Characterization and Modeling of Mismatch in MOS Transistors for Precision Analog Design

KADABA R. LAKSHMIKUMAR, MEMBER, IEEE, ROBERT A. HADAWAY, AND MILES A. COPELAND, SENIOR MEMBER, IEEE

Abstract—**This paper is concerned with the design of precision MOS analog circuits. Section II of the paper discusses the characterization and modeling of mismatch in MOS transistors. A characterization methodology is presented that accurately predicts the mismatch in drain current over a wide operating range using a minimum set of measured data. The physical causes of mismatch are discussed in detail for both p- and n-channel devices. Statistical methods are used to develop analytical models that relate the mismatch to the device dimensions. It is shown that these models are valid for small-geometry devices also. Extensive experimental data from a 3-μm CMOS process are used to verify these models.**

Section III of the paper demonstrates the application of the transistor matching studies to the design of a high-performance digital-to-analog converter (DAC). A circuit design methodology is presented that highlights the close interaction between the circuit yield and the matching accuracy of devices. It has been possible to achieve a circuit yield of greater than 97 percent as a result of the knowledge generated regarding the matching behavior of transistors and due to the systematic design approach.

I. Introduction

THE DESIGN of precision analog circuits requires a thorough understanding of the matching behavior of components available in any given technology. In MOS technology, capacitors are being widely used for designing precision analog circuits such as data converters [1] and switched-capacitor filters [2], [3] because of their excellent matching characteristics [4]. The matching behavior of MOS capacitors has been discussed in detail [5]–[7]. However, all precision analog circuits cannot be designed using capacitors alone. For applications such as high-speed data conversion, capacitive techniques tend to be too slow. Further a digital VLSI process may not offer linear capacitors. These factors motivated us to study the matching behavior of MOS transistors.

Manuscript received April 21, 1986; revised July 14, 1986. The Carleton University part of this work was supported by the Canadian Commonwealth Scholarship and Fellowship Committee and the Natural Sciences and Engineering Research Council of Canada.

K. R. Lakshmikumar was with the Department of Electronics, Carleton University, Ottawa, Ont. K1S 5B6, Canada. He is now with AT&T Bell Laboratories, Murray Hill, NJ 07974.

R. A. Hadaway is with Northern Telecom Electronics Ltd., Ottawa, Ont., Canada.

M. A. Copeland is with the Department of Electronics, Carleton University, Ottawa, Ont. K1S 5B6, Canada.

IEEE Log Number 8610606.

Section II of this paper discusses the characterization and modeling of mismatch in MOS transistors. The interest in such a study is evidenced by recent publications [7], [8]. In [8] experimental results of matching of MOS current mirrors are discussed without any reference to the physical causes of mismatch. The work reported in [7] attempts to break down the causes of mismatch but the experimental results are limited to large-area n-channel devices only. The work reported here is aimed at providing a more comprehensive understanding of the causes of mismatch in both p- and n-channel devices of large and small geometry. As the circuit designer has freedom to choose only the device dimensions, analytical models have been developed that relate the electrical mismatch to the dimensions. Extensive data to verify these models are obtained from a 5-V, 3 μm, p-well CMOS process that is in use at Northern Telecom Electronics Limited, Ottawa, Canada.

Section III of the paper demonstrates the application of the knowledge of matching behavior for the design of a high-speed digital-to-analog converter (DAC). Of late, the area of high-speed data converters in MOS technology is gaining importance (for example, [23] and [24]). However, these designs do not indicate the circuit yield obtainable for a given resolution, or the possibility of extension of these techniques to higher resolution converters. Therefore a circuit design methodology is presented here that relates the achievable linearity and yield to the matching accuracy of the components. A high-performance DAC with a circuit yield of greater than 97 percent has been realized without using any post-process trimming and yet occupying a small chip area using this design methodology [9].

II. Transistor Matching Studies

In general, there are two variations to consider in an integrated circuit process. *Global variation* accounts for the total variation in the value of a component over a wafer or a batch. *Local variation* or *mismatch* reflects the variation in a component value with reference to an adjacent component on the same chip. As the design of precision analog circuits is based on component ratios rather than their

0018-9200/86/1200-1057$01.00

absolute values, we have concentrated our study on the mismatch behavior.

The characterization of mismatch in MOS transistors is more complex than that in the case of capacitors. The drain current matching not only depends on the device dimensions but also on the operating point. In Section II-A, a characterization methodology is developed that accurately determines the mismatch in drain current over a wide operating region, using a minimum set of measurement data. The physical causes of mismatch are discussed in Section II-B, and analytical expressions to relate the mismatch to device dimensions are developed. We call these quantitative relationships *mismatch models*.

A. Characterization Methodology

Our aim is to predict the mismatch in the drain current over a wide range of operating conditions using a minimum set of measured data, and simultaneously to throw light on the detailed causes of mismatch. This problem can be best approached by measuring the mismatch in various parameters of a suitable circuit model [10]. The model chosen should be such that it gives an adequate description of the electrical behavior of the device, and at the same time should have readily measurable parameters that are amenable to statistical description. As an elaborate circuit model may greatly exceed the accuracy of measurable data or may hamper the extraction of statistically significant model parameters, we chose the simple square-law model. The current–voltage relationship in the triode region is given by

$$I = K(V_{GS} - V_T - V_{DS}/2)V_{DS} \tag{1}$$

where I is the drain current, K is the conductance constant, V_T is the threshold voltage, and V_{DS} is the drain-to-source voltage. The statistically significant parameters of this model are V_T and K. The mismatch in V_T accounts for the variations in the different charge quantities, and in the gate oxide capacitance per unit area. The variations in the dimensions, channel mobility, and gate oxide capacitance per unit area are measured as the mismatch in K. As both V_T and K are dependent on the gate oxide capacitance per unit area, we need to measure the correlation between the mismatches in V_T and K also.

The square-law model (1) is not an accurate description of the current–voltage relationship. It should be noted that we are only looking for local variations and are not so much concerned about the estimation of the absolute value of the parameters. Therefore any small model error would cancel out to a first order while estimating the mismatch, and hence the square-law model should suffice for our application. Several assumptions are made while deriving this one-dimensional model. As some of these are not strictly applicable, a further discussion to justify the use of this model is in order.

The gradual channel approximation and the assumption that the substrate is uniformly doped do not necessarily hold for small-geometry devices. An accurate analysis calls for a two-dimensional solution of Poisson's equation. However, in order to develop analytical expressions for the mismatch behavior, we use a one-dimensional circuit model and apply appropriate corrections to account for the effects of dimensional dependence on threshold voltage and nonuniform doping of the substrate. Also we have assumed a simplified picture of the oxide–silicon interface, i.e., that oxide fixed charge and interface trap charges are considered to be smeared-out uniform charge sheets. In fact the oxide fixed charge and charged interface traps are localized, and not sheets [11]. The accurate calculation of the surface potential would then be a two-dimensional problem. We are simplifying the problem by estimating the aggregate of the localized nonuniformity over the entire area of the channel by using the simplified one-dimensional circuit model. With these approximations, we develop analytical expressions for the mismatch behavior that compare remarkably well with experimental data.

Generally, MOS transistors will be operating in the saturation region in analog circuits. Therefore we should relate the measured mismatches in V_T and K to the saturation region, where the drain current is given by

$$I = \frac{K}{2}(V_{GS} - V_T)^2. \tag{2}$$

Then the variance in the drain current may be written as

$$\frac{\sigma_I^2}{\bar{I}^2} = \frac{\sigma_K^2}{\bar{K}^2} + 4\frac{\sigma_{VT}^2}{(V_{GS} - \bar{V}_T)^2} - 4r\frac{\sigma_{VT}}{V_{GS} - \bar{V}_T} \cdot \frac{\sigma_K}{\bar{K}} \tag{3}$$

following the derivation in [12] concerning the variance of a function of two random variables. Here r is the correlation coefficient between the mismatches in V_T and K, $\bar{I}$ is the expected value of the random variable I, σ_I is the standard deviation of I, and so on. Thus the mismatch in drain current at any operating point may be estimated if σ_K, σ_{VT}, and r are known.

Experimentally, V_T and K are determined by measuring the drain current versus gate voltage for a small value of V_{DS}. The maximum slope of the I versus V_{GS} curve provides the value of K. V_T is obtained from the intercept of the maximum slope on to the V_{GS} axis. If ΔK_i is the difference in the value of K for the ith matched pair of devices, the standard deviation of K is given by

$$\sigma_K = \left[\frac{1}{N-1}\left\{\sum_{i=1}^{N}(\Delta K_i)^2 - \frac{1}{N}\left(\sum_{i=1}^{N}\Delta K_i\right)^2\right\}\right]^{1/2} \tag{4}$$

where N is the number of matched pairs measured on each wafer. The second term in (4) is close to zero as the matched pairs are laid out in such a way as to minimize systematic mismatch. σ_{VT} is also computed in a similar way.

B. Factors Causing Mismatch

1. Threshold Voltage Mismatch: The threshold voltage of a transistor may be expressed as

$$V_T = \phi_{MS} + 2\phi_B + \frac{Q_B}{C} - \frac{Q_f}{C} + \frac{qD_I}{C} \tag{5}$$

where ϕ_{MS} is the gate–semiconductor work function difference, ϕ_B is the Fermi potential in the bulk, Q_B is the depletion charge density, Q_f is the fixed oxide charge density, D_I is the threshold adjust implant dose, and C is the gate oxide capacitance per unit area. The last term in (5) accounts for the threshold adjust implant where the implanted ions are assumed to have a delta function profile at the silicon–silicon dioxide interface [14]. The standard deviation of V_T may be determined if we can find the standard deviations of the various terms on the right-hand side of (5). The Fermi potential ϕ_B has a logarithmic dependence on the substrate doping, and ϕ_{MS} has a similar dependence on the doping in the substrate and in the polysilicon gate. Hence these terms may be regarded as constants not contributing to any mismatch.

Next we consider oxide fixed charge which is reported to have a Poisson distribution [11, p. 242]. Then its variance is given by

$$\sigma_{Qf}^2 = \frac{q\bar{Q}_f}{\bar{L}\bar{W}} \tag{6}$$

where L is the effective length and W is the effective width of the channel.

The depletion charge per unit area Q_B is also a random variable dependent on the distribution of the dopant atoms. This is an important difference between the treatment given here and the one reported in [7], where Q_B is treated as a constant. In fact, it is reported in [11, p. 237] that the dopant ions are nonuniformly distributed in MOS devices. No theoretical treatment of fluctuations in dopant ion density is available. However, we shall show that the physical conditions in the substrate favor a Poisson distribution [13]. The number of atoms per unit volume in silicon is $5\times10^{22}\ \mathrm{cm}^{-3}$. Only a very small fraction of these sites are occupied by the dopant atoms. The number of dopant atoms in nonoverlapping volumes is independent. Further, for domains of sufficiently small volume, the probability of finding exactly one dopant atom in a domain is proportional to the volume, and the probability of finding more than one atom is negligible. Hence the dopant ions may be considered to have Poisson distribution. Then the variance in Q_B may be shown to be [15]

$$\frac{\sigma_{QB}^2}{\bar{Q}_B^2} = \frac{1}{4\bar{L}\bar{W}\bar{W}_d\bar{N}_A} \tag{7}$$

where W_d is the depletion layer width and N_A is the substrate doping.

Similarly, assuming the implanted ions to follow a Poisson distribution, the variance of D_I is given by

$$\sigma_{DI}^2 = \frac{\bar{D}_I}{\bar{L}\bar{W}}. \tag{8}$$

The variance in C may be determined by estimating the variances in oxide thickness and permittivity [7]. It can be shown that [15]

$$\frac{\sigma_C^2}{\bar{C}^2} = \frac{1}{\bar{L}\bar{W}} A_{ox} \tag{9}$$

where A_{ox} is a parameter to be determined from measurements.

The random variables Q_f, Q_B, D_I, and C are all independent. Hence the variance in V_T may be written using (5) as

$$\sigma_{VT}^2 = \frac{1}{\bar{C}^2}\left(\sigma_{QB}^2 + \sigma_{QF}^2 + q^2\sigma_{DI}^2\right) + \frac{\sigma_C^2}{\bar{C}^2}\left(\frac{\bar{Q}_B^2}{\bar{C}^2} + \frac{\bar{Q}_f^2}{\bar{C}^2} + \frac{q^2\bar{D}_I^2}{\bar{C}^2}\right). \tag{10}$$

Substituting (6)–(9) into (10) we have

$$\sigma_{VT}^2 = \frac{1}{\bar{L}\bar{W}\bar{C}^2}\left[q\left(\bar{Q}_B + \bar{Q}_f + q\bar{D}_I\right) + A_{ox}\left(\bar{Q}_B^2 + \bar{Q}_f^2 + q^2\bar{D}_I^2\right)\right]. \tag{11}$$

Now let us examine the importance of the various terms on the right-hand side of (11), for both p- and n-channel devices. Consider n-channel devices first. In our process, the threshold adjust implant is carried out for p-channel transistors only. Therefore $qD_I = 0$ for n-channel devices. The depletion charge density per unit area is

$$Q_B = 7.7\times10^{-8}\ \mathrm{C/cm^2}. \tag{12}$$

In a well-controlled process the number of fixed oxide charges can be reduced to about $2\times10^{10}/\mathrm{cm^2}$, and hence

$$Q_f = 3.2\times10^{-9}\ \mathrm{C/cm^2}. \tag{13}$$

Comparing (12) and (13) we may infer that the contribution of the variability of the fixed oxide charges to threshold voltage mismatch (11) may be neglected.

The measured relative standard deviation of the threshold voltage ($\sigma_{VT}/\bar{V}_T$) is plotted against the reciprocal of the square root of the effective channel area in Fig. 1 for n-channel devices with six different W/L (drawn) values. $\sigma_{VT}/\bar{V}_T$ is chosen as the ordinate so as to express the variation as a percentage, independent of the operating point. However, if one is interested in current mismatch only, then $\sigma_{VT}/(V_{GS} - \bar{V}_T)$ should be plotted so that it may be directly used in (3).

For collecting statistics, 128 device pairs of each size were measured on every wafer. The vertical error bars reflect the spread in measured values over four wafers. Although the error bars appear large in this figure, in reality they represent relatively small deviations. For ex-

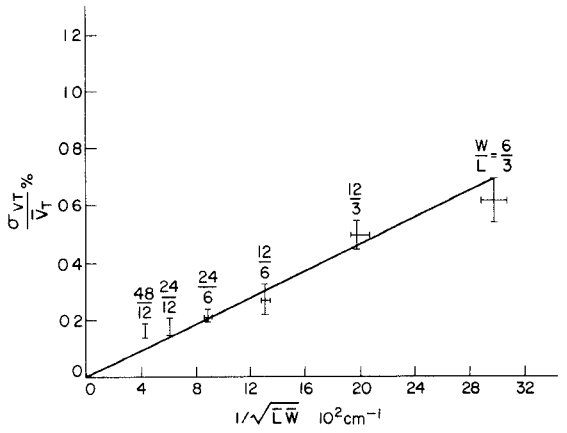


Fig. 1. Threshold voltage mismatch versus dimensions for n-channel devices.

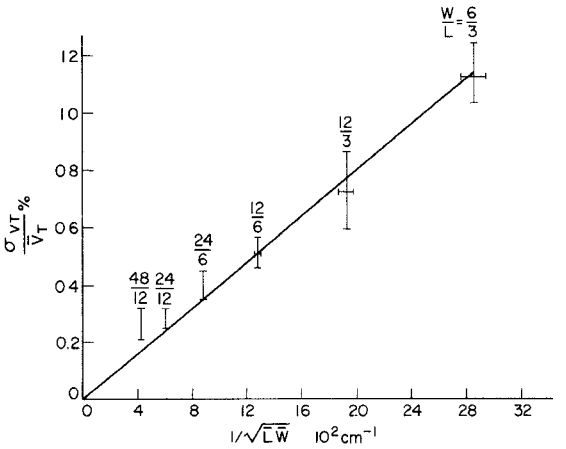


Fig. 2. Threshold voltage mismatch versus dimensions for p-channel devices.

ample, devices with $W/L = 6\ \mu m/3\ \mu m$ have a spread in the standard deviation of matching of threshold voltage of 3.5–4.9 mV. This spread is partly due to the nominal process variations from wafer to wafer, and partly due to the dependence of the matching on the electrical dimensions of the device. The effective channel length and width of devices were measured electrically at different places on a wafer and also on different wafers. The spread in the values is indicated by the horizontal error bars. The measured data fit well with the theoretical straight line relationship given by (11), which may be approximated for n-channel devices as

$$\sigma_{VT}/\bar{V}_T = \frac{1}{\sqrt{\bar{L}\bar{W}}}\left(2.5875\times10^{-12} + 1.2421\ A_{ox}\right)^{1/2}/\bar{V}_T. \tag{14}$$

Comparing the slope of the line in Fig. 1 with that in (14), it is found that

$$A_{ox} = 6.4631\times10^{-14}\ \text{cm}^2. \tag{15}$$

Then from (9), $\sigma_C/\bar{C} = 0.02$ percent for a $24\times6\text{-}\mu\text{m}^2$ gate. This low value agrees with the extremely uniform nature of the gate oxide thickness observed in other measurements [16].

Now we consider p-channel devices. As the threshold adjust implant is a very shallow one, a considerable portion of the implanted ions is retained in the gate oxide. Although this results in charged states, they are readily annealed during subsequent processing [17]. However, the presence of these impurity atoms in the oxide may cause a degradation in the capacitance matching of p-channel devices as compared to n-channel ones. For our process $Q_B = 4.810\times10^{-8}$ C/cm^2 and $qD_I = 8.0\times10^{-8}$ C/cm^2. Hence the contribution of Q_f to threshold voltage mismatch may be ignored and (11) may be written as

$$\sigma_{VT}/\bar{V}_T = \frac{1}{\sqrt{\bar{L}\bar{W}}}\left(4.2945\times10^{-12} + 1.8463\ A_{ox}\right)^{1/2}/\bar{V}_T. \tag{16}$$

The numerical coefficients in (16) are larger than the corresponding ones in (14) indicating a larger mismatch in p-channel devices, owing to the additional threshold adjust implant. This may be physically interpreted as a larger variation in the surface concentration due to the differential doping occurring at the surface.

The mismatch in threshold voltage of p-channel devices is plotted in Fig. 2. The data fit very well into the theoretical straight line relationship, and it is found that

$$A_{ox} = 3.0369\times10^{-12}\ \text{cm}^2. \tag{17}$$

Comparing (15) and (17) we may infer that the gate oxide capacitance matching is poorer for p-channel devices than that for n-channel ones.

We will now summarize the threshold voltage mismatch behavior. These findings are particular to this work.

a) The standard deviation of mismatch is inversely proportional to the square root of the effective channel area.

b) In a well-controlled process the nonuniform distribution of the fixed oxide charges has negligible effect on threshold voltage mismatch.

c) The nonuniform distribution of the dopant atoms in the bulk is a major contributor to the threshold voltage mismatch. The assumption that these atoms follow a Poisson distribution has resulted in excellent agreement with measurements.

d) Devices which use a compensating threshold adjust implant have a higher mismatch in threshold voltage due to the differential doping occurring at the surface. This is the major reason for the significantly larger mismatch noticed in p-channel devices as compared to n-channel transistors.

e) The gate oxide capacitance is quite uniform and hence has little influence on the threshold voltage mismatch. However, between n- and p-channel devices, the gate oxide capacitance of the latter has slightly poorer matching characteristic. This could be due to the nonuniform distribution of the threshold adjust implant atoms in the gate oxide.

2. Conductance Constant Mismatch: The conductance constant is given by

$$K = \mu C\ W/L \tag{18}$$

where μ is the channel mobility. We can express the variance in K in terms of the variances in μ, C, W, and L. Let us first consider the length of the device.

Electrically, the channel length is the average distance between the source and the drain diffusions. Any raggedness in the definition of the polysilicon may not be exactly reproduced in the source and drain diffusion edges. Further, we are not so much concerned about this raggedness; rather we are interested in the difference in the electrical length from one device to the next. Such a mismatch in length may not be due to the nonuniformity of the edge alone. In the absence of a complete knowledge of the causes of variation in length, we will simply indicate the variance of the length by σ_L^2 and make no attempt to derive any expression for it. In [7] the nonuniformity of the edges is the only cause considered for the mismatch and an expression for σ_L^2 is derived which is inversely proportional to the width of the device. This would mean that the mismatch in length would tend to zero for very wide devices. We have observed results that contradict this. For example, the conductance constant matching of devices with $W=100$ μm and $L=2$ μm is not all that different from devices with $W=200$ μm and $L=2$ μm. In fact we have noticed that σ_L is more or less independent of the device width.

The width of the device may be treated similarly and thus we let the variance in W be σ_W^2. The definitions of the length and width occur during different stages of processing and under different conditions. Hence L and W may be treated as independent random variables.

To determine the variance in mobility, a knowledge of the factors that affect it is required. It is reported in [18] that at room temperature and moderate gate bias the electron mobility is mainly governed by scattering due to interface charge centers and phonons. An empirical relationship for μ is [18]

$$\mu = \frac{\mu_0(N_A)}{1+\alpha(N_A)N_f} \tag{19}$$

where $\mu_0(N_A)$ and $\alpha(N_A)$ are empirical constants with very little dependence on the dopant concentration. Thus the mismatch in μ may be approximated to be entirely due to the nonuniformity of Q_f. As the fixed oxide charges have a Poisson distribution, we may write

$$\sigma_\mu^2 = \frac{\mu_0^2\alpha^2}{\left(1+\alpha\bar{Q}_f\right)^4}\cdot\frac{\overline{N_f}}{\overline{L}\,\overline{W}}. \tag{20}$$

The discussion given above for the mobility mismatch is for electrons only. We are not aware of any model that relates the mobility of holes to the doping concentration in the bulk and the fixed oxide charge density. The situation in the case of p-channel devices is further complicated by the threshold adjust implant. This could cause some damage in the substrate which may not be completely annealed, resulting in a poorer mobility matching than in the case of n-channel devices. In spite of these uncertainties, it is still reasonable to assume that the mobility mismatch has a similar dependence on channel dimensions as given by (20). Then

$$\frac{\sigma_\mu}{\bar{\mu}} = \left(\frac{A_\mu}{\overline{L}\,\overline{W}}\right)^{1/2} \tag{21}$$

where $\sqrt{A_\mu} = 4.95\times10^{-7}$ cm for n-channel devices and is not known for p-channel transistors.

The factors on the right-hand side of (18) are all independent. Thus

$$\frac{\sigma_K^2}{\bar{K}^2} = \frac{\sigma_L^2}{\bar{L}^2} + \frac{\sigma_W^2}{\bar{W}^2} + \frac{\sigma_\mu^2}{\bar{\mu}^2} + \frac{\sigma_C^2}{\bar{C}^2}. \tag{22}$$

From (9) and (21)

$$\frac{\sigma_K^2}{\bar{K}^2} = \frac{1}{\overline{L}\,\overline{W}}\left(A_\mu + A_{ox}\right) + \frac{\sigma_L^2}{\bar{L}^2} + \frac{\sigma_W^2}{\bar{W}^2}. \tag{23}$$

After substituting the values of A_μ and A_{ox} for n-channel devices, (23) may be solved for σ_L and σ_W using the measured values of σ_K of different sized devices. It is found that σ_L and σ_W are approximately the same and in the range of 0.01–0.03 μm. To provide a feel for the relative importance of the factors causing mismatch in K, we may substitute $\sigma_L = \sigma_W = 0.02$ μm in (23). Then

$$\frac{\sigma_K^2}{\bar{K}^2} = \left(2.46\times10^{-13} + 0.646\times10^{-13}\right)\cdot\frac{1}{\overline{L}\,\overline{W}} + 4\times10^{-12}\left(\frac{1}{\bar{L}^2} + \frac{1}{\bar{W}^2}\right) \tag{24}$$

where the effective dimensions $\bar{L}$ and $\bar{W}$ have the units of centimeters. $\sigma_K/\bar{K}$ is plotted against $(1/\bar{L}^2 + 1/\bar{W}^2)^{1/2}$ in Fig. 3. The plotted relationship is not linear as shown by (24), with the curvature increasing for smaller geometry devices due to the increasing contribution of the $1/\overline{L}\,\overline{W}$ term. A similar plot for p-channel devices is shown in Fig. 4. The p-channel devices have a larger mismatch in conductance constant. One reason for this is the poorer gate oxide capacitance matching as has already been pointed out in connection with threshold voltage mismatch. Another factor could be a larger mobility variation.

We will now summarize the mismatch behavior in the conductance constant. These results are particular to this work.

a) The mismatch in K due to edge variations is proportional to $(1/\bar{L}^2 + 1/\bar{W}^2)^{1/2}$. The standard deviation of mismatch in length and width is in the range 0.01–0.03 μm. For n-channel devices, this is the dominant source of mismatch in K.

b) The larger gate oxide capacitance variation observed in p-channel devices in connection with V_T mismatch agrees with the larger mismatch in K.

c) For n-channel devices, the variation in mobility has little effect on the mismatch in K. The corresponding quantity for p-channel transistors, however, could be larger

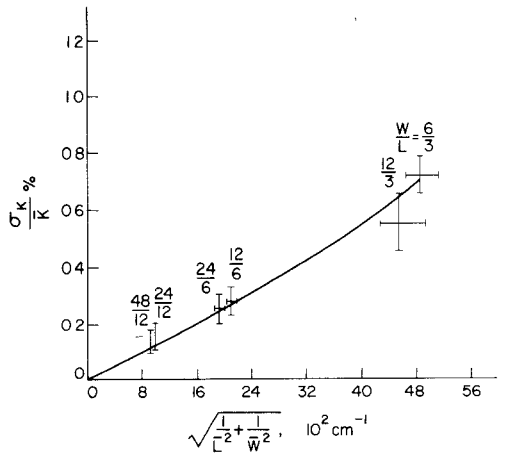


Fig. 3. Conductance constant mismatch versus dimension for n-channel devices.

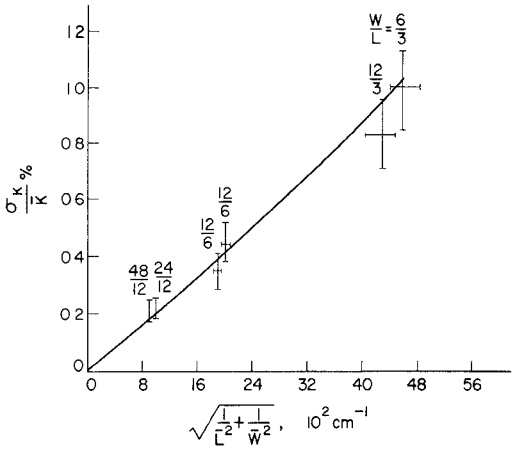


Fig. 4 Conductance constant mismatch versus dimension for p-channel devices.

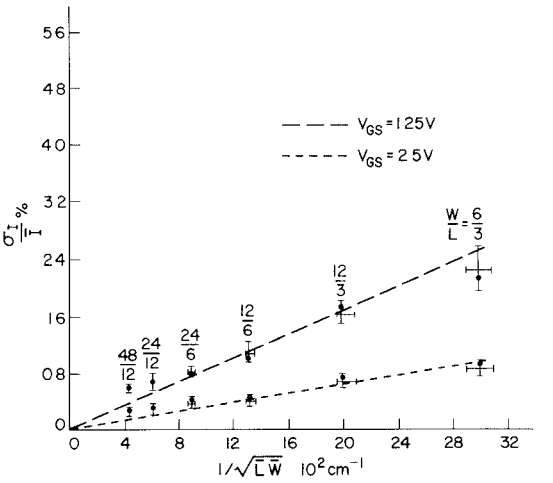


Fig. 5. Drain current mismatch versus dimension for n-channel devices The dots are the estimated values using (25).

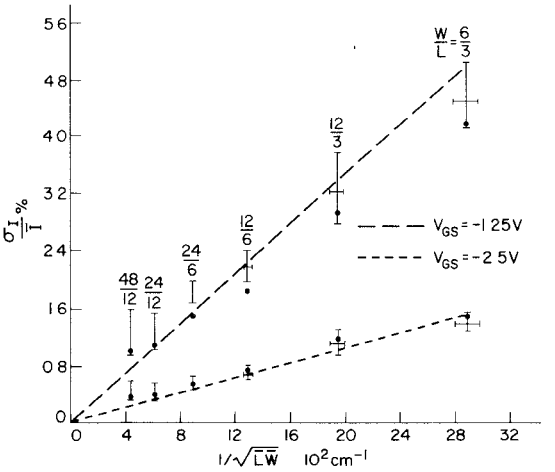


Fig. 6. Drain current mismatch versus dimension for p-channel devices. The dots are the estimated values using (25).

due to any damage in the substrate caused by the threshold adjust implant.

3. Correlation Between Mismatches in V_T and K: A common contributing factor to the mismatches in V_T and K is the variation in the gate oxide capacitance. Hence we can expect a dependence between the mismatches in V_T and K. A theoretical expression for the correlation coefficient is derived in [15]. Also the value has been experimentally measured. The agreement is excellent for n-channel devices and fair for p-channel ones. However, both the theoretical and experimental values are close to zero indicating that the mismatches in V_T and K are almost independent.

4. Mismatch in Drain Current: The drain current mismatch in the saturation region is given by (3). As the correlation coefficient is nearly equal to zero, we have

$$\frac{\sigma_I^2}{\bar{I}^2} = \frac{\sigma_K^2}{\bar{K}^2} + 4\frac{\sigma_{VT}^2}{(V_{GS}-\bar{V}_T)^2}. \tag{25}$$

At low values of gate-to-source voltage the dominant factor causing the mismatch in drain current is the threshold voltage variation. For bias levels approaching the mid-rail, the conductance constant and the threshold voltage mismatches have almost equal contributions to the drain current mismatch. From (25) σ_I may be estimated from the measured values of σ_K and σ_{VT}. Also we have actually measured σ_I at different gate biases to compare with the estimated values. Fig. 5 is a plot of $\sigma_I/\bar{I}$ versus $1/\sqrt{\bar{L}\bar{W}}$ for n-channel devices for two gate voltages. The estimated values obtained from (25) using the measured average values of σ_{VT} and σ_K are indicated by the dots. A similar result is shown for the p-channel devices in Fig. 6. The excellent agreement between the measured values and the estimated ones for both p- and n-channel devices validates the characterization methodology and also verifies the mismatch models.

5. Range of Applicability: It is important to consider the dimensional range over which the mismatch models we have developed in the preceding sections are accurate. In our analysis we have assumed that the dimensional variations are accounted for entirely by the mismatch in conductance constant and have no influence on the threshold voltage mismatch. As the threshold voltage of small-geom-

etry devices is a function of channel length and width, it is necessary to estimate the mismatch in threshold voltage brought about by the dimensional variation to validate the above assumption and hence the mismatch models. To this end, the shift in threshold voltage brought about due to short-channel and narrow-width effects was estimated for a device with effective dimensions 2×2 μm^2 fabricated in our process, using the expressions in [14]. It was found that the mismatch component of the threshold voltage brought about by the dimensional variations is only 10 percent of the total threshold voltage mismatch, in the worst case [15]. We also verified this fact for even smaller device geometries using the process parameters given in [19]. Hence we may attribute the dimensional variations entirely to the mismatch in K and not to V_T. Thus we may conclude that the characterization methodology and the mismatch models are valid for small-geometry devices also. However, as new processes emerge permitting smaller geometry devices, further experimental work is needed to characterize the mismatch.

6. Effect of Temperature: As the threshold voltage and conductance constant vary with temperature, it is interesting to know their matching behavior as a function of temperature. In the case of the threshold voltage, as expressed by (5), the only terms that are dependent on temperature are ϕ_{MS} and ϕ_B. We have seen that the contribution of these terms to the threshold voltage mismatch is negligible. Therefore we may expect the matching behavior of threshold voltage to be almost independent of temperature.

The only factor through which the conductance constant matching can be affected is the temperature dependence of mobility. For n-channel devices we have seen that the mismatch in conductance constant is largely due to photolithographic edge variations, and mobility variations have the least effect. Thus temperature variations should have very little effect on the conductance constant matching of n-channel devices. Since the mismatch in drain current is due to mismatches in threshold voltage and conductance constant, we can expect the current mismatch in n-channel devices to be almost unaffected over a wide temperature range. Limited experimental results seem to agree with this prediction. As far as the p-channel devices are concerned, since the mobility behavior of holes is not clearly understood, no theoretical explanation of the temperature effect is possible.

III. Design Methodology

To demonstrate the usefulness of the study of the matching behavior of transistors and the related models, we took up the task of designing an 8-bit current-steering CMOS DAC. Circuit details of the DAC have already been presented [9]. Here we only indicate the design methodology. In Section III-A, a brief description of the DAC configuration is presented. Section III-B discusses statistical error analysis. The close interaction between the DAC configuration, the matching accuracy of devices, and the circuit yield is brought out here. Finally, yield results are presented in Section III-C.

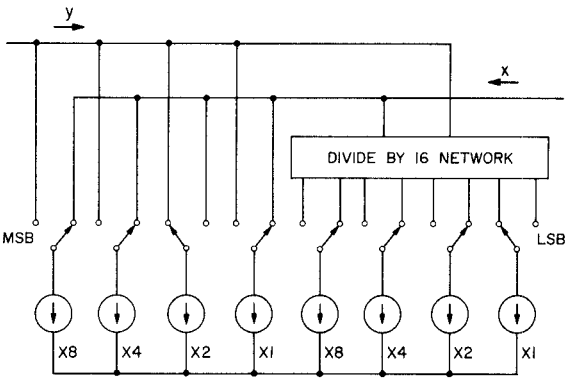


Fig. 7. Schematic of a multiple current-source DAC.

A. DAC Configuration

A multiple current-source approach was chosen to realize the binary-weighted currents. This was done primarily to overcome the problems of nonlinear relationship between the drain current and aspect ratio of small-geometry MOS devices [14], [15], [20], and the voltage coefficient of resistance of R-$2R$ networks. The configuration is shown in Fig. 7 and is similar to that reported in [21]. The least significant bit (LSB) has one *unit* current source, the next significant bit has two unit current sources connected in parallel, and so on. The exponential growth in the number of unit current sources is overcome by having an interstage 16:1 resistive current divider.

B. Statistical Error Analysis

In general, the errors generated by a DAC consist of linearity, offset, and gain errors. Usually, a DAC may be calibrated for zero gain and offset errors. Linearity error, however, occurs due to the random mismatch in the conversion elements. Hence, the circuit yield is a function of the matching accuracy of the unit current sources.

Integral nonlinearity of a DAC is generally defined as the difference between the actual output to the desired output normalized to the full-scale output of the DAC. This enables the nonlinearity to be expressed in terms of fractions of LSB or as a percentage.

Let x be the output of the DAC for a given input word and y be the analog complement of the output. The full-scale output of the converter is the sum $x+y$. To express the error as a fraction of the full scale, first we normalize the output to the full scale as follows:

$$z(x,y)=\frac{x}{x+y} \tag{26}$$

where z is the normalized output, and x and y are dependent on the input digital word and the DAC configuration.

For the 8-bit interstage divider DAC shown in Fig. 7

$$x=\left[8D_1+4D_2+2D_3+D_4+\frac{1}{16}(8D_5+4D_6+2D_7+D_8)\right]I_{\text{unit}} \quad (27)$$

and

$$y=\left[8(1-D_1)+4(1-D_2)+2(1-D_3)+(1-D_4)+\frac{1}{16}\{8(1-D_5)+4(1-D_6)+2(1-D_7)+(1-D_8)\}\right]I_{\text{unit}} \quad (28)$$

where

$$D_i=0 \text{ or } 1$$
$$i=1,2,\cdots,8$$

D_1 is the most significant bit (MSB), D_8 is the LSB, and I_{unit} is the unit current source with a mean value $\bar{I}$ and standard deviation of matching σ. As the unit current sources are random and uncorrelated, we may treat x and y to be independent random variables with standard deviations σ_x and σ_y, respectively. Now we may determine the standard deviation of z in terms of σ_x and σ_y [12]

$$\sigma_z^2=\frac{\bar{y}^2\sigma_x^2+\bar{x}^2\sigma_y^2}{(\bar{x}+\bar{y})^4} \quad (29)$$

where σ_x and σ_y are evaluated as follows:

$$\bar{x}=\left[8D_1+4D_2+2D_3+D_4+\frac{1}{16}(8D_5+4D_6+2D_7+D_8)\right]\bar{I}$$

and

$$\sigma_x^2=\left[8D_1+4D_2+2D_3+D_4+\frac{1}{16}(8D_5+4D_6+2D_7+D_8)\right]\sigma^2=\frac{\bar{x}}{\bar{I}}\cdot\sigma^2. \quad (30)$$

Similarly

$$\sigma_y^2=\frac{\bar{y}}{\bar{I}}\cdot\sigma^2. \quad (31)$$

Substituting (30) and (31) into (29)

$$\sigma_z^2=\frac{1}{\bar{I}}\cdot\frac{\bar{x}\bar{y}}{(\bar{x}+\bar{y})^3}\cdot\sigma^2. \quad (32)$$

The expected value of z may be shown to be [12]

$$\bar{z}=\frac{\bar{x}}{\bar{x}+\bar{y}}. \quad (33)$$

Using (33) in (32), we have

$$\sigma_z^2=\frac{\bar{z}(1-\bar{z})}{\bar{I}(\bar{x}+\bar{y})}\cdot\sigma^2. \quad (34)$$

The above result may also be derived by determining the joint probability density function of z as shown in [22]. Equation (34) expresses the variances of the D/A output for different digital words. The function $\bar{z}(1-\bar{z})$ will have a maximum value when $\bar{z}=1/2$, i.e., when the output is halfway through the full scale, and falls off towards zero for minimum and maximum input word combinations. This observation suggests that the MSB current could be the most critical and should have the highest accuracy. Error contributions of the bits taper off towards the LSB, and hence the relative error contributions of all the bit current sources need not be the same. Such an error distribution is indeed a natural consequence in the multiple current-source approach where the relative accuracy of the bits improves towards the MSB. This may be shown as follows. If the unit current source has a mean value $\bar{I}$ and standard deviation of matching σ, connecting n such sources in parallel would produce an equivalent current source with mean value $n\bar{I}$ and standard deviation $\sqrt{n}\,\sigma$, as the current sources are uncorrelated. Thus there is an improvement in accuracy by a factor $\sqrt{n}$.

The analysis so far has shown that maximum error occurs halfway through the full scale and the error contributions of the individual bits taper off towards the LSB. Now we proceed further to relate the circuit yield to the standard deviation of the unit current sources. Here we define circuit yield as the percentage of functional devices that have integral nonlinearity less than 1/2 LSB. In other words, we are eliminating catastrophic device failures due to defects, etc. With this definition, a theoretical estimate of the circuit yield of the DAC is obtained by multiplying the probabilities that each of the 256 outputs of the DAC have less than 1/2 LSB error. For normal distribution with variances given by (34), the yield is

$$G=\prod_{i=2}^{255}\frac{1}{\sqrt{2\pi}\,\sigma_z}\int_{\bar{z}-1/512}^{\bar{z}+1/512}\exp\left\{-\frac{(z-\bar{z})^2}{2\sigma_z^2}\right\}\cdot dz=\prod_{i=2}^{255}\operatorname{erf}\left(Q/\sqrt{2}\right) \quad (35)$$

where 1/512 is the normalized 1/2 LSB value and

$$Q=\frac{1}{512\left[\dfrac{\bar{z}(1-\bar{z})}{15+15/16}\right]^{1/2}}\frac{\sigma}{\bar{I}}. \quad (36)$$

The method used to derive (35) is quite general and may be easily extended to converters of different resolutions and accuracies. The circuit yield as given by (35) is plotted as a function of $\sigma/\bar{I}$ in Fig. 8.

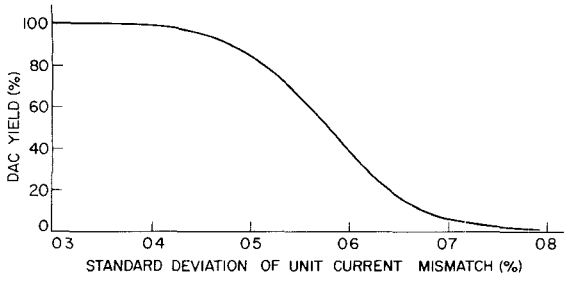


Fig. 8. DAC yield versus current-source mismatch.

TABLE I
DAC YIELD

Wafer Number	Devices Tested	Functional Devices	Good Devices	Circuit Yield %
1	32*	32	31	97
2	55*	55	55	100
3	64	60	60	100
4	64	52	52	100
5	64	60	60	100

*Broken Wafers. Hence the number of devices tested are less than 64.

Because of the error function nature of the relation between yield and current-source matching, there is an almost flat region close to the 100-percent yield level, followed by a very steep region and finally the yield asymptotically approaches zero. To avoid the possibility of any process variations from batch to batch affecting the yield adversely, the design should be such as to avoid the region where the yield is very sensitive to the matching accuracy. On the other hand, a very conservative matching tolerance is also not desirable because the improvement in yield is marginal with improvement in matching accuracy beyond a certain point. Therefore we chose 95-percent yield level as an optimum value that would not appreciably shift due to process variations from batch to batch. From the theoretical relationship shown in Fig. 8, the standard deviation of matching of the current sources should be about 0.45 percent to achieve this yield level. It may be noted that for an 8-bit DAC, an integral nonlinearity of $\pm 1/2$ LSB is equivalent to 0.2 percent of full scale. Thus, with this configuration, it is possible to provide good integral linearity associated with a high circuit yield without requiring an equivalent degree of component matching. Further it is shown in [15] that the divide-by-16 network is highly tolerant to component mismatch and hence is not a potential source of yield loss.

The analysis given above provides a systematic design approach for data converters. A similar approach may be used to design any precision analog circuit in general.

C. Results

Based on the understanding of the matching behavior of MOS devices and the systematic design methodology, an 8-bit high-speed DAC has been designed. The electrical performance results are reported in [9]. The unit current source used in the DAC is made up of a 24-μm-wide and 6-μm-long n-channel transistor in combination with a 4.7-kΩ source degradation resistor. This configuration has a better matching accuracy than a transistor alone for the same current value, owing to the better matching of resistors and the local negative feedback offered by the resistor [9]. The combination has a standard deviation of current matching of 0.45 percent when biased at a current of 128 μA. This should result in a circuit yield of approximately 95 percent. It may be noted that the same degree of matching may be obtained without using the source degradation resistor by choosing larger area devices, and/or operating the devices with a larger gate-to-source voltage. Knowledge of the mismatches in V_T and K in conjunction with (25) may be used to obtain curves such as those in Figs. 5 and 6, for any process and operating condition. This information when used with (35) or Fig. 8 completes the design cycle.

Circuit yield statistics of the DAC are presented in Table I. Column 2 indicates the number of devices that are tested on each wafer. The next column shows the number of devices that are functional. In other words, we are eliminating catastrophic failures here. The fourth column shows the number of devices with integral nonlinearity less than $\pm 1/2$ LSB. Finally, the circuit yield is shown in the last column. In most cases the circuit yield is 100 percent, demonstrating the accuracy of the device characterization and the circuit design methodology. We have been able to achieve this high level of yield using relatively small devices and without using any trimming because of the knowledge we generated regarding the matching behavior of the devices as a function of dimensions, and the systematic design methodology we have followed.

IV. Conclusion

The design of precision analog circuits presents challenges in the areas of device matching characterization and circuit design. Novel methodologies relevant to both aspects of the design are presented in this paper. Section II is devoted to the study of transistor matching behavior. The overall objective has been not only to provide a clear understanding of the random mismatch, but also to develop a comprehensive design approach for precision analog circuits. The parameters a circuit designer will have freedom to choose are the dimensions of the devices. Therefore analytical models have been developed that relate the mismatch to device dimensions.

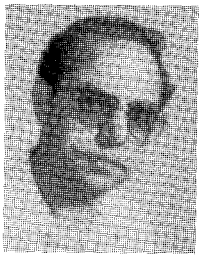
Section III of the paper discusses the application of the matching characterization in precision analog design. Design methodology for a high-performance DAC is illustrated. This is presently important because of the need for high-speed data converters in MOS technology. The close interaction between device matching and circuit yield is discussed. Experimental results of circuit yield are also presented.

ACKNOWLEDGMENT

The authors are indebted to the management of Northern Telecom Electronics Ltd., for extending their facilities to carry out this research. They would like to thank Dr. M. Simard-Normandin for her help with device characterization. Thanks are also due to M. King for many suggestions and discussions. The assistance of N. Prasad concerning statistical analysis is gratefully acknowledged.

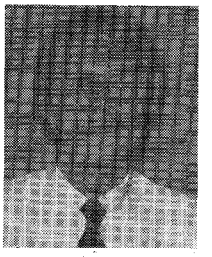
REFERENCES

[1] J. L. McCreary and P. R. Gary, "All MOS charge redistribution analog-to-digital conversion techniques—Part I," *IEEE J. Solid-State Circuits*, vol. SC-10, pp. 371–379, Dec. 1975.
[2] J. T. Caves, M. A. Copeland, C. F. Rahim, and S. D. Rosenbaum, "Sampled analog filtering using switched capacitors as resistor equivalents," *IEEE J. Solid-State Circuits*, vol. SC-12, pp. 592–599, Dec. 1977.
[3] B. J. Hosticka, R. W. Brodersen, and P. R. Gray, "MOS sampled data recursive filters using switched capacitor integrators," *IEEE J. Solid-State Circuits*, vol. SC-12, pp. 600–608, Dec. 1977.
[4] D. A. Hodges, P. R. Gray, and R. W. Brodersen "Potential of MOS technologies for analog integrated circuits," *IEEE J. Solid-State Circuits*, vol. SC-13, pp. 285–294, June 1978.
[5] J. L. McCreary, "Matching properties and voltage and temperature dependence of MOS capacitors," *IEEE J. Solid-State Circuits*, vol. SC-16, pp. 608–616, Dec. 1981.
[6] J. B. Shyu, G. C. Temes, and K. Yao, "Random errors in MOS capacitors," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 1070–1076, Dec. 1982.
[7] 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.
[8] M. Akyia and S. Nakashima, "High-precision MOS current mirror," *Proc. Inst. Elec. Eng.*, vol. 131, pt. I, pp. 170–175, Oct. 1984.
[9] K. R. Lakshmi Kumar, R. A. Hadaway, M. A. Copeland, and M. I. H. King, "A high-speed 8-bit current steering CMOS DAC," in *Proc. IEEE 1985 Custom Integrated Circuits Conf.*, pp. 156–159.
[10] J. Logan, "Characterization and modelling for statistical design," *Bell Syst. Tech. J.*, vol. 50, pp. 1105–1147, Apr. 1971.
[11] E. H. Nicollian and J. R. Brews, *MOS Physics and Technology*. New York: Wiley, 1982.
[12] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. Tokyo: McGraw-Hill, Kogakusha, 1965.
[13] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. I. New York, Wiley, 1957, p. 146.
[14] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: Wiley, 1981.
[15] K. R. Lakshmi Kumar, "Characterization and modelling of mismatch in MOS devices and application to precision analog design," Ph.D. dissertation, Carleton Univ., Ottawa, Ont., Canada, 1985.
[16] M. Simard-Normandin, private communication, 1985.
[17] S. K. Ghandhi, *VLSI Fabrication Principles Silicon and Gallium Arsenide*. New York: Wiley, 1983, p. 353.
[18] S. C. Sun and J. D. Plummer, "Electron mobility in inversion and accumulation layers on thermally oxidized silicon surfaces," *IEEE J. Solid-State Circuits*, vol. SC-15, pp. 562–573, Aug. 1980.
[19] J. R. Brews, W. Fitchner, E. H. Nicollian, and S. M. Sze, "Generalized guide for MOSFET miniaturization," *IEEE Electron Device Lett.*, vol. ED.-1, pp. 2–4, Jan. 1980.
[20] P. Yang and P. K. Chatterjee, "SPICE modelling for small geometry MOSFET circuits," *IEEE Trans. Computer-Aided Des.* vol. CAD-1, pp. 169–182, Oct. 1982.
[21] P. H. Saul *et al.*, "An 8-bit, 5-ns monolithic D/A converter subsystem," *IEEE J. Solid-State Circuits*, vol. SC-15, pp. 1033–1039, Dec. 1980.
[22] S. Kuboki *et al.*, "Nonlinearity analysis of resistor string A/D converters," *IEEE Trans. Circuits Syst.*, vol. CAS-29, pp. 383–390, June 1982.
[23] P. H. Saul *et al.*, "An 8b CMOS video DAC," in *ISSCC Dig. Tech. Papers*, 1985, pp. 32–33.
[24] T. Miki *et al.*, "An 80 MHz 8b CMOS D/A converter," in *ISSCC Dig. Tech. Papers*, Feb. 1986, pp. 132–133.



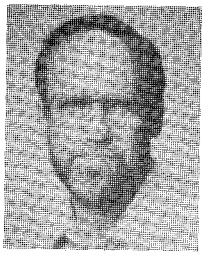
Kadaba R. Lakshmikumar (S'82–M'86) received the B.E. and M.E. degrees in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1974 and 1976, respectively, and the Ph.D. degree in electrical engineering from Carleton University, Ottawa, Ont., Canada, in 1985.

From 1976 to 1981 he worked at Bharat Electronics Limited, Bangalore, India, as a designer of bipolar analog integrated circuits. In 1981 he received the Canadian Commonwealth Scholarship to pursue doctoral studies in Canada. His thesis was a cooperative effort with Northern Telecom Electronics Limited, Ottawa, Ont., Canada. Since 1985 he has been a Member of the Technical Staff at AT&T Bell Laboratories, Murray Hill, NJ, where he is engaged in the design of high-frequency MOS analog integrated circuits.



Robert A. Hadaway was born in Toronto, Ont., Canada, on April 26, 1953. He is an Engineering Technology Graduate from Humber College of Applied Arts and Technology, Toronto, Ont., Canada.

In 1976 he joined Bell Northern Research, Ottawa, Ont., Canada, where he was involved in the design of analog/digital telephony circuits using bipolar and NMOS technologies. In 1980 his interests turned to technology development, specifically high-voltage n-channel transistors for display drivers. In 1982 he joined Northern Telecom Electronics, Ottawa, Ont., Canada, where he was engaged in the implementation of CMOS compatible high-voltage and nonvolatile memory devices. He is currently responsible for the development of enhancements to CMOS ASIC technologies.



Miles A. Copeland (M'65–SM'85) was born in Quebec City, Que., Canada, on May 5, 1934. He received the B.Sc. E.E. degree from the University of Manitoba, Man., Canada, in 1957, and the M.Eng. and Ph.D. degrees from the University of Toronto, Ont., Canada, in 1962 and 1965, respectively.

He joined Carleton University, Ottawa, Ont., Canada, in 1965 where he is now Chairman of the Department of Electronics. He has had continuing interaction with industry, including cooperative research and consulting with Bell-Northern Research, Northern Telecom, Mitel Corporation, and General Electric Corporate R&D Labs, and has been involved with many graduate students in cooperative thesis work with industry. His research interests include the design of MOS integrated circuits, both analog and digital, with particular interest in sampled analog circuits such as switched-capacitor filters. Recently, he has been interested in high bandwidth SC filters, CAD techniques for design of mixed digital and sampled-analog systems, and analog circuit design techniques for short-channel VLSI processes.