Short Papers

On an Efficient CAD Implementation of the Distance Term in Pelgrom's Mismatch Model

B. Linares-Barranco and T. Serrano-Gotarredona

***Abstract*—In 1989, Pelgrom *et al.* published a mismatch model for MOS transistors, where the variation of parameter mismatch between two identical transistors is given by two independent terms: a size-dependent term and a distance-dependent term. Some CAD tools based on a nonphysical interpretation of Pelgrom's distance term result in excessive computationally expensive algorithms, which become nonviable even for circuits with a reduced number of transistors. Furthermore, some researchers are reporting new variations on the original nonphysically interpreted algorithms, which may render false results. The purpose of this paper is to clarify the physical interpretation of the distance term of Pelgrom *et al.* and indicate how to model it efficiently in prospective CAD tools.**

***Index Terms*—Analog design, mismatch gradient planes, mismatch modeling, Pelgrom model, sigma-space analysis.**

I. Introduction

The mismatch model proposed by Pelgrom *et al.* [1] models the standard deviation in the mismatch of property P between two identical MOS transistors of width W and length L separated by a distance D (from center to center) in the layout, as

$$\sigma^2(\Delta P) = \frac{A_P^2}{WL} + S_P^2 D^2. \tag{1}$$

This model was experimentally verified by Pelgrom by measuring the mismatch on many dies, fabricated on many runs, and including many identical transistors per die, as well as many transistor sizes per die (see Fig. 2 in [1]). The model was verified for different foundries and technologies. Theoretical derivations of this model (see Appendix A) reveal that the two terms originate by two different means. The size-dependent term is caused by random fluctuations of material and technological properties of transistors. For example, dopant concentrations along the wafers and dies can be considered to be modeled by a constant term around which there exist random fluctuations. These random fluctuations are averaged over a transistor area WL contributing to a specific transistor electrical property (like threshold voltage or beta). The larger the area, the smaller the impact of the random fluctuations. Therefore, this random-induced mismatch term decreases with the transistor area. On the other hand, the distance term in (1), which is size independent (common for all sizes), is originated by gradients along the dies and wafers. For example, dopant concentrations follow smooth systematic surfaces along the wafers. Such surfaces are usually well known by IC manufacturers, although rarely made public. For a particular die, the surface can be approximated many times by a plane. Depending on the die position within the wafer, a different gradient plane will result. Consequently, in practice, this gradient plane has a random nature from die to die. Since Pelgrom measured many dies, his model includes the random characterization of these random gradient planes. Layout techniques, like common centroids, can eliminate the impact of gradient-induced mismatch and consequently eliminate the distance term from (1). However, layout techniques can never eliminate the random-induced size-dependent term because of its purely random nature.

In the next section, we describe how one can model both terms of the Pelgrom model in a circuit simulator. Then, in Section III, we explain another implementation based on nonphysical interpretations. In Section IV, we give some hints on how to further improve gradient-induced mismatch in CAD tools, and finally, in Section V, we give our conclusions.

II. Implementing Pelgrom's Mismatch Model

The random-induced size-dependent term is quite straightforward to add in a circuit simulation tool. First, one needs to know the critical transistor mismatch parameters, whose random fluctuations impact transistor currents. They are usually a small number of parameters. Pelgrom suggested two main ones (V_{T0} and β) and a secondary one (γ) [1]. Bastos *et al.* added mobility degradation [2], [3], Serrano *et al.* suggested a total of five relevant parameters [4], [5], and recent models extending from weak to strong inversion have been proposed with no more than five parameters [6]–[8]. For more sophisticated transistor models like BSIM, about 16 parameters would be required [9]. Consider a generic mismatch relevant parameter P. Let us assume $P_{\rm mean}$ is the mean value predicted by the manufacturer. Usually, the manufacturer also characterizes global interdie variations $\Delta P_{\rm global}$ common for all transistors in the same die, whose standard deviation $\sigma(\Delta P_{\rm global})$ would characterize variations from die to die. Finally, a local term ΔP_i has to be added for each transistor in the circuit, such that its standard deviation is characterized by (1). This local mismatch term includes two components: a random size-dependent component $\Delta P_{{\rm rand\_}i}$ and a gradient-induced size-independent component $\Delta P_{{\rm grad\_}i}$. The standard deviation of the random component is given by $\sigma(\Delta P_{{\rm rand\_}i}) = A_P/(\sqrt{2W_iL_i})$. Here, W_i and L_i are width and length of transistor i. The factor 2 in the denominator accounts for the fact that each transistor deviates from a nominal mismatchless transistor. Parameters A_P are provided by the manufacturer for the mismatch relevant parameters. Sometimes, correlations between these parameters are also characterized. In these cases, it is convenient to reflect them when generating the random numbers $\Delta P_{{\rm rand\_}i}$ for each parameter [4].

Now, let us add the distance term of (1). Let us assume we have the layout of our circuit and we know the central coordinates of each transistor i in the layout (x_i, y_i). Let us assume also that the manufacturer provides parameter S_P of the distance term for parameter P in (1). Let us assume also that for each fabricated die, we can approximate the gradient of P along the die by a plane $P(x, y) = Ax + By + C$, where $C = P_{\rm mean} + \Delta P_{\rm global}$, and A and B are random numbers. Consider now two transistors i and j located at coordinates (x_i, y_i) and (x_j, y_j) (see Fig. 1). The mismatch in property P caused by the gradient plane of the die is $\Delta P_{{\rm grad\_}ij} = A(x_i - x_j) + B(y_i - y_j)$. Repeating this for many dies by generating random numbers A and B for each die, we can compute $\sigma^2(\Delta P_{{\rm grad\_}ij}) = \sigma^2(A)(x_i - x_j)^2 + \sigma^2(B)(y_i - y_j)^2$. Assuming symmetry of the random planes $\sigma(A) = \sigma(B)$ (no preferred directions) results in

$$\sigma^2(\Delta P_{{\rm grad\_}ij}) = \sigma^2(A)\left[(x_i - x_j)^2 + (y_i - y_j)^2\right] = \sigma^2(A) D_{ij}^2 \tag{2}$$

Manuscript received September 13, 2006; revised December 5, 2006. This paper was recommended by Associate Editor H. E. Graeb.

The authors are with the Instituto de Microelectrónica de Sevilla, Centro Nacional de Microelectrónica, Consejo Superior de Investigaciones Científicas and Universidad de Sevilla, 41012 Seville, Spain (e-mail: bernabe@imse.cnm.es).

Digital Object Identifier 10.1109/TCAD.2007.893546

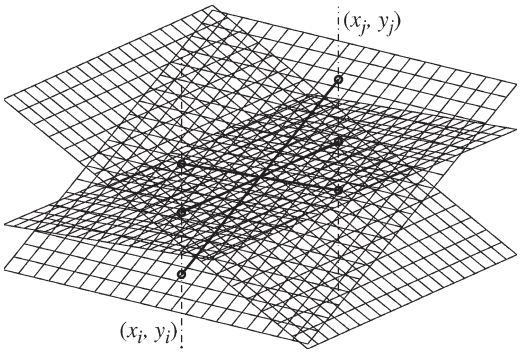

Fig. 1. Illustration of random gradient planes for transistor property P that gives rise to a distance-dependent mismatch term ΔP between the two transistors at coordinates (x_i, y_i) and (x_j, y_j).

where D_{ij} is the distance between transistors i and j. Comparing (2) with the distance term in (1) reveals that

$$S_P = \sigma(A) = \sigma(B). \tag{3}$$

Consequently, if the manufacturer provides parameter S_P, we can generate the random gradient planes for each simulated die.

Summarizing, according to the 1989 findings of Pelgrom [1], a CAD tool implementing all these mismatch components of a MOS transistor should compute for each transistor i and for each of its mismatch relevant parameters P_i a deviation including the following terms: $P_i = P_{\text{mean}} + \Delta P_{\text{global}} + \Delta P_{\text{rand\_}i} + \Delta P_{\text{grad\_}i}$. To compute $\Delta P_{\text{rand\_}i}$, a random number needs to be generated for each transistor i. However, to compute $\Delta P_{\text{grad\_}i}$, only two random numbers need to be computed for all NMOS transistors (and another two for all PMOS) in the same circuit, which are parameters A and B, characterized by (3). With these two random numbers, and the central coordinate of each transistor i, the term $\Delta P_{\text{grad\_}i}$ would be given by

$$\Delta P_{\text{grad\_}i} = Ax_i + By_i. \tag{4}$$

Note that this way of interpreting Pelgrom's distance term cancels out gradient effects when using common-centroid layout techniques. Assume a differential pair where each transistor is split into two. Assume that transistors 1 and 4 in parallel form the first equivalent transistor of the differential pair, and transistors 2 and 3 in parallel form the second one. If they are laid out in a common-centroid configuration, their respective (central) coordinates can be written in the form $(x_1, y_1) = (-x_d, y_d)$, $(x_2, y_2) = (x_d, y_d)$, $(x_3, y_3) = (-x_d, -y_d)$, and $(x_4, y_4) = (x_d, -y_d)$. Applying (4) to this differential pair will result in a gradient-induced mismatch of $\Delta P_{\text{grad\_14–23}} = (\Delta P_{\text{grad\_1}} + \Delta P_{\text{grad\_4}}) - (\Delta P_{\text{grad\_2}} + \Delta P_{\text{grad\_3}}) = 0$, which is consistent with Pelgrom's distance term prediction for common-centroid layouts [1] [see (19) in Appendix A].

III. Nonphysical Interpretation

Shortly after Pelgrom's seminal paper was published, Michael *et al.* developed a means to implement the distance term in (1) [9]–[11], which is called sigma-space analysis or design. This methodology has been further developed [12]. However, here, the distance term is not considered to model the statistics of the possible gradient planes from die to die. It is considered that within the same die, there is another random component per transistor ΔP_{d_i}, such that when computing the difference between two transistors separated by a distance D_{ij}, the standard deviation of this difference obeys

$$\sigma^2_{Dij} = \sigma^2(\Delta P_{d\_i} - \Delta P_{d\_j}) = S_P^2 D_{ij}^2. \tag{5}$$

In principle, there could be many ways to generate random numbers for an arbitrary number of transistors in a circuit satisfying (5). The way proposed in [9] would be as follows. Consider that there are N transistors. Then, for each of them, we compute

$$\begin{aligned}
\Delta P_{d_1} &= 0 \\
\Delta P_{d_2} &= A_{22} R_2 \\
\Delta P_{d_3} &= A_{32} R_2 + A_{33} R_3 \\
&\cdots \\
\Delta P_{d_N} &= \sum_{i=2}^{N} A_{Ni} R_i
\end{aligned} \tag{6}$$

where R_i is a random number normally distributed with zero mean and standard deviation equal to 1. Coefficients A_{ji} are computed by using (5) for each transistor pair. Since there are a total of $N(N-1)/2$ transistor pairs, there is a total of $N(N-1)/2$ nonlinear quadratic equations to solve with a total of $N(N-1)/2$ parameters A_{ji} in (6). If (5) were linear, there would be one unique solution. However, since they are nonlinear, we should expect many possible solutions. The computational cost of solving these equations grows exponentially with the number of transistors.

Obviously, the random gradient plane solution discussed in Section II [see (4)] should be one of the solutions of the formulation of (6), since both interpretations satisfy the distance statistics of (5). However, the computational cost of (3) and (4) is much smaller than for (5) and (6), and especially when there are a large number of transistors in the circuit. Surprisingly, a more detailed analysis of the solutions of (6) (see Appendix B) reveals that the gradient plane of Section II is the only possible solution.

IV. Further Refinements for Implementing Gradient-Induced Mismatch in a CAD Tool

So far, we have seen that implementing Pelgrom's distance-dependent mismatch through a single random gradient plane or through the formulation of (6) is equivalent, although the latter is much more expensive computationally. Consequently, both formulations model Pelgrom's distance term by gradient planes. This is correct as long as the gradients on the wafers have spatial constants much larger than die sizes. However, this assumption cannot always be taken as true. For example, in Fig. 2, we show measurements of 50 NMOS transistors of size 5×5 μm^2 spaced evenly over a distance of 2.6 mm fabricated in a CMOS 0.35-μm process. We can very clearly see the random mismatch component from transistor to transistor [which is the one modeled by the size-dependent term in Pelgrom's model of (1)] and the gradient component. The central smooth curve in Fig. 2 was computed by fitting the experimental data to a third-order polynomial. Note that the gradient between transistors 20 to 50 fits very well a linear plane. Also, transistors 1 to 20 could eventually be fitted by a linear plane as well. However, all 50 transistors do not fit very well into a single gradient plane. This means that for this technology, dies over 2 mm (approximately) may have gradients that are not fitted well by planes. Dies of 2-mm size are usually small dies. In modern designs, it is very common to fabricate large dies of up to and over 1 cm^2. However, it is also true that the analog part (and therefore, more mismatch sensitive part of the die) is usually a small fraction. Consequently, hopefully, this analog part could be modeled using gradient planes. Nonetheless, in modern submicrometer technologies, where mismatch effects are very strong, and digital circuits may

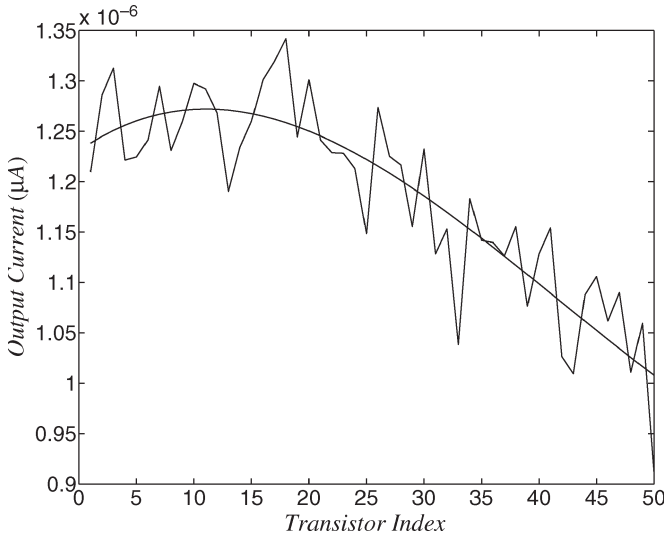

Fig. 2. Measured dc current of 50 identical current sources over a distance of 2.6 mm in a CMOS 0.35-μm technology.

need to consider its effect, it might not be a good approximation to consider a single gradient plane for the whole die. This would result in too pessimistic gradient-induced mismatch predictions. In this case, one might consider tiling the die into several gradient planes, with continuous transitions from plane to plane. Another option, if one has available wafer maps from the foundry, would be to place each die in random positions within the wafer. However, foundries rarely provide such data.

Another interesting observation regarding gradient mismatch modeling is the following. Looking carefully at the steps of Pelgrom's model theoretical derivation (Appendix A), at some point [between (16) and (17)], it is assumed that the spatial constant of the gradients (D_w) is much larger than the transistor sizes and the intertransistor distances. This is true if D_w is of the order of wafer size. However, observations like the one in Fig. 2 reveal that this spatial constant could be much less (at least in the order of a few millimeters). This has two consequences. The first one is that if one estimates constant S_P in (1) from its theoretical expression [see (18)], one should not use the wafer diameter for parameter D_w but rather a length in the order of a few millimeters. The second consequence is that for transistor sizes or intertransistor distances approaching D_w, one can no longer assume the approximations after (16) in Pelgrom's model theoretical derivation. This implies that we may no longer expect to decouple the random mismatch component from the gradient component into two independent terms, as in (1).

V. Conclusion

We have shown that the distance-dependent term $(S_P^2 D^2)$ in Pelgrom's mismatch model is equivalent, to consider for each die a random gradient plane $Ax + By$, such that $\sigma(A) = \sigma(B) = S_P$. We have shown that the solution of the formulation known as sigma-space analysis for predicting this mismatch term in a CAD tool is precisely this random plane. However, the mathematical formulation used is extremely complex and grows exponentially with the number of transistors, resulting in a nonviable solution for moderate and large circuits. However, the random-plane physical interpretation of this mismatch component results in a very simple mathematical formulation, very easy to implement in a CAD tool, and without almost any computing penalty. Hints on refining gradient-induced mismatch in a CAD tool are also given.

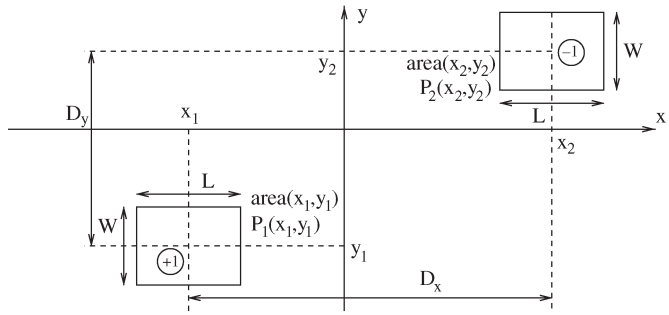

Fig. 3. Position and coordinates of two transistors.

Appendix A
Pelgrom Model Derivation

Fig. 3 shows two transistors of size $W \times L$ located at coordinates (x_1, y_1) and (x_2, y_2), respectively. Let us define the position of the pair as its middle point $x_{12} = (x_1 + x_2)/2$, $y_{12} = (y_1 + y_2)/2$. Let us assume that property P of a transistor can be obtained by averaging a certain density function $\mathtt{P}$ over its area

$$P_1(x_1, y_1) = \frac{1}{WL} \int \int_{\text{area}(x_1, y_1)} \mathtt{P}(x', y') dx' dy'. \tag{7}$$

The density function $\mathtt{P}(x', y')$ is assumed to reflect wafer gradients as well as pure random components. Under these assumptions, the mismatch in property P for the transistor pair located at (x_{12}, y_{12}) is given by

$$\begin{aligned}\Delta P(x_{12}, y_{12}) &= P_1(x_1, y_1) - P_2(x_2, y_2) \\ &= \frac{1}{WL} \int \int_{\text{area}(x_1, y_1)} \mathtt{P}(x', y') dx' dy' \\ &\quad - \frac{1}{WL} \int \int_{\text{area}(x_2, y_2)} \mathtt{P}(x', y') dx' dy' \\ &= \frac{1}{WL} \int \int_{\Re^2} G(x_{12} - x', y_{12} - y') \mathtt{P}(x', y') dx' dy'\end{aligned} \tag{8}$$

where G is "1" when (x, y) is inside the area of transistor centered at (x_1, y_1), G is "-1" for the area of transistor centered at (x_2, y_2), and G is "0" elsewhere. In general, G is a geometry function which depends on the specific layout of the transistor pair. Taking the Fourier transform in (8) yields $\Delta P(\omega_x, \omega_y) = (1/(WL)) G(\omega_x, \omega_y) P(\omega_x, \omega_y)$, where $\Delta P(\omega_x, \omega_y)$ is the Fourier Transform of $\Delta P(x_{12}, y_{12})$, $G(\omega_x, \omega_y)$ is the one of $G(x, y)$, and $P(\omega_x, \omega_y)$ is the one of $\mathtt{P}(x, y)$.

For the layout of Fig. 3, it can be shown that

$$G(\omega_x, \omega_y) = \frac{\sin\left(\frac{\omega_x L}{2}\right)}{\frac{\omega_x}{2}} \frac{\sin\left(\frac{\omega_y W}{2}\right)}{\frac{\omega_y}{2}} (-2j) \sin\left(\frac{\omega_x D_x + \omega_y D_y}{2}\right). \tag{9}$$

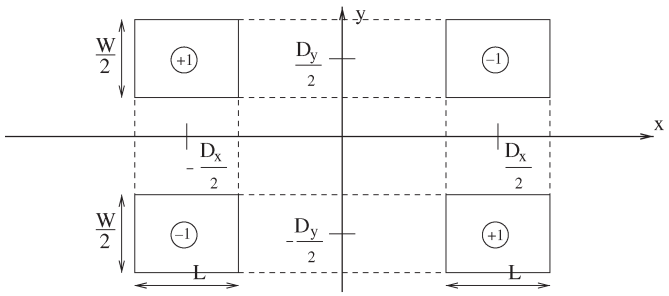

Fig. 4. Layout configuration for a transistor pair using common centroid.

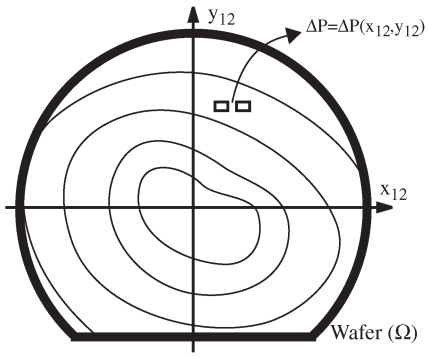

Fig. 5. Wafer gradients.

If $D_y = 0$, it follows that

$$\left|\frac{1}{WL}G(\omega_x,\omega_y)\right| = \frac{\sin\left(\frac{\omega_x L}{2}\right)}{\frac{\omega_x L}{2}}\frac{\sin\left(\frac{\omega_y W}{2}\right)}{\frac{\omega_y W}{2}}\left\{2\sin\left(\frac{\omega_x D_x}{2}\right)\right\}. \tag{10}$$

For a pair of transistors in a common-centroid configuration, as shown in Fig. 4, it would be

$$\left|\frac{1}{WL}G(\omega_x,\omega_y)\right| = \frac{\sin\left(\frac{\omega_x L}{2}\right)}{\frac{\omega_x L}{2}}\frac{\sin\left(\frac{\omega_y W}{4}\right)}{\frac{\omega_y W}{4}}\left\{2\sin\left(\frac{\omega_x D_x}{2}\right)\sin\left(\frac{\omega_y D_y}{2}\right)\right\}. \tag{11}$$

Fig. 5 shows a wafer in which contour lines of constant property P have been drawn. In the wafer, at coordinate (x_{12}, y_{12}), a pair of transistors is drawn.

Assuming that when averaging $\Delta P(x_{12}, y_{12})$ all over the wafer we have $\overline{\Delta P}|_{\text{Wafer}} \approx 0$, we can write that

$$\sigma^2(\Delta P) = \frac{1}{\Omega}\int\int_{\Omega} \Delta P^2(x_{12}, y_{12})dx_{12}dy_{12} \tag{12}$$

where Ω is the area of the wafer. Applying Poisson's theorem to (12) results in

$$\sigma^2(\Delta P) = \frac{1}{4\pi^2\Omega}\int_{-\infty}^{\infty} d\omega_x \int_{-\infty}^{-\infty} d\omega_y \left|\frac{1}{WL}G(\omega_x,\omega_y)P(\omega_x,\omega_y)\right|^2. \tag{13}$$

Let us now make the following assumption: $P(\omega_x, \omega_y) = P_o + \mathcal{W}(\omega_x, \omega_y)$, where P_o is a constant (frequency independent) representative of white noise and $\mathcal{W}(\omega_x, \omega_y)$ is a wafer map component

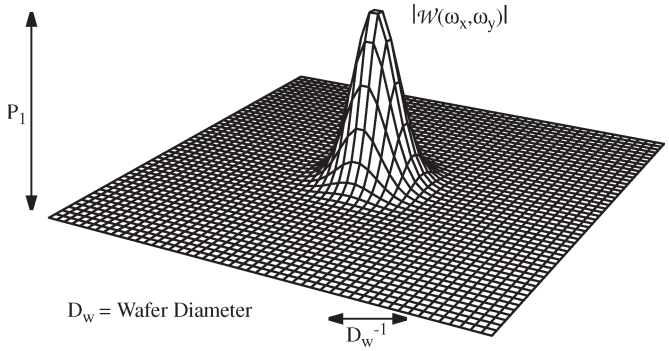

Fig. 6. Approximate shape of frequency domain function $\mathcal{W}()$.

responsible for long-distance gradients along the wafer. The spatial frequency content of function $\mathcal{W}(\omega_x, \omega_y)$ is for frequencies of the order of D_w^{-1}, where D_w is the wafer diameter. Therefore, function $\mathcal{W}(\omega_x, \omega_y)$ can be assumed to have a shape of the type shown in Fig. 6, and consequently, we can assume that

$$\mathcal{W}(\omega_x,\omega_y) = \begin{cases} \backsim P_1, & \text{if } \left(\frac{-1}{D_w} \le \omega_x \le \frac{1}{D_w}\right), \left(\frac{-1}{D_w} \le \omega_y \le \frac{1}{D_w}\right) \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

Therefore, (13) can be written as

$$\sigma^2(\Delta P) = \frac{1}{4\pi^2\Omega W^2L^2}\int_{-\infty}^{\infty} d\omega_x \int_{-\infty}^{\infty} d\omega_y |G|^2|P_0+\mathcal{W}|^2 = \frac{1}{4\pi^2\Omega W^2L^2}\left\{|P_0|^2Y_1 + Y_2\right\}. \tag{15}$$

Assuming a transistor pair as in Fig. 3, it would be

$$Y_1 = \int_{-\infty}^{\infty} d\omega_x d\omega_y |G|^2 = 8\pi^2 WL$$

$$Y_2 = \int_{-\infty}^{\infty} d\omega_x \int_{-\infty}^{\infty} d\omega_y |G|^2\left[P_0^*\mathcal{W} + P_0\mathcal{W}^* + |\mathcal{W}|^2\right]. \tag{16}$$

Since $D_w \gg D_x, W, L$, then $|G(\omega_x, \omega_y)| \approx \omega_x D_x LW$. Thus

$$Y_2 \approx \int_{-\frac{1}{D_w}}^{\frac{1}{D_w}} d\omega_x \int_{-\frac{1}{D_w}}^{\frac{1}{D_w}} d\omega_y \omega_x^2 D_x^2 L^2 W^2\left[P_0^*\mathcal{W} + P_0\mathcal{W}^* + |\mathcal{W}|^2\right] = D_x^2L^2W^2k_0'$$

$$k_0' = \int_{-\frac{1}{D_w}}^{\frac{1}{D_w}} d\omega_x \int_{-\frac{1}{D_w}}^{\frac{1}{D_w}} d\omega_y \omega_x^2\left[P_0^*\mathcal{W} + P_0\mathcal{W}^* + |\mathcal{W}|^2\right] \tag{17}$$

where $P_o^*\mathcal{W} + P_o\mathcal{W}^* + |\mathcal{W}|^2 \approx P_o^*P_1 + P_oP_1^* + |P_1|^2 = k_o$, and $Y_2 = (4k_oD_x^2L^2W^2)/(3D_w^4)$. This results in

$$\sigma^2(\Delta P) = \frac{2|P_0|^2}{\Omega WL} + \frac{k_0D_x^2}{3\pi^2\Omega D_w^4} = \frac{A_p^2}{WL} + S_p^2D_x^2,$$
$$A_p^2 = \frac{2|P_0|^2}{\Omega}, \quad S_p^2 = \frac{k_0}{3\pi^2\Omega D_w^4}. \tag{18}$$

For a common-centroid configuration, by changing the geometry function, it can be shown that the result is

$$\sigma^2(\Delta P) = \frac{2|P_o|^2}{\Omega WL} + \frac{k_o D_x^2 D_y^2}{36\pi^2 \Omega D_w^6} = \frac{A_P^2}{WL} + S_P^2 \frac{D_x^2 D_y^2}{D_w^2}. \quad (19)$$

Note that the distance term has been reduced an amount of the order $(D_y/D_w)^2$, which is very small. Consequently, from a practical point of view, the distance term can be considered to have disappeared. Perfect gradient planes would cancel out exactly this distance term. However, for the model in Fig. 6, the gradients are not always perfect planes and might have a higher order curvature component. This is why in (19), there results a residual distance term.

APPENDIX B

Let us demonstrate that the solution to the set of (5) and (6) is exactly the random gradient plane described Section II. First, we will compute the solution for a three-transistor circuit. Then, we will show that if it is verified for a circuit with $N-1$ transistors, it will be verified also by a circuit with N transistors. Therefore, it is true for any number of transistors.

Let us start with a circuit of three transistors and consider a generic mismatch sensitive parameter P. According to the formulation of (6)

$$\Delta P_{d\_1} = 0 \quad \Delta P_{d\_2} = A_{22}R_2 \quad \Delta P_{d\_3} = A_{32}R_2 + A_{33}R_3. \quad (20)$$

Let us choose, without any loss of generality, a coordinate system such that the coordinates of the three transistors are (0, 0), $(x_2, 0)$, and (x_3, y_3), respectively. Applying (5) and (20) to these three transistors results in

$$\begin{aligned}
\sigma_{12}^2 &= S_p^2 x_2^2 = \sigma^2(\Delta P_{d_2} - \Delta P_{d_1}) = A_{22}^2 \\
\sigma_{13}^2 &= S_p^2 \left(x_3^2 + y_3^2\right) = \sigma^2(\Delta P_{d_3} - \Delta P_{d_1}) = A_{32}^2 + A_{33}^2 \\
\sigma_{23}^2 &= S_p^2 \left[(x_3 - x_2)^2 + y_3^2\right] = \sigma^2(\Delta P_{d_3} - \Delta P_{d_2}) \\
&= (A_{32} - A_{22})^2 + A_{33}^2. \quad (21)
\end{aligned}$$

The solutions of this set of equations are

$$A_{22} = s_1 S_p x_2 \quad A_{32} = s_1 S_p x_3 \quad A_{33} = s_2 S_p y_3 \quad (22)$$

where $s_1 = \pm 1$ and $s_2 = \pm 1$. Consequently, there are four possible solutions, depending on the signs s_1 and s_2 (parameter S_p is supposed to be always positive). Substituting (22) into (20) results in

$$\Delta P_{d\_1} = 0 \quad \Delta P_{d\_2} = Ax_2 \quad \Delta P_{d\_3} = Ax_3 + By_3 \quad (23)$$

with $A = s_1 S_p R_2$ and $B = s_2 S_p R_3$. Note that (23) defines all ΔP_{d_i} on a random plane defined by the two random parameters A and B, such that $\sigma(A) = \sigma(B) = S_p$. Consequently, the distance mismatch solution of (5) and (6) lies on a random plane for all transistors.

If we now add a fourth transistor at coordinates (x_4, y_4), with $\Delta P_{d\_4} = A_{42}R_2 + A_{43}R_3 + A_{44}R_4$, we obtain the same solutions than in (22) plus $A_{42} = s_1 S_p x_4$, $A_{43} = s_2 S_p y_4$, and $A_{44} = 0$. Consequently, $\Delta P_{d\_1} = 0$, $\Delta P_{d\_2} = Ax_2$, $\Delta P_{d\_3} = Ax_3 + By_3$, $\Delta P_{d\_4} = Ax_4 + By_4$, and again, the distance mismatch solution of (5) and (6) lies on a random plane for all transistors. Let us now consider a circuit with $N-1$ transistors, and let us assume that the solution also lies on a random plane. This means that

$$A_{n2} = s_1 S_p x_n \quad A_{n3} = s_2 S_p y_n \quad A_{nk} = 0, \qquad k = 4, \ldots, n-1 \quad (24)$$

for $n = 4, \ldots, N-1$. If we now consider the general case of a circuit with N transistors, each at coordinate (x_n, y_n), then following the formulation of (6), we have for the last transistor $\Delta P_{d\_N} = \sum_{j=2}^{N} (A_{Nj} R_j)$. Following the formulation of (21) and assuming (24), we have $N-1$ equations for the parameters A_{Nj} of the Nth transistor

$$\begin{aligned}
\sigma_{1N}^2 &= S_p^2 \left(x_N^2 + y_N^2\right) = \sigma^2(\Delta P_{d_N} - \Delta P_{d_1}) = \sum_{j=2}^{N} A_{Nj}^2 \\
\sigma_{2N}^2 &= S_p^2 \left[(x_N - x_2)^2 + y_N^2\right] = \sigma^2(\Delta P_{d_N} - \Delta P_{d_2}) \\
&= (A_{N2} - A_{22})^2 + \sum_{j=3}^{N} A_{Nj}^2 \\
\sigma_{kN}^2 &= S_p^2 \left[(x_N - x_k)^2 + (y_N - y_k)^2\right] = \sigma^2(\Delta P_{d_N} - \Delta P_{d_k}) \\
&= (A_{N2} - A_{k2})^2 + (A_{N3} - A_{k3})^2 + \sum_{j=4}^{N} A_{Nj}^2. \quad (25)
\end{aligned}$$

Note that the last equation in (25) can be written for $k = 3, \ldots, n-1$ because of the assumption in (24). Subtracting the first two equations in (25) and using (24) yields $A_{N2} = s_1 S_p x_N$. Subtracting the third equation from the first two ones results in two equations which when solved using (24) yields $A_{N3} = s_2 S_p y_N$ and $\sum_{j=4}^{N} A_{Nj}^2 = 0$. Consequently, (24) is also satisfied for $n = N$, and again, the distance mismatch solution of (5) and (6) lies on a random plane for all transistors.

REFERENCES

[1] M. Pelgrom, A. Duinmaijer, and A. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, Oct. 1989.

[2] J. Bastos, "Characterization of MOS transistor mismatch for analog design," Ph.D. dissertation, Katholieke Universiteit Leuven, Leuven, Belgium, 1998.

[3] M. Steyaert, J. Bastos, R. Roovers, P. Kinget, W. Sansen, B. Graindourze, A. Pergoot, and E. Janssens, "Threshold voltage mismatch in short-channel MOS transistors," *Electron. Lett.*, vol. 30, no. 18, pp. 1546–1548, Sep. 1, 1994.

[4] T. Serrano-Gotarredona and B. Linares-Barranco, "Systematic width-and length dependent CMOS transistor mismatch characterization and simulation," *Int. J. Analog Integr. Circuits Signal Process.*, vol. 21, no. 3, pp. 271–296, Dec. 1999.

[5] T. Serrano-Gotarredona and B. Linares-Barranco, "A new five-parameter MOS transistor mismatch model," *IEEE Electron Device Lett.*, vol. 21, no. 1, pp. 37–39, Jan. 2000.

[6] J. Croon, M. Rosmeulen, S. Decoutere, and W. Sansen, "An easy-to-use mismatch model for the MOS transistor," *IEEE J. Solid-State Circuits*, vol. 37, no. 8, pp. 1056–1064, Aug. 2002.

[7] P. G. Drennan and C. C. McAndrew, "Understanding MOSFET mismatch for analog design," *IEEE J. Solid-State Circuits*, vol. 38, no. 3, pp. 450–456, Mar. 2003.

[8] T. Serrano-Gotarredona and B. Linares-Barranco, "CMOS transistor mismatch model valid from weak to strong inversion," in *Proc. ESSCIRC*, Sep. 2003, pp. 627–630.

[9] C. Michael and M. Ismail, "Statistical modeling of device mismatch for analog MOS integrated circuits," *IEEE J. Solid-State Circuits*, vol. 27, no. 2, pp. 154–166, Feb. 1992.

[10] C. Michael and M. Ismail, *Statistical Modeling for Computer-Aided Design of MOS VLSI Circuits*. Boston, MA: Kluwer, 1993.

[11] C. Michael, C. Abel, and M. Ismail, "SMOS: A CAD-compatible statistical model for analog MOS integrated circuit simulation," *Int. J. Circuit Theory Appl.*, vol. 20, no. 3, pp. 327–347, Mar. 1992.

[12] H. Zhang, Y. Zhao, and A. Doboli, "ALAMO: An improved σ-space based methodology for modeling process parameter variations in analog circuits," in *Proc. Conf. Des. Autom. Test Eur.*, 2006, pp. 156–161.