A Fuzzy Optimization Approach for Variation Aware Power Minimization During Gate Sizing

V. Mahalingam, Student Member, IEEE, Nagarajan Ranganathan, Fellow, IEEE, and Justin E. Harlow, Member, IEEE

Abstract—Technology scaling in the nanometer era has increased the transistor’s susceptibility to process variations. The effects of such variations are having a huge impact on the yield of the integrated circuits and need to be considered early in the design flow. Traditional corner based deterministic methods are no longer effective and circuit optimization methods require reinvention with a statistical perspective. In this paper, we propose a new gate sizing algorithm using fuzzy linear programming in which the uncertainty due to process variations is modeled using fuzzy numbers. The variations in gate delay which is a function of the size and the fan-outs of the gates are represented using triangular fuzzy numbers with linear membership functions. Initially, as a preprocessing step for fuzzy optimization, we perform deterministic optimizations by fixing the fuzzy parameters to the worst and the average case values, the results of which are used to convert the fuzzy optimization problem into a crisp nonlinear problem. The crisp problem with delay and power as constraints is then formulated to maximize the robustness, i.e., the variance of the circuit. The fuzzy optimization approach was tested on ITC’99 benchmark circuits and the results were validated for timing yield using Monte Carlo simulations. The proposed approach is shown to achieve better power reduction than the worst case deterministic optimization as well as the stochastic programming based gate sizing methods, while having comparable runtimes.

Index Terms—Design automation, fuzzy optimization, gate sizing, linear programming, process variations.

I. INTRODUCTION

Variability in nanometer very large scale integrated circuits has continued to increase with the decrease in the feature size of transistors. The environmental effects like changes in power supply voltage and temperature and physical effects like changes in transistor width, channel length, oxide thickness, and interconnect dimensions cause transistors’ power and performance to vary. The physical effects due to the imprecision in the fabrication process control leads to randomness in the number of dopant atoms in transistors [1]. Identically designed circuits can have a large difference in timing and power characteristics leading to loss in the parametric yield of circuits [2]. Hence, capturing the uncertainty due to these variations early during the analysis and optimization phase is crucial. Several circuit optimization methods have been developed over the years, towards improving the area, power, and timing of microprocessors. These optimizations achieve an optimal balance between power consumption and the timing specification by making most paths of the circuits timing critical. However, with the increasing effects of process variations in the nanometer era, such optimization can worsen the timing yield, as any of these critical paths can fail [3]. Here, timing yield is defined as the percentage of chips meeting the timing specification. A guarded approach 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 yield, but leads to high overhead in terms of circuit area and power. On the other hand, the average case values for these parameters correspond to less overheads, but result in unacceptable timing yield. Hence, in the nanometer technology era, new methodologies are needed which can guarantee high yield without any compromise on the area (power) overheads.

Gate sizing has been a simple yet powerful technique to improve the power/delay ratio of VLSI circuits. Several modeling schemes and solution methodologies have been proposed over the years to optimize power and performance through gate sizing. The process of gate sizing can be defined as finding the optimal drive strengths of individual gates of a circuit for a given objective function and constraints. For example, the objective function can be to minimize power or area for a specified timing target. Taxonomy of various gate sizing approaches found in the literature, categorized as: i) deterministic [4]–[10] and ii) variation aware gate sizing [3], [11]–[23] approaches, is shown in Fig. 1. The works listed in the figure have used iterative sizing, linear programming, geometric programming, game theory, and several other interesting formulations to identify the optimal gate sizes for minimizing the power and/or timing of circuits. 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 interdie and intradie variations. The interdie variation occurs across different dies and affects all the transistors in the chip in a similar fashion. The intradie 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 intradie variations are correlated with respect to the position of the transistor in the die.

Initially these process variations modeling was limited to statistical static timing analysis (SSTA) [18], [24], [25], where continuous distributions are propagated instead of deterministic

Manuscript received February 4, 2007; revised September 16, 2007. This work was supported in part by the Semiconductor Research Corporation (SRC) under Grant 1596.001 and in part by the National Science Foundation (NSF) Computing Research Infrastructure under Grant CNS-0551621. A preliminary version of this paper was published as “A novel approach for variation aware power minimization during gate sizing,” in Proc. Int. Symp. Low Power Electron. Design, 2006, pp. 174-179.

The authors are with the Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: mvenkata@csee.usf.edu; ranganat@csee.usf.edu; harlow@csee.usf.edu).

Digital Object Identifier 10.1109/TVLSI.2008.2000597

1063-8210/$25.00 © 2008 IEEE
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 [3], [11], [13], [15]–[17]. In [3], the optimization uses a penalty function to improve the slacks of critical paths to improve yield. An SSTA engine is used in the iterative optimization framework [13] 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 [16] and [17] to incorporate yield in the gate sizing problem formulation. However, the SSTA based approaches [3], [11], [13] use continuous distributions, which require a number of operations to be performed iteratively at each node and hence involve higher runtimes [26], [27]. The stochastic programming based statistical optimization technique, on the other hand, is reasonably fast, but it is claimed in [28] that it can produce less optimized solutions compared with fuzzy programming when tested with Monte Carlo simulations.

In this paper, we propose 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 [29]. In VLSI design automation, fuzzy logic has been applied to model imprecise coefficients in VLSI testing and scheduling in high-level synthesis [27]. To the best of our knowledge, this is the first 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 [30] and linear membership functions [31]. However, more complex models including nonlinear or other posynomial models [32] can be easily incorporated into the fuzzy optimization flow.

Fig. 2 shows an outline of the fuzzy optimization approach. 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. The solution bounds and a variation parameter \( \lambda \) is used to transform the uncertain fuzzy problem into a crisp nonlinear programming problem [29], [31]. The term crisp problem in fuzzy mathematical programming context refers to a nonfuzzy or non-interval-based real value. The variation parameter \( \lambda \) in this crisp problem ranges from (0,1) and implicitly models the interval values (fuzzy variables) of the original problem. This transformation is referred to as symmetric relaxation [31]. The additional parameter \( \lambda \) introduced during the transformation implicitly captures the uncertainty due to process variations and a maximization of
this variable leads to high process variation resistance. Fuzzy numbers with nonlinear membership functions can be modeled by replacing this linear parameter $\lambda$ with a nonlinear function in terms of the variation parameter. The solution of this crisp nonlinear problem represents the optimal value in the presence of variations.

The crisp problem, in general, has been proven to provide the most satisfying solution in the presence of variations in the coefficients of the constraints or objective function in the optimization [32], [33]. In the context of variation aware circuit optimization, the above crisp model with delay and power as constraints, can be used to maximize the robustness, i.e., the variation resistance of the circuit and thus the yield. The proposed approach has been tested on ITC’99 benchmark circuits and the results indicate sizable savings in power compared to the worst case deterministic gate sizing approach. The proposed fuzzy programming based gate sizing also provides better results than the stochastic programming based gate sizing approach, [17], in terms of power savings with a comparable runtime. The results are validated using Monte Carlo simulations, which indicate a high timing yield for the circuits designed with fuzzy gate sizing methodology. The rest of the paper is organized as follows. In Section II, we discuss the motivation as to why fuzzy programming is suitable for variation aware gate sizing. Some preliminaries of fuzzy mathematical programming and the methodology of handling variations in the context of gate sizing are given in Section III. The details of the proposed modeling and the methodology for fuzzy gate sizing approach is given in Section IV, followed by experimental results and conclusions in Sections V and VI, respectively.

II. WHY FUZZY PROGRAMMING FOR VARIATION AWARE GATE SIZING?

In this section, we discuss why fuzzy mathematical programming is well suited for modeling the uncertainty due to process variations in VLSI circuits. The impact of process variations in the nanometer-era are completely nondeterministic and the degree of uncertainty is expected to be worse in future generations [3]. A common approach to handling the uncertainties due to process variations has been to use probabilistic models, in which the uncertain parameters are represented in terms of probability distributions. However, the probabilistic way of evaluating and optimizing the uncertainties is computationally expensive due to the need for complicated multiple integration techniques needed for continuous distributions [26], [27] or due to the large number of scenarios for the corresponding discrete representation. Variation aware gate sizing, proposed in [13], using statistical static timing analysis requires very high execution times.

Furthermore, certain probabilistic modeling requires exhaustive description of uncertain parameters to build probabilistic distributions from historic (empirical) data. When such a description is not available (for example, in a new technology or for a new variation parameter, where extensive details of uncertainty is not known), we do not have enough information for deriving or obtaining the probabilistic distributions. Also, exhaustive Monte Carlo simulations are needed to generate probability distributions for all the varying parameters. An alternative treatment of uncertainty is needed in the situations, where in an expert can predict or obtain only the mean and worst case values of an uncertain parameter. Fuzzy mathematical programming and interval arithmetic can be used to make decisions in the above conditions.

In addition to the above arguments, Buckley has also shown in [28] that fuzzy programming based optimization guarantees solutions that are better or at least as good as their stochastic counterparts. The author provides a comparison of the stochastic and fuzzy programming methodologies using Monte Carlo simulations. The fuzzy optimization, in uncertain environments, finds the best solution (supremum operation over all feasible solutions) as opposed to averaging (integrals over all feasible solutions) in stochastic programming based optimization. Hence, fuzzy programming selects a solution which is better than or at least as good as the stochastic solution. The above arguments provided us with the motivation to investigate fuzzy mathematical programming approach to model uncertainty due to process variations in VLSI design automation. The performance of the proposed algorithm is compared with that of the stochastic programming based gate sizing in order to illustrate the efficiency of fuzzy programming for optimizing in presence of variations. It is shown in Section V that the proposed approach yields better power savings than stochastic gate sizing with comparable execution times under the assumptions of same models, setup, parameters, and objective function for their implementations.
III. FUZZY LINEAR PROGRAMMING METHODOLOGY

In this section, we briefly discuss important concepts in uncertainty aware optimization using fuzzy programming and variation modeling using fuzzy numbers. The reader is referred to [32], [33], and [29]) for a detailed treatment of the concepts of fuzzy mathematical programming. Zadeh in [29] introduced the concept of fuzzy sets and systems in which an element belonging to a set need not be binary valued, but could be any value in between [0, 1]. The membership value of the element in the interval [0, 1] is decided on how much it belongs to it; and the higher the degree of belonging, then, the higher the membership value is.

A. Fuzzy Numbers

Fuzzy set theory and fuzzy optimization techniques provide an efficient mechanism for modeling and optimizing systems that exhibit imprecision. The theory and methodology of fuzzy programming based optimization has been popular since the inception of decision making in fuzzy environments by Bellman and Zadeh, in 1970 [29]. Several models and approaches have been proposed for uncertainty management using fuzzy linear programming, fuzzy multiobjective programming, fuzzy dynamic programming, fuzzy integer programming, possibilistic programming and fuzzy nonlinear programming. An extensive list of references can be found in [32] and [34]. A recent survey on fuzzy linear programming based optimization from an application perspective has been provided by Inuiguchi and Ramik [35]. In order to solve a complex uncertain problem using fuzzy optimization, fuzzy modeling is critical.

A simple information such as the processing time for a task is around 23 minutes or within the range of 21 and 25 can be expressed by means of the following membership function:

\[ \mu_{23}(x) = \begin{cases} 
\frac{(x-21)}{2}, & \text{if } 21 \leq x \leq 23 \\
\frac{(x-25)}{2}, & \text{if } 23 < x \leq 25 \\
0, & \text{otherwise.}
\end{cases} \]

The uncertainty due to process variations can be modeled as normally distributed random variables with mean \( E \) and standard deviation \( \sigma \). In this work, instead of using normal distribution, we model these variations as interval valued fuzzy numbers in the range \( E - 3\sigma \) and \( E + 3\sigma \). The \( 3\sigma \) value is assumed to be the deterministic worst case variation value, meaning all uncertain process parameters are set to \( 3\sigma \), for maximum timing yield in worst case deterministic optimization. Interval valued fuzzy numbers were first explained by Zadeh [29] using possibilistic distributions. In the context of fuzzy mathematical programming, possibilistic distributions \( \pi(\cdot) \) are analogous to linear membership functions [27]. Triangular and trapezoidal memberships are commonly used possibilistic distributions in solving fuzzy linear programming problems.

Fuzzy optimization with nonlinear membership functions have also been attempted for improving modeling accuracy [32], where the variation parameter \( \lambda \) is replaced with a nonlinear function.

The triangular fuzzy number is usually denoted by a triple \( X = (x^m, x^l, x^u) \), where \( x^m \) is the most possible value or the mean value and \( x^l, x^u \) are the lower and upper bounds, denoting the pessimistic and optimistic value of the number. Depending on the context, the value \( x^l \) can be pessimistic or optimistic variation from the mean value \( x^m \) and the same holds for the value \( x^u \). In the context of VLSI circuit optimization, the triple \( L_{\text{eff}} = (L_{\text{eff}}^m, L_{\text{eff}}^l, L_{\text{eff}}^u) \) can be used to model variations in channel length. Since the general objective in circuit optimization is to minimize delay or power, the pessimistic value in this context for effective channel length is \( L_{\text{eff}}^m + 3\sigma \). Similarly, if we model the gate oxide thickness as a fuzzy triple, the pessimistic value is once again the upper bound value.

B. Solution Technique

In this section, we explain the solution methodology of variation aware optimization using fuzzy linear programming. Fuzzy linear programming (FLP) is a special case of fuzzy mathematical programming, where objective function and constraints of the optimization problem are linear. The varying coefficients are assumed to be varying linearly in the specified interval. The FLP problem shown here is a maximization problem with uncertain coefficient in the constraints

\[
\text{maximize } \sum_{i=1}^{n} a_i x_i \\
\text{subject to } \sum_{i=1}^{n} b_{ji} x_i \leq c_j, \quad 1 \leq j \leq m \quad (1)
\]

where \( m \) is the number of constraints, \( n \) the number of variables, \( c_j \) is the limit of the constraint, \( b_{ji} \), and \( a_i \) are constant regression coefficients, \( x_i \) is the variable in the optimization formulation and at least one \( x_j > 0 \). In the above optimization problem, the coefficient \( b_{ji} \) is the interval valued fuzzy number, which has a mean value \( b_{ji} \) and a maximum variation of \( d_{ji} \). The upper bound is assumed to be the pessimistic variation for this fuzzy number. The fuzzy number \( b_{ji} \) is also assumed to vary linearly with the value of the variable \( x_i \). To defuzzify the FLP we need to identify the lower and upper bounds of the optimal solution. Defuzzify or defuzzification refers to the process of converting the fuzzy problem into a crisp or real-valued nonlinear problem. The upper bound value for the fuzzy optimization problem can be estimated by setting the fuzzy coefficients fixed to the average case of the triangular fuzzy number as shown in the following equation. The upper bound value here refers to the value of the objective function in the optimization. The fuzzy interval \( b_{ji} \) is in the left hand side of the constraint and the mean value for \( b_{ji} \) allows a higher value for variable \( x_i \) and hence a larger objective value (upper bound)

\[
\text{Obj}_1 = \max \sum_{i=1}^{n} a_i x_i \\
\text{subject to } \sum_{i=1}^{n} b_{ji} x_i \leq c_j, \quad 1 \leq j \leq m. \quad (2)
\]

Similarly, the lower bound value is found by setting the fuzzy coefficients to the pessimistic \( b_{ji} + d_{ji} \) value, which constrains
the maximization problem more resulting in a smaller objective value (lower bound)

\[
\text{Obj}_2 = \max \sum_{i=1}^{n} a_i x_i
\]

subject to \( \sum_{i=1}^{n} (b_{ji} + d_{ji}) x_i \leq c_j, \quad 1 \leq j \leq m \). (3)

Now, with these bound objective values and a new variation parameter, we can formulate a crisp problem which will represent an optimal solution in presence of variations. The objective function for the fuzzy programming problem takes values between this lower \( \text{Obj}_l = \min(\text{Obj}_1, \text{Obj}_2) \) and upper \( \text{Obj}_u = \max(\text{Obj}_1, \text{Obj}_2) \) bound values. Using these bound values and the symmetric definition of fuzzy decision proposed by Bellman and Zadeh [29], the fuzzy problem can be defuzzified into a crisp nonlinear problem as shown in the following equation:

\[
\lambda (\text{Obj}_l - \text{Obj}_u) - \sum_{i=1}^{n} a_i x_i + \text{Obj}_u \leq 0,
\]

\[
\sum_{i=1}^{n} (b_{ji} + \lambda d_{ji}) x_i - c_j \leq 0, \quad 1 \leq j \leq m
\]

\[
x_j \geq 0, \quad 0 \leq \lambda \leq 1, \quad 1 \leq i \leq n.
\]

where \( \lambda \) is the variation parameter introduced in the crisp problem which is to be maximized for a fuzzy optimal solution between the lower and upper bound values. The solution of this nonlinear programming problem can be interpreted as representing an overall degree of satisfaction in presence of varying parameters [32]. In the general formulation the variation parameter \( \lambda \) can take values between 0 and 1. However, for our gate sizing problem, we can restrict it to a much smaller range. The smaller range does not affect the quality of the solution, but improves the runtime of the optimization process. In the next section, we explain the modeling and formulation of the variation aware gate sizing problem.

IV. PROPOSED VARIATION AWARE FUZZY GATE SIZING

In this section, we describe our formulation of the gate sizing problem in the presence of uncertainty due to process variations. The problem of gate sizing can be defined as finding the optimal drive strengths such that the specified critical path timing is met and the overhead (power in this work) is minimized. We determine the size of the gates with the goal of minimizing dynamic power with delay as constraints. We use linear programming due to simplicity of modeling, faster execution time, and availability of well-developed fuzzy linear programming techniques for variation aware optimization [33]. However, it should be noted that fuzzy programming based optimization can also solve optimization problems in a nonlinear programming setup [32]. Next, we present the power and delay models used in this work.

A. Power and Timing Models

The dynamic power consumption of a gate \( (i) \) is given as

\[
P_2 = \frac{1}{2} f V_{dd}^2 E_i (C_i + C_{\text{wire}}) + P_{sc}
\]

where \( P_2 \) is the total dynamic power consumed by gate \( i \), \( f \) is the clock frequency, \( V_{dd} \) is the supply voltage for the gate, \( E_i \) is the average switching activity of the gate, \( C_i \) is the intrinsic gate capacitance internal to the gate and \( C_{\text{wire}} \) is the sum of all the interconnects that fan out from gate \( i \). Thus, reducing the size \( (s_i) \) of the gate reduces the intrinsic gate capacitance of gate \( i \), power consumption, and fan-in load capacitance of the gate. Secondly, in this work we model gate delay as a linear function of gate size. The linear delay model proposed in [30] is given by

\[
d_i = a_i - b_i s_i + c_i \sum_{j \in \text{fan}(i)} s_j
\]

where \( s_i \) refers to the size of gate \( i \), \( \text{fan}(i) \) is the set of gates that fan out from gate \( i \), constant coefficients \( a_i, b_i, c_i \) are empirically determined by extensive SPICE simulations for each gate in the library for various sizes and fanout counts. A similar nominal delay model has also been used in a recent stochastic programming based statistical gate sizing approach [17].

B. Variation Modeling With Spatial Correlation

The increasing influence of process variations on circuit yield is making variation aware optimization a requirement early in the design flow. The uncertainty due to process variations has been modeled in most works using the following equation:

\[
D = d_i + \sum_{j=1}^{n} d_j X_{j} + d_r X_r
\]

where \( d_i \) is the nominal delay and \( X_{j} \) and \( X_r \) are the random parameters representing correlated and independent variations respectively. The magnitude of these variations is given by the variables \( d_j \) and \( d_r \), which is determined from extensive simulations. The correlations of different areas of transistors are different and are usually high for gates close to each other in the die. The correlations in this work were included in the delay model as proposed by the authors in [36]. The lowest level (Level 0) of the die is divided into 16 regions and they are grouped into subblocks on the upper levels. The correlation of connected gates in the die is directly proportional to the number of common regions the gates share in levels 0 and 1. The gates placed closer to each other have a similar variation characteristic and hence high correlation. The magnitude of variation in a gate is determined by the count of its fan-out gates, placed in the same Level 0 region. A gate with all its fan-out gates in the same region is modeled to have a smaller variation value.

We capture these variations using the concept of fuzzy numbers. We assume the delay of the gate as an interval with lower and upper bound values. In other words, each gate’s delay is now a triangular value (average, low, high), instead of a single discrete value. We focus on the intradie variations, meaning that each transistor can have a different amount of variation. In the pioneering work on process variations by Sani Nassif [37], it
has been pointed out that in the absence of real statistical data on a process run it’s reasonable to assume a variation parameter value of 25% on the delay due to process variations. The uncertainty in the delay values are transferred to the coefficients $b_i$ and $c_i$ of the linear delay model shown in (6). Following the above assumption, in the work reported in [16] and [17], it has been pointed out that the regression coefficients $b_i$ and $c_i$ closely approximate the variation effects of $L_{eff}$ and $t_{oX}$ based on simulation experiments with statistical data. They observed that the regression coefficients $b_i$ and $c_i$ to have variation values of 8% and 10%, which in turn corresponded to the 25% variation effect in circuit delay. These coefficients are the fuzzy numbers of the triangular form with a linearly varying membership function. Since these observations were an outcome of significant experiments based on statistical data from real process runs, we followed the same assumptions in our work. This allows us to make a fair comparison with the work reported in [17]. Next, we explain the proposed fuzzy gate sizing approach for optimization in presence of process variations.

C. Variation Aware Fuzzy Gate Sizing

In this paper, we use delay constrained dynamic power minimization for the gate sizing problem. If minimizing power is our sole interest, then all the gates can be set to minimum size. However, the problem objective is to achieve minimum power for a specified timing target. Hence, the cost function of the deterministic optimization formulation must include both delay and power. The deterministic formulation of the sizing problem is given by

$$\min \sum_i P_i \quad \text{s.t.} \quad D_p \leq T_{spec} \forall p \in P$$

and

$$D_p = \sum_{i \in P} \left( a_i - b_i s_i + c_i \sum_{j \in f(i)} s_j \right) \quad (8)$$

where $T_{spec}$ is the specified timing target of the circuit, $P$ denotes a particular path in a circuit which belongs to the set of all paths $P$ and $D_p$ is the sum of the delays of all gates in path $p$. The summation of the dynamic power of all the gates is used as the objective function. The dynamic power model in (5) is substituted here. The fuzzy version of the above deterministic optimization problem with uncertain parameters is given by

$$\min \sum_i P_i \quad \text{s.t.} \quad D_p \leq T_{spec} \forall p \in P$$

and

$$D_p = \sum_{i \in P} \left( a_i - (b_i - g_i)s_i + (c_i + h_i) \sum_{j \in f(i)} s_j \right) \quad (9)$$

where $s_i$ is bounded by minimum and maximum gate size, the coefficients $b_i$ and $c_i$ are the uncertain parameters. The uncertain parameters are modeled as fuzzy number triplets of the form $(b_i, \tilde{b}_i - g_i, \tilde{b}_i + g_i)$ and $(c_i, \tilde{c}_i - h_i, \tilde{c}_i + h_i)$, where $g_i$ and $h_i$ are the maximum variations for the coefficients $b_i$ and $c_i$, respectively. The respective $b_i$ and $c_i$ closely approximate the variation in effective channel length ($L_{eff}$) and oxide thickness ($t_{oX}$). The authors in [17] also follow a similar modeling for gate sizing in the presence of uncertainty using chance constrained programming. The fuzzy gate sizing problem is then transformed into a crisp nonlinear problem using the following steps. A deterministic optimization is performed initially with the varying coefficients set to worst and average case values of the fuzzy number. In the worst case optimization, the fuzzy gate delay equation in the fuzzy problem is replaced with the following equation:

$$d_i = \left( a_i - (b_i - g_i)s_i + (c_i + h_i) \sum_{j \in f(i)} s_j \right) \quad (10)$$

The gate delay in the above equation is the most pessimistic estimate, resulting in the worst possible delay for the gate. It can also be seen that the worst case estimate corresponds to the lower bound in coefficient $b_i$, since $b_i$ is inversely proportional to the effective channel length and upper bound in coefficient $c_i$ as it is directly proportional to gate oxide thickness. Similarly, the typical or nominal case of the gate delay is the case where the fuzzy numbers are fixed to their average case values. In the nominal case optimization, the fuzzy delay equations in the fuzzy problem is replaced with the following equation:

$$d_i = \left( a_i - (b_i)s_i + (c_i) \sum_{j \in f(i)} s_j \right) \quad (11)$$

The deterministic optimization problem (8) is solved with the delay equations set to the worst case and nominal case equations. The KNITRO optimization solver available through the NEOS optimization server is used to solve the linear programming problems. The results of these optimization correspond to worst case gate sizing ($w_{sizing}$) and nominal case gate sizing ($n_{sizing}$) values. The above values and a new variation parameter $\lambda$ are used to transform the fuzzy optimization problem into a crisp nonlinear programming problem using the symmetric relaxation method [29]. The crisp nonlinear problem for gate sizing in the presence of process variations is given by

$$\max \lambda \quad \text{s.t.} \quad \lambda(n_{sizing} - w_{sizing}) - \sum_i P_i + w_{sizing} \leq 0,$$

and

$$D_p \leq T_{spec} \forall p \in P$$

and

$$D_p = \sum_{i \in P} \left( a_i - (b_i - g_i \ast \lambda) s_i + (c_i + h_i \ast \lambda) \sum_{j \in f(i)} s_j \right) \quad (12)$$

where the parameter $\lambda$ is bounded by 0 and 1. Even though the parameter $\lambda$ can take any values between 0 and 1, for the gate sizing problem, it can be easily bounded to a smaller value. In this paper, we bound the $\lambda$ value to be between 0.5 and 0.75. We estimated that such a smaller bound is sufficient due to the dual requirement of high yield and low overhead for the gate sizing optimization in presence of variations. The smaller bound speeds up the fuzzy gate sizing procedure by 2–3 times, without affecting the final solution. The crisp optimization problem has three variables, power ($P_i$), delay ($D_p$) and process variations.
(λ) in the above formulation. The parameter λ is the variation resistance (robustness) property of the circuit, meaning the ability to meet the timing constraint even in the presence of variations. The problem tries to maximize variation resistance, constraints delay value even with variations to be less than specified timing, and bounds the power value to be in between \( w_{\text{siz}} \) and \( n_{\text{siz}} \) values. One can favor the power value to be close to the \( n_{\text{siz}} \) value by maximizing the variation resistance value. Hence, the crisp optimization problem tries to satisfy all the three requirements to the maximum degree. It has been shown for problems in other application domains that the above formulation provides the most satisfying optimization solution in the presence of uncertainty [32], [33].

Another issue in the above optimization formulation is that the number of paths in the circuit grows exponentially in the number of gates. Hence, the path based formulation is converted to a node based optimization problem [16], [17]. The node based formulation is a widely used technique [3], [6], [7], [11], [16], [17] to improve the computational efficiency of optimizing large circuits. The gate sizing problem with the node based formulation can be shown as

\[
\begin{align*}
\min & \sum_i P_i \\
\text{s.t.} & \quad a_j \leq T_{\text{spec}} \quad \forall j \in \text{input}(PO) \\
& \quad a_j + D_i \leq a_i \quad \forall i \text{ and } j \in \text{input}(i) \\
& \quad D_i = \left( a_i - b_i s_i + c_i \sum_{j \in \text{fan}(i)} s_j \right),
\end{align*}
\]

Here \( a_j \) is the arrival time at node \( j \) and \( T_{\text{spec}} \) is the timing specification. In the node based approach, the path based constraints are broken by using arrival time variables at each node. The number of such constraints is linear and is proportional to the number of interconnects in the circuit. The node based formulation introduces some suboptimality (decrease in power optimization in this work). However, the decrease is negligible for circuits with less than 20 levels of logic [17]. Since the trend in the nanometer era is towards higher clock speeds and lesser levels of logic, we believe that the overall impact is very less and the computational benefit in terms of running time of the optimization justifies such a modification. In the next section, we present the simulation steps and the experimental results of the fuzzy gate sizing approach tested on ITC’99 benchmark circuits.

V. EXPERIMENTAL RESULTS

The proposed fuzzy linear programming optimization for gate sizing was tested on ITC’99 benchmark circuits. The complete simulation flow is shown in Fig. 3. First, the RTL level VHDL netlists are converted to structural level verilog netlist using the synopsys design compiler tool. The gate level netlist is completely flattened to the basic gates in the standard cell library. The output verilog file from the design compiler is then placed and routed using the cadence design encounter tool. The benchmark circuits are synthesized using the TSMC 90 nm db library of cell delay information and the verilog file are given as an input to a C script (def2ampl), which converts the netlist into an AMPL based mathematical program format for power minimization using fuzzy gate sizing. AMPL is a widely used modeling language for large scale mathematical programming problems. The average switching activity in each line was calculated by simulating each of the benchmark circuits with 100,000 random vectors. The equation coefficients for power and delay models for the standard cell library cells in the TSMC 90 nm libraries are characterized for various gate sizes and fan-outs using hspice simulations. The fit is justified as the rms error is less than 7% for a restricted range of gate sizes (1 x to 4 x). The Def2ampl script uses these delay equations to generate the linear programming models for the benchmarks with delay coefficients set to mean and the maximum possible variation (worst case). The maximum variation in gate delay is assumed to be 25% from its mean value [37]. This is translated into appropriate values for the coefficients \( b_i \) and \( c_i \) in the
delay model. The linear optimization problems are solved using the KNITRO Solver available through the NEOS server for optimization. The def2ampl script uses the results of these optimizations and generates a fuzzy nonlinear AMPL model. The fuzzy nonlinear optimization problem is also solved using the KNITRO solver to find the optimal gate sizes in presence of variations in gate delay. The proposed fuzzy sizing approach is compared with the stochastic programming based gate sizing under uncertainty [17]. The latter method was also implemented with the same setup, parameters, and objective functions for fairness in comparison.

The power reduction achieved by the fuzzy sizing approach compared to worst case deterministic sizing and the stochastic programming approach is documented in Table I. The worst case sizing results correspond to the delay coefficients set to their maximum variation case. $T_{spec}$ corresponds to the minimum required delay for the circuit, obtained by unconstrained delay optimization. Optimized power values for worst case gate sizing, stochastic programming based gate sizing and fuzzy programming based gate sizing is shown in columns 4, 5, and 6 and the percentage reduction in power of fuzzy approach compared to worst case and stochastic approach is given in columns 7 and 8 respectively. The percentage savings compared to deterministic worst case sizing in power is calculated using the following equation:

$$PR_1 = \frac{\text{Power}_{\text{DWC-GS}} - \text{Power}_{\text{F-GS}}}{\text{Power}_{\text{DWC-GS}}} \times 100 \%.$$  \hfill (14)

Similarly, the percentage improvement of fuzzy sizing compared to stochastic sizing is calculated as

$$PR_2 = \frac{\text{Power}_{\text{S-GS}} - \text{Power}_{\text{F-GS}}}{\text{Power}_{\text{S-GS}}} \times 100 \%.$$  \hfill (15)

It can be seen that there is a sizable savings in power by using the fuzzy sizing approach as compared to deterministic worst case gate sizing and stochastic programming approach. The execution time of the fuzzy optimization approach is also shown in Table I. We also studied the impact of our algorithm on leakage power using the leakage models proposed in [14]. The fuzzy approach and the stochastic approach on the average improve leakage power by close to 17% and 8% respectively compared with the deterministic worst case (DWC) approach during variation aware gate sizing. It should be noted that the above result only indicates the impact of our proposed algorithm formulation on leakage power. However, a multiobjective optimization framework will need to be formulated to efficiently optimize both dynamic and leakage power.

The runtime of the fuzzy logic based optimization is comparable to the stochastic programming approach as seen from Table I. Fig. 4 also illustrates that the runtime complexity of fuzzy linear programming for gate sizing is close to linear in the number of gates in the circuit. Secondly, the effects of spatial correlation are considered as mentioned in Section IV. The variation magnitudes $g_k$ and $h_k$ of gate delay coefficients $b_k$ and $c_k$ are discretized and the contribution of each fanout gate is weighed inversely with respect to the number of sharing regions between the gates. Fig. 5 shows the percentage savings of the correlation aware variation modeling when compared to base fuzzy gate sizing approach. It can be clearly seen that the spatial correlation model eliminates further pessimism in the variation modeling and achieves a average savings of 3.4% power reduction compared to the base F-GS approach. Finally, to verify the result of the fuzzy sizing approach, we generated 10000 samples of the ITC’99 benchmarks. The circuits were fixed with gate size outputs from the fuzzy sizing method and the gate coefficient values $b_k$ and $c_k$ were assumed to have random variation value in the range $b_k$ to $b_k - g_k$ and $c_k$ to $c_k + h_k$ respectively.

<table>
<thead>
<tr>
<th>ITC ’99 Circuit</th>
<th>Number of gates</th>
<th>$T_{spec}$ (ns)</th>
<th>Gate Sizing Power (µW)</th>
<th>CPU time (sec)</th>
<th>% Reduction of F-GS over S-GS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DWC-GS</td>
<td>S-GS</td>
<td>F-GS</td>
</tr>
<tr>
<td>b11</td>
<td>385</td>
<td>0.25</td>
<td>288</td>
<td>254</td>
<td>232</td>
</tr>
<tr>
<td>b12</td>
<td>834</td>
<td>0.39</td>
<td>465</td>
<td>397</td>
<td>357</td>
</tr>
<tr>
<td>b14</td>
<td>4232</td>
<td>2.62</td>
<td>1826</td>
<td>1695</td>
<td>1524</td>
</tr>
<tr>
<td>b15</td>
<td>4585</td>
<td>2.98</td>
<td>1774</td>
<td>1521</td>
<td>1397</td>
</tr>
<tr>
<td>b20</td>
<td>8900</td>
<td>2.68</td>
<td>3797</td>
<td>3423</td>
<td>3120</td>
</tr>
<tr>
<td>b17</td>
<td>21191</td>
<td>3.48</td>
<td>8423</td>
<td>7812</td>
<td>7044</td>
</tr>
<tr>
<td>b18</td>
<td>43151</td>
<td>4.11</td>
<td>14176</td>
<td>13045</td>
<td>11753</td>
</tr>
</tbody>
</table>

Average Savings Percent = 18.57% 9.2%
The variation value was generated from a uniform distribution between these ranges. We then performed Monte Carlo simulation with these random samples to determine the frequency of timing violations. The fuzzy logic approach had an timing yield between these ranges. We then performed Monte Carlo simulation.

The variation value was generated from a uniform distribution.

In this paper, we proposed a new approach for gate sizing considering process variations using fuzzy linear programming. The variations in channel length and oxide thickness are modeled as fuzzy numbers with linear membership functions. The proposed fuzzy gate sizing approach maximizes variation resistance (robustness) of the circuit, with delay and power as constraints in the formulation. Experimental results on ITC ’99 benchmark circuits indicate sizable savings in power and a run-time comparable with that of the stochastic sizing approach. The results validated using Monte Carlo simulations confirm the high variation resistance of the circuits sized using the fuzzy programming approach.

VI. CONCLUSION

In this paper, we proposed a new approach for gate sizing considering process variations using fuzzy linear programming. The variations in channel length and oxide thickness are modeled as fuzzy numbers with linear membership functions. The proposed fuzzy gate sizing approach maximizes variation resistance (robustness) of the circuit, with delay and power as constraints in the formulation. Experimental results on ITC ’99 benchmark circuits indicate sizable savings in power and a run-time comparable with that of the stochastic sizing approach. The results validated using Monte Carlo simulations confirm the high variation resistance of the circuits sized using the fuzzy programming approach.

REFERENCES


V. Mahalingam (S’03) received the B.E. degree in computer science from Sri Venkateswara College of Engineering (SVCE), University of Madras, India, in 2003 and the M.S. degree in computer engineering from the University of South Florida, Tampa, in 2005, where he is currently working toward the Ph.D. degree in computer engineering. In 2002 and 2003, he worked as a part-time Research Assistant at Waran Research Foundation (WARF), India. He has coauthored over 10 papers in refereed journals and conferences and has been a reviewer for the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, ACM Transactions on Design Automation of Electronic Systems, and the Design Automation Conference. His research interests include design automation, circuit design, VLSI testing, and DFM.

Mr. Mahalingam is a student member of the IEEE Computer Society and has received the IEEE Richard E. Merwin scholarship award in 2006 for excellence in USF-IEEE-CS leadership.
Nagarajan (Ranga) Ranganathan (S’81–M’88–SM’92–F’02) received the B.E. degree (Honors) in electrical and electronics engineering from Regional Engineering College, Tiruchirapalli, University of Madras, India, in 1983, and the Ph.D. degree in computer science from the University of Central Florida, Orlando in 1988.

During 1998–1999, he was a Professor of Electrical and Computer Engineering at the University of Texas at El Paso. He is a Distinguished University Professor of Computer Science and Engineering at the University of South Florida, Tampa. He has developed many special purpose VLSI circuits and systems for computer vision, image and video processing, pattern recognition, data compression, and signal processing applications. He has coauthored over 220 papers in refereed journals and conferences, four book chapters, and has received six U.S. patents with one pending. He has edited three books titled VLSI Algorithms and Architectures: Fundamentals, VLSI Algorithms and Architectures: Advanced Concepts, (IEEE CS Press, 1993), and VLSI for Pattern Recognition and Artificial Intelligence, (World Scientific Publishers, 1995). His research interests include VLSI circuit and system design, VLSI design automation, multimeteric optimization in hardware and software systems, biomedical information processing, and parallel computing.

Dr. Ranganathan is a member of the IEEE Computer Society, IEEE Circuits and Systems Society, and the VLSI Society of India. He has served on the editorial boards for the journals: Pattern Recognition (1993–1997), VLSI Design (1994–present), the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS (1995–1997), the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS (1997–1999), and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (1997–2000). He was the chair of the IEEE Computer Society Technical Committee on VLSI during 1997–2001. He served on the steering committee of the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS during 1999–2001, the steering committee chair during 2002–2003, and the Editor-in-Chief for two consecutive terms during 2003–2007. He served as the program cochair for ICVLSID’94, ISVLSI’96, and ISVLSI’05 and as general cochair for ICVLSID’95, IWLSI’98, ICVLSID’98, IWLSI’98, and ISVLSI’05. He has served on technical program committees of international conferences including ICCD, ICCP, IPPS, SPDP, ICHPC, GLSVLSI, ASYNC, ISQED, ISLPED, CAMP, ISCAS, MSE, and ICCAD. He received the USF Outstanding Research Achievement Award in 2002, the USF President’s Faculty Excellence Award in 2003, USF Theodore-Venette Askounes Ashford Distinguished Scholar Award in 2003, and the SIGMA XI Scientific Honor Society Tampa Bay Chapter Outstanding Faculty Researcher Award in 2004. He was a corecipient of three Best Paper Awards at the International Conf. on VLSI Design in 1995, 2004, and 2006.

Justin E. Harlow III (S’62–M’64–SM’92–LSM’07) received the B.S. degree in electrical engineering from the University of Florida, Gainesville, in 1964, the MBA degree from Florida Institute of Technology, Melbourne, in 1983, and the M.S. degree in computer engineering from Duke University, Durham, NC, in 1994.

He served in management and technical positions at National Semiconductor, Fairchild Semiconductor, Harris Semiconductor, ITT Semiconductors, and Honeywell Corporate Research Center. From 1990 through 2004, he was Director of Integrated Circuits and Systems research at the Semiconductor Research Corporation (SRC) in Durham, NC. He is currently an Instructor and Research Associate at the University of South Florida, Tampa. His research interests include computer architectures for emerging nanotechnologies, design automation, VLSI design, testability, and EDA tool performance metrics. He has published a number of papers in these areas, and holds 2 U.S. patents.

Mr. Harlow organized and served on the Architectures study committee for the for Silicon Nanoelectronics and Beyond initiative of the National Science Foundation and the Semiconductor Research Corporation. He has served on numerous National Science Foundation and Defense Advanced Research Projects Agency (DARPA) panels, and has been a consultant to DARPA on future trends in design automation.