15.0 Custom Integrated Circuits

Academic and Research Staff

Prof. J. Allen, Prof. B. Musicus, Prof. J. White, Prof. J.L. Wyatt, Jr.

Graduate Students

R.C. Armstrong, D.G. Baltus, C.S. Bamji, H. Jeong, K. Kundert, A. Lumsdaine, L.M. McCormick, S.P. McCormick, H. Miyanaga, K. Nabors, P. O'Brien, G.N.S. Prasanna, M. Reichelt, W.S. Song, D. Standley, F. Van Aelten

Undergraduate Students

J. Deroo, K. O'Connor

15.1 Custom Ingegrated Circuits

Analog Devices, Inc.
International Business Machines, Inc.
Joint Services Electronics Program (Contract DAAL03-86-K-0002)
U.S. Air Force - Office of Scientific Research (Grant AFOSR 86-0164)

Jonathan Allen, Robert C. Armstrong, Donald G. Baltus, Cyrus S. Bamji, Lynne M. McCormick, Steven P. McCormick, Hiroshi Miyanaga, Mark Reichelt, Filip Van Aelten.

The overall goal of VLSI CAD research is to provide the means to produce custom integrated circuits correctly, quickly, and economically. In the past, correctness applied only to the desired function, but there is increasing need to design to a performance specification, expressed in terms of speed, circuit area, and power. In this research group, the main emphasis is on CAD tools for performance-directed synthesis, with particular emphasis on digital signal processing applications. This goal implies the development of algorithms for optimizing performance of the total design. These complete designs, however, are specified at several levels of abstraction, ranging from function through architecture, logic, circuit, and layout. Traditionally, optimization techniques have been applied within a single such abstraction, but total optimization implies the simultaneous specification of all levels of representation such that the desired performance goal is realized. To facilitate this process, each such representation must be constructed so that optimization algorithms can be effectively designed, often utilizing well-understood methods. Furthermore, these representations must be coordinated so that each represents a projection of a single overall design. This "consistency" requirement guarantees that the distinct levels of abstraction can all be regarded as views of one abstract underlying design object.

The major research emphases in performance directed synthesis follow from the above observations. Algorithms for design at the several levels of representation are needed, but means for characterizing performance in several aspects is also required. Furthermore, the role of formal representational techniques is central, both for specify-

ing well-formedness at each design level, and for providing the requisite framework for consistency maintenance techniques. In the following paragraphs, several research projects are described which address these concerns.

Given a circuit specification, there is a need for a flexible compiler which will produce optimized layout automatically subject to several parameter settings. These parameters specificy the location of input and output pins, as well as feedthrough busses, and the desired aspect ratio. Baltus[1,2] has constructed an algorithm for this purpose that converts both NMOS and CMOS circuits to compact layouts, while observing the pin location and aspect ratio constraints. These techniques also allow for widely varying device size, grouping devices conveniently in terms of topological connectivity, device type (e.g., N or P) and device size. Diffusion breaks are minimized, and unequal numbers of N and P devices are allowed for CMOS circuits. This program is of great interest because it is highly flexible and not constrained to a rigid layout architecture, but still manages to produce high-quality layouts. The output of the program is in symbolic form, which can then be compacted to final geometrical specifications. This algorithm can be seen as a major contribution to the overall goal of moving the designer's focus away from the detailed layout to higher levels of design, which are more directly meaningful in terms of design goals. As such it encapsulates powerful algorithmic methods for the manipulation of layout artwork, exceeding the capability of human designers for medium and large sized cells.

As the minimum dimensions of layouts decrease, noise coupling between lines becomes an increasing problem, requiring accurate modeling of electromagnetic coupling and estimation of noise signal magnitudes. McCormick,[3] building on earlier models for resistance and capacitace associated with layout, has developed a new general representation. In the past, waveform bounding, higher order approximations (up to the second order moment of the impulse response), and macromodeling have been used to capture a variety of aspects of waveforms, but this new generalized representation subsumes all of these, permitting a variable order of representation. In particular, noise waveforms require third and fourth order moments of the impulse response. The generalized representation provides a uniform representation for all CAD analyses, providing both flexibility and variable accuracy. Use of the waveform moment representation permits efficient computation of noise coupling "hot spots" in a circuit layout, thus permitting the analysis of layouts for quality of (or relative lack of) noise coupling. Clearly, advances in technology now make this additional well-formedness check mandatory, and a part of the overall assessment of an integrated circuit design.

Just as the waveform moment representation is useful for noise coupling analysis, macromodeling techniques have been introduced for circuit delay modeling. Techniques have been developed by Brocco[4] for modeling of CMOS circuits for time delay, which take into account the rise (or fall) time of the input waveform(s), the output capacitive loading, and the nature of the circuit itself. Such modeling had been introduced before, giving accuracy within 5% of SPICE results, but the modeling of transmission gates (which introduce two time constants) is novel as is the method by which all gate models are combined. Gate resistances themselves are macromodeled, instead of being represented as constant values. In fact, four different types of macromodeled resistances are introduced in order to provide the desired accuracy. The resulting models are both general and accurate, and represent the desired tradeoff between accuracy and computational efficiency.

When extremely high accuracy of simulation is needed, especially for high frequency circuits, then even macromodeling of device and circuit action is insufficient, and the drift-diffusion based partial differential equation approximation for electron transport must be used rather than lumped models usually employed in circuit simulators. Unfortunately, the computations required for such analyses are very extensive. Nevertheless, Reichelt, White and Allen[5] have recently investigated the possibility of accelerating the transient simulation of MOS devices by using waveform relaxation. Convergence has been shown for practical cases, and the discretization needed for two-dimensional MOS transient device simulation is now being investigated.

For some time, we have been searching for appropriate formalisms for the verification of VLSI layouts and circuit schematics. By introducing formal grammars, correctness is tied to a rigorous set of composition rules which govern how blocks of layout and circuit schematic may be combined. These layouts and circuit schematics are represented as graphs, so that the composition rules are defined as graph transformations. Bamji[6] has shown how individual composition rules can span both layout and circuit schematic graphs, making verification of the layout to circuit schematic correspondence possible. These grammatical composition rules permit incremental verification, and the introduction of modifications during the analysis process with minimal overhead.

Van Aelten[7] has demonstrated the use of the grammatical formalism for a variety of circuit styles. Now he is starting to extend these techniques to the logic domain through functions for mapping objects at the circuit level to objects on the logic level. These mappings are similar to valuation functions used in denotational semantics. These functions efficiently describe the abstraction of all legitimate circuit schematic structures to the logic level, utilizing the recursiveness of the grammar. This approach also captures the notion of tolerant well-formedness requirements by moving some constraints now employed at the circuit schematic level to the logic level and its behavioral representation.

While formal grammatical techniques are useful for characterizing (and recognizing) well-formed structures at the various levels of representation, there remains the problem of consistency maintainence between these levels in a design database. Armstrong[8] is developing a formal model which can serve as the basis for incremental database consistency maintainence, utilizing basic set theoretic mappings. The consistency problem can be formally stated using these mappings, and graph grammars can describe the fundamental structures to be manipulated in maintaining consistency. These techniques are expected to provide the basis for correct translations in design synthesis, so that all design views remain consistent at all points of the design evolution, including the exploration of the space of all correct designs in order to find desired performance configurations.

From the above project descriptions, it is clear that many representational techniques are being studied as part of the overall task of building CAD programs for high performance circuit design. While these studies address the area of circuit performance, architectural performance (or parallelism) must also be used when the intended task algorithms permit. Miyanaga is studying the design of a new basic element for multiprocessor systems aimed at digital signal processing tasks. Mapping algorithms for the distribution of the overall algorithm on the several processors are also being developed in a coordinated way, so that optimal performance can be readily obtained for a wide variety of digital signal processing tasks using specialized mapping transformations.

Starting from the given functional task, the studies described above aim at providing the needed design representations, the means to specify their correctness and inter-domain consistency, and the context for design exploration needed to find high performance solutions to computationally demanding tasks. Present goals do not include the construction of an overall design system, but are instead focused on fundamental formal representational issues central to efficient and correct CAD systems aimed at high-performance design.

References

[1] D.G. Baltus, *Generating Efficient Layouts from Optimized MOS Circuit Schematics.* RLE Technical Report No. 535. MIT, 1988.

[2] D.G. Baltus and J. Allen, "SOLO: A Generator of Efficient Layouts From Optimized MOS Circuit Schematics" *Proceedings of the 25th Design Automation Conference,* June 1988.

[3] S.P. McCormick, *Cross-Talk Noise Modeling for VLSI Interconnections,* Ph.D. diss. proposal, Dept. Electr. Eng. and Comp. Sci., MIT, 1987.

[4] L.M. Brocco, *Macromodeling CMOS Circuits for Timing Simulation.* RLE Technical Report No. 529. MIT, 1987.

[5] M. Reichelt, J. White, J. Allen, and F. Odeh, "Waveform Relaxation Applied to Transient Device Simulation," *Proceedings of the International Symposium on Circuits and Systems,* 1988.

[6] C.S. Bamji, *Graph-Based Representation and Coupled Verification of VLSI Schematics and Layouts*, Ph.D. diss. proposal, Dept. of Electr. Eng. and Comp. Sci., MIT, 1987.

[7] F. Van Aelten, *Constraining Mechanisms for Circuit and Logic Level VLSI Representations,* S.M. thesis proposal, Dept. of Electr. Eng. and Comp. Sci., MIT, 1988.

[8] R.C. Armstrong, *A Formal Approach to Incremental Consistency Maintainence in Multirepresentation VLSI Databases,* Ph.D. diss. proposal, Dept. Electr. Eng. and Comp. Sci., MIT, 1987.

15.2 Extracting Masks from Optical Images of VLSI Chips

Rockwell International Corporation
OKI Semiconductor
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)

Bruce R. Musicus, Hong Jeong

One of the chief difficulties in studying image modeling and image understanding is that it is difficult to find useful models to aid in interpreting unconstrained images. To better understand the role of image modeling, we have focused on the particular problem of reverse-engineering a VLSI chip given a micro-photograph of the chip. The in-

teresting feature of this problem is that it requires combining conventional image processing with a rule-based image model. An enormous amount of modeling information is available concerning the design of VLSI chips: they are fabricated in layers of known composition and optical appearance; the images are formed from strips of material delineated by clear, though ragged, boundaries; the strips must form electrical circuits with known characteristics. Given all this *a priori* information, including knowledge of minimum line and feature widths as well as rules about the composition of layers making up the chip, our goal has been to build an efficient analysis system for reconstructing the masks that were used to manufacture the chip.

Our initial work in this area focused on low-level image processing issues, such as compensating for improper focusing, imbalanced lighting, and texture, while trying to accurately segment the image into line strips. It was found that local analysis methods, analyzing small windows to decide if they contained an edge or not, worked reasonably well and were relatively insensitive to lighting and texture fluctuations.

Our latest work has been to address the back end of the analysis system.[1–4] Given a clean line drawing representing a section of the VLSI chip, how do we piece together the various strips into the masks that formed the chip? We have developed various algorithms which deduce all possible legal interpretations of a given line drawing using only information about the edges. Our programs are written in a mixture of LISP, C, and PEARL, and run under UNIX 4.3. Using a database of rules of VLSI design, the programs start with edges and vertices, piece together paths marking the edges of a strip in some mask, assign the paths to layers, and label the layers. The most difficult part is to correctly infer where mask strips cross over each other, and to properly interpret accidental edges, where several strip edges coincide in the image.

Our most efficient algorithm utilizes a depth-first search strategy which constructs the final mask layers through a series of stages, trying out all possible interpretations of each line in the original image. Pruning this search tree quickly is vital, because of the exponentially increasing number of interpretations. Geometric reasoning based on the multi-layer strip model limits the growth in the number of interpretations until we reach the stages where we must assign potential strip contours to layers, choose which part of each layer is filled, and name the layers. At this point, our package invokes elementary circuit reasoning to rule out clearly infeasible circuit constructs. At present, our package only uses rules about contact cuts and wiring.

There are two major extensions to this work which we are now contemplating. The first is to increase the level of circuit reasoning ability in the software package to prune out more infeasible circuits. In particular, we would like to incorporate more careful reasoning about the shapes of the strips, the location of contact cuts within the shapes, and strip crossings which form transistors. A second major extension would be to use feedback from the symbolic reasoning system to improve the edge and segment extraction in the front-end. At present, our software assumes that the initial line drawing is perfect, and relies heavily on this assumption to rule out invalid interpretations. Ideally, we would like to modify the system to be more robust to mistakes in the line drawing, and to be able to isolate approximately where an error may have occurred in the initial segmentation. Such feedback could be used to change the processing used in the front end, to produce a better line drawing of the chip. In addition, we would like to use more information from the original image, such as brightness, color, and texture, to further improve the interpretation process.

References

[1] H. Jeong and B.R. Musicus, "Mask Extraction from Optical Images of VLSI Circuits," *IEEE ICASSP '87,* Dallas, Texas, April 5-9, 1987.

[2] H. Jeong, "Mask Extraction from Line Drawings of Optical VLSI Images," Ph.D. diss., MIT, 1987.

[3] H. Jeong and B.R. Musicus, "Extracting Masks from Optical Images of VLSI Circuits," In *Advances in Machine Vision, Algorithms, and Architectures,* ed. Jorge L.C. Sanz. New York: Springer-Verlag, 1988.

[4] H. Jeong and B.R. Musicus. "Extracting Masks for VLSI Circuits Using the Geometrical Configurations Around Contact Cuts," to be submitted 1988.

15.3 Cellular Array for Image Processing

Rockwell International Corporation
OKI Semiconductor
U.S. Navy - Office of Naval Research (Contract N00014-81-K-0742)
U.S. Air Force - Office of Scientific Research (Contract AFOSR 86-0164)

Bruce R. Musicus, G.N. Srinivassa Prasanna, Hong Jeong, Edward Schembor, John Deroo, Kevin O'Conner

Low-level image processing operations, such as contrast stretching, compensation for lighting variation, noise suppression, or edge enhancement, often rely on highly repetitive processing of the pixels in the image. In conventional image processing architectures, this characteristic is exploited by pipelining the image data through a computational pipeline which repetitively executes the same instruction on all the data flowing through it. An alternative approach, which we are exploring, is to build a large number of small processors, and use these processors in parallel to execute the same instructions on different parts of the image. The Connection Machine is the best-known commercial implementation of this architectural idea. Our goal is to explore much simpler and cheaper implementations, which can be carefully matched to the algorithmic domain in order to achieve high performance at low cost.

To better understand hardware, software, and algorithmic issues involved in this approach, we have built a small 16 by 16 array of 256 single-bit processors, packaged on 2 VME boards with data memory, a horizontally microcoded sequencer, and a host interface.[1,2,3] Combined with a frame grabber and a 68000 controller card, we have a very high performance machine capable of extremely high speed computation for a particular class of signal processing problems. The array is built from four AAP chips from OKI Semiconductor, and operates at a 6.5 MHz rate, performing 256 bit operations on every clock tick. Both bit-serial and bit-parallel arithmetic are supported. Data memory is specially designed to supply overlapping frames of bit-serial or bit-parallel data to the processor array. The machine is programmed with a microP-assembler with high-level control constructs and expression evaluation. Various algorithms for low-

level image processing and matrix arithmetic are under development, and we are finishing final system debugging.

The purpose of building this system was to better understand the consequences of using a high degree of parallelism in signal processing tasks which would appear to be highly parallelizable. As would be expected, writing software for this array is non-trivial. The user must be highly familiar with the byzantine structure of the processor architecture, and with what operations can and cannot be performed in parallel. New algorithms must be developed for standard image processing tasks which can utilize all 256 processors in parallel. Algorithms must be carefully partitioned so that they deal with 16 by 16 chunks of the image. Much programming effort and run time must be devoted simply to moving the image data in and out of the array. High -level languages are difficult to design for the machine, because the hardware capabilities are highly non-orthogonal and are tightly bound to particular settings of various flag bits. In part, it was our familiarity with the difficulties of programming this machine that has motivated us to consider compilation systems capable of automating at least some part of this programming.

References

[1] J. DeRoo, *Image Processing Software Package,* S.B. thesis, Dept. Electr. Eng. and Comp. Sci., MIT, 1987.

[2] K.G. O'Conner, *Low Level System Interface for the OKI 16416 Cellular Array Project,* S.B. thesis, Dept. Electr. Eng. and Comp. Sci., MIT, 1987.

[3] E.A. Schembor, *Software Kernel for Array Processor Based Parallel Image Processing System,* S.B. thesis, Dept. Electr. Eng. and Comp. Sci., MIT, 1987.

15.4 Algorithmic Fault Tolerance in Digital Signal Processing

Charles Stark Draper Laboratory
U.S. Air Force - Office of Scientific Research (Contract AFOSR 86-0164)

Bruce R. Musicus, William S. Song

Conventional methods for achieving fault tolerant computer architectures rely on triplicating computational resources and using voter circuitry to reject incorrectly computed results. From an information theory viewpoint, however, a much better error control philosophy would be to use coding techniques to achieve a high level of error recovery with minimal overhead. Essentially, a coder distributes information across a noisy channel bandwidth in such a way that individual noise spikes may destroy a portion of many bits, but not an entire bit. A decoder at the receiver can combine information from the full channel bandwidth to reconstruct the original message, with a very high degree of reliability.

Coding techniques are used heavily in high performance memory systems and in communication networks. These techniques are excellent at protecting modules where data entering at one end is expected to arrive intact and unchanged at the other end.

However, coding techniques are not usually used to protect actual computation. Instead, high levels of fault tolerance within CPU's are traditionally achieved by duplicating or triplicating processor resources, and voting on the results. Another problem with coding is that the coding and decoding procedure adds to the latency of the channel, slowing down any machine using the protected component.

In this project, we are developing a new approach to fault tolerance, in which we can protect certain types of linear computation against processor failure, by using a small number of redundant processors to protect each other and to protect the real processors. One design uses a bank of analog-to-digital converters operating in round-robin fashion to achieve an overall sampling rate somewhat above the Nyquist rate for the signal. A dither system and digital low-pass filter combine to reduce quantization errors in the front end. This same low-pass, however, can be used to detect and correct temporary or permanent errors in any of the converters, without substantially increasing the total amount of computation. The system is able to trade off additional hardware for greater accuracy and higher levels of fault protection. As converters fail, all that happens is that the effective quantization error increases.

Another application is to the FFT processor system used in range and velocity doppler sonar processing. Here we use a stack of processors to process multiple scans of sonar data from a phased-array antenna. Each processor does the same linear FFT processing, but on different sets of range cells. Adding extra processors working on linear combinations of the inputs to the other processors allows simple fault detection and correction. Regardless of the number of processors in the system, detecting K simultaneous failures requires only K extra processors; detecting and correcting K simultaneous failures requires only 2K extra processors. When conventional truncation or rounding arithmetic is used, however, then the error checking can only be approximate. In this case, adding more processors improves the accuracy of the fault checking and correction. Generalized likelihood ratio tests are used to select the most likely failure hypothesis, and to perform the most likely fault correction. Realistic systems result which use comparatively small computational overhead (<50%) to achieve 100% single fault detection and correction. We are presently working with Draper Labs on the design of a sonar system incorporating these concepts.

Publication

Song, W., and B.R. Musicus, "A Fault-Tolerant Architecture for a Parallel Digital Signal Processing Machine," In *Proceedings 1987 International Conference on Computer and Circuit Design,* Rye, New York, 1987.

15.5 Simulation of VLSI Circuits

DARPA/U.S. Navy - Office of Naval Research (Contracts N00014-80-C-0622 and N00014-87-K-0825)
National Science Foundation (Grant ECS-83-10941)

John L. Wyatt, Jr., Keith S. Nabors, Peter O'Brien, David Standley, Andrew Lumsdaine

We have been finishing up our previous project on delay estimation for digital integrated circuits and beginning a new one on parallel simulation methods for regular analog arrays.

One goal of this last phase of our CAD project has been to extend the waveform bounding results that Penfield et al. developed for MOS circuits so that they work for high-speed ECL as well. Our recent work consists of two major parts: 1) macromodelling of ECL logic gates acting both as drivers and as loads; and 2) delay estimation for individual nets using the gate macromodel parameters and RC tree models for metal interconnect. The success of the macromodelling approach relies on repetitive use of a library of modelled cells. A fixed computational cost (several mainframe CPU hours per cell) is paid to obtain parameter values for the simplified macromodels. The resultant timing estimates are typically within 5% - 10% of SPICE and are obtained with roughly 1000x less CPU time per run. This work has taken a very practical turn and has been extensively tested on an industrial ECL process and cell library. It is now in use in two industrial settings and will appear in Peter O'Brien's S.M. thesis this spring.

Another goal at the final phase of this project is to explore the possibility of building a very fast, arbitrarily accurate delay simulator for digital MOS circuits that is highly optimized and specialized to exploit the special mathematical structure of linear RC networks in which all capacitors are connected to ground, a common class of models for signal propagation delay in digital MOS. This work is nearly complete and will appear in Keith Nabors' M.S. thesis this spring.

We are beginning a new project on the parallel simulation of large, regular analog arrays. The goal is to produce a simulation tool that can be used for the design of smart sensors for machine vision, such as those currently being developed at Carver Mead's laboratory at the California Institute of Technology. These chips typically consist of large arrays, e.g., from 32x32 to 128x128, of moderately simple analog cells that perform a collective analog computation by communication with nearest neighbors. No currently available simulation tool is of any use on circuits of this size.

Systems of this type have two features that should influence the design of a tailor-made simulator. One is the natural hierarchy: devices in circuits in cells in arrays. The other is the designer's concern with circuit sensitivity in such systems: how do individual component variations affect overall system performance? We are studying ways to adapt the classical adjoint network approach to a hierarchical framework to solve this latter problem with acceptable computational efficiency.

Publications

Wyatt, J.L., Jr., "Signal Propagation Delay in RC Models for Interconnect," *Circuit Analysis, Simulation and Design. Part II: VLSI Circuit Analysis and Simulation,* ed. A. Ruehli, 254-291, North-Holland, 1987. (Also published as VLSI Memo 87-381, 1987, under the title *The Practical Engineer's No-Nonsense Guide to On-Chip Signal Delay Calculations.*)

Wyatt, J.L., Jr., *Nonlinear Dynamic Maximum Power Theorem:* VLSI Memo 87-371. MIT, 1987.

Wyatt, J.L., Jr. and D.L. Standley, "A Method for the Design of Stable Lateral Inhibition Networks that is Robust in the Presence of Circuit Parasitics," Proceedings of the IEEE Conference on Neural Information Processing Systems, Denver, Colorado, November, 1987.

15.6 Mixed Circuit and Device Simulation

International Business Machines, Inc.

Mark Reichelt, Jacob White, Jonathan Allen

Both digital and analog MOS circuit designers rely heavily on circuit simulation programs like SPICE to insure the correctness and to test the performance of their designs. For most applications, the lumped MOS models used in these programs accurately reflect the behavior of terminal currents and charges, but in some cases, these models are not adequate. In particular, charge redistribution between source and drain during device switching cannot easily be modeled by a lumped device, but the details of this charge redistribution can have an important effect on circuit behavior. In circuits like dynamic memory cells, sense amplifiers, analog-to-digital converters, and high frequency operational amplifiers, charge redistribution effects may not only degrade performance, but can inhibit proper function.

For these critical applications, sufficiently accurate transient simulations can be performed if, instead of using a lumped model for each transistor, some of the transistor terminal currents and charges are computed by numerically solving the drift-diffusion based partial-differential equation approximation for electron transport in the device. However, simulating a circuit with even a few of the transistors treated by solving the drift-diffusion equations is very computationally expensive, because the accurate solution of the transport equations of an MOS device requires a two-dimensional mesh with more than a thousand points.

One approach to accelerating this kind of mixed device and circuit simulation is to apply waveform relaxation to accelerating the transient simulation, not just at the circuit level, but inside the devices being simulated with a drift-diffusion description. In the present investigation, the WR algorithm is being applied to the sparsely-connected system of algebraic and ordinary differential equations in time generated by standard spatial discretization of the drift-diffusion equations that describe MOS devices. It has been proved (with J. Wyatt) that the WR algorithm contracts in a uniform norm on a model of the device simulation problem, and the result was verified on a one-dimensional experiment.[2] The implementation of the method for 2-D device simulation is in progress.

15.7 Detailed Simulation of Phase-Locked Loops

AT&T Bell Laboratories

K. Kundert, Jacob White

Switching filters and phase-locked loops are computationally expensive circuits to simulate using conventional circuit simulators like SPICE or ASTAP. This is because these kinds of circuits are all clocked at a frequency whose period is orders of magnitude smaller than the time interval of interest to the designer. To construct such a long time solution, a program like SPICE or ASTAP must calculate the behavior of the circuit for many high frequency clock cycles. The focus of this research is to develop and implement more efficient techniques for the circuit simulation of switching filter and phase-locked loop designs. The basic approach to simulating these circuits is to exploit only the property that the behavior of such a circuit in a given high frequency clock cycle is similar, but not identical, to the behavior in the preceding and following cycles. Therefore, by accurately computing the solution over a few selected cycles, an accurate long time solution can be constructed.

Simulating switched analog systems is an old problem but this novel approach has led to a recent success. The most structured of these switching problems, and therefore the easiest, is the switched-capacitor filter. A very efficent algorithm for the steady state analysis of switched-capacitor filters has been developed.[1] The idea is based on simulating selected cycles of the high-frequency clock accurately with a standard discretization method, and pasting together the selected cycles by computing the low frequency behavior with a truncated Fourier series. If carefully constructed, the nonlinear system that must be solved for the Fourier coefficients is almost linear and can be solved rapidly with Newton's method.

15.8 Numerical Algorithms for Hydrodynamics-Based Device Simulation

DARPA/U.S. Navy - Office of Naval Research (Contract N00014-87-K-825)
Analog Devices, Inc.

Keith S. Nabors, Jacob White

The model used in conventional device simulation programs is based on the drift-diffusion model of electron transport, and this model does not accurately predict the field distribution near the drain in small geometry devices. This is of particular importance for predicting oxide breakdown due to penetration by "hot" electrons. An approach for improving the accuracy of the drift-diffusion model of electron transport is to take an additional moment of the Boltzman equation, which yields a system of equations for electron transport that is similar to the drift-diffusion model, but includes the electron energies. The model is referred to as the hydrodynamic model and has been implemented in several simulators.

The hydrodynamic model is not as numerically tame as the standard drift-diffusion model, and other hydrodynamic simulators either circumvent this problem by ignoring difficult terms, or occasionally produce oscillatory results. A. research goal in this area is to try to develop a simulation program, based on the complete hydrodynamic model, whose numerical methods are efficient and are as robust as those used for the drift-diffusion model. Present work in this direction has been to implement a 2-D hydrodynamics-based simulator using standard numerical techiques in order to characterize the instabilities. Results so far demonstrate that the instabilities occur in the direction orthogonal to that of the dominant current flow.

15.9 Parallel Numerical Simulation Algorithms

DARPA/U.S. Navy - Office of Naval Research (Contract N00014-87-K-825)

Andrew Lumsdaine, Jacob White.

The key problem in parallelizing many of the numerical algorithms used in circuit and device simulators is finding efficient techniques for solving large sparse linear systems in parallel. Most parallel matrix solution algorithms fall into one of two general categories, the direct (Gaussian-elimination based) and the iterative, and each presents quite different problems. The computation in the direct approach is not very structured, and is therefore difficult to parallelize. Iterative methods are easily parallelized, but are not numerically robust.

The direct solution of circuit simulation matrices is particularly difficult to parallelize, in part because methods like parallel nested dissection are ineffective due to the difficulty of finding good separators. For this reason, the interaction between sparse matrix data structures and computer memory structure is being investigated (with Professor W. Dalley) to see how to store sparse matrices for effective parallel computation. One interesting recent result is that it is possible to store matrices in a completely scattered form, which allows for access of the sparse entries by fast indexing, in only three times the storage. This data structure is static, constructed once the matrix structure is known, and is therefore very attractive for parallel implementation.

In order to improve the reliability of relaxation methods for circuit simulation, approaches are being investigated based on extracting bands from a given sparse matrix, solving the bands directly, and relaxing on the rest of the matrix. This approach is efficient because band matrices can be solved in order $log(n)$ time on order n processors, and this approach is more reliable than standard relaxation, because "less" relaxation is being used. This banded relaxation has been tested on circuit simulation matrices and does converge faster and more often than standard relaxation on all examples tried so far. The present work is on selecting the ordering of the matrix to best exploit the direct solution of the band, and to automatically select the band size. In addition, an implemention of the method on the Connection Machine is in progress.

Studying the behavior of parallel iterative algorithms for sparse matrix solution has lead to an interesting theoretical result indicating an optimality of Gauss-Jacobi over Gauss-Seidel relaxation, given enough processors. Of particular interest is that the result connects the spectral radius of the iteration matrices to their graphical properties.[3] The results are being extended to the waveform relaxation case, where in practice, this limiting result seems to show on as few as eight processors.

As mentioned above, relaxation algorithms for solving matrices are easily parallelized. It is also possible to apply relaxation directly to the differential equation, referred to as waveform relaxation (WR), and easily develop a parallel algorithm in which different differential equations are solved on different processors. A recently developed variant of the WR algorithm, referred to as waveform-relaxation Newton (WRN), allows for additional parallelism in that most of the computation for each of the discretization timepoints for a single differential equation can be computed in parallel. In recent

theoretical work, it has been proved that WRN converges globally even when applied to circuits with nonlinear capacitors.[4]

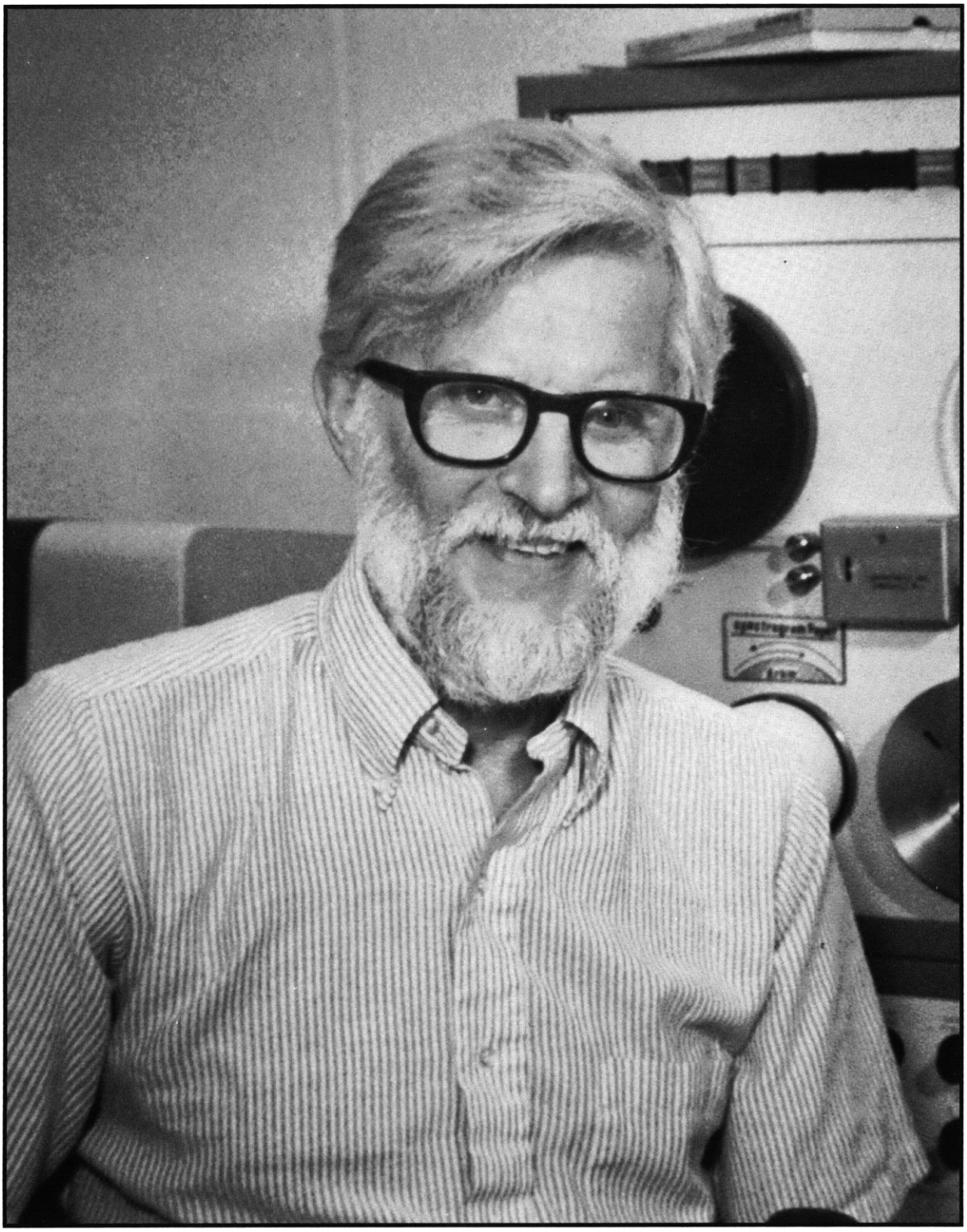
References

[1] K. Kundert, J. White and A. Sangiovanni-Vincentelli, "A Mixed Frequency-Time Approach for Finding the Steady State Solution of Clocked Analog Circuits," Custom International Circuits Conference, Rochester, New York, 1988.

[2] M. Reichelt, J. White, J. Allen and F. Odeh, "Waveform Relaxation Applied to Transient Device Simulation," International Symposium on Circuits and Systems, Espoo, Finland, 1988.

[3] D. Smart and J. White, "Reducing the Parallel Solution Time of Sparse Circuit Matrices using Reordered Gaussian Elimination and Relaxation," International Symposium on Circuits and Systems, Espoo, Finland, 1988.

[4] R. Saleh, J. White, A. S. Vincentelli, and A. R. Newton, "Accelerating Relaxation Algorithms for Circuit Simulation Using Waveform-Newton and Step-Size Refinement," submitted to the *IEEE Trans. Comp. Aided Design*.



Professor Kenneth N. Stevens