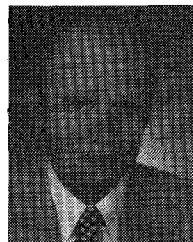




Rupert Kling was born in Bayrischzell, Germany, on December 12, 1944. He received the M.S. degree in electrical engineering from the Technical University of Munich, Munich, West Germany, in 1973.

In 1973 he joined the Siemens Component Division, Munich, West Germany. He was working in the field of test program generation for various MOS-LSI circuits. At present, he is a member of a design team for NMOS IC's.



Heinz Schulte was born in Meschede, Germany, on August 17, 1937. He received the diploma in physics from the Technische Hochschule, Aachen, West Germany, in 1968.

Since 1968 he has been employed as a Development Engineer at the semiconductor factory of Siemens Company, Munich, West Germany. He has been engaged in the development of silicon-gate technology for memory devices.

NORA: A Racefree Dynamic CMOS Technique for Pipelined Logic Structures

NELSON F. GONCALVES, STUDENT MEMBER, IEEE, AND HUGO J. DE MAN, SENIOR MEMBER, IEEE

Abstract—This paper describes a new dynamic CMOS technique which is fully racefree, yet has high logic flexibility. The circuits operate racefree from two clocks ϕ and $\bar{\phi}$ regardless of their overlap time. In contrast to the critical clock skew specification in the conventional CMOS pipelined circuits, the proposed technique imposes no restriction to the amount of clock skew. The main building blocks of the NORA technique are dynamic CMOS and C^2 MOS logic functions. Static CMOS functions can also be employed. Logic composition rules to mix dynamic CMOS, C^2 MOS, and conventional CMOS will be presented. Different from Domino technique, logic inversion is also provided. This means higher logic flexibility and less transistors for the same function. The effects of charge redistribution, noise margin, and leakage in the dynamic CMOS blocks are also analyzed. Experimental results show the feasibility of the principles discussed.

I. INTRODUCTION

IN the conventional CMOS technique there is an inherent redundancy of information. For each n-type device there is a corresponding p-type device. In fact, a complete logic function is built with the n devices and repeated with the p devices. As a consequence of this approach, substantial amounts of silicon are wasted, especially for complex logic. Also, power dissipation and speed are degraded by the extra area and extra transistors.

Another important problem of CMOS technique is clock races in pipelined circuits. To latch the information between two pipelined sections, transmission gates are usually employed. In

Manuscript received November 8, 1982; revised January 18, 1983. This work was supported in part by Fundação de Amparo a Pesquisa do Estado de São Paulo, Brazil.

The authors are with the Department Elektrotechniek-ESAT, Katholieke Universiteit Leuven, B 3030-Heverlee, Belgium.

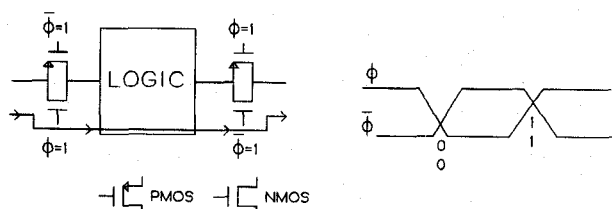


Fig. 1. Signal races in CMOS pipelined circuits.

CMOS logic, these transmission gates are generally implemented with p-n gates in parallel and controlled by clocks ϕ and $\bar{\phi}$, as shown in Fig. 1. The use of single gates (p- or n-type) is to be avoided in CMOS due to power dissipation and low noise margin as a result of clock feedthrough and bulk effect. CMOS p-n transmission gates, controlled by clocks ϕ and $\bar{\phi}$, suffer from signal races. As depicted in Fig. 1, this results from unavoidable overlap of the clock phases during the clock transitions. During the phase overlaps, all the transmission gates are switched on, which may cause illegal flow of information, depending on the ratio between the gate delay and the clock skew.

This race problem is usually bypassed by a careful synchronization of the two clock phases within a small fraction of the gate delay (a few nanoseconds). This skew clock control is extremely difficult, especially for high speed technologies, for unmatched clock loads or for distributed clock VLSI circuits [1]. This leads to highly critical and unstable designs. A possible solution to the clock race is the use of four clock phases which, however, requires too much silicon area.

To overcome the redundancy of information in the conventional CMOS, dynamic circuit schemes have been proposed in

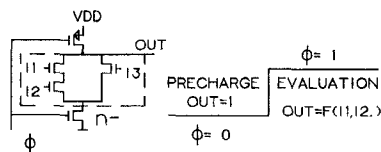


Fig. 2. An n-type dynamic CMOS logic block.

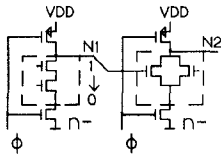


Fig. 3. Internal delay race problem.

the literature [2]–[4]. Fig. 2 shows the dynamic CMOS building block. The desired logic function is implemented using only n-type devices. The logic tree is connected to V_{DD} and ground through clocked transistors. There are two modes of operation. First, for phase $\phi = 0$ the output node is precharged to a high level while the current path to ground is turned off. Then, for phase $\phi = 1$, the path to the high level is turned off by the clock and the path to ground is turned on. Therefore, depending on the state of the inputs, the output node will either float at the high level or will be pulled down.

A clear advantage of this CMOS dynamic block is the reduced silicon area. Whereas there are $2n$ transistors in a conventional n-input CMOS gate, the dynamic configuration needs only $n + 2$. Also due to the smaller area and consequently smaller capacitances, power dissipation and speed are, in principle, improved by the dynamic approach.

A strong limitation of this dynamic structure is the impossibility of cascading the logic blocks for implementing complex logic. Consider, for instance, the circuit in Fig. 3. During the precharge phase, nodes $N1$ and $N2$ are set up to the high level "1." In the evaluation phase ($\phi = 1$), internal delay in block 1, associated with a "1" \rightarrow "0" transition of node $N1$, can cause an incorrect discharge of node $N2$. This occurs because, during the evaluation phase and while node $N1$ is still "1," there is a direct path between node $N2$ and ground. When this path is eliminated by the effective transition of node $N1$ to "0," the precharge information of node $N2$ could already be gone. We define such a race as the "internal delay problem."

In the Domino technique, Krambeck *et al.* [4] have solved the internal race by placing a static inverter after every dynamic block, as indicated in Fig. 4. During the precharge phase, the outputs of all the static inverters are set up to a low level. Consequently, all the n-type transistors driven by these inputs are set up to an OFF condition. Now, during the evaluation phase, internal delays cannot incorrectly discharge the dynamic storage nodes since during the entire delay period the path to ground is turned off.

A limitation of the Domino technique is the lack of inverted signals. The combination of the dynamic block with the static inverter gives a noninverted signal. This decreases logic flexibility and, therefore, usually requires more transistors for a given logic function. Besides this inconvenience, no provisions are made to overcome the clock race problem.

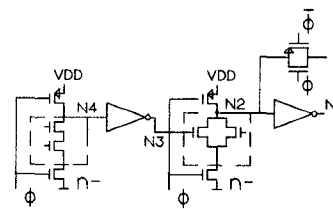


Fig. 4. Domino circuit.

In the next section, a new technique called NORA is presented which overcomes the above deficiencies. In Section III, the properties of the NORA CMOS technique are analyzed and proved. Logic composition rules to mix dynamic, static, and C^2 MOS [5] logic functions are also derived. The dynamic CMOS limitations are described in Section IV. Experimental results and major conclusions are presented in Sections V and VI, respectively.

II. NORA CMOS TECHNIQUE

The main building blocks of NORA technique are shown in Fig. 5. The logic functions are implemented using n-type and p-type dynamic CMOS and C^2 MOS blocks. Conventional (static) CMOS function blocks can also be eventually employed. Logic composition rules to combine these functions, preserving the racefree properties, will be presented in Section III. As it will further be shown, to guarantee a fully racefree operation in pipelined circuits, the storage of information must always be performed by a C^2 MOS function block (C^2 MOS latch stage). In a previous paper [6], the NORA (*NO RAcce*) technique was called n-p-CMOS, due to the possible employment of n- and p-dynamic blocks. We decided to change the name because the p-dynamic block is not essential to the racefree principle; it is only used to increase the logic flexibility.

The pipelined circuit in Fig. 5 is defined as a ϕ -section. For phase $\phi = 0$ ($\bar{\phi} = 1$), the ϕ -section is in the precharge phase. The outputs of all the n- and p-dynamic blocks are precharged to "1" and "0," respectively. Also during this phase, the ϕ -section inputs are in a sampling mode, i.e., these inputs are set up.

For phase $\phi = 1$ ($\bar{\phi} = 0$), the ϕ -section is in the evaluation phase. The ϕ -section inputs are held constant, and the outputs of all the dynamic blocks are evaluated as a function of the ϕ -section inputs and of the internal inputs¹. From these output results, those which must be transferred to the next pipelined section are stored in C^2 MOS latch stages.

In the circuit of Fig. 5, notice the following characteristics (see Section III).

1) Inverted and noninverted signals are provided. When direct coupling between dynamic blocks is desired, the logic function is implemented by alternating p- and n-logic blocks. If the inverter is required, a Domino like connection is employed, i.e., sequences of the same block type are used (n-

¹For convenience, the inputs of a dynamic block have been separated into section inputs and internal inputs. The section inputs are set up during the precharge phase. The internal inputs are set up during the evaluation phase. For instance, in Fig. 5:

IN —section inputs
 $N 2, N 3$ —internal inputs.

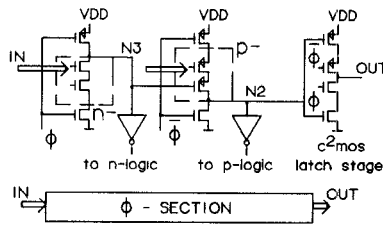
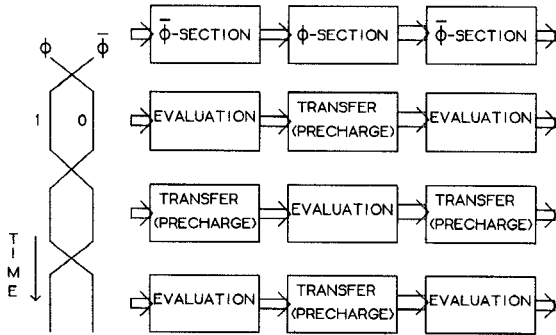
Fig. 5. NORA-CMOS pipelined circuit— ϕ -section.

Fig. 6. Pipelined system.

inverter-n or p-inverter-p). Compared with the Domino technique, this means higher logic flexibility and less transistors for the same function.

2) n-p as well as p-n sequences are possible and the sequences can be of arbitrary logical depth. Therefore, many logic levels can be operated in only half a clock period.

By interchanging ϕ and $\bar{\phi}$ in the circuit of Fig. 5, a $\bar{\phi}$ -section is obtained. A sequence of ϕ - and $\bar{\phi}$ -sections makes a pipelined system, as shown in Fig. 6. For phase $\phi = 0$ $\bar{\phi} = 1$, the ϕ -sections are precharged while the $\bar{\phi}$ -sections are in the evaluation phase. The ϕ -section outputs are held constant by the C^2 MOS latch stages. Then, for phase $\phi = 1$ $\bar{\phi} = 0$, the ϕ -sections are in the evaluation phase and the $\bar{\phi}$ -sections are precharged. Now, the $\bar{\phi}$ -section outputs, evaluated in the previous phase, are held constant in such a way that the ϕ -sections can use this information to compute the corresponding results. In this way, there is a complete flow of information; with the information travelling from one $\bar{\phi}$ -section to the next ϕ -section, from this to the next $\bar{\phi}$ -section, and so on.

III. NORA RACEFREE PROPERTIES AND LOGIC COMPOSITION RULES

In this section the racefree properties of the NORA technique will be carefully analyzed. Logic composition rules to combine dynamic, conventional, and C^2 MOS function blocks will also be derived.

A. Internal Delay Racefree Property

The internal delay racefree property is defined as the capability of the dynamic block to keep its precharge signal during the delay time of the previous blocks to set up the internal inputs. It is easy to prove that a dynamic block will have the internal delay racefree property if the following conditions occur:

- 1) During the precharge phase, the internal inputs are set up in such a way they cut off their corresponding transistors.
- 2) During the evaluation phase, the internal inputs are glitch-free, i.e., these inputs can make only one transition.

From the above conditions the following results can be derived:

a) When the number of “static” inversions between two dynamic blocks is even, complementary type of logic blocks must be used for these two blocks (n-p or p-n). For instance in Fig. 5, this corresponds to alternate p- and n-logic blocks when the direct coupling between dynamic blocks is desired.

b) The same type of dynamic blocks (n-n or p-p) must be used when the number of “static” inversions is odd. In Fig. 5, this corresponds to Domino-like connections: n-inverter-n or p-inverter-p.

c) Normally, after mixing dynamic blocks with static CMOS, the circuit should be kept static up to the C^2 MOS latch stage. Static functions can also be used after the C^2 MOS stage. This should be done because in general “static” functions driven by dynamic blocks are not glitch-free. (Exceptions are the inverter and in some cases the NAND, NOR . . .)

These logic composition rules can easily be implemented in a CAD system like Dialog [7] for automatic checking of the logic design consistency.

B. Clock Racefree Properties

As indicated in Fig. 6, to have a working pipelined system the results generated during the evaluation phase must be held constant until the end of the transfer phase. The latched information should not be altered by the precharge signal or by input variations. It will now be proven that after the evaluation phase a NORA pipelined section keeps its output results in spite of high-high or low-low clock overlaps (clock skew).

For simplicity, let us initially consider that all the circuits in the pipelined section are built only with dynamic blocks; the two exceptions being the C^2 MOS latch stage and the static inverter for connecting complementary dynamic blocks. For this circuit, two possible cases should be analyzed.

Case I—Precharge Racefree

During the evaluation phase, the dynamic block which precedes the C^2 MOS latch stage has its precharge signal modified by the inputs. Such a situation is indicated in Fig. 7 for an n-type and a p-type dynamic block.

As indicated in Fig. 7, the alteration of the output information is controlled by only one of the phases ϕ or $\bar{\phi}$. Therefore, these outputs are not influenced by the other phase. The outputs are, for instance, completely immune to the overlap of the phases. This kind of output latch control by only one phase (ϕ or $\bar{\phi}$) is completely different from the conventional case with transmission gates, where the output latch is controlled simultaneously by the two phases ϕ and $\bar{\phi}$. In contrast to the critical clock skew specification of the conventional transmission gates (few nanoseconds), the NORA technique imposes no restriction to the amount of clock skew.

Note: Although the NORA circuit is immune to the overlap

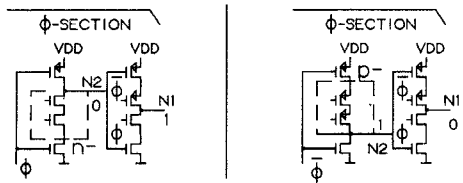


Fig. 7. Precharge racefree—precharge signal altered by the inputs:

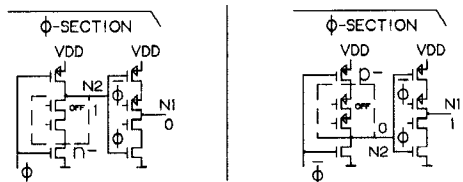
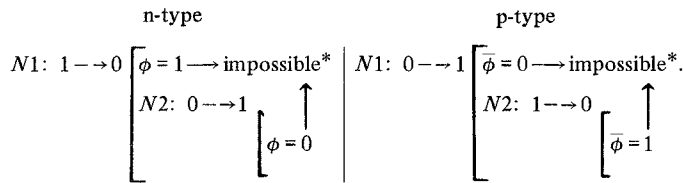


Fig. 8. Input variation racefree—precharge signal kept by the inputs.

of the clock phases, there could still be signal races for clock signals with very slow rise and fall times (10 to 20 times the gate delay). In contrast to the clock overlap race this kind of race is easily eliminated since it does not require control of the clock skew and, also, the available time margin is much larger.

Case II—Input Variation Racefree

The other possible case, i.e, when the dynamic block keeps the precharge signal, is illustrated in Fig. 8.

If the dynamic block keeps the precharge signal, at least one of the logic transistor should be driven off. If this transistor is controlled by an internal input, the dynamic block which generates this input has also kept its precharge signal. This occurs because the internal inputs are precharged in such a way that the corresponding driven transistors are off. Therefore, there must be at least one sequence of dynamic blocks with precharge signals preserved. Fig. 9 depicts this sequence. Again, as shown in Fig. 9, the alteration of the output information is controlled by only one of the phases ϕ or $\bar{\phi}$. Therefore, they are not influenced by the overlap of the phases.

For the case being analyzed, the racefree property has been derived from the interrelation between a dynamic CMOS block and a C²MOS latch stage. Let us now show that the input variation racefree property can also be derived by the action of two C²MOS stages: "A NORA pipelined circuit is input variation racefree if the total number of inversions (static and dynamic) between two C²MOS latch stages is even." The proof is indicated in Fig. 10. This racefree property can also be used to solve the clock race condition of some conventional CMOS circuits. An important circuit which can be built using the above property is the shift register.

Combining the racefree properties derived from two C²MOS functions and from C²MOS with dynamic block, the following result can be proven.

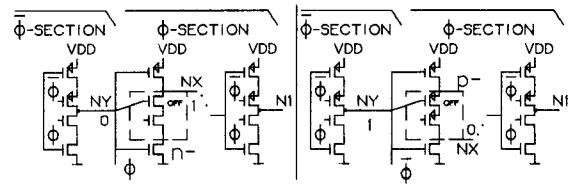


Fig. 9. Input variation racefree—sequence of dynamic blocks with precharge signals kept by the inputs:

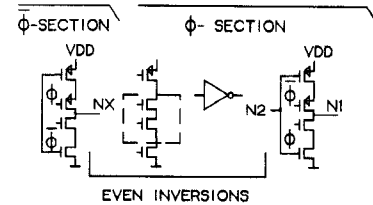
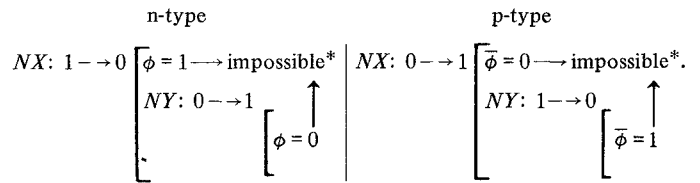
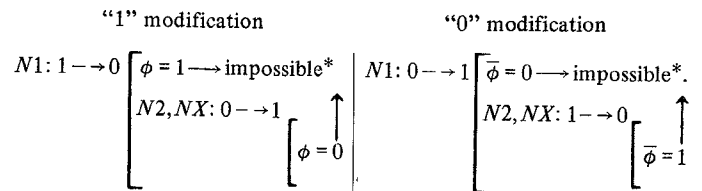


Fig. 10. Input variation racefree—even inversions between two C²MOS latch stages:



Consider a NORA pipelined section, built with dynamic, conventional, and C²MOS function blocks. Consider all the chains of function blocks of this pipelined section, starting in a C²MOS input stage (C²MOS latch stage of the previous pipelined section) and ending in a C²MOS output latch stage. The NORA pipelined section is clock racefree if, for every chain, the following conditions are satisfied:

- 1) Precharge racefree:
 - a) There is an even number of inversions between the C²MOS output stage and the last dynamic block (see Fig. 11).
 - 2) Input variation racefree:
 - b1) There is a dynamic block in such a way that there is an even number of inversions between this dynamic block and the C²MOS input stage (see Fig. 12); or
 - b2) the total number of inversions between the two (input, output) C²MOS stages is even (see Fig. 10).

If the pipelined section does not satisfy the clock race conditions, generally, circuit modifications can be easily included. By way of example, consider the nonracefree pipelined section indicated in Fig. 13(a). For this example, the following circuit modifications would eliminate the race condition:

- 1) conversion of one static function to dynamic function [see Fig. 13(b)];
- 2) conversion of one static function to C²MOS function [see Fig. 13(c)];
- 3) placement of one static function after the C²MOS latch

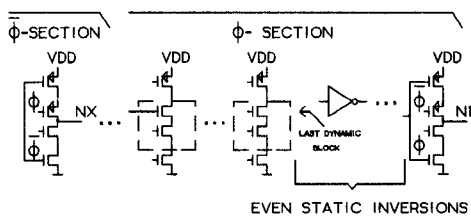


Fig. 11. Precharge racefree—even inversions between the C²MOS output stage and the last dynamic block.

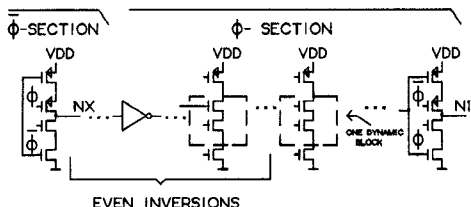


Fig. 12. Input variation racefree—even inversions between the C²MOS input stage and one dynamic block.

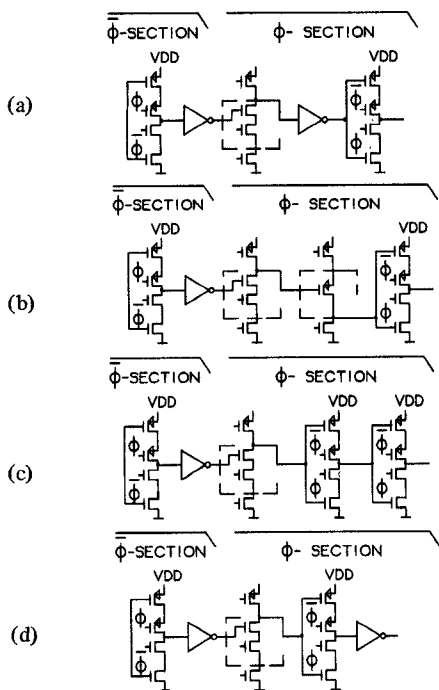


Fig. 13. Elimination of signal races. (a) Circuit with race of signals. (b) Conversion to dynamic CMOS function. (c) Conversion to C²MOS function. (d) Placement after the C²MOS output stage.

stage, provided that the racefree property of the next pipelined section would not be destroyed [see Fig. 13(d)].

IV. DYNAMIC CMOS LIMITATIONS

In this section the limitations of the NORA technique will be presented. These limitations are directly related to the dynamic storage of information and, therefore, they are common to all the dynamic techniques.

A. Charge Redistribution

The output signal of the dynamic blocks relies on storage nodes. As indicated in Fig. 14, by commutation of an OFF transistor to an ON state, a charge redistribution effect may ap-

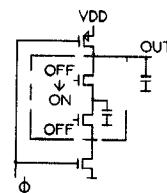


Fig. 14. Charge redistribution in dynamic blocks.

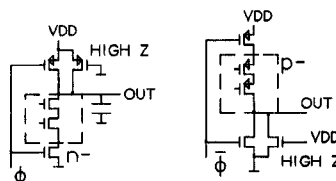


Fig. 15. Dynamic CMOS for low operating frequency.

pear between the output capacitance and the parasitic logic tree capacitances.

Normally, there will be no charge redistribution between the precharged node and the logic tree nodes controlled by section inputs. This occurs because these inputs are set up during the precharge phase and, therefore, the logic tree nodes will also be precharged. Yet, some charge redistribution effect will exist, if the precharge period after input set up is too small. This extra period of precharge generally does not result in speed limitation for the pipelined system due to the small capacitances of the logic trees.

For the internal inputs, such attenuation of the charge redistribution does not exist, since these inputs are set up only after the precharge period. In this case, the charge redistribution must be minimized by layout and by proper logic tree arrangement. The transistors driven by internal inputs must be placed as far as possible from the output storage node.

B. Leakage and Noise Margin

Another limitation of the dynamic CMOS techniques is the leakage of the storage nodes. Due to clock feedthrough, power supply variation, noise, etc., the inputs of the dynamic block can be altered from the ideal zero and V_{DD} values. Consequently, the logic transistors are driven to weak inversion. This leakage effect imposes a limit to the lowest operating frequency and to the noise margin of the circuit.

For lower frequency applications, a possible solution [8] is the addition of a high impedance transistor, as shown in Fig. 15.

V. RESULTS

In Fig. 16, a microphotograph of the chip designed to characterize the NORA technique is shown. It contains serial full adders, subtractors, shift registers, a 4-bit serial-parallel multiplier and some special structures to analyze charge redistribution, leakage, and clock feedthrough. Fig. 17 shows a NORA serial full adder containing only 32 transistors. The pipelined output sum is generated using only 20 transistors, compared with 28 if conventional CMOS is employed. The circuit area is 130 × 318 μm², giving a density of 770 transistors/mm² in a 5 μm technology, which compare favorably with an NMOS solution. The threshold levels of the n-type and p-type devices

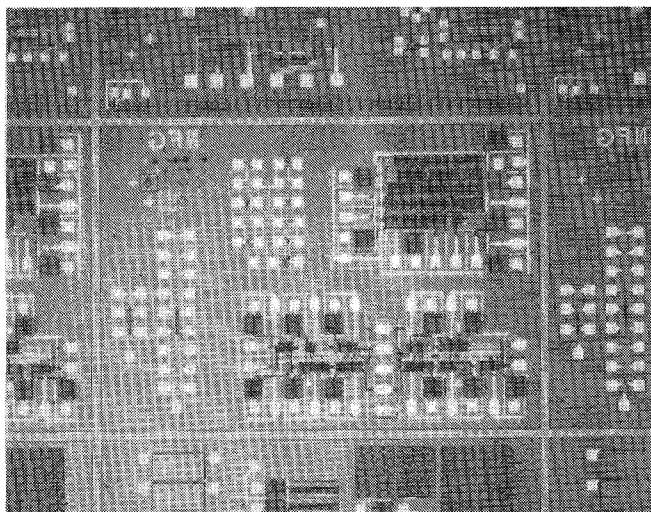


Fig. 16. Microphotograph of the chip to characterize the NORA-CMOS technique.

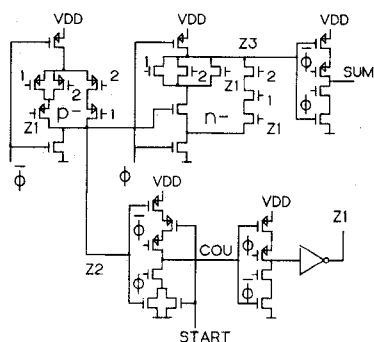


Fig. 17. NORA-CMOS serial full-adder-circuit diagram.

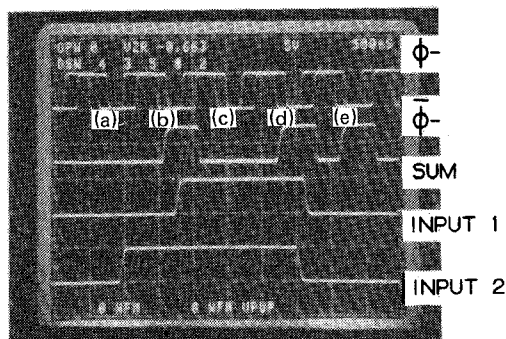


Fig. 18. NORA-CMOS serial full-adder-experimental results for very large clock skew. (a) $0+0+0$. (b) $1+0+0$. (c) $1+1+0$. (d) $1+1+1$. (e) $0+0+1$.

are $+1$ V and -1 V, respectively. Fig. 18 shows experimental results for very large clock skew = 150 ns at 1 MHz clock frequency, without disturbing the circuit operation. Notice that during the evaluation phase $\phi = 1$ $\bar{\phi} = 0$, the results are obtained and then hold constant until the end of the transfer phase $\phi = 0$ $\bar{\phi} = 1$. Also from experimental results, the minimum working frequency was less than 1 kHz at room temperature, indicating that the current leakage due to weak inversion is not a critical limitation. The measured power dissipation of the serial full-adder was $17 \mu\text{W}/\text{MHz}$ at a supply voltage of 5 V. The circuits were designed for a maximum operating frequency of 14 MHz, and the devices have been tested on the wafer probe up to 10 MHz. More careful measurements about speed and noise

margin are under investigation and will be presented in a later publication.

VI. CONCLUSION

A new dynamic CMOS technique has been presented. The NORA technique provides high logic flexibility, high speed circuits, and compact chip areas. A new concept of latch control by only one clock phase was theoretically and experimentally demonstrated. By this concept the critical clock skew specification of the conventional CMOS technique is completely eliminated. This simplifies the design and greatly increases the reliability, feasibility, and testability of CMOS circuits. The NORA technique also provides very high density layouts, which compare favorably with NMOS solutions.

REFERENCES

- [1] M. Shoji, "Electrical design of BELLMAC-32A microprocessor," in *Proc. IEEE Int. Conf. Circuits Comput.*, 1982, pp. 112-115.
- [2] W. M. Pensey and L. Lau, *MOS Integrated Circuits*. New York: Van Nostrand, 1972, pp. 260-282.
- [3] E. Hebenstreit and K. Horninger, "High-speed programmable logic arrays in ESFI SOS technology," *IEEE J. Solid-State Circuits*, vol. SC-11, pp. 370-374, June 1976.
- [4] R. H. Krambeck, C. M. Lee, and H. S. Law, "High-speed compact circuits with CMOS," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 614-619, June 1982.
- [5] Y. Suzuki, K. Odagawa, and T. Abe, "Clocked CMOS calculator circuitry," *IEEE J. Solid-State Circuits*, vol. SC-8, pp. 462-469, Dec. 1973.
- [6] N. F. Goncalves and H. De Man, "n-p-CMOS: A racefree dynamic CMOS technique for pipelined logic structures," in *ESSCIRC Dig. Tech. Papers*, Sept. 1982, pp. 141-144.
- [7] H. De Man, D. Dumlugol, P. Stevens, G. Schrooten, and I. Bolsens, "Logmos: A transistor oriented logic simulator with assignable delays," in *Proc. IEEE Int. Conf. Circuits Comput.*, 1982, pp. 42-45.
- [8] R. G. Stewart, "High density CMOS ROM arrays," *IEEE J. Solid-State Circuits*, vol. SC-12, pp. 502-506, Oct. 1977.



Nelson F. Goncalves (SM'81) was born in Brazil on September 26, 1952. He received the B.S. and M.S. degrees from the University of São Paulo, São Paulo, Brazil, in 1975 and 1978, respectively.

From 1976 to 1978, he worked on the characterization of the interface Si/SiO₂. In September 1980, he joined ESAT, Katholieke Universiteit Leuven, where he is studying for the Ph.D. degree in electrical engineering. He is presently working on the design of dynamic CMOS circuits. He has a Fellowship from FAPESP, Brazil.



Hugo J. De Man (M'81-SM'81) was born in Boom, Belgium, on September 19, 1940. He received the electrical engineering degree and the Ph.D. degree in applied science from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1964 and 1968, respectively.

In 1968 he became a Member of the Staff of the Laboratory for Physics and Electronics of Semiconductors at the University of Leuven, working on integrated circuit technology. From 1969 until 1971, he was at the Electronic Research Laboratory, University of California, Berkeley, as an ESRO-NASA Postdoctoral Research Fellow, working on computer-aided devices and circuit design. In 1971 he returned to the University of Leuven as a Research Associate of the Belgian National Science Foundation (NFWO). In 1974 he became a Professor at the University of Leuven. From 1974-1975, he was Visiting Associate Professor at the University of California, Berkeley. His current field of research is the design of integrated circuits and computer-aided design.