_{m-1}$ to l bits right of point;
end.

To prove te correctness of this algorithm we must show that v gives the $(l+2)$ leading bits of $1/(1-u)$ (of which the rightmost l bits represent the fractional part). To this end, we must show that the error of the underapproximation is $< 2^{-l}$.

For any variable a used by the above algorithm let $\tilde{a}$ denote the corresponding exact value (note that, since all numbers are nonnegative, the truncation mechanism gives $\tilde{a} \geqslant a$), and $\delta(a)$ the absolute error on a, such that $a = \tilde{a} - \delta(a)$. Recall also that $\delta(a \cdot b) < \delta(a)\,\tilde{b} + \delta(b)\,\tilde{a}$ and that $\delta(a+b) = \delta(a) + \delta(b)$. Using these relationships, we readily have

$$\delta(\sigma_0 \cdots \sigma_{m-1}) < \tilde{\sigma}_0 \cdots \tilde{\sigma}_{m-1} \left(\frac{\delta(\sigma_0)}{\tilde{\sigma}_0} + \cdots + \frac{\delta(\sigma_{m-1})}{\tilde{\sigma}_{m-1}} \right).$$

Since $\tilde{\sigma}_0 \cdots \tilde{\sigma}_{m-1} < 3$ and $\tilde{\sigma}_i > 1$ $(i = 0, \ldots, m-1)$, we obtain

$$\delta(\sigma_0 \cdots \sigma_{m-1}) < 3(\delta(\sigma_0) + \cdots + \delta(\sigma_{m-1})).$$

From $\tilde{\sigma}_i = \sum_{j=0}^{s-1} \tilde{u}_i^j$ we have

$$\delta(\sigma_i) = \sum_{j=0}^{s-1} \delta(u_i^j) < \sum_{j=0}^{s-1} j \tilde{u}_i^{j-1} \delta(u_i) < \delta(u_i)/(1 - \tilde{u}_i)^2 < 4\delta(u_i),$$

since $\tilde{u}_i < \frac{1}{2}$ for $i = 1, \ldots, m-1$. (Obviously $\delta(u_0) = 0$.)

Thus $\delta(\sigma_0 \cdots \sigma_{m-1}) < 12m \max \delta(u_i)$ and the condition

$$12m \max \delta(u_i) < 2^{-l}$$

ensures the correctness of the method. We claim that $\delta(u_i) < 2^{-q}$ as a result of truncating to q bits right of the point. Indeed $\delta(u_1) < 2^{-q}$, trivially. For $i > 1$, assuming $\delta(u_i) < 2^{-q}$, let $u_{i+1}^* = u_i^s$ (before the truncation). Then

$$\delta(u_{i+1}^*) < s\tilde{u}_i^{s-1} \delta(u_i) < \frac{s}{2^{s-1}} \delta(u_i),$$

since $u_i < \frac{1}{2}$ for $i > 1$. If we assume $s \geqslant 2$, then $\delta(u^*_{i+1}) < 2^{-q}$, which shows that its q bits to the right of the point are correct. Thus, the prescribed truncation yields $\delta(u_{i+1}) < 2^{-q}$, and the induction step is complete. In conclusion, we choose

$$q \geqslant l + \log 12m.$$

(Note that for any choice of s, $\lceil \log 12m \rceil < 4 + \log \log l$ by the definition of m.)

Noting that $m \cdot O(\log l) = O(\log^2 l / \log s)$, we have:

LEMMA 2. *For any $2 \leqslant s \leqslant l$ there exists a circuit computing the l-bit inverse of an l-bit number in time $O(\log^2 l/\log s)$ and with area $O((ls)^2)$.*

The AT^2-performance of the above circuit is given by

$$AT^2 = O\left(l^2 \log^4 l \cdot \frac{s^2}{\log^2 s}\right). \tag{1}$$

By choosing the depth s as $s = l^{\varepsilon}$ $(\varepsilon > 0)$, the resulting circuit achieves $T = O((1/\varepsilon) \log l)$ and $AT^2 = O(l^{2(1+\varepsilon)})$, i.e., it is a moderately AT^2-suboptimal divider still achieving $T = O(\log l)$, for fixed ε. We are aware that this result had been previously obtained by Leighton (1985), presumably by a similar argument.

We close this section by noting that if v is an l-bit underinverse of $x \in [\frac{1}{2}, 1)$ then $v + 2^{-l}$ is an l-bit overinverse of x.

3. A Technique of Successive Refinements

We now describe an alternative approach to the computation of the inverse of an l-bit number, which uses the BCH method as a subroutine. Informally, this approach begins by computing a (small length) coarse overapproximation of the inverse of x, and subsequently refines it by multiplicative factors, which are all inverses of numbers very close to 1 (from above). Therefore, the first approximation takes advantage of the small operand length, whereas the subsequent refinements exploit the presence of leading zeros in the representation of the number to be inverted. This method is best described for an l-bit integer $x \in [1, 2)$. (Note the modified range of normalization.)

The number $x \in [1, 2)$ can be written as

$$x = x_1 + 2^{-l_1 - 2} \cdot w,$$

where x_1 is an $(l_1 + 2)$-bit number (the leading $l_1 + 2$ bits of x) and w is an $(l - l_1 - 2)$-bit number (the trailing $l - l_1 - 2$ bits of x). Then $x_1 \in [1, 2)$ and

$w \in [0, 2)$. Let $v_1 \cdot 2$ be an (l_1+1)-bit overinverse to $x_1 \cdot 2^{-1}$ (i.e., $x_1 v_1 = 1 + \eta$, $\eta < 2^{-l_1-1}$). Then

$$v_1 x = v_1 x_1 + v_1 w 2^{-l_1-2} = 1 + \eta + v_1 w 2^{-l_1-2} < 1 + 4 \cdot 2^{-l_1-2} = 1 + 2^{-l_1},$$

since $\eta < 2^{-l_1-2}$, $v_1 \leqslant 1$, and $w < 2$. This means that $v_1 x$ has at least $l_1 - 1$ consecutive 0's immediately to the right of the point. Define

$$z_2 = v_1 x.$$

Then, if v_2 denotes an approximation of $1/z_2$, we have $v_2 z_2 \simeq 1/x$. Also, if $v_2 z_2 = 1 + \eta'$ then $v_1 v_2 x = 1 + \eta'$, i.e., $v_1 v_2$ is an overapproximation of $1/x$ of precision η'. The process can be iterated thereby obtaining

$$1/x \simeq v_1 v_2 \cdots v_k.$$

This leads to the following algorithm:

ALGORITHM INVERSE 3(x)

Input: an l-bit number $x \in [1, 2)$, and an integer sequence $l_1 < l_2 < \cdots < l_k = l$.
Output: an l-bit number $v \in (1/2, 1]$, such that $vx = 1 + \varepsilon$, $\varepsilon < 2^{-l}$

```
(1)  begin  v := 1
(2)         for  i = 1 to k do
(3)              begin  t_i := leftmost (l_i + 2) bits of x;
(4)                     z_i := v t_i;
(5)                     x_i := leftmost (l_i + 1) bits of z_i;
(6)                     v_i := 2^{-1}((l_i + 1)-bit overinverse of x_i 2^{-1});
(7)                     v := v · v_i
                 end;
     end
```

The correctness of the method is established by showing that the error is bounded from above by 2^{-l}. Indeed, note that $t_k = x$, so that (line (4)) $z_k = v_{k-1} \cdots v_1 x$, and $z_k v_k = (v_k \cdots v_1) x$. But (line (5)) $x_k = z_k - \eta_k$, $\eta_k < 2^{-l-1}$ and (line (6)) $x_k v_k = 1 + \delta_k$, $\delta_k < 2^{-l-1}$. We conclude

$$z_k v_k = x_k v_k + \eta_k v_k = 1 + \delta_k + \eta_k v_k = 1 + \gamma, \gamma < 2^{-l},$$

since $\delta_k + \eta_k v_k < 2^{-l-1} + 2^{-l-1} = 2^{-l}$. This shows that $v_k \cdots v_1$ is the desired overapproximation of the inverse of x.

Step 6 is the crucial action in the above algorithm; we realize it by making use of the BCH method. To analyze its performance, we need

LEMMA 3. *If an l-bit number* $x \in [\frac{1}{2}, 1)$ *has* $l' - 1$ *zeros immediately to the right of the leading* 1, *the l-bit inverse of x can be computed in time* $T = O(\log(l/l') \cdot \log l/\log s)$ *and area* $A = O((ls)^2)$, *for any* $2 \leqslant s < l/l'$. (*Note that this result subsumes Lemma* 2 *for* $l' = 1$.)

Proof. Indeed $u = 1 - 2x$ is a (nonpositive) proper fraction whose absolute value has l' zeros immediately to the right of the point. This implies that $|u^{\lceil l/l' \rceil}| < 2^{-l}$, so that only the first $\lceil l/l' \rceil$ consecutive powers of u need to be computed. ∎

The numbers x_i, $i = 1,..., k$, used in Step 6 meet the conditions of Lemma 3, since $vt_i - 1$ is a (nonnegative) number with l_{i-1} leading zeros ($l_0 = 0$, by convention). Step 6 is therefore carried out by applying Algorithm INVERSE 2 so that the ith iteration is characterized by length l_i and depth s_i. An implementation of this technique is therefore completely specified by the two sequences:

$$l_1, l_2,..., l_k$$

and

$$s_1, s_2,..., s_k.$$

Before closing this section we note that Step 7 involves a multiplication of $O(l_i)$-bit numbers at the ith iteration; thus this operation is no more complex than the execution of the homologous Step 6, and will not be further mentioned in this discussion.

4. THE DIVIDER NETWORK

We have all the premises to illustrate in detail the structure of the divider sketched in Fig. 1.

The first J stages are collectively designed to implement the successive refinement technique; each module implements the modified BCH algorithm. For $i = 1, 2,..., J$, let l_i be the (output) operand length, s_i the depth, $A_{1,i}$ the area, and $T_{1,i}$ the time of the ith module. We seek a solution where all such modules have identical area (i.e., $A_{1,i} = A'$ for $i = 1,..., J$) and identical computation time, equal to the target time (i.e., $T_{1,i} = \theta((\log n)^{1+\varepsilon})$, $i = 1,..., J$). By the requirement of optimality, we have

$$\sqrt{A_{1,i}} = \frac{n}{T_{1,i}} = \frac{n}{(\log n)^{1+\varepsilon}}. \tag{2}$$

We also choose

$$l_i = \frac{n}{(\log n)^{1+\varepsilon} s_i}, \tag{3}$$

$$s_i = \left(\frac{n}{(\log n)^{1+\varepsilon}}\right)^{1/2(\log n)^{\varepsilon(i-1)}} \qquad (i = 1,..., J), \tag{4}$$

The parameter J is chosen as the largest value of i for which $s_i \geqslant 2$, and is readily found to be $\Theta(1/\varepsilon)$. Also note that $2 \leqslant s_J = s_{J+1}^{(\log n)^\varepsilon} < 2^{(\log n)^\varepsilon}$. Since the area of the ith module is $\Theta((l_i s_i)^2)$, condition (2) is obviously verified. From (3) and (4) we obtain $l_1 = (n/(\log n)^{1+\varepsilon})^{1/2}$, and from Lemma 2:

$$\begin{aligned} T_{1,1} &= O\left(\log l_1 \frac{\log l_1}{\log s_1}\right) \\ &= O((\log n)^2/(\log n)^{1-\varepsilon}) \\ &= O((\log n)^{1+\varepsilon}). \end{aligned}$$

From Lemma 3, for $i = 2,..., J$,

$$\begin{aligned} T_{1,i} &= O\left(\log \frac{l_i}{l_{i-1}} \cdot \frac{\log l_i}{\log s_i}\right) \\ &= O\left(\log \frac{s_{i-1}}{s_i} \cdot \frac{\log l_i}{\log s_i}\right) && \text{since } l_i s_i = l_{i-1} s_{i-1} \\ &= O\left(\log l_i \cdot \frac{\log s_{i-1}}{\log s_i}\right) && \text{since } s_i \geqslant 2 \\ &= O\left(\log n \cdot \frac{2(\log n)^{\varepsilon(i-1)}}{2(\log n)^{\varepsilon(i-2)}}\right) \\ &= O((\log n)^{1+\varepsilon}) \end{aligned}$$

thus verifying the objective for the computation time.

With these choices, each module of the chain is AT^2-optimal, and the global computation time is $c_1(1/\varepsilon)(\log n)^{1+\varepsilon} = \Theta((\log n)^{1+\varepsilon})$, for some constant c_1. The value of l_J, the number of bits of the result, is approximately

$$l_J \geqslant \frac{n}{(\log n)^{1+\varepsilon} \cdot 2^{(\log n)^\varepsilon}}.$$

This value l_J represents the length of the operand supplied to the cascade of the two Newton approximators, to be described next. Notice that, since each Newton iteration doubles the number of accurate bits, if we start with l_J accurate bits, only $(1+\varepsilon)\log\log n$ Newton iterations are needed to complete the task.

Starting with the downstream approximator, we recall (see Fig. 2) that this module is in turn the cascade of p submodules (p is an integer to be defined shortly), where the ith submodule has area and time $A_{3,i}$ and $T_{3,i}$, respectively, and

$$A_{3,i} = 2A_{3,i-1}, \qquad T_{3,i} = \sqrt{2}\, T_{3,i-1}, \qquad i = 2, 3,..., p.$$

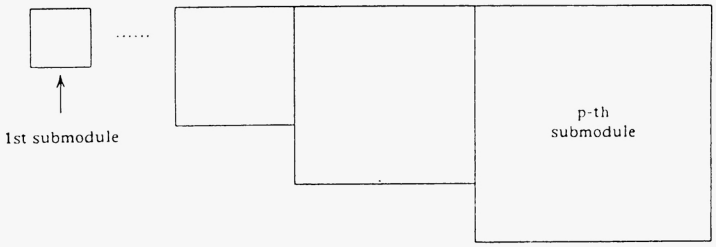

FIG. 2. The module structure of the slow "Newton approximator."

With this choice (originally proposed in (Mehlhorn, 1984)), the global area and time of the slow approximator are respectively proportional to the area $A_{3,p}$ and time $T_{3,p}$ of the pth (last) submodule. Since we are aiming for an AT^2-optimal network with computation time $O(T)$, we must have

$$A_{3,p} T_{3,p}^2 = O(n^2)$$

and

$$T_{3,p} = T.$$

This condition enables us to specify the parameter p. Indeed, the speed of the submodules increases as we proceed upstream (by decreasing submodule index), and each submodule must satisfy the condition that its multiplication time is at least logarithmic in the operand length. Since the operand length is halved in going from index i to index $i-1$ (due to the mechanism of the Newton approximation), and the most stringent condition occurs for $i=1$, we have

$$\frac{T}{(\sqrt{2})^{p-1}} \geqslant \log\left(\frac{n}{2^{p-1}}\right),$$

which is certainly satisfied if we select p as

$$p - 1 = 2 \log\left(\frac{T}{\log n}\right) = 2\varepsilon \log\log n, \tag{5}$$

or $p = 1 + 2\varepsilon \log\log n$.

Finally we turn our attention to the "fast approximator." This module receives an approximation of length $l_f \geqslant n/(\log n)^{1+\varepsilon} \cdot 2^{(\log n)^{\varepsilon}}$ and delivers an approximation of length $n/2(\log n)^{2\varepsilon}$. (Note that this is exactly the input operand length of the first module of the slow Newton approximator discussed earlier.) Thus, this module must execute at most $(\log n)^{\varepsilon} + (1-\varepsilon)\log\log n$ iteration steps, each of them within time $\Theta(\log n)$. The module essentially consists of a "fastest" multiplier, i.e., time $O(\log n)$,

of numbers of length $n/(\log n)^{2\varepsilon}$, and can be realized with area A_2 such that $A_2(\log n)^2 = \Theta((n/(\log n)^{2\varepsilon})^2)$ and hence $A_2 = \Theta((n/(\log n)^{1+2\varepsilon})^2)$. Thus, the resulting AT^2-measure for this module is

$$A_2 T^2 = \Theta\left(\left(\frac{n}{(\log n)^{1+2\varepsilon}} \cdot (\log n) \cdot (\log n)^{\varepsilon}\right)^2\right) = O(n^2)$$

and the optimality condition is clearly satisfied.

Since each of the three major units of our divider—the chain of modified BCH dividers, the fast Newton approximator and the slow Newton approximator—has area $O((n/(\log n)^{1+\varepsilon})^2)$ and time $O((\log n)^{1+\varepsilon})$, we conclude with

THEOREM 1. *For any fixed* $1 \geqslant \varepsilon > 0$, *the n-bit inverse of an n-bit number can be calculated with optimal* AT^2*-performance for any* $T \in [\Omega((\log n)^{1+\varepsilon}), O((\log n)^2)]$.

5. CONCLUSION

We constructed an AT^2-optimal divider with computation time $(\log n)^{1+\varepsilon}$ for any $\varepsilon > 0$. The reader may wonder whether one can choose ε as a decreasing function of n (tending to zero as n goes to infinity). This is indeed the case if the construction is slightly modified. In the construction as it is now we use a chain of modified BCH dividers each with the same area and speed. Thus both area and time grow as $1/\varepsilon$ and hence AT^2 grows (at least) as $(1/\varepsilon)^3$.

If ε is chosen as a function of n, then this simple chain of equally sized modules does not suffice. Rather one has to use a chain of increasingly larger (and slower) modules as we did for the Newton iteration. Omitting the tedious and not particular illuminating details we have

THEOREM 2. *There is an* AT^2*-optimal divider for n-bit integers for any* $T \in [\Omega(\log n \cdot 2^{(\log\log n)^{3.4}}), O((\log n)^2)]$.

Note that $2^{(\log\log n)^{3.4}} = O((\log n)^{\varepsilon})$ for any $\varepsilon > 0$.

RECEIVED August 1985; ACCEPTED October 1986

REFERENCES

MEHLHORN K. (1984), AT^2-optimal VLSI integer division and integer square rooting, *Integration* **2**, 163–167.

REIF, J. (1983), Logarithmic depth circuits for algebraic functions, *in* "Proceedings, IEEE 24th Found. of Comput. Sci.," pp. 138–145.

BEAME, P. W., COOK, S. A., AND HOOVER, H. J. (1984), Log depth circuits for division and related problems, *in* "Proceedings, IEEE 25th Found. of Comput. Sci.," pp. 1–6.

LUK, W. K., AND VUILLEMIN, J. (1983), Recursive implementation of optimal time VLSI integer multipliers, *in* "VLSI 83," Trondheim, Norway.

SPANIOL, O. (1976), "Arithmetik in Rechenanlagen," Teubner, Stuttgart.

MEHLHORN, K., AND PREPARATA, F. P. (1983), AT^2-optimal VLSI integer multiplier with minimum computation time, *Inform. and Control* **58**, Vol. 1–3, 137–156.

KNUTH, D. E. (1981), "The Art of Computer Programming, Vol. 2., Seminumerical Algorithms," 2nd. ed., Addison–Wesley, Reading, Ma.

LEIGHTON, F. T. (1985), personal communication, May.

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium