A Linear Statistical Analysis for Full-Chip Leakage Power with Spatial Correlation *

Ruijing Shen
Dept. of Electrical Engineering
University of California
Riverside, CA, USA
rshen@ee.ucr.edu

Sheldon X.-D. Tan
Dept. of Electrical Engineering
University of California
Riverside, CA, USA
stan@ee.ucr.edu

Jinjun Xiong
IBM Thomas J. Watson
Research Center
Yorktown Heights, NY, USA
jinjun@us.ibm.com

ABSTRACT

In this paper, we present an approved linear-time algorithm for statistical leakage analysis in the present of any spatial correlation condition (strong or weak). The new algorithm adopts a new set of uncorrelated variables over virtual grids to represent the original physical random variables and the grid size (thus of number of new random variables) is determined by the spatial correlation length. In this way, each physical variable is always represented by virtual variables locally. We prove that the number of neighboring virtual grids for each grid is not related to condition of spatial correlation, which leads to linear time complexity in terms of number of gates. We compute the gate leakage by the orthogonal polynomials based collocation method. The total leakage of a whole chip can be computed by simply summing up the coefficients of corresponding orthogonal polynomials for each grid. Furthermore, look-up table can be created to cache statistical information for each **type** of gates in library instead of calculating leakage for every single gate on chip. As a result, we end up with $O(N)$ time complexity, where N is the number of grids on chip. The proposed method has no restrictions on static leakage models, types of statistical distributions for leakage currents. Experimental results show that the proposed method is about 1000X faster than the recently proposed grid-based method [2] with similar accuracy and many orders of magnitude times over the Monte Carlo method.

Categories and Subject Descriptors

I.6.5 [**Simulation and Modeling**]: Model Development; B.7.2 [**Integrated Circuits**]: Design Aids—*simulation*

General Terms

Algorithms

*This work is funded in part by NSF grant under No. CCF-0448534, in part by NSF grant under No. OISE-0929699 and in part by National Natural Science Foundation of China (NSFC) grant under No. 60828008

GLSVLSI'10, May 16–18, 2010, Providence, Rhode Island, USA.

1. INTRODUCTION

Leakage current has become a major component of total chip power consumption due to aggressive scaling of devices [1]. The dominant factors in the leakage currents are the subthreshold leakage current I_{sub} and gate oxide leakage current I_{gate}. And both kinds of leakage currents are highly sensitive to process variations due to the exponential relation between the leakage current and process parameters like channel lengths and threshold voltages. As process-induced variability becomes more pronounced in the deep sub-micro regime [8], leakage variations become prominent, and traditional worst case approach will lead to extremely pessimistic and expensive design solutions. Verifying all the possible corners for various process parameters becomes prohibitively time consuming if possible at all. Statistical estimation and analysis of leakage powers considering process variability become necessary in various chip design steps to improve design yield and robustness.

As verifying powers at various design steps is an important issue, many earlier works have been proposed for full-chip statistical leakage analysis (SLA) considering the process variations such as [7,12,15]. When the process-induced variability is spatially correlated, the computational cost of the distribution of total leakage of the chip becomes quadratic – $O(n^2)$, where n is number of gates as we have to consider all the other correlated gates for the current of one gate. To mitigate this problem, several approaches have been proposed recently [2, 5, 6, 13]. The work in [2] partitions the whole chip into many grids to reduce the number of variables at the loss of accuracy. This method works well when correlation is strong. The approach in [13] also explores the strong spatial correlation by reducing the number of variables using principal component analysis (PCA) and orthogonal polynomials to represent leakage currents at gate and whole chip levels. The PCA-based approach in [6] only works when random variables are uncorrelated. In [5], a linear time complexity method is proposed, but it assumes symmetric spatial correlation with simplified leakage models (considering only the channel length variations). Recently an efficient method was proposed [17], which works for weak spatial correlations. The method formulates the major computation task into matrix-vector multiplications based on simplified gate leakage models via Taylor's expansion. It then applies fast computing methods like Fast Multi-pole method or pre-corrected FFT to compute the multiplication. However, this method assumes that leakage currents are purely lognormal, which is not the case as we show in the paper. Also it can only give the mean and variance, not the complete distribution of the leakage powers.

In this paper, we present a new grid-based linear-time algorithm for SLA. The new algorithm exploits the fact that

leakage of a gate in the presence of spatial correlation is only affected by gates in a few neighbor grids when the grid sizes are properly selected based on the correlation. Our new method is inspired by the recent spatial correlation models for statistical timing analysis [3]. The new algorithm adopts a similar set of uncorrelated variables over virtual grids to represent the original physical random variables via fitting and the grid size (thus of number of new random variables) is determined by the spatial correlation length. In this way, each physical variable is always represented by virtual variables locally. We prove that the number of neighboring virtual grids for each grid is not related to condition of spatial correlation (strong or weak), which leads to linear time complexity in terms of number of gates.

We compute the gate leakage by the orthogonal polynomials based collocation method and the variational leakage of one gate is represented in an analytic form in terms of the random variables, which can give complete distribution. Leakage of a whole chip can be computed by simply summing up the coefficients of corresponding orthogonal polynomials for each grid. Further more, we can cache statistical leakage results for each type of gates (by building a look-up table for the coefficients of orthogonal polynomials of each kind of gates), which can be computed just once in standard cell library. As a result, we can achieve even faster time complexity – $O(N)$, where N is the number of grids on chip without library construction costs. The new method can work with any condition of spatial correlations (strong or weak). Unlike the existing approaches [2, 5], the new method does not make any assumptions about the distributions of leakage currents on gate-level or chip-level. Experimental results show that the proposed method is about 1000X faster than the recently proposed grid-based method [2] with similar accuracy.

The rest of this paper is organized as follows: Section 2 presents the process variational models and gate leakage models used in this work. Section 3 presents the new spatial correlation models. Section 4 presents our proposed method. Section 5 presents the experimental results and Section 6 concludes this paper.

2. VARITIONAL AND LEAKAGE MODELS

In this section, we present static leakage models and the spatial correlated variational models used in this work.

2.1 Static leakage modeling

Full-chip leakage current has two major components, I_{sub} and I_{gate}. I_{sub} is exponentially dependent on V_{th}, which is observed to be most sensitive to T_{ox} and L due to short-channel effects. ITRS'08 [1] indicates that the variation of L is a primary factor for device parameter variation, and the number of dopants in channel results in an unacceptably large statistical variation of V_{th}. When the change in L or T_{ox} is small, the precise relationship shows an exponential dependent effect on I_{sub}, with the effect of T_{ox} being relatively weak. For I_{gate}, both L and T_{ox} have strong impacts on the leakage currents, which are exponential functions of the two variables.

In our work, we also follow the analytical expressions given in [2], which estimate the subthreshold and gate oxide leakage currents as follows:

$$I_{sub} = e^{a_1+a_2L+a_3L^2+a_4T_{ox}^{-1}+a_5T_{ox}}, \quad (1)$$

$$I_{gate} = e^{a_1+a_2L+a_3L^2+a_4T_{ox}+a_5T_{ox}^2}, \quad (2)$$

where a_1 through a_5 are the fitting parameters for each unique input combination of a gate. Notice that the exponents of I_{sub} and I_{gate} are not linear functions of L and T_{ox}, which is required by the existing approach [17]. The actually gate leakage currents are taken as the average of the leakage currents from all the input states in this work. We notice that some works like [2] keep only the values from dominate states.

2.2 Spatial variational models

Following the existing approaches, we also assume that the process variations of L and T_{ox} follow multivariate normal distributions [14], and both of them include inter-die and intra-die components. Since T_{ox} is in vertical layout feature dimension, and is caused by chemical mechanical polishing processes, it only depends on local layout density and has no spatial correlation [10]. Therefore, we focus on the spatial correlation of L. In general, the number of process parameters that exhibit spatial correlation can be more than one, and it is understood that this is not a limitation of our approach.

For a gate/module in a chip, we model the variations of L and T_{ox} as the following random variables:

$$\begin{aligned} \Delta L &= \Delta L_{inter} + \Delta L_{intra} = \sigma_{L,inter}\xi_{L,inter} + \Delta L_{intra}, \\ \Delta T_{ox} &= \Delta T_{ox,inter} + \Delta T_{ox,intra} \\ &= \sigma_{ox,inter}\xi_{ox,inter} + \sigma_{ox,intra}\xi_{ox,intra}, \end{aligned} \quad (3)$$

where $\xi_{L,inter}$, $\xi_{ox,inter}$, $\xi_{ox,intra}$ are independent random variables following $N(0,1)$, since ΔL_{inter} and $\Delta T_{ox,inter}$ are the same for all gates in all grids, and T_{ox} has no spatial correlation. On the other hand, ΔL_{intra} is different for each gate, and has spatial correlations. The spatial correlation used in this paper is given by the following empirical exponential model [16].

$$\rho(d) = e^{-d^2/\eta^2}, \quad (4)$$

where d is the distance between two grid centers and η is called the *correlation length*. Large η means the spatial correlation is strong, vice versa. The spatial correlation can be captured by the spatial covariance matrix $\Omega_{n,n}$, where n is the number of gates on chip. The elements in $\Omega_{n,n}$ are modeled as (4), which are only related to d.

As mentioned in Section 1, traditional SLA has $O(n^2)$ complexity in the presence of spatial correlation. There are mainly two schemes proposed to partially mitigate this problem. The first is the grid-based method [2] where all the correlated variables in a grid will be represented by one grid variable. But this method works only for strong correlation, as the number of grids will be much smaller than number of gates. The second approach utilizes PCA or PCA-like techniques to reduce the number of variables [13]. The errors can be easily controlled during the numerical processes. But the new set of independent variables has no physical meaning. Both methods achieve linear time complexity $O(n)$ in case of strong spatial correlation. In this paper, we show that we can achieve linear time complexity in term of number of grids for both strong and weak spatial correlation at similar errors of existing approaches. Our new approach is based on a new set of grid-based *independent* variables, which is used to model spatial correlation and has more physical meanings, as shown in next section.

3. NEW VIRTUAL GRID BASED SPATIAL CORRELATION MODELING

Existing approaches to handle spatial correlation for fast statistical analysis by grids or by PCA-based reduction only

work well for strong spatial correlation. In this paper, we introduce a new virtual-grid based spatial correlation model, which was proposed recently for fast statistical timing analysis [3] to address the computational efficient modeling for weak spatial correlation. In Section 4, we will prove that this model works for both strong and weak cases.

For simplicity of presentation, only one parameter L is considered to be spatial correlated here. But it is understood that multiple variables with spatial correlation can also be handled by our techniques to be presented.

The new modeling is based on the observation that the leakage current of a gate in the presence of weak spatial correlation only correlates to a few neighbor gates. If we can introduce a new set of uncorrelated variables which can catch the localized correlation, computing the leakage current of one gate can be done in a constant time by only considering its neighbor gates. Hence total full-chip statistical leakage currents can then be computed by simply adding all the gate leakage currents together in terms of the new set of variables in linear time. The idea is similar to the PCA-based approach [13] but with different set of new independent variables.

Specifically, the chip area is still divided into a set of grid cells. In our case, the grid can be very small so that every grid can even contain one gate. Then we introduce a "virtual" random variable for each grid for one source of process variation.

These virtual random variables are *independent* and will be the basis for statistical leakage current calculation that concern spatial correlation. Then we can express the original physical random variable of a gate in a grid as a linear combination of the virtual random variables of its own grid as well as its nearby neighbors. Since each virtual randome variable is defined for each grid, which has specific location in a chip, such location-dependent correlation model still retains the important spatial physical meaning (compared to PCA-based models). The grid can be rectangle or hexagonal or other shapes. We use hexagonal grids [3] in this paper since they have minimum anisotropy for 2-D space. And the proposed approach can be applied on any shape-based partition.

Here we define d_c as the grid length. Gates located in the same grid have strong correlation (larger than a given threshold value ρ_{high}) and are assumed to have the same parameter variations. And "spatial correlation distance" d_{max} is defined as the minimum distance beyond which the spatial correlation between any two grids is sufficiently small (or smaller than a given threshold value ρ_{low}) so that we can ignore it.

In this new model, the ith grid is associated with one virtual random variable $\xi_i \sim N(0,1)$, which is independent of all other virtual random variables. $\Delta L_{intra,i}$ can then be expressed as its k closest neighboring grids. We introduce the concept of *correlation index neighbor* set $T(i)$ for grid i, which defines a set of the indices of ξ_i to model the spatial correlation of $\Delta L_{intra,i}$ as

$$\Delta L_{intra,i} = \sum\nolimits_{q \in T(i)} \alpha_q \cdot \xi_q. \tag{5}$$

For example, if $T(i)$ for each grid is defined as its closest $k = 7$ neighboring grids, then ΔL located at grid (x_i, y_i) can be represented as a linear combination of seven virtual random variables located in its neighbering set. The grid model are shown in Fig. 1. Take $\Delta L_{intra,1}$ in Fig. 1 for instance, we have $\Delta L_{intra,1} = \alpha_1\xi_1 + \alpha_2\xi_2 + \ldots + \alpha_7\xi_7$.

This concept of correlation index helps to model the spatial correlation. Two grids close to each other will share more common spatial random variables, which means the correlation is strong. On the other hand, two grids physically far away from each other will share less or no common spatial random variables. In this way, the spatial correlation is modeled as a *homogeneous and isotropic random field* so that the spatial correlation is only related to distance. That is to say, spatial correlation can be fully described by $\rho(d)$ in (4).

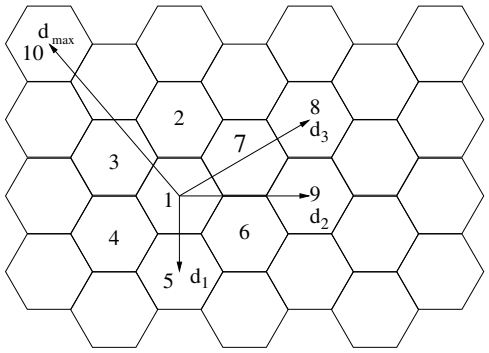

Figure 1: Location-dependent modeling with the $T(i)$ of grid i defined as its seven neighboring cells.

Since $\rho(d)$ is only a function of distance, the number of unique distance values between two correlated grids equals the number of unique element values in $\Omega_{N,N}$. Take the $T(i)$ set we mentioned in Fig 1 for example. In this case, the spatial correlation distance $d_{max} = \sqrt{21}d_c$, and there are only three unique correlation distances d_1 to d_3. Correspondingly, there are only three unique correlation coefficients in $\Omega_{N,N}$, without including 0 for $d \geq d_{max}$ or $\sigma^2_{L,intra}$ for distance within one grid.

Furthermore, the same correlation index can be used for all grids and the coefficient α_k should be the same for the same distance because of the homogeneousness and isotropy of spatial correlation. For the grid marked 1 in Fig. 1, we only have two unique values among the seven coefficients, i.e., we set $p_0 = \alpha_1$, $p_1 = \alpha_i, i = 2, 3, \ldots, 7$. In other words, we have $\Delta L_{intra,1} = p_0\xi_1 + p_1(\xi_2 + \ldots + \xi_7)$. According to (4), a nonlinear over-determined system can be built to determine the two unique values of p_0, p_1 as follows,

$$\begin{aligned} p_0^2 + 6p_1^2 &= \sigma_{L,intra}\rho(0) \\ 2p_0p_1 + 2p_1^2 &= \sigma_{L,intra}\rho(d_1) \\ 2p_1^2 &= \sigma_{L,intra}\rho(d_2) \\ p_1^2 &= \sigma_{L,intra}\rho(d_3) \end{aligned} \tag{6}$$

In matrix form, we can rewrite (5) for whole chip as

$$\Delta L_{intra} = P_{N,N} \cdot \vec{\xi}, \tag{7}$$

where N is the number of grids, and $\vec{\xi} = [\xi_1, \xi_2, \ldots, \xi_N]$. According to (5), the correlation index set contains only k neighboring spatial random variables, which is a very small fraction of the total spatial random variables. As a result, $P_{N,N}$ is a sparse matrix. Every gate only relates to k random variables, which has specific location information.

Fundamentally, PCA-based method performs a similar process and has a similar new transformation matrix between the original and new set of variables:

$$\Delta L_{intra} = V_{n,n} \cdot \vec{\xi}, \tag{8}$$

where $V_{n,n}$ is the transformation matrix obtained from eigenvalue decomposition of the correlation matrix in PCA. The

major difference is that $V_{n,n}$ is a dense matrix even though the original correlation matrix is sparse. This makes a huge difference especially when the spatial correlation is weak as eigen-decomposition will take almost $O(n^3)$ to compute. We remark that the new independent spatial correlation model also works for medium and strong correlation cases, which will be shown in the next section.

4. THE PROPOSED LINEAR-TIME ANALYSIS METHOD

In this section, we will present the new full-chip statistical leakage analysis method. The new algorithm consists of two parts. The first part is pre-characterization as shown in Fig. 2. We build analytic leakage current expressions for each kind of gate on top of a small set of independent virtual random variables. For fixed values of ρ_{high}, ρ_{low} and one library, a look-up table can be set up for coefficients of Hermite polynomials of I_{sub} and I_{gate} for the leakage analytic expressions for each kind of gate using (15) and (16). This process only need to be done once for one LIBRARY, given ρ_{high} and ρ_{low}. Besides, it deals with a tiny-size nonlinear over-determined system, which can be solved fast with any least-square algorithm.

Algorithm: BUILD LOOK-UP TABLE

Input: standard cell lib, ρ_{high}, ρ_{low}.
Output: look-up table for coefficients of Hermite polynomials of I_{sub} and I_{gate} for the leakage analytic expressions for each kind of gate using (15) and (16).

1. Generate fitting parameter matrices a_{sub} and a_{gate} of I_{sub} and I_{gate} in (1) and (2) for each type of gates (after SPICE simulation on each input pattern) (Section 2.1).
2. Calculate d_{max}/d_c from ρ_{high} and ρ_{low} to determine the neighboring set. And then solve (6) to determine coefficients in (5).
3. Generate Smolyak quadrature points set Θ_z^2 with corresponding weights.
4. Calculate the coefficients of Hermite polynomials of I_{sub} and I_{gate} for the leakage analytic expressions for each kind of gate in library using (15) and (16).

Figure 2: Flow of building the look-up table.

When we deal with full-chip SLA, the coefficients of local Hermite polynomials in neighboring grid set for each grid can be simply calculated by the look-up table. After transferring the local coefficients to corresponding global positions, we can compute the final full-chip leakage expressions by simple polynomial additions. From the expression, we can calculate other statistical information (like mean, variance, and even the whole distributions).

4.1 Review of orthogonal polynomial method

A random variable $x(\vec{\xi})$ with limited variance can be approximated by truncated Hermite PC expansion as follows [4]:

$$x(\vec{\xi}) = \sum_{q=0}^{Q} a_q H_q(\vec{\xi}), \tag{9}$$

where $\vec{\xi} = [\xi_1, \xi_2, ..., \xi_z]$. $\xi_i \sim N(0,1)$, and are orthogonal to each other. $H_q(\vec{\xi})$ is Hermite polynomial and a_q is the deterministic coefficient. For example, 2nd order Hermite polynomials set includes

$$H_{0th} = 1, \quad H_{1st} = \xi_i,$$
$$H_{2nd,type1} = \xi_i^2 - 1, \quad H_{2nd,type2} = \xi_i \xi_j, \quad (i \neq j). \tag{10}$$

a_q can be determined by

$$a_q = \frac{< x(\vec{\xi}), H_q(\vec{\xi}) >}{< H_q^2(\vec{\xi}) >} \approx \sum x(\gamma_l) H_q(\gamma_l) w_l, \tag{11}$$

where γ_l and w_l are Gaussian Hermite quadrature abscissas (quadrature points) and weights, respectively. The kernel of this technique is calculating the coefficients of Hermite polynomials efficiently. For multiple random variables, which require multi-dimensional quadrature, Smolyak quadrature [9] is used as an efficient method to reduce the number of quadrature points. For z-dimensional integration, the size of Level-Q Smolyak quadrature point set is $O(z^Q/(Q)!)$.

In our problem, $x(\vec{\xi})$ will be the leakage current for each gate, and eventually for the whole chip. For the jth gate, which is in the ith grid, from (5), $\Delta L_{intra,j}$ only relates to k independent virtual random variables in $T(i)$, and from (3) and (3), the corresponding variable vector, $\vec{\xi}_{g,j}$, is defined as

$$\vec{\xi}_{g,j} = [\xi_q, q \in T(i), \ \xi_{ox,i}, \xi_{L,inter}, \xi_{ox,inter}], \tag{12}$$

which only includes $z = k + 3$ independent variables.

To compute leakage current for one gate, we need to present both I_{sub} and I_{gate} of each gate in Hermite polynomials, respectively:

$$I_{sub}(\vec{\xi}_{g,j}) = \sum_{q=0}^{Q} I_{sub,q,j} H_q(\vec{\xi}_{g,j}), \tag{13}$$

$$I_{gate}(\vec{\xi}_{g,j}) = \sum_{q=0}^{Q} I_{gate,q,j} H_q(\vec{\xi}_{g,j}), \tag{14}$$

$I_{sub,q,j}$ and $I_{gate,q,j}$ are then computed by the numerical Smolyak quadrature method. In this paper, we use 2nd order Hermite polynomials for SLA. Let S be the size of z-dimensional second order (level-2) quadrature point set Θ_z^2. From [9], $S \sim O(k^2)$. $I_{sub,q,j}$ and $I_{gate,q,j}$ can be computed as the following:

$$I_{sub,q,j} = \sum_{l=1}^{S} I_{sub}(\vec{\gamma_l}) H_q(\vec{\gamma_l}) w_l / < H_q^2(\vec{\xi}_{g,j}) >, \tag{15}$$

$$I_{gate,q,j} = \sum_{l=1}^{S} I_{gate}(\vec{\gamma_l}) H_q(\vec{\gamma_l}) w_l / < H_q^2(\vec{\xi}_{g,j}) >, \tag{16}$$

where $\vec{\gamma_l}$ is Smolyak quadrature sample, $I_{sub}(\vec{\gamma_l})$ and $I_{gate}(\vec{\gamma_l})$ are computed using (1) and (2).

As a result, their coefficients for qth Hermite polynomial at jth gate can be added directly as

$$I_{leakage,q,j} = \sum I_{sub,q,j} + \sum I_{gate,q,j}. \tag{17}$$

Notice that the time complexity of calculating leakage for a gate is $O(k^2)$. And the number of involved independent random variables k is very small (compared to total number of gates), which is the case here. The analytic expression is also in terms of those involved random variables.

4.2 Acceleration by look-up table approach

One observation is that the variables in $T(i)$ actually represent the relative location, not the absolute location. In other words, a local neighboring set T and a local set of variables $\vec{\xi}_{loc} = [\xi_1, \ldots, \xi_k]$ can be shared by all the gates in all the grids for calculating the coefficients of Hermite

polynomials in (15) and (16) as the grids are distributed homogenously throughout the whole chip.

Obviously, the neighboring set T and the coefficients in (5) are determined by d_{max}/d_c. From the specific spatial correlation model in (4), the ratio of spatial correlation distance d_{max} over grid length d_c becomes

$$d_{max}/d_c = \sqrt{ln(\rho_{low})/ln(\rho_{high})}. \quad (18)$$

We observe that once the threshold values ρ_{high} and ρ_{low} are set, d_{max}/d_c is not related to the correlation length η. This means the spatial correlation (strong or weak) has nothing to do with T and the virtual random variables used in our model.

At the same time, the fitting parameters of static leakage in (1) and (2) is only related to the types of gate in library. As a result, the coefficients of Hermite polynomials for the leakage of one gate in (15) and (16) are only functions of the type of the gate, ρ_{high} and ρ_{low}. Therefore, a simple look-up table can be used to store the coefficients of Hermite polynomials of all kinds of gate in library. In other words, we do not need compute the coefficients of Hermite polynomials for each gate, just for each *type* of gates instead. Actually this makes a big difference as the time complexity is no longer number of gates, $O(n)$. Instead it is a linear function of number of grids, $O(N)$.

For the look-up table, suppose Q is the number of Hermite polynomials involved and m is the number of gate type in library, then it includes two tables as follows: $C_S = \{I_{sub,q,j}\}$, where $I_{sub,q,j}$ represents the coefficient of H_q for jth kind of gate in libary for subthreshold leakage; and $C_G = \{I_{gate,q,j}\}$, where $I_{sub,q,j}$ represents the coefficient of H_q for jth kind of gate in library for gate oxide leakage. C_S and C_G are $Q \times m$ matrices. Notice that the table needs to be built up only once and can be used for different designs with different conditions of spatial correlations as the new algorithm is independent of spatial correlation length η.

4.3 Computation of full-chip leakage currents

Here we define a gate map matrix $G_{N\times m} = \{g_{i,j}\}$, where $g_{i,j}$ represents the number of jth kind of gate in library located in ith grid on chip. Then the coefficients of local Hermite polynomials in neighboring set for all the grids on chip can be easily calculated by the look-up table as follows,

$$I_{sub,loc} = G \cdot C_S^T, \quad I_{gate,loc} = G \cdot C_G^T \quad (19)$$

In order to get the full-chip leakage current, the local coefficients need to be transferred to their corresponding global position

$$T(i) = (x_i, y_i) + T. \quad (20)$$

For the ith grid, the local set of random variables $\vec{\xi}_{loc}$ should be transferred to the corresponding positions in $T(i)$. Therefore, $I_{sub,loc}$ and $I_{gate,loc}$ can be transferred to the corresponding global coefficients based on the global virtual random variable set $\vec{\xi}$. For example, the coefficient of ξ_i in the ith grid is

$$I_{sub}(\xi_i) = \sum_{k,i\in T(k)} I_{sub,loc}(\xi_{T(k)-(x_k,y_k)}) \quad (21)$$

After that, we can proceed to compute the leakage current for the whole chip as follows,

$$I_{leakage}(\vec{\xi}) = \sum I_{sub}(\vec{\xi}) + I_{gate}(\vec{\xi}) \quad (22)$$

The summation is done for each coefficient of global Hermite polynomials to obtain the analytic expression of the final leakage currents in terms of $\vec{\xi}$. We can then obtain the mean value, variance, PDF and CDF of the leakage current very easily. For instance, the mean value and variance for the full-chip leakage current are

$$\mu_{leakage} = I_{leakage,0th}, \quad (23)$$

$$\sigma^2_{leakage} = \sum I^2_{leakage,1st} + 2\sum I^2_{leakage,2nd,type1} + \sum I^2_{leakage,2nd,type2}, \quad (24)$$

where $I_{leakage,ith}$ is the leakage coefficient for ith Hermite polynomial of second order defined in (10).

5. EXPERIMENTAL RESULTS

The proposed method has been implemented in Matlab 7.8.0. Since the leakage model for method in [17] has to be purely log-normal (linear terms in exponent parts), we did not choose it for comparing purpose. All the experimental results are carried out in a Linux system with quad Intel Xeon CPUs with 2.99Ghz and 16GB memory.

These methods were tested on PDWorkshop91 benchmark. The circuits were synthesized with Nangate Open Cell Library and the placement is from MCNC. The technology parameters come from the 45nm FreePDK Base Kit and PTM models [11]. According to [1], L and T_{ox} for high performance logic in 45nm technology will be 18nm and 1.8nm, respectively. And the physical variation should be controlled within +/-12%. So the 3σ values of variations for L and T_{ox} were set to 12% of the nominal values, of which inter-die variations constitute 20% and intra-die variations, 80%. L is modeled as sum of spatial correlated sources of variations, and T_{ox} is modeled as an independent source of variation. The same framework can be easily extended to include other parameters of variations. Both L and T_{ox} are modeled as Gaussian random variables. For L, the spatial correlation was modeled based on (4). The test cases are given in Table 1 (all length units in μm). In the fourth column, "s" and "w" stands for strong and weak spatial correlations, respectively.

Table 1: Summary of Test cases used in this paper.

Circuit	*Gate #*	*Area*	*Testcase*	*Grid #*
SC0	125	1459×1350	Case 1 (s)	4
			Case 2 (w)	175
SC1	1888	4892×4874	Case 3 (s)	27
			Case 4 (w)	1903
SC2	6417	10092×10466	Case 5 (s)	99
			Case 6 (w)	8167

For comparison purposes, we performed Monte Carlo (MC) simulations with 50,000 runs, the method in [2] and the proposed approach on the benchmarks. The results of the comparison of mean value and standard deviations of full-chip leakage current are shown in Table 2, where *New* is the proposed method. The average errors for mean and standard variance (σ) values of the new technique are 6.68% and 8.49%, respectively. While for the method in [2], the average errors for mean value and σ are 6.10% and 10.41%, respectively. From Table 2, it is shown that these two algorithms have almost the same accuracy. The results also show that our method can handle both strong and weak spatial correlations by adjusting grid size.

Table 3 compares the CPU times of the three methods, which shows that the new method, *New*, is much faster than the method in [2] and MC simulation. On average, the proposed technique has about 2012X speedup over [2] and many

Table 2: SLA accuracy comparison based on Monte Carlo.

Test Case	*Grid #*	*Mean Value (μA)*			*Errors (%)*		*Standard Deviation (μA)*			*Errors (%)*	
		MC	Method [2]	New	Method [2]	New	*MC*	Method [2]	New	Method [2]	New
Case1	4	3.313	3.105	3.118	-6.27	-5.89	0.9873	0.8540	0.9115	-13.5	-7.68
Case2	175	3.313	3.105	3.118	-6.26	-5.88	0.5951	0.5506	0.5586	-7.48	-6.13
Case3	27	43.04	40.73	39.68	-5.35	-7.81	9.311	8.574	7.887	-7.92	-15.3
Case4	1903	43.10	40.73	39.68	-5.47	-7.93	7.775	6.924	7.065	-10.94	-9.12
Case5	99	191.7	179.0	179.6	-6.64	-6.30	34.24	30.62	31.84	-10.57	-7.00
Case6	8167	191.7	179.0	179.6	-6.62	-6.28	32.66	28.72	30.79	-12.06	-5.74

orders of magnitudes over the MC method. And the speed of our approach is not affected by the total number of grids. If the spatial correlation is strong, which means d_{max} is large, d_c can be increased at the same time without loss of accuracy. So the number of neighbor grids in $T(i)$ will still be much smaller than the number of gates. The new method will be efficient and linear in both cases.

Table 3: CPU time comparison.

Test Case	*CPU time(s)*			*Speedup over*	
	MC	[2]	New	MC	[2]
Case1	83.14	19.54	0.0197	4220	991
Case2	87.09	17.72	0.0236	3690	750
Case3	828.42	141.74	0.0330	25103	4295
Case4	869.12	74.50	0.4559	1906	163
Case5	4571.1	832.63	0.1522	30033	5411
Case6	8460.3	3972.2	8.5462	989	464

6. CONCLUSION

In this paper, we have presented a linear-time algorithm for full-chip statistical analysis of leakage currents in the presence of any condition of spatial correlation (strong or weak). The new algorithm adopts a new set of uncorrelated variables over virtual grids to represent the original physical random variables with spatial correlation and the grid size is determined by the spatial correlation length. In this way, each physical variable is always represented by virtual variables in local neighboring set. Furthermore, we propose to use look-up table to cache the statistical leakage information of each type of gate in library so that we do not need to compute leakage currents for each gate. As a result, the full-chip leakage can be calculated with $O(N)$ time complexity, where N is the number of grids on chip. The new method maintains the linear complexity for both strong and weak spatial correlation and has no limitation of leakage current model or variation model. Experimental results show that the proposed method is about 1000X faster than the recently proposed method [2] with similar accuracy and many orders of magnitude times over the Monte Carlo method.

7. REFERENCES

[1] International technology roadmap for semiconductors(ITRS), 2008. http://public.itrs.net.

[2] H. Chang and S. S. Sapatnekar. Full-chip analysis of leakage power under process variations, including spatial correlations. In *Proc. Design Automation Conference (DAC)*, pp. 523–528, 2005.

[3] R. Chen, L. Zhang, V. Zolotov, C. Visweswariah, and J. Xiong. Static timing: back to our roots. In *Proc. Asia South Pacific Design Automation Conf. (ASPDAC)*, pp. 310–315, Jan. 2008.

[4] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Dover Publications, 2003.

[5] K. R. Heloue, N. Azizi, and F. N. Najm. Modeling and estimation of full-chip leakage current considering within-die correlation. In *Proc. Design Automation Conference (DAC)*, pp. 93–98, 2007.

[6] X. Li, J. Le, and L. T. Pileggi. Projection-based statistical analysis of full-chip leakage power with non-log-normal distributions. In *Proc. Design Automation Conference (DAC)*, pp. 103–108, 2006.

[7] S. Narendra, V. De, S. Borkar, D. A. Antoniadis, and A. P. Chandrakasan. Full-chip subthreshold leakage power prediction and reduction techniques for sub-0.18-μm CMOS. *IEEE J. Solid-State Circuits*, 39(2), 2004.

[8] S. Nassif. Delay variability: sources, impact and trends. In *Proc. IEEE Int. Solid-State Circuits Conf.*, pp. 368–369, San Francisco, CA, Feb 2000.

[9] E. Novak and K. Ritter. Simple cubature formulas with high polynomial exactness. *Constructive Approximation*, 15(4):449–522, 1999.

[10] D. O. Ouma, D. S. Boning, J. E. Chung, W. G. Easter, V. Saxena, S. Misra, and A. Crevasse. Characterization and modeling of oxide chemical-mechanical polishing using planarization length and pattern density concepts. *IEEE Trans. on Semiconductor Manufacturing*, 15(2):232–244, 2002.

[11] Predictive technology model. `http://www.eas.asu.edu/~ptm/`.

[12] R. Rao, A. Srivastava, D. Blaauw, and D. Sylvester. Statistical analysis of subthreshold leakage current for VLSI circuits. *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, 12(2):131–139, Feb 2004.

[13] R. Shen, N. Mi, S. X.-D. Tan, Y. Cai, and X. Hong. Statistical modeling and analysis of chip-level leakage power by spectral stochastic method. In *Proc. Asia South Pacific Design Automation Conf. (ASPDAC)*, pp. 161–166, Jan. 2009.

[14] R. Teodorescu, B. Greskamp, J. Nakano, S. R. Sarangi, A. Tiwari, and J. Torrellas. A model of parameter variation and resulting timing errors for microarchitects. In *Workshop on Architectural Support for Gigascale Integration (ASGI)*, Jun 2007.

[15] J. M. Wang, B. Srinivas, D. Ma, C. C.-P. Chen, and J. Li. System-level power and thermal modeling and analysis by orthogonal polynomial based response surface approach (OPRS). In *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 727–734, Nov. 2005.

[16] J. Xiong, V. Zolotov, and L. He. Robust extraction of spatial correlation. *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, 26(4), 2007.

[17] Z. Ye and Z. Yu. An efficient algorithm for modeling spatially-correlated process variation in statistical full-chip leakage analysis. In *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 295–301, Nov. 2009.