

EMBEDDED MODEL PREDICTIVE CONTROL FOR SYSTEM-ON-A-CHIP APPLICATIONS

Leonidas G. Bleris* and Mayuresh V. Kothare**,¹
Jesus Garcia*** and Mark G. Arnold***

* *Electrical and Computer Engineering Department*

** *Chemical Engineering Department*

*** *Computer Science and Engineering Department*

Lehigh University, Bethlehem, PA 18015, U.S.A.

Abstract: We propose a framework for embedding model predictive control for Systems-on-a-Chip applications. In order to allow the implementation of such a computationally expensive controller on chip, we propose reducing the precision of the operations coupled with using logarithmic number system arithmetic. Two particular control problems are examined. We provide the methodology for choosing the design parameters; we emulate the performance of the embedded controller for the examined cases and we give the microprocessor architecture details.

Keywords: Embedded Systems, Model Predictive Control, Systems-on-a-Chip.

1. INTRODUCTION

The realization of autonomous Systems-on-a-Chip (SoC) applications has been a very significant research area during the past decade. Moving on from simple prototype applications, new generations of more complex SoC are currently being developed for a variety of applications. The market for SoC is expanding rapidly to areas like medicine and bioengineering (DNA and genetic code analysis and synthesis, drug delivery, diagnostics and imaging), information technologies, avionics and aerospace (nano and microscale actuators and sensors, smart reconfigurable geometry wings and blades, microgyroscopes), automotive systems and transportation (accelerometers), microreactors for in-situ and on-demand chemical production. The functionality and performance of such a versatile SoC is directly related to the reliability and quality of the control logic used. For example processing DNA molecules in the

microscale requires the fabrication of microfluidic devices capable of handling, mixing, thermal cycling and separating liquid micro-samples. Therefore efficient custom-made controllers have to be designed and integrated on chip. Furthermore the performance of the overall system has to be examined and verified through advanced simulation techniques.

With this paper we provide the framework for applying Model Predictive Control (MPC) and designing the microcontroller architecture that will allow optimal sensing-control-actuation performance for microchemical system applications. Microchemical systems are a new generation of miniature chemical systems that carry out chemical reactions and separations, in precisely fabricated three dimensional microreactor configurations. Our research effort is centered on eventually developing and embedding a controller to a catalytic reformer and separator microchemical system that can operate as a sustained source of hydrogen fuel for Proton Exchange Membrane (PEM) fuel cells. This SoC will be used poten-

¹ Corresponding author: phone (610) 758-6654, fax (610) 758-5057, e-mail: mayuresh.kothare@lehigh.edu

tially as an alternative to conventional portable sources of electricity such as batteries for laptop computers and mobile phones due to its ability to provide an uninterrupted supply of electricity as long as a supply of methanol, water and heat can be provided.

The application of real-time embedded model predictive control for microscale devices presents new technological challenges. We propose reducing the precision of the microprocessor to the minimum while maintaining stable control performance. Taking advantage of the low precision, a Logarithmic Number System (LNS) based microprocessor architecture is used, providing further energy and computational cost savings. In order to choose the optimal design parameters for this high-performance embedded controller we utilize computational tools to simulate the plant and we emulate the microcontroller arithmetic operations.

Recently a multi-parametric programming method was proposed (Bemporad *et al.*, 2002) to solve off-line the quadratic optimization problem associated with MPC. The constrained quadratic optimization is shown to be piecewise affine, using partitions of the state space determined by the constraints. The feedback laws are pre-computed and online calculations are greatly simplified, consisting of a table-lookup in memory and an affine transformation. However, the upper bound on the number of required regions is $N_r \leq \sum_{k=0}^{2^q-1} k!q^k$, where q is the number of constraints. For an input constrained problem, increasing the number of controlled variables (thus the number of constraints), yields prohibitive memory requirements for Systems-on-Chip.

As an alternative, we propose a framework for embedding MPC by reducing the precision, thereby increasing the optimization speed while reducing the power consumption and the overall chip area. The power consumption in the arithmetic units is related to the number of switching transistors. Therefore by reducing the transistor count we expect not only a lower cost for the embedded microcontroller but also a significant reduction in the consumed power. Additionally by integrating all the memory on-chip and reducing its size, the power dissipation can be further reduced. This provides a solution for embedding MPC for complex large applications.

With the rapid development of computers and the software tool capabilities, novel approaches can be used for the analysis of SoC. In this paper we use FEMLAB (COM, 2001), a Partial Differential Equation (PDE) solver. We use the ability to export complex micro-geometries into MATLAB environment and to analyze rigorously their dynamic behavior by solving problems such as

heat transfer, convection-conduction and Navier-Stokes flows. Two control problems that are critical for the target microchemical system are provided. The control of the temperature distribution across a wafer geometry is initially examined. Secondly, we look into the problem of controlling the temperature of a fluid flowing in a microchannel and the flow velocity at the outlet of the microchannel by adjusting the inlet velocity and regulating the provided distributed energy supply.

2. THEORETICAL BACKGROUND

2.1 Model Predictive Control Theory

Model predictive control is also known as receding horizon control or moving horizon control. Generally controllers belonging to the MPC family are characterized by the following steps. Initially the future outputs are calculated at each sample interval over a predetermined horizon N , the prediction horizon, using the process model. These outputs $y(t+k|t)$ for $k=1\dots N$ depend up to the time t on the past inputs and on the future signals $u(t+k|t)$, $k=0\dots N-1$ which are those to be sent to the system. The next step is to calculate the set of future control moves by optimizing a determined criterion in order to keep the process as close as possible to a predefined reference trajectory. This criterion is usually a quadratic function of the difference between the predicted output signal and the reference trajectory. In some cases the control moves $u(t+k|t)$ are included in the objective function in order to minimize the control effort. Finally, the first control move $u(t|t)$ is sent to the system while the rest are rejected. This is because at the next sampling instant the output $y(t+1)$ is already known from the optimization results and the procedure is repeated with the new values so that we get an updated control sequence.

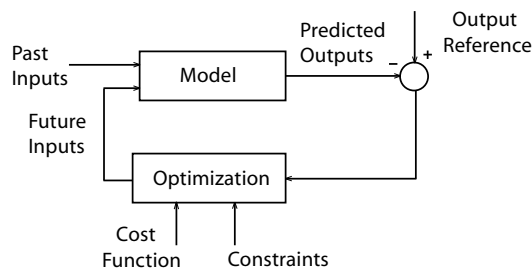


Fig. 1. Model Predictive Control block diagram.

Figure 1 shows how to implement the basic controller structure. A model is used to predict the future plant outputs based on past and current values and on the calculated optimal future control moves. These actions are calculated by the optimization that also takes into account the constraints. There are many models used in different formulations of the model predictive control.

These include step response models, impulse response models, transfer function models and the state space models. In this paper we build the geometry, define the PDE models and the boundary conditions in FEMLAB and we use the impulse responses as a model for the predictive control problem formulation.

2.2 Logarithmic Number System

One of the main advantages of floating point (FP) is its standardization, but for an architecture oriented towards embedded systems and therefore susceptible to being tailored for each particular application the need for novel approaches arises. The logarithmic number system (LNS) represents the value of the real number X using a sign bit and the base- b logarithm of $|X|$, x , so that

$$x = \text{round}(\log_b |X|), \quad (1)$$

where x is a fixed-point signed binary number, consisting of K integer bits and F fraction bits (Swartzlander and Alexopoulos, Dec. 1975). The $\text{round}()$ operation approximates the value of $\log_b |X|$ so that it is representable in $N = K + F$ bits. The values for K and F define respectively the dynamic range (largest and closest to zero representable values), and the precision (distance between consecutive representable values) that are available. LNS provides a representation with numeric characteristics similar to those offered by FP (Arnold *et al.*, Aug. 1992). The first reason for preferring LNS is the constant relative error characteristic (as opposed to FP) that gives a superior worst-case error in LNS for the same word size. FP architectures can be scaled in a more efficient way for high precisions, but for low-to-moderate precisions (below 32 bits), LNS implementations offer an advantage in term of cost, power consumption and speed, that increases as the word size decreases.

3. CONTROL OF TEMPERATURE DISTRIBUTION ACROSS A WAFER

Consider the geometry of Figure 2 that illustrates a wafer with a distributed array of resistive heaters placed on top. We measure the temperature of this geometry by placing sensors on the bottom of the wafer at the exact opposite location of the actuators. We consider the 2D cross section of this wafer with nine distributed heaters and nine sensors. We are interested in heating and maintaining the temperature of the wafer at a predefined set-point, while constraining the heat supply. This is a significant control objective for our examined SoC application since the catalytic reformer and separator need to be heated to a specific temperature to ensure catalytic activation. Additionally the optimal working point for these

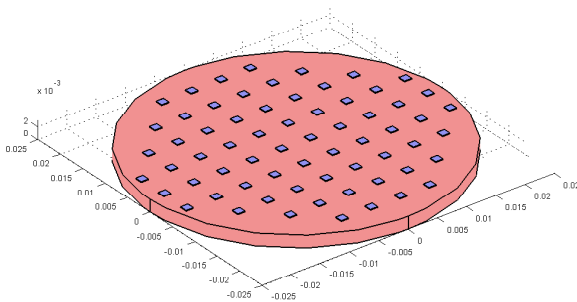


Fig. 2. Wafer geometry (dimensions in meters).

two reactions is different; therefore, we need to maintain a temperature gradient within the wafer. The temperature distribution in the wafer is described by the time-dependent two-dimensional heat equation

$$\frac{\partial T(t, x, y)}{\partial t} = \frac{1}{a} \nabla^2 T(t, x, y) \quad (2)$$

with initial conditions $T(0, x, y) = 300\text{K}$. Here $a = \rho C / \kappa$, κ is the thermal conductivity, ρ is the density and C is the heat capacity. The resistive actuator heat losses depend on the exposed surface $Q = hA(T - T_{air})$ where $T_{air} = 300\text{K}$, the natural convection heat-transfer coefficient $h=25\text{W}/\text{m}^2\text{K}$ and A is the exposed heater area. With appropriate use of materials we have the ability (Bleris and Kothare, 2003) to control and maintain temperature gradients in the proposed geometry. The material used for the following FEM simulations is pyrex. We initially set the temperature set-point (T_{sp}) at the sensors placed at the boundaries at 370K and for the rest at 390K . The heat supply is constrained for each actuator at $5\text{W}/\text{m}^2$. We apply this problem on a 2D geometry but the proposed approach can be easily extended in 3D, using an array of sensors and actuators.

The optimization is a fundamental part of MPC since it computes the optimal control input for the process. The computational effort required in MPC derives almost entirely from the optimization algorithm. Logically the choice of the optimization technique (Fletcher, 1987) is decisive to the performance of the controller. The algorithm that is implemented for the minimization of the cost function is a direct application of Newton's method (incorporating the constraints in the cost function using barrier functions). It has to be noted that from the architecture perspective the operations that are responsible for the majority of instructions that are executed in the processor are matrix inversions and abundant dot products. To

estimate the performance of MPC using reduced precision, the basic arithmetic operations required by the optimization procedure were emulated using LNS. The design parameters of MPC as well as the number representation K and F (integer and fractional part of the fixed point exponents) were varied, allowing the selection of the desired compromise between precision reduction and response quality of the system.

The control problem of reaching the set-point is solved initially using $K = 7$ (integer bits) and $F = 20$ (fraction bits) for different control horizons (ch), Figure 3. The prediction horizon is fixed at 100. We define the error of the controller performance using

$$Error = \frac{\sum_{i=1}^9 |Tsp(i) - T(i)|}{\sum_{i=1}^9 (Tsp(i) - 300)} \times 100 \quad (3)$$

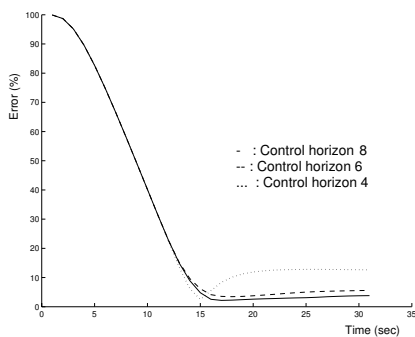


Fig. 3. Percentage of error using different control horizons.

From Figure 3 it is obvious that the performance of MPC is not satisfactory for $ch = 4$ and it is closest to the set-point using $ch = 8$. The total energy consumed using $ch = 8$ was $567\text{J}/\text{m}^2$ while using $ch = 6$ yields a 3.4% reduction. Additionally, since we are interested in minimizing the architectural complexity of the embedded controller we choose to use $ch = 6$. In Figure 4 we provide the simulation results. The objective now is to reduce further the implementation complexity, while maintaining a performance close to that of $K = 7$, $F = 20$ and $ch = 6$.

The next step for the proposed framework is to reduce the size of the integer part K . Emulation results of the coupled system of the MPC microcontroller and the FEMLAB models show exactly the same performance for $K = 7, 6, 5$ and significantly worse for $K = 4$. Therefore $K = 5$ was chosen. Subsequently the fractional part F was gradually reduced and in Figure 5 we provide the performance for different cases. From Figure 5 we notice that for $K = 5$, $F = 10$ and $ch = 6$ the error increases significantly. This can be explained by the fact that the subtraction operation at low precisions can be inaccurate when the result is close to zero. As the plant output

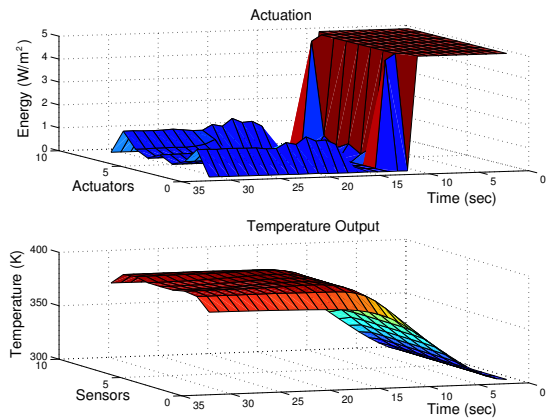


Fig. 4. Control variables and resulting temperature profile for $K=7$, $F=20$ and $CH=6$.

approaches the set-point, the values of the manipulated variables start oscillating (Figure 6) to keep the output close to the set-point. Furthermore this phenomenon creates added undesirable power consumption. To overcome this problem a hybrid MPC approach is under investigation, where the MPC switches to a Proportional Integral (PI) controller close to the set-point.

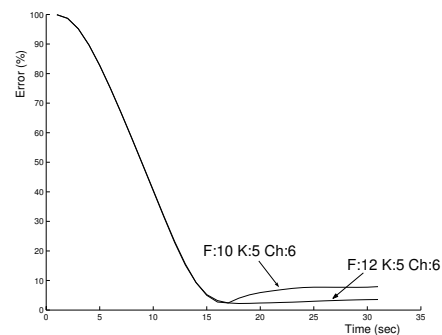


Fig. 5. Percentage of error reducing F (fraction bits).

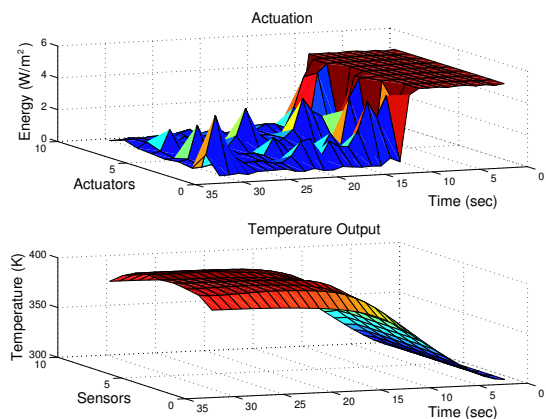


Fig. 6. Control variables and resulting temperature profile using $K=5$, $F=10$ and $CH=6$.

The preceding simulations show that for the examined problem, the minimum values for K and F were 5 and 12, respectively, which requires a 19-bit LNS representation (adding two bits for the signs

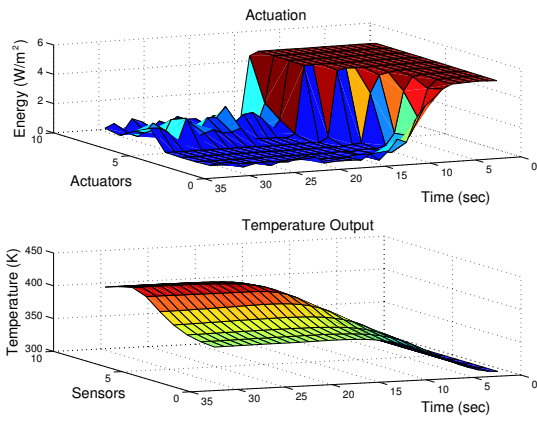


Fig. 7. Control variables and resulting temperature profile using $K=5$, $F=12$ and $CH=6$.

of K and F). We test this configuration using now a different set-point for the output temperature. The set-point is at 360K for the first 4 sensors, the 5th sensor is free to take any value and the final 4 sensors at 400K. From the simulation results of Figure 7, we notice the successful performance of the proposed reduced-precision controller.

4. NON-ISOTHERMAL FLOW

In this section we examine the coupled problem of microfluidic flow and convection-conduction in the 2D geometry of Figure 8. The microchannel has depth of 3mm and the arrows indicate the inlet, the outlet and the location of one of the ten distributed resistive heaters. We measure the temperature by placing ten sensors on the bottom of the geometry at the exact opposite location of the actuators. The dynamic model of the flow



Fig. 8. Microchannel geometry.

can be described by the Navier-Stokes and the convection-conduction equations

$$\rho \frac{\partial v}{\partial t} - \nabla \cdot \eta (\nabla v + (\nabla v)^T) + \rho (v \cdot \nabla) v + \nabla p = 0 \quad (4)$$

$$\nabla \cdot (\rho v) = 0 \quad (5)$$

$$\rho C_p \frac{\partial T}{\partial t} + \nabla \cdot (-k \nabla T + \rho C_p T v) = 0 \quad (6)$$

where k is the thermal conductivity and C_p the heat capacity. Furthermore we use the ideal gas law to relate the changes of temperature to density $\rho = \frac{pM}{RT}$, where R is the gas constant and M is the molar mass. For the Navier-Stokes equations we define the boundary conditions to be no-slip at the walls and straight out at the outlet. For the heat balance the temperature at the inlet is 300K, at the top wall it is the energy provided by the resistive heaters, it is insulated at the bottom

and at the outlet we assume that the transport is dominated by convection (the gradient of T perpendicular to the outlet is zero). We are interested in heating and maintaining the temperature of the fluid to a predefined set-point, while constraining the heat supply. Moreover due to the changes of density, the velocity of the fluid increases upon heating. We are therefore interested in controlling the velocity of the fluid at the outlet. Because of the fact that we have laminar flows in microscale (due to the low Reynolds number), we measure and control the maximum velocity which we expect to have at the middle of the channel. We solve the proposed control problem by decentralizing the system. Using proportional control the inlet velocity is adjusted to control the outlet velocity and MPC is applied for the regulation of the heat supply. Because of the fact that the fluid enters the microchannel at 300K, controlling the temperature profile becomes more important close to the inlet, especially for high velocities and elevated temperature set-point. We initially set the desired temperature at 350K throughout the microchannel and the maximum velocity at the outlet at 2mm/sec. The simulation results of applying MPC for controlling the temperature, with $K=7$, $F=20$, control horizon of 6 and prediction horizon 21, are illustrated in Figure 9. We

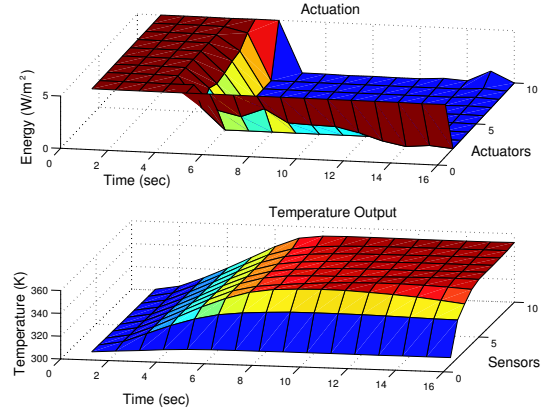


Fig. 9. Control variables and resulting temperature profile using $K=7$, $F=20$ and $CH=6$.

now simulate the system by gradually reducing the integer bits (K) and then the fraction bits (F) (Figure 10). We are interested in calculating the minimum precision while maintaining the same control performance. From the simulation results it becomes evident that we can reduce the controller size down to 17-bits without compromising the performance. As the LNS sign bit is treated separately, this requires only 16-bit addition to implement multiplication.

5. PROPOSED ARCHITECTURE

The requisites that an embedded MPC has to satisfy are: high speed, low power consumption

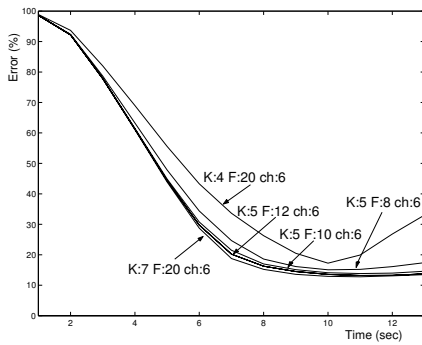


Fig. 10. Percentage of error using different design parameters.

and low cost. The last two are related, since power consumption is related to the transistor count and therefore the cost. The proposed architecture shown in Figure 11 utilizes macrocell technology. Commercially available designs of macrocells such as microprocessor cores and embedded DRAM exist. By using LNS custom arithmetic units, realistic MPC problems can be solved online with a system that integrates every component on the same chip. High speed in the embedded system

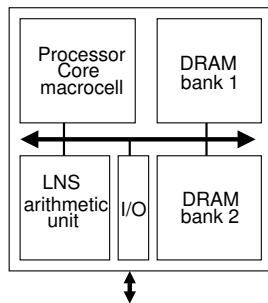


Fig. 11. Architecture block.

will depend both on fast arithmetic computations, and fast access to the data stored in memory. The main computing complexity comes from the abundant dot products and the matrix inversion. With double-precision floating point data the required ALU is expensive and operations would be slow and power consuming. Using a pipelined 16-bit LNS adder, neither scalar-vector multiplication nor vector addition represent a problem; these operations can be computed with a throughput of 1 operation/cycle. Long dot products can be unrolled in the adder pipeline, resulting in 1 op/cycle with a small overhead. To quantify the advantage of reducing the precision, estimations for both 64-bit FP and 16-bit LNS circuits show that for an arithmetic unit that computes addition, subtraction, multiplication and division, the size required is about 17 times larger for 64-bit FP, while the clock cycle is at least 3.23 times faster in 16-bit LNS. Having functional units that allow computing the majority of the operations with a throughput close to 1 operation/cycle, the other main obstacle to obtain high speed is memory access. Fast access to the RAM is a major

problem for most microprocessors (Hennessy and Patterson, 1996). However, current technology allows embedding DRAM macrocells of several Mb, combining the lower DRAM cost (compared with SRAM as used in caches), with an access time of just one clock cycle. MPC algorithms use vectors and matrices of a fixed length (within the optimization loop, which represents the majority of the operations to be computed). Very regular loops of a fixed size, known *a priori*, allow customizing the memory system for these accesses. Page-access to the DRAM will maximize the data rate, by reading and writing vectors by rows. The data size reduction not only increases speed, but also reduces the memory and the size of the arithmetic units.

6. CONCLUSIONS

A framework for embedding model predictive control logic on chip has been provided. Reducing the precision of the operations coupled with the use of a logarithmic-number-system arithmetic allows a very efficient implementation. We have emulated the performance of the reduced precision LNS-based MPC for two examined cases and we have provided brief microprocessor architecture details.

Future work includes obtaining theoretical bounds for convergence as a function of precision and formulating a detailed embedded system architecture. Also incorporating a hybrid logic to the controller to allow an early termination of the optimization algorithm so that further computational and energy savings will be investigated.

REFERENCES

- Arnold, M. G., T. A. Bailey, J. R. Cowles and M. D. Winkel (Aug. 1992). Applying Features of IEEE 754 to Sign/Logarithm Arithmetic. *IEEE Transactions on Computers* **41**, 1040–1050.
- Bemporad, A., M. Morari, V. Dua and E. N. Pistikopoulos (2002). The explicit linear quadratic regulator for constrained systems. *Automatica* **38**(1), 3–20.
- Bleris, L. G. and M. V. Kothare (2003). Model based control of temperature distribution in integrated microchemical systems. In: *Proceedings of the 2003 American Control Conference*. Denver, CO. pp. 1308–1313.
- COM (2001). *FEMLAB Reference Manual*.
- Fletcher, R. (1987). *Practical methods of optimization*. New York, Wiley Interscience.
- Hennessy, J. L. and D. A. Patterson (1996). *Computer Architecture: A Quantitative Approach*. Morgan Kaufmann. San Francisco, CA.
- Swartzlander, E. E. and A. G. Alexopoulos (Dec. 1975). The Sign/Logarithm Number System. *IEEE Transactions on Computers* **24**(12), 1238–1242.