

A FUZZY PROGRAMMING BASED TECHNIQUES FOR VARIATION AWARE MULTI-METRIC OPTIMIZATION

¹T.S.KARTHIK, ²V. JAWAHAR SENTHILKUMAR

^{1,2} Department of Electronics and Communication Engineering, College of Engineering-Guindy, Anna University, Chennai-600025

E-mail: ¹ writetotsk@gmail.com, ² veerajawahar@annauniv.edu

ABSTRACT

The uncertainty due to process variations is modeled using interval valued fuzzy numbers and a fuzzy programming based optimization is proposed to improve circuit yield without significant over design. In addition to the statistical optimization methods, we have proposed a novel technique that dynamically detects and creates the slack needed to accommodate the delay due to variations. The variation aware gate sizing technique is formulated as a fuzzy linear program and the uncertainty in delay due to process variations is modeled using fuzzy membership functions. The timing based placement technique; on the other hand, due to its quadratic dependence on wire length is modeled as nonlinear programming problem. The variations in timing based placement are modeled as fuzzy numbers in the fuzzy formulation and as chance constraints in the stochastic formulation. In the context of dynamic variation compensation, a delay detection circuit is used to identify the uncertainty in critical path delay. The delay detection circuit controls the instance of data capture in critical path memory flops to avoid a timing failure in the presence of variations.

Keywords: *Gate Sizing, Process Variation, Timing, Fuzzy Programming and Clocking*

1. INTRODUCTION

Devices that are close together in the layout have a higher probability of being alike in characteristics than devices placed far apart. CMP (chemical-mechanical polishing) effects and optical proximity effects also increase the magnitude of intra-die variation in nanometer technology. Recently, several researchers have attempted to optimize power, delay and noise in the presence of process variations [1, 4, 7, 8, 9, 12,]. The works in [1, 7] mainly focused on circuit optimization schemes with a statistical perspective. In other words, a statistical delay model or SSTA is used to guide timing analysis. The authors in [8,9], presented a statistical optimization approach that takes into account randomness in gate delays by formulating an efficient mathematical program.

The critical path approach when tested on random designs, save power with a small timing penalty. Since, several circuit optimization techniques like, gate sizing, buffer insertion and incremental placement are inherently suited to be modeled as a mathematical program. In the next section, we discuss briefly the basics of mathematical programming and uncertainty aware optimization schemes.

The major part of the work in this paper focuses on statistical optimization of yield, delay and power

as a variation aware mathematical program formulation. This paper is presented as follows. Section 2 presents the Variation aware gate sizing. Section 3 presents the Variation aware timing based placement. Section 4 about the Dynamic Clock Stretching. Section 5 then combines the results discussions and updated the valid models. Finally Section 6 Concludes this paper.

2. VARIATION AWARE GATE SIZING

The variation aware gate sizing works models the process, voltage and temperature (PVT) variation using a statistical approach. The PVT variations in the nanometer era, can be categorized as inter-die and intra-die variations. The inter-die variation occurs across different dies and affects all the transistors in the chip in a similar fashion. The intra-die variations, on the other hand, refer to variability within a single chip resulting in the gate lengths of some transistors larger and some others smaller than the intended sizes. The characteristics of the intra-die variations are correlated with respect to the position of the transistor in the die. The modeling of process variations, initially was limited to statistical static timing analysis (SSTA) [2,4], where continuous distributions are propagated instead of deterministic values to find closed form expressions for performance in presence of

variations. More recently, statistical design optimization for improving power and area for an acceptable yield has been investigated in [6, 7, 8, 9]. In the optimization uses a penalty function to improve the slacks of critical paths to improve yield.

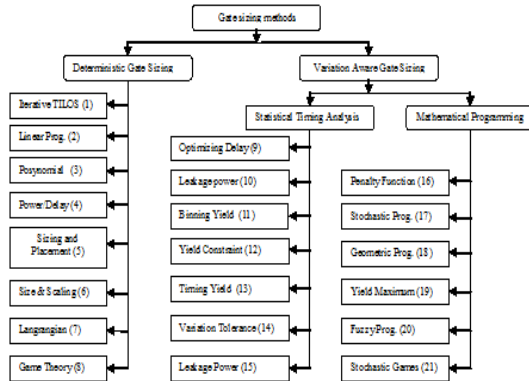


Figure 1. Taxonomy Diagram Of Optimization Methods For Gate Sizing

An SSTA engine is used in the iterative optimization framework [7] to find the most critical gates to size in terms of power/delay sensitivity. A stochastic programming approach with chance (probabilistic) constraints is used in [8] and [9] to incorporate yield in the gate sizing problem formulation. However, the SSTA based approaches [6,7] use continuous distributions, which require a number of operations to be performed iteratively at each node and hence, involve higher runtimes. The stochastic programming based statistical optimization technique, on the other hand, is reasonably fast, but is claimed in [5] that it can produce less optimized solutions compared to fuzzy programming, when tested with Monte-Carlo simulations.

In figure 1 shows the various methods (1-21) gate sizing are available, proposed a new variation aware gate sizing algorithm considering the uncertainty due to process variations using the concept of fuzzy linear programming. In the context of fuzzy set theory, imprecision is defined as an uncertainty where it is difficult to even predict the average behavior of the outcome. Probability theory can be used to model situations in which the average behavior is predictable (situations that obey the law of large numbers) and enough information is available to model the probability distribution functions. The theory of fuzzy sets and systems on the other hand, has been used to model imprecision in different applications such as vision and robotics. In VLSI design automation, fuzzy logic has been applied to model imprecise

coefficients in VLSI testing and scheduling in high level synthesis. To the best of our knowledge, this is the second time the concepts of fuzzy sets and systems and fuzzy mathematical programming is being used to model the uncertainty due to process variations in nanometer VLSI circuits. For simplicity, we use linear delay models [3] and linear membership functions. However, more complex models including nonlinear or other polynomial models can be easily incorporated into the fuzzy optimization flow. The fuzzy optimization is a two step process. Initially, a deterministic optimization is performed assuming the worst and the average case values for the variation parameters to identify the bounds of the uncertain problem.

3. VARIATION AWARE TIMING BASED PLACEMENT

Circuit optimization techniques such as, gate sizing, incremental placement, buffer insertion, is commonly used to improve the performance of integrated circuits. Timing based incremental placement is crucial in nanometer circuits to meet the high performance requirement. The process finds the optimal locations of cells in a critical sub circuit such that the delay of the circuit is minimized. Circuit designers over the years, have used corner case models to optimize and analyze designs. The idea is to meet the timing specification at the best, worst and typical case model values. However, with process variations, the above test results can be far from the actual values. A guarded approach in terms of yield, to eliminate the effects of variability, is to perform deterministic optimization at the worst case values of the varying parameters. The worst case approach guarantees high timing yield, but leads to sub-optimal solutions in terms of performance. Timing yield in this context, is defined as the percentage of chips meeting the timing specification. Typical case value, on the other hand, guarantees optimal solutions but can result in unacceptable timing yield. It is clear that new methodologies are needed, which can guarantee a high timing yield and at the same time provide a solution with a high power/performance ratio.

Several researchers have investigated the effects of variations in timing analysis and statistical design optimization. Static timing analysis was replaced with statistical static timing analysis (SSTA) [2], where continuous distributions are propagated instead of deterministic values to find closed form expressions for performance in the presence of variations. Recently, variation aware gate sizing for improving power and area for an acceptable yield has been investigated in [7, 9]. Thus, the consideration of process variations is important in the design and the optimization of circuits. In this paper, we propose the use of fuzzy mathematical programming (FMP) and stochastic chance constrained programming for variation aware timing based incremental placement problem. The un-

certainty due to process variations are modeled using fuzzy numbers in the FMP case and using probabilistic constraints in the chance constrained programming formulation. Recently, the authors in [5, 11, 12] have considered process variations, while solving the placement problem. The authors in [11], have considered the effects of variations during placement in FPGAs. Variations due to lens aberrations have been considered in [5] and a fuzzy optimization flow for timing variations in [12,13].

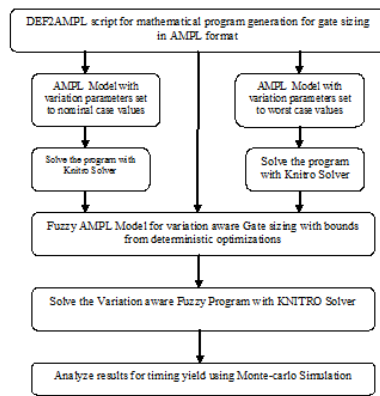


Figure 2. Fuzzy Gate Sizing: Simulation Flow

The problem of timing based incremental placement is an important part of the timing convergence flow. It can be formally defined as the process of finding the optimal locations of cells in a critical sub circuit such that the delay of the circuit is minimized. In timing based placement, the length of interconnects in the critical paths need to be minimized by changing the locations of certain cells. The timing of a circuit is usually measured in terms of the worst negative slack and the total negative slack. Slack in this context, is defined as the difference between the required time and actual arrival time of the signal. Timing driven placement approaches can be categorized into net-based and path based [6, 9] approaches. The net-based approach translates the timing requirements into sensitivity coefficients of timing critical nets and performs a weighted wire length minimization. Hence, modeling the effects of process variations in these net-based approaches is not straightforward. On the other hand, the path based approaches hold an accurate timing view and minimize critical path delay more directly by involving path delay constraints in the optimization problem. A problem with the path based approach is their high computational complexity due to the exponential number of paths. But path based delay constraints can be transformed into node-based constraints [9] to improve the feasibility of optimizing large circuits. The transformation only introduces a sub-optimality of 1-2% [9].

The theory of fuzzy sets and systems over the years, has been applied in VLSI design automation for high level synthesis [30] and for modeling variations in

gate sizing. The uncertainty due to variations can be modeled using fuzzy numbers with linear membership functions. The proposed timing based placement approach is formulated to minimize the worst negative slack of the circuit in the presence of process variations. The fuzzy optimization approach starts with a deterministic optimization assuming the worst and the average case values for the variation parameters. The results of these deterministic optimizations are used to convert the fuzzy optimization problem into a crisp nonlinear problem using the symmetric relaxation method. The crisp problem formulation, in general, has been shown to provide satisfactory solution in the presence of imprecision or variations in coefficients of the constraints or objective function in the optimization problem. We show that the fuzzy optimization approach improves the variation resistance of the circuit without compromising on the achievable performance. The stochastic chance constrained programming (CCP) approach is again a well established technique for performing uncertainty aware optimization. It has previously been applied to model process variations during the gate sizing problem [8, 9]. Here, we also perform variation aware nonlinear timing based placement using stochastic CCP. The stochastic CCP is cast as a robust mathematical program with varying parameters in the constraints of the formulation. The proposed approach uses probabilistic constraints to capture the uncertainty due to process variations. The optimization as a pre-processing step, converts these probabilistic constraints into an equivalent second-order conic program (SOCP) by explicitly using the mean, variance and the inverse-distribution of the varying parameters. Similar to the crisp-fuzzy problem, the translated stochastic-SOCP is solved using an interior point nonlinear optimization solver.

4. DYNAMIC CLOCK STRETCHING

The power-performance trade-off in the nanometer era, has only exacerbated with the inception of parameter variations in nanometer technology. Parameter variations comprise process deviation due to doping concentration, temperature fluctuations, power supply voltage variations and noise due to coupling. Variations can cause frequency and power dissipated to vary from the specified target and hence can result in parametric yield loss. Parametric yield, in this context, is defined as a design's sensitivity to variations, and is expected to cause 60-70% of all yield losses in the impending technology generations. To ensure correct operation under all possible variations (process, voltage and temperature), circuits are often designed with a conservative margin. The margins are added to the voltage and/or device structures to account for the uncertainty due to worst case combination of variations. However, such a worst case



combination is very rare or even impossible in most situations making this design strategy overly conservative.

In this context, several researchers have proposed the use of statistical timing analysis and statistical optimization mechanism to meet timing in the presence of variations without significant over design [2, 7]. The variation aware optimization methodologies use stochastic or fuzzy methodology to minimize the impact of uncertainty due to process variations on performance, power and other design overheads. Statistical timing analysis (SSTA) was investigated in, [2], where continuous distributions are propagated instead of deterministic values to find closed form expressions for performance in presence of variations. Variation aware solutions have also been developed for circuit optimization problems like gate sizing, buffer insertion and incremental placement [7, 12]. The main objective of these works has been to improve yield, without compromising on performance, power and area. The variation aware optimization techniques have shown to improve design overheads without loss in parametric yield. However, the statistical optimization methods still over consume resources irrespective of whether the circuit is affected by variations or not. Hence, to facilitate more aggressive power-performance-yield tradeoff improvement, dynamic schemes to detect and correct the uncertainty due to process variations are becoming necessary. Further, the authors in, proposed a novel design paradigm which achieves robustness with respect to timing failure by using the concept of critical path isolation. The methodology isolates critical paths by making them predictable and rare under parametric variations. The top critical paths, which can fail in single cycle operation, are predicted ahead of time and are avoided by providing two cycle operations.

One of the popular methods to dynamically combat process variation's impact on design has been to use adaptive voltage scaling (AVS). The voltage scaling systems tracks the actual silicon behavior with an on-chip detection circuit and scales voltage in small increments to meet performance without high overheads in the presence of process variations. In [3], the critical path of the system was duplicated to form a ring oscillator and actual performance requirement of the circuit is co-related to the speed of the oscillator and appropriate voltage scaling is performed. However, in the nanometer era, it is not feasible to use a single reference for a critical path and the variations spread, can make the close to critical

delay paths critical on actual implementation. Recently the authors in, proposed an AVS system which can emulate critical paths with different characteristics. With increasing amount of on-chip variations and spatial correlation the methodology can have severe discrepancies. In a bid to reduce such margin and remove the dependency of feedback mechanism on a single path, a novel on-chip timing checker was proposed in to test a set of potential critical paths. The method uses a shadow latch with a delayed clock to capture data in all potential critical paths. An error signal is generated if the value in original and shadow latch is different due to a timing violation caused by process variations. The methodology however aims at correcting (not preventing) errors caused by aggressive dynamic voltage scaling. To guarantee high timing yield and low overheads in the presence of variations, the ultimate solution is to dynamically alter the clock signal frequency.

In this work proposed a technique to control and adjust clock phase dynamically in the presence of variations. The methodology focused on the design of a dynamic delay buffer cell that senses voltage and temperature variations and alters clock phase proportionately. However, it is not generic to all types of variations and does not include spatial correlation between the delay buffer and the gates in the critical path. Plus, the methodology is not input data dependent and hence changes the clock capture trigger in more than required number of instances. In this chapter, we propose a new approach for dynamic clock stretching by dynamically detecting delay due to process variations. Here, instead of modulating the clock duty cycle, we delay the capture clock edge to critical memory cells to accommodate the increased signal propagation delay due to variations. The methodology captures the signal transition halfway in the critical path in a positive level triggered latch. However, if the signal transition on the critical path, which is expected to occur before time $T/2$ (positive level of clock) is delayed due to process variation. The latch in the detection circuit holds an opposite value compared to the signal line and a delay-flag is set. Here, T is the clock cycle time. The delay-flag dynamically stretches the clock at the destination memory flop, to accommodate the extra signal propagation delay due to variations. Thus the clock stretching methodology avoids a mismatch in the data being captured and hence prevents timing error. The detection circuitry needs to be added to the top "n" critical paths and an error signal from any of these paths can stretch the clock in the



appropriate destination memory cell. The clock is stretched (the capture edge trigger is delayed) considering both spatial correlations between closely spaced critical path gates and an average variation range is verified.

Hence, we can safely assume the irrespective of the gate's level in a critical path, if one gate is affected by variations, there is high probability other gates in the critical path are affected in similar fashion. For example, if there is a variation in the second half of the path, due to the property of spatial correlations some gates in the first half of the path will also be affected by variations and vice versa. The magnitude of these variations, will be hard to predict, and can be different based on their location in the chip area. However, the presence or absence of variations can be safely assumed with the property of spatial correlations.

Further, the use of a delayed clock edge trigger for the destination flops, can result in timing inconsistency. The main issues in this context, are the short paths and consecutive critical paths. A short path raises the possibility of data corruption (failures) in a destination memory cell. However, in nanometer designs short paths are usually rare due to the multiple objectives of power, performance and yield. Plus, with a small margin of clock stretching (10-15%), it is easy to perform sizing to eliminate short path failures. Secondly in pipeline circuits if a critical path is followed by another critical path in the following pipeline stage, the CSL methodology can cause timing failures. This is because the delayed clock circuitry reduces the data capture time available in subsequent pipeline stage. Hence, in pipeline circuits with consecutive critical paths, the delay-flag has to be propagated to subsequent stages to create the necessary slack by automatic clock stretching. In the next section, we validate the proposed CSL technique on derived circuits.

5. RESULTS AND DISCUSSIONS

In this section, we present the simulation flow and experimental results of the proposed fuzzy programming based timing placement and compare it with stochastic and worst case process variation approaches. First, the RTL level VHDL netlists are converted to structural level Verilog netlist using the synopsys design compiler tool. The output Verilog file from the design compiler is then placed and routed using the cadence design encounter tool. The design encounter is also used to perform clock synthesis and timing analysis of

the input netlist. The figure 2 shows the placed and routed netlist (DEF File), timing analysis report (TARPT) and the Verilog file is given as an input to a C script (DEF2AMPL), which converts the netlist into an AMPL based mathematical program format for timing based placement optimization. AMPL is a widely used modeling language for large scale mathematical programming problems. The DEF2AMPL script, on different options, generates the worst case deterministic, typical case deterministic, stochastic or the fuzzy version of the timing based placement problem. The *DEF2AMPL* script, as pre-processing step also generates the coefficients A_0 A_1 A_2 B_0 B_1 and B_2 using interpolation. The script selects the maximum allowable displacement for cells depending on the circuit area and the gate's criticality. The list of values for the interpolation is generated from the design encounter tool. The maximum variation in gate and interconnect delay is assumed to be 25% from the mean value due to the varying process parameters, which is in accordance with the results obtained.

The mathematical programming problems are solved using the KNITRO nonlinear optimization solver available through the NEOS server for optimization. The results of the deterministic nominal and worst case optimizations are also fed to DEF2AMPL script for generating the bound's constraint in the fuzzy nonlinear AMPL model. The fuzzy and the stochastic optimization problem find the optimal. To evaluate the proposed methodology, we simulated an example circuit using Cadence NCVerilog simulator. The purpose of this simulation is to elucidate the functionality of the methodology in the presence of variations in circuit elements. The efficiency of the methodology, however is calculated with Monte-Carlo based timing yield simulations. A chain of inverters in between two flip-flops stages is chosen as the example circuit. In this circuit, all interconnects in the path makes a transition. Hence, the net halfway in the path becomes the necessary critical interconnect.

The clock cycle time is selected in reference to the critical path delay at the nominal corner. In other words, the clock cycle time is chosen without adding a margin for uncertainty in delay due to process variations.

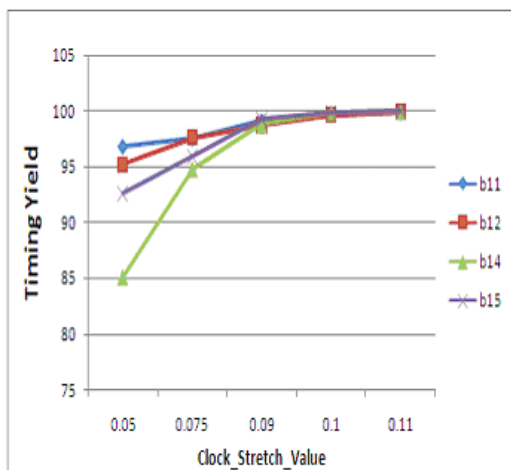


Figure 3. Clock Stretch Range Versus Timing Yield

The improvement in timing yield for the circuits was estimated using Monte-Carlo simulations. The gate and net delay of the circuit elements (b11-b15) as shown in the figure 3 were assumed to have a variation range of around 20% from the nominal value. In the absence of real statistical data, it has been pointed out in, that it's reasonable to assume a variation parameter value of around 20-25% on the delay due to process variations. The RTL level VHDL netlists of the benchmark circuits were converted to a flattened gate level Verilog netlist using the Synopsys design compiler. The output Verilog file from the design compiler is then placed and routed using the cadence encounter tool. A timing analysis report (TARPT) file is then generated to identify the critical paths whose delay value is within 15% of the most critical path. A Monte-Carlo simulation framework is created in a C-program environment with placed and routed (DEF), parasitics file (SPEF), timing analysis report (TARPT) and standard cell delay libraries as input. The Monte-Carlo simulation creates 20000 instance of the circuit with varied delay between nominal and maximum variation range to estimate the timing yield. The clock was stretched to create an extra timing slack of 10% only if delay due to process variations are activated in the worst case critical paths. In the context of timing failures due to short paths, it is crucial to keep the clock stretching range as short as possible. Hence, we performed a simple analysis on selected circuits to see the impact of clock stretch range on timing yield. A smaller value for clock stretch range, for example 5% is shown to impact the timing yield significantly. Hence, with the dual objective of

near perfect timing yield and zero short path failures, we have selected the clock stretch range to be 10% of the clock period. In addition to the timing yield improvement results, we have also specified the benchmark characteristics (number of gates and interconnects), the number of near critical paths and the area overhead due to CSL logic. The proposed CSL methodology also incurs an average area overhead of 5%. The area overhead can be further reduced, if we resort to isolating critical paths similar to the previous works on dynamic clock stretching [10].

6. CONCLUSION

In this chapter, we have proposed a dynamic clock stretching technique to improve the timing yield of circuits in the presence of uncertainty due to process variations. Statistical optimization based techniques due to their over design property, consume extra resources (performance and/or power) even in the absence of variations. The proposed methodology on the other hand, adds timing slack/margin (clock stretching) only in the presence of variations. Further, even in the presence of variations, the proposed methodology activates clock stretching logic only on input patterns that enable the worst critical paths. The dynamic delay detection circuitry, improves yield by controlling the instance of data capture in critical path memory flops. Experimental results based on Monte-Carlo simulation indicate sizeable improvement in average timing yield with a negligible area overhead.

The device and interconnect scaling in CMOS circuits with the objective to follow Moore's curve have brought out numerous issues for design, test and manufacturing engineers. The level of miniaturization and integration of billion transistors on a single chip, gives a clear picture of the nanometer circuit complexity. The increasing integration levels is introducing new issues, which is making multi-metric circuit optimization more complex. The downward scaling of technology is also gradually reaching the limits of ballistic transportation. Hence, it is crucial to develop circuit optimization techniques in the nanometer era, that can achieve high performance, low power dissipation and high reliability. The optimization objectives are highly correlated and conflicting in nature. Further, with increasing levels of variations in process parameters, performance is greatly affected leading to yield loss. It is a challenging task to address all these issues in a



single framework. The focus of this paper is to address these concerns, by proposing new techniques for modeling and optimization of nanometer VLSI circuits considering process variations.

REFERENCES:

- [1] A. Davoodi and A. Srivastava. Variability Driven Gate Sizing for Binning Yield Optimization. In Design Automation Conference, pp. 294–299, 2005.
- [2] A. Devgan and C. Kashyap. Block-based Static Timing Analysis with Uncertainty. In International Conference on Computer Aided Design, pp. 607–614, 2003.
- [3] M Berkelaar and J. Jess. Gate Sizing in MOS Digital Circuits with Linear Programming. In European Design Automation Conference, pp. 217–221, 1990.
- [4] D. Sinha, N.V. Shenoy and H. Zhou. Statistical Timing Yield Optimization by Gate Sizing. In IEEE Transactions on VLSI Systems, pp. 1140–1146, 2006.
- [5] J.J. Buckley. Stochastic versus Possibilistic Programming. In International Journal on FuzzySets and Systems, pp. 173–177, 34(2), 1990.
- [6] K. Chopra, S. Shah, A. Srivastava, D. Blaauw and D. Sylvester. Parametric Yield Maximization using Gate Sizing Based on Efficient Statistical Power and Delay Gradient Computation. In International Conference on Computer Aided Design, pp. 1023–1028, 2005.
- [7] M. Hashimoto and H. Onodera. A Performance Optimization Method by Gate Sizing using Statistical Static Timing Analysis. In International Symposium on Physical Design, pp. 111–116, 2000.
- [8] M. Mani, A. Devgan, and M. Orshansky. An Efficient Algorithm for Statistical Minimization of Total Power under Timing Yield Constraints. In Design Automation Conference, pp. 309–314, 2005.
- [9] M. Mani and M. Orshansky. A New Statistical Optimization Algorithm for Gate Sizing. In International Conference on Computer Design, pp.272–277, 2004.
- [10] S. Ghosh, S. Bhunia and K. Roy. CRISTA: A New Paradigm for Low-Power, Variation-Tolerant, and Adaptive Circuit Synthesis Using Critical Path Isolation. In IEEE Transactions on Computer- Aided Design of Integrated Circuits and Systems, pp. 1947–1956, 26(11), 2007.
- [11] S. Srinivasan and V. Narayanan. Variation Aware Placement for FPGAs. In International Symposium on VLSI, pp. 422–425, 2006.
- [12] V. Mahalingam and N. Ranganathan. Variation Aware Timing Based Placement Using FuzzyProgramming. In International Symposium on Quality Electronic Design, pp. 327–332, 2007.
- [13] V. Mahalingam and N. Ranganathan. Timing Based Placement Considering Uncertainty due to Process Variations. In *IEEE Transactions on VLSI Systems*, pp. 174-179 , 2008.