Modeling Small MOSFETs Using Ensemble Devices

Josef Watts, Robinson Pino, Henry Trombley
IBM Microelectronics, Essex Junction VT, USA. jswatts@us.ibm.com

Abstract

When characterizing small area FETs for modeling, it is common practice to measure an ensemble of many devices in parallel and then use the average current as the typical behavior of a single small device. [1] This reduces the sampling error caused by random dopant fluctuation and line edge roughness. However, using this average current introduces a distortion of the drain current versus gate voltage characteristics. Specifically in the subthreshold and low overdrive regions the average current in the ensemble is higher than the typical current, but in high the overdrive region the average current equals the typical current. This paper presents experimental verification of this assertion, explains the statistical origin of the effect, and proposes a compact modeling method for accurate extraction of typical behavior from ensemble measurements and reproducing either single finger or ensemble currents for circuit simulation.

Keywords: mosfet, circuit modeling, statistics

1. Introduction

Consider a small FET composed of two fingers. In the absence of manufacturing variations each finger has the same IV characteristics and the current of the two fingers in parallel is twice the current of one finger. In practice small FET fingers display variation in threshold voltage due to random dopant fluctuations (RDF).[2] Suppose the threshold voltage of one finger in our hypothetical FET is 30 millivolts higher than typical and that of the other finger is 30 millivolts lower. On average the fingers have the typical Vt and so in some sense this is a typical FET subject to RDF. One finger will have more current than typical and the other less. Well above threshold, the differences in current from typical will be roughly proportional to the differences of Vt from typical because the current is roughly linear with gate voltage. In our example, the total current will be approximately twice the typical value for one finger. On the other hand, below threshold, the current is exponentially related to the threshold voltage and the two fingers will not have equal and opposite current deltas. If one finger has twice the typical current and the other will have roughly half the typical current. The total current will be 2.5 times the typical for one finger, not two times as we might expect from a typical device.

If we measure the Vt of the total device by the single point method we will find a lower Vt than the average of the individual finger Vt's. This effect has been observed when trying to estimate the quiescent current for CMOS SRAMs [3] and logic chips [4,5]. These researchers have noted that the distribution of off currents is lognormal because of the logarithmic relationship between off current and both threshold voltage and FET gate length.

This effect must be considered when extracting a compact model from measurements of multiple devices in parallel. If the model is adjusted to match line targets it is important to understand the structures used to establish and monitor the line targets. Ensemble devices will produce higher off current targets and lower Vt targets than single finger devices for the same manufacturing process. Finally since circuit designers use small FETs both in parallel arrangements and as single FETs, the compact model needs to be able to model both cases correctly.

2. Experimental Details

The IV characteristics of single and ensemble n-type FETs were measured at room temperature with 50 mV drain bias. Two device geometries were considered: 0.1μm/0.05μm (W/L) and 0.12μm/0.06μm. The ensemble devices consisted of 20 identical devices wired in parallel. For each geometry a total of 152 die were measured. From the IV data the single point threshold voltage (Vt) and extrapolated Vt were obtained. The single point Vt was measured at a drain current of 300× (W/L) nA. The maximum g_m method was used for the extrapolated Vt.

Figure 1, shows the drain current distribution of single devices, 0.1μm/0.05μm (W/L), at a gate bias of 0.3 V. We see that in the subthreshold region, the drain current distribution is skewed which results in a difference between the mean and the median. Figure 2 shows the distribution of single point Vt values for the same devices, which has a Gaussian distribution.

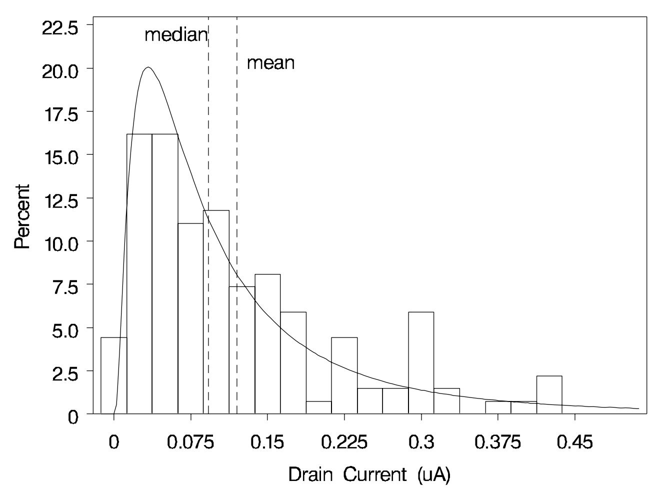

Figure 1. Drain current distribution of single devices at Vg = 0.3 V, 0.1um/0.05um (W/L).

To construct Figure 3, all the single finger measurements for one geometry were pooled and the mean and median of the drain current were calculated for each gate bias. Measuring an ensemble device with many fingers and dividing by the number of fingers is equivalent to measuring many single finger devices and finding the mean value. So the mean value of current at any bias condition represents the current measured in an ensemble device. On the other hand the median current corresponds to a single finger device with the median Vt (Ioff is monotonic in Vt). Since the Vt distribution is Gaussian we can reasonably call this FET the "typical" FET. More importantly for our purposes if we call this FET the typical FET we can model the Vt variation with a Gaussian distribution and correctly reproduce the whole range of off and on currents.

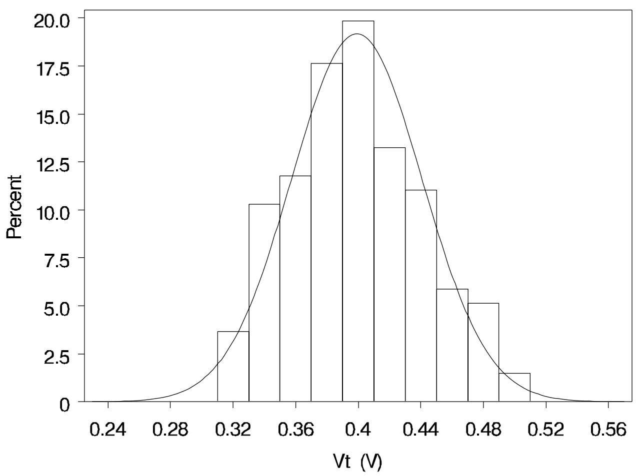

Figure 2. Single point Vt distribution of single devices, 0.1um/0.05um (W/L).

In Figure 3, we observe that a Vt shift exists between the mean current (ensemble device) and the median current (single device). This is expected from the drain current in subthreshold distribution shown in Figure 1 in which the mean and median of the distribution are clearly different. We show in Section 3.1 that the magnitude of the offset is proportional to the second derivative with respect to Vt and this is observed in Figure 4 where the two curves come together as the curvature decreases at high overdrive.

In addition, in Figure 5 we observe degradation in the effective g_m measured from the ensemble devices. The median current displays higher peak g_m since it must catch up with mean current at high overdrive. This difference in g_m explains why the extrapolated Vt which is calculated based on above Vt currents shows an offset between single an ensemble devices.

Table 1 shows the single point and extrapolated Vt, for the single and the 20 parallel devices. These threshold voltages were calculated from the mean and median currents plotted in Figure 3 and the terms mean and median in the table refer to which curve was used, not to statistics for a population of individual Vt measurements. Similar results were observed for the 0.12µm/0.06µm (W/L) single and 20 parallel devices.

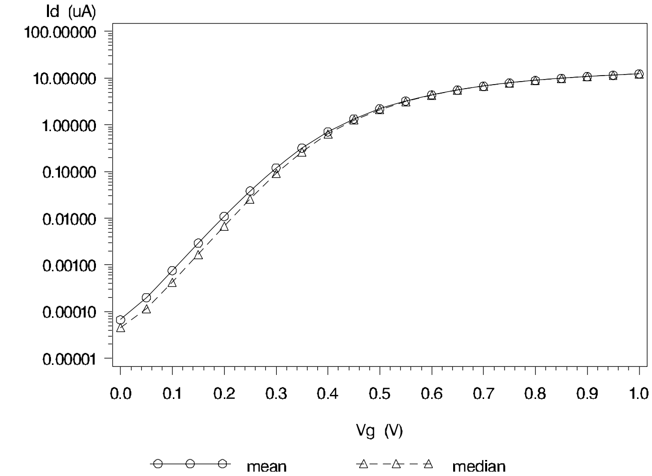

Figure 3. Drain current mean (ensemble) and median (single) on logarithmic scale.

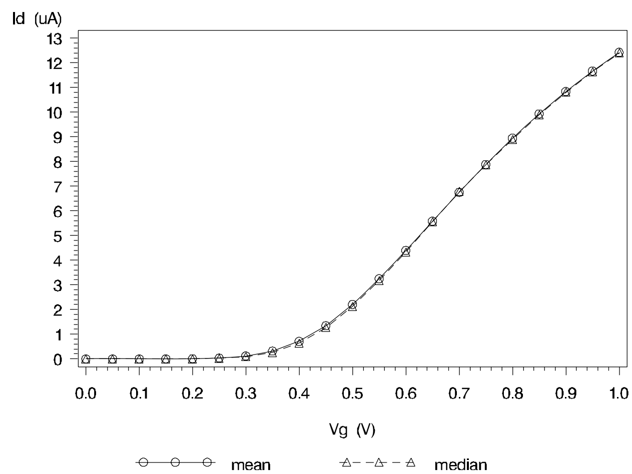

Figure 4. Drain current mean (ensemble) and median (single) on logarithmic scale.

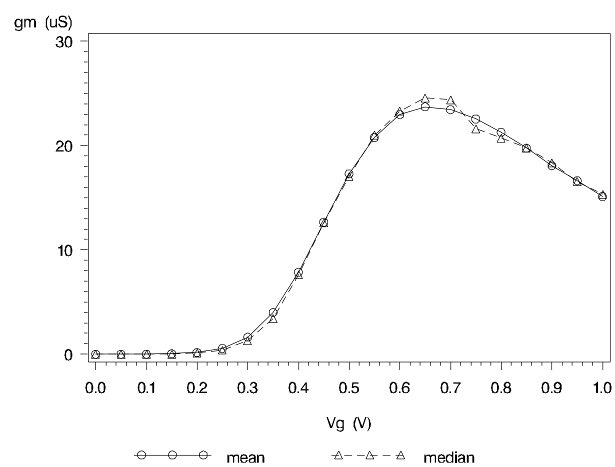

Figure 5. Transconductance calculated from mean (ensemble) and median (single) current.

Table 1. Threshold voltage (Vt) and Ion/Ioff results.

	1 Device: 0.10/0.05 (W/L)	20 Device: 0.10/0.05 (W/L)
Single Point Vt Mean (mV)	389.18	380.72
Single Point Vt Median (mV)	396.09	381.74
Delta (mV)	**6.91**	**1.02**
Extrapolated Vt Mean (mV)	414.52	395.42
Extrapolated Vt Median (mV)	423.45	400.47
Delta (mV)	**8.93**	**5.05**
Mean Ion/Ioff	1.86×10^5	1.34×10^5
Median Ion/Ioff	2.73×10^5	1.68×10^5
Ratio (median/mean)	**1.47**	**1.25**

For the results from 20 parallel devices shown in Table 1, it is important to point out the reduced Vt shift observed as compared to the single devices. There is minimal Vt shift observed between the mean and median within the 20 parallel devices data set because the output current in these ensemble devices is already an averaged current. Thus, the median output corresponds to an average current between 20 devices which minimizes the statistical effect. Also, Table 1 displays the Ion/Ioff ratio where the median Ion/Ioff ration is larger than the mean Ion/Ioff ratio since both have the same on current but the median ratio has a lower Ioff current. In addition, the Ion/Ioff mean and median ratios for the 20 parallel devices appear closer to each other since these currents are coming from ensembles where the effect is reduced. Thus, when performing FET modeling utilizing ensemble devices, unknowingly, we are modeling a higher current in the subthreshold and weak inversion regions than is real for a typical single device.

3. Modeling Procedure

The statistical distortion of the measured IV characteristics of ensembles of devices needs to be accounted for in model extraction and circuit simulation. Our method, which is shown in Figure 6 does not require any new measurements, however, it does require new extraction steps and extra time independent calculations during circuit simulation. DC measurements of geometric scaling FETs, including ensembles of small area FETs, and model extraction are done as usual, as is measurement of match FET pairs and extraction of a random dopant fluctuation (RDF) model.

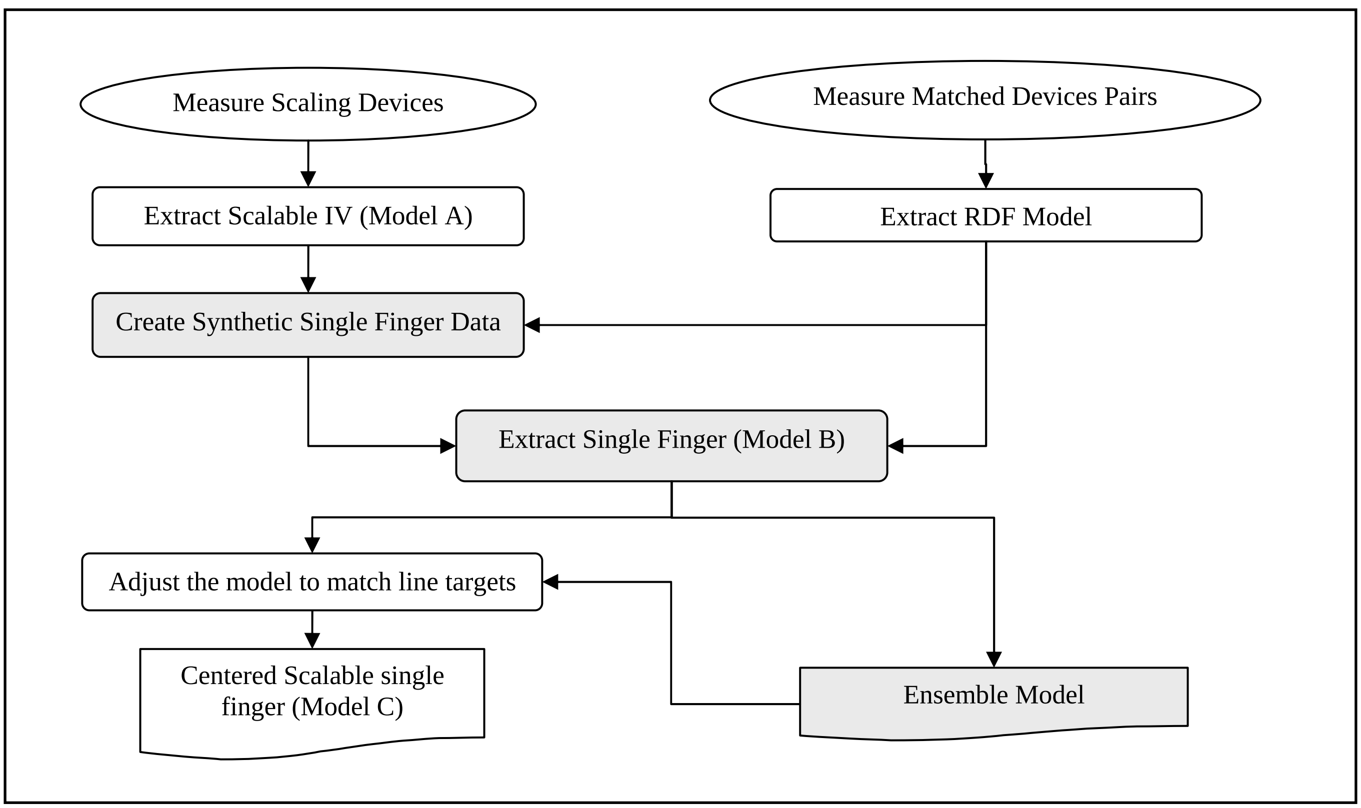

Figure 6. Modeling procedure which accounts for ensemble device distortion

The extracted model (model A) which represents the measured ensemble devices is used to create a synthetic dataset. This data is adjusted to estimate single device typical data using the RDF model and the procedure in Section 3.1.

The DC model is fit to the single device data by adjusting a few parameters as discussed in Section 3.2 creating model B. The parameter differences between models A and B are used to create the ensemble model. When the model is run, the ensemble model examines the instance parameters and adjusts the model card parameters for the actual number of parallel fingers or devices. Model B is adjusted to match line targets, using the ensemble model to account for the actual number of fingers in the FETs used for line monitoring. The resulting model (model C) represents single finger FETs but in combination with the ensemble model can simulate FETs with any number of fingers.

3.1 Recovering Single Finger Data

Measurement of an ensemble device is equivalent to estimating the mean current at each applied voltage of a finite sample of single fingers. If we assume the fingers are identical except for threshold voltage and that the threshold voltages are independent and normally distributed, we can recover the current of a typical single finger. A typical finger is one with the typical or average Vt. The same procedure applies to both measured and simulated data. We begin by expressing the current of any finger as a function of (Vg-V_t+Δ_i,) and expanding in a Taylor series. (Δ_i the difference between the Vt of this finger and the mean Vt and we note that the derivative with respect to Δ_I *is the same as with respect to* V_g).

$$I_i = I(V_g - V_t + \Delta_i)$$

$$I_i = I_0 + \frac{dI}{dV_g}\Delta_i + \frac{1}{2}\frac{d^2I}{dV_g^2}\Delta_i^2 + \ldots \tag{1}$$

To calculate the total current of an ensemble, we sum over all fingers and use the properties of the normal distribution to evaluate the sum over powers of Δ.

$$I_e = \frac{1}{n}\sum_i I_i = I_0 + \frac{dI}{dV_g}\frac{1}{n}\sum_i \Delta_i + \frac{1}{2}\frac{d^2I}{dV_g^2}\frac{1}{n}\sum_i \Delta_i^2 + \ldots$$

$$I_e = \sum_i I_i = I_0 + 0 + \frac{1}{2}\frac{d^2I}{dV_g^2}\sigma_{Vt}^2 + \ldots \tag{2}$$

In this expression σ_{Vt}^2 is the variance of threshold voltage between identical FETs close to one another, exactly what is typically measured for RDF modeling. I_e is the current we have measured; in order to solve for the typical single finger current we must make one more assumption. We assume that the ratio of single finger and ensemble currents and the ratios of their derivatives of any order are equal to a single number we call *x*. In the subthreshold region where this effect is most important, the current is approximately exponential in V_g and this is a good assumption. Above threshold the difference in currents is small, making the value of the ratios of currents and derivatives all close to one. With this assumption we can rewrite (2) and solve for *x.* We have shown only the first two terms of the expansion for clarity but we find in practice that terms up through the sixth power may be needed. We show them in the final result only.

$$I_e = xI_e + 0 + x\frac{1}{2}\frac{d^2I_e}{dV_t^2}\sigma_{Vt}^2 + \ldots$$

Solving for *x:*

$$x = \frac{48}{48 + 24\beta\sigma_{Vt}^2 + 6\beta^2\sigma_{Vt}^4 + \beta^3\sigma_{Vt}^6}$$

Where we use the properties of the exponential function to estimate the fourth and sixth derivatives as:

$$\frac{d^4I}{dV_t^4} = \beta\frac{d^2I}{dV_t^2}\ ;\ \frac{d^6I}{dV_t^6} = \beta^2\frac{d^2I}{dV_t^2} \text{ where } \beta = \frac{\frac{d^2I}{dV_t^2}}{I}$$

Above threshold the higher order terms are small so using the above formulas does not introduce significant error.

3.2 Creating the Ensemble Model

To invoke a device model that is dependent on the number of parallel ensemble devices requires that we create a parameterized model that we can skew in a continuous mode from the typical single device model to the typical ensemble model with infinite devices in parallel. We create this model by first extracting the model parameters for the large ensemble of devices. We then use this model as a starting point to fit the single typical device data set we calculated in Section 3.1. For practical reasons, we want to skew as few model parameters as possible. We have chosen three parameters that have strong influence on the model behavior in the subthreshold region.

Figures 7 and 8 show the model fit to the ensemble device data and the skewed model fit to the single device typical data. Figure 7 has a linear scale for current to highlight the fact that there is minimal change in the typical current between the single device and large ensemble of devices in this region. Figure 8 shows the same curves with a log scale for current. Here it is clearly seen that there is a shift between the single device typical current and the typical current of a large ensemble of devices and that both sets of data can be accurately fit with models that differ only by a few parameters.

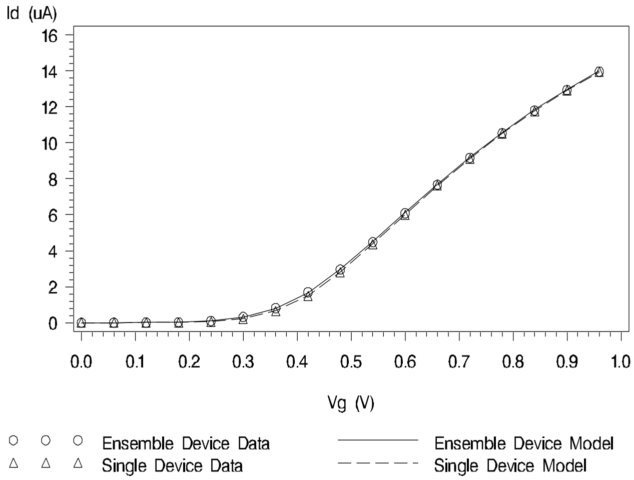

Figure 7. Typical single device and typical large ensemble IdVg data and fit model. Minimal difference is seen between the curves at Vg higher than the threshold voltage.

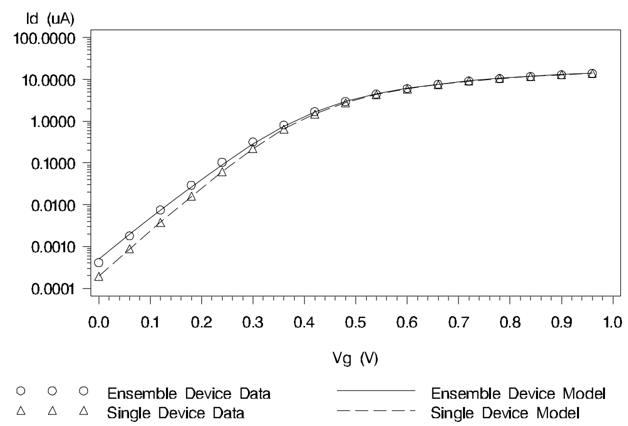

Figure 8. Typical single device and typical large ensemble IdVg data and fit model. BSIM4 model parameters voff, nfactor, and vth0 have been adjusted.

To skew between the typical single device model and the typical ensemble model, we use an empirically expression:

$$s = 1 - \left(\frac{2}{n} - \frac{1}{n^2} \right) \quad (3)$$

where, *n* the number of devices in the ensemble. Model parameters are now modified on an instance basis using

$$Px_n = Px_1 + s\left(Px_e - Px_1\right)$$

where Px_e is the model parameter extracted using the ensemble data set and adjusted for $n \rightarrow \infty$ and Px_1 is the model parameter extracted using the synthetic single device data set.

3.3 Ensemble model in use

Our final model implementation works as follows. We will keep a global model card which describes a nominal, single finger device. The distinction between ensemble and single device behavior will take place is applied by the ensemble model on an instance level when a netlist is simulated. Then, the model card will be invoked, and global process variation is applied based on options selected by the user. For each FET instance in the netlist additional adjustments are made for local process variation and layout dependent effects. The ensemble model is one of the time independent instance specific adjustments. For example, in BSIM4, some model parameters which we have identified and that can be adjusted for ensemble effects are VOFF, NFACTOR, and VTH0.

The parameter n in (3) is determined from the instance parameters and the type of analysis. In some case the netlist specifies values for nf (number of fingers) and/or m (number of parallel FET represented by this one instance) and these are used to calculate n.

In some cases multi-fingered devices are netlisted as a separate instance for each finger because of layout effects which are not identical for all fingers. In this case an additional instance parameter is added to specify the number of parallel instances. This parameter is used as n unless a Monte Carlo analysis is being run. If a Monte Carlo analysis is run each finger instance will get a unique Vt offset and the sum of the instance currents will automatically model the ensemble effect. Even in Monte Carlo the ensemble model is needed for FET instance with nf or m greater than 1 because only one Vt offset is used for all fingers.

Conclusions

We have shown that measuring ensemble devices gives a different typical result than measuring single finger devices. We have shown a comprehensive methodology for measuring and modeling devices with any number of fingers from data measurement through model extraction to generation of instance specific model cards based on the number of parallel instances in the netlist.

References

[1] H. Lee, et al, IEEE ED Letters, Vol. 23, No. 5, May 2002

[2] J. T. Watt and J. D. Plummer, IEEE Trans. on ED, Vol. 35, No. 12

[3] T. B. Hook, et al, IEEE Trans. on ED, Vol. 49, No. 8, Aug. 2002

[4] R. Rao, et al, IEEE Trans. on VLSI, Vol. 12, No. 2, Feb. 2004

[5] I. A. Ferzli and F. N. Najam, IEEE Trans. on CAD of IC, V25, I1, Jan. 2006