Delta–Sigma Cellular Automata for Analog VLSI Random Vector Generation

Gert Cauwenberghs, *Member, IEEE*

***Abstract*—We present a class of analog cellular automata for parallel analog random vector generation, including theory on the randomness properties, scalable parallel very large scale integration (VLSI) architectures, and experimental results from an analog VLSI prototype with 64 channels. Linear congruential coupling between cells produces parallel channels of uniformly distributed random analog values, with statistics that are uncorrelated both across channels and over time. The cell for each random channel essentially implements a switched-capacitor delta–sigma modulator, and measures 100 μm $\times$ 120 μm in 2 μm CMOS technology. The 64 cells are connected as a MASH cascade in a chain or ring topology on a two-dimensional (2-D) grid, and can be rearranged for use in various VLSI applications that require a parallel supply of random analog vectors, such as analog encryption and secure communications, analog built-in self-test, stochastic neural networks, and simulated annealing optimization and learning.**

***Index Terms*—Random generation, noise, delta–sigma modulation, cellular automata, analog VLSI, neural networks, switched-capacitor circuits.**

I. Introduction

ON-LINE random analog signal generation is an essential component in many of today's analog very large scale integration (VLSI) systems for signal or information processing. An on-line supply of random analog vectors comes handy, for instance, to support testing and characterization of the hardware, or as part of the implemented algorithms. Examples of applications include encryption and secure communications [1]–[3], analog VLSI built-in self-test [4]–[6], and neural computation [7], [8], simulated annealing optimization [9], [10] and stochastic model-free learning [11]–[14].

Most commonly used in parallel VLSI are arrays of random binary sources implemented with linear feedback shift registers (LFSR) [15], [16] or cellular automata (CA) [17], [18], which yield compact and scalable parallel VLSI architectures [20]–[22]. Analog random vectors with a near-gaussian amplitude profile can be obtained from the binary random vectors through low-pass filtering [19], [23]. Sequential correlations over time caused by filtering can be elimitated by subsampling, at the expense of bandwidth.

Manuscript received July 31, 1997; revised May 30, 1998. This work was supported by the National Science Foundation (NSF) Career Award MIP-9702346 and ARPA/ONR MURI N00014-95-1-0409. Chip fabrication was provided through MOSIS.

The author is with the Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: gert@bach.ece.jhu.edu).

Publisher Item Identifier S 1057-7130(99)01775-9.

High-bandwidth, low-power analog noise generators in VLSI are obtained by means of chaotic oscillators [24]–[26], or through recursion of a nonlinear map such as the logistic map or a linear congruential map [27]–[30].

The most natural way to generate analog noise in VLSI is to amplify existing circuit noise, which usually is more of a nuisance than an aid to circuit design. Analog and binary random sources in VLSI have been demonstrated using amplified analog noise and a high-gain comparator [31], or using a latch initialized at the metastable operating point [32]. The challenge of this approach is to control sensitivity to physical parameters such as temperature, and minimize correlation effects across cells in an array due to unavoidable capacitive and power supply coupling.

Parasitic coupling between cells presents a problem to most other analog approaches as well, such as mode-locking phenomena in arrays of nonlinear oscillators, which cause strong correlations across cells.

In this paper, we demonstrate that a particular form of nonlinear coupling between cells not only avoids correlations across cells, but in addition produces a truly random sequence in the sense that the outcome of a cell at a given time is statistically independent of its history. This remarkable property is impossible by construction in an isolated cell with deterministic chaotic state recursion, regardless of the nonlinearity in the map, and emerges from interactions with neighbors. The interactions are nearest-neighbor as in cellular automata, and permit a simple scalable and parallel VLSI architecture. Our motivation to study these structures is inspired by remarkable noise-shaping properties observed in MASH cascade structures of delta–sigma modulators [33]–[36], as used for stable higher order oversampled A/D conversion [37]–[39]. As a particular case, we consider cellular arrays of cascaded delta–sigma modulators for the purpose of random analog vector generation, arranged on a two-dimensional (2-D) grid or in a linear array for scalable VLSI implementation.

The following section introduces the basic cellular architecture and its variants, and relates delta–sigma modulation to a congruential linear analog version of additive cellular automata. In Section III, the statistical properties of randomness are explored in theory. Section IV presents a compact analog VLSI implementation, and Section V includes experimental results from a 64-channel (and 65-channel) CMOS prototype. Finally, Section VI concludes the paper.

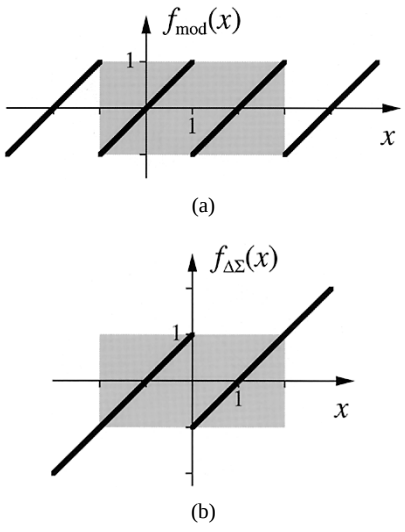

Fig. 1. Two cases considered for the nonlinear map $f(.)$. (a) Congruent map (modulo 2). (b) Quantization residue map (as used in single-bit delta–sigma modulation). The shaded inset shows the region where both maps coincide (except for a unit offset).

II. Nonlinear Noise-Shaping and Cellular Architecture

The general structure we consider combines additive cellular automata [17] and cellular neural networks [18], together with linear congruential maps [27], [28] or, as shown to be equivalent [34], delta–sigma modulation [39]. The interactions between cells are of the form

$$x_i(k+1) = f\left(\alpha + \beta \sum_{j\in\mathcal{N}(i)} x_j(k)\right) \tag{1}$$

where $\mathcal{N}(i)$ defines a neighborhood of cells interacting with cell i including itself, and where $f(.)$ defines a nonlinear map. Besides careful choice of the constants α and β, the form of $f(.)$ is critical to the randomness properties of the sequence $x_i(k)$.

A. Nonlinear Mappings

Two particular forms of the nonlinear map $f(.)$ are of interest, depicted in Fig. 1: the congruential map (or modulo operation) defined by the recursion

$$\begin{aligned} f_{\mathrm{mod}}(x) &= x && \text{if } -1 < x \le 1; \\ &= f_{\mathrm{mod}}(x-2) && \text{if } x > 1; \\ &= f_{\mathrm{mod}}(x+2) && \text{if } x \le -1 \end{aligned} \tag{2}$$

and the quantization residue map (as used in single-bit delta–sigma modulation) defined as

$$\begin{aligned} f_{\Delta\Sigma}(x) &= x - \mathrm{sign}(x) \\ &= x-1 && \text{if } x > 0 \\ &= x+1 && \text{if } x \le 0. \end{aligned} \tag{3}$$

The noise-shaping effect by the congruent map (2) can be intuitively understood from the scrambling of state variables that results from the modulo operations, and will be the subject of theoretical study in the following section. First, we establish the conditions for equivalence between the two maps $f_{\mathrm{mod}}(x)$ and $f_{\Delta\Sigma}(x)$, which is important because the first is easier to analyze (Section III) and the latter easier to implement (Section IV). The following lemmas derive from [33], [34]:

Lemma 1: Let a first vector sequence $x_i(k)$ be defined by (1) with constants α, β, and map $f_{\Delta\Sigma}(.)$ in (3), and a second sequence $x'_i(k)$ with constants $\alpha' = \alpha + 1$, same β, and map $f_{\mathrm{mod}}(.)$ in (2). If $|\alpha|+|\beta|\#\mathcal{N} \le 2$ and the initial values satisfy $-1 \le x_i(0) = x'_i(0) \le 1$, then $-1 \le x_i(k) = x'_i(k) \le 1$ for all $k \ge 0$.

This lemma states that the two mappings $f_{\Delta\Sigma}(.)$ and $f_{\mathrm{mod}}(.)$ (offset by one) generate identical sequences, contained in the $[-1, 1]$ interval, given identical initial conditions contained in that interval. The proof follows by induction $k \to k+1$, and from asserting that $-1 \le f_{\Delta\Sigma}(x) = f_{\mathrm{mod}}(x+1) \le 1$ for any $|x| \le 2$ (illustrated by Fig. 1). The condition $|\alpha|+|\beta|\#\mathcal{N} \le 2$ [where $\#\mathcal{N}$ denotes the cardinality of the neighborhood $\mathcal{N}(.)$] ensures that the argument of $f_{\Delta\Sigma}(.)$ in (1) is always between -2 and 2 as needed.

Lemma 2: Let a first vector sequence $x'_i(k)$ be defined by (1) with constants α, β and map $f_{\mathrm{mod}}(.)$ in (3), and a second sequence $x''_i(k)$ by the same constants but an identity map $f(x) \equiv x$. If β is integer and the initial values satisfy $x'_i(0) = f_{\mathrm{mod}}(x''_i(0))$, then $x'_i(k) = f_{\mathrm{mod}}(x''_i(k))$ for all $k \ge 0$.

The last lemma states that the usual algebraic rules of modulo arithmetic can be applied to the map $f_{\mathrm{mod}}(.)$ in (3). In particular, modulo and linear operations commute under the conditions satisfied by an integer choice for β, and the dynamics of a modulo system can conveniently be analyzed from a linear system, which is equivalent under subsequent modulo transformation. The proof is by induction on k, and follows from the fact that for each $x, f_{\mathrm{mod}}(x) = x + 2l$ for some integer l. Notice that the lemma is valid even for noninteger values of α. The lemma is used in the theory of Section III.

B. Cascade and Cellular Stuctures

The general template of nearest-neighbor interactions (1) allows to formulate cellular networks of various topologies. The simplest case to be possibly considered is a neighborhood of two cells, including one neighbor besides the cell itself. The largest absolute value possible for β to satisfy the conditions of Lemma 1 equals 1; any value lower than that would be undesirable for the purpose of generating random sequences as explained in the following section. With $\alpha = 0$ and $\beta = 1$, we obtain

$$x_i(k+1) = f_{\Delta\Sigma}(x_i(k) + x_{i-1}(k)) \tag{4}$$

shown in Fig. 2(a). This is functionally equivalent to a MASH cascade of first-order, single-bit delta–sigma modulators [40], where the quantization "noise" of the integrator of one stage feeds into the next [37], [38]. To visualize this equivalence, notice that the integration loop ("sigma") and the quantization residue ("delta") operations in Fig. 2(a) are permuted relative to the usual way a delta–sigma modulator is depicted [40].

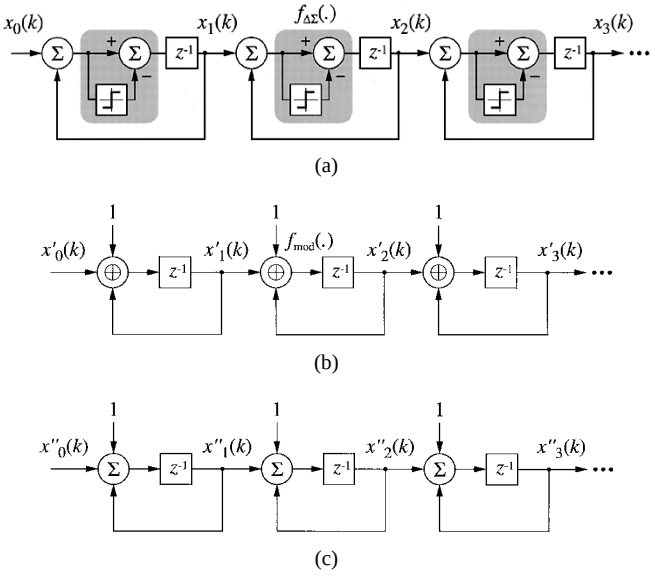

Fig. 2. MASH cascaded delta–sigma modulation. (a) System-level architecture. (b) Equivalent linear congruential additive cellular automaton according to Lemma 1. Additions are modulo 2 as shown in Fig. 1(a). (c) Corresponding linear additive cellular automaton according to Lemma 2.

Technicalities of compensating for systematic analog offsets aside [41], [42], cascaded structures of the MASH type are attractive for stable higher order oversampled A/D conversion, since the modulators do not overload and the "noise" does not appear to correlate with the input, at least for constant and sinusoidal inputs [33] and iid random inputs [35], [36].

As Lemma 1 asserts, the cascaded structure can alternatively be viewed as an analog extension on Wolfram's rule 120 cellular automata [17] $x_i(k+1) = x_i(k) \oplus x_{i-1}(k)$, where the exclusive-or operator is replaced with a modulo summation operator, shown in Fig. 2(b). Interestingly, the same modulo sum operator has also been used to construct a self-synchronizing analog encryption/decryption system [3].

The corresponding linear model of the cascade of delta–sigma modulators, according to Lemma 2, is shown in Fig. 2(c). The equivalence under the modulo operator proves especially useful in analyzing the randomness properties of the sequence $x_i(k) = f_{\mathrm{mod}}(x''_i(k))$ from linear analysis of the sequence $x''_i(k)$.

The quality (or "randomness") of random vectors generated in an array implementing (1) generally depends on the specifics of the neighborhood template $\mathcal{N}(\cdot)$ and the constants α and β, besides the form of the boundary conditions applied at the periphery of the array. Our experience with alternative structures has shown that not much is to be gained over the simple linear structure (4) in Fig. 2(a) by increasing the complexity of implementation with a template size larger than two. We limit the analysis to this linear cascade structure, which is conveniently implemented in analog VLSI as shown in the following sections.

We consider two special cases of boundary conditions for the cascade structure of N cells $x_1 \cdots x_N$: a "chain" topology with a constant input supplied to the first element x_1, and a "ring" topology with cyclic boundary conditions where the output of the last element x_N feeds into the input of the first x_1 [$x_0 \equiv x_N$ in (4)]. The ring structure is preferrable because of symmetry which provides more uniform random noise properties across the array, although stability of noise shaping in the feedback loop is an issue which will be addressed below.

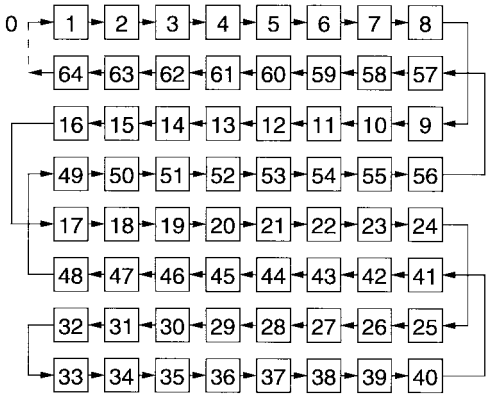

Fig. 3. Array of 64 MASH random generating cells. Linear cascaded chain or ring topology implemented on a 2-D grid.

The linear cascaded chain and ring topologies can be implemented in scalable cellular VLSI architectures on either a one-dimensional (1-D) and 2-D grid. To realize the chain and ring topologies on a 2-D grid, shown in Fig. 3, two sets of linear cascade segments are interleaved in opposing directions, and external connections at the periphery of the array span no more than two adjacent cell spacings on the grid.

Before proceeding to the experimental results from a VLSI implementation, we first formally establish the randomness properties, serving as a theoretical foundation of the "why" and "how" of this work and of some of the arguments made above.

III. Noise-Shaping Properties

Although the following analysis is performed in the special case of the cascade of delta–sigma modulators (4), it is equally applicable and straightforwardly extended to the general case (1) for integer values of β. The analysis addresses both spatial and temporal aspects of randomness: the statistical properties of a cell's outcome conditional on the history of its neighbors and itself, and the randomness of the time series $x_i(k)$ from linear analysis of $x''_i(k)$.

A. Inter-Cell Statistics

The joint cell statistics can be analyzed using mathematical tools operating on *densities*, such as outlined in [43], and extended to the multivariate case involving conditional probablities. The probability density of a cell i's outcome at time k conditional on cell j's outcome at time l is formally defined as

$$p_{i,j}^{k,l}(x|y) \stackrel{\mathrm{def}}{=} \lim_{\Delta x \to 0} \frac{\Pr(x < x_i(k) < x + \Delta x | x_j(l) = y)}{\Delta x}. \quad (5)$$

Similarly, the unconditional probability density is defined as

$$p_i^k(x) \stackrel{\mathrm{def}}{=} \lim_{\Delta x \to 0} \frac{\Pr(x < x_i(k) < x + \Delta x)}{\Delta x} \quad (6)$$

which also can be expressed as $\int_{-1}^{1} p_{i,j}^{k,l}(x|y)p_j^l(y)\,dy$ by application of Bayes' rule.

Clearly, the outcome of cell i at time $k+1$ only depends on those of cells i and $i-1$ at time k as determined by (4), or equivalently

$$x_i(k+1) = f_{\mathrm{mod}}(x_i(k) + x_{i-1}(k) + 1) \tag{7}$$

which after applying rules of modulo arithmetic can be reformulated as

$$\begin{aligned} x_{i-1}(k) &= f_{\mathrm{mod}}(x_i(k+1) - x_i(k) + 1) \\ x_i(k) &= f_{\mathrm{mod}}(x_i(k+1) - x_{i-1}(k) + 1) \end{aligned} \tag{8}$$

and thus

$$\begin{aligned} p_{i,i}^{k+1,k}(x|y) &= p_{i-1}^k(f_{\mathrm{mod}}(x-y+1)) \\ p_{i,i-1}^{k+1,k}(x|y) &= p_i^k(f_{\mathrm{mod}}(x-y+1)). \end{aligned} \tag{9}$$

In particular, if element $i-1$ has uniform probability density $p_{i-1} \equiv 1/2$, then so does element i, and furthermore the conditional probability densities of i, both on the previous values of i and $i-1$, are uniform as well

$$\begin{aligned} p_{i,i}^{k+1,k}(x|y) &= \tfrac{1}{2} \equiv p_i^{k+1}(x) \quad \forall x,y \in [-1,1]; \quad \forall k \\ p_{i,i-1}^{k+1,k}(x|y) &= \tfrac{1}{2} \equiv p_i^{k+1}(x) \quad \forall x,y \in [-1,1]; \quad \forall k. \end{aligned} \tag{10}$$

In other words, a uniform density (not necessarily random) at the input $i-1$ implies a uniform density at the output i, *unconditional* on previous values of the input as well as the output itself. This establishes, in a *statistical* sense, that "source" i is "random," and independent of the input $i-1$. (Dynamical aspects of randomness are the subject of next section.)

The same result can also be obtained directly by applying the Frobenius–Perron density operator [43] of transformation (7) to the conditional probabilities, for which uniform densities are a stationary solution. Property (10) is then easily extended to the general case of $p_{i,j}^{k,l}$. Because of causality, already $p_{i,j}^{k,l}(x|y) \equiv p_i^k(x) = 1/2$ for $j<i$ or for $l \le k$ by construction. The other combinations of i,j and k,l are obtained by induction on the markov chain that characterizes the cascade of cells over time, by repeated application of the Frobenius–Perron operator. As said, a cell's outcome depends directly on the outcomes of its immediate predecessor and itself, and so the markov property is a valid assumption. This induction, across the chain and over time, then establishes the proof of the following Theorem:

Theorem 1: The vector sequence $x_i(k), i = 1\cdots N,\ k = 1\cdots\infty$ obtained from a cascade of modulators according to (4) with initial conditions $-1 \le x_i(0) \le 1$ and boundary conditions $x_0(k)$ drawn from a uniform random distribution, i.e.: $p_{0,0}^{k,l}(x|y) \equiv p_0^k(x) = 1/2, \forall x,y \in [-1,1]; k \ne l$, follows a uniform random distribution with mutually statistically independent components, i.e.: $p_{i,j}^{k,l}(x|y) \equiv p_0^k(x) = 1/2, \forall x,y \in [-1,1]$; where either $k \ne l$ or $i \ne j$.

The theorem states that when a uniform random input is supplied to the first element in a chain of MASH cells, all cells in the chain also produce uniform random outputs, which are statistically uncorrelated not only with the supplied input, but also across different cells and over time. In other words, a single random supply generates an array of random sources, with the desirable property of spatial and temporal statistical independence across all channels.

Of course, a system generating white noise from white noise at the input defeats the purpose of a random *generator*. The following observation relaxes the requirement of a random input to the first stage. Let p_{i-1}^k be an arbitrary probability density distribution, feeding into cell i. Then the distribution p_i^k is determined by the following recursive integral equation:

$$\begin{aligned} p_i^{k+1}(x) &= \int_{-1}^{1} p_{i,i}^{k+1,k}(x|y)p_i^k(y)\,dy \\ &= \int_{-1}^{1} p_{i-1}^k(f_{\mathrm{mod}}(x-y+1))p_i^k(y)\,dy \\ &= \int_{-1}^{1} p_{i-1}^k(z)p_i^k(f_{\mathrm{mod}}(x-z+1))\,dz \end{aligned} \tag{11}$$

where $z = f_{\mathrm{mod}}(x-y+1)$. Clearly, a valid steady-state solution of the recursion is $p_i^k = p_i^{k-1} \equiv 1/2$. It is easy to show that this is the only solution of the linear integral equation (the asymptotic solution reached for $k \to \infty$) in case p_{i-1}^k is strictly positive over the entire interval.

Thus, for any strictly positive, continuous-density distribution $p_0^k(x)$ at the input, the first cell in the cascade is already "random" with a uniform probability density 1/2. Furthermore, the first cell's outcome is then independent of the input as in the case of a random input, from (9): $p_{1,0}^{k+1,k} \equiv p_1^{k+1} = 1/2$. The only effects of nonrandomness in the input x_0 on the statistics of the other x_i are sequential correlations in x_1. Subsequent channels $i>1$ are random, uncorrelated, and independent as under the conditions of Theorem 1. This is summarized in the following corollary.

Corollary 1: The vector sequence $x_i(k), i = 1\cdots N, k = 1\cdots\infty$ obtained from a cascade of modulators according to (4) with initial conditions $-1 \le x_i(0) \le 1$ and boundary conditions $-1 \le x_0(k) \le 1$, drawn from an everywhere strictly positive but otherwise arbitrary density distribution p_0, follows a uniform random distribution with mutually statistically independent components except for $i = 1$, i.e.: $p_{i,j}^{k,l}(x|y) \equiv p_0^k(x) = 1/2, \forall x,y \in [-1,1]$; where either $k \ne l$ or $i \ne j$, excluding the singular case $i = j = 1$ and $k = l+1$.

An example of a distribution p_0 which violates this condition is a DC zero input, $x_0(k) \equiv 0$ which clearly produces statistical anomalies in the outputs $x_i(k)$, especially if the initial conditions $x_i(0)$ are zero as well. Such artifacts are not of much concern in analog implementations where zero-valued regions in probability distributions are excluded because of noise and parameter fluctuations.

B. Dynamics of Chain and Ring Topologies

Clearly, a purely statistical analysis of a purely deterministic system cannot capture all effects which distinguish a "good" random generator from a lousy one. The emphasis in the previous section was on statistical independence across channels and over time. Next we study the dynamics of the channel outputs, as determined by by the topology of the cascade of cells.

Equation (4) forms a dynamical system of which the noise-shaping properties can be studied in closed form using standard linear analysis techniques. The analysis here generalizes previous results on quantization noise in a cascade of MASH delta–sigma modulators [33] mainly in two ways: an arbitrary input presented to a chain of cells, and a closed ring of cells.

Lemma 2 allows rendering of (4) and (7) into an equivalent linear form, illustrated in Fig. 2. To further simplify analysis, the unit offset term shown in Fig. 2(c) is eliminated by the substitution of variables in (7)

$$x_i(k) = f_{\mathrm{mod}}(\tilde{x}_i(k) - 1) \tag{12}$$

yielding a set of homogeneous linear equations

$$\tilde{x}_i(k+1) = \tilde{x}_i(k) + \tilde{x}_{i-1}(k) \tag{13}$$

transformed into the z-domain as

$$\tilde{X}_i(z) = \frac{1}{z-1}\tilde{X}_{i-1}(z). \tag{14}$$

Each cell i is thus represented by an accumulator with transfer function $1/z - 1$.

1) Linear Chain: For a chain topology, the output of cell i is determined by

$$\tilde{X}_i(z) = \left(\frac{1}{z-1}\right)^i \tilde{X}_0(z) \tag{15}$$

which, transformed back in the time-domain, can be expressed in terms of the input sequence $\tilde{x}_0(k)$ and initial conditions $\tilde{x}_i(0)$ as

$$\tilde{x}_i(k) = \sum_{l=0}^{k-i} \tilde{x}_0(l) C_{i-1}^{k-l-1} + \sum_{j=1}^{i} \tilde{x}_j(0) C_{i-j}^{k} \tag{16}$$

where $C_p^q \stackrel{\mathrm{def}}{=} q!/p!(q-p)!$, denoting binomial coefficients.

An important observation here is that the series (16) diverges for $k \to \infty$, and more strongly so for larger i: $\tilde{x}_i(k)$ tends either to $\mathcal{O}(k^i)$ or $\mathcal{O}(k^{i-1})$, depending on the particular input sequence $\tilde{x}_0(k)$. The effect of the folding nonlinearity f_{mod} in (12) on (16) produces an output $x_i(k)$, highly sensitive to initial conditions. This might seem to indicate the presence of chaos, although this is not the case since the chain topology has a zero Lyapunov exponent (all poles are located at $z = 1$, or $s = 0$).

To further quantify the effect, analytical partial derivatives of $x_i(k)$ to initial and boundary values can be evaluated directly from (16). Let $S_{i,j}^{k,l}$ denote the sensitivity of $x_i(k)$ to $x_j(l)$, defined as the absolute value of the partial derivative. Note that any small change in $\tilde{x}$ corresponds to an identical change in x under transformation (12) except where $x = 0$, since the derivative of f_{mod} is unity everywhere but at its discontinuities. Thus

$$S_{i,0}^{k,l} \stackrel{\mathrm{def}}{=} \left|\frac{\partial x_i(k)}{\partial x_0(l)}\right| = C_{i-1}^{k-l-1} \approx \mathcal{O}((k-l-1)^{i-1})$$

$$S_{i,j}^{k,0} \stackrel{\mathrm{def}}{=} \left|\frac{\partial x_i(k)}{\partial x_j(0)}\right| = C_{i-j}^{k} \approx \mathcal{O}(k^{i-j}) \tag{17}$$

is valid where $x_i(k) \neq 0$ and $x_j(0) \neq 0$ or $x_0(l) \neq 0$. As expected, the sensitivity of the channel output $x_i(k)$ to previous outputs $x_j(0)$ and distant outputs $x_0(l)$ increases strongly with time k and with distance i.

A second important observation is that because all binomial coefficients are integers, a *continuously* uniform distribution of $x_i(k)$ requires that at least one of the terms $\tilde{x}_0(l)$ or $\tilde{x}_j(0)$ in (16) be an irrational number. When this condition is met (by making at least one $x_0(l)$ or $x_j(0)$ irrational within the $[-1, 1]$ range), it can be shown (using similar methods as employed in [33]) that the output modulation sequence $x_i(k)$ is uniformely distributed and uncorrelated with the input $x_0(k)$.

2) Closed Ring: A closed ring topology with N modulator cells can be considered as a special case of a linear chain where $x_0 \equiv x_N$. In principle, since the above analysis for the chain topology does not assume any structure for $x_0(k)$ (other than it be contained in the modulation range interval), the above conclusions apply here as well. However, the modulation dynamics for a ring topology is entirely different by nature of the feedback. Closing the loop implies

$$\tilde{X}_i(z) = \left(\frac{1}{z-1}\right)^N \tilde{X}_i(z) \tag{18}$$

for any cell i. The complex eigenvalues z_n in the z-domain, satisfying $((1/z_n - 1))^N = 1$, are given by

$$z_n = 1 + e^{j2\pi(n/N)} \qquad n = 0, \cdots N-1 \tag{19}$$

and the eigenvectors, satisfying $X_i^n = [1/(z_n - 1)]X_{i-1}^n$, are correspondingly

$$X_i^n = e^{-j2\pi(in/N)}. \tag{20}$$

The general solution in the time-domain is then given by

$$\tilde{x}_i(k) = \sum_{n=0}^{N-1} c_n X_i^n (z_n)^k \tag{21}$$

where the complex constants c_n are expressed in terms of the initial conditions $\tilde{x}_j(0)$ as

$$\tilde{x}_i(k) = \sum_{j=0}^{N-1} \tilde{x}_j(0) \left(\frac{1}{N} \sum_{n=0}^{N-1} X_i^n X_j^{n*} (z_n)^k\right). \tag{22}$$

The dominant dominant eigenvalue is $z_0 = 2$, and the series diverges rapidly if at least one of the $\tilde{x}_j(0)$ is nonzero: $\tilde{x}_i(k) \approx \mathcal{O}(2^k)$ as $k \to \infty$. The ring topology has clearly a strict positive Lyapunov exponent, and therefore exhibits "deterministic chaos" according to the standard definition. The sensitivity to the initial values is

$$S_{i,j}^{k,0} = \frac{1}{N} \sum_{n=0}^{N-1} X_i^n X_j^{n*} (z_n)^k \approx \mathcal{O}(2^k). \tag{23}$$

Notice that for any i, there is a k' for which $2^k > k^i$ for all $k > k'$, and thus any channel in a ring is more sensitive to previous and distant channels, on the long term, than any channel in a chain of arbitrary length.

As long as at least one real or imaginary components of z_n is irrational, so is the sequence $\tilde{x}_i(k)$ regardless of the initial values $\tilde{x}_j(0)$ (provided one of them is nonzero), and the modulation output is uniformly distributed. Such is the case for $N = 3$ and $N > 4$. There is, however, one artifact that potentially terminates the random output under certain circumstances. This artifact exists only for even values of N, and is due to the presence of a zero eigenvalue for $n = N/2$. The corresponding "terminator" eigenvector is real with constant amplitude, alternating sign from one cell to the next. If at any time k the state of the ring modulator $\tilde{x}_i(k)$ reduces to any scaled version of this eigenvector or any of the equivalent vectors on the grid invariant to the f_{mod} transformation, then the modulator halts in the zero state, starting in the next cycle $\tilde{x}_i(l) \equiv 0$ and $x_i(l) \equiv \pm 1$, for all $l > k$. The chance of this artifact actually occuring is virtually zero in an analog implementation, but is significant in a digital implementation with fixed-point arithmetic.

C. Analog versus Digital

Generating truly random numbers from a deterministic, infinitely precise, limited resolution system such as a digital computer is clearly utopia. Generating "pseudo-random" sequences with random-like properties has been a challenge to computer scientists that involves intricate mathematics such as number theory and the like [27]. The random properties of analog counterparts of these random generators are drastically different, owing to the infinite resolution of the analog representation, and the intrinsic randomness of additive noise in the physical implementation.

All of the complications that have arisen in the above study of the dynamics of coupled modulators are due to artifacts specific to discrete (digital) rather than continuous (analog) systems. For instance, in analog implementations, the realization of rational numbers has zero probability, and physical noise washes away any of the effects of close proximity to a rational number (analogous, say, to the role of noise in flipping or flopping a metastable flipflop). On the other hand, *any* number represented in a digital computer is rational, and "random" sequences are guaranteed to repeat themselves, with a period which in the *best* case equals the number of different discrete states (minus one) [28]. Similarly, the artifact of the zero eigenvalue in a ring topology with even number of cells affects digital implementations only. In analog, the chance of precisely hitting a member of the family of the "terminator" eigenvector is again zero, and even if one gets close enough, the devastating effect of the aftermath is washed away quickly by virtue of the additive physical noise injected into the system.

The role of additive noise and dithering in delta–sigma modulation has been studied extensively, e.g. [35], [36]. For the ring and chain topologies of MASH modulators studied above, the effect of additive noise in an analog implementation can be quantified as follows. Let $n_j(l)$ represent the root-mean-square noise contributed physically to the $x_j(l)$ variable at time l. Then the long-term, cumulative effect onto variable $x_i(k)$ at time k, in terms of the root-mean-square deviation $d_i(k)$ from the noiseless case, can be estimated from the sensitivities $S_{i,j}^{k,l}$ as

$$d_i(k) = \left(\sum_{l=0}^{k} \sum_{j=0}^{N-1} \left(S_{i,j}^{k,l} n_j(l) \right)^2 \right)^{1/2} \tag{24}$$

increasing sharply with k, and especially so for the ring topology. Even under arbitrary low noise conditions, the deviation from the nominal dynamics in $x_i(k)$ as studied above grows rapidly, and renders the output unpredictable after a number of cycles k, as the cumulative noise amplitude $d_i(k)$ outgrows the nominal signal level $x_i(k)$.

Notice that the compressive nonlinearity f_{mod} folds any excess signal back into the $[-1, 1]$ range. Since $d_i(k)$ quickly goes to infinity, the folding effectively *shapes* the amplified noise $d_i(k)$ into a uniform distribution regardless of the distribution of the generating $n_j(l)$ noise sources. This type of noise-shaping is in addition to the deterministic noise-shaping of quantization noise in the delta–sigma modulators, and adds a truly random component to the output, absent in digital implementations. In analog implementations, noise shaping of this type avoids problems that arise in systems which directly amplify the noise contributed by physical sources, i.e., a nonuniform amplitude distribution affected by $1/f$ noise and temperature variations, and correlation across channels due to parasitic coupling. By nature of the properties studied above, an analog array of MASH modulators shapes the noise, contributed by the quantizers and by physical sources, into mutually independent, uniformly distributed, random channels.

IV. Implementation

The array of MASH modulators can be implemented in a variety of VLSI technologies, digital and analog. The architecture using modulo arithmetic depicted in Fig. 2(b) is most suitable for digital implementations, using carry-free adders and registers. Analog implementations are more served by the structure of Fig. 2(a), since the nonlinearity $f_{\Delta\Sigma}$ is readily implemented using a high-gain comparator and differencing circuitry.

The discussion in the foregoing section has made the advantages of an analog implementation clear. In addition, analog VLSI implementation offers potentially higher integration density and higher energy efficiency than equivalent digital VLSI implementations. Unlike more convential analog designs, where high precision and high noise rejection are primary design constraints, the circuits can be implemented with minimum-size, noisy and imprecise components biased at lower currents. Effects of component noise and mismatches in the implementation are discussed below.

A. Architecture

The analog VLSI implementation presented here targets applications of on-line random generation integrated in VLSI

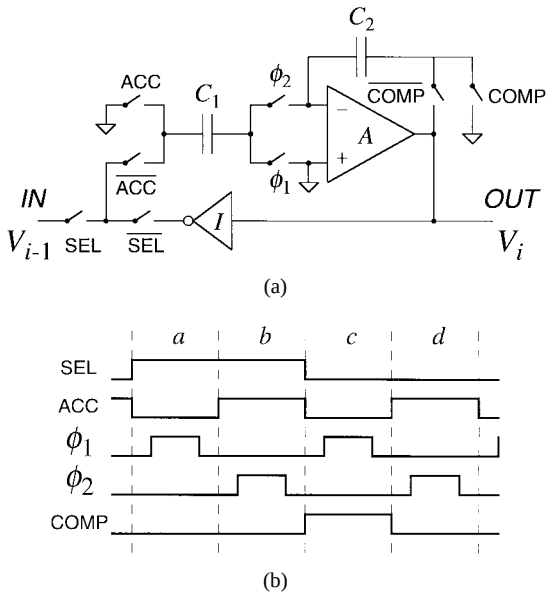

Fig. 4. Switched-capacitor MASH modulator cell. (a) Simplified circuit diagram. (b) Timing diagram.

systems for information and signal processing, where a steady stream of random variables are required locally, in parallel, and in analog format. We have adopted an implementation style using low-power, high-density switched-capacitor circuits, producing a voltage output format $V_i(k)$. Compact alternative realizations using current-mode technology, such as used in [30], can be derived as well.

The switched-capacitor architecture implementing the MASH cell in Fig. 2(a) is shown in Fig. 4(a), and the corresponding signal timing diagram is given in Fig. 4(b). The state $V_i(k)$, corresponding to $x_i(k)$ in Fig. 2(a), is stored across capacitor C_2. To save power and silicon real estate, the amplifier A serves the dual purpose of accumulator $1/(z-1)$ and quantizer $f_{\Delta\Sigma}$, controlled by the COMP signal. When COMP is active (high), amplifier A compares V_i with zero, and presents the result (sign of V_i) to the accumulator input through the inverting of one-bit D/A converter I. When COMP is inactive (low), $V_i(k)$ is presented to the output OUT. The accumulator functions as a standard switched-capacitor noninverting integrator [39], where in the sampling phase the capacitor C_1 is precharged to the input, and in the accumulate phase this charge is transferred onto capacitor C_2. This operation is controlled with signals ACC, ϕ_1 and ϕ_2, repeated twice as shown in Fig. 4(b) (phases *a–b*, and *c–d*). The input to the accumulator is controlled by the SEL signal, which first selects the output from the preceding stage $V_{i-1}(k)$ presented to IN, and then the output from the comparator. The four-phase operation is summarized as follows:

1) sample input $V_{i-1}(k)$ from previous stage;
2) accumulate;
3) compare with zero and sample inverted result;
4) accumulate, yielding $V_i(k+1)$.

Functionally, the first accumulate produces $x_i(k)+x_{i-1}(k)$, and the second accumulate subtracts the sign of the first. The net operation thus yields $f_{\Delta\Sigma}(x_i(k)+x_{i-1}(k))$ as desired.

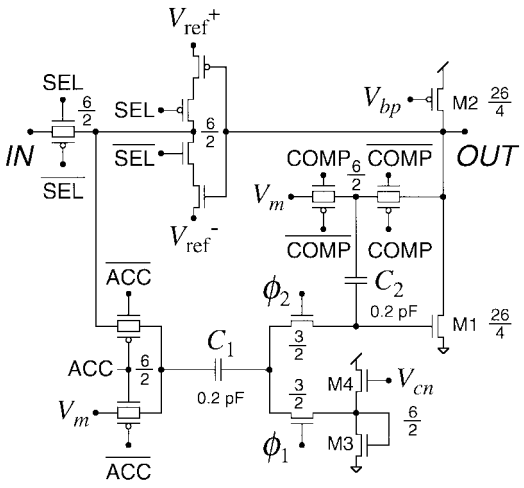

Fig. 5. CMOS switched-capacitor circuit diagram of MASH modulator cell.

B. CMOS Implementation

The transistor-level circuit diagram of the MASH cell is shown in Fig. 5. For low-power operation and compatibility with digital interface circuitry, the circuit uses a single supply V_{dd}, set to 5 V for the experiments. The signal ground level is set to $V_m = 2$ V, and the signal range is ± 1 V as determined by the D/A levels, $V_{\text{ref}}^- = 1$ V and $V_{\text{ref}}^+ = 3$ V, symmetric around V_m. Thus

$$V_i(k) = V_m + V_{\text{range}} x_i(k) \tag{25}$$

where $V_{\text{range}} = 1$ V.

The amplifier A is implemented as a single, noncascoded pseudo-nMOS inverter M1–M2. The relatively low gain of this design is adequate for the purpose of a random generator, where linearity and gain errors are less important than power dissipation and size. The virtual ground voltage of the amplifier, used for the precharge in the sampling phase of the accumulator, is obtained from circuit M3–M4, of which the V_{cn} bias is generated from V_{bp}. The reason for not precharging directly from the unity gain connected amplifier is because the accumulator output is needed simultaneously to precharge the next cell, occupying the amplifier. This introduces $1/f$ noise in the accumulator, which otherwise would have been cancelled by a correlated double-sampling technique. Finally, the capacitances C_1 and C_2 are 0.2 pF in 2 μm technology, enough to provide adequate matching, and to avoid excessive switch-injection and clock-feedthrough noise contributed by the switches ϕ_1 and ϕ_2.

The layout of the cell, measuring $100\lambda \times 120\lambda$ in a double-metal, double-poly CMOS process, is shown in Fig. 6. The second poly is used for capacitors only, which can be replaced by MOS gate oxide capacitors. In 0.25 μm CMOS technology, this layout supports the integration of over half a million cells on a single 1 cm^2 chip, although this obviously excludes any circuitry actually using the array of random numbers.

C. Sources of Imprecision and Noise

The scaling of the technology brings up issues of circuit noise and mismatch, and their effect on performance. Sources

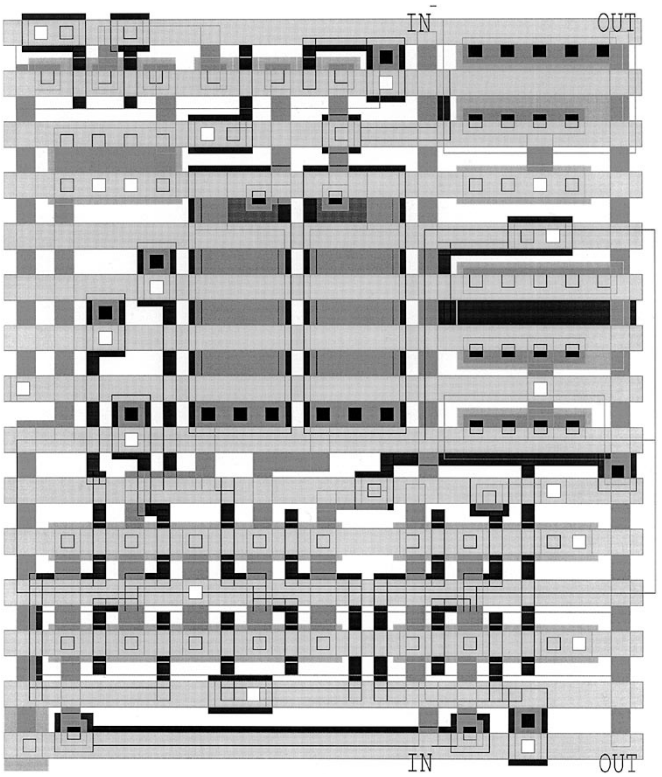

Fig. 6. Layout of switched-capacitor MASH modulator cell. Dimensions are $100\lambda \times 120\lambda$ in MOSIS scalable CMOS technology.

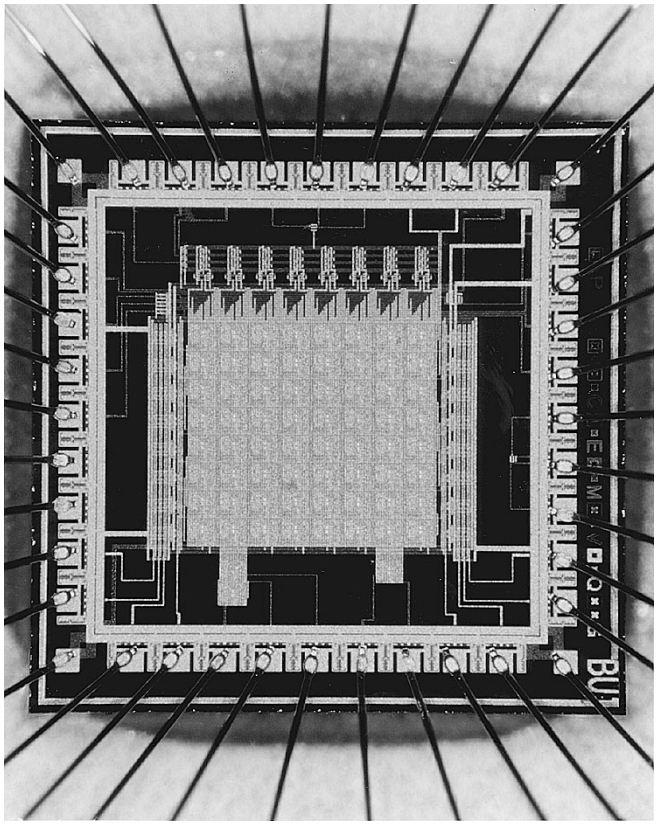

Fig. 7. Micrograph of the $64(+1)$-channel VLSI parallel random analog vector generator, including an 8×8 array of MASH modulators plus one extra modulator. Dimensions are 2.22 mm $\times$ 2.25 mm in 2 μm CMOS.

of noise are mainly thermal noise as contributed by the capacitances C_1 and C_2, and $1/f$ noise of the amplifier M1–M2 and the virtual ground circuit M3–M4. Both scale inversely with λ. Noise is an issue mostly for applications requiring reproducible "random" sequences for given initial state conditions, in particular for analog encryption and secure communications. However, increased noise is desirable (or at least not undesirable) when quality of randomness is the only concern.

The most significant imprecisions in the implementation are the finite gain of the amplifier and the mismatch between the two capacitors, which affect the gain and linearity of the cell transfer function. The results of Theorem 1 and Corollary 1 hold only for integer values of the gain β in (1), and a small deviation from $\beta = 1$ introduces slight nonuniformities in the conditional probabilities $p_{i,i}^{k+1,k}$ and $p_{i,i-1}^{k+1,k}$ in (9), even though the unconditional probabilities p_i^k are mostly unaffected. The effect of β mismatch on statistical dependence is thus constrained locally in time and space, and can be virtually eliminated by "oversampling" in time or space. The effect of gain nonlinearity is qualitatively similar. Effects of small gain and nonlinearity errors on the dynamics of the modulator chain and ring topologies are of minimal impact. The effect of gain variations on the dynamics can be formally analyzed from sensitivities (17) and (17) after inclusion of the appropriate β terms.

Switch-injection noise and clock-feedthrough in the switches and transistor mismatches contribute an offset error α which does not affect performance. As a matter of fact, any of the properties studied above still apply when an arbitrary offset α is included in (1), under a transformation of variables similar to (12).

V. Experimental Results

Fig. 7 shows a micrograph of the tiny (2.22 mm $\times$ 2.25 mm) 2 μm CMOS chip prototyped through MOSIS, which integrates a 2-D array of 64 MASH cells configured as shown in Fig. 3, plus one extra MASH cell and additional test circuitry. Of the 64 channels, two can be randomly accessed at the same time by means of two independent sets of horizontal and vertical address-decoded multiplexers. This allows characterization of any pair of channels simultaneously.

All experimental results reported in this paper were obtained from this chip. Experiments were performed on chain and ring topologies with 64 and 65 cells, using the array and the extra modulator. Although the theory predicts differences in dynamical properties for rings with even and odd number of cells N, these effects are unobservable in the data obtained for 64 and 65 cells. The data shown is limited to the case $N = 65$, and the complete data set is available on request.

Iterative Map: The measured iterative map and transfer characteristic of a single MASH modulator cell, implementing (4), is shown in Fig. 8, for a spectrum of input and output voltages in the range of the $[V_{\text{ref}}^-, V_{\text{ref}}^+]$ interval. The combined gain errors are in the order of 5%, and their effect on the random statistics, as anticipated above, is evaluated next.

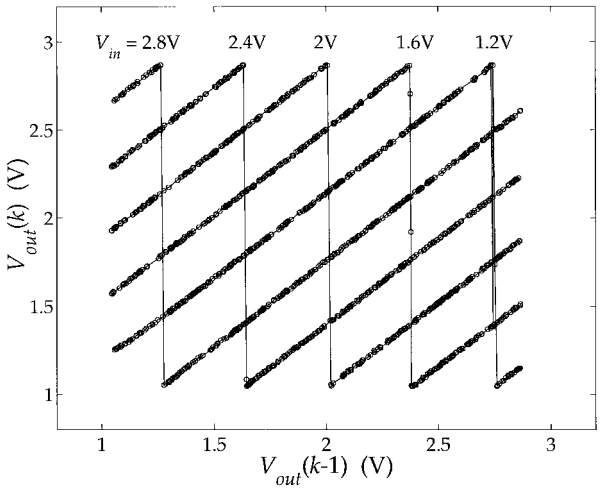

Fig. 8. Measured Iterative map of a single MASH modulator cell.

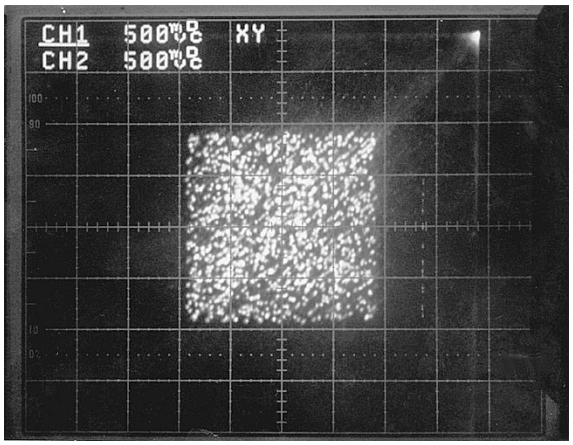

Fig. 9. Oscilloscope X–Y plot showing outputs from two neighboring channels in the 65-channel ring.

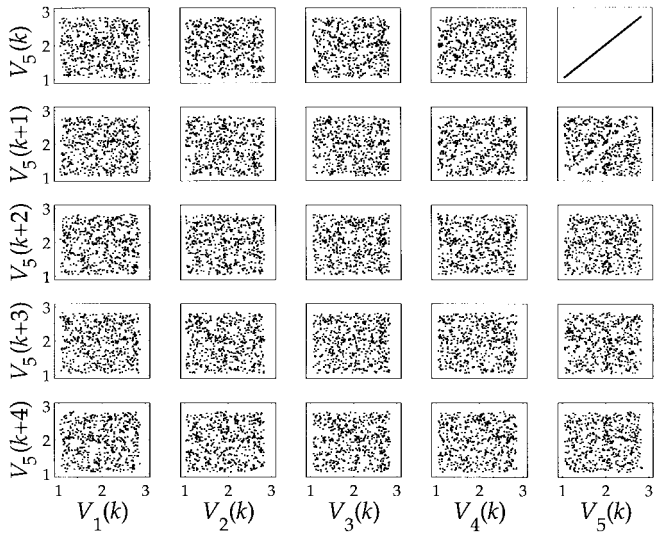

Fig. 10. Measured time-space correlogram. Scatter plots of data from the 65-cell ring, across five neighboring cells and five consecutive time delays.

Statistics: The hypothesis of statistical independence $p_{i,j}^{k,l}$ across channels and over time, theorized in Section III-A, was tested experimentally, illustrated in Fig. 9 for two concurrent neighboring outputs, and shown in further detail in Fig. 10,

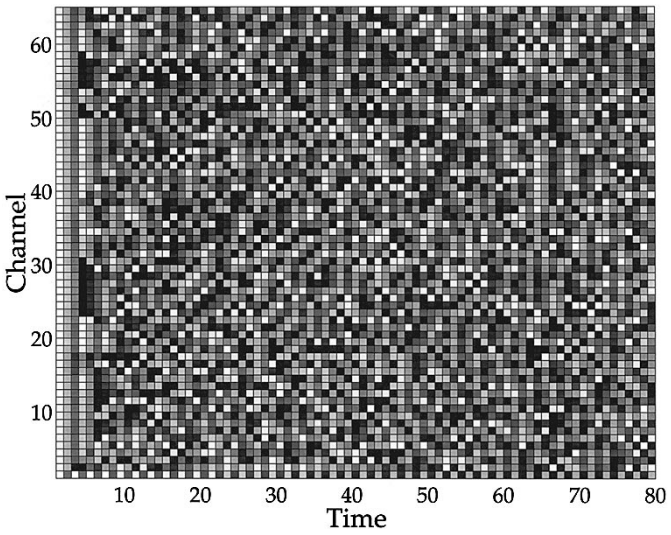

Fig. 11. Measured transient dynamics of the 65-cell chain, from zero-level (V_m) initial conditions, and for zero-level (V_m) input to the first cell.

recorded over several consecutive channels and time delays on the 65-cell ring. The graphs show $x_i(k)$ versus $x_j(l)$, producing a scatter plot corresponding to the joint probability density which ideally should be uniform. The only perceptable effects of any (nontrivial) correlations across outputs appear to be between consecutive values in two neighboring channels, $x_5(k+1)$ versus $x_4(k)$ and $x_5(k)$. As anticipated in Section IV-C, the linear gaps in the otherwise uniform distribution in the $x_5(k)$-$x_5(k+1)$ and $x_4(k)$-$x_5(k+1)$ graphs are due to a value of the gain β less than one, which cause certain values in the $[-1,1]$ interval to be inaccessible to $x_i(k+1)$ from a given $x_i(k)$ or $x_{i-1}(k)$ initial value. If this nonuniformity is a serious issue, it can be reduced by increasing the open-loop gain of the amplifier and compensate for various mismatches. Alternatively, a simpler remedy is *oversampling* either in space or in time, e.g., skipping every other sample in the sequence or every other cell in the cascade. Notice that even without such methods, the quality of random vector generation obtained "as is" should be more than adequate for most purposes, especially given that conventional designs based on congruent linear recursions or other iterative maps are by construction entirely deterministic over time (the $x_i(k)$-$x_i(k+1)$ scatter plot condenses to a solid curve).

For a 65-cell chain topology with constant zero-level input to the first cell, the obtained results were qualitatively similar, except for the first few channels which showed systematic correlations, due to transient effects studied next.

Dynamics: The effect of chain and ring topologies on the transient and steady modulation dynamics, recorded across the entire 65-cell cascade over 80 time steps, is illustrated in Figs. 11 and 12. In both cases, transient effects due to near-zero initial conditions for all cells, are clearly visible for the first few ($\approx$5) cycles. For the chain topology in Fig. 11, the first few ($\approx$3) cells display a tendency to limit cycle oscillations, due to the degenerate effect of a near-zero DC input to the first cell as discussed in Sections III-A and IV-C. Power density spectral analysis of experimental data over a 1024-point rectangular time window, shown in Fig. 13, reveals that effects of limit

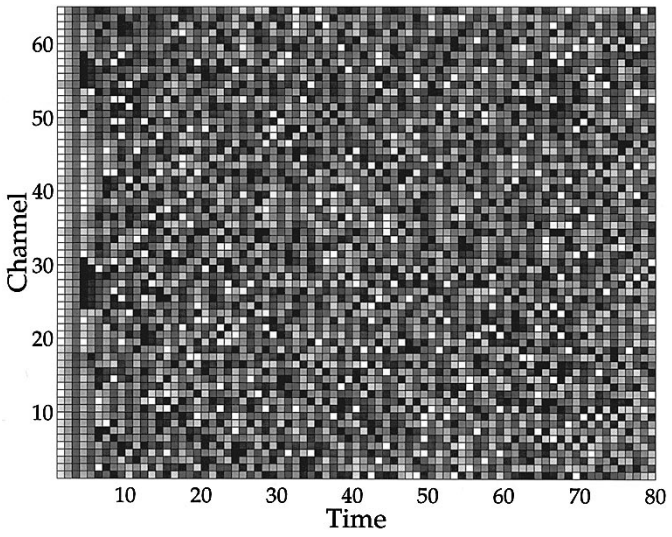

Fig. 12. Measured transient dynamics of the 65-cell ring, from zero-level (V_m) initial conditions.

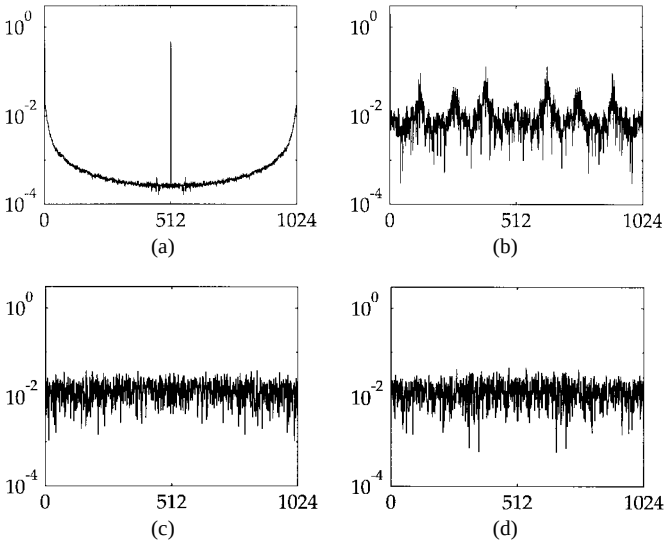

Fig. 13. Discrete Fourier power density spectra of recorded cell sequences. (a) First cell in the 65-cell chain, with zero-level (V_m) input. (b) Second Cell. (c) Fourth Cell. (d) 65-cell ring.

cycle oscillations or other colored spectral features are limited to no more than the first three stages of the chain, and are absent in the ring.

System-Level Issues: The operation of the chip has been verified over a range of speeds from 2 Ksamples/s to 50 Ksamples/s per channel. The maximum of 50 Ksamples/s per channel obtained is affected by external capacitive loading of the (multiplexed) output which has not been buffered. Measurements of supply currents yield power dissipation levels ranging from 16 μW to 245 μW per cell, corresponding to 6 nJ of energy dissipated per sample. Extrapolation of these results are technology dependent; scaling of the technology (from 2 μm to deep submicron feature sizes) would allow to further reduce energy consumption and increase available bandwidth at least proportionally. Increased circuit noise levels caused by the downscaling of the capacitors are a concern only to the extent to which reproducibility of the modulation output sequences is more important than their randomness properties. In any case, by the folding nature of the modulation as discussed in Section III, the noise at the output is shaped into a uniform random distribution independent of the structure of the injected circuit noise.

VI. Conclusions

Delta–sigma modulation and modulo arithmetic have been combined in a cellular architecture for parallel analog random vector generation, with statistical properties that are unique to the functional form of interactions across cells and not found in arrays of separate analog random generators: statistical independence over time and across channels. We have formally analyzed these properties for two architectures (chain and ring topologies), and experimentally verified the results on a 65-channel VLSI prototype.

As in the linear congruential map, the modulo operation produces a uniform random amplitude distrubution for each of the cell state variables. However, the contribution by the neighboring cells to the modulo sum scrambles the dependency of the cell's state on its previous value, and we have shown that this produces a uniform random sequence void of sequential correlations so characteristic of the linear congruential generator and other iterative chaotic maps. We have also shown that the modulo sum nature of the coupling between cells avoids any correlations between cells across the array. The randomizing effect of this strong nonlinear coupling is fundamentally different from mode-locking and synchronization phenomena that arise in arrays of weakly coupled oscillators.

The functional equivalence between linear congruent modulation and delta–sigma modulation offers elegant circuit implementations both in analog and digital VLSI technologies. Owing to noise-shaping properties similar to quantization noise in delta–sigma modulation, the amplification and modulation of physical noise present in an analog implementation generates truly random, nonperiodic sequences with statistics that are guaranteed uniform, void of some anomalies that potentially occur when using limited-resolution, infinite-precision arithmetic as in a digital implementation. The amount of injected physical noise clearly determines the extent in time to which the random sequences are reproducible from identical initial conditions. The tradeoff between randomness and reproducibility is an important issue for applications of analog encryption. We have performed a sensitivity study based on an analytical model of the dynamics, quantifying uncertainty in the sequence as a function of time and distance from given initial or boundary conditions.

Finally, results from a fabricated 64(+1)-channel prototype in 2-μm CMOS technology confirm the theoretical results, and indicate that effects of component mismatches and other circuit imperfections are not detrimental to the statistics and dynamics of the modulation sequences. The implementation architecture is equivalent to a MASH cascade of delta–sigma modulators, of which the noise-shaping properties inspired much of this work. While the particular switched-capacitor design used to implement the MASH modulator serves for demonstration purposes only, the cell layout supports the integration of over half a million random generators on a 1 cm^2 die in

0.2-μm CMOS technology. Besides the excellent statistical properties, the small size and low energy consumption of the random cell make it particularly well suited for large-scale integrated applications of parallel distributed analog signal and information processing, where an on-line supply of random values is embedded locally with each processing element.

REFERENCES

[1] S. C. Kak, "Overview of analog signal encryption," *Proc. Inst. Elect. Eng. F*, vol. 130, no. 5, pp. 399–404, 1983.

[2] L. J. Kocarev, K. S. Halle, K. Eckert, U. Parlitz, and L. O. Chua, "Experimental demonstration of secure communications via chaotic synchronization," *Int. J. Bifurcations and Chaos*, vol. 2, pp. 709–713, 1992.

[3] M. Göetz, K. Kelber, and W. Schwartz, "Discrete-time chaotic encryption systems—Part I: Statistical design approach," *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 963–970, 1997.

[4] V. D. Agrawal, "Testing in a mixed-signal world," in *Proc. 9th IEEE Int. ASIC Conf.* Piscataway, NJ: IEEE Press, 1996, pp. 241–244.

[5] C. L. Wey, "Alternative built-in self-test (BIST) structures for analog circuit fault-analysis," *Electron. Lett.*, vol. 27, no. 18, pp. 1627–1628, 1991.

[6] A. P. Strole and H. J. Wunderlich, "Testchip—A chip for weighted random pattern generation, evaluation, and test control," *IEEE J. Solid-State Circuits*, vol. 26, no. 7, pp. 1056–1063, 1991.

[7] J. Alspector, B. Gupta, and R. B. Allen, "Performance of a stochastic learning microchip," in *Advances in Neural Information Processing Systems*. San Mateo, CA: Morgan Kaufman, vol. 1, pp. 748–760, 1989.

[8] C. A. Mead and M. Ismail, Eds., *Analog VLSI Implementation of Neural Systems*. Norwell, MA: Kluwer, 1989.

[9] B. W. Lee and B. J. Sheu, "Hardware annealing in electronic neural networks," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 134–137, Jan. 1991.

[10] G. Premont, P. Lalanne, P. Chavel, M. Kuijk, and P. Heremans, "Generation of sigmoid probability functions by clipped differential speckle detection," *Opt. Commun.*, vol. 129, nos. 5–6, pp. 347–352, Sept. 1996.

[11] A. Dembo and T. Kailath, "Model-free distributed learning," *IEEE Trans. Neural Networks*, vol. 1, pp. 58–70, Jan. 1990.

[12] M. Jabri and B. Flower, "Weight perturbation: An optimal architecture and learning technique for analog VLSI feedforward and recurrent multilayered networks," *IEEE Trans. Neural Networks*, vol. 3, pp. 154–157, Jan. 1992.

[13] D. Kirk, D. Kerns, K. Fleischer, and A. Barr, "Analog VLSI implementation of gradient descent," in *Advances in Neural Information Processing Systems*. San Mateo, CA: Morgan Kaufman, vol. 5, pp. 789–796, 1993.

[14] G. Cauwenberghs, "An analog VLSI recurrent neural network learning a continuous-time trajectory," *IEEE Trans. Neural Networks*, vol. 7, **(author: please supply pp.)** Mar. 1996.

[15] R. C. Tausworthe, "Random numbers generated by linear recurrence modulo two," *Math. Comput.*, vol. 19, pp. 201–209, 1965.

[16] S. W. Golomb, *Shift Register Sequences*. San Francisco, CA: Holden-Day, 1967.

[17] S. Wolfram, "Statistical mechanics of cellular automata," *Rev. Mod. Phys.*, vol. 55, pp. 601–644, 1983.

[18] P. L. Venetianer, P. Szolgay, K. R. Crounse, T. Roska, and L. O. Chua, "Analog combinatorics and cellular automata—Key algorithms and layout design," *Int. J. Circuit Theory Appl.*, vol. 24, no. 1, pp. 145–164, 1996.

[19] J. Alspector, J. W. Gannett, S. Haber, M. B. Parker, and R. Chu, "A VLSI-efficient technique for generating multiple uncorrelated noise sources and its application to stochastic neural networks," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 109–123, Jan. 1991.

[20] J. Saarinen, J. Tomberf, L. Vehmanen, and K. Kaski, "VLSI implementation of Tausworthe random number generator for parallel processing environment," *Proc. Inst. Elect. Eng. E: Computers And Digital Techniques*, vol. 138, no. 3, pp. 138–146, 1991.

[21] A. J. Al-Khalili and D. M. Al-Khalili, "A controlled probability random pulse-generator suitable for VLSI implementation," *IEEE Trans. Instrum. Meas.*, vol. 39, no. 1, pp. 168–174, 1990.

[22] P. D. Hortensius, R. D. Mcleod, and H. C. Card, "Parallel random number generation for VLSI systems using cellular automata," *IEEE Trans. Comp.*, vol. 38, pp. 1466–1472, Oct. 1989.

[23] A. Dupret, E. Belhaire, and P. Garda, "Scalable array of Gaussian white-noise sources for analog VLSI implementation," *Electron. Lett*, vol. 31, no. 17, pp. 1457–1458, Aug. 1995.

[24] A. Rodriguez-Vazquez and M. Delgado-Restituto, "CMOS design of chaotic oscillators using state variables—A monolithic Chua circuit," *IEEE Trans. Circuits Systems II*, vol. 40, pp. 596–613, Oct. 1993.

[25] L. O. Chua, "Chua's circuit 10 years later," *Int. J. Circuit Theory Appl.*, vol. 4, pp. 279–305, 1994.

[26] J. E. Neely and J. G. Harris, "A chaotic oscillator cell in subthreshold CMOS for spatio-temporal simulation," in *Proc. Workshop on Spatiotemporal Models in Biological and Artificial Systems*, Sintra, Portugal, Nov. 6–9, 1996.

[27] D. Knuth, *Seminumerical Algorithms*, 2nd ed., vol. 2 of *The Art of Computer Programming*. Reading, MA: Addison-Wesley, 1981, ch. 3.

[28] W. H. Press, B. P. Flannery, S. A. Teulolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*. Cambridge, MA: Cambridge University Press, 1986, ch. 7.

[29] A. Rodriguez-Vazquez, M. Delgado, S. Espejo, and J. L. Huertas, "Switched-capacitor broad-band noise generator for CMOS VLSI," *Electron. Lett.*, vol. 27, no. 21, pp. 1913–1915, 1991.

[30] M. Delgado-Restituto, F. Medeiro, and A. Rodriguez-Vazquez, "Nonlinear switched-current CMOS IC for random signal generation," *Electron. Lett.*, vol. 29, no. 25, pp. 2190–2191, 1993.

[31] W. T. Holman, J. A. Connelly, and A. B. Dowlatabadi, "An integrated analog/digital random noise source," *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 521–528, June 1997.

[32] M. J. Bellido, A. J. Acosta, M. Valencia, A. Barriga, and J. L. Huertas, "Simple binary random number generator," *Electron. Lett.*, vol. 28, no. 7, pp. 617–618, 1992.

[33] W. Chou, P. W. Wong, and R. M. Gray, "Multi-stage delta-sigma modulation," *IEEE Trans. Inform. Theory*, vol. 35, pp. 784–796, July 1988.

[34] W. Chou and R. M. Gray, "Modulo sigma-delta modulation," *IEEE Trans. Commun.*, vol. 40, pp. 1388–1395, Aug. 1992.

[35] C. Wu and R. M. Gray, "Dithering and its effects on sigma-delta and multistage sigma-delta modulation," *IEEE Trans. Inform. Theory*, vol. 37, pp. 500–513, Mar. 1991.

[36] I. Galton, "Granular quantization noise in a class of delta-sigma modulators," *IEEE Inform. Theory*, vol. 40, pp. 848–859, May 1994.

[37] T. Hayashi, Y. Inabe, K. Uchimura, and T. Kimura, "A multistage delta-sigma modulator without double integration loop," *ISSCC Tech. Dig. Pap.*, vol. 39, pp. 182–183, 1986.

[38] K. Uchimura, T. Hayashi, T. Kimura, and A. Iwata, "Oversampled A-to-D and D-to-A converters with multistage noise shaping modulators," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP–36, pp. 1899–1905, Dec. 1988.

[39] J. C. Candy and G. C. Temes, "Oversampled methods for A/D and D/A conversion," in *Oversampled Delta-Sigma Data Converters*. Piscataway, NJ: IEEE Press, pp. 1–29, 1992.

[40] O. J. A. P. Nys and E. Dijkstra, "On configurable oversampled A/D converters," *IEEE J. Solid-State Circuits*, vol. 28, pp. 736–742, July 1993.

[41] Y. H. Yang, R. Schreier, G. C. Temes, and S. Kiaei, "Online adaptive digital correction of dual-quantization delta-sigma modulators," *Electron. Lett.*, vol. 28, no. 16, pp. 1511–1513, July 1992.

[42] G. Cauwenberghs and G. C. Temes, "Adaptive calibration of multiple quantization oversampled A/D converters," in *Proc. IEEE Int. Symp. Circuits and Systems*, Atlanta, GA, vol. I, pp. 512–515, 1996.

[43] A. Lasota and M. C. Mackey, "Chaos, fractals, and noise," in *Applied Mathematical Sciences 97*. Berlin/Vienna/New York: Springer-Verlag, 1984, ch. 1–3.

Gert Cauwenberghs (S'89–M'92) received the Engineer's degree in applied physics from the University of Brussels, Belgium, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1989 and 1994, respectively.

In 1994, he joined The Johns Hopkins University, Baltimore, MD, where he is now an Associate Professor in electrical and computer engineering. He is presently on sabbatical leave at the Center for Computational and Biological Learning, Massachusetts Institute of Technology, Cambridge, and at the Center for Adaptive Systems, Boston University, MA. His research covers analog and digital VLSI circuits, systems, and algorithms for parallel signal processing and adaptive neural computation.

Dr. Cauwenberghs received the National Science Foundation Career Award in 1997.