Area Efficient Device-Parameter Estimation using Sensitivity-Configurable Ring Oscillator

Shoichi Iizuka*, Yuma Higuchi*, Masanori Hashimoto*, Takao Onoye*
*Dept. Information Systems Engineering, Osaka University, Japan
{iizuka.shoichi, hasimoto}@ist.osaka-u.ac.jp

***Abstract*—This paper proposes an area efficient device parameter estimation method with sensitivity-configurable ring oscillator (RO). This sensitivity-configurable RO has a number of configurations and the proposed method exploits this property for reducing sensor area and/or improving estimation accuracy. The proposed method selects multiple sets of sensitivity configurations, obtains multiple estimates and computes the average of them for accuracy improvement exploiting an averaging effect. Experimental results with a 32-nm predictive technology model show that the proposed method can reduce the estimation error by 49% or reduce the sensor area by 75% while keeping the accuracy.**

I. Introduction

Manufacturing variability is becoming more influential on circuit performance and parametric yield, and the traditional worst-case design with guard-band is considerably spoiling performance improvement which is supposed to be attained by device miniaturization. To tune circuit performance after fabrication and sustain parametric yield, several adaptation methods, such as voltage scaling and adaptive body bias, have been proposed [1], [2]. For efficient post-silicon adaptation, it is required to estimate for every chip how device parameters varied from their nominal values during the fabrication process. For example, when the magnitude of PMOS threshold voltage is high and NMOS threshold voltage is typical, forward body bias should be given to PMOSs, not to NMOSs. Otherwise, an increase in leakage current would be introduced.

RO-based sensors have been studied for extracting within-die ramdom variation [3] and die-to-die variation including wafer-to-wafer and lot-to-lot [4]–[7]. Essentially, for extracting n device parameters, n types of sensors that have different sensitivities to device parameters are necessary. On the other hand, ROs that have high sensitivity to a single device parameter have been proposed to decompose the measured variations, which are the mixtures of variations contributed by each device parameter, into individual device-parameter variations [6], [7]. With a set of these ROs with different sensitivities, device parameters can be estimated. An important point here is that at least n types of ROs with different sensitivities must be implemented on silicon.

A fundamental obstacle to accurate device-parameter estimation is within-die random variation since the sensors are affected by not only die-to-die variation but also random variation and the measured frequencies include uncertainties. As within-die variation becomes significant, larger-stage ROs or a number of the same ROs are required to reasonably mitigate the uncertainty of the measured frequencies. In this case, larger area overhead for the sensor implementation is inevitable.

A device-parameter extraction method aiming at smaller area overhead is proposed in [8]. This method uses a single type of RO whose sensitivities to device parameters are reconfigurable in measurement time instead of various types of ROs. By changing the configuration, we obtain more than n measured frequencies with different sensitivities from a single RO. This reconfiguration capability helps reduce silicon area necessary for the sensor since the necessary number of RO types is reduced from n to 1. Reference [8] proposed an objective function for selecting the best set of n RO configurations with the least prospective estimation error. However, the area overhead to mitigate the uncertainty originating from within-die random variation remains unsolved in [8]. For example, 100 ROs per chip were supposed in the experiments.

This paper tackles the uncertainty due to within-die random variation and proposes a device-parameter extraction method that exploits more configurations of the sensitivity-reconfigurable RO to mitigate the uncertainty. We obtain $m(>1)$ estimates for each device parameter with m sets of n sensitivity configurations, and average these m estimates to exclude the uncertainty originating from within-die random variation. Exploiting more sensitivity configurations makes the parameter estimation more robust to within-die variation and reduces the number of ROs necessary to mitigate within-die variation. Figure 1 illustrates the feature of the proposed method, where x-axis represents the number of ROs on a chip

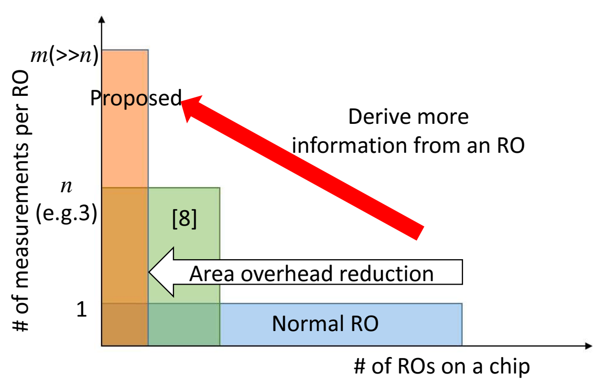

Fig. 1. Area overhead reduction by using more sensitivity-configurations.

978-1-4799-7792-5/15/$31.00 ©2015 IEEE

corresponding to silicon area, and y-axis represents the number of frequencies per RO measured for device-parameter estimation. The area of each rectangle represents the overall amount of information derived from the sensors. While keeping the overall information from the sensors roughly unchanged, [8] reduced the number of ROs by using n configurations per RO. The proposed method derives more information from a sensitivity-configurable RO and reduces the number of ROs further by using many ($\gg n$) sensitivity-configurations.

The rest of this paper is organized as follows. Section 2 explains the conventional device-parameter estimation that uses a sensitivity-configurable RO. Section 3 introduces the proposed method to attain high estimation accuracy. In Section 4, experimental results are shown to validate our approach. Finally, Section 5 concludes the discussion.

II. Conventional parameter extraction using sensitivity-configurable RO

This section describes the sensitivity-configurable ring oscillator and conventional device-parameter extraction method using a sensitivity-configurable RO.

A. Sensitivity-configurable ring oscillator

Figure 2 shows a single inverting stage composing the sensitivity-configurable RO. In this structure, the voltages given to the four terminals (INVN/P and CAPN/P) change the sensitivity of oscillation frequency to device-parameter variations. The voltages for individual terminals can be selected from, for example, V_{dd}, V_{bn}, V_{bp}, and V_{ss}, where V_{bn} is the voltage generated by the bias generator shown in Figure 3 and V_{bp} is similarly generated by the complementary circuit, which are proposed in [7]. Excluding invalid voltage assignments (e.g. stopping oscillation), 144 ($= 3^2 4^2$) data can be obtained using only the sensitivity-configurable RO in this setup. It is possible to increase the number of data further by increasing the number of assignable voltages, but in this work 144 combinations are supposed to available. In addition, we can obtain ROs that have different sensitivities to device parameters by altering the supply voltage V_{dd} given to the sensitivity-configurable RO and the bias generators.

B. Parameter extraction in [8]

We here explain how device parameters are estimated with the sensitivity-configurable RO in [8]. The upside of

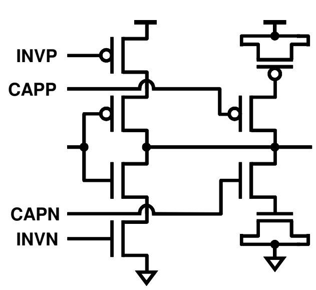

Fig. 2. Single inverting stage for sensitivity-configurable RO.

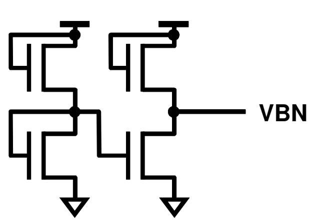

Fig. 3. V_{bn} bias generator.

Figure 4 illustrates the estimation procedure of [8]. There are s sensitivity-configurable ROs on a chip, and for each sensitivity configuration, we measure s ROs and average their oscillating frequencies to mitigate the impact of within-die variation. The average frequencies will be used in the successive estimation process. The necessary number of ROs on a chip will be experimentally discussed in Section IV.D. Now, we can obtain $144\times l$ configurations and corresponding oscillating frequencies, and we use n measured frequencies for estimating n device parameters. The simultaneous equations to be solved for estimating $\boldsymbol{\Delta G_x}$ using n configurations with corresponding measured frequencies a_1 to a_n are as follows.

$$\begin{cases} a_1 = f_1(\boldsymbol{\Delta G_x}) \\ a_2 = f_2(\boldsymbol{\Delta G_x}) \\ \quad\vdots \\ a_n = f_n(\boldsymbol{\Delta G_x}) \end{cases} \tag{1}$$

$\boldsymbol{\Delta G_x}$ is a vector representation of ΔG_x, where ΔG_x is the global variation component of parameter x. Parameter x includes, for example, threshold voltage of PMOS/NMOS V_{thp}/V_{thn} and channel length L. Besides, there are various numerical methods for solving non-linear simultaneous equations, such as Newton-Raphson method, and $\boldsymbol{\Delta G_x}$ is derived with one of them.

An important factor that determines the estimation accuracy is the selection of n equations, i.e. n sensitivity configurations. To obtain a good set of n configurations, [8] proposed an objective function that represented the prospective estimation error, and used the best set of configurations with the minimum value of the objective function. This objective function is derived based on the fact that the condition number of the sensitivity matrix represents the upper bound of the ratio of relative error norm. See [8] for the details of the objective function. In this work, this objective function is used to find not only the best set but also top m sets of configurations.

III. Proposed Method

In this section, we propose a device-parameter extraction method robust to within-die variation that derives more information from each sensitivity-configurable RO. The proposed method is expected to improve the accuracy compared to [8]. From another aspect, with the same accuracy, the proposed method can reduce the number of sensitivity-configurable ROs physically implemented on silicon.

The proposed method selects top m sets of sensitivity configurations. From each set, we obtain an estimate consisting of a set of n device parameters, and hence we obtain m estimates in total. The proposed method then outputs the average of m estimates as the final estimation result. The detailed procedure illustrated in the underside of Figure 4 is as follows.

For s ROs on a chip, we measure oscillating frequencies with the sensitivity configurations included in the top m sets. The proposed method then computes the average of s measured oscillating frequencies for each sensitivity configuration, and this average frequency is given to the simultaneous equations of Eq. (1), which is similar to the conventional method. Functions f_1 to f_n are constructed beforehand with circuit simulations and curve fitting. The proposed method has m sets of sensitivity configurations, i.e. m sets of simultaneous equations. By solving each set of simultaneous equations, we obtain an estimate which consists of n device parameters. We finally calculate the average value of m estimates for each device parameter.

An assumption of this approach is that the estimation errors of m estimates are reasonably random and the estimation errors can be averaged out. This averaging effect is often used to mitigate random variation effects; for example a larger number of ROs are used or a larger stage RO is used. Intuitively speaking, the impact of the variation of individual transistors on the oscillating frequency depends on the RO configuration. For example, the channel length variation of MPL1 affects the operating frequency when MPL2 is on. On the other hand, when MLP2 is off, this channel length variation does not impact the oscillating frequency. Consequently, the impact of each transistor variation might be able to be averaged by using a larger number of configurations. Besides, the randomness of m estimates is difficult to theoretically clarify, and hence we will experimentally demonstrate how many estimates are adequate to obtain the averaging effect in the next section.

Another condition necessary to make the proposed method effective is that m sets of sensitivity configurations should have small prospective estimation errors, i.e. good values of the objective function. If the m-th set of sensitivity configuration has large prospective estimation error, its estimates of device parameters are not helpful to mitigate the uncertainty of estimated device parameters. On the other hand, in our 32 nm experiments which will be shown in the next section, there are 71 sets of sensitivity configurations in 5% range of the best objective function value. Thus, we conclude that this second condition on the availability of good sets of sensitivity configuration is satisfied.

IV. Experimental results

This section experimentally confirms the accuracy improvement and sensor area reduction thanks to the proposed device-parameter estimation.

A. Experimental setup

One hundred chips were virtually fabricated on the basis of Monte Carlo simulation, and their die-to-die variations were estimated. Here, a 32-nm predictive technology model [9], [10] with nominal threshold voltage was used for evaluation. The number of stages of the sensitivity-configurable RO (Figure 2) was set to seven. The number of ROs on a chip, s, was set to 100, and the relation between s and estimation accuracy will be discussed in Section IV.D. Also, three discrete voltages (0.7, 0.9, 1.1[V]) were given to the sensitivity-configurable RO and bias generators as the supply voltage (l = 3). In this setup, there were 144×3=432 configurations available for each RO. All the sensor outputs in each chip were evaluated with circuit simulator [11].

It is assumed that the variational device-parameters to be extracted are threshold voltages of NMOS and PMOS, V_{thn} and V_{thp}, and channel length, L, which means n = 3, and the variability is composed of two components, which are global and random variations. The global variation here means die-to-die variation which causes the same variation offset to all the transistors on a chip, while the random variation corresponds to within-die variation which is different for transistor by transistor. Then, the offset of a device-parameter x from its nominal value ΔV_x is expressed by Eq. (2),

$$\Delta V_x = \Delta G_x + \Delta R_x, \tag{2}$$

where ΔR_x denotes random variability of parameter x, and x could be V_{thn}, V_{thp} or L. In what follows, it is assumed that $\sigma_{\Delta G_{Vthn/p}} = \sigma_{\Delta R_{Vthn/p}} = 20\text{mV}$ and $\sigma_{\Delta G_L} = \sigma_{\Delta R_L} = 1\text{nm}$. The amounts of random variations are size-dependent [12], and the standard deviations above are set for $L = 32\text{nm}$ and $W = 256\text{nm(NMOS)}/328\text{nm(PMOS)}$. In this experiment, spatially correlated variations are not included. We assume that all ROs are placed in a small area enough to ignore spatially correlated variations. In the Monte Carlo simulation, the size-dependency is considered by scaling the standard deviation according to the Pelgrom model.

B. Validation of the proposed estimation method

As described in Section III, the proposed method averages m estimates obtained from m sets of sensitivity configurations. The assumption behind the proposed method is that m estimates are randomly distributed. If all the m estimates are shifted in the same direction, taking the average is not a help to mitigate the estimation error. We here exemplify the error distribution of m ΔG_{Vthn} estimates, where m is 50.

Figure 5 shows the estimation error between the estimated ΔG_{Vthn} and the true value for the top 50 sets of sensitivity configurations. X-axis is the objective function value for each set of sensitivity configurations. Here, we have 100 chips, and then the estimation error is the average among 100 chips. We can see the device-parameter extraction using the best set with the minimum objective function value does not necessarily give the most accurate estimate. This is because the objective function represents not the average but the upper bound of the estimation error. More important thing to observe here is that the estimates are randomly distributed in both positive and negative directions. Therefore, we can expect that computing the average of m estimates improves the estimation accuracy.

From the opposite point of view, the device-parameter estimation using a single set of sensitivity configuration is risky, i.e. a sensitivity configuration included in the set for device parameter estimation might be excessively affected by random variation. To quantitatively show this risk, we evaluated the maximum, average and minimum estimation errors among 50 ΔG_{Vthn} estimates (Figure 6). In some chips, the maximum estimation errors were large and one of them exceeded 9 mV while the corresponding average error among 50 estimates was 3.5 mV. Thus, the device parameter estimation based on a

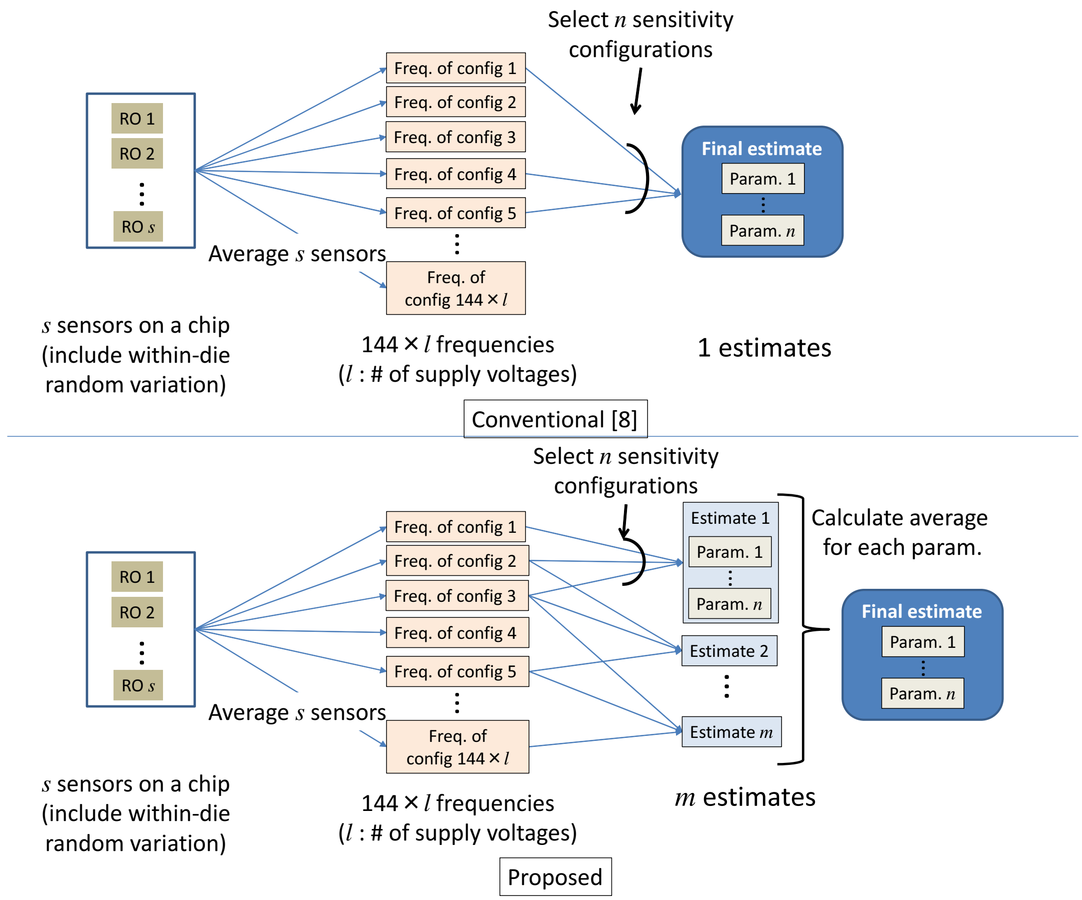

Fig. 4. Estimation procedures of conventional [8] and proposed methods.

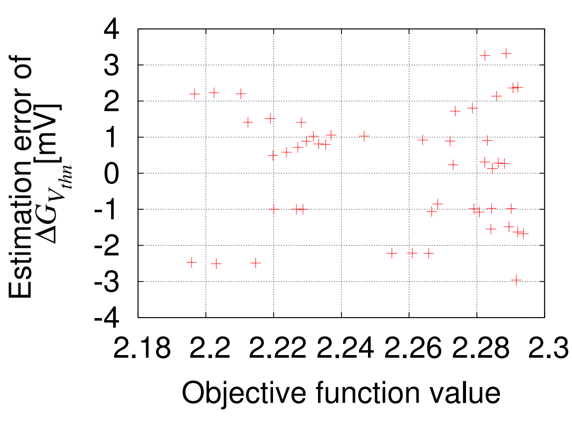

Fig. 5. Error distribution of ΔG_{Vthn} estimates for top 50 sets of sensitivity configurations.

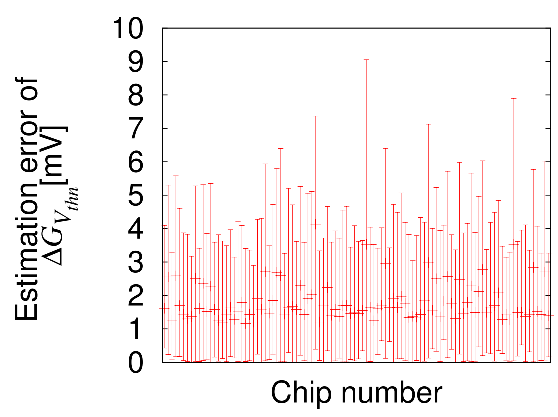

Fig. 6. Maximum, average and minimum errors of ΔG_{Vthn} for 100 chips.

single set of sensitivity configuration is not robust, and the averaging is effective to avoid such large estimation errors.

Next, we evaluate how many estimates are adequate for averaging in the proposed method. As mentioned before, the top 50 sets of sensitivity configurations have similar values of the objective function. Therefore, in this experiment, we picked up m sets randomly from the 50 sets, computed m estimates and obtained the average of m estimates. We repeated this 100 times and calculated the mean and standard deviation (σ) of the estimation errors. Figure 7 shows the mean and the mean+3σ of the estimation error. Here, the estimation error is the L2-norm of ΔG_L, ΔG_{Vthn} and ΔG_{Vthp} estimation errors

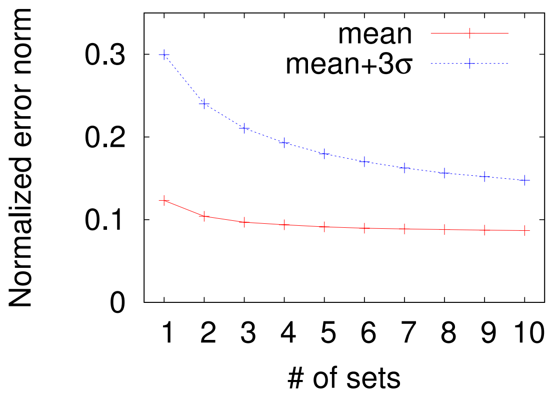

Fig. 7. Mean and mean+3σ of the estimation error.

normalized by $\sigma_{\Delta G_L}$, $\sigma_{\Delta G_{Vthn}}$ and $\sigma_{\Delta G_{Vthn}}$, respectively. We can see that the averaging is effective to reduce both the mean and the mean+3σ, and especially the mean+3σ is significantly reduced. Just focusing on the mean, 3 sets could be adequate, but when placing on emphasis on the worst case, 10 sets or more can be reasonable.

C. Estimation accuracy evaluation

Next, we evaluated the accuracy improvement from [8]. Table I shows the averages of the absolute estimation errors of $\boldsymbol{\Delta G_x}$ from the variations given to each chip. Here, 10 sets of sensitivity configurations were used in the proposed method ($m = 10$). There are two rows of C1 and C2 for [8]. [8] proposed an iterative estimation method for coping with non-linearity of functions f_1 to f_n in Eq. (1), and the iterative method uses another set of simultaneous equations depending on the first-round estimates. C2 corresponds to the iterative estimation method. Besides, this iterative estimation was not applied to the proposed method similar to C1.

We can see the normalized error norm, which represents the overall estimation error to be minimized, is reduced from 0.185 to 0.0943 by 49% compared to C1 and the error norm of the proposed method is smaller than that of C2 by 23%. This indicates that the proposed method derives more information from sensitivity-reconfigurable ROs and efficiently exploits it for achieving more accurate device-parameter extraction.

D. Area overhead reduction

The accuracy improvement shown above can be translated into the reduction of the sensor area. To evaluate this, the device-parameter extraction was conducted supposing different numbers of sensors were available on a chip. Figure 8 shows the relation between the estimation error norms and the number of sensors on a chip, s. When s is larger than 10, the proposed method improved the estimation accuracy. Focusing on the estimation error of [8] with 100 sensors, the proposed method can achieve the same accuracy with 25 sensors, which results in 75% silicon area reduction.

TABLE I
ESTIMATION RESULTS.

Estimate condition	$\Delta G_{V_{thn}}$ [mV]	$\Delta G_{V_{thp}}$ [mV]	ΔG_L [nm]	Normalized error norm
C1: Conventional [8] Initial	2.48 (0.1240)	1.26 (0.0630)	0.1094 (0.1094)	0.185
C2: Conventional [8] Second	1.39 (0.0695)	1.36 (0.0680)	0.0489 (0.0489)	0.123
Proposed	1.09 (0.0544)	1.05 (0.0523)	0.0383 (0.0383)	0.0943
Err. Reduction from C1	56%	17%	65%	49%
Err. Reduction from C2	22%	23%	22%	23%

Values in brackets are the errors normalized by respective standard deviations $\sigma_{\Delta G_x}$.

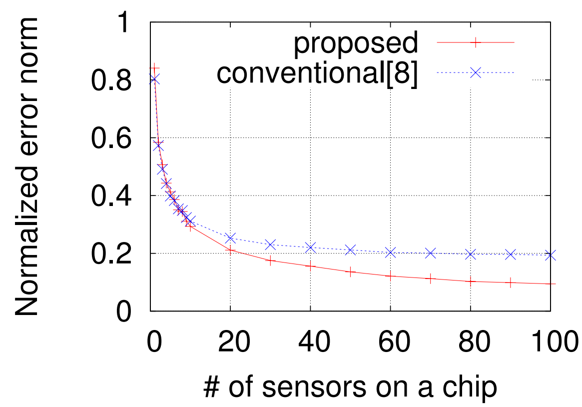

Fig. 8. Sensor area reduction by using more sensitivity-configurations.

V. CONCLUSION

This paper proposed an accuracy improvement method for device-parameter extraction with the sensitivity-configurable ring oscillator. The proposed method selects m sets of sensitivity configurations and computes the average of m estimates for accuracy improvement thanks to the averaging effect. Experimental results using virtually fabricated test chips in 32-nm process showed that the proposed method reduced the estimation error by 49% or the sensor area by 75% with the same accuracy.

ACKNOWLEDGEMENTS

This work was supported by NEDO and partly by STARC.

REFERENCES

[1] S. Martin, K. Flautner, T. Mudge, and D. Blaauw, "Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads," in *Proc. ICCAD*, 2002, pp. 721–725.

[2] J. Tschanz, J. Kao, S. Narendra, R. Nair, D. Antoniadis, A. Chandrakasan, and V. De, "Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage," *IEEE JSSC*, vol. 37, no. 11, pp. 1396–1402, 2002.

[3] B. Das, B. Amrutur, H. Jamadagni, N. Arvind, and V. Visvanathan, "Within-die gate delay variability measurement using reconfigurable ring oscillator," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 22, no. 2, pp. 256–267, May 2009.

[4] M. Bhushan, M. Ketchen, S. Polonsky, and A. Gattiker, “Ring oscillator based technique for measuring variability statistics,” in *Proc. ICMTS*, 2006, pp. 87–92.
[5] L. Pang and B. Nikolic, “Measurements and Analysis of Process Variability in 90 nm CMOS,” *IEEE JSSC*, vol. 44, no. 5, pp. 1655–1663, 2009.
[6] I. A. K. M. Mahfuzul, A. Tsuchiya, K. Kobayashi, and H. Onodera, “Variation-sensitive monitor circuits for estimation of global process parameter variation,” *Semiconductor Manufacturing, IEEE Transactions on*, vol. 25, no. 4, pp. 571–580, Nov 2012.
[7] B. Wan, J. Wang, G. Keskin, and L. T. Pileggi, “Ring Oscillators for Single Process-Parameter Monitoring,” in *Proc. Workshop on Test Structure Design for Variability Characterization*, 2008.
[8] Y. Higuchi, K. Shinkai, M. Hashimoto, R. Rao, and S. Nassif, “Extracting Device-Parameter Variations using a Single Sensitivity-Configurable Ring Oscillator,” *Proc. ETS*, pp. 106–111, May. 2013.
[9] “Predictive Technology Model (PTM),” http://ptm.asu.edu/.
[10] W. Zhao and Y. Cao, “New generation of Predictive Technology Model for sub-45nm early design exploration,” *IEEE Transactions on Electron Devices*, vol. 53, no. 11, pp. 2816–2823, 2006.
[11] *Spectre Circuit Simulator*, Cadence Design Systems, Inc.
[12] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, “Matching properties of MOS transistors,” *IEEE JSSC*, vol. 24, no. 5, pp. 1433–1439, 1989.