A Rigorous Error Analysis of D/A Conversion with Dynamic Element Matching

Ian Galton, *Member, IEEE*, and Paolo Carbone, *Student Member, IEEE*

***Abstract*—A known approach to reducing harmonic distortion in D/A converters involves a technique called dynamic element matching. The idea is not to reduce the power of the overall conversion error but rather to give it a random, noise-like structure. By reducing the correlation among successive samples of the conversion error, harmonic distortion is reduced. This paper presents the first rigorous and quantitative theoretical analysis of the conversion error introduced by an important type of D/A converter with dynamic element matching. In addition to supporting previously published experimental results that indicate the conversion error consists of white noise instead of harmonic distortion, the analysis provides an expression for the power of the white noise in terms of the power of the input sequence and the component matching errors. A yield estimation technique based on the expression is presented that can be used to estimate how the power of the white noise varies across different copies of the same D/A converter circuit for any given component matching error statistics.**

I. Introduction

IDEALLY, a digital-to-analog (D/A) converter transforms a sequence of values represented as b-bits numbers to exactly the same sequence of values represented as analog voltages. Consequently, from a signal processing point of view, an ideal D/A converter implements the identity operation on its input and is therefore a linear system. However, practical D/A converters introduce errors that cause the values represented as analog voltages to differ from the corresponding values represented as b-bits numbers. Thus, the sequence of output values can be written as $y(n) = x(n) + \epsilon(n)$ where $x(n)$ is the sequence of input values and $\epsilon(n)$ is a sequence representing the D/A conversion error. In general, $\epsilon(n)$ is a nonlinear function of the input sequence, so practical D/A converters are generally nonlinear devices.

The nonlinearity of $\epsilon(n)$ gives rise to harmonic distortion that in many applications proves to be the limiting factor in overall system performance. For example, direct digital synthesizers have been shown to be performance limited by the harmonic distortion arising from the nonlinearity of currently available D/A converters [1], [2]. Similarly, the accuracy of delta-sigma A/D converters are often limited by D/A converter linearity [3].

Interestingly, in many applications that require D/A conversion and are sensitive to harmonic distortion, the bit width, b, of the sequence to be D/A converted is not large. Therefore, relatively coarse D/A converters often suffice provided they do not introduce too much harmonic distortion. For example, it has recently been shown that direct digital synthesizers with highly linear 8-b D/A converters (i.e., D/A converters wherein $b = 8$) can achieve 90 dBc spur suppression [1]. Even coarser D/A converters are typically used in delta-sigma A/D converters [3], [4].

In principle, reducing the power of $\epsilon(n)$ reduces the harmonic distortion, but for a given D/A converter architecture the power of $\epsilon(n)$ can only be reduced at the expense of increasing the precision of the circuit technology. Hence, the limitations of the circuit technology generally place a lower bound on the power of $\epsilon(n)$. An alternative approach to reducing harmonic distortion in D/A converters involves a technique called *dynamic element matching* [5]–[9]. The idea is not so much to reduce the power of $\epsilon(n)$, but rather to give it a random, noise-like structure. By reducing the correlation among successive samples of $\epsilon(n)$, harmonic distortion is reduced. A particular version of the approach introduced by Carley [8], [9] is an extension of the well known flash technique often used for fast A/D conversion [10]. Carley's approach, which will be reviewed in detail below, involves the use of a randomizing digital encoder to effectively decorrelate successive samples of $\epsilon(n)$. For lack of an existing name, this paper will refer to D/A converters based on Carley's approach as *stochastic flash* D/A converters.

The main contribution of this paper is a rigorous and quantitative theoretical analysis of the conversion error introduced by stochastic flash D/A converters. Although Carley demonstrated a 3-b stochastic flash D/A converter for use as part of a 15-b delta-sigma modulator D/A converter and showed experimentally that the system introduces very little harmonic distortion, no mathematical analysis of the stochastic flash D/A converter has been published previously. The analysis developed in this paper indicates that, under mild circuit performance assumptions, the only distortion introduced by $\epsilon(n)$ is that of a gain error, a dc offset, and an additive white noise component. Hence, in principle, stochastic flash D/A converters introduce no harmonic distortion whatsoever. In addition to supporting Carley's observations, the analysis provides expressions for the gain error, dc offset, and the power of the white noise in terms of the mean and power of the

Manuscript received October 25, 1993; revised January 4, 1995. The research described in this paper was performed in part for the Jet Propulsion Laboratory, California Institute of Technology, under a contract for the National Aeronautics and Space Administration. This paper was recommended by Associate Editor L. R. Carley.

I. Galton is with the Department of Electrical and Computer Engineering, University of California, Irvine, CA 92717 USA.

P. Carbone was with the Department of Electrical and Computer Engineering, University of California, Irvine, CA 92717 USA on leave from the Department of Electronics and Information, University of Padova, 35131 Padova, Italy. He is now with the Third University of Rome, Italy.

IEEE Log Number 9415864.

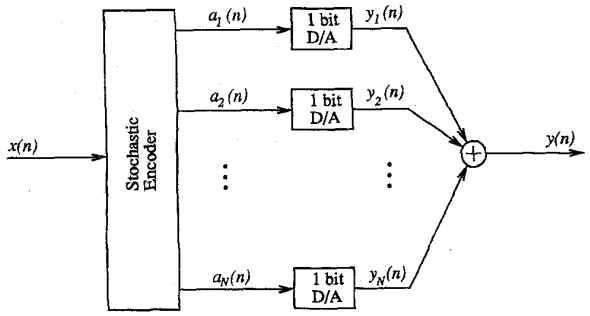

Fig. 1. The stochastic flash D/A converter.

input sequence, and the specific component matching errors. A general yield estimation technique based on the expression for the power of the white noise is presented that can be used to estimate the distribution of the white noise power across different copies of the same D/A converter circuit for any set of component matching error statistics.

The remainder of the paper is divided into four main sections. Section II reviews the stochastic flash D/A converter approach in detail. Additionally, an efficient digital encoder architecture based on a *Beneš network* [11] is proposed that is appropriate for a monolithic 8-b stochastic flash D/A converter. Theoretical results are developed in Section III and the Appendix that characterize the conversion error in terms of the time-average mean and autocorrelation of the D/A converter output. The yield estimation technique is also presented in Section III. Simulation results are presented in Section IV. Section V describes the application of an 8-b stochastic flash converter to direct digital synthesis, and presents simulation results of the overall system.

II. The Stochastic Flash D/A Converter Architecture

The stochastic flash D/A converter architecture is shown in Fig. 1. It operates on a sequence of digital numbers, $x(n)$, and produces a sequence of analog output values, $y(n)$. In a manner that will be described below, the *stochastic encoder* maps each input value to N 1-b output values, $a_1(n), a_2(n), \cdots, a_N(n)$. These are converted into analog voltages or currents by the N 1-b D/A converters and then summed to produce $y(n)$.

The input sequence, $x(n)$, is assumed to be a deterministic sequence of numbers with values $x(n) \in \{x_{\min} + k\Delta : k = 0, 1, \cdots, N\}$ where Δ is the *step-size* of the D/A converter and $N > 1$. Thus, $x_{\min} \le x(n) \le x_{\max}$ where $x_{\max} = x_{\min} + N\Delta$. For each value of $x(n)$, the stochastic encoder sets $K(n)$ of its output lines to one and the remaining $N - K(n)$ of its output lines to zero where

$$K(n) = \frac{x(n) - x_{\min}}{\Delta}. \tag{1}$$

For each value of $x(n)$, this can be done in $\binom{N}{K(n)}$ different ways. The idea behind dynamic element matching is to have the stochastic encoder choose among the $\binom{N}{K(n)}$ possibilities randomly.

Formally, the operation of the stochastic encoder can be described as follows. Let $\mathbf{a}(n) = [a_1(n) \cdots a_N(n)]^T$. Then the stochastic encoder chooses its outputs such that

$$\mathbf{a}(n) \in \{0, 1\}^N, \tag{2}$$

and

$$\mathbf{a}^T(n)\mathbf{a}(n) = K(n), \tag{3}$$

where $\mathbf{a}(n)$ is a sequence of independent random variables such that for every value of $x(n)$ each of the $\binom{N}{K(n)}$ possible values of $\mathbf{a}(n)$ are equiprobable.

The 1-b D/A converters each operate according to

$$y_r(n) = \begin{cases} w_h + e_{h_r}, & \text{if } a_r(n) = 1 \\ w_l + e_{l_r}, & \text{if } a_r(n) = 0 \end{cases}$$

where $w_h = x_{\max}/N$, $w_l = x_{\min}/N$ (hence, $w_h - w_l = \Delta$), and e_{h_r}, e_{l_r} represent D/A conversion errors.

If the 1-b D/A converters are ideal, $e_{h_r} = e_{l_r} = 0$ for all r, and

$$y(n) = \mathbf{a}^T(n)\begin{bmatrix} w_h \\ \vdots \\ w_h \end{bmatrix} + \overline{\mathbf{a}}^T(n)\begin{bmatrix} w_l \\ \vdots \\ w_l \end{bmatrix}$$

where $\overline{\mathbf{a}}^T(n)$ represents the one's-complement of $\mathbf{a}^T(n)$. This can be rewritten as

$$y(n) = \mathbf{a}^T(n)\mathbf{a}(n)w_h + [N - \mathbf{a}^T(n)\mathbf{a}(n)]w_l.$$

By factoring and using (3), this becomes

$$\begin{aligned} y(n) &= \mathbf{a}^T(n)\mathbf{a}(n)[w_h - w_l] + Nw_l \\ &= \frac{x(n) - x_{\min}}{\Delta}\Delta + x_{\min} = x(n). \end{aligned} \tag{4}$$

Therefore, in the case of ideal 1-b D/A converters, the analog output equals the digital input signal $x(n)$.

Now suppose the 1-b D/A converters are not ideal. Then, $\mathbf{e}_h \neq \mathbf{0}$ or $\mathbf{e}_l \neq \mathbf{0}$ where

$$\mathbf{e}_h = \begin{bmatrix} e_{h_1} \\ \vdots \\ e_{h_N} \end{bmatrix}, \quad \text{and} \quad \mathbf{e}_l = \begin{bmatrix} e_{l_1} \\ \vdots \\ e_{l_N} \end{bmatrix}.$$

In this case, the output $y(n)$ can be written as

$$\begin{aligned} y(n) = \mathbf{a}^T(n)&\left(\begin{bmatrix} w_h \\ \vdots \\ w_h \end{bmatrix} + \begin{bmatrix} e_{h_1} \\ \vdots \\ e_{h_N} \end{bmatrix}\right) \\ &+ \overline{\mathbf{a}}^T(n)\left(\begin{bmatrix} w_l \\ \vdots \\ w_l \end{bmatrix} + \begin{bmatrix} e_{l_1} \\ \vdots \\ e_{l_N} \end{bmatrix}\right). \end{aligned}$$

Applying (4), this becomes

$$y(n) = x(n) + \mathbf{a}^T(n)\mathbf{e}_h + \overline{\mathbf{a}}^T(n)\mathbf{e}_l. \tag{5}$$

Thus the conversion error introduced by the nonideal 1-b D/A converters is

$$\epsilon(n) = \mathbf{a}^T(n)\mathbf{e}_h + \overline{\mathbf{a}}^T(n)\mathbf{e}_l. \tag{6}$$

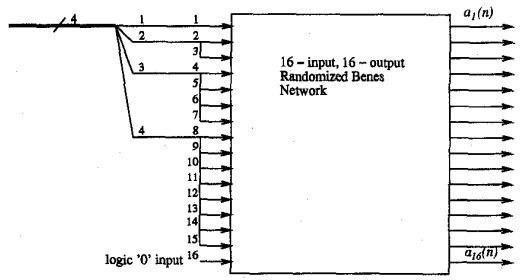

Fig. 2. A 4-b stochastic encoder based on a randomized Beneš network.

Equation (5) holds regardless of whether $\mathbf{e}_h$ and $\mathbf{e}_l$ are fixed vectors, deterministic functions of $x(n)$, or random variables. However, if the 1-b D/A converters and the N-input summing node are designed carefully, then to a very good approximation $\mathbf{e}_h$ and $\mathbf{e}_l$ can be considered to be fixed vectors. For example, such would be the case if each 1-b D/A conversion were performed by selecting or shunting the output of regulated gate cascode current source [12]. The outputs of the D/A converters would simply be connected and passed through a resistive load to implement the N-input summing node. Therefore, throughout the remainder of this paper $\mathbf{e}_h$ and $\mathbf{e}_l$ will be assumed to be fixed vectors.

At this point, it is worth digressing briefly to discuss an efficient method of implementing the stochastic encoder. In order to conserve digital hardware, Carley used a *butterfly-type randomizer* that only partially implemented the operation of a stochastic encoder in his 3-b stochastic flash D/A converter. However, hardware-efficient approaches exist in the field of parallel computing with which to implement the full stochastic encoder. In particular, a *Beneš network* [11] is an efficient N-input, N-output structure capable of connecting its inputs to its outputs in any of the $N!$ possible permutations. Such a network consists of $N \log_2 N - N/2$ binary switches. Each binary switch connects two 1-b inputs to two 1-b outputs either with or without the two inputs swapped. By controlling each of these switches with 1-b of a pseudorandom number generator, the Beneš network randomly permutes its N-inputs to its N-outputs. A 4-b stochastic encoder based on such a randomized Beneš network is shown in Fig. 2.

It appears reasonable that at least an 8-b monolithic version of the stochastic encoder could be practically implemented using high-density CMOS circuit technology. For an 8-b system, the required number of binary switches is 1920. To implement the randomization, each of these switches could be controlled by one bit of a linear feedback shift register (LFSR). For example, 62 31-b LFSR's would then be required to generate the necessary control signals. Although pipelining circuitry would be necessary and routing of the control circuitry would be nontrivial, the complexity of the overall circuit is not excessive by current digital circuit standards.

III. Second-Order Statistics of the Stochastic Flash D/A Converter

Before analyzing the conversion error introduced by the stochastic flash D/A converter, it is illustrative to first consider a similar system except with an encoder that does not perform randomization. Suppose for a moment that instead of the stochastic encoder a deterministic encoder were used that chooses its outputs according to (2) and (3) but for each value of $x(n)$ only chooses one of the $\binom{N}{K(n)}$ possible output values. It is easy to verify that (5) would still be valid. Such a system can be thought of as a (nonstochastic) flash D/A converter because it is analogous to a flash A/D converter. It is easy to show that this system introduces harmonic distortion. For example, suppose that $x(n)$ is a sinusoid with period M quantized to b-bits. Because $\mathbf{e}_h$ and $\mathbf{e}_l$ are fixed and $\mathbf{a}^T(n)$ is a memoryless, deterministic function of $x(n)$, $\epsilon(n)$ must be a memoryless, deterministic function of $x(n)$ and therefore must be periodic with maximum period M. Therefore, it consists solely of harmonics of the input sinusoid which, together, constitute harmonic distortion.

In contrast, with a stochastic encoder, $\epsilon(n)$ is a discrete-valued random variable. As will be shown, the randomization performed by the stochastic encoder decorrelates $\epsilon(n)$ such that harmonic distortion is theoretically eliminated although a gain error, an offset error, and white noise are introduced. This statement will be quantified in the remainder of this section, and an example will be presented that uses the result to estimate the statistics with which the noise power varies across different copies of the same circuit.

In accordance with the usual definitions, the time-average mean and autocorrelation of the input sequence $x(n)$ are defined as

$$\overline{M}_x = \lim_{P\to\infty} \frac{1}{P} \sum_{n=1}^{P} x(n)$$

and

$$\overline{R}_{xx}(k) = \lim_{P\to\infty} \frac{1}{P} \sum_{n=1}^{P} x(n)x(n+k)$$

respectively. The time-average mean and autocorrelation of the output sequence, $y(n)$, are defined analogously with x replaced by y in the above definitions.

As proven in the Appendix, the time-average autocorrelation of the output of the stochastic flash D/A converter is

$$\overline{R}_{yy}(k) = (1+\alpha)\overline{R}_{xx}(k) + \overline{\beta} + \overline{\sigma}^2\delta(k) \tag{7}$$

with probability 1, where α, $\overline{\beta}$, and $\overline{\sigma}^2$ are constants that depend on the 1–b D/A converter errors, $\mathbf{e}_l$ and $\mathbf{e}_h$. Therefore, no matter how much error is introduced by the 1-b D/A converters or how these errors are distributed, the stochastic flash D/A converter behaves exactly like an ideal D/A converter aside from a white noise term and fixed gain and dc offsets. The practical implication is that there is no theoretical lower bound on the harmonic distortion suffered by physical stochastic flash D/A converter circuits.

This property of stochastic flash D/A converters has been reported by Carley on the basis of heuristic reasoning and experimental results. The analysis presented in the Appendix not only provides rigorous theoretical support for Carley's

observation, but it also quantifies the result. In particular, the analysis yields the following expressions for α, $\overline{\beta}$, and $\overline{\sigma}^2$:

$$\alpha = e_0(2+e_0), \qquad \overline{\beta} = 2e_1(1+e_0)\overline{M}_x + e_1^2$$

and

$$\overline{\sigma}^2 = e_2\left[-x_{\min}x_{\max} - \overline{R}_{xx}(0) + (x_{\min}+x_{\max})\overline{M}_x\right] \quad (8)$$

where

$$e_0 = \frac{1}{N\Delta}\sum_{k=1}^{N}(e_{h_k} - e_{l_k}),$$

$$e_1 = \frac{1}{N\Delta}\sum_{k=1}^{N}(x_{\min}e_{h_k} - x_{\max}e_{l_k}), \quad (9)$$

and

$$e_2 = \frac{1}{(N-1)\Delta^2} \cdot \left[\frac{1}{N}\sum_{k=1}^{N}(e_{h_k} - e_{l_k})^2 - \left(\frac{1}{N}\sum_{k=1}^{N}(e_{h_k} - e_{l_k})\right)^2\right]. \quad (10)$$

As will be shown in the remainder of this section, these expressions provide useful information regarding the theoretical performance of the stochastic flash D/A converter. In particular, various properties of the stochastic flash D/A converter will be deduced from the expression for $\overline{\sigma}^2$. As will be evident, similar arguments can be applied to α and $\overline{\beta}$. To simplify the notation, it will be assumed for the remainder of the section that $x_{\min} = -x_{\max}$ so (8) reduces to

$$\overline{\sigma}^2 = e_2\left[x_{\max}^2 - \overline{R}_{xx}(0)\right]. \quad (11)$$

Let

$$\widetilde{e}_k = e_{h_k} - e_{l_k} - \overline{e},$$

where

$$\overline{e} = \frac{1}{N}\sum_{k=1}^{N}(e_{h_k} - e_{l_k}).$$

Using the terminology of [13], $\widetilde{e}_k$ and $\overline{e}$ can be viewed as the *local variation* and *global variation*, respectively, of $e_{h_k} - e_{l_k}$ across a particular copy of the stochastic flash D/A converter circuit under consideration. With these definitions, (10) and (11) reduce to

$$\overline{\sigma}^2 = \frac{1}{N(N-1)\Delta^2}\left[x_{\max}^2 - \overline{R}_{xx}(0)\right]\sum_{k=1}^{N}\widetilde{e}_k^2, \quad (12)$$

from which it is clear that $\overline{\sigma}^2$ only depends on the local variation of $e_{h_k} - e_{l_k}$.

Before considering how $\overline{\sigma}^2$varies as a function of its input sequence or how it varies across different copies of the same circuit, it is interesting to calculate a practical upper bound on $\overline{\sigma}^2$. Since $x_{\max} = N\Delta/2$ and $\overline{R}_{xx}(0) \geq 0$, it follows from (12) that

$$\overline{\sigma}^2 \leq \frac{N^2}{4(N-1)}\max_k\{\widetilde{e}_k^2\} \approx \frac{N}{4}\max_k\{\widetilde{e}_k^2\}.$$

For example, suppose the 1-b D/A converters are based on current sources as described in the previous section. Then $\widetilde{e}_k$ depends on how well the current source transistors can be matched in the given circuit technology. For carefully designed CMOS current sources, $\widetilde{e}_k$ can be modeled as a random function of k that has zero mean and is normally distributed with standard deviation $\sigma_e \approx 0.005 \cdot \Delta$ [13], [14]. Because $\widetilde{e}_k$ has zero mean and is normally distributed, the probability that $|\widetilde{e}_k| < 4\sigma_e$ for all k on all copies of the circuit is extremely high. Hence, with a very high likelihood, every copy of the circuit will introduce white noise with power $\overline{\sigma}^2 < 0.0001 \cdot N\Delta^2$.

For certain applications, this practical upper bound on $\overline{\sigma}^2$ is negligibly low. For example, this is often the situation in applications such as direct digital synthesis where $x(n)$ is a uniformly quantized representation of a continuous-amplitude signal. In such cases, $x(n)$ can be viewed as consisting of the continuous-amplitude signal plus quantization noise, and the quantization noise often is well modeled as white noise with power $\Delta^2/12$. Since the white noise introduced by the stochastic flash D/A converter is independent from the quantization noise, it will only have a significant effect on the mean squared error of the output if $\overline{\sigma}^2$ is at least of the same order of magnitude as $\Delta^2/12$. However as shown above, even for relatively large values of N (e.g., $N = 255$), under reasonable circuit assumptions the upper bound on $\overline{\sigma}^2$ is small compared to $\Delta^2/12$. Therefore, in such applications, the theoretical performance of the stochastic flash D/A converter is essentially equivalent to that of an ideal D/A converter. Simulation results performed in the next two sections support this conclusion.

Nevertheless, in some applications the noise introduced by the stochastic flash D/A converter can be problematic. For example, such is the case with delta-sigma modulator A/D converters. Delta-sigma modulators typically contain a coarse A/D converter in the feed-forward path, and a coarse D/A converter in the feedback path. Although the quantization noise introduced by the A/D converter as seen at the output is highly attenuated in the band-of-interest, the noise introduced by the D/A converter is passed to the output without any attenuation [3]. Consequently, the power of the noise introduced by the D/A converter often must be several orders of magnitude lower than that of the quantization noise lest it limit the overall A/D conversion accuracy. In such cases, it is desirable to know how $\overline{\sigma}^2$ varies as a function of the input sequence, and how it varies across different copies of the same stochastic flash D/A converter circuit.

It follows from (11) that the power of the white noise is greatest when $x(n) = 0$ for all n and is zero when $x(n) = x_{\max}$ for all n. A sinusoidal input: $x(n) = A\sin(\omega n)$ has power $\overline{R}_{xx}(0) = A^2/2$. Therefore, the white noise floor of a stochastic flash D/A converter will be 3 dB below its maximum value for a full-scale sinusoidal input (i.e., $A = x_{\max}$), and reducing the power of the sinusoidal input increases the noise floor proportionally.

To the extent that the statistics of $e_{h_k}$and e_{l_k} are known for the given circuit technology, (8) and (10) can be used as the basis of a yield estimation technique to predict how

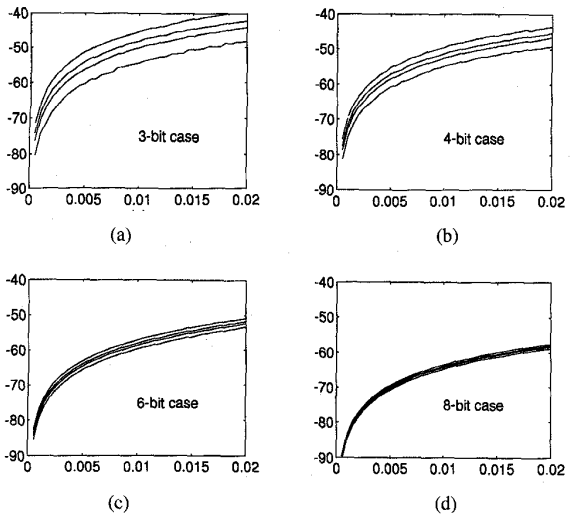

Fig. 3. Yield estimation data for (a) 3-b, (b) 4-b, (c) 6-b, and (d) 8-b stochastic flash D/A converters with values of σ_e/Δ ranging from 0.0005 to 0.02. In each case, from top to bottom the four curves show the largest of the smallest 95%, 65%, 35%, and 5% of the calculated values of $\overline{\sigma}^2$ (in dB relative to $x_{\max}^2$), respectively.

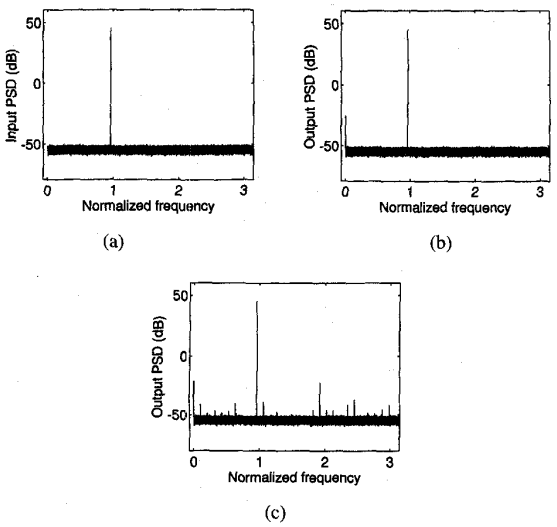

Fig. 4. Estimated power spectral density (PSD) in dB relative to the power of a full scale dc input of (a) the input sequence, (b) the stochastic flash D/A converter output sequence, and (c) the deterministic flash D/A converter output sequence.

$\overline{\sigma}^2$ will vary across different copies of the same circuit. The most general approach involves computer analysis. The idea is to repeatedly generate samples of e_{h_k} and e_{l_k} with the appropriate statistics and use them in (10) and (8) to calculate samples of $\overline{\sigma}^2$. Provided a large number of samples of $\overline{\sigma}^2$ are so calculated, the distribution of the calculated values will closely approximate the underlying statistical distribution. The calculated distribution thus provides an estimate of the number of copies of the circuit that will achieve a $\overline{\sigma}^2$ below any given value of interest.

For example, Fig. 3(a)–(d) shows yield estimation data corresponding to 3-b, 4-b, 6-b, and 8-b stochastic flash D/A converters, respectively. From top to bottom, the four curves shown in each figure represent the largest of the smallest 95%, 65%, 35% and 5% of the calculated values of $\overline{\sigma}^2$ (plotted in units of dB as $10 \cdot \log_{10}(\overline{\sigma}^2/x_{max}^2)$), respectively. Each figure shows data from from 40 yield estimations with values of σ_e/Δ ranging from 0.0005 to 0.02. For example, as can be seen from Fig. 3(a), the yield estimation for a 3-b D/A converter with $\sigma_e = 0.01 \cdot \Delta$ predicts that 95% of the devices will have $\overline{\sigma}^2$ below -45 dB, and that 5% of the devices will have $\overline{\sigma}^2$ below -55 dB. Each yield estimation was based on 1000 calculated values of $\overline{\sigma}^2$. The values of $e_{h_k} - e_{l_k}$ were calculated as samples of independent normally distributed random variables each with a standard deviation σ_e, although any other statistics for $e_{h_k} - e_{l_k}$ could have been used without otherwise affecting the calculations. A different random number seed was used for each yield estimation, so the relative smoothness of the curves indicates that there was little variance in the yield estimation results. It follows that the data are good approximations of the underlying statistical distribution of $\overline{\sigma}^2$.

As is evident from Fig. 3(a)–(d), the predicted spread of the white noise floor introduced by the stochastic flash D/A converter becomes small as the number of bits is increased. This is reasonable because as the number of bits is increased, the 1-b D/A converter errors, which are measured relative to Δ, become smaller relative to $x_{\max}$. If the yield estimations were performed with errors measured relative to $x_{\max}$, this phenomenon would not be observed. However, the local variation of CMOS current sources is not strongly dependent on the nominal magnitude of the current [14], so for the case where the 1-b D/A converters are based on CMOS current sources it is more realistic to present the data as shown.

IV. Simulation Results

This section presents simulation data that support the theoretical results derived above. As will be shown, for an 8-b stochastic flash D/A converter with a dithered sinusoidal input, no harmonic distortion is evident in the simulated output sequence.

The simulations used an input sequence that was generated by adding a dither sequence to a sinusoid and then quantizing the result to 8-b. The sinusoid had an amplitude of 127Δ. The dither sequence was an independent identically distributed (iid) sequence with a triangular probability density function supported on $(-\Delta, \Delta)$. As shown in [15], the power spectral density of such a dithered sinusoid is equal to that of the original sinusoid plus an independent white noise sequence with power $\frac{\Delta}{4}^2$. Thus, by design the input sequence behaved as an ideal sinusoid plus white noise.

Fig. 4(a) shows the estimated power spectral density of the dithered sinusoid input sequence in dB/radian relative to 127Δ. The data were obtained by averaging 15 periodograms [16] each corresponding to $2^{19} = 524288$ points of the dithered sinusoid input sequence. The frequency of the sinusoid was set to $160000\pi/524288$ radians so as to avoid spectral leakage. As expected, no harmonics of the sinusoid are evident in the figure and the noise caused by the operations of adding dither and quantizing appears white with a power spectral density of

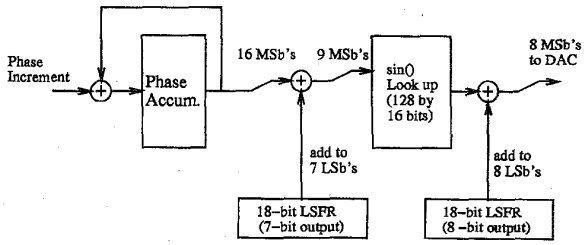

Fig. 5. Direct digital synthesizer block diagram as presented in [1].

approximately -54 dB/radian (which corresponds to a total power of $\frac{\Delta^2}{4}$).

Fig. 4(b) shows the estimated power spectral density of the output of an 8-b stochastic flash D/A converter operating on the dithered sinusoidal input sequence. The spectral estimation was performed in an identical manner to that of Fig. 4(a). The 1-b D/A converter errors, e_{h_r} and e_{l_r}, $1 \leq r \leq 256$, were chosen as normally distributed independent random variables with a standard deviation of $0.005 \cdot \Delta$. As predicted by the theory outlined in Section III, no harmonic distortion is evident in the figure. Indeed, there is little visible difference between the estimated power spectral densities of the input and output sequences. For comparison, Fig. 4(c) shows the estimated power spectral density of the output of the same flash D/A converter except with the stochastic encoder replaced by a deterministic encoder as described in Section III. As predicted in Section III, the figure clearly indicates the presence of numerous harmonics of the sinusoid frequency.

V. Application to High Spectral Purity Frequency Synthesis

In principle, a high spectral purity direct digital synthesizer could be implemented by proceeding as in the simulations described above. A sequence consisting of a sinusoid and an iid dither sequence with a triangular probability density function could be accurately generated by a floating point processor and then quantized to b-bits. As discussed above and supported by the simulation results shown in Fig. 4(a), the resulting sequence, which will be referred to below as an *ideally dithered and quantized sinusoid*, would have a power spectral density equal to that of an ideal sinusoid plus white noise. The problem with this approach is that to achieve a wide range of sinusoid frequencies, the hardware required to implement the floating point processor would be excessive.

To circumvent this problem, Flanagan and Zimmerman have proposed an alternate approach to generating a dithered and quantized sinusoid that requires significantly less digital hardware than the approach outlined above [1]. An 8-b example of their approach is shown in Fig. 5. The system consists of a dithered phase accumulator and a dithered sinusoid look-up table. They showed that by dithering and quantizing both phase and amplitude, the power of each harmonic in the final 8-b sequence is at least 90 dB below the power of the sinusoid (i.e., the sequence has 90 dBc spur suppression), and the noise power is only a few dB above that of the corresponding ideally dithered and quantized sinusoid. This result is supported by

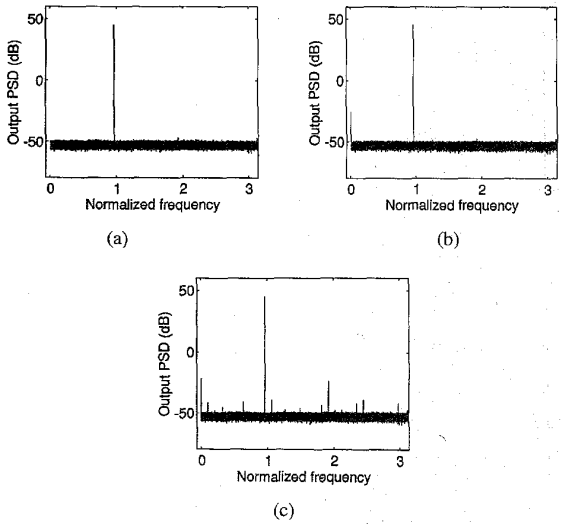

Fig. 6. Estimated power spectral density (PSD) in dB relative to the power of a full scale dc input of (a) the sequence generated by the system of Fig. 5, (b) the corresponding stochastic flash D/A converter output sequence, and (c) the corresponding deterministic flash D/A converter output sequence.

the simulation data presented in Fig. 6(a) which shows the estimated power spectral density of the output of the system shown in Fig. 5.

Typically, direct digital frequency synthesis is used to generate analog signals. Therefore, the system of Fig. 5 would usually be followed by an 8-b D/A converter. Because the digital portion of the system has 90 dBc spur suppression, it would be desirable to use a D/A converter that introduces no more than -90 dB harmonic distortion relative to its maximum signal power. For applications with sample rates above 20 MHz, it does not appear that D/A converters with such a low level of harmonic distortion are commercially available at the present time. Nevertheless, the results of this paper indicate that provided the 1-b D/A converters can be implemented such that $\mathbf{e}_h$ and $\mathbf{e}_l$ are essentially static and the N-input summing node can be implemented with sufficient linearity, it may be possible to design a stochastic flash D/A converter with the required specifications.

Assuming for the moment that such a design is feasible, the theoretical results of Section III indicate that a properly designed stochastic flash D/A converter will operate almost as well as an ideal D/A converter except that the analog output signal will have a slightly higher than ideal noise floor and will contain a dc offset. Fig. 6(b) shows the estimated power spectral density of the output of a stochastic flash D/A converter operating on the 8-b sequence generated by the system of Fig. 5. As in the previous section, the 1-b D/A converter errors, e_{h_r} and e_{l_r}, $1 \leq r \leq 256$, were chosen as normally distributed independent random variables with a standard deviation of $0.005 \cdot \Delta$. As expected, aside from a dc offset and a slight increase in the noise floor, the data in Fig. 6(b) are similar to those in Fig. 6(a). In particular, no harmonic distortion is evident. For reference, Fig. 6(c) shows the estimated power spectral density of the stochastic flash D/A converter with the stochastic encoder

replaced by a deterministic encoder. As is evident from a comparison of Fig. 5(b) and (c), the randomization performed by the stochastic encoder is in this case sufficient to suppress harmonic distortion.

VI. Conclusion

A rigorous and quantitative analysis of the conversion error introduced by an important type of D/A converter with dynamic element matching has been presented. The analysis supports previously published experimental results that indicate dynamic element matching introduces white noise instead of harmonic distortion. Moreover, the analysis results in an expression for the power of the white noise in terms of the mean and power of the input sequence and the specific component matching errors. A general yield estimation technique is presented that uses the expression to estimate the statistics of the white noise power across different copies of a given D/A converter circuit for arbitrary component matching error statistics. Simulation results are presented that support the theory, and an application to the design of a practical, high spectral purity, direct digital synthesizer is discussed. Additionally, an efficient digital encoder architecture based on a Beneš network is proposed that appears appropriate for a monolithic 8-b version of the D/A converter.

Appendix

The detailed analysis of the conversion error of the stochastic flash D/A converter is presented in this appendix. First, a lemma is presented that provides intermediate results required to support the subsequent analysis. Then, a theorem is presented that provides the results used in the body of the paper.

Lemma: Suppose that $K(n)$ is an integer-valued function such that $0 \le K(n) \le N$, and that $\mathbf{a}(n) = [a_1 \cdots a_N]^T$ is a sequence of independent random variables satisfying

$$\mathbf{a}(n) \in \{0,1\}^N,$$

and

$$\mathbf{a}^T(n)\mathbf{a}(n) = K(n),$$

where for every value of $K(n)$ each of the $\binom{N}{K(n)}$ possible values of $\mathbf{a}(n)$ are equiprobable. Then, for $N > 1$,

$$\mathrm{E}[\mathbf{a}(n)] = \frac{K(n)}{N}\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}, \tag{13}$$

$$\mathrm{E}[\mathbf{a}(n)\mathbf{a}^T(n)] = \frac{K(n)[K(n)-1]}{N(N-1)}\mathbf{A_N} + \frac{K(n)[N-K(n)]}{N(N-1)}\mathbf{I_N}, \tag{14}$$

and,

$$\mathrm{E}[\mathbf{a}(n)\bar{\mathbf{a}}^T(n)] = \frac{K(n)[N-K(n)]}{N(N-1)}(\mathbf{A_N} - \mathbf{I_N}). \tag{15}$$

where $\mathbf{I_N}$ represents the $N \times N$ identity matrix and $\mathbf{A_N}$ the $N \times N$ all-ones matrix.

Proof: Let N_K be the number of different vectors $\mathbf{a}(n)$ such that the sum of their elements is equal to $K(n)$. By definition $K(n) \in \{0, \cdots, N\}$, so $N_K = \binom{N}{K(n)}$. From the definition of the expectation operator,

$$\mathrm{E}[\mathbf{a}(n)] = \sum_{i=1}^{N_K} \mathbf{a}^{(i)}(n) P\Big(\mathbf{a}(n) = \mathbf{a}^{(i)}(n)\Big), \tag{16}$$

where $\mathbf{a}^{(i)}(n)$ represents the i-th vector $\mathbf{a}(n)$ that obeys the constraint. By hypothesis, the different values of $\mathbf{a}(n)$ are equiprobable, so (16) becomes

$$\mathrm{E}[\mathbf{a}(n)] = \frac{1}{N_K}\sum_{i=1}^{N_K} \mathbf{a}^{(i)}(n). \tag{17}$$

Suppose $K(n) \neq 0$. Consider the r-th element, $a_r^{(i)}(n)$, of $\mathbf{a}^{(i)}(n)$. The number of distinct values of i for which $a_r^{(i)}(n)$ is equal to 1 corresponds to the number of times the sum of the $N-1$ elements with index different from r is equal to $K(n)-1$, i.e., $\binom{N-1}{K(n)-1}$. Since this reasoning applies for each $0 \le r \le N$, expression (17) becomes

$$\mathrm{E}[\mathbf{a}(n)] = \frac{\binom{N-1}{K(n)-1}}{\binom{N}{K(n)}}\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix},$$

which can be further simplified as

$$\mathrm{E}[\mathbf{a}(n)] = \frac{K(n)}{N}\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}.$$

If $K(n) = 0$ then $N_K = 1$, $\mathbf{a}^T(n) = [0 \cdots 0]$ and (13) follows trivially.

By definition,

$$\mathrm{E}[\mathbf{a}(n)\mathbf{a}^T(n)] = \sum_{i=1}^{N_K} \mathbf{a}^{(i)}(n)\mathbf{a}^{(i)T}(n) P\Big(\mathbf{a}(n)\mathbf{a}^T(n) = \mathbf{a}^{(i)}(n)\mathbf{a}^{(i)T}(n)\Big).$$

By virtue of the equiprobability of the vectors $\mathbf{a}^{(i)}(n)$, this becomes

$$\mathrm{E}[\mathbf{a}(n)\mathbf{a}^T(n)] = \frac{1}{N_K}\sum_{i=1}^{N_K} \mathbf{a}^{(i)}(n)\mathbf{a}^{(i)T}(n).$$

The principal diagonal of the matrix $\mathbf{a}(n)\mathbf{a}^T(n)$ is equal to $\mathbf{a}(n)$ and, therefore, by the reasoning employed in the proof of (13), it follows that

$$\mathrm{diag}\Big\{\frac{1}{N_K}\sum_{i=1}^{N_K} \mathbf{a}^{(i)}(n)\mathbf{a}^{(i)T}(n)\Big\} = \frac{K(n)}{N}\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}. \tag{18}$$

Now consider an off-diagonal element $a_{j,k}^{(i)}$ of matrix $\mathbf{a}^{(i)}(n)\mathbf{a}^{(i)T}(n)$. Assume $K(n) > 1$. Consider the number of times $a_{j,k}^{(i)} = 1$ as $i = 1, \cdots, N_K$. If the element $a_{k,k}^{(i)}$ on the principal diagonal is equal to zero, then $a_{j,k}^{(i)} = 0$. Therefore, the number of times $a_{j,k}^{(i)} = 1$ and $a_{k,k}^{(i)} = 1$, is

equal to the number of times the sum of the $N-2$ elements of column k with row index different from j and from k, is equal to $K(n)-2$, i.e., $\binom{N-2}{K(n)-2}$. Therefore, the expectation of each off-diagonal element is equal to

$$\begin{aligned}&\frac{1}{N_K}\sum_{i=1}^{N_K}\mathbf{a}_i(n)\mathbf{a}_i^T(n)(\mathbf{A_N}-\mathbf{I_N})\\&=\binom{N-2}{K(n)-2}\binom{N}{K(n)}^{-1}(\mathbf{A_N}-\mathbf{I_N})\\&=\frac{K(n)[K(n)-1]}{N(N-1)}(\mathbf{A_N}-\mathbf{I_N}).\end{aligned}\tag{19}$$

By combining (18) and (19), (14) follows for $K(n)>1$. If $K(n)=0$ then $\mathrm{E}[\mathbf{a}(n)\mathbf{a}^T(n)]$ equals the all-zeros matrix and (14) trivially follows. If $K(n)=1$ then the off-diagonal elements of matrix $\mathbf{a}(n)\mathbf{a}^T(n)$ are equal to zero. As a consequence $\mathrm{E}[\mathbf{a}^T(n)]=\frac{K(n)}{N}\mathbf{I_N}$ and (14) again follows.

By definition,

$$\begin{aligned}&\mathrm{E}[\mathbf{a}(n)\bar{\mathbf{a}}^T(n)]\\&\quad=\sum_{i=1}^{N_K}\mathbf{a}^{(i)}(n)\bar{\mathbf{a}}^{(i)T}(n)P\Big(\mathbf{a}(n)\bar{\mathbf{a}}^T(n)=\mathbf{a}^{(i)}(n)\bar{\mathbf{a}}^{(i)T}(n)\Big).\end{aligned}$$

Using the equiprobability hypothesis, this becomes

$$\mathrm{E}[\mathbf{a}(n)\bar{\mathbf{a}}^T(n)]=\frac{1}{N_K}\sum_{i=1}^{N_K}\mathbf{a}^{(i)}(n)\bar{\mathbf{a}}^{(i)T}(n).\tag{20}$$

By construction, the principal diagonal elements of matrix $\mathbf{a}(n)\bar{\mathbf{a}}^T(n)$ equal zero. Suppose $0<K(n)<N$. The number of times an off-diagonal element $a_{j,k}^{(i)}$ of matrix $\mathbf{a}(n)\bar{\mathbf{a}}^T(n)$ is equal to 1 corresponds to the number of times the sum of the $N-2$ elements of column k with row index different from j and from k, is equal to $K(n)-1$, i.e., $\binom{N-2}{K(n)-1}$. Thus expression (20) becomes

$$\begin{aligned}\mathrm{E}[\mathbf{a}(n)\bar{\mathbf{a}}^T(n)]&=\frac{\binom{N-2}{K(n)-1}}{\binom{N}{K(n)}}(\mathbf{A_N}-\mathbf{I_N})\\&=\frac{K(n)[N-K(n)]}{N(N-1)}(\mathbf{A_N}-\mathbf{I_N}).\end{aligned}$$

If $K(n)=0$ or $K(n)=N$ then, by direct inspection, the matrix $\mathbf{a}(n)\bar{\mathbf{a}}^T(n)$ is identically equal to the all-zeros matrix. As a consequence (15) still holds. ■

Theorem: If $\overline{M}_x$ and $\overline{R}_{xx}(k)$ exist, then[1]

$$\overline{M}_y=\overline{M}_x(1+e_0)+e_1,\tag{21}$$

and

$$\overline{R}_{yy}(k)=(1+\alpha)\overline{R}_{xx}(k)+\overline{\beta}+\overline{\sigma}^2\delta(k),\tag{22}$$

[1]The function $\delta(k)$ is the Kronecker delta defined as

$$\delta(k)=\begin{cases}1, & \text{if } k=0;\\ 0, & \text{otherwise;}\end{cases}$$

and the symbol $\mathbf{1}^T$ is defined as the all-ones vector: $[1\cdots1]$.

with probability 1, where

$$\begin{aligned}e_0&=\frac{1}{N\Delta}\Big(\mathbf{1}^T\mathbf{e}_h-\mathbf{1}^T\mathbf{e}_l\Big),\\e_1&=\frac{1}{N\Delta}\Big(x_{\min}\mathbf{1}^T\mathbf{e}_h-x_{\max}\mathbf{1}^T\mathbf{e}_l\Big),\end{aligned}$$

$$\alpha=e_0(2+e_0),\qquad\overline{\beta}=2e_1(1+e_0)\overline{M}_x+e_1^2,$$

and

$$\begin{aligned}\overline{\sigma}^2=\frac{1}{(N-1)}&\left[e_0^2-\frac{(\mathbf{e}_h-\mathbf{e}_l)^T(\mathbf{e}_h-\mathbf{e}_l)}{N\Delta^2}\right]\\&\cdot\left[x_{\min}x_{\max}+\overline{R}_{xx}(0)-(x_{\min}+x_{\max})\overline{M}_x\right].\end{aligned}$$

Proof: Because $\mathbf{e}_h$ and $\mathbf{e}_l$ are fixed, it follows from (5) that

$$\mathrm{E}[y(n)]=\mathrm{E}[x(n)]+\mathrm{E}[\mathbf{a}^T(n)]\mathbf{e}_h+\mathrm{E}[\bar{\mathbf{a}}^T(n)]\mathbf{e}_l,$$

where $\mathrm{E}[\cdot]$ represents the statistical expectation operator. By applying the lemma presented above and using the fact that $x(n)$ is deterministic, this becomes

$$\mathrm{E}[y(n)]=x(n)+\frac{K(n)}{N}\mathbf{1}^T\mathbf{e}_h+\frac{N-K(n)}{N}\mathbf{1}^T\mathbf{e}_l.\tag{23}$$

Expanding $K(n)$ using (1) and collecting terms gives

$$\mathrm{E}[y(n)]=x(n)(1+e_0)+e_1.\tag{24}$$

Because $\overline{M}_x$ exists, it follows that

$$\lim_{P\to\infty}\frac{1}{P}\sum_{n=1}^{P}\mathrm{E}[y(n)]=\overline{M}_x(1+e_0)+e_1.$$

To deduce that (21) holds with probability 1, it is sufficient to show that the sequence $y(n)$ obeys the strong law of large numbers. Note that there must exist some number C such that $\mathrm{Var}[y(n)]<C$ for of all n because $x(n)$ is bounded and $\mathbf{a}(n)\in\{0,1\}^N$. Therefore, $y(n)$ satisfies the Kolmogorov Criterion [17] which implies that it satisfies the strong law of large numbers.

Now consider $R_{yy}(n,k)$, the statistical autocorrelation of $y(n)$ defined as $R_{yy}(n,k)=\mathrm{E}[y(n)y(n+k)]$. Because $x(n)$ is deterministic and $\mathbf{a}(n)$ is a sequence of independent random variables, it follows from (5) that $y(n)$ is also a sequence of independent random variables. Therefore,

$$R_{yy}(n,k)=\mathrm{E}[y(n)]\,\mathrm{E}[y(n+k)],$$

when $k\neq0$. It follows that, in general,

$$R_{yy}(n,k)=\mathrm{E}[y(n)]\,\mathrm{E}[y(n+k)]+\sigma^2(n)\delta(k),$$

where

$$\sigma^2(n)=\mathrm{E}[y^2(n)]-\Big[\mathrm{E}[y(n)]\Big]^2.\tag{25}$$

Applying (24) and the definition of α gives

$$\begin{aligned}R_{yy}(n,k)=&\,x(n)x(n+k)(1+\alpha)\\&+[x(n)+x(n+k)]e_1(1+e_0)\\&+e_1^2+\sigma^2(n)\delta(k).\end{aligned}\tag{26}$$

To complete the calculation of $R_{yy}(n,k)$, it remains to find an equation for $\sigma^2(n)$. From (5) , it follows that

$$\mathrm{E}\big[y^2(n)\big] = x^2(n) + 2x(n)\,\mathrm{E}\big[\mathbf{a}^T(n)\mathbf{e}_h + \overline{\mathbf{a}}^T(n)\mathbf{e}_l\big] + \mathrm{E}\big[\mathbf{a}^T(n)\mathbf{e}_h + \overline{\mathbf{a}}^T(n)\mathbf{e}_l\big]^2. \tag{27}$$

It is convenient to evaluate the second two terms in (27) separately. Applying the first equation of the lemma presented in the appendix, the second term becomes

$$\mathrm{E}\big[\mathbf{a}^T(n)\mathbf{e}_h + \overline{\mathbf{a}}^T(n)\mathbf{e}_l\big] = \frac{K(n)}{N}\mathbf{1}^T\mathbf{e}_h + \frac{N-K(n)}{N}\mathbf{1}^T\mathbf{e}_l. \tag{28}$$

Expanding the third term in (27) gives

$$\mathrm{E}\big[\mathbf{a}^T(n)\mathbf{e}_h + \overline{\mathbf{a}}^T(n)\mathbf{e}_l\big]^2 = \mathbf{e}_h^T\,\mathrm{E}\big[\mathbf{a}(n)\mathbf{a}^T(n)\big]\mathbf{e}_h + \mathbf{e}_l^T\,\mathrm{E}\big[\overline{\mathbf{a}}(n)\overline{\mathbf{a}}^T(n)\big]\mathbf{e}_l + 2\mathbf{e}_h^T\,\mathrm{E}\big[\mathbf{a}(n)\overline{\mathbf{a}}^T(n)\big]\mathbf{e}_l.$$

With $\mathbf{A}_N$ defined as the $N\times N$ all-ones matrix and $\mathbf{I}_N$ defined as the $N \times N$ identity matrix, note that for any N-length vectors $\mathbf{a}$ and $\mathbf{b}$, $\mathbf{a}^T\mathbf{A}_N\mathbf{b} = (\mathbf{1}^T\mathbf{a})(\mathbf{1}^T\mathbf{b})$, and $\mathbf{a}^T\mathbf{I}_N\mathbf{b} = \mathbf{a}^T\mathbf{b}$. Using these relationships and the second and third equations of Lemma presented in the appendix gives

$$\begin{aligned}\mathrm{E}\big[\mathbf{a}^T(n)\mathbf{e}_h + \overline{\mathbf{a}}^T(n)\mathbf{e}_l\big]^2 =& \frac{K(n)[K(n)-1]}{N(N-1)}(\mathbf{1}^T\mathbf{e}_h)^2 + \frac{K(n)[N-K(n)]}{N(N-1)}\mathbf{e}_h^T\mathbf{e}_h \\ &+ \frac{[N-K(n)][N-K(n)-1]}{N(N-1)}(\mathbf{1}^T\mathbf{e}_l)^2 + \frac{K(n)[N-K(n)]}{N(N-1)}\mathbf{e}_l^T\mathbf{e}_l \\ &+ 2\frac{K(n)[N-K(n)]}{N(N-1)}\big[(\mathbf{1}^T\mathbf{e}_h)(\mathbf{1}^T\mathbf{e}_l) - \mathbf{e}_h^T\mathbf{e}_l\big].\end{aligned} \tag{29}$$

Substituting (28) and (29) into (27), substituting the result and (23) into (25) , collecting terms and using the identity: $\mathbf{e}_h^T\mathbf{e}_h + \mathbf{e}_l^T\mathbf{e}_l - 2\mathbf{e}_h^T\mathbf{e}_l = (\mathbf{e}_h - \mathbf{e}_l)^T(\mathbf{e}_h - \mathbf{e}_l)$ gives

$$\sigma^2(n) = \Big[N(\mathbf{e}_h - \mathbf{e}_l)^T(\mathbf{e}_h - \mathbf{e}_l) - (\mathbf{1}^T\mathbf{e}_h - \mathbf{1}^T\mathbf{e}_l)^2\Big] \cdot \frac{K(n)(N-K(n))}{N^2(N-1)}.$$

Expanding $K(n)$ using (1) and applying the definition of e_0 gives

$$\sigma^2(n) = \frac{1}{(N-1)}\left[e_0^2 - \frac{(\mathbf{e}_h - \mathbf{e}_l)^T(\mathbf{e}_h - \mathbf{e}_l)}{N\Delta^2}\right] \cdot \Big[x_{\min}x_{\max} + x^2(n) - (x_{\min} + x_{\max})x(n)\Big]. \tag{30}$$

It follows from (26), (30), and the definitions of $\overline{\beta}$ and $\overline{\sigma}^2$ that

$$\lim_{P\to\infty}\frac{1}{P}\sum_{n=1}^{P} R_{yy}(n,k) = (1+\alpha)\overline{R}_{xx}(k) + \overline{\beta} + \overline{\sigma}^2\delta(k).$$

Therefore, to finish the proof it is sufficient to show that

$$\overline{R}_{yy}(k) = \lim_{P\to\infty}\frac{1}{P}\sum_{n=1}^{P} R_{yy}(n,k) \tag{31}$$

with probability 1, or, equivalently, that for each k the sequence $y(n)y(n+k)$ obeys the strong law of large numbers. As above, the Kolmogorov Criterion will be used. Because of symmetry, it is sufficient to prove the result for $k \geq 0$. Choose any fixed $k \geq 0$. For each $p = 0,\cdots,k-1$, consider the subsequence $r_p(n) = y\big(n(k+1)+p\big)y\big(n(k+1)+p+k\big)$. Because $y(n)$ is a sequence of independent random variables, it follows that, for each p, $r_p(n)$ is a sequence of independent random variables. Moreover, since $y(n)$ is a bounded sequence, $r_p(n)$ must also be a bounded sequence. Therefore, there exists some number C' such that $\mathrm{Var}\big[r_p(n)\big] < C'$ for all n and all p. It follows from the Kolmogorov Criterion that for each p, $r_p(n)$ obeys the strong law of large numbers.

Since, by definition, $\mathrm{E}\big[r_p(n)\big] = R_{yy}\big(n(k+1)+p,k\big)$,

$$\lim_{P\to\infty}\frac{1}{P}\sum_{n=1}^{P} r_p(n) = \lim_{P\to\infty}\frac{1}{P}\sum_{n=1}^{P} R_{yy}\big(n(k+1)+p,k\big)$$

with probability 1. This holds for $p = 0,\cdots,k-1$, so it follows that

$$\lim_{P\to\infty}\frac{1}{P}\sum_{n=1}^{P}\sum_{p=0}^{k-1} r_p(n) = \lim_{P\to\infty}\frac{1}{P}\sum_{n=1}^{P}\sum_{p=0}^{k-1} R_{yy}\big(n(k+1)+p,k\big)$$

with probability 1 which is equivalent to (31) . ■

Note that the theorem includes an ergodic result. Although it is convenient to define the behavior of the stochastic encoder in terms of its statistics, in most applications it is the time-averages rather than the statistical averages that are of interest. Hence, the theorem was framed in terms of time-averages rather than statistical averages. Nevertheless, similar results concerning statistical averages are contained in the proof of the theorem. In particular, suppose that $x(n)$ is a wide-sense-stationary (wss) random process that is independent of $\mathbf{a}(n)$. Then the theorem would hold if all of the time averages were replaced by the corresponding statistical averages (e.g., (21) would be valid if $\overline{M}_y$ and $\overline{M}_x$ were replaced by $\mathrm{E}\big[y(n)\big]$ and $\mathrm{E}\big[x(n)\big]$, respectively). Similarly, the theorem can be trivially modified to apply to the class of quasi-stationary input sequences (for a discussion of quasi-stationary sequences, see [18]). The quasi-stationary version of the theorem explicitly contains both the theorem as presented and the version pertaining to wss sequences as special cases. However, it is the opinion of the authors that this slight generalization unnecessarily complicates an otherwise simple and elegant result.

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers and Dr. G. Zimmerman of the Jet Propulsion Laboratory for their suggestions on how to improve the paper. The authors are also grateful to Professor T. Lang of the University of California, Irvine for introducing them to the Beneš Network.

REFERENCES

[1] M. J. Flanagan, G. A. Zimmerman, "Spur-reduced digital sinusoid synthesis," *IEEE Trans. Commun.*, to be published.
[2] D. Buchanan, "Choose DAC's for DDS system applications," *Microwaves & RF*, Aug. 1992.
[3] J. C. Candy and G. C. Temes, "Oversampling methods for A/D and D/A conversion," *Oversampling Delta-Sigma Data Converters Theory, Design and Simulation*. New York: IEEE Press, 1992, pp. 1-25.
[4] I. Galton, "Granular quantization noise in a class of delta-sigma modulators," *IEEE Trans. Inform. Theory*, vol. 40, pp. 848-859, May 1994.
[5] K. B. Klaassen, "Digitally controlled absolute voltage division," *IEEE Trans. Instrum. Meas.*, vol. IM-24, pp. 106-112, June 1975.
[6] R. J. van de Plassche, "Dynamic element matching for high-accuracy monolithic D/A converters," *IEEE J. Solid-State Circuits*, vol. SC-11, pp. 795-800, Dec. 1976.
[7] R. J. van de Plassche and D. Goedhart, "A monolithic 14-bit D/A converter," *IEEE J. Solid-State Circuits*, vol. SC-14, pp. 552-556, June 1979.
[8] L. R. Carley, "A noise shaping coder topology for 15+ bits converters," *IEEE J. Solid-State Circuits*, vol. 24, pp. 267-273, Apr. 1989.
[9] L. R. Carley and J. Kenney, "A 16-bit 4th order noise-shaping D/A converter," in *IEEE Proc. CICC*,1988, pp. 21.7.1-21.7.4.
[10] R. van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*. Boston, MA: Kluwer Academic, 1994.
[11] R. T. Leighton, *An Introduction to Parallel Algorithms and Architectures: Arrays · Trees · Hypercubes*. San Mateo, CA: Morgan Kaufmann, 1992.
[12] B. J. Hosticka, "Improvement of the gain of MOS Amplifiers," *IEEE J. Solid-State Circuits*, vol. SC-14, pp. 1111-1114, Dec. 1979.
[13] J. B. Shyu, G. C. Temes, and F. Krummenacher, "Random error effects in matched MOS capacitors and current sources," *IEEE J. Solid-State Circuits*, vol. SC-19, pp. 948-955, Dec. 1984.
[14] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, pp. 1433-1440, Oct. 1989.
[15] R. M. Gray, T. G. Stockham Jr., "Dithered quantizers," *IEEE Trans. Inform. Theory*, vol. 39, pp. 805-812, May 1993.
[16] A. V. Oppenheim, R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[17] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed. New York: Wiley, 1950.
[18] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

Ian Galton (M'92) for a photograph and biography, see p. 630 of the October 1995 issue of this TRANSACTIONS.

Paolo Carbone was born in Bolzano, Italy in 1965. He received the laurea degree and the Ph.D. degree in electronic engineering from the University of Padua in 1990 and 1994, respectively.

Since 1994 he has been a Researcher with the Third University of Rome. His research interests include digital signal processing and A/D–D/A conversion techniques.