Variogram based Robust Extraction of Process Variation Model

Kaviraj Chopra, Narendra Shenoy and David Blaauw

Abstract— **Aggressive device scaling has made it imperative to account for process variations in the design flow. A robust model of process variations is an essential requirement for any meaningful variation aware design analysis and optimization. Unfortunately the previous approaches on extracting spatial correlation function assume ergodicity and isotropy while estimating the inter-die(global) component of variation and spatial correlation function, respectively. We find that making such simplifying assumptions may result in significant estimation error. In this work to address these issues, we propose an alternative approach to extract spatial variation models based on the theory of spatial statistics. The proposed approach uses the concept of variogram function that represents how parameters can covary as a function of spatial distance. The variogram function provides us with a representation that is independent of the global component of variation. This allows us to directly estimate the within die component of variations and thus circumvents the need for making ergodicity assumption. We further show that using two dimensional variogram functions allows us to model geometrically anisotropic process variation data. Additionally, for extracting process variation models in the presence of significant measurement noise, we employ *weighted least squares* regression technique, which is known to be statistically more robust technique than the previously used ordinary least square technique. Our experimental results on extracting the process variation model model from a Monte-Carlo generated data-set corrupted with significant random noise validates the robustness of the proposed approach.**

I. INTRODUCTION

The semiconductor manufacturing process has become more complex, while process control precision is struggling to maintain its accuracy with continuous process scaling. As a result, a variety of steps throughout the manufacturing process are prone to fluctuations. These include effects due to chemical mechanical polishing which is used to planarize oxides surrounding metal lines and optical proximity effects which are a consequence of patterning features smaller than the wavelength of light [13], [11]. These in combination with numerous other effects cause an increase in variation of device and interconnect physical parameters such as gate length (or Critical Dimension - CD), gate-oxide thickness, channel doping concentration, interconnect thickness and height, etc which in turn affects their electrical characteristics. These variations in electrical characteristics of circuit components have led to increased variability in circuit performance, resulting in yield degradation. Unfortunately, as listed in [1], no manufacturable solutions are available for controlling the variability of several parameters (e.g. CD) for 45 nm technology nodes. Thus, for achieving reasonable yield it has become important to model variability during design. Acknowledging these issues, statistical performance analysis and optimization approaches [5], [16], [12], have emerged as a possible solution for statistically quantifying the variability in performance. A model of parameter variation is an essential input for both statistical circuit analysis and optimization, therefore it is important to understand and characterize parameter variations.

Numerous research efforts [5], [16], [15], have focused on developing statistical analysis and optimization techniques assuming thatg of a model of parameter variation is known a priori. However not much attention has been paid on extracting such a model from experimental measurements. In [14] the authors focused on extracting the deterministic component of CD variation for 0.18μm CMOS technology. But as random variations were insignificant for the 0.18μm CMOS technology the random component of spatial variations was ignored. A simple computation of the spatial correlation coefficient as a function of distance from wafer-scale measurement data was presented in [7]. From their measurement results, it was found that spatial correlation was significantly different in the horizontal and vertical direction. In the field of spatial statistics, this difference in spatial dependence across all die with respect to the direction is referred to as *anisotropy*.

Recently the authors in [17], [18] introduced a formal approach to model and extract correlation functions. They noted an important result from random field theory that a correlation function is valid only if any correlation matrix derived from it is a positive semi-definite and therefore, any arbitrary function such as those derived from polynomial curve fitting cannot be a valid spatial correlation function. They also presented an algorithm to extract such a valid correlation function from measurement data in the presence of random noise. In cases when measured variation data is found not to be stationary, they further proposed a technique to extract a valid spatial correlation matrix by employing a modified alternative projection algorithm. A study of different correlation models and their associated timing methods using measured critical dimension (CD) measurement data is given in [3]. A heuristic algorithm to fit quad-tree based correlation models was also presented in [3]. Recently, we found out independent of our work another approach [10] was simultaneously developed based on the spatial statistics concepts [6]. In [10], the authors proposed a geo-statistics based mathematical method to extract a model of spatial correlation trend of a single die.

In this work, we present an alternative approach to address the above limitations. We appeal to spatial statistics methods to develop a method for robust extraction of an anisotropic process variation model. Our approach is based on the concept

of variogram functions, which is an alternative representation of covariance function. The variogram function can be understood as the variance of the difference of two random variables at different locations in a die. By construction itself, the variogram representation becomes independent of the global component of variation. This approach also side-steps the need to compute the global variance prior to computing the correlation model. Apart from this key motivating property, it has been shown that the variogram function can be estimated more reliably from measurement data than the covariance function [6]. A further study of spatial statistics revealed that two dimensional valid variogram functions can be used to model anisotropic process variation data [9]. Using such a two dimensional anisotropic valid variogram function we develop a method for extracting the process variation model. First, we compute an empirical variogram from the given measurement data set. Then, we extract an estimate of a valid anisotropic variogram function from the empirical variogram using nonlinear regression. For nonlinear regression, the weighted least square is chosen as the objective, which is known to be more robust to outliers then the previously used least square objective. Finally, the global component is computed by subtracting the within die component of variation, found from the variogram function, from the overall variation. We applied the proposed approach on measured gate length data and cross-validated the extracted process correlation model using a spatial statistics method of Kriging (explained later).

The rest of the paper is organized as follows: in Section III we explain the concept of variogram. Section IV explains the proposed algorithm for extracting the process variation. Experimental results that show the accuracy of the proposed approach are given in Section V. We conclude the paper in Section V.

II. Problem Description

A decomposition of process variations can be understood as shown in Figure 1. The physical variation can be classified based on whether they are deterministic or statistical variations and based on the spatial scale over which they operate. *Systematic variations* are the components of process variability that follows a well understood behavior and can be predicted upfront by analyzing the design layout. Therefore, the impact of such variations can be accounted by using deterministic analysis at later stages of the design process, particularly at timing sign-off by simply adjusting their nominal values accordingly. *Non-Systematic variations* represent the truly uncertain component of parameter variation that is caused by those processes that are orthogonal to design implementation. Non-systematic variations can further be analyzed by observing that different sources of variations act on different spatial scales. Due to large scale shifts in the process steps that occur from lot-to-lot, wafer-to-wafer, reticle-to-reticle, and across a reticle, all the devices on the same die observe the same amount of fluctuations. This die-to-die component of variations is commonly referred to as *global* or *inter-die variation*. Furthermore, several underlying process steps that give rise to device parameter variation change gradually from one location to another within a die. Hence, these process steps will tend to affect all devices that are placed close to each other in a similar manner. Therefore, closely spaced devices are more likely to have similar characteristics than those placed far apart. The component of variation that exhibits such spatial dependence is known as *spatially correlated variation*. For accurate statistical analysis, it is necessary to capture the dependence between parameters of two devices due to global and spatial component of variations. The residual variation resulting from certain device scale effects such as random dopant fluctuations, line edge roughness, etc. is referred to as *independent variation*.

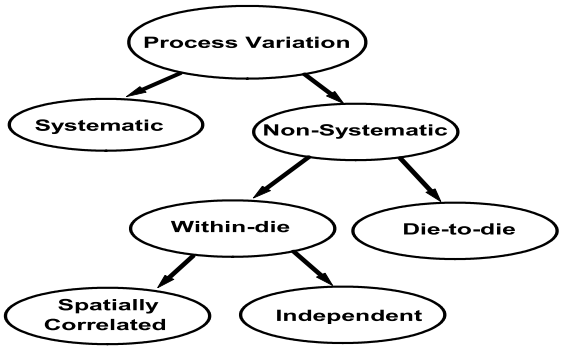

Fig. 1. Components of process variations.

Our goal is to extract a process variation model for some parameter of interest, which can be either a process parameter such as depth of focus and dose, a device parameter such as gate length and oxide thickness, or an electrical parameter like drive current, gate delay and threshold voltage. Due to global variations, manufactured values of a parameter vary from die to die whereas due to within die variation, they also vary with its location within a die. Therefore, for modeling within die variations, a distinct random variable has to be associated corresponding to each location in the die. Thus, all components of process variation can be modeled using a two dimensional random function that associates a random variable for each location within the die. To set notation, a model of process variation at location $\mathbf{x} = (x, y)$ can be given as follows:

$$F(\mathbf{x}) = f_0 + X + Y(\mathbf{x}) + Z(\mathbf{x}),$$

where (i) f_0 is the deterministic mean of the random function, (ii) X represents the global component of variation that is same for all locations in the die, (iii) $Y(\cdot)$ is a zero-mean random function defined over each location $\mathbf{x} = (x, y)$ in the die, representing the parameter's spatial dependence (after mean is removed) (iv) $Z(\cdot)$ is also a zero-mean random function, but it represents the residual independent component of variation. For any location $\mathbf{x}$, the corresponding random variables $X, Y(\mathbf{x})$ and $Z(\mathbf{x})$ are independent of each other. Recall, that the mean f_0 is also dependent on location due to design dependent systematic variations. In [14], the authors

have presented a detailed method for extracting systematic variation, using which mean trends resulting from systematic variations can be removed from the measured characterization data. Therefore, in this work, we focus on extracting the non-systematic components of variation. In particular, we wish to find the global variance σ_X^2, the spatial variance σ_Y^2, σ_Z^2 the variance of independent component and the correlation function.

III. Variogram Function

In this section, we present the concept of variogram function for modeling spatial variations. As mentioned earlier, the variogram function models the spatial dependence by quantifying the amount by which the parameters of two devices can vary from each other as a function of their spatial distance. The definition of variogram function is given as follows:

Definition 1: For a random function $F(\mathbf{s})$, **the variogram function** between any two points $\mathbf{s_1}$ and $\mathbf{s_2}$ is defined as

$$2\gamma(\mathbf{s_1},\mathbf{s_2}) = Variance(F(\mathbf{s_1}) - F(\mathbf{s_2})),$$

where $\gamma(\cdot)$ is referred to as the semi-variogram.
Similar to the second-order stationary assumption required for extracting correlation function, a more general intrinsic stationary assumption is required for characterizing random functions using variogram functions.

Definition 2: If the mean of the random function $F(\mathbf{s})$ is constant and its variogram function $2\gamma(\mathbf{s_1},\mathbf{s_2})$ between any two points $\mathbf{s_1}$ and $\mathbf{s_2}$ depends only on the difference vector $\mathbf{v} = \mathbf{s_1} - \mathbf{s_2}$, i.e.

$$2\gamma(\mathbf{s_1},\mathbf{s_2}) = 2\gamma(\mathbf{s_1} - \mathbf{s_2}) = 2\gamma(\mathbf{v}),$$

then it is called an **intrinsically stationary** random process.
To understand the variogram function, consider an example semi-variogram plot shown in Figure 2. In the presence of spatial dependence, two random variables (device parameters) located close to each other tend to vary in a similar manner. Therefore, the variogram function, which represents the variance of their difference, monotonically increases with the separation distance. Due to the device scale independent component of variation, the y intercept of the variogram function can be greater than zero.

All monotonically decreasing functions cannot be considered as valid candidates for a covariance function. Similarly, not all monotonically increasing functions can qualify for a valid variogram function. A spectral theory also exists for valid variogram functions that defines negative definiteness as the necessary condition for any arbitrary function to be a valid variogram function. A detailed description of the spectral theory of variogram functions can be found in [6]. For our purpose, a valid class of functions, known as Matern-class, can be used to extract a valid variogram function. A parametric family of valid variogram functions based on the Matern-class can be given as

$$2\gamma(\mathbf{v}) = p_0 + p_1\left(1 - \frac{1}{2^{p_3-1}\Gamma(p_3)}\left(\frac{\mathbf{v}}{p_2}\right)^{p_3} K_{p_3}\left(\frac{\mathbf{v}}{p_2}\right)\right),$$

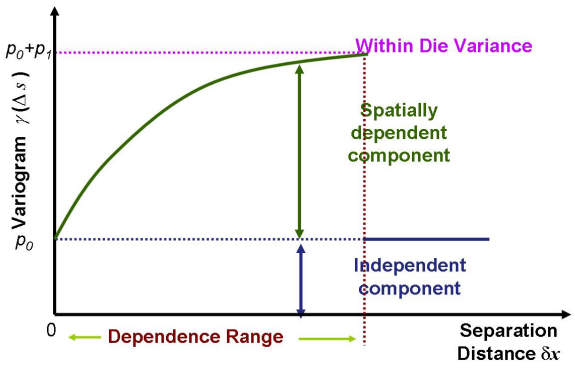

Fig. 2. A typical Variogram in 1-dimension

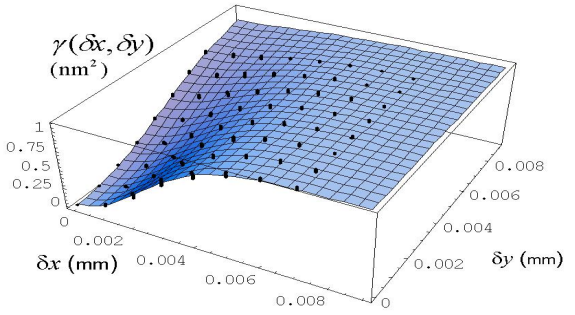

Fig. 3. A typical anisotropic Variogram in 2-dimension

where p_0 represents the variance of the independent component (i..e σ_Z^2), p_1 represents the variance of the spatially dependent component (i.e. σ_Y^2), p_2 is the spatial scale parameter, p_3 is the shape parameter, $\Gamma(\cdot)$ is the gamma function and $K_{p_3}(\cdot)$ is the modified Bessel function of the second kind and order p_3. The flexibility of the above mentioned Matern-class has been demonstrated on a variety of spatial data sets [8].

If the variogram depends only on the Euclidean distance (i.e. $||s_1 - s_2||$) between locations, then the variogram is said to be isotropic; otherwise, it is considered to be anisotropic. Anisotropy is caused by the underlying physical sources of variation evolving differentially in space. A classification of types of anisotropy arising in spatial analysis and methods for modeling them using appropriate variogram functions can be found in [9]. A common theme in most of the methods is to generalize the isotropic function to model the asymmetries. The simplest and the most common form of anisotropy is the *geometric anisotropy*, where the random function is not isotropic in the original space, but it is in some linearly transformed space. For modeling geometric anisotropy, a valid variogram function can be achieved from the isotropic variogram function by linearly transforming the distance vector $\mathbf{v}$. Based on this concept, the geometrically anisotropic Matern-

1. Compute the experimental variogram using (2)
2. Using nonlinear regression solve (IV), to find all parameters $\mathcal{P}$
3. From (4), find $\bar{\rho}(\mathbf{v})$ by substituting all parameters $\mathcal{P}$
4. Evaluate σ_X^2 from (5)

Fig. 4. Proposed algorithm for process variation extraction

class of variogram functions can be given as

$$2\gamma(\mathbf{v}) = p_0 + p_1\left(1 - \frac{1}{2^{p_3-1}\Gamma(p_3)}\left(\frac{||\mathbb{P}\mathbf{v}||}{p_2}\right)^{p_3}\left(\frac{||\mathbb{P}\mathbf{v}||}{p_2}\right)\right), \tag{1}$$

where $\mathbb{P}$ is 2×2-matrix of parameters that models the axes and scale of anisotropy. For simplicity, we denote the set of all parameters with $\mathcal{P} = \{p_0, p_1, p_2, p_3, \mathbb{P}\}$. Figure 3 shows a surface plot of a possible anisotropic semi-variogram as a function of distance vector $\mathbf{v} = (\delta x, \delta y)$. If $\mathbb{P}$ is an identity matrix then the above expression reduces to the isotropic case.

IV. Process Variation Extraction

In this section, we present the overall algorithm for extracting the process variation model (see Figure 4). As mentioned earlier, by construction the semi-variogram function is independent of the global component of variation. Therefore, by estimating the variogram representation, we can directly find the spatially dependent within-die component of variation. The measurement data of N die at M locations per die represents a set of samples of the random function $F(\cdot)$. For distinguishing between the actual function and measurement data, we use the over-line to denote a sampling, for example, the set of samples of random variable at location $\mathbf{s}$ is denoted by $\overline{F}(\mathbf{s})$. The k^{th} measured sample of the random variable $F(\mathbf{s})$ at location $\mathbf{s}$ is given by $f_k(\mathbf{s})$. From the measurement data, first, we find the unbiased estimate of the sample semi-variogram, using the following expression:

$$\begin{aligned}\overline{\gamma}(\overline{F}(\mathbf{s_i}),\overline{F}(\mathbf{s_j})) &= \frac{1}{2}Variance(\overline{F}(\mathbf{s_i}) - \overline{F}(\mathbf{s_j})) \\ &= \frac{1}{2(N-1)}\sum_{k\in N}(f_k(\mathbf{s_i}) - f_k(\mathbf{s_j}))^2 \\ &- \frac{\left(\sum_{k\in N}(f_k(\mathbf{s_i}) - f_k(\mathbf{s_j}))\right)^2}{2N(N-1)}.\end{aligned} \tag{2}$$

The sample semi-variogram $\overline{\gamma}(\overline{F}(\mathbf{s_i}),\overline{F}(\mathbf{s}_j))$ is computed for every location pair $(\mathbf{s_i},\mathbf{s_j})$ separated by the distance vector $\mathbf{v} = \mathbf{s_i} - \mathbf{s_j}$. This gives us the so-called *empirical semi-variogram* as a finite set of ordered pairs $(\mathbf{v}, \overline{\gamma}(\mathbf{v}))$.

Now, for estimating a valid variogram function, we wish to find the set of parameters $\mathcal{P}$ of the anisotropic Matern-class given in (III). This can be formulated as a nonlinear estimation problem. It is known from nonlinear estimation theory [2] that the method of least squares is not statistical; it is purely a numerical criterion. When the response variable (i.e. the experimental variogram $\overline{\gamma}(\mathbf{v})$) are random variables with a non-scalar variance matrix, the method of least squares is found to be an inefficient estimator [2]. Due to the spatial dependence amongst each of the measured locations, the variance-matrix of the empirical variogram data set is not a diagonal matrix. An alternative objective, known as generalized least squares provides statistically robust estimates, however it is found to be computationally more expensive. Nevertheless, an intermediate approach, the method of weighted least squares, provides reasonable compromise between robustness and computational efficiency. For extracting valid variogram function, a nonlinear estimation formulation based on the weighted least square objective can be given as follows:

$$\min_{\mathcal{P}} || \frac{(\overline{\gamma}(\mathbf{v}) - \gamma(\mathbf{v}))}{Variance(\overline{\gamma}(\mathbf{v}))} ||, \tag{3}$$

where $\overline{\gamma}(\mathbf{v})$ is the empirical semi-variogram, $\gamma(\mathbf{v})$ is the family of variogram functions given in (III), and $Variance(\overline{\gamma}(\mathbf{v}))$ is the weight representing the diagonal of the covariance matrix of estimates in the empirical variogram.

Solving the above optimization problem gives statistically robust estimates of parameters $\mathcal{P}$. The intra-die variation components $\overline{\sigma_Y^2} \approx p_1$ and $\overline{\sigma_Z^2} \approx p_0$ are obtained directly by estimating the variogram function. The estimate of spatial correlation function can be obtained from the variogram function and $\overline{\sigma_Y^2}$ and $\overline{\sigma_Z^2}$ using the following relation: (see Appendix for the derivation):

$$\bar{\rho}(\mathbf{v}) = \frac{\sigma_Z^2 + \sigma_Y^2 - \bar{\gamma}(\mathbf{v})}{\sigma_Y^2}. \tag{4}$$

Finally, the only component remaining to be estimated is the global variance σ_X^2. Now at any location s, an unbiased estimate of overall sample variance of random variable $F(\mathbf{s})$ can be found as follows:

$$Variance(\overline{F}(\mathbf{s})) = \frac{1}{N-1}\left(\sum_{k\in N} f_k(\mathbf{s_i})^2 - \frac{1}{N}\sum_{k\in N} f_k(\mathbf{s_i})\right)$$

As all components of variation, X,Y and Z are independent of each other, we know that the total variance of every random variable $F(\mathbf{s})$ is the sum of all three components out of which we already know $\overline{\sigma_Y^2}$ and $\overline{\sigma_Z^2}$. Thus for each random variable the global variance be estimated by subtracting the $\overline{\sigma_Y^2}$ and $\overline{\sigma_Z^2}$ from its overall variance. An estimate of the overall global variance can be found by taking the average over all M measurement locations on the die, i.e.

$$\overline{\sigma_X^2} \approx \frac{1}{M}\sum_{\mathbf{s}\in \text{Die}} Variance(\overline{F}(\mathbf{s}) - \sigma_Y^2 - \sigma_Z^2) \tag{5}$$

V. Experimental Result

In our first experiment, we applied our algorithm to extract a process variation model from 130nm Electrical Line-width Measurement (ELM) data taken from horizontal polysilicon lines that have optical proximity corrections included [4]. The data-set included measurements from 5 different wafers, each wafer contained 23 fields, and each field included 308 measurement points: 14 points in the horizontal direction and 22 points in the vertical direction. Individual measurement points were spaced horizontally by 2.19mm and vertically

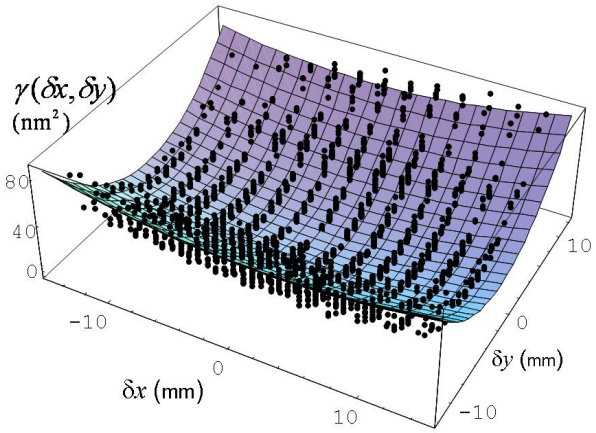

Fig. 5. The Variogram of measured gate length data

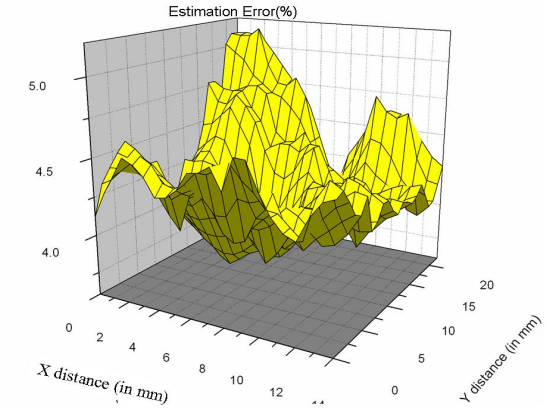

Fig. 6. A plot of Estimation Error $(\%)$ for a typical die.

by 1.14mm. For these 5 wafers, we divided the reticles into various die sizes in order to investigate the effects of die size on the gate-length variation. We diced a reticle into 4 die, a 2-die by 2-die configuration where each die is 15mm x 12mm. A surface plot of anisotropic variogram function extracted from the measurement data and the scatter plot of exact empirical variogram (dots) numerically computed from the ELM data are shown in Figure 5. From visual inspection, it can be seen that the proposed algorithm can extract good fits of valid anisotropic variogram functions from measured ELM data. Similar results were observed when the entire reticle was considered and when the reticle was divided into 16 die.

To further cross-validate the variogram fit of the ELM data a more rigorous validation was performed using the concept of Kriging [6]. Kriging, named after its inventor Danie G. Krige, is another useful concept developed in the field of geostatistics. It can be understood as a linear estimation technique for spatially dependent random variables. The prior knowledge about the spatial dependence of a parameter (say gate length) is first captured in the form of a variogram function. Then for any arbitrary die, given a subset of measured gate lengths on a specific die, Kriging theory gives the optimal conditional estimates of gate lengths at unobserved/unmeasured points using the extracted variogram function. The estimates are given as a weighted linear combination of the known gate lengths and the optimality criterion for Kriging is to minimize the least squares error.

The ELM data of 5 wafers was divided into two sub-sets: measurement of wafers $1-4$ which was used to a extract variogram function and (5^th wafer measurement data) which was used to corss-validate the extracted variogram function. For each die on the 5^{th} wafer, out of the total 11×7 measurement locations, half of the points were randomly chosen to be as observed points and the other half were estimated using the Kriging theory and extracted variogram function. The Kriging estimation error with respect to the extracted variogram function indicates the accuracy of the extracted variogram function and the process variation model. Several different Kriging estimation algorithms are developed in spatial statistics literature [6]. For this work, we implemented the ordinary Kriging approach and computed the average estimation error for each die. The worst-case error over all 23×4 die was observed to be 5%. In general, we observed that the extracted variogram function accurately estimates the measured CD trend for all 92 die. The maximum error was typically observed for die placed at the corners of the wafer. A typical surface plot of estimation error $(\%)$ across the two die-dimensions is given in Figure 6.

In the second experiment, we validate the accuracy and robustness of our algorithm using a Monte Carlo model of process variation. For the sake of comparison, we closely follow the experimental setup used in [17]. A valid correlation function $\rho(\mathbf{s})$ and known variance of all variation components are used by the Monte Carlo model. To model the measurement errors inherent in real data, different amounts of gaussian noise is added to the Monte Carlo model. A representative measurement data-set is generated using Monte Carlo simulations for N number of sample die and M number of measurement locations on each die. By applying the proposed algorithm, all components of process variation are extracted from the measurement data. A comparison of extracted values of σ_X^2, σ_Y^2, and $\rho(\mathbf{s})$ with the known variation components used in the Monte Carlo model, will determine the robustness and accuracy of the proposed extraction algorithm.

Similar to [17], we report the percentage error for the global variation and spatial variation relative to the Monte Carlo model values, but not for the random variation as it is indistinguishable from the added measurement noise. The error in spatial correlation function, was measured using the same metric as [17], given by $err(\rho\mathbf{s}) = \frac{\tilde{\rho}(\mathbf{s})-\rho(\mathbf{s})}{\rho(\mathbf{s})}$. The results of our experiment are given in Table I. The number of sample die N, the number of measurement location M and the amount of random noise added into the Monte Carlo model relative to the total variation $(\sigma_X^2+\sigma_Y^2+\sigma_Z^2)$ are listed in the first 3 columns. Similar to [17], by varying M and N different data-sets were generated and for each data-set we tested our algorithm for 10

TABLE I

ACCURACY AND ROBUSTNESS OF PROPOSED APPROACH

M	N	%Noise	% $err(\sigma_X^2)$	% $err(\sigma_Y^2)$	% $err(\rho(\mathbf{s}))$
64	2000	10	0.35	2.20	2.95
		50	2.61	-0.24	2.91
		100	-2.66	2.17	3.02
64	1500	10	-1.41	-0.76	3.82
		50	5.11	-0.01	3.54
		100	-1.30	2.13	3.57
64	1000	10	-1.53	2.01	4.01
		50	-0.11	-0.94	4.20
		100	2.08	1.33	4.31
64	500	10	-8.91	2.26	6.13
		50	0.37	3.27	5.84
		100	11.76	-2.08	6.40
49	1000	10	-5.98	1.94	4.21
		50	-5.22	3.85	4.30
		100	8.82	-1.12	4.34
36	1000	10	5.72	-3.39	4.06
		50	-6.44	-2.02	4.18
		100	-9.38	4.33	4.82

%, 50 % and 100% noise. It is evident from Table I that the proposed algorithm gives accurate and robust estimates of all variation components.

A comparison between the previously published results[17] on a similar experimental setup favors the proposed algorithm in terms of overall accuracy. For example, the worst case error of $err(\sigma_X^2)$, $err(\sigma_Y^2)$ and $err(\rho(\mathbf{s}))$ observed in the previous approach were 18%, 11% and 9% but it should be noted that different data-sets were used in both the cases.

VI. CONCLUSION

Based on the key concept of variogram function, we have presented a new approach to extract spatial variation models. The key advantage of the variogram function is that it provides us with a representation that is independent of the global component of variation. It allows us to directly estimate the within die component of variations and thus circumvents the need for making ergodicity assumption. In this work, we further showed that using two dimensional variogram functions allows us to model geometrically anisotropic process variation data. Additionally, for extracting process variation models in the presence of significant measurement noise, we employ *weighted least squares* regression technique, a technique which is known to be statistically more robust than the previously used ordinary least square technique. The experiment results on both Monte-Carlo models and ELM measurement data confirm the validity of the proposed approach.

APPENDIX

From the definition of variogram function, the relation between correlation function and variogram function can be derived as follows:

$$\begin{aligned} 2\gamma(\mathbf{s_1},\mathbf{s_2}) &= Variance(F(\mathbf{s_1})-F(\mathbf{s_2})) \\ &= Variance(F(\mathbf{s_1}))+Variance(F(\mathbf{s_2})) \\ &- 2Covariance(F(\mathbf{s_1}),F(\mathbf{s_2})) \end{aligned}$$

Under the stationarity assumption, we can write both variogram function and covariance function in terms of separation vector $\mathbf{v}$, therefore we have

$$\begin{aligned} \gamma(\mathbf{v}) &= Covariance(0)-Covariance(\mathbf{v}) \\ &= (\sigma_X^2+\sigma_Y^2+\sigma_Z^2)-(\sigma_X^2+\sigma_Y^2\rho(\mathbf{v})) \\ &= \sigma_Z^2+\sigma_Y^2(1-\rho(\mathbf{v})) \\ \therefore \rho(\mathbf{v}) &= \frac{\sigma_Z^2+\sigma_Y^2-\gamma(\mathbf{v})}{\sigma_Y^2} \end{aligned}$$

REFERENCES

[1] International technology roadmap for semiconductors 2005.

[2] D. M. Bates and D. G. Watts. *Nonlinear Regression and Its Applications*. New York: Wiley, 1988.

[3] D. B. Brian Cline, Kaviraj Chopra. Analysis and modeling of cd variation for statistical static timing. In *ICCAD '06 Proceedings of the Computer-Aided Design*, 2006.

[4] J. Cain and C. Spanos. Electrical linewidth metrology for systematic cd variation characterization and causal analysis. In *Proceedings of SPIE Int. Soc. Opt. Eng.*, pages 516–521, 2003.

[5] H. Chang and S. Sapatnekar. Statistical timing analysis considering spatial correlations using a single pert-like traversal. In *IEEE International Conference on Computer Aided Design*, 2003.

[6] N. A. Cressie. *Statistics for Spatial Data*. Wiley Series in Probability and Statistics, 1993.

[7] P. Friedberg, Y. Cao, J. Cain, R. Wang, J. Rabaey, and C. Spanos. Modeling within-die spatial correlation effects for process-design co-optimization. In *Int. Symp. on Quality of Electronic Design*, 2005.

[8] M. S. Handcock and J. R. Wallis. An approach to statistical spatial-temporal modeling of meteorological fields. *Journal of the American Statistical Association*, pages 368–390, 1994.

[9] Z. D. L. Another look at anisotropy in geostatistics. *Mathematical geology (Math. geol.)*, pages 453–470, 1993.

[10] F. Liu. A general framework for spatial correlation modeling in vlsi design. In *in TAU '07 international workshop on Timing issues in the specification and synthesis of digital systems*, 2007.

[11] C. A. Mack. Understanding focus effects in submicrometer optical lithography: a review. *Optical Engineering*, 32(10):2350–2362, 1993.

[12] M. Mani, A. Devgan, and M. Orshansky. An efficient algorithm for statistical minimization of total power under timing yield constraints. In *DAC '05 Proceedings of the 42nd annual conference on Design automation*, pages 309–314, June 2005.

[13] G. Nanz and L. E. Camilletti. Modeling of chemicalmechanical polishing: A review. *IEEE Trans. on Semico. Manuf.*, 8(4):382–389, 1995.

[14] M. Orshansky, L. Milor, P. Chen, K. Keutzer, and C. Hu. Impact of spatial intrachip gate length variability on the performance of high-speed digital circuits. *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, pages 544–553, May 2002.

[15] R. R. Rao, A. Devgan, D. Blaauw, and D. Sylvester. Parametric yield estimation considering leakage variability. In *DAC '04 Proceedings of the 41st annual conference on Design automation*, pages 442–447, New York, NY, USA, June 2004. ACM Press.

[16] C. Visweswariah, K. Ravindran, K. Kalafala, S. G. Walker, and S. Narayan. First-order incremental block-based statistical timing analysis. In *DAC '04 Proceedings of the 41st annual conference on Design automation*, 2004.

[17] J. Xiong, V. Zolotov, and L. He. Robust extraction of spatial correlation. In *ISPD '06: Proceedings of the 2006 international symposium on Physical design*, pages 2–9, New York, NY, USA, 2006. ACM Press.

[18] J. Xiong, V. Zolotov, and L. He. Robust extraction of spatial correlation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, April 2007.