School of Electrical, Electronic & Computer Engineering

Variability: Analysis and Impact on Circuit Response

By

Hiran Ramakrishnan

Technical Report Series

NCL-EECE-MSD-TR-2009 -144

March 2009

Contact: H.Ramakrishnan@ncl.ac.uk

NCL-EECE-MSD-TR-2009 -144

School of Electrical, Electronic & Computer Engineering,
Merz Court,
Newcastle University,
Newcastle upon Tyne, NE1 7RU, UK.

http://async.org.uk/

1. Introduction

Most modern digital VLSI designs operate with GHz frequencies. Hence the precision in the design and manufacturing process is very critical. During chip manufacturing, random process variation can affect the size of the transistor, which becomes a greater percentage of the overall transistor size as the dimension shrink. With MOSFETS becoming smaller, the number of atoms in the semiconductor material that produce many of the transistor properties is becoming fewer thereby amplifying the dependence of the transistor characteristics with process parameters. The transistor characteristics become less deterministic, but more statistical. This report includes various methods to analyze and model variability. In addition, several circuits are used to study the impact of process parameter fluctuations on its output.

1.1What is Variability?

Variability can generally be referred as the spread of a distribution is.

1.2 How Variability is Measured?

Variability is seen everywhere in day to day life and the most frequently used measures of variability are:-

1. *Range*

   The range is the simplest measure of variability to calculate; it is simply the highest score minus the lowest score.

2. ***Interquartile Range***

   The interquartile range (IQR) is the range of the middle 50% of the scores in a distribution.

3. ***Variance***

   Variability can also be defined in terms of how close the scores in the distribution are to the middle of the distribution.

4. Standard Deviation

The standard deviation is simply the square root of the variance. The standard deviation is an especially useful measure of variability when the distribution is normal or approximately normal (see Chapter 6) because the proportion of the distribution within a given number of standard deviations from the mean can be calculated. For example, 68% of the distribution is within one standard deviation of the mean and approximately 95% of the distribution is within two standard deviations of the mean.

1.3 Importance of Variability in IC manufacturing and its impact on circuit applications

Variations in general can be classified as controllable (systematic) variations and uncontrollable (random) variations. These two types of variations can be seen in most of the man-made systems. Controllable or systematic variations can be tuned for the desired optimum performance of the system, whereas random variations are uncontrollable and can be spontaneous which changes the function of the system; the changes may even lead to the failure of the system. Variability holds primary importance when artifacts are created from an assembly of millions of components comprising minute quantities of materials as in the manufacture of integrated circuit.

In the design of circuits with deep sub micron technologies the analysis of variability has become a very important tool, , to predict the response variation very early in the design cycle due to process parameter spreads. Furthermore, for several decades the output from semiconductor manufacturers has been high volume products with process optimisation being continued throughout the lifetime of the product to ensure a satisfactory yield. However, product lifetimes are continually shrinking to keep pace with market demands. In addition, there is an increase in foundry business where product volumes are low; subsequently it is no longer feasible to optimise the process during the product lifetime resulting in an increase in parametric yield loss. Consequently, there is an increasing need for semiconductor process versatility which can be adopted by a diverse customer base. Central to ensuring customer satisfaction of the manufactured product is an understanding of the effects of process variation on customer designs.

Integrated circuit manufacturing is one of the processes which involve many high precision steps. Most modern digital VLSI designs which operate with very high frequencies consist of transistors with dimensions in the range of few nano-meters. Hence precision in the design and manufacturing process is very critical. During chip manufacturing, random process variation can affect the size of the transistor, which becomes a greater percentage of the overall transistor size as the dimensions shrink. With MOSFETS becoming smaller, the number of atoms in the semiconductor material that produce many of the transistor properties is becoming fewer thereby amplifying the dependence of the transistor characteristics with process parameters. The transistor characteristics become less deterministic. This statistical variation increases design difficulty. Design corners based on circuit applications can be identified once the distribution of the variation of parameters is known. The variation of the threshold voltage for technology generations ranging from 250nm down to 45nm are shown in Table 1.

L (nm)	Nominal Vth (mV)	% Change from the nominal
250	450	4.7
180	400	5.8
130	330	8.2
60	300	9.3
65	280	10.7
45	200	16

Table. 1 Higher relative variability of threshold voltage for newer technologies [1].

The unnecessary parameter fluctuations can be differentiated based on the variability sources which can be physical or environmental. The physical source can be the manufacturing process, manufacturing equipment, electron migration, changes in the characteristics of devices and/or interconnect lines due to ageing. The environmental sources can be due to the change in supply voltage, temperature (operating conditions), local coupling, peculiar design implementation etc. The

process variation which is a physical variation can occur from wafer to wafer or lot to lot (both are inter process variations). It can also be between die to die (intra process variations) [1]. The process parameter fluctuations in general can be organized as global variations and local variations.

Wafer to wafer or lot to lot variations are global variations which can be due to materials and gas flow (linear variation) or thermal or the wafer spin process (radial variation). This type of variation mainly affects the dynamic performance of the digital systems for example switching speed, dynamic power and gain in analog systems. There can be reticle variation (during lithography) due to the optical processes. Other sources of variations include position and proximity variations arising from the layout. These variations can be extracted and modeled and since they are systematic they are called predictable variations. Variations arising from layout can be optimized by several methods, for example mask compensation.

- **Global Variations are [2] :**
  - Material properties of wafer, resists, etc.
  - Lens aberration, flow turbulence, oven temperature, etc.
  - Implant dose, diffusion time, focus, exposure energy, etc.
  - Lg, W, oxide thickness, layer thickness, doping, etc.

Local variations between identically laid-out devices arise from random microscopic processes variations. These variations are randomly distributed and are unpredictable. For 130nm technology, local variations constitute around 30% of the overall variations [24]. The random distributions arise from the variation of process parameters for example impurity concentration densities, oxide thickness and diffusion depths which result from varying operating or environmental conditions during the deposition or diffusion of the impurity dopants. The fluctuations in the process parameters may result in the variation of sheet resistance and threshold voltage. However, the variation of threshold voltage can also be due to the variation of surface charge or change in the oxide thickness. The varying resolution of the photolithographic process may lead to the variation of geometric dimensions which may result in the variation of transistor output characteristics for example, drain current. These variations are not correlated because the width of the transistor (W) is

determined in the field oxide step while the effective length of the channel (L_{eff}) is defined in the poly and source-drain diffusion steps [24].

- **Local Variations are [2] :**
  - Line Edge Roughness (LER).
  - Discrete oxide thickness.
  - Random dopant fluctuations.
    - Statistical fluctuations in number and position of dopant atoms in the channel – spatially uncorrelated.

The variations will impact the performance of circuit, which may exhibit wider variability leading to the degradation of yield in modern technologies and applications. The circuit parametric variations for a 130nm technology process are shown in the Table 2.

Process	Change	Interconnect	% change
L_{eff}	16.70%	ε	3
V_{th} (V)	30%	ρ	30
t_{ox} (^{o} A)	10%	*w*	20
R_{ds} (Ω)	10%	s (nm)	20
Run-time		t (nm)	10
V_{dd} (V)	10%	*h*	10
T (^{o}C)	25 -100	R_{via} (Ω)	20

Table. 2 Sources of major variation at 130nm technology node [1].

For a given operating temperature L_{eff}, V_{th} and V_{dd} are the most dominant variation sources of a logic gate. V_{dd} and temperature are run-time variation sources. Figure 1 shows the impact of process variations on the delay of a 4 bit adder fabricated using 130nm technology. Using the variation values shown in Table 2 it can be seen from Figure 1 that the performance variability $3\sigma/\mu$ (where σ is the standard deviation and μ is the mean) is as large as 15% at a supply voltage of 1.2V. However, when the supply voltage is scaled to 0.5V in an aim to reduce the power consumption, the performance variability increases to 45% which shows that circuit yield degrades with voltage scaling.

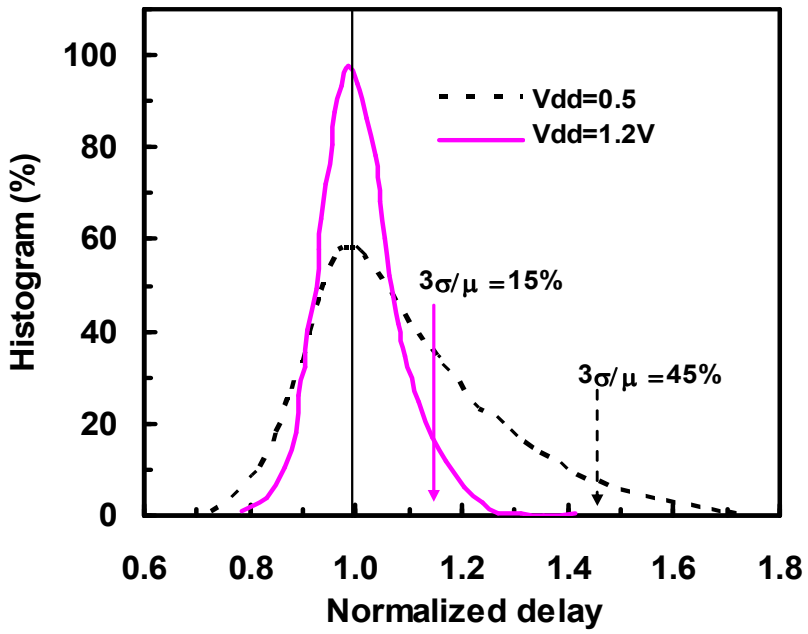


Fig. 1 Monte Carlo simulations of a 4-bit adder – Impact of variations on delay [1].

The impact of variations on leakage current for a 4KB SDRAM realized with 130nm technology generation is shown in Figure 2. It can be observed from the figure that in comparison to the leakage current values simulated without considering channel length and threshold voltage variations, the leakage current obtained by measurement, rises exponentially at large values of supply voltages which may lead to the malfunctioning of the SRAM. This dramatic increase of leakage current in the

measurement data shown in Figure 2 is caused by the transistors with shorter channel length and lower threshold voltage.

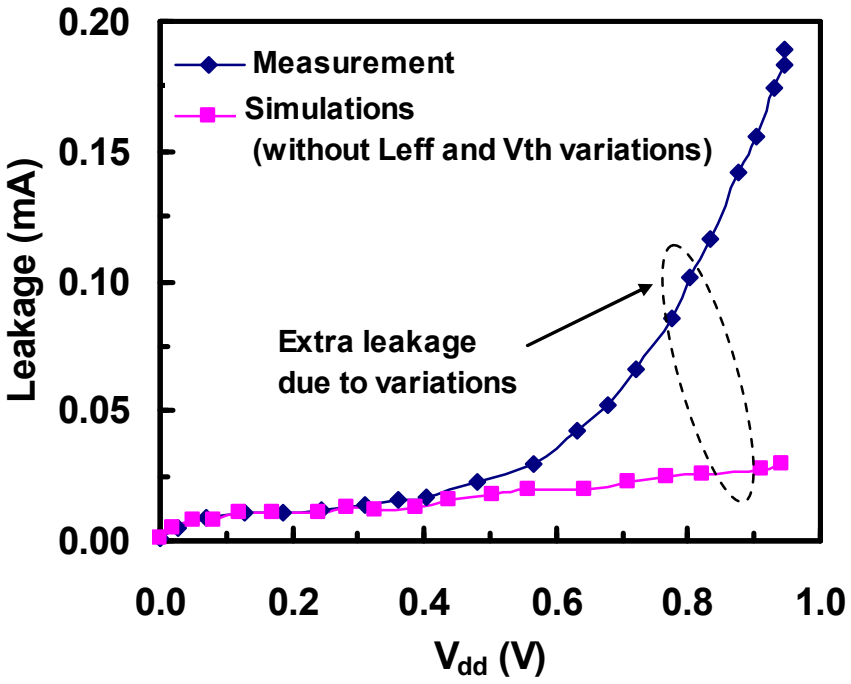


Fig. 2 Leakage measured from a 4KB SDRAM- Impact of variations on leakage power consumption [1].

Given that compensating for random variations is difficult, containing the effects of the random parameter variations can be considered as the most significant challenge the microelectronics industry will face in the coming technology generations. Hence it is important to have an organized and collective effort of technology and design engineers to mitigate the effect of variations on modern day circuits. Designers need EDA tools to accurately model different sources of variations to guarantee improved yield which could be called as 'Process Aware Design' (PAD). PAD could exploit the issues of variability to achieve trade-offs between power and circuit throughput thereby saving the cost from using the worst-case design approach. Hence the first step is to study a given variation and its impact on circuit performance. However the amount of variation expected from various processing technologies for example strained-Si, SOI, SGOI etc are different [3, 4, 5].

The answers for the following questions could lead us to variability tolerant or variability aware designs, keeping up the pace of microelectronics industry to follow Moore's Law.

- What is Variability?
- What are the sources of Variability?
- How serious is the impact of Variability on the response?
- How is the Variability measured?
- What are the parameters involved in the variation?
- Are there any other global factors which are significant for the variation?
- What's the best statistical technique available for modelling variation?
- Is the statistical technique used for analysis is efficient?
- What's the accuracy of the statistical technique?
- Will it lead to overdesign just like deterministic techniques?
- Is the statistical technique computationally efficient?
- Can it be applied in different platforms?
- Can it be used for different technologies?
- Is it user friendly?
- Are there tools available?
- Are the existing TCAD/EDA tools restricting research?
- Can we adapt the existing tools to do the research?
- Is it better to keep the older technology to save costs?
- Are there any technologies (process) which is variability tolerant?
- Can we exploit variability (if it is unavoidable) for certain applications?

This report presents some of the answers to the concerns listed above. Some of the statistical techniques used for the analyses of variability are explained in detail. The methodology used for the variability analysis and optimisation in IC manufacturing is presented. A statistical method called Design of Experiments (DOE) is used to analyze the process variability. Statistical screening method is performed to determine the most significant parameters. The disadvantages associated with the basic DOE are explained and another technique called the multi-partitioned DOE is presented. Comparison of these statistical techniques with the conventional Monte Carlo methods is discussed in detail. Basic asynchronous blocks for e.g., Jamb Latch, MUTEX, realized with strained-Si technology under different operating conditions are discussed. The impact of process and operational variability on the characteristics of the response of these circuits are presented.

1.4 Corner Case Design Vs Statistical Design

However, to study the impact of variability in the manufacture of transistors statistical methods seem to be an obvious choice as corner case methods, previously used, do not seem to work for deep sub-micron manufacturing technologies. Corner cases would prove to be an overdesign for new technologies which will incur severe cost and poor yield for the manufactured product. Table 3 summarizes the advantages of statistical design techniques over the conventional worst case design (corner case) approach. Previously the effects of process variations were accounted for by incorporating artificial distributions in SPICE models.

Monte Carlo Analysis is a widely used technique to analyse parameter variability to generate response distributions. The simulation flow for the technique is shown in Figure 3. Although the technique is very accurate it is computationally inefficient particularly when a large number of variables is involved, for example, in analysing the effect of parameter variation in semiconductor processes; however, if appropriate, accuracy could be traded-off against computation times. A further disadvantage of this method is that every time the manufacturing process changes the simulation is to be performed again to achieve a proper understanding of the statistical variations.

Corner case design	Statistical design
Logic gates in a circuit have the same process corners	Different logic gates in the same circuit have different process corners
Suffers from overdesign	Realistic worst case extremes
Significant performance loss	Well designed circuits with no performance loss
Computationally efficient	Computationally inefficient

Table. 3 Corner case design Vs Statistical design techniques [6].

Another approach to analysing the effect of process variation is to use Design of Experiment (DOE) and Response Surface Modelling (RSM) techniques to determine process combinations that should be simulated to achieve the corresponding response. The resulting response surfaces are then used to statistically produce process distributions that result from control factor variations. A comparison between the efficiency of Monte Carlo Analysis and the DOE technique in analysing the effect of parameter variations on the threshold voltage of n- and p-MOS has been discussed in the literature. It has been observed from literatures that the threshold voltage distributions obtained using Monte Carlo technique and DOE/ RSM approach analyses fits well with simulations lasting for a few seconds in the case of DOE/ RSM approach compared to 10 minutes for Monte Carlo technique. The analyses is limited to DOE/RSM approach due to its computational efficiency compared to Monte Carlo technique and the large number of process parameters in the technology to be analysed.

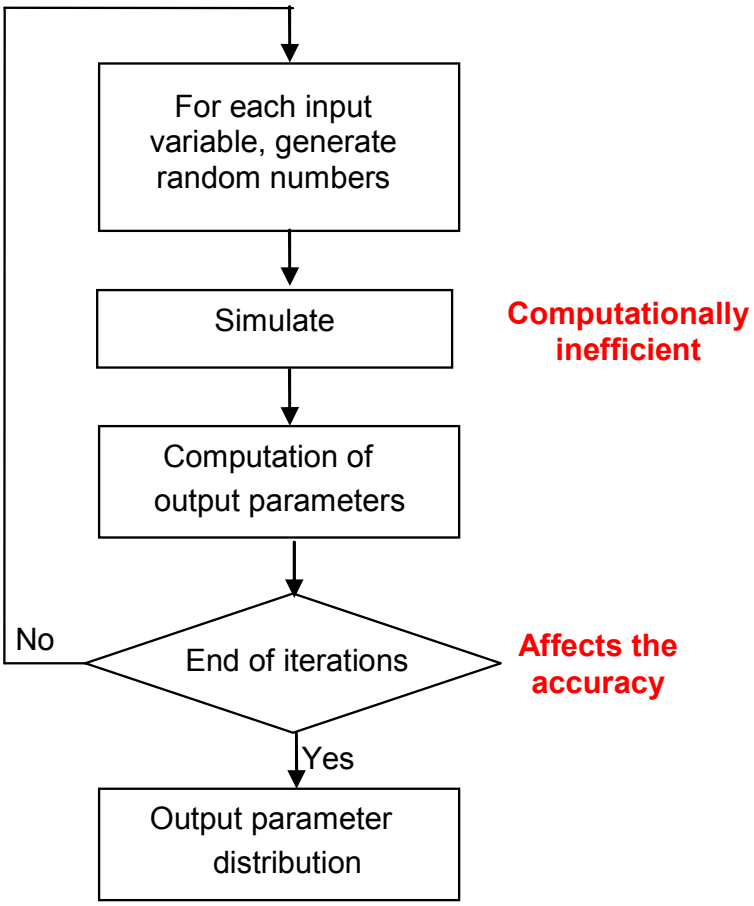


Fig. 3 Monte Carlo Analysis simulation flow [6].

2. Importance of TCAD in the analysis: An Introduction

CAD tools are essential to perform optimisation, calibration and the creation of the models to accurately predict the effects of process variability, thereby increasing the overall productivity of the designs. The entire work related to the analysis of variability is performed by developing a standard Si and strained-Si technology library by simulations using TSUPREM4 (process simulator) [7], MEDICI (device simulator) [8], AURORA (parameter extraction program) [9], MINITAB (Design of experiments) [10] and PSPICE (circuit simulator) [11] subjected to different operating conditions and varying amounts of strain in the

channel. High performance (HP) strained-Si and standard Si devices on a 65nm technology node are considered for the analyses. ITRS/MASTAR [12, 13] device parameters have been used for the device and circuit dimensions and operating conditions for simulation.

By interfacing TCAD simulations with a SPICE model extraction strategy [9] and applying variations to parameters which are specific to the selected manufacturing process, designers are able to predict the effect of variation of process parameters thereby improving the yield or matching the manufactured devices, according to the characteristics to particular applications.

The ensuing discussion will outline the methodology used to identify the process parameters of a typical 65nm technology node and the methodology to study the impact of these process parameters on circuit output response. The demonstration circuits used for the analyses discussed in section 7 comprise two basic asynchronous logic building blocks, namely, the Mutual Exclusion Element (MUTEX) and Jamb Latch synchronizer. In the analysis, the significant parameters which impact on the circuit outputs are identified and the response surface modelled. Finally, a comparison is made between strained and conventional Si implementations of the demonstration circuits with respect to an important design parameter, namely, "metastability resolution time".

2.1 Introduction to TCAD

For a large number of transistors per chip and owing to the development of new structures and materials for devices, verification of circuits is becoming more complex and the use of analytical simulators for studying and designing the transistors are therefore necessary. The commercially available TCAD tools for semiconductor device and process simulations are MEDICI and TSUPREM4 from Synopsys Corporation, ATHENA and ATLAS from SILVACO, TCAD-studio etc. For circuit model extractions AURORA from Synopsys and UTMOST from SILVACO are the most commonly used tools.

MEDICI is a device simulator, which allows user to create a two dimensional structure of a semiconductor device, including the definition of oxide and silicon

regions, doping profiles to simulate the current – voltage characteristics of the device. TSUPREM4 can be used to replicate the IC manufacturing process. These simulators can be used to model any semiconductor device. AURORA is general purpose optimisation tool for fitting analytical models to data and extracting parameters for circuit simulation. The mathematical models used in the suite of TCAD tools comprise a set of fundamental equations, for example, Poisson, the continuity and transport equations. Poisson's Equation relates variations in electrostatic potential to local charge densities. The continuity and the transport equations describe the way that the electron and hole densities evolve as a result of transport, generation, and recombination processes. A number of physical models are incorporated into the simulator for accurate simulations, including the models for carrier recombination, photo generation, impact ionization, band-gap narrowing, band to band tunnelling, mobility and carrier life time. The TCAD tool suite has the capability to analyse current density, electric field temperature etc. MEDICI features include transient and AC small-signal analysis, impact ionisation, gate current and ionisation integrals. There are different transport models available in device simulators for example, drift-diffusion and hydrodynamic models. In this thesis for the device simulations using TCAD, hydrodynamic simulations were preferred over conventional drift-diffusion simulation due to the degradation of accuracy when smaller feature sizes for the devices are considered [7, 8]. However, the accuracy also depends on the mesh density used to define the device. A very dense mesh was defined in places of interest in the device for example, the channel of the transistor. There is a variety of mobility models provided by the simulator which need to be selected and this selection was done during the process calibration. The calibration process is discussed in the sub-sections.

The interaction of TCAD tools of interest is shown in the Figure 3.1.

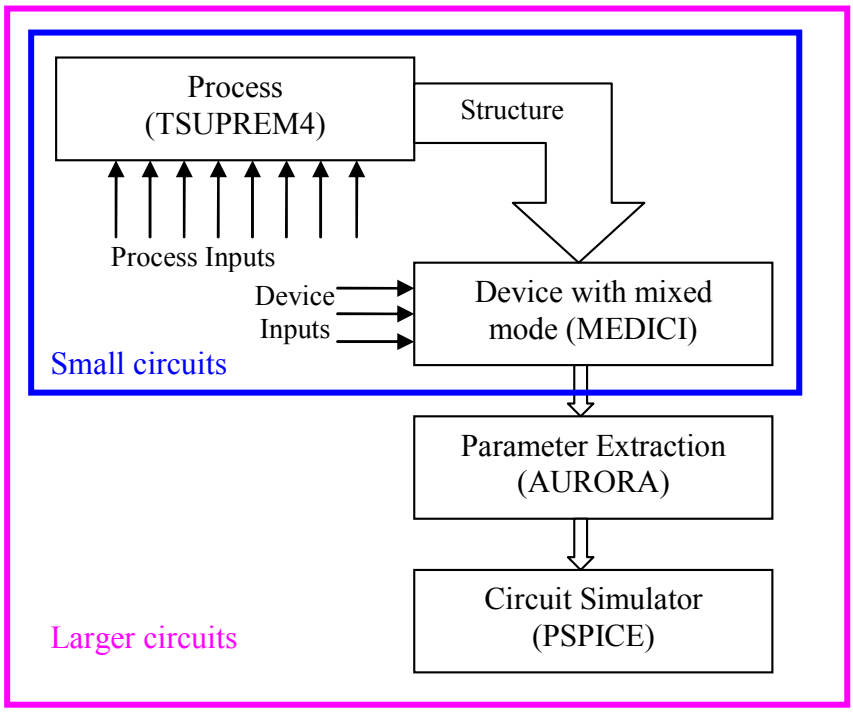


Fig. 4 TCAD tools.

As shown in Figure 4, if a larger circuit has to be analysed (For example, Jamb latch and MUTEX) then the mixed mode simulation could not be used due to the limitation on the number of nodes used to define the device. For larger circuits, the device characteristics are passed to the AURORA program which would generate the appropriate device models for a circuit level simulator such as PSPICE.

2.2 Method of simulator calibration

For accurate device characterisation, the process and device simulators in the TCAD suite of tools must be calibrated. The calibration procedure for the process simulator comprises creating a virtual two dimensional device structure by selecting the appropriate process coefficients and process models, for example diffusion and

ion implantation in the simulator. Thereafter the very tedious task is undertaken of 'matching' various profiles in the virtual device with those from the actual device; for example, doping profile extracted from a physical device using Secondary Mass Ion Spectroscopy (SIMS). The matching procedure, which effectively calibrates the simulator for a given process, is a lengthy task due to the large number of undefined coefficients which can be tuned in the empirical based models in the process simulator and the length of the subsequent process simulation runs; the procedure is also iterative. As an exact match will be rarely obtained there will also be slight discrepancies between the structure generated by the process simulator and the approximately extracted SIMS and Transmission Electron Microscope (TEM) data. A two dimensional mesh structure of a conventional Si MOSFET constructed using MEDICI is shown in Figure 5.

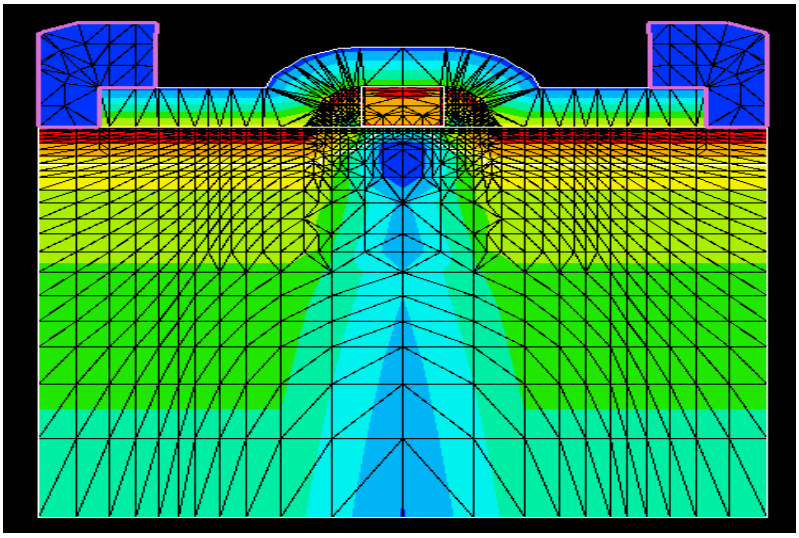


Fig. 5 Structure of a conventional MOSFET constructed using MEDICI.

The method of calibration adopted in this report is shown in Figure 6. The experimental data obtained using device electrical characterisation is compared with an example TCAD deck. The data included current voltage characteristics and dimension related parameters. The process of matching experimental data with example TCAD deck was continued until a near accurate match is obtained. In the matching process, different mobility models along with different doping profiles and

dimensions were analysed. The mobility models available in TCAD are classified into two categories namely, Low field and Inversion layer mobility models.

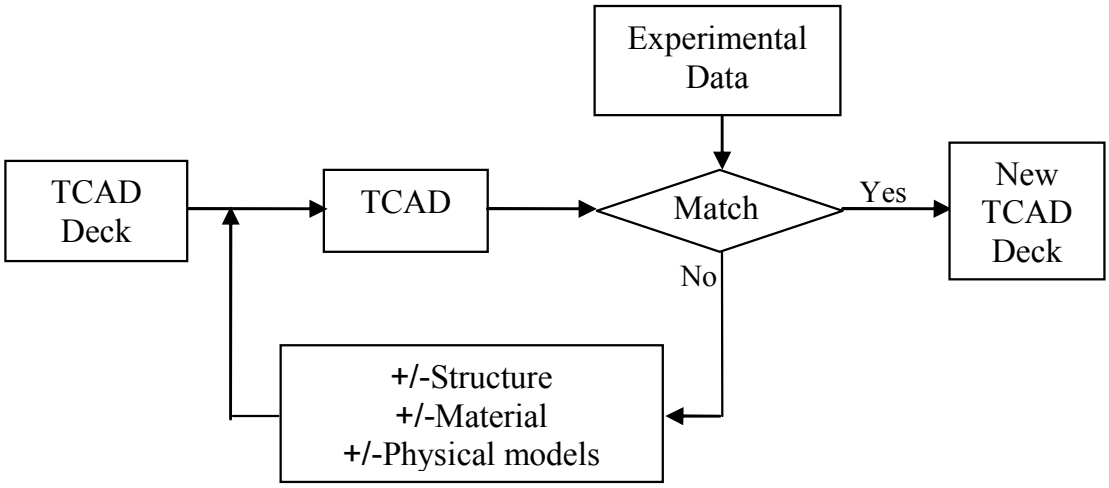


Fig. 6 Method of Calibration

Initially, mobility models namely, CONMOB, PRPMOB and FLDMOB which represents bulk, parallel and perpendicular fields respectively [9] were chosen due to the fact that operation of MOSFET's range from zero electric field to high electric field. To obtain a reasonable match of current voltage characteristics with that of the experimental data, doping profile, width of source-drain junctions etc were tweaked. For deep sub-micron devices, an enhanced mobility model SRFMOB2 which takes into account phonon scattering, surface roughness scattering and charged impurity scattering was selected. In the case of self-heating analysis, a temperature dependent mobility model TMPMOB was included in the TCAD deck along with hydrodynamic models [7, 8].

3. Identification of Process Parameters

The Monte Carlo analysis was the generally adopted approach used to study the effects of parameter variability in a wide range of applications for example, chemical industries, flight simulations etc; where the number of parameters is limited. However, when the number of parameters considered increases, this approach, although accurate, is computationally inefficient. The technique described in this report to identify semiconductor process parameters whose variability would impact

most on the circuit characteristics, considered critical by the designer, is realised through a two phase process using DOE and RSM statistical tools. DOE is a strategy used to create a set of experiments in which the range of variables can be altered systematically to enable any correlation between variables to be ascertained to determine the main contributing factors to the variation in, for example, design parameters of interest. In this particular application the 'experiments' in DOE are simulations. Within the DOE procedure, the Plackett-Burman screening technique is incorporated to screen the most significant parameters which will subsequently be used in the second phase to model the response surface [14].

Experimental Design formally represents a sequence of experiments to be performed, expressed in terms of factors or design variables set at specified levels that is, predefined values. Having identified the most influential parameters, it is essential to reflect the effects of the variations onto the output response of the system under investigation. In this instance the 'Variations' are the semiconductor process parameters and its effect on some aspect of the output response of a circuit which is of interest.

Variability analysis is performed by developing a standard Si and strained-Si technology library. In creating a technology library a number of CAD tools are used in sequence as shown in Figure 7; namely TSUPREM4, MEDICI, AURORA and PSPICE, DOE analysis is performed using the statistical package MINITAB.

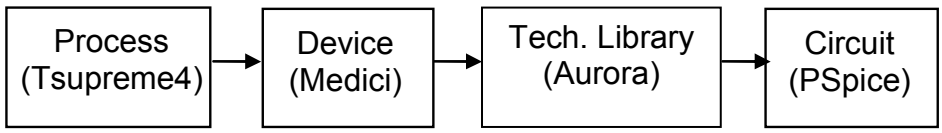


Fig. 7 CAD simulator flow.

4. Methodology to Study the Impact of Process Variations on Circuit Parameters

As a simple demonstration of the above procedures, a Mutual Exclusion Element (MUTEX) and a Jamb Latch synchronizer were selected to study the impact of process variations on particular circuit parameters. The MUTEX and Jamb Latch circuits are important building blocks which are used extensively in asynchronous

circuits. In the design of a MUTEX two critical timing parameters, t_{off} and t_m [15-17], which are discussed in detail later are important for the correct operation of the circuit, they are also susceptible to the effects of process variations; whereas in the case of Jamb Latch, t_m is the output response on which the process variations are affected.

The implementation technology chosen for the demonstration circuits is strained-Si. The amount of strain in the strained-Si device can be considered as a process parameter and hence gives an easy way of comparing the performance of the circuits implemented in the standard silicon process technology (0% strain) and one implemented in the strained-Si technology.

The electrical characteristics of strained-Si and conventional Si n- and p-MOSFETs are applied to AURORA parameter extraction simulator for optimization and extraction of PSPICE models. The strained-Si and Si PSPICE transistor library based on different strain conditions and technology nodes were developed for the analyses. Subsequently these transistors were used in the design of the demonstration circuits in PSPICE and simulations were performed to evaluate the variations of t_m and t_{off} of strained-Si and Si based MUTEX circuits under different strain and process conditions.

The procedure followed to study the impact of process variations on the chosen circuits is shown in Figure 8 in which there are two paths. The result from the first path is the benchmark for the circuit, and the second is the stochastic output based on the variability of the process and device parameters using DOE. The best process parameters can then be selected, depending upon the application of the circuit, to ensure a good yield.

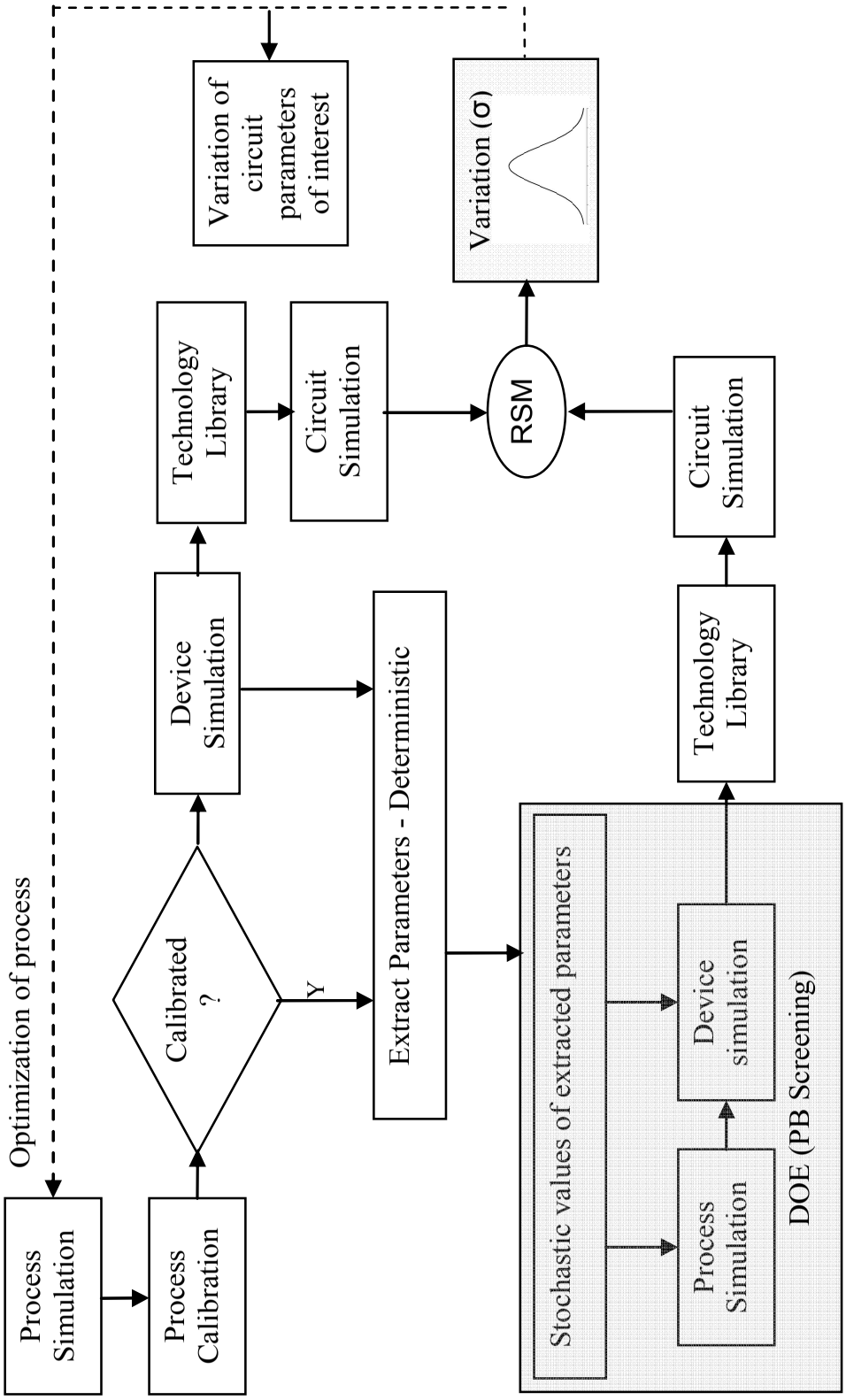


Fig. 8 Methodology to study the effect of process variations.

To study the impact of the variability of process parameters on circuit characteristics the main process parameters were selected from a chosen 65nm process (with ITRS specifications). State of the art 65nm technology was selected as the technology node of choice as its variability is seen to be larger in comparison to 300nm technology node. Monte Carlo Analysis would have been the obvious choice to study the effects of variability as it is the conventional statistical approach when considering large amounts of data. However, to study the effects of the variation of large number of processing parameters on a circuit a huge number of process, device and circuit simulations would need to be performed which is highly complicated, computationally inefficient and time consuming but accurate using the Monte Carlo approach. This is due to the fact that the accuracy of the results from Monte Carlo Analysis depends on the number of random samples, a large number of finite samples are necessary to achieve reasonable accuracy for a complex system which was shown in Figure 3. Consequently, a different statistical method called DOE was used which when compared to Monte Carlo approach is very efficient computationally. Figure 9 shows the steps undertaken in the analysis of the effects of process variation using DOE/RSM.

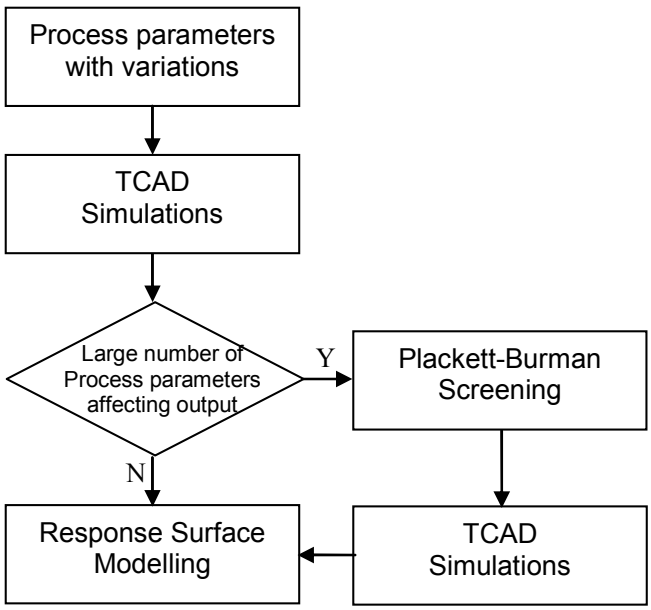


Fig. 9 Flow chart of the variability analysis.

Figure 10 shows various techniques and methods available at each stage of operation of DOE/RSM

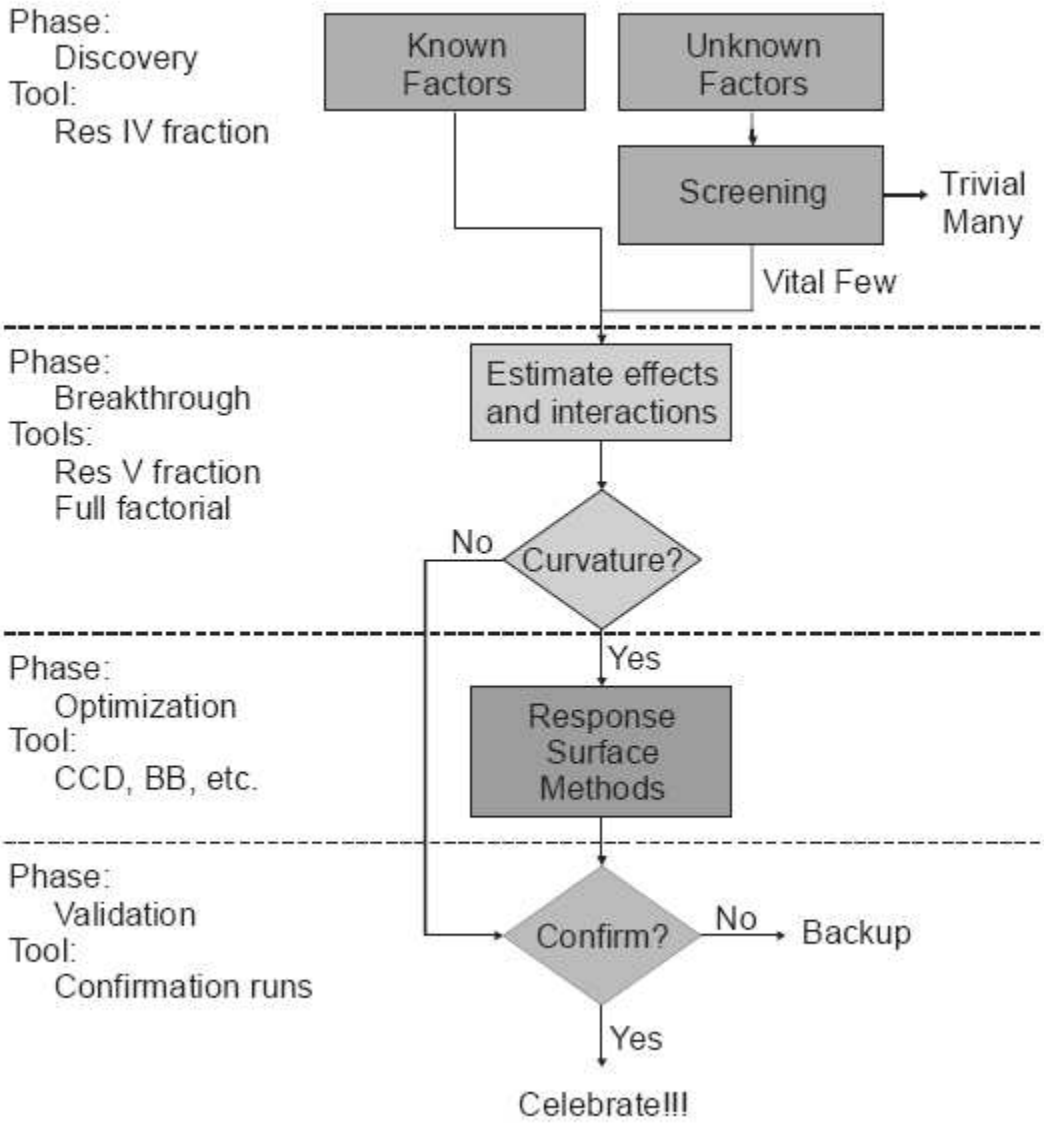


Fig. 10 Techniques available to perform DOE/RSM [13].

The main disadvantage of DOE is that the whole system becomes complex when more parameters impact significantly on the response of the system. To reduce the complexity of DOE, multi-partitioned DOE is used for the analyses.

5. Example to show how multipartitioning is performed.

Figure 11 shows an acyclic graph of an arbitrary process technology assuming that A, B, C, D, E, F, G and H are the input variables which impact on the process output parameter X. There can be individual effects (output response is affected by the individual input parameters) or interaction effects (inputs which combine to give an impact on the output response) between the input process variables which have a greater or lesser effect on the output parameter X. As explained in Figure 12, to find the significant input parameters from the list (A…H) which are seen to be significant by monitoring the response X, we performed screening and folding procedure (to identify the individual input parameters which affects the response). Correlation analysis is performed and the parameters are listed according to the ranking.

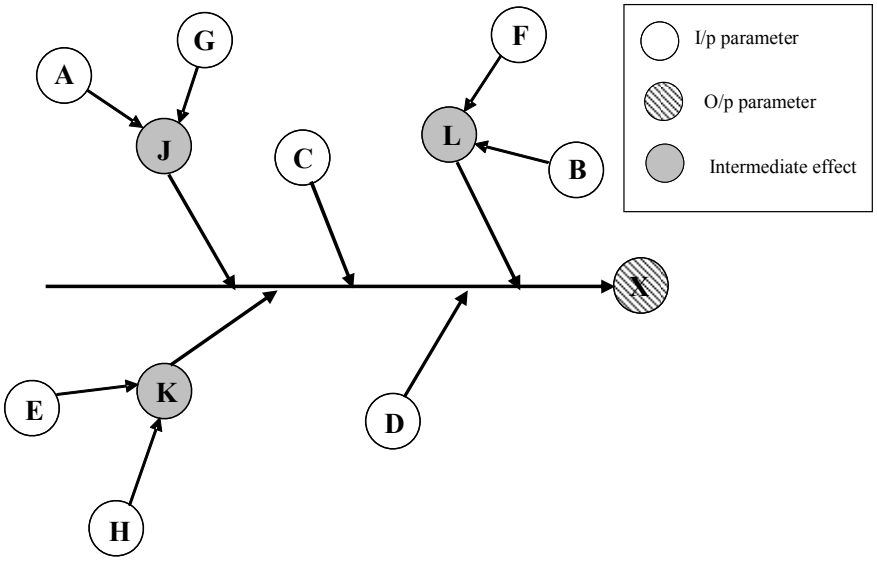


Fig. 11 Acyclic graph showing an arbitrary process technology

Assuming the parameters is sorted in the descending order according to their statistical significance as shown in the Table1.

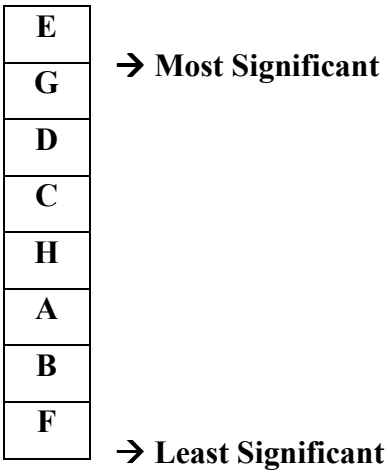


Figure 12

If the number of significant parameters is large, we apply a method called multi partitioning where the parameters which are ranked are put in different groups to reduce the complexity in doing Response Surface Modelling (RSM) to model the first order, second order or square effects.

Let the groups be partitioned as shown in the Figure 13 (a) and Figure 13(b).

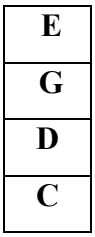


Fig. 13 (a)

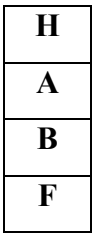


Fig. 13(b)

RSM can be done for both the groups to get the interaction effects on the responses by each group. However, this ranking and grouping is based on statistical advantage, it can be seen from the process acyclic graph that G and A always interact to give an intermediate effect called 'J' which in turn impacts on the response X. Also E and H interacts together to give an effect called 'K' which then determines X. So from the process knowledge the previous partitions can be regrouped as shown in the Figure 14 (a) and 14 (b).

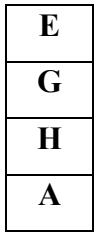


Fig. 14 (a)

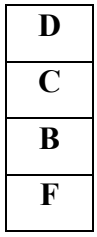


Fig. 14 (b)

It is also possible to have different groups by compromising the computational efficiency as shown in Figure 15 (a) and Fig 15 (b).

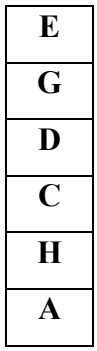


Fig. 15(a)

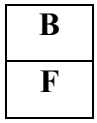


Fig. 15(b)

Having ranked and partitioned the parameters RSM can be performed on each group to create a complete process variability models. Computational efficiency of an RSM can be quantified as 2^n+2n+1, where 'n' is the no of parameters used for modelling [14].

As an example the influence of input process parameters on the threshold voltage of a 65nm generation device is shown in Figure 16.

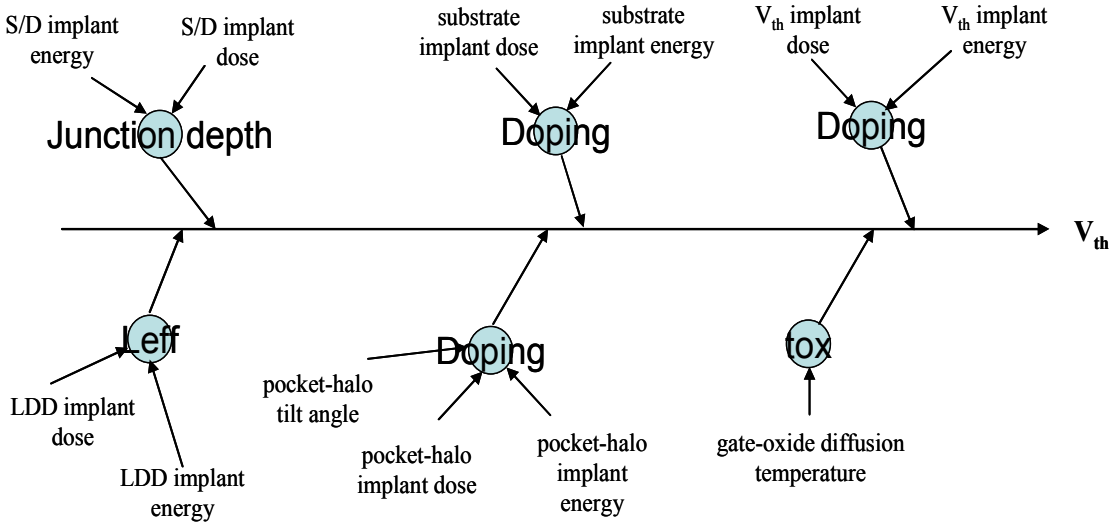


Fig. 16 Acyclic graph of a 65nm technology device which shows the influence of input process parameters on the threshold voltage.

5. Multi-Level Partitioned Response Surface Model

[Taken from relevant publications - 2]

An approach for building multi-level partitioned response surface models is presented in this section. Response surface models are constructed to replace the computationally expensive simulation analysis and facilitate fast analysis and exploration of the design space. The most widely used response surface approximating functions are low-order polynomials relating a predicted response, $\hat{y}$ to a set of design variables x. If little curvature appears to exist, a two level fractional factorial experiment is designed, and the first-order polynomial to approximate the response is given as

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i \qquad (1)$$

If significant curvature exists, the second-order polynomial including all two-factor interactions, is commonly used, is given as

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \sum_{j=1}^{k} \beta_{ij} x_i x_j \quad (2)$$

In (1) and (2), *k* is the number of input variables, x_i is the *i*th input variable and β is the RSM coefficient calculated using least squares regression analysis to fit the response surface approximation $\hat{y}$ [14].

The multi-level partitioned response surface approach helps to overcome the problem of modelling large number of variables. The variables or factors and the responses are grouped and the partitioned response surface models are developed which takes into consideration, the effect of all factors. This technique was first developed by Koch [14] and he demonstrated it on commercial turbofan engine application.

Consider the number of factors, n, and the number of responses, y to be modelled for a particular complex simulation code. The factors are grouped or partitioned into say two sets with *k* number of factors in set 1 and *(n-k)* number of factors in set 2. The responses are also grouped into two sets with the first set of *s* responses $\hat{y}_1$ and the second set of *(r-s)* responses $\hat{y}_2$. The partitioning of factors and responses is ideally based on the process knowledge about the problem under consideration. Since, each response will be a function of all the factors; the accurate partitioning is not required. The first set of experiments is designed and run and the first set of responses $\hat{y}_1$ are fitted as the function of the *k* factors of set 1, as in (3). The second set of experiment is designed and run and the second set of responses $\hat{y}_2$ are fitted as the function of *(n-k)* factors of set 2, as in (5). In both these experiments, $\hat{y}_1$ and $\hat{y}_2$ are measured. The effect of second set of factors *(n-k)* on the response $\hat{y}_1$ is modelled by fitting the mean term a_0 of (4) as the function of the second set of factors. The same procedure is repeated for the second set of responses $\hat{y}_2$ using the first set of factors as in (6). Hence the two-level response surface models are created without any additional experiments performed. When the experiments are performed for first set, the factors of second set are kept at their nominal values and vice versa and for both these experiments y_1 and y_2 are measured. The data used to fit $\hat{y}_1$ in (3) is used to fit c_0 in (6) and the data used to fit $\hat{y}_2$ in (5) is used to fit a_0 in (4).

$$\hat{y}_1 = a_0 + \sum_{i=1}^{k} a_i x_i + \sum_{i=1}^{k} a_{ii} x_i^2 + \sum_{i<j} \sum a_{ij} x_i x_j \qquad (3)$$

where,

$$a_0 = b_0 + \sum_{s=k+1}^{n} a_s x_s + \sum_{s=k+1}^{n} a_{ss} x_s^2 + \sum_{s<t} \sum a_{st} x_s x_t \qquad (4)$$

$$\hat{y}_2 = c_0 + \sum_{s=k+1}^{n} c_s x_s + \sum_{s=k+1}^{n} c_{ss} x_s^2 + \sum_{s<t} \sum c_{st} x_s x_t \qquad (5)$$

where,

$$c_0 = d_0 + \sum_{i=1}^{k} d_i x_i + \sum_{i=1}^{k} d_{ii} x_i^2 + \sum_{i<j} \sum d_{ij} x_i x_j \qquad (6)$$

The advantage of using this method is that the computational or the experimentation cost for modelling the responses is reduced immensely. For example, if there are 16 significant factors to be modelled using a standard central composite design (CCD) in RSM, it would take 65,569 simulation runs [14]. If the same 16 factors are partitioned into two sets of eight factors each, two experiments of 273 runs each are required to fit the two part response surface models using the above mentioned approach. Also, the approach is not restricted to two groups of factors, but the factors and responses can be partitioned for more than two groups.

The major disadvantage in this approach is that the interaction effects between the set of factors partitioned cannot be modelled. From the partitioned response surface modelled in (3)-(6), the $x_t x_s$ interaction terms are the missing terms. Hence, the assumption to be made while using this approach is that the interactions terms are from the partitioned groups are negligible. But this can be overcome by partitioning the factors and responses such that the interactions which are significant are included in one group.

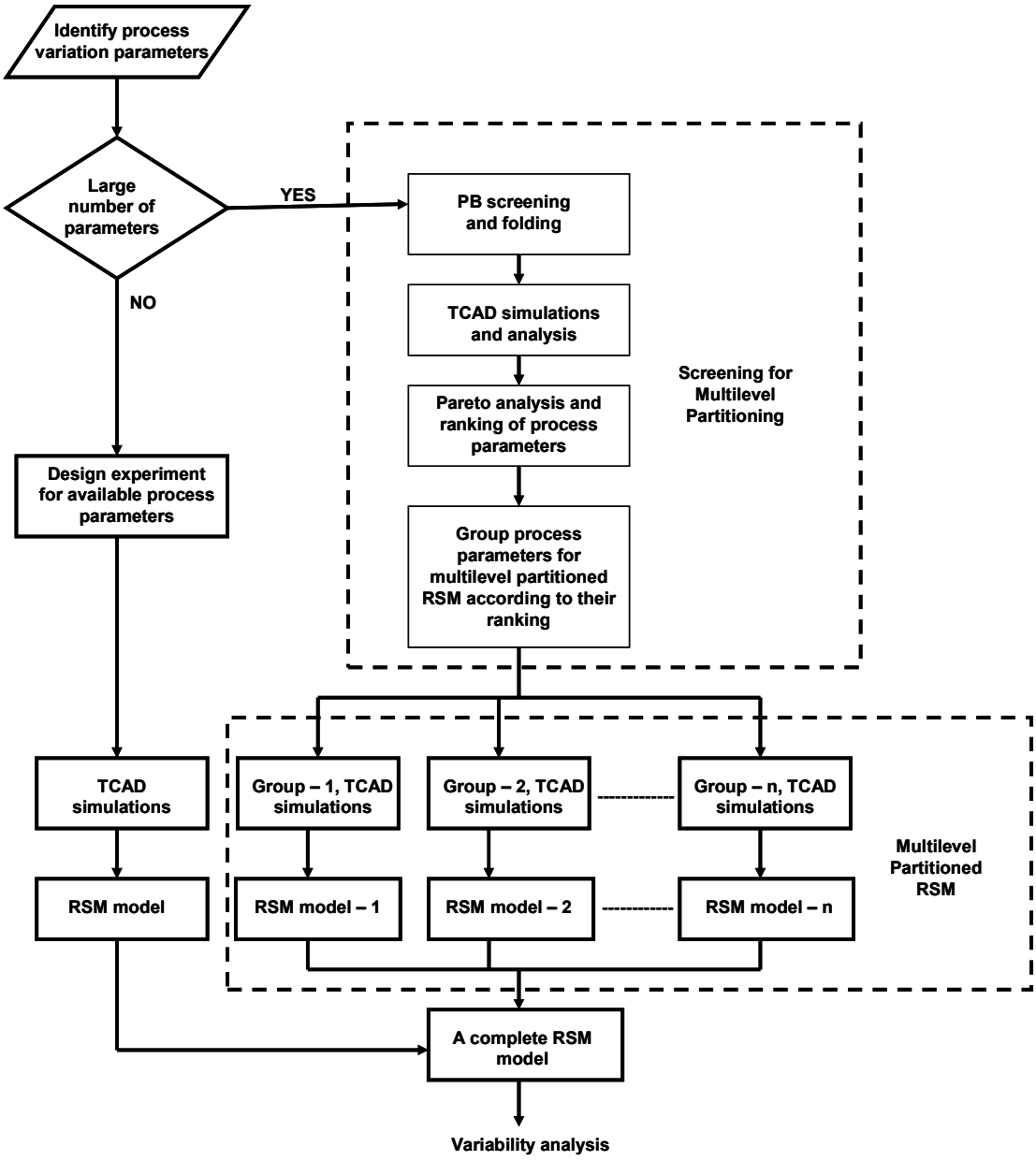


Fig. 17 Flow diagram of the basic RSM and Multilevel Partitioned Response Surface Approach

The overall approach to model and analyse the process variability, is shown in Figure 17. Our methodology consists of TCAD simulations and employing DoE techniques such as RSM for statistical analysis and modelling. For a given process and technology, the first step is to identify or consider the variation parameters which have a significant impact on the device or even circuit response parameters. If the number of parameters to be considered for the process under study is large, screening is employed. Screening is usually used when there is large number of parameters (say > 10). It filters out the process variation parameters which are less significant. The

drawback of screening is the assumption that the main factor effects dominate the response and all the interaction effects are negligible. In our approach, we have used folding technique to overcome this drawback. The folding experiment is performed using TCAD simulations and a Pareto analysis of device responses is carried out to rank the significant parameters. The parameters are then grouped according to their ranking. This multilevel partitioning screening using folding technique allows us to consider all the process parameters for the second order response surface modelling of variability. TCAD simulations are then performed according to the partitioned groups and using the results of simulations, response surface models are constructed for each group. Finally, a complete RS model is obtained using the multilevel partitioning technique described earlier.

If the number of process parameters identified is small, then there is no necessity of screening and the experiment is designed directly to perform RSM which then gives a model in terms of all the process parameters.

6.1 Modelling Variability for 65nm Technology

The multi-level partitioned response surface modelling technique is applied to study the device parameter variations as a function of manufacturing process variations. In this section we present the NMOS and PMOS device characteristics obtained from calibrated TCAD simulations for 65nm technology. The published experimental data from [16 13] and the latest ITRS data for 65nm technology node were used for the calibration of process and the device characteristics simulations. The process and device simulations are performed using two dimensional process and device simulators – TSUPREM4 and MEDICI. The structural features of the NMOS and PMOS devices included poly-silicon gate length of 25nm, equivalent gate-oxide thickness of 1.1nm and side-wall spacer thickness of 30nm. The process simulation included the necessary steps such as transient enhanced diffusion (TED), retrograde channel doping, super-halo doping. The device simulation included drift-diffusion model, carrier continuity, energy balance equations and reverse short channel effects. Also, the band to band tunnelling model is turned on and quantum mechanical effects are taken into consideration to obtain the accurate device characteristics.

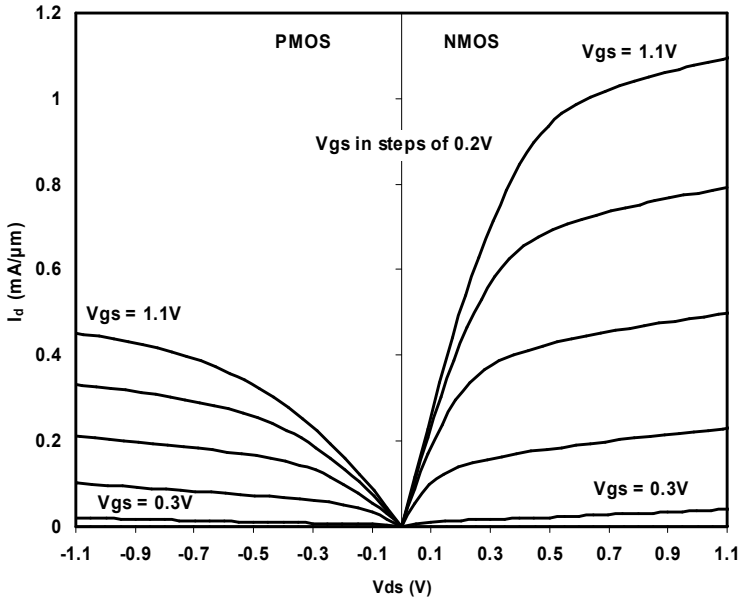


Fig. 18 I_d-V_{ds} characteristics for the calibrated NMOS and PMOS devices at different gate biases.

The device characteristic results for these devices are shown in Table 3.

Table 3: Nominal device characteristics for 65nm technology

Device	I_{on} (mA/µm)	V_{th} (V)	$I_{sd,leak}$ (µA/µm)	SS (mV/dec)
NMOS	1.1	0.158	0.072	89.1
PMOS	0.465	0.14	0.014	75.81

The drain saturation current (I_{dsat}), threshold voltage (V_{th}), leakage current ($I_{sd,leak}$) and subthreshold slope (SS) were targeted as the device performance measure of interest for modelling variability. Here, the drain saturation current is defined as I_d obtained for different gate biases and at $V_d = \pm 1.1V$. The threshold voltage is extracted using the linear extrapolation method. The leakage current is defined as I_d observed for gate voltage, $V_{gs} = 0V$.

For the study and modelling of variability of these device performance parameters, 17 relevant process parameters were identified as the sources of uncontrollable variation for the 65nm NMOS technology, as shown in Table 4. The process parameters identified for PMOS are similar to that of NMOS. All the process parameters variation range was set to ±10% of their mean values with the exception of temperatures. The temperature variation range was set to ±10 ^oC. This is because the

temperature values are very high and practically it would not vary in the range of ±10%. It is assumed that $\pm 10\%$ and ±10 ^{o}C, corresponds to $\pm 3\sigma$ variation for this process under study.

Table 4: Uncontrollable Process parameters variation for 65nm NMOS technology

Symbol	Description	Unit	Mean Value	Lower level value (-1)	Higher level value (+1)
x_1	gate length	nm	25	22.5	27.5
x_2	P – well implant dose	atoms/cm^2	1x10^{12}	0.9x10^{12}	1.1x10^{12}
x_3	P – well implant energy	keV	40	36	44
x_4	P – well diffusion temp.	^{o}C	1000	990	1010
x_5	V_{th} implant dose	atoms/cm^2	4x10^{13}	3.6x10^{13}	4.4x10^{13}
x_6	V_{th} implant energy	keV	8.5	7.65	8.85
x_7	gate – oxide diffusion temp.	^{o}C	800	790	810
x_8	poly Si. Thickness	nm	60	54	66
x_9	pocket halo dose	atoms/cm^2	8x10^{12}	7.2x10^{12}	8.8x10^{12}
x_{10}	pocket halo energy	keV	11	9.9	12.1
x_{11}	pocket halo tilt angle	^{o}	35	31.5	38.5
x_{12}	LDD implant dose	atoms/cm^2	1.4x10^{14}	1.26x10^{14}	1.54x10^{14}
x_{13}	LDD implant energy	keV	6	5.4	6.6
x_{14}	side–wall spacer thickness	nm	30	27	33
x_{15}	S/D implant dose	atoms/cm^2	1.2x10^{15}	1.08x10^{15}	1.32x10^{15}
x_{16}	S/D implant energy	keV	10	9	11
x_{17}	RTA temp.	^{o}C	1000	990	1010

Thus to investigate 17 process parameters using standard central composite designed experiment [14], it would take 131,107 process as well as device simulations. However, constructing response surface models for 17 process parameters from these 131,107 simulations and analyses will be computationally inefficient and would not be manageable. One approach, here, can be to use a screening technique and filter out the most significant parameters from those 17 process parameters and do RSM only on the screened significant parameters. But as mentioned earlier, the major limitation of screening, such as PB screening is that – it screens out only the main effects and moreover, the main effects are confounded with the second order interactions. Also, while modelling variability for very small devices it is important to consider all the process parameters instead of just few screened out. Thus, the multi-level partitioned response surface approach is appropriate here. In our study we have used the PB

screening, but not for screening but as an aid to partition or group the process parameters according to their relative importance. The partitioning of 19 process parameters using PB technique is explained in the next section.

6. 2 Multi-Partitioning of 17 process parameters using PB-screening technique.

The next problem to address was to partition or group the initial 17 process parameters. Screening experiment is a part of DoE, where, the less significant parameters are screened out. This step is commonly used to reduce the dimensionality of the input parameter space, required for a detailed DoE/RSM study. This is because performing RSM for large input space becomes computationally inefficient. But here, we have used the PB screening technique, not to screen the significant/insignificant process parameters but to partition the process parameters in different groups according to their relative importance.

For the 17 process parameters, we have 20-run PB design [10, 14] requiring a total of 20 simulation runs. The limitation of PB screening is that the main effects screened out are often confounded with the two-factor interactions giving a wrong estimation of the screened out significant parameters; which themselves can be significant two factor interaction but not the individual or main effect. Confounding occurs when one or more effects in fractional factorial designs such as PB cannot be estimated separately. The effects that cannot be separated are said to be aliased. To overcome this limitation, we have used a technique called 'folding' which reduces the confounding. In case of PB screening, it entangles all the second order interactions from the main effects. Folding is obtained by reversing all the signs (reversing the factor values) from original runs and hence would require additional runs. In this case, it took additional 20 runs to perform folding. Hence, 40 process and device simulation experiments, each for NMOS and PMOS, were carried out as part of PB design with folding. Pareto analysis is carried out to analyse these experimental results to rank the importance of the process parameters for each response. The responses are device performance parameters as in Table 3.

The Pareto plots of standardized effect of I_{dsat} and $I_{sd,leak}$ for 65nm NMOS device, are presented in Fig. 19 and 20 respectively. The Pareto plot compares the relative magnitude and the statistical significance of all the main effects (process parameters) and ranks the parameters accordingly. The effect plots are in the

decreasing order of the absolute value of the effects. It is important to note here that the ranking of process parameters is different for different responses. From Fig. 19 it can be seen that for the response, I_{dsat}, *x_7 (gate-oxide diffusion temperature)*, *x_1 (gate-length), x_{17} (RTA temperature), x_5 (V_{th} implant dose)* and *x_{12} (LDD implant dose)* are the top five significant parameters. But for the response $I_{sd,leak}$, the top five significant parameters are *x_1 (gate-length), x_{17}(RTA temperature), x_{12} (LDD implant dose)* and *x_2 (P-well implant dose)* and *x_8(poly silicon thickness).*

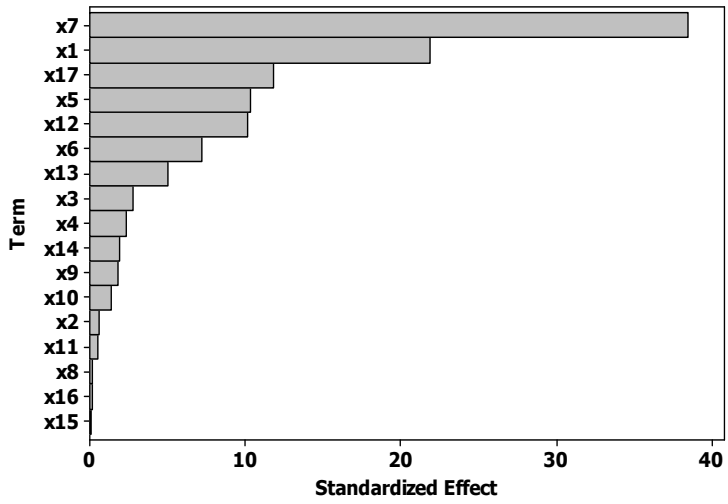


Fig. 19 Pareto plot for I_{dsat} for NMOS device

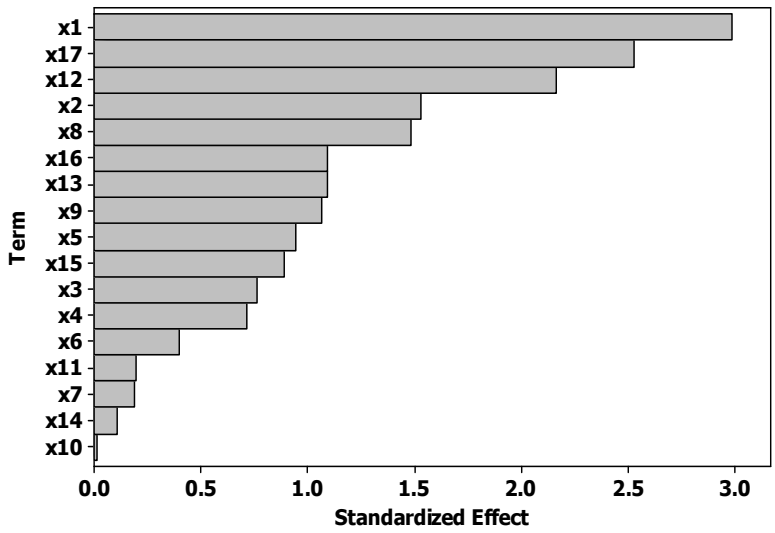


Fig. 20 Pareto plot for $I_{sd,leak}$ for NMOS device

A similar analysis was also done for V_{th} and SS. Hence, for all the multiple responses, the ranking of the process parameters was different. One approach for higher order building such as RSM, from these screening results, is to partition the parameters separately and build the response surface models for each response, varying only the parameters partitioned for that response. But with this approach, the experimentation time is increased by the factor of responses and also the resulting models are not consistent – parameters varied in one model are fixed in other.

The partitioning of process parameters for the 65nm technology is done considering the significant process parameters from PB screening and their ranking for all the responses from the analysis of Pareto charts. The 17 process parameters are partitioned into three groups by combining the set of significant parameters for each response. Tables 5-7 shows the partitioned process parameters for NMOS. The process parameters for all the responses are partitioned into groups according to their significance or rankings. One important consideration while grouping the process parameters was to put the implant dose and implant energy in one group. This is because, at the device level, the implant dose and energy parameters directly affect the doping in the device. For example, x_{13} *(LDD implant energy)* is not ranked in the top five significant parameters for any of the device responses (Fig. 19 and 20), but it is still included in the group 1 (Table 5). This is because x_{12} *(LDD implant dose)* is the significant parameter for most of the device responses and both x_{12} and x_{13} constitute towards the overall LDD doping of the device.

Table 5: Group – I, Partitioned NMOS process parameters from Pareto analysis

Symbol	Description	Unit	Mean Value	Lower level value (-1)	Higher level value (+1)
x_1	gate length	nm	25	22.5	27.5
x_4	P - well diffusion temp.	^oC	1000	990	1010
x_7	gate-oxide diffusion temp.	^oC	800	790	810
x_{12}	LDD implant dose (As)	atoms/cm^2	1.4x10^{14}	1.26x10^{14}	1.54x10^{14}
x_{13}	LDD implant energy	keV	6	5.4	6.6
x_{17}	RTA temp.	^oC	1000	990	1010

Table 6: Group – II, Partitioned NMOS process parameters from Pareto analysis

Symbol	Description	Unit	Mean Value	Lower level value (-1)	Higher level value (+1)
x_5	V_{th} implant dose (B)	atoms/cm^2	4x10^{13}	3.6x10^{13}	4.4x10^{13}
x_6	V_{th} implant energy	keV	8.5	7.65	9.35
x_9	pocket halo dose (B)	atoms/cm^2	8x10^{12}	7.2x10^{12}	8.8x10^{12}
x_{10}	pocket halo energy	keV	11	9.9	12.1
x_{11}	pocket halo tilt angle	^{o}	35	31.5	38.5
x_{14}	side-wall spacer thickness	nm	30	27	33

Table 7: Group – III, Partitioned NMOS process parameters from Pareto analysis

Symbol	Description	Unit	Mean Value	Lower level value (-1)	Higher level value (+1)
x_2	P - well implant dose (B)	atoms/cm^2	1x10^{12}	0.9x10^{12}	1.1x10^{12}
x_3	P - well implant energy	keV	40	36	44
x_8	poly Si. thickness	nm	60	54	66
x_{15}	S/D implant dose (As)	atoms/cm^2	1.2x10^{15}	1.08x10^{15}	1.32x10^{15}
x_{16}	S/D implant energy	keV	10	9	11

The Pareto plots of the standardized effects of I_{dsat} and V_{th} for the PMOS device are shown in Fig. 21 and Fig. 22 respectively.

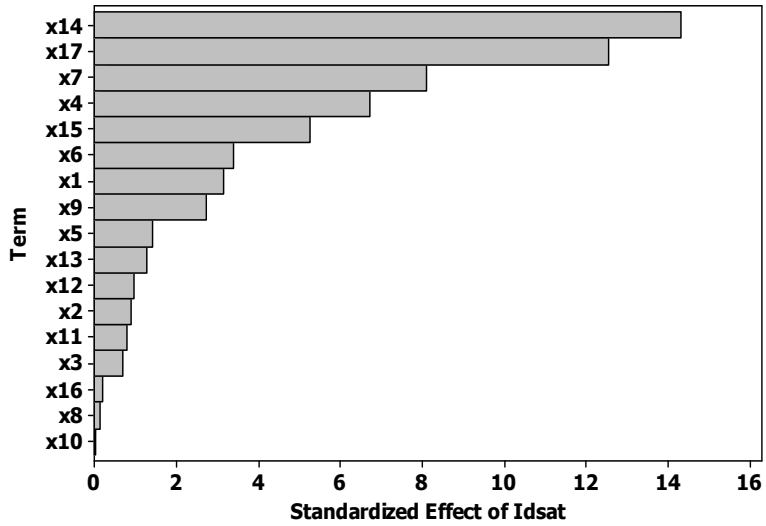


Fig. 21 Pareto plot of I_{dsat} for PMOS device

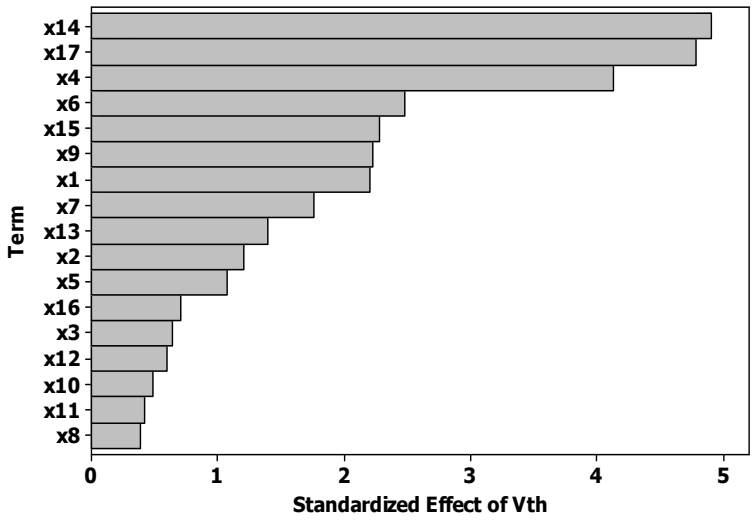


Fig. 22 Pareto plot of V_{th} for PMOS device

The partitioned process parameters for PMOS are shown in Tables 8-10. The parameters are partitioned in groups in similar way, as it was done for NMOS. It is important to note here that the parameters for PMOS are different than NMOS. For example, For example, side wall spacer thickness has the highest impact on the PMOS device as compared to the NMOS device.

The partitioning approach makes the problem of modelling large number of parameters more controllable and allows all the device responses to be modelled in all 17 process parameters and reduces the. Moreover, it significantly reduces the computational efficiency for second order response surface modelling. The next section explains the building of multilevel partitioned response surfaces for the 65nm CMOS technology.

Table 8: Group – I, Partitioned PMOS process parameters from Pareto analysis

Symbol	Description	Unit	Mean Value	Lower level value (-1)	Higher level value (+1)
x_4	N - well diffusion temp.	^oC	1100	1090	1110
x_7	gate-oxide diffusion temp.	^oC	800	790	810
x_{14}	side-wall spacer thickness	nm	30	27	33
x_{15}	S/D implant dose (BF2)	atoms/cm^2	$4.8x10^{14}$	$4.32x10^{14}$	$5.28x10^{14}$
x_{16}	S/D implant energy	keV	7	6.3	7.7
x_{17}	RTA temp.	^oC	1000	990	1010

Table 9: Group – II, Partitioned PMOS process parameters from Pareto analysis

Symbol	Description	Unit	Mean Value	Lower level value (-1)	Higher level value (+1)
x_1	gate length	nm	25	22.5	27.5
x_5	V_{th} implant dose (As)	atoms/cm^2	5x10^{12}	4.5x10^{12}	5.5x10^{12}
x_6	V_{th} implant energy	keV	7	6.3	7.7
x_9	pocket halo dose (As)	atoms/cm^2	9x10^{12}	8.1x10^{12}	9.9x10^{12}
x_{10}	pocket halo energy	keV	25	22.5	27.5
x_{11}	pocket halo tilt angle	^{o}	35	31.5	38.5

Table 10: Group – III, Partitioned PMOS process parameters from Pareto analysis

Symbol	Description	Unit	Mean Value	Lower level value (-1)	Higher level value (+1)
x_2	N - well implant dose (P)	atoms/cm^2	3x10^{15}	2.7x10^{15}	3.3x10^{15}
x_3	N - well implant energy	keV	40	36	44
x_8	poly Si. thickness	nm	60	54	66
x_{12}	LDD implant dose (BF2)	atoms/cm^2	1x10^{13}	9x10^{12}	1.1x10^{13}
x_{13}	LDD implant energy	keV	7	6.3	7.7

6.3 Multi-Level Response Surface Modelling for 65nm Technology

The process parameters partitions are used for the multi-level second order response surface modelling of the device responses - drain saturation current (I_{dsat}), threshold voltage (V_{th}), leakage current ($I_{sd,leak}$) and subthreshold slope (SS). We have used the Central Composite Design (CCD) [11] for approximating a second-order response surface model and studying the second-order effects. CCD is the first order design augmented with an additional axial or star points and centre points making it a second-order surface with lesser design points than would be required for a three-level full factorial design. The CCD is designed for six NMOS process parameters partitioned for group – I (Table 4), with a total of 77 experiments (2^6 factorial points, 2·6 axial points and one centre point). Hence 77 process and device simulations were

performed for the second order RSM. While varying these six parameters from group – I, the parameters from other groups were kept at their nominal values. Similarly, 77 and 43 process and device simulations were performed for group – II (Table 5) and group – III (Table 6) respectively, as a part of this multilevel partitioned response surface modelling. For each of the responses for all the groups, regression analyses were then carried out to obtain the coefficients for all the device responses. The equations (7) – (9) give the RS models for I_{dsat} (NMOS), in terms of the process parameters. The model in equation (7) is build from the group – I parameters and that of (8) and (9) from group – II and group – III respectively. Here, to capture the effects of the parameters of groups – II and III in group I, firstly the mean (a_0) of the responses of group – I parameters is fitted as a function of group – II parameters. The mean term (b_0) of the responses of group – II is fitted as a function of group – III parameters, thus creating the three-level RSM. Hence the response, I_{dsat} is a function of all the process parameters. The models for the other device parameters responses for NMOS and PMOS were also constructed using the same approach.

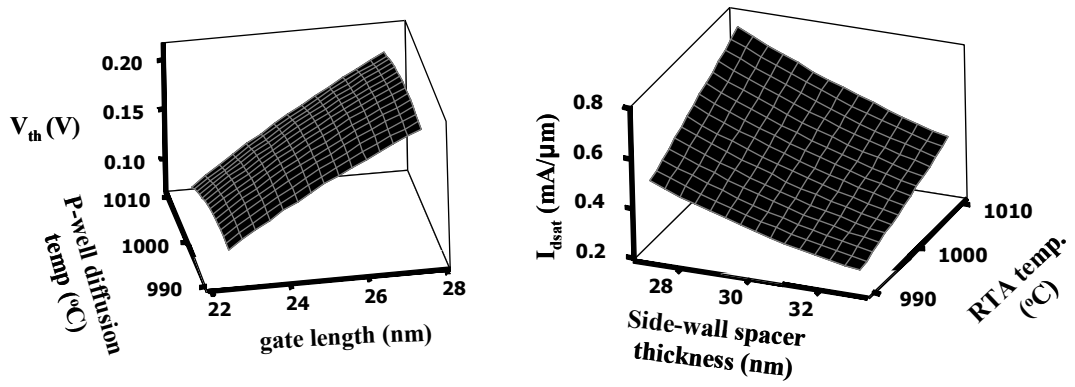
$$\begin{aligned} I_{d_{sat}}(mA/\mu m) = a_0 &- 0.1225x_1 + 0.0316x_4 - 0.2205x_7 + 0.0562x_{12} \\ &+ 0.0301x_{13} + 0.0617x_{17} + 0.0123x_1^2 - 0.0058x_1x_4 \\ &+ 0.0059x_1x_7 - 0.0108x_1x_{12} - 0.0056x_1x_{13} \\ &- 0.0136x_1x_{17} + 0.03x_4^2 - 0.005x_4x_7 + 0.0027x_4x_{12} \\ &- 0.0016x_4x_{13} + 0.006x_4x_{17} + 0.0053x_7^2 - 0.006x_7x_{12} \\ &+ 0.0022x_7x_{13} - 0.004x_7x_{17} - 0.001x_{12}^2 + 0.002x_{12}x_{13} \\ &+ 0.0091x_{12}x_{17} - 0.0022x_{13}^2 + 0.001x_{13}x_{17} + 0.0078x_{17}^2 \end{aligned} \tag{7}$$

$$\begin{aligned} a_0 = b_0 &- 0.0557x_5 + 0.0567x_6 - 0.0095x_9 + 0.0106x_{10} - 0.0012x_{11} \\ &- 0.0102x_{14} + 0.0026x_5^2 - 0.0001x_5x_6 + 0.0008x_5x_9 - 0.001x_5x_{10} \\ &+ 0.0008x_5x_{14} + 0.0016x_6^2 - 0.0008x_6x_9 + 0.001x_6x_{10} - 0.0007x_6x_{14} \\ &+ 0.0001x_9^2 + 0.001x_9x_{10} + 0.0002x_9x_{14} + 0.0001x_{10}^2 + 0.0003x_{10}x_{11} \\ &- 0.0001x_{10}x_{14} + 0.0006x_{11}^2 + 0.0011x_{14}^2 \end{aligned} \tag{8}$$

$$b_0 = 1.0947 - 0.0001x_2 - 0.029x_3 - 0.0002x_8 + 0.0016x_{15} + 0.0011x_{16} - 0.0002x_2^2 - 0.0001x_2x_8 + 0.031x_3^2 + 0.0001x_3x_{15} + 0.0001x_3x_{16} + 0.0003x_8^2 - 0.0002x_{15}^2 - 0.0002x_{16}^2$$

(9)

In the above model equations, x_1, x_2, x_3....x_{17} represents the process parameters for groups I – III with its mean and ±3σ deviations given in Tables 5-7. The variability analysis was performed based on the model equations. RS plots were generated for all the responses as function of process parameters. Fig. 23 (a) shows the RS plot for V_{th} with respect to the most significant process parameters of group – I (NMOS) – x_1 (gate length) and x_4 (P-well diffusion temperature). As depicted from the figure, the variation in x_1 and x_4 causes a variation of +27% and −37% in V_{th} from its mean value.



(a) (b)

Fig. 23. (a) RS plot of V_{th} (NMOS) as a function of P-well diffusion temperature and gate length and (b) RS plot of I_{dsat} (PMOS) as a function of side-wall spacer thickness and RTA temperature.

The variation in I_{dsat} (PMOS) for group – I process parameters is shown in Figure 23 (b). The figure shows the RS plot as a function of x_{14} (side-wall spacer thickness) and x_{17} (RTA temperature). There is +70% and -36% variation in I_{dsat} from its mean value, due to ±10% and ±10°C variation in x_{14} and x_{17} respectively.

The accuracy and the validity of the models are checked using the goodness of the second order fit, R^2 (R-square) and R^2_{adj} (Adjusted R-square) [14]. R^2 is the

coefficient of determination and is a statistical measure of how well the regression line approximates the actual data points. The R^2 value can be artificially increased by including additional parameters to the model which may not be statistically significant. Hence models with large R^2 values can give poor functional fit and hence poor predictions of the estimates of the response. R^2_{adj} is modified R^2 for the number of terms in the model. Unlike R^2, R^2_{adj} may get smaller when unnecessary or additional terms are added to the model.

The resulting models fits for the NMOS and PMOS device responses for different groups are summarised in Tables 11 and 12 respectively. While fitting the model for leakage current ($I_{sd,leak}$), initially the R^2 and R^2_{adj} values for this model were very poor. This was because the $I_{sd,leak}$ values were very low and the range of variation were very large. Hence to improve the model performance, logarithmic transformation was performed on these values.

Table 11: Response Surface Model Fits for NMOS device responses

Device Responses	Group - I		Group - II		Group - III	
	R^2 (%)	R^2_{adj} (%)	R^2 (%)	R^2_{adj} (%)	R^2 (%)	R^2_{adj} (%)
I_{dsat}	100	100	100	100	100	100
V_{th}	99.9	99.8	100	100	99.9	99.9
$log(I_{sd,leak})$	98.4	97.6	99.6	99.4	99.6	99.2
SS	94.4	91.4	98.8	98.1	99.8	99.5

Table 12: Response Surface Model Fits for PMOS device responses

Device Responses	Group - I		Group - II		Group - III	
	R^2 (%)	R^2_{adj} (%)	R^2 (%)	R^2_{adj} (%)	R^2 (%)	R^2_{adj} (%)
I_{dsat}	99.9	99.9	100	100	99	98.1
V_{th}	97.8	97.5	100	100	94.5	89.6
$log(I_{sd,leak})$	97.1	95.5	100	100	96.1	92.6
SS	99.3	99.8	99.6	99.3	99.6	99.3

The R^2 and R^2_{adj} values close to 100% are desired in fitting the RS models. The high values of R^2 and R^2_{adj} in most of the responses shows that the second order RS models capture a large portion of the observed variance. It can also be seen that, performing logarithmic transformation of log ($I_{sd,leak}$) has improved the model efficiency for leakage current.

Figure 24 shows the graphical inspection of response model fits for the threshold voltage (V_{th}) of group – I (Table III) NMOS and leakage current log ($I_{sd,leak}$) of group – III (Table 10) PMOS parameters. The plot shows the model fit V_{th} against the actual V_{th} response. The angled straight line (45^o straight line) represents the ideal fit where the fitted values are scattered. The vertical line in the plot represents the mean value of the response, V_{th}. As the values of the $I_{sd,leak}$ for the PMOS device were negative, its absolute values were taken before performing the logarithmic transformation. In the response model fit plot of log ($I_{sd,leak}$), the model fit values are slightly deviating from the actual values. This is also reflected in the lower R^2 (96.1%) and R^2_{adj} (92.6%) values from Table 12. This is because the response values of $I_{sd,leak}$ are changing largely and in discrete steps. The model is still accepted because it if for group – III parameters in which lower ranked or less significant parameters are modelled.

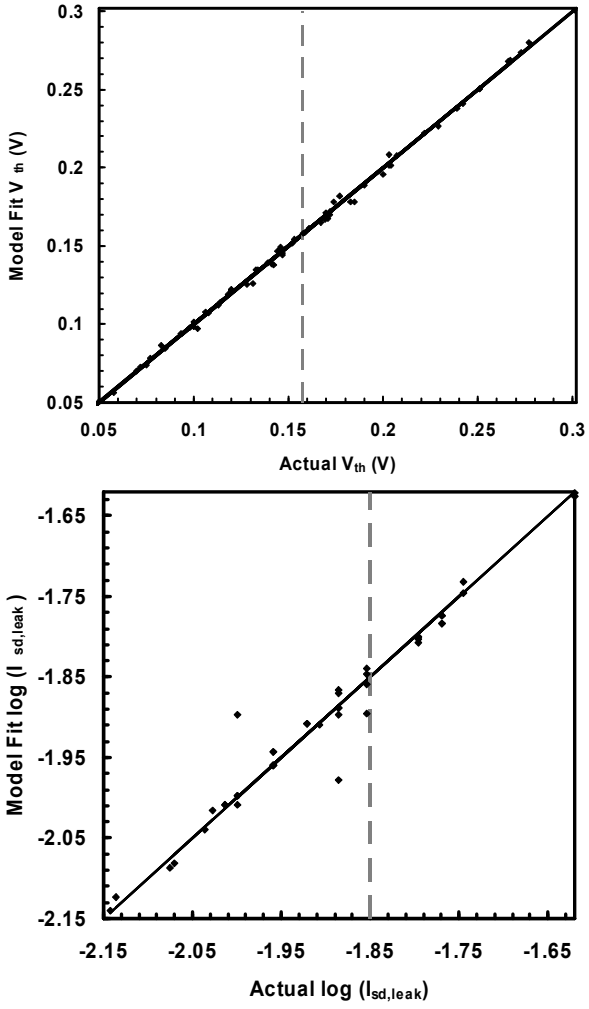


Fig. 24 RS model fit vs. Actual fit for V_{th} and log ($I_{sd,leak}$)

7. Analysis of circuit design parameters with process variability

This section discusses the impact of process parameter fluctuations on the output variables of interest for the MUTEX and Jamb Latch synchronizer realised with strained-Si technology. These circuits are also realised with standard Si technology to see the advantages of using strained-Si technology in asynchronous circuits.

7.1 Mutual Exclusion Element (MUTEX)

Asynchronous communication systems require its input channels to be mutually exclusive. In this type of systems there are several situations where a resource is shared between several independent processes [17]. Instrumental in this sharing process is a basic circuit called a MUTEX whose block diagram is shown in Figure 25.

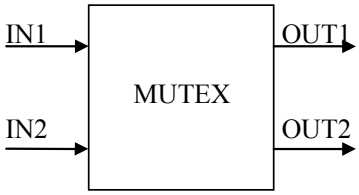


Fig. 25 Block diagram of a MUTEX with two input signals.

The input signals IN1 and IN2 shown in Figure 25 are two signals that originate from two independent sources, and the task of the MUTEX is to pass these inputs to the corresponding outputs OUT1 and OUT2 in such a way that only one of the outputs is active at any given time [18]. However in the case where there is only one input request which is active then the operation is insignificant. If the first input request arrives earlier than the second one, the latter request is blocked until the first request is de-asserted. However there is a problem when both the input signals IN1 and IN2 are asserted at the same time. In this instance the MUTEX has to decide upon which input should be given priority; the length of the time taken to resolve the situation is called 'metastability resolution time' denoted by t_m. A similar situation can arise in synchronous memory element when data and clock signals violate setup and hold times. The minimum separation time between the input signals to ensure stable operation of the MUTEX is called the offset time t_{off}. Both t_m and t_{off} are critical parameters in the design of a MUTEX and are affected by process variations.

The implementation of the MUTEX is shown in Figure 26 involves a pair of cross coupled NAND gates and a metastability filter.

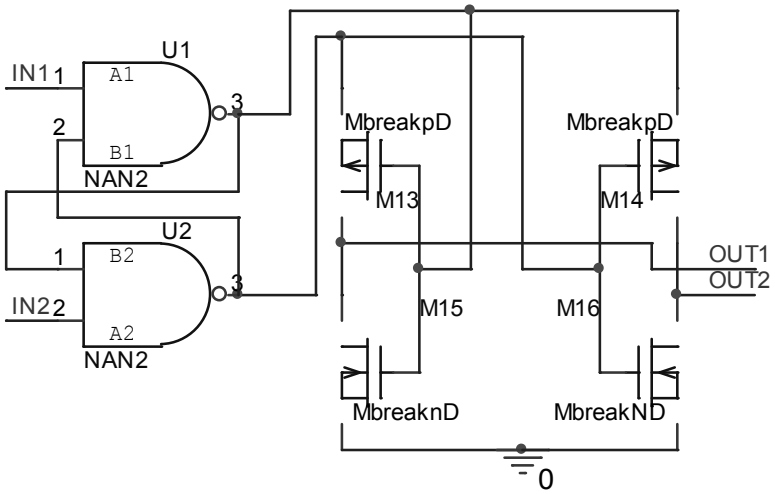


Fig. 26 MUTEX implemented using cross coupled NAND gates with metastability filter.

The cross coupled NAND gates enable one input to block the other. If both inputs IN1 and IN2 are asserted at the same time, the circuit becomes metastable with both signals U1 and U2 sitting at a voltage halfway between the supply voltage and ground. The metastability filter prevents these undefined values from propagating to the outputs; OUT1 and OUT2 are both kept low until signals U1 and U2 differ by more than a transistor threshold voltage. The metastability filter can be implemented using two buffers whose logic thresholds have been made particularly high (or low) by trimming the strengths of the pull-up and pull-down transistor paths.

7.1.1 Variability analysis: MUTEX

The flow chart of the analysis to study the variability of the process parameters in a MUTEX realised with strained-Si and standard Si technology is shown in Figure 27. The most significant parameters have to be identified using Plackett-Burman screening method. The significant parameters are chosen based on a threshold value below which it is considered that the input parameter is insignificant. Considering strain as an important parameter, it is added with the significant parameters obtained after screening to model the variability using RSM. Here the widths of transistors in MUTEX were kept constant to study the variation of t_m and t_{off}

based on process variation. Even though the variation of widths of transistors in MUTEX is very important as it is typical for different applications using MUTEX, it is not considered here for the analyses.

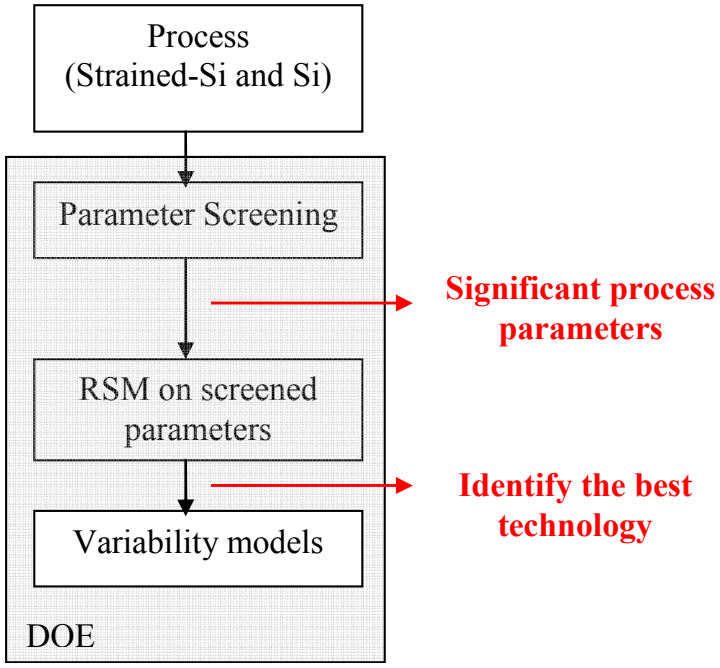


Fig. 27 Flow chart of the variability analysis of MUTEX.

7.1.2 Impact of Process variation on the behaviour of the MUTEX

The transistors used to implement a MUTEX are subjected to process variability conditions to study its effects on t_m and t_{off} using DOE/RSM methods. The supply voltage is fixed at 1V. The first phase of the variability analysis is to find out the significant process parameters (and their variation) which have the greatest impact on t_m and t_{off}. All the process parameters shown in Table 13 were varied by ± 10%. However, the amount of strain is varied from 0% (conventional Si technology) to 0.99% (strained-Si). From the parameter screening technique (Placket-Burman method), Oxide thickness (t_{ox}), N_{sub} (substrate doping) and the amount of strain are found to be the most significant parameters which impact mostly on t_m and t_{off} [15-17]. RSM is then performed for the three parameters. Since only three parameters are identified as being the most significant parameters in the case of MUTEX after the

screening technique, only fifteen (2^n+2n+1) simulations were required to generate the response surface model of the MUTEX outputs. The statistical p-values were obtained from the analysis. P-values provide the way of testing the relationship between the predictor (input variables which are significant process parameters) and response (output variables which are t_m and t_{off}) and are used to determine statistically significant terms in the model. The p-values for linear, square and interaction effects obtained by surface modelling are shown in Table 13.

Effects	Process parameter	t_m	t_{off}
		p- values	p- values
Linear	t_{ox}	0.454	0.746
”	N_{sub}	0.140	0.016
”	strain (%)	0.008	0.196
Square	t_{ox} x t_{ox}	0.989	0.742
”	N_{sub} x N_{sub}	0.999	0.800
”	Strain (%) x strain (%)	0.028	0.181
Interaction	t_{ox} x N_{sub}	0.580	0.336
”	t_{ox} x strain (%)	0.912	0.628
”	N_{sub} x strain (%)	0.004	0.010

Table. 13 The linear, square and interaction effects of process parameters on t_m and t_{off} of MUTEX.

The terms with p-values ≤ 0.5 are statistically significant. From the linear effects it can be seen that strain is the most significant process parameter which impacts t_m, and for t_{off} it is N_{sub}. For square effects; the strain value has the highest impact on t_m and t_{off} of the MUTEX. However, among the interaction effects the interaction of N_{sub} and strain are also significant parameters for t_m and t_{off} variation.

Figure 28 shows the surface plot of the impact of interaction effects of N_{sub} and strain on t_{off} of a MUTEX. It can be seen that t_{off} is giving a large range of variation for 0% strain (conventional Si) when N_{sub} is changing from -10% to +10% which is undesirable. When the strain in the channel of the transistors is increasing it can seen that the variation of t_{off} with N_{sub} is not high compared to the lowest strain

(0%) which gives an edge for the strained-Si technology over the conventional Si. This can be attributed to the fact that the enhancement of the mobility of electrons and holes are not the same in the case of strained-Si [20] and hence the transconductance ratio and the change in threshold voltage is not the same as in conventional Si. As mentioned above the strain is found to be the most significant parameter which impacts on t_{off}. This is advantageous due to the fact that the amount of stress applied to the channel to create strain can be controlled by the Ge% in the wafer (in the case of bi-axial strain) or by the process steps in local strain (uni-axial strain).

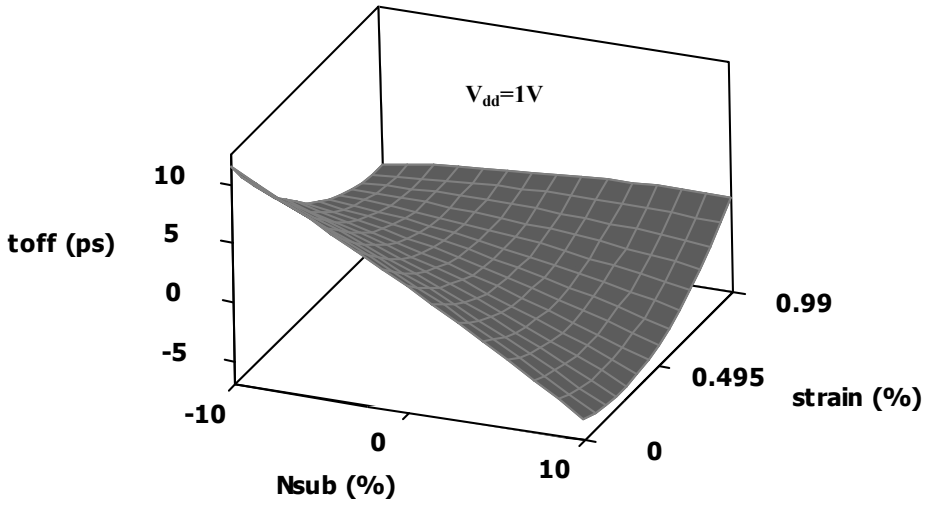


Fig. 28 Surface plot of the impact of t_{off} due to the variation of N_{sub} by 10% and strain varying from 0.0% to 0.99% of MUTEX.

The surface plot to show the impact of interaction effects of N_{sub} and strain on t_m of MUTEX is shown in Figure 29. Similar to t_{off}, t_m is also giving a large range of variation for 0% strain (conventional Si) when N_{sub} is changing from -10% to +10%. t_m of MUTEX is found to decrease when the amount of strain applied in the channel of transistors is increasing. It can also be seen that t_m is almost 100% less when N_{sub} is +10% from the nominal at 0.99% strain compared to the lowest strain (0%) which again indicates that strained-Si technology is a better candidate to realise a MUTEX compared to the conventional Si MUTEX.

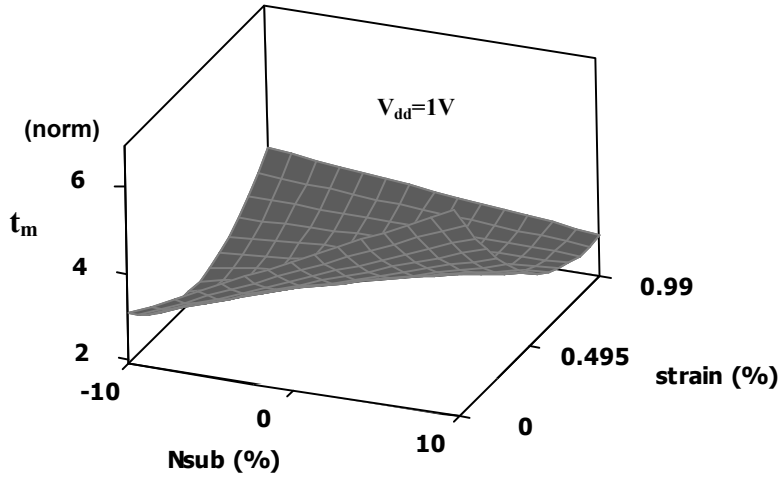


Fig. 29 Surface plot to show the impact of t_m due to the variation of N_{sub} by 10% and strain varying from 0.0% to 0.99% of MUTEX.

Table 14 shows the validity of the model developed to study the impact of process variation on t_{off} and t_m using DOE/ RSM methods. The validity of the model can be checked using the R-square statistics. R-square is known as the coefficient of determination and is a statistical measure of how well the regression line approximates to the real data points. R-square and adjusted R-square statistics represent the amount of variation in the response that is explained by the model. The R-square value always increases with the addition of a variable to the model, regardless of whether the additional variable is statistically significant or not. Hence, models with large R-square values give poor predictions of new observations or estimates of the response. Adjusted R-square is a modified R-square for the number of terms in the model. Unlike R-square, adjusted R-square may get smaller when unnecessary or additional terms are added to the model. It can be seen from Table 14 that the adjusted R-square value is less for both t_m and t_{off}.

	R-sq	R-sq (adj)
t_m	93.36%	81.41%
t_{off}	87.49	64.98%

Table. 14 The validity of the model.

7.2 Jamb Latch Synchronizer

Whenever there is a signal transfer between two systems operating at different frequencies or the same frequency with different phases, the signals coming from the first system become asynchronous to the second system thereby increasing the probability of the output of first system going into metastability and propagating the asynchronous signal into the second system [21]. The most common approach to minimize the metastability problem is to use a synchronizing circuit to take the asynchronous signal and align it to the timing regimen of the rest of the system. The Jamb Latch is a simple circuit commonly used as a synchronizer because of its relatively better performance [19] compared to the conventional two stage synchronizers; its basic configuration is shown in Figure 30.

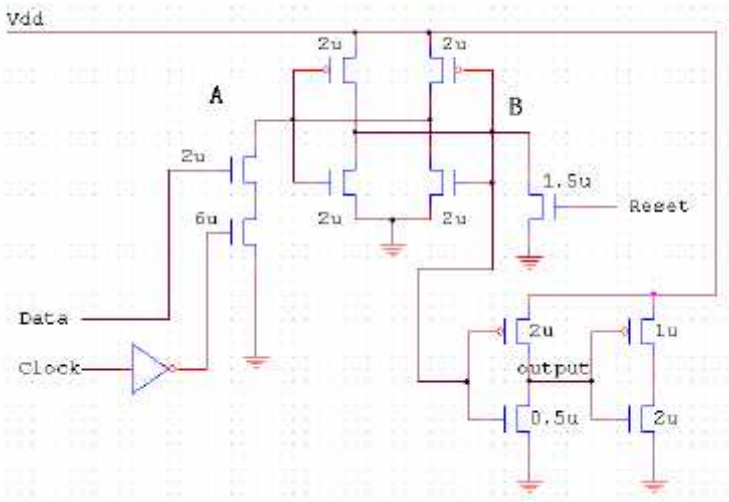


Fig. 30Schematic of Jamb Latch [17].

7.2.2.1 Variability Analysis: Jamb Latch Synchronizer

In this section we explore the possibility of performance enhancement in terms of metastability resolution time in synchronizers using strained-Si transistors. The analysis is based on the Jamb Latch which is a commonly used synchronizer in real time circuits. Although the studies are based on the behaviour of the Jamb Latch synchronizer using strained-Si as the base material, emphasis is given to the comparison of strained-Si technology against that of conventional Si technology with

similar device dimensions. The device dimensions and other physical parameters for DSM high performance (HP) devices are taken from *MASTAR*.

The electrical characteristics of strained-Si and conventional Si n- and p-MOSFETs are applied to AURORA parameter extraction simulator for optimization and extraction of PSPICE models. These transistors are configured for the design of Jamb Latch in PSPICE. Simulations are performed to evaluate the metastability resolution time of strained-Si and Si based Jamb Latch under different operating conditions and dimensions. To study the impact of variability of the process parameters on t_m, DOE/RSM is applied to develop corresponding variability models. Plackett- Burman screening is performed to identify the most significant process parameters which have the greatest influence on t_m. The amount of strain in the channel, the operating temperature and the supply voltage are found to be the most significant parameters affecting the metastability resolution time of the Jamb Latch.

The ratio of widths of p- MOS to n- MOS in Jamb Latch is taken to be 1μm:1μm, which is considered as the best as reported. The simulations are performed at supply voltages ranging from 0.9V to 1.3V for the transistors under different process conditions. The supply voltage is limited to the above range as is specified by ITRS for the HP devices in 65nm technology node. The variation of t_m with respect to the various technology nodes is plotted in Figure 31 for supply voltages of 1.1V and 1.2V. It can be seen from Figure 31 that t_m reduces with scaling of devices.

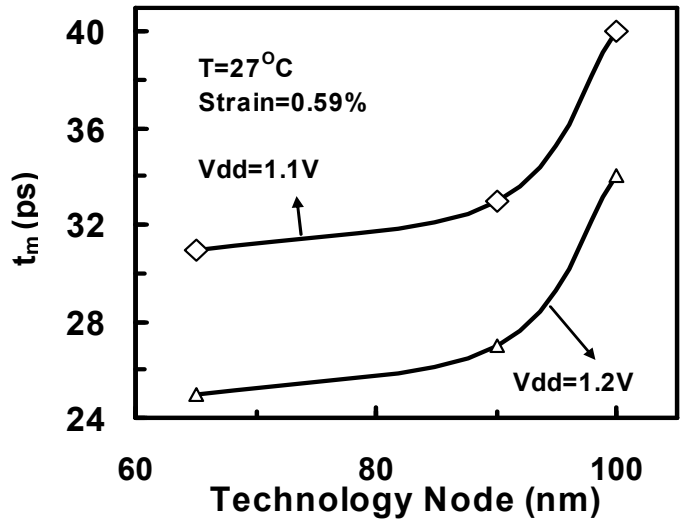


Fig. 31 Metastability resolution time Vs technology node for strained-Si CMOS based Jamb Latch at 1.1V and 1.2V.

Figure 32 shows the variation of t_m against the amount of strain applied in the channel of transistors in same technology node (65nm). For comparison, supply voltages of 1.1V and 1.2V are considered. It can be seen from the figure that the resolution time decreases almost linearly with the increase in the amount of strain applied to the transistors that is, standard Si devices (with 0% strain) are taking more time to resolve the metastability problem compared to the strained-Si devices. The value of t_m is almost 66% less for strained-Si based Jamb Latch compared to that of standard Si Jamb Latch. These results are attributed to the fundamental differences between standard Si and strained-Si technology. The strained-Si device inherently has higher transconductance and lower threshold voltage compared to that of the standard Si due to the presence of strain and smaller band gap. However, it is not expected that the same trend to continue with the increase in the amount of strain.

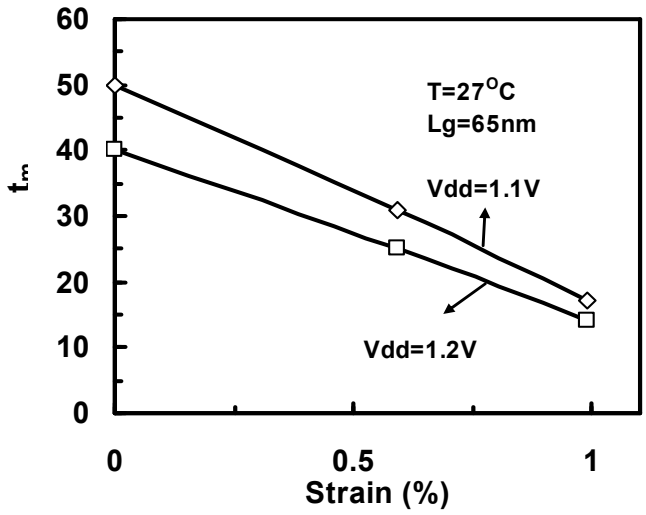


Fig. 32 Metastability resolution time Vs amount of strain in the channel for supply voltages of 1.1V and 1.2V at 65nm technology node.

Figure 33 shows the comparison of t_m of the Jamb Latch realized with transistors with 0.99% and 0.59% amounts of strain in the channel.

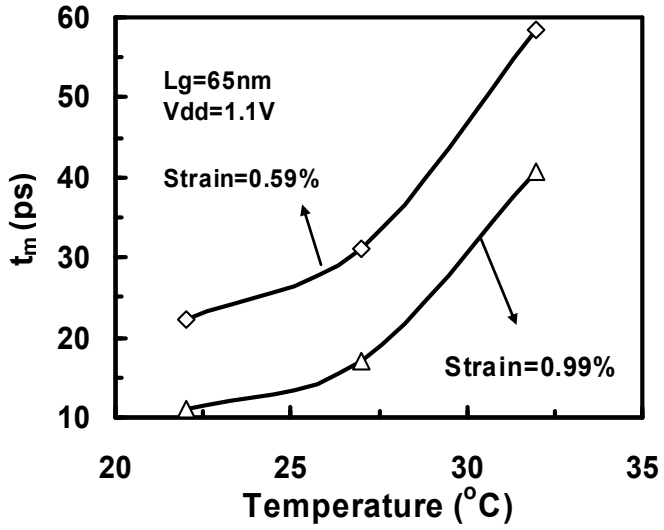


Fig. 33 Variation of t_m against temperature for transistors with different amount of strain in the channel.

A supply voltage of 1.1V is given for the circuit operation. It can be seen from the figure that the strained-Si with 0.99% strain in the channel exhibits a lower metastability time compared to that of strained-Si Jamb Latch with 0.59% for different ranges of temperature. It can also be seen from figure that t_m reduces with the reduction in temperature. This again indicates that strained-Si is a better candidate compared to Si technology for asynchronous circuit applications.

RSM has been used to study the impact of the variation of the amount of strain, supply voltage and temperature on t_m. Figure 34 shows the response surface plot of t_m (ps) as the function of nominal temperature and power supply voltage. The impact of temperature is evident from the response surface. At the low levels of temperatures and supply voltages, t_m has decreased and has its maximum value at the high level of temperature and lower level of supply voltage. Hence the response surface plot shows the range of variations caused in t_m due to the variations in the electrical parameters.

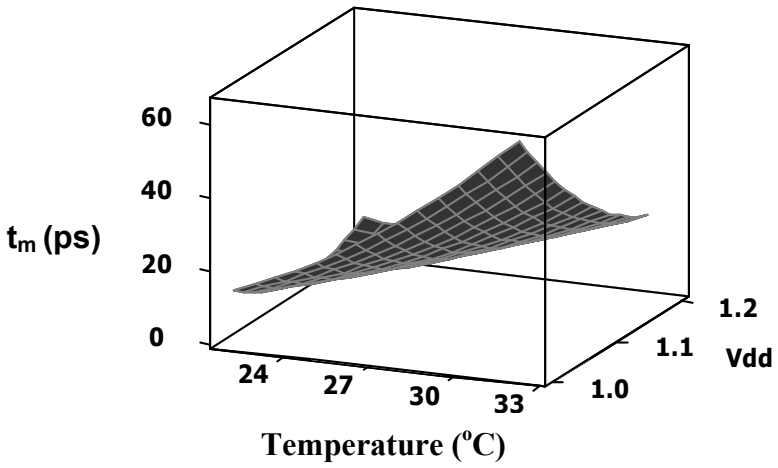


Fig. 34 Response surface plot of t_m as function of operating temperature and supply voltage.

Figure 35 shows the response surface of t_m with respect to amount of strain and supply voltage. The total change in t_m is about 40ps due to the variations of the strain applied in the channel and the supply voltage. Metastability resolution time is lower with greater amounts of strain and lower values of supply voltage. It can be seen from the figure that the impact of the amount of strain is more significant on t_m than the supply voltage variation.

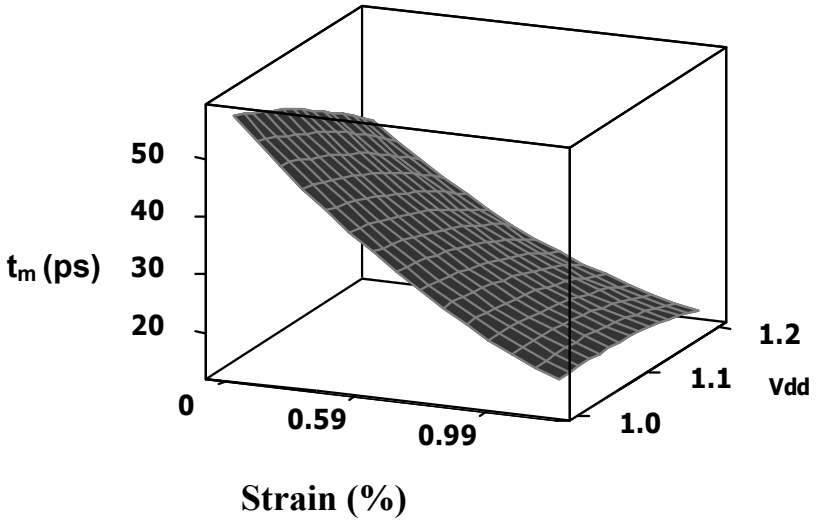


Fig. 35 Response surface plot of t_m as function of amount of strain and supply voltage.

The results obtained from the analyses of MUTEX and Jamb Latch indicate the potential and flexibility of strained-Si technology for asynchronous circuit applications. Strained-Si is found to be a better technology than standard Si for making synchronizers. Besides, it is possible to achieve significantly better low-power, high-speed, overall noise margin and better timing closure performance from strained-Si based circuits than conventional Si circuits. These facts coupled with the availability of higher supply voltage reduction window makes strained-Si based circuits an ideal candidate for futuristic synchronous and asynchronous circuit design. Variability analysis using DOE to study the variation of metastability resolution time with process and operational parameters (V_{dd} and nominal temperature) in three dimensional spaces indicates that the variation of strain in the channel is found to be a significant parameter compared to the supply voltage variation. This information can be used by the designer to accommodate the variations hence ensuring real time operations without fault.

8. Future Work: Reduce Variability Vs Variability Aware??

It's high time to determine whether we need to get around with variability using new technologies to build transistors or build new design so that the circuit is aware of variability and act accordingly. One such idea put forward from the technology side is FINFET.

FinFETs

The FinFET is based on a single gate transistor design and is a non-planar double-gate transistor built on an SOI substrate. The distinguishing characteristic of the FinFET is that the conducting channel is wrapped around a thin Si "fin", which forms the body of the device. The dimensions of the fin determine the effective channel length of the device [19]. A 25-nm transistor which operates on 0.7 Volt was demonstrated in December 2002 by Taiwan Semiconductor Manufacturing Company. The structure of FinFET is shown in Figure 36.

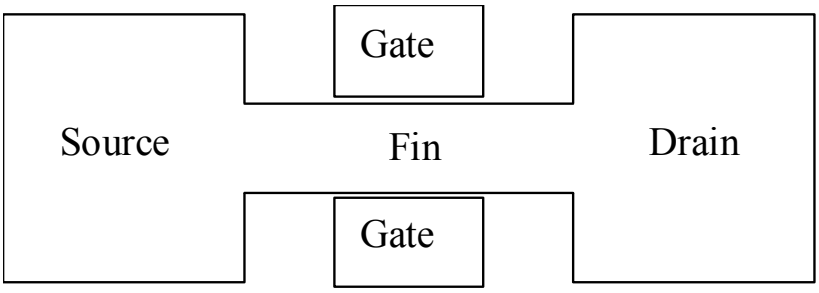


Fig. 36 Structure of a double-gate FinFET device [45].

To build variability tolerant or variability aware designs its important to have tools for the designers to understand the parameters which are open to variability. Automation of different tools is highly desirable as the analysis of variability needs better understanding and knowledge from multiple fields of expertise. The concept of one such tool called "VARMA" is presented here.

8.1 VARMA: VARiability Modelling and Analysis support

Overview of the tool

Parameter variability in IC manufacturing process is becoming increasingly important when it comes to deep submicron devices due to feature scaling. Besides, the impact on circuit performance due to these variations is also increasing and affecting the yield. Even though the introduction of new materials and structures, instead of conventional silicon, is helping to keep the Moore's law alive, the variability effects arising from new technologies, e.g. strain/stress variation arising from self-heating, Ge diffusion (bi-axial strain effects) etc., has to be analysed to help circuit designers to optimize their design for maximum yield. In digital and analogue circuits the output parameters which need to be optimised are different. In digital circuits power and delay are the two parameters which are to be optimised whereas in analogue circuits, depending on the application, there will be more than 2 to 10 parameters. Due to unpredicted variability in the process all these output parameters can be badly affected. Statistical approaches rather than corner case approach will be better suited for designers to analyse performance and productivity due to variability.

CAD tools are vital to perform optimisation and model the variability for accurate variability prediction, thereby increasing the overall productivity of the design.

Methods and Benefits

The flow chart of the simulation flow is shown in Figure 37.

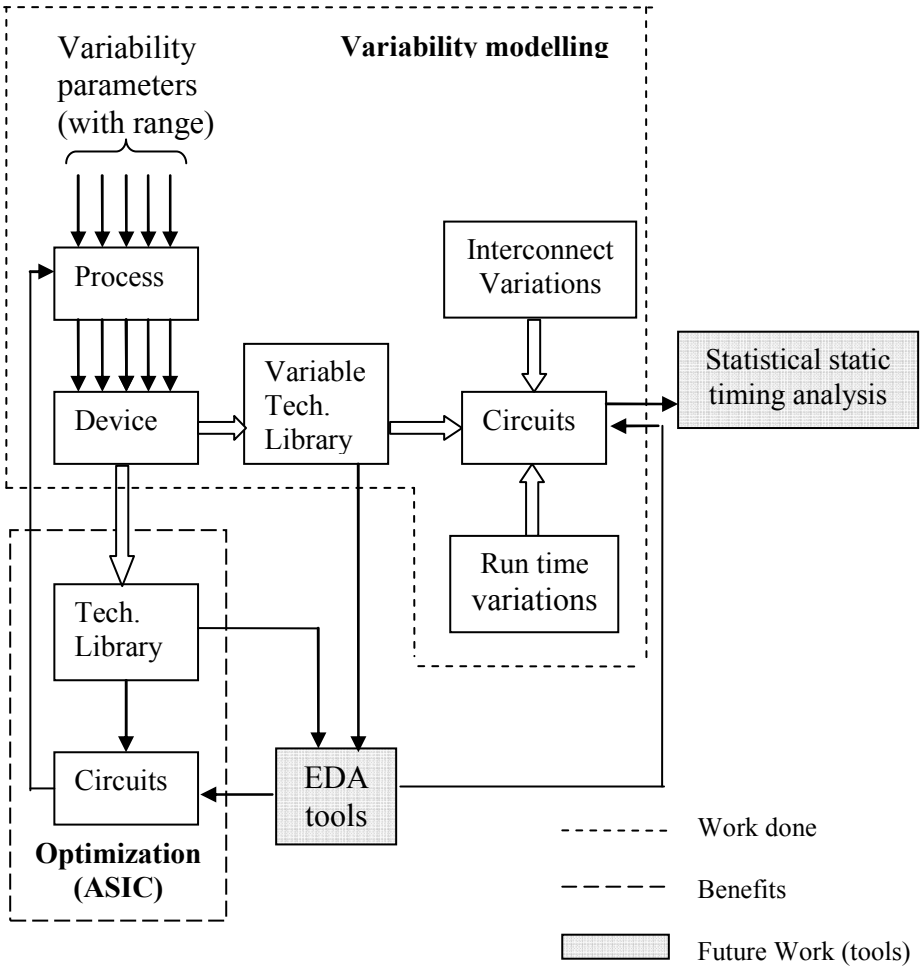


Fig. 37 VARMA Project flow

This project flow can be included in the conventional IC design flow. The VARMA project involves

1. Variability Modelling Part
2. Optimization
3. EDA tools.

VARMA automates 1 and 2 and EDA tool box forms the front-end of the project where the user can work on. The screen shots of VARMA are shown in Part-2 (Appendix A). The variability modelling part includes process; device and circuit simulation or characterisation of CMOS from the wafers, and finding the significant parameters of the process which impacts on the response of circuit outputs. The method underlying VARMA flow involves the variability study for the purpose of optimisation based on applications and variability study to improve the yield. The future work includes the incorporation of the variability analysis method into the normal EDA flow. The Figure 37 shows the general flow of the whole analysis process.

Conclusion

The variability analysis performed here using the statistical technique DOE/RSM seems to be inefficient when large number of process parameters is considered due to the interaction of process parameters. New statistical technique need to be introduced to perform the variability analysis of deep sub-micron devices which are more prone to process variability. The analysis needs to be extended to CMOS devices based on new materials and transistors based on new structures.

References

[1] B. P. Wong, *Nano-CMOS Circuit and Physical Design.* New Jersey: John Wiley, 2005.

[2] S. R. Nassif, A. J. Strojwas, and S. W. Director, "FABRICS II: A Statistically Based IC Fabrication Process Simulator", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 3, pp. 40-46, 1984.

[3] M. T. Currie, C. W. Leitz, T. A. Langdo, G. Taraschi, and E. A. Fitzgerald, “Carrier mobilities and process stability of strained-Si n and p-MOSFETs on SiGe virtual substrates”, *Journal of Vacuum Science and Technology*, Vol. 19, pp. 2268–2279, Nov. 2001.

[4] S. Chattopadhyay, L. D. Driscoll, K. S. K. Kwa, S. H. Olsen, et.al., "Strained Si MOSFETs on relaxed SiGe platforms: performance and challenges", *Solid-State Electronics*, vol. 48, pp. 1407-1416, 2004.

[5] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Fabrication and analysis of deep submicron strained-Si n-MOSFET's", *IEEE Transactions on Electron Devices*, vol. 47, pp. 1406-1415, 2000.

[6] C. Jacoboni, and P. Lugli, *The Monte Carlo method for semiconductor device simulation.* New York: Springer, 1989.

[7] TSUPREM4 user manual: *Two-Dimensional Process simulation program.* Sunnyvale: Synopsys Inc., 2000.

[8] MEDICI user manual: *Two-Dimensional Device simulation program.* Sunnyvale: Synopsys, Inc., 2000.

[9] AURORA user guide: *AURORA simulation program.* Sunnyvale: Synopsys Inc., 2006.

[10] Minitab, *Minitab Release 15, Statistical Software*. Minitab Inc., 2006.

[11] PSPICE user manual: *ORCAD PCB design suite.* San Jose: Cadence design systems Inc., 2006

[12] International Technology Roadmap for Semiconductors, “Process Integration, Devices and Structures”, 2007 Ed.

[13] Model for assessment of cmos technologies and roadmaps, *MASTAR.* ST Microelectronics*,* Advanced devices research group, 2000-2005.

[14] D. C. Montgomery, *Design and Analysis of Experiments.* New York: Wiley, 1991.

[15] J. Zhou, D. Kinniment, G. Russell, and A. Yakovlev, "A robust synchronizer", *IEEE Computer Society Annual Symposium on Emerging VLSI Technologies and Architectures*, 2006.

[16] C. Dike and E. Burton, "Miller and noise effects in a synchronizing flip-flop", *IEEE Journal of Solid-State Circuits,* vol. 34, pp. 849-855, 1999.

[17] A. M. Abas, A. Bystrov, D. J. Kinniment, O.V. Maevsky et.al., “Time Difference Amplifier”, *IEEE Electronic Letters*, Vol. 38, pp. 1437-1438, Nov. 2002.

[18] J. SparsÃ¸ and S. B. Furber, *Principles of Asynchronous Circuit Design: A Systems Perspective*. Boston: Kluwer Academic Publishers, 2001.

[19] Risch, L. "Pushing CMOS Beyond the Roadmap", Proceedings of ESSCIRC, pp. 63, 2005.

Relevant Publications:

1. H. Ramakrishnan, S. Shedabale, G. Russell, and A. Yakovlev, “Stacked Strained Silicon Transistors for Low-Power High-Performance Circuit Applications”, *IEEE ECTC08*, pp. 1793-1798, Florida, May. 2008.
2. S. Shedabale, H. Ramakrishnan, G. Russell, A. Yakovlev, and S. Chattopadhyay, "Statistical Modelling of the Variation in Advanced Process Technologies using a Multi-level Partitioned Response Surface Approach", *IET Circuits, Devices and Systems Journal*.
3. H. Ramakrishnan, S. Shedabale, G. Russell, and A.Yakovlev, "Analysing the Effect of Process Variation to Reduce Parametric Yield Loss", *IEEE ICICDT 2008*, pp. 171-175, Grenoble, Jun. 2008.
4. H. Ramakrishnan, S. Shedabale, G. Russell and A. Yakovlev, “Statistical Variability Analysis of a Mutual Exclusion Element for Strained Silicon Processes”, *Workshop on Impact of Process Variability on Design and Test DATE’ 08*, Munich, Mar. 2008
5. H. Ramakrishnan, S. Shedabale, J. Zhou, G. Russell, and A. Yakovlev, "Variability analysis of a high performance strained silicon Jamb latch synchronizer", *19th UK Asynchronous Forum 2007*, London, Sept. 2007.
6. H. Ramakrishnan, K. Maharatna, S. Chattopadhyay, and A. Yakovlev, "Impact of strain on the design of low-power high-speed circuits", *IEEE ISCAS 2007*, pp. 1153-1156, New Orleans, May. 2007.
7. H. Ramakrishnan, S. Chattopadhyay, K. Maharatna, and A. Yakovlev, "Exploration of potential of strained silicon CMOS for low-power circuit design”, *Int. Conf. on Ultimate Integration on Silicon (ULIS)*, pp. 29-32, Mar. 2007.
8. B. Halak, S. Shedabale, H. Ramakrishnan, A. Yakovlev and G. Russell, “The Impact of Variability on the Reliability of Long on-chip Interconnect in the

Presence of Crosstalk", *System Level Interconnect Prediction (SLIP 2008)*, Newcastle, pp. 65-72, Apr. 2008

9. H. Ramakrishnan, Strained Silicon Technology for Low-Power High-Speed Circuit Applications, NCL-EECE-MSD-TR-2008-140, PhD Thesis, Microelectronic System Design Group, School of EECE, Newcastle University, July 2008.

Appendix A

EDA Tools for Strained Silicon CMOS Cell Libraries with Variability Models

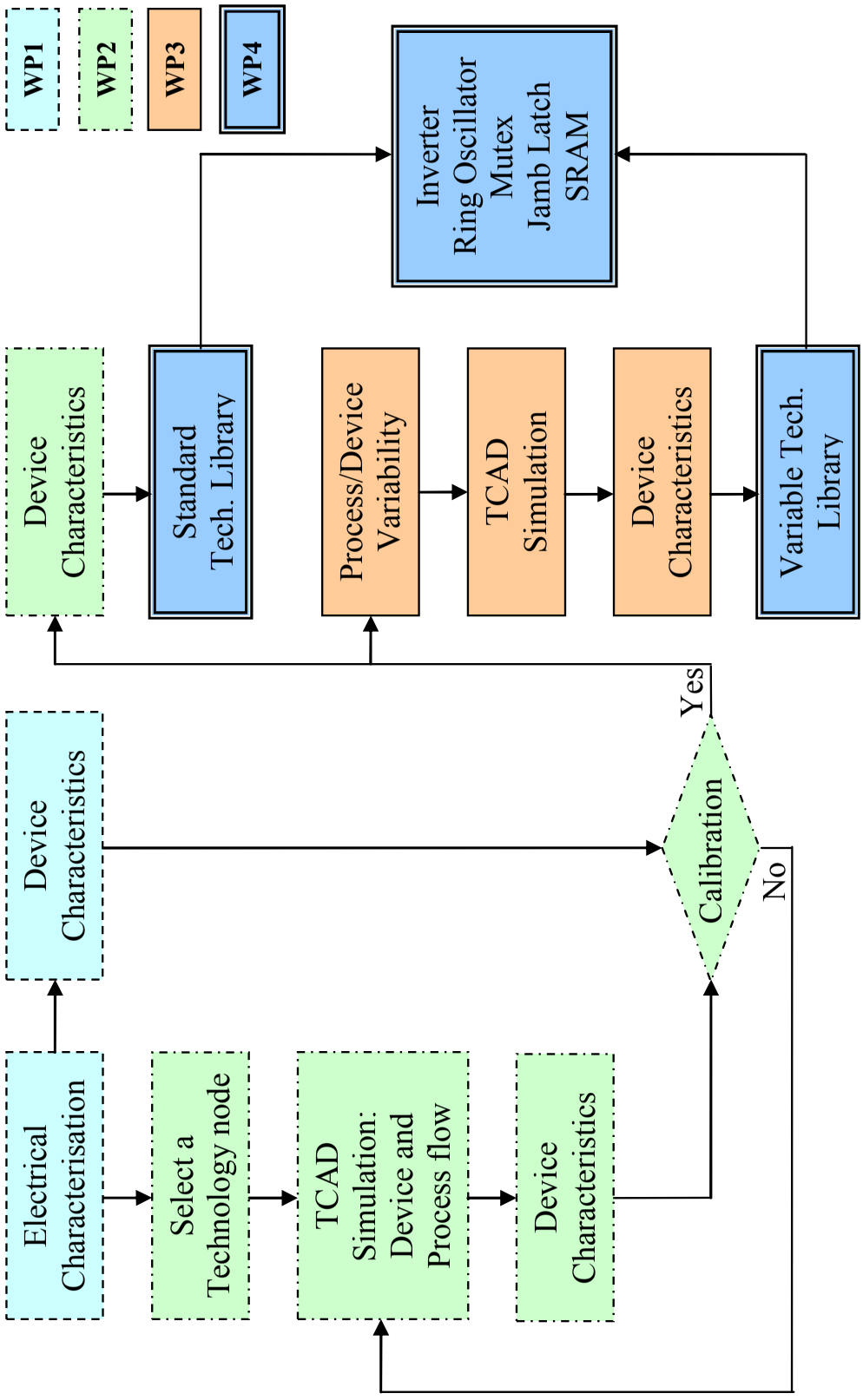


Fig B.1 EPSRC Project

Relevant Publications

WP1: Electrical Characterisation of strained Si and Conventional Si DSM CMOS devices

1. S. Chattopadhyay, L. D. Driscoll, K. S. K. Kwa, S. H. Olsen, et.al., "Strained Si MOSFETs on relaxed SiGe platforms: performance and challenges", *Solid-State Electronics*, vol. 48, pp. 1407-1416, 2004.

WP2: TCAD Simulation: Design of Process and Device flow for n-and p-MOSFET's

1. H. Ramakrishnan, Strained Silicon Technology for Low-Power High-Speed Circuit Applications, NCL-EECE-MSD-TR-2008-140, PhD Thesis, Microelectronic System Design Group, School of EECE, Newcastle University, July 2008 [Chapter 3].

WP3: Incorporation of device and process Variability in TCAD simulator

1. S. Shedabale, H. Ramakrishnan, G. Russell, A. Yakovlev, and S. Chattopadhyay, "Statistical Modelling of the Variation in Advanced Process Technologies using a Multi-level Partitioned Response Surface Approach", *IET Circuits, Devices and Systems Journal,* vol.5, pp.451-464, 2008.

WP4: Design Cell Libraries for strained-Si Technology

1. H. Ramakrishnan, S. Shedabale, G. Russell, and A. Yakovlev, "Stacked Strained Silicon Transistors for Low-Power High-Performance Circuit Applications", *IEEE ECTC08*, pp. 1793-1798, Florida, May. 2008.
2. H. Ramakrishnan, S. Shedabale, G. Russell, and A.Yakovlev, "Analysing the Effect of Process Variation to Reduce Parametric Yield Loss", *IEEE ICICDT 2008*, pp. 171-175, Grenoble, Jun. 2008.
3. H. Ramakrishnan, S. Shedabale, G. Russell and A. Yakovlev, "Statistical Variability Analysis of a Mutual Exclusion Element for Strained Silicon Processes", *Workshop on Impact of Process Variability on Design and Test DATE' 08*, Munich, Mar. 2008

4. H. Ramakrishnan, S. Shedabale, J. Zhou, G. Russell, and A. Yakovlev, "Variability analysis of a high performance strained silicon Jamb latch synchronizer", *19th UK Asynchronous Forum 2007*, London, Sept. 2007.
5. H. Ramakrishnan, K. Maharatna, S. Chattopadhyay, and A. Yakovlev, "Impact of strain on the design of low-power high-speed circuits", *IEEE ISCAS 2007*, pp. 1153-1156, New Orleans, May. 2007.
6. H. Ramakrishnan, S. Chattopadhyay, K. Maharatna, and A. Yakovlev, "Exploration of potential of strained silicon CMOS for low-power circuit design", *Int. Conf. on Ultimate Integration on Silicon (ULIS)*, pp. 29-32, Mar. 2007.
7. B. Halak, S. Shedabale, H. Ramakrishnan, A. Yakovlev and G. Russell, "The Impact of Variability on the Reliability of Long on-chip Interconnect in the Presence of Crosstalk", *System Level Interconnect Prediction (SLIP 2008)*, Newcastle, pp. 65-72, Apr. 2008