Process-sensitive Monitor Circuits for Estimation of Die-to-Die Process Variability

Islam A.K.M Mahfuzul
Department of Communications and Computer Engineering
Kyoto University
mahfuz@vlsi.kuee.kyoto-u.ac.jp

Akira Tsuchiya
Department of Communications and Computer Engineering
Kyoto University
tsuchiya@vlsi.kuee.kyoto-u.ac.jp

Kazutoshi Kobayashi
Department of Electronics
Kyoto Institute of Technology
kobayasi@kit.ac.jp

Hidetoshi Onodera
Department of Communications and Computer Engineering
Kyoto University
onodera@vlsi.kuee.kyoto-u.ac.jp

ABSTRACT

In this paper, we propose a set of ring oscillators (ROs) to estimate D2D variation of MOS threshold voltage and gate length. First, we show a design guideline on designing ROs with enhanced process sensitivities. We propose a set of ROs for process parameter estimation. We then develop an estimation method using a linear model to extract the variations from the measured frequencies. Simulation results confirm that our proposed circuits are able to extract process parameters in the conditions of error in the measurement. We build test chips to confirm the validity of our proposed ROs. We extracted process parameter variation using our monitor circuits. Variation results satisfy the variation range in Process Control Module (PCM) data and are within the corner model. The proposed ROs can be built on-chip to monitor process parameter variation in real time.

1. INTRODUCTION

As the scaling of Silicon CMOS technology progresses, variation in transistor performance has been becoming serious problem. In the 65nm process and beyond, this variability plays a major role in chip performance. Typically, circuit designers consider the process variations into account and design their circuits so that the performances satisfy the worst corners. Such a design methodology often results in "over design" due to unnecessary margining in defining the corners. Thus, the performance of the resulted chip becomes suboptimal [1].

The variation of CMOS transistor performance can be divided into die-to-die (D2D) and within-die (WID) variation. As the technology scaling continues, WID variation is becoming more significant [2]. But WID variation is mainly random and thus it is possible to reduce WID variation by designing the circuits to suppress the random effect (for example, by increasing the number of stages). On the other hand, D2D variation effects the performances of all the transistors in a chip in the same direction (fast or slow). Therefore, the effect of D2D variation is more significant in determining the chip performance such as speed and power consumption.

Several methods have been proposed such as Adaptive Body Bias (ABB) and Adaptive Supply Voltages (AVS) in order to reduce the design margin and control the performance of the chip after production [3, 4]. D2D Variation in chip performance occurs mainly due to MOSFET gate length and threshold voltage variation [5]. In order to apply body bias to each MOSFET, the amount of variations of these process parameters is needed. For on-chip body biasing we need monitor circuits for on-chip parameter extraction.

Recently, several monitor circuits have been proposed to monitor the variation of a particular process variation [6, 7, 8]. These monitor circuits use either device arrays [6, 7] or op-amps [8] and thus require considerable area and measurement access that makes them unsuitable for on-chip parameter extraction. A method to calculate the saturated current of each MOS is also proposed [9]. However, to compensate the performance of the chip optimally, we need to know the variation of each process parameter.

In this paper, we propose a set of RO monitor circuits to estimate D2D variation of MOS threshold voltage (V_{THP} and V_{THN}) and gate length (L) from the measured value of frequency from chip. The sensitivities to process parameter variation of our monitor circuits are being enhanced by

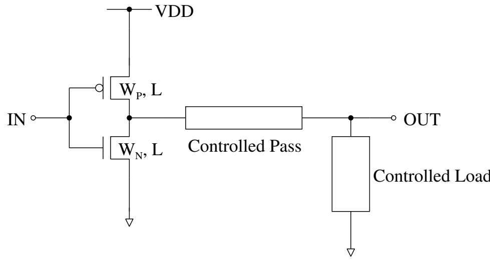

Figure 1: Ways of modifying the standard inverter cell

modifying the structure of the inverter. We show a general guideline on designing ROs with process-enhanced sensitivities. We then develop a linear model for the frequency of RO and use the model to extract process parameters by solving a reverse equation. To prove our monitor circuits, we design test chips. We are able to extract the variation of each process parameter from the measured values. Our estimation results fit within the range of PCM data and are within the corner model.

The remainder of this paper is constructed as follows. In chapter 2 we discuss several techniques to design ROs with enhanced process sensitivity and propose a set of ROs to estimate process parameter variation. In chapter 3 we discuss the estimation method to extract process parameter variation using our proposed circuits. In chapter 4 we discuss our test chip and the estimation results. Finally, in chapter 5 we conclude our discussion.

2. RO MONITOR CIRCUITS FOR PROCESS PARAMETER ESTIMATION

ROs are widely used to monitor WID and D2D variation. However, frequency of a RO is effected by many process parameter variations simultaneously. But, if we can design ROs who have different sensitivities to each process parameter then we can get the amount of process parameter variation by comparing the frequency changes of the ROs using an appropriate model. In these section, we show a general guideline on how to design a RO with enhanced process sensitivity and propose a RO set for process variation estimation.

2.1 General Guideline to Design ROs with Enhanced Sensitivities

A general guideline to design ROs with enhanced sensitivities is demonstrated in Fig.1. We can modify the transistors in a inverter or control the passing current while charging and discharging the output load to get enhanced process sensitivity. We can control the output load also to change the sensitivities. Changing the gate length will effect the sensitivity to gate length variation. Modifying gate width changes the current flow while charging and discharging and thus the sensitivities change. Below are some examples on how to realize inverter structure in Fig.1.

2.1.1 RO with Parallel MOS

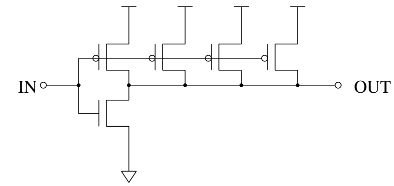

Figure 2: PMOS RICH inverter

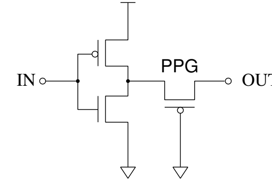

Figure 3: Inverter with a PMOS pass transistor

To make RO frequency more sensitive to NMOS (PMOS), we can increase the PMOS (NMOS) current so that the frequency becomes more sensitive to NMOS (PMOS) current. We can change MOS current by changing gate width. But, considering the standard cell structure, we propose to increase finger numbers or make several PMOSs (NMOSs) parallel. But, increasing finger numbers may include unexpected strain effect. So using parallel MOSs are better choices. [9] uses inverters with parallel MOSs to successfully extract the MOS current. Fig.2 shows an example of designing an inverter where PMOS is 4 times larger than the standard cell. We call this cell as PMOS RICH inverter. From simulation results we get a 20% increase in the sensitivity to NMOS threshold voltage comparing to a standard inverter cell RO.

2.1.2 RO with Pass Transistor

We can control the passing current while charging and discharging by putting a pass transistor in output as shown in Fig.3. Fig.3 shows an inverter cell with a PMOS pass transistor in the output. In order to increase the threshold voltage sensitivity, the pass transistor in Fig.3 is kept on so that it performs in the linear region. In this way, current through the pass transistor becomes highly sensitive to threshold voltage change. From simulation result, we get 10 times higher sensitivity to V_{THP} variation and 3 times higher sensitivity to gate length variation comparing to the standard inverter cell RO. In case of V_{THN} variation, as the voltage of output node of the inverter does not go down to zero due to the voltage drop in PMOS, leak current occurs. So, while loading the output load leak current flows through NMOS and it takes longer time to charge the output load. Thus, we get a reversed relationship between output frequency and V_{THN} variation.

2.1.3 RO with MOS Controlled Load

We can design a controlled load inverter by adding a pass transistor and a load in the output as shown in Fig.4. This kind of structure partially cancels the effect of the threshold voltage variation of MOS. In Fig.4 when the threshold voltage of NMOS (PMOS) decreases, the value of NMOS

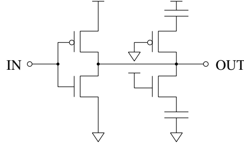

Figure 4: Inverter with MOS controlled Loads

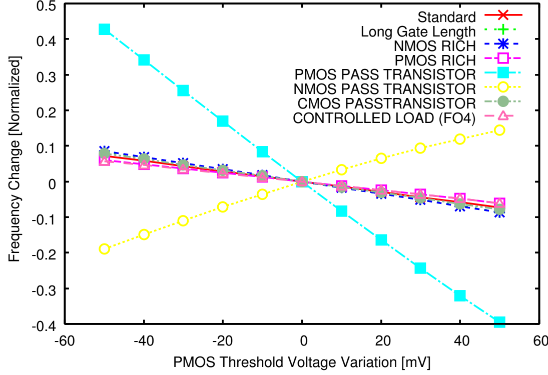

Figure 5: Frequency change for several ROs according to PMOS threshold voltage variation

(PMOS) resistance decreases. As a result, the effective load to the inverter decreases and delay becomes longer. Thus, the effect of decrease in the threshold voltage will be canceled. From simulation result, we get 15% decrease in the sensitivity of V_{THP} comparing to the standard inverter cell RO.

2.2 Combining ROs for Process Parameter Estimation

We can combine several ROs and compare their frequency changes to extract each process parameter variation. We can get the sensitivity vectors from the above discussed ROs. If we can get three ROs whose sensitivity vectors are perpendicular to each other then we can get each process parameter variation from the frequency changes of the ROs. Fig.5 shows the frequency change according to V_{THP} variation for different ROs. From Fig.5, we can get ROs with different sensitivities to V_{THP}. ROs with long gate length inverter, NMOS RICH inverter and PMOS pass transistor inverter have higher sensitivities while ROs with PMOS RICH inverter and controlled load inverter have lower sensitivities than the standard inverter RO. We get similar sensitivity changes for V_{THN} and L. Table 1 shows the sensitivities of various types of ROs to the process parameters. These sensitivities cab be calculated through circuit simulation. Table 2 shows the angles between the sensitivity vectors of the ROs. From table 2, we see that RO set of standard cell RO, RO with PMOS pass transistor and NMOS pass transistor form larger angles between them. RO set of CMOS pass transistor, PMOS pass transistor and NMOS pass transistor also form large angles between them. By considering the area and complexity of design we prefer standard cell RO instead of CMOS pass transistor. Another reason for choosing standard cell RO is that we do not want to modify standard

Table 1: Sensitivity vectors of various ROs

RO Type	Gate Length [nm]	K_P	K_N	K_L
Standard	60	−0.036	−0.032	−0.027
NMOS RICH	60	−0.043	−0.031	−0.030
PMOS RICH	60	−0.030	−0.038	−0.027
NMOS Passgate	60	0.084	−0.41	−0.053
PMOS Passgate	60	−0.21	0.019	−0.081
CMOS Passgate	60	−0.039	−0.036	−0.026
Controlled Load	60	−0.029	−0.030	−0.030

cells too much. So, our proposed set of ROs to estimate the process variations are standard RO, RO with PMOS pass transistor and RO with NMOS pass transistor. We name this set of RO as RO set #1. As the angles between the vectors are large, we will show in the next chapter that our monitor circuits can extract process variation correctly even in the presence of some measurement error.

3. METHOD TO EXTRACT PROCESS PARAMETER VARIATION

An appropriate method is needed to extract the process parameter variation from the ROs discussed in the previous chapter. In this section, we propose an estimation method based on a linear model for the frequency of RO. Based on our method, we show that correct estimation of process parameter variation is possible with our proposed set of ROs even under the influences of some other effects.

3.1 RO Frequency Model

As the frequencies in the Fig.5 changes smoothly and almost linearly, we have developed a linear model for the frequency of RO to the variation of V_{THP}, V_{THN} and L. Eq.(1) is our linear model.

$$\frac{\Delta f}{f_0} = K_{\text{P}} \frac{\Delta V_{\text{THP}}}{\Delta_0 V_{\text{THP}}} + K_{\text{N}} \frac{\Delta V_{\text{THN}}}{\Delta_0 V_{\text{THN}}} + K_{\text{L}} \frac{\Delta L}{\Delta_0 L} \tag{1}$$

In Eq.(1), K_{P}, K_{N} and K_{L} are the sensitivity coefficients for V_{THP}, V_{THN} and L variation respectively. Here, Δ refers to the change of the process parameter and Δ_0 refers to a fixed amount of change to normalize the variation. Δf is the shift of frequency according to a nominal value which gives the D2D variation. f_0 is the nominal value for a particular RO.

Although, this linear model is very simple, it has several merits. First of all, the model is easy to apply and the sensitivities can be calculated through circuit simulation. Secondly, in case of the nonlinear nature of the circuit frequency, this model can be applied as many times to improve the accuracy.

3.2 Extraction of Process Parameter Variation

We need three equations to extract the variation of three process parameters of our interest. We get the three equations from our three different ROs. The amount of variation of each process parameter can be extracted by solving the reverse problem presented by Eq.(2).

$$\vec{V} = \mathbf{S}^{-1}\vec{F} \tag{2}$$

Table 2: Angles between the sensitivity vectors of two ROs [Unit: degree]

	NMOS RICH	PMOS RICH	Standard	NMOS Passgate	PMOS Passgate	Controlled Load	CMOS Passgate
NMOS RICH	0.0	13.9	4.9	64.5	38.9	9.8	5.3
PMOS RICH	13.9	0.0	10.3	50.6	52.8	6.4	8.9
Standard	4.9	10.3	0.0	60.8	43.0	5.1	1.9
NMOS Passgate	64.5	50.6	60.8	0.0	103.1	56.0	59.5
PMOS Passgate	38.9	52.8	43.0	103.1	0.0	48.1	44.1
Controlled Load	9.8	6.4	5.1	56.0	48.1	0.0	4.7
CMOS Passgate	5.3	8.9	1.9	59.5	44.1	4.7	0.0

where

$$\vec{V} = \begin{pmatrix} \Delta V_{\mathrm{THP}} \\ \Delta V_{\mathrm{THN}} \\ \Delta L \end{pmatrix}, \mathbf{S} = \begin{pmatrix} k_{\mathrm{P1}} & k_{\mathrm{N1}} & k_{\mathrm{L1}} \\ k_{\mathrm{P2}} & k_{\mathrm{N2}} & k_{\mathrm{L2}} \\ k_{\mathrm{P3}} & k_{\mathrm{N3}} & k_{\mathrm{L3}} \end{pmatrix}, \vec{F} = \begin{pmatrix} \frac{\Delta f_1}{f_{01}} \\ \frac{\Delta f_2}{f_{02}} \\ \frac{\Delta f_3}{f_{03}} \end{pmatrix}$$

Here, $\vec{V}$ is the vector for the variations of V_{THP}, V_{THN} and L. $\mathbf{S}$ is the sensitivity matrix and $\vec{F}$ is the vector for the frequency shift from the nominal value. $(K_{\mathrm{P1}}, K_{\mathrm{N1}}, K_{\mathrm{L1}})$ represents the sensitivity vector for the first RO. Similarly, $(K_{\mathrm{P2}}, K_{\mathrm{N2}}, K_{\mathrm{L2}})$ and $(K_{\mathrm{P3}}, K_{\mathrm{N3}}, K_{\mathrm{L3}})$ represents sensitivity vectors for the second and the third RO. By considering the effects of other parameters and the error in the measurement, the vectors should separate from each other sufficiently. In case of nonlinear nature in the frequency changes of the RO, model Eq.(2) can be applied several times to improve the accuracy.

3.3 Validity Check

We made several experiments by simulation in 65nm process to confirm the validity of our proposed ROs. To show the validness of our proposed set of ROs, we demonstrate the estimation result of process parameter variations under the influence of some other effects and show that our circuits are able to extract correct value even when there are other effects. We then compare the results with the results obtained from a different set ROs whose angles between the vectors are not so large. We choose standard cell RO, PMOS RICH RO and NMOS RICH RO to form a different set of ROs. We name this RO set as RO set #2.

3.3.1 Experiment Setup

First, we assume that we know the amount of variation of each process parameter. We simulate the RO frequencies under the condition of known process parameter variations. We then apply our method presented in Eq.(2) and extract process parameter variations with our two sets of ROs. If the estimated variations match with the known variations, then the estimation becomes correct.

To demonstrate the validness of our proposed ROs, we show the estimation results under the effects of parameters other than our parameter of interest. In order to do that, we add 1% error to each RO frequency and check whether our proposed circuits are able to extract the process variation correctly.

3.3.2 Experiment Results

We checked the validity of our circuits at all over the process space. Here, we show the results for the points at

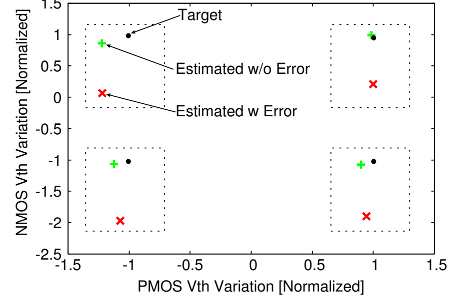

Figure 6: Estimated results of V_{TH} variation at four points of process space with RO set #2

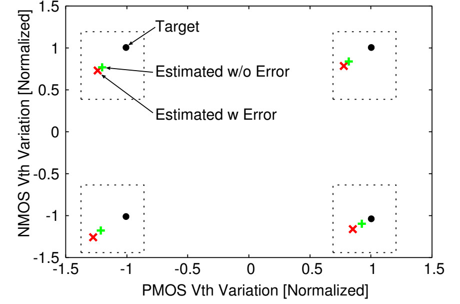

Figure 7: Estimated results of V_{TH} variation at four points of process space with RO set #1

$(\pm 1, \pm 1, \pm 1)$ where each indicates the σ variation for V_{THP}, V_{THN} and L variation. Fig.6 shows the estimation result for our first set of ROs. In Fig.6 X-axis refers to V_{THP} variation and Y-axis refers to V_{THN} variation. "•" points in the figure are the target values that we want to estimate. "+" points are the estimated values of process variation when there is no error in the RO frequency. "×" points are estimated results when there is 1% error in RO frequency. We want the estimated results to match with the "•" points. From Fig.6 we can see that the target variation is achieved correctly in an ideal condition where there is no error in the frequency. However, with an error of 1% in the frequency, the estimated results deviate from the real variation largely. This is because the angles between the sensitivity vectors in RO set #2 are not large enough. So, using single MOS RICH inverter cell ROs are not suitable for correct process parameter extraction.

Next, we make the same experiment as above with RO set #1. Estimation results are shown in Fig.7. From Fig.7 we can see that the estimated results for both the cases are quite similar. Although there are some errors initially when when there is no error in the frequency, this error is due to the non-linear nature of the frequency change according to process parameter variation. We can overcome this problem by applying our model several times. The important thing here is how the estimated results variate when there are effects of other parameters rather than our interest. We can see that in spite of the same error in the frequency, process variations are correctly estimated. So, it is shown that our proposed circuits are able to extract process variations correctly even in the presence of some error in the RO frequency. In the next chapter, we show the result from real chip and confirm the validity of our circuits.

4. TEST CHIP IN 65NM PROCESS

One of the major challenges for our monitor circuits is to demonstrate their validity in real chips. We designed test chips in 65nm process to check the validation of our proposed monitor circuits. In this section we describe the structure of our test chip. Then, we show the validity of our proposed cirucits by estimating the D2D variation from our test chip.

4.1 Chip Design

We designed the ROs of various types shown in Table 1. We put our ROs in an array of 20×20. Fig.9 shows the conceptual layout of our chip. Each die contains 400 ROs of the same type. Fig.10 shows the block diagram of our test chip. We used on-chip counter to measure RO frequency. Each RO is 7 staged RO. The purpose of this kind of design is to check the random variation first, and then average the frequencies to cancel the random effects. In this way, we can get both the WID and D2D variation and conclude on the number of stages needed for on-chip process monitoring.

4.2 Analysis of Measured Frequency

Table.3 shows the WID variation and D2D variation for each RO in our chip. From Table.3 we see that the amount of variation is enhanced in ROs with pass transistor. This confirms that our design is correct.

4.3 Estimation from Measured Frequency

We estimated the global variation of V_{THP}, V_{THN} and L using our proposed monitor circuits for 20 chips. Fig.11 shows the estimation results . In Fig.11, we set the variation range of V_{THP} and V_{THN} in PCM data from -1 to $+1$ respectively and normalized the estimated variations with the range in order to compare them with PCB data. From Fig.11, we can see that our estimated results almost fit within the range of variation in PCM data. We have two samples of estimation result for V_{THN} which is outside the range. As the number of sample in the PCM data is very limited, we are thinking that some of our chips are located far from the PCMs. The variation range of V_{THP} and V_{THN} are 1/3 of the variation range in the corner model. This confirms that our estimation results are correct. For the gate length, we have got the estimation result from -2nm to 1.3nm for L of 60nm.

Another way to confirm our estimation results is to check whether the frequencies of all the ROs can be reproduced

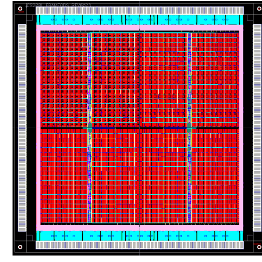

Figure 8: Test chip in 65nm process

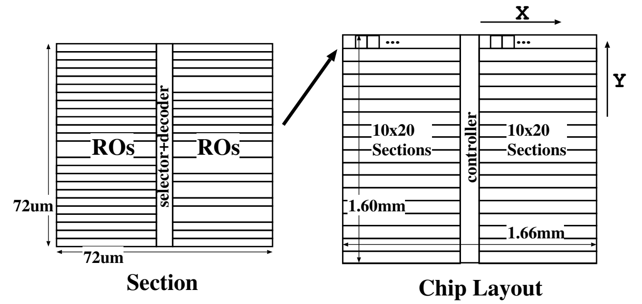

Figure 9: Conceptual layout of the test chip

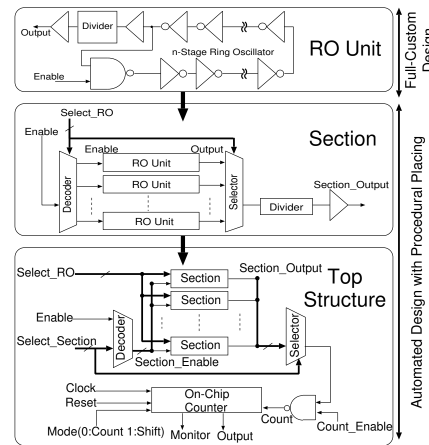

Figure 10: Block diagram of the test chip

Table 3: WID and D2D variation of RO frequency

RO Type	L [nm]	WID Variation (σ/μ) [%]	D2D Variation (σ/μ) [%]
Standard	60	1.44 - 1.71	1.27
NMOS RICH	60	1.17 - 1.60	1.13
PMOS RICH	60	1.41 - 1.67	1.34
NMOS Passgate	60	8.65 - 15.6	5.82
PMOS Passgate	60	2.80 - 5.88	2.78
Controlled Load	60	1.47 - 1.85	1.21
CMOS Passgate	60	1.21 - 1.54	1.41

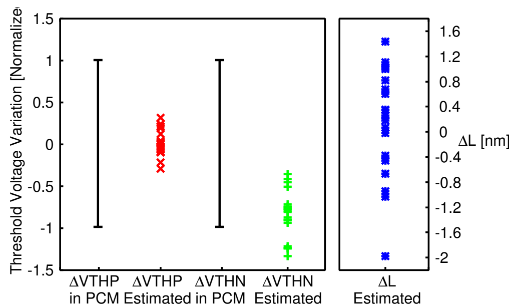

Figure 11: Estimated V_{THP}, V_{THN} and L variations for all the chips

correctly by simulation using our estimation results. In order to do that, we simulated the RO frequencies of various types at our estimated process space and compared the simulated frequencies with the measured values from the chip. Table 4 shows the comparison between the measured values (average) and the simulated values for a single chip for different ROs. From Table 4, we see that error between measured value and simulated value for all the ROs are within an error of maximum 1.2% except for the PMOS RICH RO. The reason for this mismatch is due to layout problem. In our layout for the PMOS RICH RO, distance between pwells of NMOS is longer than the other cells. Therefore, unexpected strain effect is being involved here. We need to make sure that the distance between the wells are identical to all the cells. Besides this, our estimated result could reproduce the measured frequency for all the chips and thus it can be confirmed that our monitor circuits work well. For on-chip monitoring we need long stage ROs. So, our future work is on determining the number of stages needed for on-chip D2D process monitoring.

Table 4: Comparison between measured values and simulated values using estimated process variation [Unit: MHz]

RO Type	Measurement	Simulation	Error[%]
NMOS RICH	767.1	765.2	−0.25
PMOS RICH	748.4	785.4	4.9
Standard	1236	1236	0.00
NMOS Passgate	203.6	204.0	0.20
PMOS Passgate	264.8	264.6	−0.08
Controlled Load	584.5	591.6	1.2
CMOS Passgate	574.2	578.5	0.75

5. CONCLUSIONS

In this paper, We show a design guideline on designing ROs with enhanced process sensitivity. We propose a set of RO for estimation of PMOS threshold voltage, NMOS threshold voltage and gate length variation. We develop a method based on linear model to extract process parameter variation from our proposed set of RO. We confirm the validity of our proposed circuits from simulation based experiments. We then design test chip to verify our circuits. Our estimation results from real chip satisfy the PCM data and are within the corner model. In future, we will give a guideline on the number of stages needed for on-chip measurement.

6. ACKNOWLEDGMENTS

The authors acknowledge the support of VLSI Design and Education Center (VDEC), the University of Tokyo. The authors are also grateful to STARC, E-shuttle Inc. and Fujitsu Ltd.

7. REFERENCES

[1] B. Calhoun, "Digital circuit design challenges and opportunities in the era of nanoscale cmos," *Proceedings of the IEEE*, vol. 96, pp. 343–365, March 2008.

[2] H. Onodera, "Variability: Modeling and its impact on design," *Proceedings of the IEEE*, pp. 701–704, 2008/12.

[3] T. Mudge, K. Flautner, D. Blaauw, and S. M. Martin, "Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads," *Computer-Aided Design, International Conference on*, vol. 0, pp. 721–725, 2002.

[4] J. Tschanz, S. Narendra, R. Nair, and V. De, "Effectiveness of adaptive supply voltage and body bias for reducing impact of parameter variations in low power and high performance microprocessors," *Solid-State Circuits, IEEE Journal of*, vol. 38, pp. 826–829, May 2003.

[5] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and microarchitecture," in *DAC '03: Proceedings of the 40th annual Design Automation Conference*, pp. 338–342, 2003.

[6] K. Agarwal, S. Nassif, F. Liu, J. Hayes, and K. Nowka, "Rapid characterization of threshold voltage fluctuation in mos devices," *Microelectronic Test Structures, 2007. ICMTS '07. IEEE International Conference on*, pp. 74–77, March 2007.

[7] M. Bhushan, M. Ketchen, S. Polonsky, and A. Gattiker, "Ring oscillator based technique for measuring variability statistics," *Microelectronic Test Structures, 2006 IEEE International Conference on*, pp. 87–92, March 2006.

[8] B. Ji, D. Pearson, I. Lauer, F. Stellari, D. Frank, L. Chang, and M. Ketchen, "Operational amplifier based test structure for transistor threshold voltage variation," *Microelectronic Test Structures, 2008. ICMTS 2008. IEEE International Conference on*, pp. 3–7, March 2008.

[9] H. Notani, M. Fujii, H. Suzuki, H. Makino, and H. Shinohara, "On-chip digital idn and idp measurement by 65 nm cmos speed monitor circuit," in *Solid-State Circuits Conference, 2008. A-SSCC '08. IEEE Asian*, pp. 405–408, Nov. 2008.