

Multi-Corner, Energy-Delay Optimized, NBTI-Aware Flip-Flop Design

Hamed Abrishami, Safar Hatami, Massoud Pedram
Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, CA 90089
{habrisha, shatami, pedram}@usc.edu

Abstract

With the CMOS transistors being scaled to sub 45nm and lower, Negative Bias Temperature Instability (NBTI) has become a major concern due to its impact on PMOS transistor aging process and the corresponding reduction in the long-term reliability of CMOS circuits. This paper investigates the effect of NBTI phenomenon on the setup and hold times of flip-flops. First, it is shown that NBTI tightens the setup and hold timing constraints imposed on the flip-flops in the design. Second, an efficient algorithm is introduced for characterizing the codependent setup and hold time (CSHT) contours. Third, we introduce a multi-corner optimization problem to minimize the energy-delay product of the flip-flops. The optimization relies on mathematical programming to find the best transistor sizes. Finally, we apply our proposed optimization formulation on True Single-Phase Clock (TSPC) flip-flops and show the simulation results.¹

Keywords

Static timing analysis, setup and hold times, NBTI, circuit reliability, device aging, multi-corner optimization, mathematical programming, polynomial modeling.

1. Introduction

As CMOS transistors are scaled toward ultra deep submicron technologies, circuit reliability cannot be ignored. Device aging processes such as the Negative Bias Temperature Instability (NBTI) can have a huge impact on the circuit performance over time. Indeed the NBTI effect has proven to be a rising threat to the circuit reliability in nanometer scale technology. Due to NBTI effect, the threshold voltage of the PMOS transistors increases over time, resulting in reduced switching speeds for logic gates, and the corresponding degradation in circuit performance and increased probability of circuit failure due to timing constraint violations [1][2].

The effect of NBTI on digital CMOS circuit performance has been methodically studied in [1][3]. Recently, some techniques have been proposed to alleviate the degradation of the CMOS circuit performance with time. In [4], for example, it was shown that the speed degradation of the CMOS circuit can be offset by cell-level up-sizing during the initial design to compensate for the NBTI-induced decrease in speed of the PMOS device. The authors of [5]

proposed the use of soft-edge flip-flops which would in turn allow compensating for the delay increase in the combinational logic by introducing a transparency window for the signal launching and receiving flip-flops.

Although these works address the NBTI effect on circuit performance, they did not consider the effect of NBTI on the setup/hold time characteristics of the sequential circuit elements (i.e., latches and flip-flops). More recently researchers have begun to investigate the effect of NBTI on the timing characteristics of flip-flops. In [6] it was claimed that in the presence of NBTI, the setup and hold time of the flip-flops remain nearly constant. In [7], the effect of NBTI on different low power and high performance flip-flops was studied; however, no solution was offered to alleviate the problem. The authors of [8] introduced an ad-hoc selective transistor-level sizing to combat the NBTI effect without considering energy consumption as part of the objective.

In this paper, we show that setup and hold times of flip-flops change due to NBTI and the codependency between them tightens timing constraints over time. Moreover, we consider the energy consumption in our proposed technique which has not been properly addressed in NBTI-related prior works.

As the clock period decreases in modern ICs, inaccuracy in setup/hold times caused by corner-based static timing analysis (STA) tools becomes less acceptable. Therefore, accurate characterization of the setup and hold times of latches and registers is critically important for timing analysis of digital circuits. Setup and hold times are codependent [9] in the sense that there are multiple pairs of setup and hold times that result same clock-to-q. All pairs of setup/hold times that correspond to a constant clock-to-q delay are placed on a contour of clock-to-q delay surface. Hence, study of NBTI effect on CSHT contour becomes a vital task which is done in this paper.

In this paper, we first show how the NBTI effect alters the setup/hold time codependency characterization. Next we present an algorithm to extract the CSHT contour. Finally we introduce a multi-criteria optimization problem to size transistors of a flip-flop to minimize its energy consumption and delay product while satisfying the constraints on its timing characteristics due to NBTI effect.

The remainder of the paper is organized as follows. Section 2 provides some background on NBTI effect and flip-flop characterization. It also defines the terminology which will be used in subsequent sections. The algorithm to extract the CSHT contour is proposed in section 3. The

¹ This research was sponsored in part by a grant from the National Science Foundation under award number CCF-0811876.

effect of NBTI on codependent Setup/Hold Time (CSHT) characterization is described in Section 4. The problem formulation and mathematical program are introduced in section 5. Section 6 gives the simulation results and Section 7 concludes the paper.

2. Background

This section provides the terminology, reviews the manifestation of NBTI on threshold voltage of a PMOS transistor, the CSHT characteristic contour for a given clock-to-q delay, and explains how to utilize this contour in a STA tool for timing verification.

2.1 NBTI effect

The recent aggressive scaling of CMOS technology makes NBTI one of the dominant reliability concerns in nanoscale designs [10]. It is believed that NBTI is caused by broken Si-H bonds, which are induced by positive holes from the channel. Then H, in a neutral form, diffuses away; positive traps are left, which cause the increase of voltage threshold of the PMOS transistors [11].

For a PMOS transistor, there are two phases of NBTI, depending on its bias condition. In phase I, when $V_G=0$ (i.e., $V_{GS}=-V_{DD}$), positive interface traps are accumulating during the stress time with H atoms diffusing towards the gate. This phase is usually referred to as “stress” or “static NBTI”. In phase II, when $V_G=V_{DD}$ (i.e., $V_{GS}=0$), holes are not present in the channel, and thus, no new interface traps are generated; instead, H atoms diffuse back and anneal the broken Si-H. As a result, the number of interface traps is reduced during this stage and some of the NBTI effect is reversed. Phase II is referred to as “recovery” and can have a significant impact on NBTI effect estimation in VLSI circuits. The stress and recovery phases together are called “dynamic NBTI”.

NBTI effect on the threshold voltage is highly dependent on the temperature. Threshold voltage severely degrades in high temperatures. The huge impact of temperature is shown in section 6 through our simulation results. In addition, NBTI effect also depends on oxide thickness (technology node dependency), duty cycle, supply voltage and the voltage value of the signal applied to the gate of PMOS transistor [3].

In this paper, we consider the circuit under dynamic NBTI to model realistic circuit operation. There are some analytical models to express the change in V_{th} under dynamic NBTI [1][3][11]. In this paper in order to predict the threshold voltage degradation due to the NBTI effect at a time t and also considering duty cycle of stress vs. recovery phases, we adopt the model of reference [3].

2.2 Codependent setup and hold time

Latches and flip-flops are sequential circuit elements used in synchronous designs where a clock edge is used to sample and store a logic value on a data line. The setup time, τ_s , is the minimum time before the active edge of the clock that the input data line must be valid for reliable latching. Similarly, the hold time, τ_h , represents the minimum time that the data input must be held stable after the active clock edge. The clock-to-q delay refers to the

propagation delay from the 50% transition of the active clock edge to the 50% transition of the output, q , of the latch/register. The setup skew refers to the delay from the latest 50% transition edge of the data signal to the 50% active clock transition edge; similarly, the hold skew denotes the delay from the 50% active clock transition edge to the earliest 50% transition edge of the data signal. Figure 1 illustrates the setup and hold skews, which are denoted by τ_{sw} and τ_{hw} , respectively.

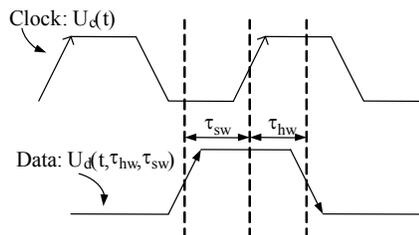


Figure 1: Setup and hold skews shown on the data and clock waveforms.

The setup (hold) time is a particular setup (hold) skew point, for which the *characteristic clock-to-q*¹, t_{cc2q} , delay increases by say 10%. (We shall denote as t_{cc2q} the clock-to-q delay which is 10% higher than t_{cc2q} .)

A common technique for setup/hold time characterization is first to generate the clock-to-q delay for various setup and hold skews via a series of transient simulations. This process in turn produces a clock-to-q *delay surface*. This is followed by extraction of a contour in the setup/hold time plane that contains all points that result in a given increase (e.g., 10% is typical) in t_{cc2q} . Figure 2 and Figure 4 show a typical clock-to-q surface and a CSHT contour plot.

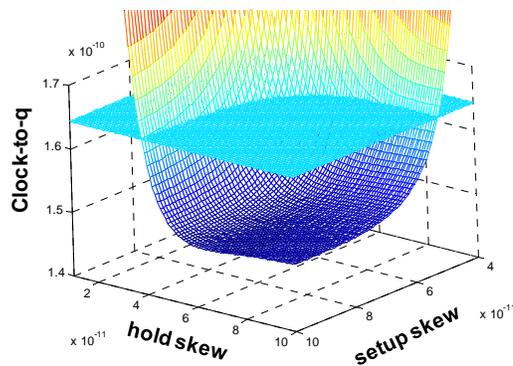


Figure 2: A clock-to-q surface.

The required setup time (RST) for a given flip-flop is defined as the minimum value of τ_{sw} for that flip-flop which results in a non-negative setup slack (i.e., the minimum setup skew needed to eliminate setup time violations for the flip-flop). The required hold time (RHT) is defined

¹ If the setup skew is larger than a certain value, then the clock-to-q delay of a flip-flop will become independent of the setup skew; this constant clock-to-q delay which is achieved for large setup skews is called the “characteristic clock-to-output delay” of the flip-flop.

similarly. On the other hand, the area above the CSHT contour is a pessimistic area where the flip-flop can correctly work in while the area under the CSHT contour is an overly optimistic area. Optimism is not permissible in STA, because it may result in failing chips. Therefore, the feasible working area for the flip-flop is the area above the CSHT contour. In addition, RST and RHT constraints must be satisfied. Hence, the flip-flop should be designed in a way to work in the shaded region in Figure 3 which is called the Feasible Region (FR).

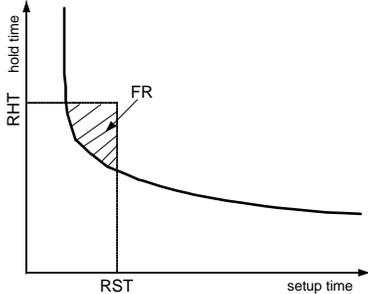


Figure 3: RST, RHT and FR in CSHT contour.

3. CSHT characterization

As mentioned the conventional method of extracting CSHT contour requires series of transient simulations to generate the t_{cc2q} surface which is not efficient when we need to obtain many contours. Authors in [13] proposed a method which numerically extracts the contour. In this section another efficient algorithm is proposed to tackle this problem which is more than two times faster than the algorithm proposed in [14]. We use Figure 4 to explain the proposed algorithm.

Definition 1 : The finite difference slope, α , of contour $\Gamma(\tau_s) = \tau_h$ at point $A = (\tau_s^A, \tau_h^A)$ is defined as: $\alpha = \frac{\tau_h^B - \tau_h^A}{\Delta\tau_s}$ where point B is a previously calculated point on Γ such that $\Delta\tau_s = \tau_s^B - \tau_s^A$. The superscript A in α^A denotes the point at which the finite difference slope α is calculated.

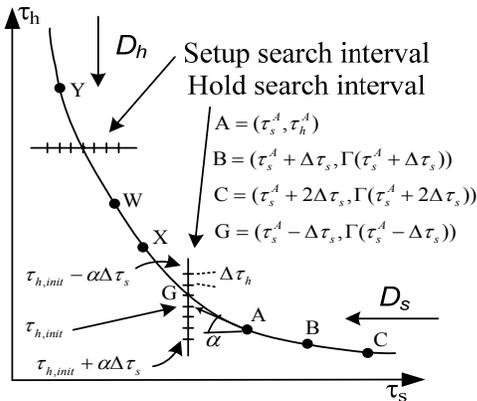


Figure 4: A setup/ hold time contour for given clock-to-q delay.

In *Definition 1*, we may want to use a point B as the reference point for slope calculation where $\tau_s^A - \tau_s^B = k\Delta\tau_s, k \geq 1$. In our experience $k = 2$ is a good

value. In the proposed algorithm (see below), we seek out the setup/hold pairs from two different directions, D_s and D_h as shown in Figure 4. The search through left-to-right direction, D_s starts from the largest setup time, τ_s^{Large} , C , and ends at point X . It is explained that for a given setup time, we look for its corresponding hold time in a hold time interval whose length is proportional to α calculated at previous setup/hold time pair. The slope at a given point A is used to guess the next point $G = (\tau_s^G, \tau_h^G)$ on Γ as follows:

$$\tau_s^G = \tau_s^A - \Delta\tau_s, \tau_h^G = \tau_h^A - \alpha\Delta\tau_s \quad (1)$$

The bounds of the search interval for hold time centered at point G is also given by $\tau_h^G \pm \alpha^A \Delta\tau_s$.

In order to reduce the time of search, the D_s search is carried out till point X where α^X is about 1 to 2 and is greatly less than α^Y which is around 20. In contrast to D_s search, the up-to-down search, D_h , is started from the largest hold time, τ_h^{Large} , and ends at point X . For the D_h search, the setup time is searched in a setup interval for a given hold time. Note that the finite difference slope for D_h search is $1/\alpha$. Since the finite difference slope for both searches, D_s and D_h , are bounded to 2, the run time for the proposed algorithm is 10X faster in comparison with the conventional method.

We next describe a backward Euler search (BES) algorithm to efficiently calculate the setup/hold time points for D_s and D_h . Let $\Delta\tau_s$ denote the setup time step resolution that the user intends to have for the CSHT characterization. The BES algorithm for D_s direction is as follows:

BES-Algorithm ($D_s, t_{cc2q}, \Delta\tau_s, \tau_s^{\text{Large}}$)

- i. Find t_{cc2q} for the flip-flop by doing a transient simulation with large setup and hold skews. Initialize $i = 1$ and τ_s^i to the largest setup time for which we want to calculate the corresponding hold time. A good guess for the largest value of setup time is half of the clock period. Next sweep the hold skew values and determine the hold time, τ_h^i .

- ii. Calculate slope α^i at (τ_s^i, τ_h^i) from *Definition 3*.

Notice $\alpha^1 = 0$ because Γ is asymptotic to a constant hold time value when $\tau_s \rightarrow \infty$.

- iii. Set $\tau_s^{i+1} = \tau_s^i - \Delta\tau_s$ and calculate the first guess for the hold time by using backward Euler (BE) method as follows (see Figure 4):

$$\tau_{h,init}^{i+1} = \tau_h^i - \alpha^i \Delta\tau_s \quad (2)$$

Sweep the hold skew values in the range of $\tau_{h,init}^{i+1} \pm \alpha^i \Delta\tau_s$ with time step $\Delta\tau_h$ (hold time step resolution) and find the hold time τ_h^{i+1} i.e., the value of hold skew which results in a clock-to-q delay equal to $1.1 \times t_{cc2q}$.

- iv. Repeat steps 2-3 for $i \geq 2$ till $\alpha \leq 2$ to compute setup/hold pairs on the contour.

To compute all the points of the contour, *BES-Algorithm* ($D_s, t_{cc2q}, \Delta\tau_s, \tau_s^{\text{Large}}$) and *BES-Algorithm* ($D_h, t_{cc2q}, \Delta\tau_h, \tau_h^{\text{Large}}$) are evaluated. For the latter one, D_h , means that all

's' subscripts are replaced by 'h' and vice versa in the body of *BES-Algorithm*. Some setup/hold time points of contour for the interval $1 \leq \alpha \leq 2$ are calculated twice (by both D_s and D_h) which can be replaced by their average. For example, two points $P_1 = (\tau_s^{P1}, \tau_h^{P1})$ and $P_2 = (\tau_s^{P2}, \tau_h^{P2})$ can be replaced by $\bar{P} = (0.5(\tau_s^{P1} + \tau_s^{P2}), \tau_h^{P1})$.

4. NBTI effect on CSHT

Increasing the threshold voltage of PMOS transistors, due to NBTI effect, results in variation in the CSHT characteristics. This means that for the same t_{c2q} , a new set of setup/hold time pairs should be obtained (cf. Figure 5 for a pictorial explanation). On the other hand, due to the NBTI effect, delay of combinational circuits itself increases. Therefore, given a fixed clock frequency, RST and RHT values will change and new STA requirements should be specified to achieve timing closure. By using NBTI-aware design techniques like [4] the delay of combinational logic blocks and clock drivers can be kept relatively unchanged. Notice that it is possible to extend our methodology to handle changes in the RST and RHT values.

In the presence of NBTI effect, a timing failure occurs when the new CSHT contour has no intersection with the FR. This means there is no setup and hold time pairs that result in non-negative setup and hold slacks. Figure 5 illustrates the effect of NBTI on the CSHT for the timing failure and non-failure cases.

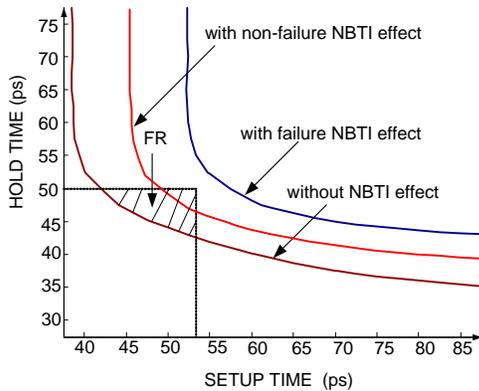


Figure 5: Setup/hold time codependency change due to the NBTI effect.

5. NBTI-aware flip-flop design

The variation in CSHT contour due to NBTI can cause a timing failure in the circuit. To overcome this failure the flip-flop must be designed in a way so as not to violate the timing constraints after aging effect. As we know all the timing characteristics of flip-flops depend on the sizing of the transistors in their circuits. Hence, we present a sizing technique for designing flip-flops to alleviate this aging problem. We also consider minimizing the energy consumption of the circuit. More precisely, NBTI effect causes increase in the t_{c2q} as well as a right upward shift of the CSHT contour. To compensate for this aging effect, we will size transistors in the flip-flop circuit in a way to shift the (new) CSHT contour below and to the left of the (original) CSHT contour. Therefore, after aging the new

CSHT contour will gradually move and approach the original CSHT contour due to NBTI effect.

5.1 Problem formulation

In this problem formulation we replace RST and RHT with maximum allowed changes in the setup and hold times of the flip-flop, respectively. These upper bounds should not be violated even after the NBTI-induced aging effect.

The objective of our optimization is to minimize the fresh state (i.e., at the beginning of circuit utilization) value of the energy-delay product of a flip-flop by imposing maximum degradation limits on the timing characteristics of the aged flip-flop. Constraints thus include upper bounds for changes in the t_{c2q} , setup and hold times of the flip flop due to the NBTI effect over a specific period of time. The solution of the optimization problem determines the transistor sizes in the flip-flop under consideration.

In addition, the target flip-flop may be operated at different voltage corners (i.e., it may be instantiated in different voltage islands in the design or, more notably, it may be subjected to different voltage levels due to employment of *dynamic voltage scaling* techniques in modern low power VLSI designs). Therefore, there are multiple supply voltage levels at which the flip-flop is desired to work correctly and energy-delay optimally. First, we introduce the optimization problem formulation for a single voltage corner and then extend it to multiple voltage corners.

5.1.1 Single corner optimization

The mathematical programming problem formulation for single corner operation may be stated as follows:

$$\begin{aligned}
 & \text{minimize} && Q(\vec{w}) = E^{fr}(\vec{w}), D^{fr}(\vec{w}) = \\
 & && E^{fr}(\vec{w}) \cdot (t_{c2q}^{fr}(\vec{w}) + \tau_s^{fr}(\vec{w})) \\
 & \text{subject to} && t_{c2q}^{aged}(\vec{w}) \leq t_{c2q,max} \\
 & && \tau_s^{aged}(\vec{w}) \leq \tau_{s,max} \\
 & && \tau_h^{aged}(\vec{w}) \leq \tau_{h,max}
 \end{aligned} \tag{3}$$

where $t_{c2q,max}$, $\tau_{s,max}$ and $\tau_{h,max}$ are maximum allowed values of t_{c2q} , setup and hold times, respectively and *fr* means the fresh state and *aged* means after aging effect happened for the specific period of time, e.g., three years. A sizing vector \vec{w} refers to set of transistors' sizes. Notice that the delay contribution of the launching flip-flop and the receiving flip-flop to the worst-case delay of the circuit is equal to t_{c2q} plus τ_s . Also note that instead of minimizing the energy-delay product of a fresh flip-flop circuit, we could have modeled and minimized the energy-delay product of a middle-aged circuit.

We refer to the solution of the optimization problem (3) as \vec{w}^* . We point out that for each sizing vector \vec{w} , there is one specific contour in the fresh state and one in the aged state since the timing characteristics of flip-flop change when the sizes of the transistors change.

5.1.2 Multi-corner Optimization

In the case of multiple voltage corners of operation, it is desired to simultaneously minimize all objective functions,

Q_i , where $i=1, \dots, m$. Q_i refers to the objective function of the i th corner. Associated with each corner i , there is a weight r_i which indicates the importance of the corner i in the multi-corner optimization, where

$$\sum_{i=1}^m r_i = 1 \quad (4)$$

Definition 2: Suppose the optimum solution for corner i is \bar{w}_i^* and $Q_i^* = Q_i(\bar{w}_i^*)$ is the *best objective value* at corner i . Clearly the objective function vector $Q^* = \{Q_1^*(\bar{w}), \dots, Q_m^*(\bar{w})\}$ is a lower bound (possibly infeasible) on the Pareto optimal set of solutions to the multi-criteria optimization problem. The *worst objective value* Q_i^{**} is defined as the $\text{Max}\{Q_i(\bar{w}_j^*)\}$ where maximization is over $j = 1, \dots, m$. Clearly $Q^{**} = \{Q_1^{**}(\bar{w}), \dots, Q_m^{**}(\bar{w})\}$ is an upper bound on the Pareto optimal set of solutions to the multi-criteria optimization problem.

Multi-corner-opt algorithm:

- i. Solve the optimization problem (3) for each corner separately to obtain \bar{w}_i^* 's, Q_i^* , and Q_i^{**} .
- ii. Solve the following nonlinear optimization problem

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^m \frac{r_i}{Q_i^{**} - Q_i^*} (Q_i(\bar{w}) - Q_i^*)^2 \\ \text{subject to} \quad & t_{c2q,i}^{aged}(\bar{w}) \leq t_{c2q,i,max} \text{ for } i=1, \dots, m \\ & \tau_{s,i}^{aged}(\bar{w}) \leq \tau_{s,i,max} \text{ for } i=1, \dots, m \\ & \tau_{h,i}^{aged}(\bar{w}) \leq \tau_{h,i,max} \text{ for } i=1, \dots, m \end{aligned} \quad (5)$$

where $t_{c2q,i,max}$, $\tau_{s,i,max}$ and $\tau_{h,i,max}$ are maximum allowed values of t_{c2q} , setup and hold times in corner i , respectively.

In fact, the optimization strategy in (5) is to minimize an L2-norm criterion. In this criterion, the distance of each function Q_i from its ideal value, Q_i^* , is weighted proportional to the priority of corner i , i.e., r_i , and normalized by the distance between worst and best objective values at that corner, i.e., $Q_i^{**} - Q_i^*$. Notice that in the absence of designer feedback about the weight of each voltage corner, we set $r_i = 1/m$.

5.2 Critical pair definition on CSHT contour

It is mentioned that each \bar{w} results in a different contour. To show the dependence of contours to the size of the transistors in the flip-flop, we define few critical points on each contour in the fresh state. These points are the critical points which can be defined by the designer. There can be two or three points as mentioned in [8]; for example, the points with minimum setup or hold times.

Definition 3: The *minimum setup plus hold times* (MSPH) point is defined as the point on a contour which has minimum $\tau_s + \tau_h$.

In most of the designs, the desired setup time is the minimum one to increase the clock frequency as much as

possible but there is a contrast between setup time and hold time in the sense that if one decreases the other one increases. In the case of minimum setup time, the hold time increases dramatically which causes hold violation in the circuit. Therefore, the desired point of operation for a flip-flop should be a point which minimizes the setup and hold times window which is MSPH point.

Hence, we choose MSPH point as the most critical point and throughout the rest of the paper we use it to do our analysis. This point can be easily found for each contour using *BES-Algorithm* which is explained in section 3.

5.3 Polynomial modeling of timing and power characteristics

After extracting each contour and finding the MSPH points on them using *BES algorithm*, we proceed by finding the polynomial functions which represents the MSPH points' setup and hold times in terms of transistor size vector, \bar{w} . These functions are the second order polynomials as follows

$$\sum_{i=1}^n \sum_{j \geq i}^n \alpha_{ij} W_i \cdot W_j + \sum_{i=1}^n \alpha_i W_i \quad (6)$$

where n is the number of the transistors in the flip-flop and W_i 's are the transistors' width.

The same technique is used to find second order polynomial functions for the aged state. So, we have $\tau_s^{fr} = f^{fr}(\bar{w})$, $\tau_s^{aged} = f^{aged}(\bar{w})$, $\tau_h^{fr} = g^{fr}(\bar{w})$ and $\tau_h^{aged} = g^{aged}(\bar{w})$.

Now, we have MSPH points which are $(\tau_s^{fr}, \tau_h^{fr}) = (f^{fr}(\bar{w}), g^{fr}(\bar{w}))$ pairs. To reduce the complexity of the problem, we find the best linear fit for MSPH points in the setup/hold time plane. We represent this line with $h=as+b$. This approximation helps removing one of the variables of the problem which is the hold time. In section 6 we show that the maximum error of this approximation is 5% in the worst case, which demonstrates the good quality of the approximation.

Energy and t_{c2q} are also modeled with the second order polynomial functions. Hence, the one corner optimization problem (3) considering second order polynomial modeling (6) can be rewritten as:

$$\begin{aligned} \text{minimize} \quad & (\sum_{i=1}^n \sum_{j \geq i}^n \alpha_{ij} W_i \cdot W_j + \sum_{i=1}^n \alpha_i W_i) \cdot (\sum_{i=1}^n \sum_{j \geq i}^n \beta_{ij} W_i \cdot W_j + \\ & \sum_{i=1}^n \beta_i W_i + \sum_{i=1}^n \sum_{j \geq i}^n \gamma_{ij} W_i \cdot W_j + \sum_{i=1}^n \gamma_i W_i) \end{aligned}$$

subject to:

$$\begin{aligned} \sum_{i=1}^n \sum_{j \geq i}^n \theta_{ij} W_i \cdot W_j + \sum_{i=1}^n \theta_i W_i & \leq t_{c2q,max} \\ \sum_{i=1}^n \sum_{j \geq i}^n \rho_{ij} W_i \cdot W_j + \sum_{i=1}^n \rho_i W_i & \leq \tau_{s,max} \end{aligned} \quad (7)$$

The multi-corner optimization formulation (5) can be rewritten by using second order polynomial functions (6), which is omitted for brevity.

5.4 Algorithm

The algorithm used to do the characterization and design flow is as follows.

- i. For each combination of transistor sizes:
 - a. Extract the MSPH point on each contour in fresh state using *BES-Algorithm*.
 - b. Measure t_{c2q} delay and energy consumption.
- ii. Find the best linear fit for MSPH points of contours in setup/hold time plane ($h=as+b$).
- iii. Find second order polynomial functions which represent the MSPH points' energy consumption, setup time and t_{c2q} in terms of size vector. We already know that hold time is linearly dependent on the setup time.
- iv. Repeat steps 1-3 for the NBTI-affected (aged state) flip-flop.
- v. Call *Multi-corner-opt* algorithm to solve the optimization problem for multi-corner optimization or (7) in the case of single corner optimization.
- vi. Make the results discrete in terms of λ .

6. Simulation results

We apply our mathematical program to TSPC flip-flop to determine the best transistor sizes for different corners of operation and also the optimum solution for multi-corner optimization. One of these corners is the corner representing the extreme NBTI effect, i.e., for high operating temperature (85 degree Celsius). It must be mentioned that the flip-flop is originally designed to have the minimum energy-delay product in the fresh state and the input signal probability is 0.5. All simulation results in this paper are obtained by HSPICE using a predictive 65nm technology model [12].

6.1 Polynomial modeling results

Timing and power characteristics of the flip-flop are modeled by using the second order polynomials. As an example the error histogram for modeling the fresh setup time is provided in Figure 6. The reported data is the relative error for data collected from HSPICE and the result of our modeling. The rest of the histograms (for modeling the other parameters) are omitted for brevity. However, Table 1 reports the error statistics of all parameter modeling results for the TSPC and master-slave (MS) flip-flops. We can see that the maximum error occurs in the modeling of fresh t_{c2q} of MSFF which is 5% (although its mean and standard deviation of error are 0.6% and 0.5%, respectively).

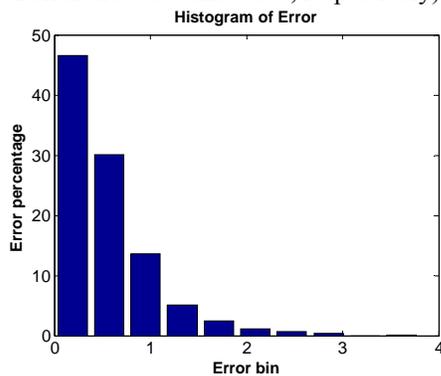


Figure 6: Histogram of error in fresh setup time modeling.

Table 1: Error statistics of modeling of flip-flops characteristics

Error (%)	Max		Mean		Standard deviation	
	TSPC	MS	TSPC	MS	TSPC	MS
Flip-Flop						
Fresh setup time	1.8	3	0.3	0.4	0.25	0.3
Fresh t_{c2q}	4	5	0.55	0.6	0.5	0.5
Aged setup time	2	3.2	0.3	0.5	0.27	0.3
Aged t_{c2q}	3.8	4.5	0.55	0.6	0.48	0.55
Energy consumption	2.8	3	0.4	0.55	0.3	0.35

6.2 Optimization results

In this section we apply the proposed optimization algorithm to TSPC flip-flop to optimally size the transistors in its circuits to overcome the NBTI effect.

The positive edge TSPC flip-flop, whose transistor-level schematics is shown in Figure 7, features positive setup and hold times. The setup time is equal to the delay of the stage 1 (clocked) inverter whereas the clock-to-q delay is related to the summation of delays of the last three stages of the flip-flop. The hold time is the difference of the falling delays of stage 1 and stage 2 inverters.

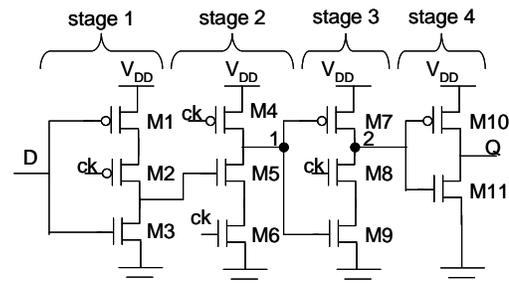


Figure 7: Positive edge-triggered TSPC flip-flop.

6.2.1 Single corner optimization experiments

The first step for the simulation is to find the optimum sizing vector for the flip-flop, which minimizes the energy-delay product in the fresh state for 25°C and 1.0V supply voltage. We denote this optimal sizing vector as \vec{w}^{1*} . Table 2 shows values of the energy-delay product in the fresh state, area, clock-to-q delay, setup time, and power consumption in the aged state for \vec{w}^{1*} .

The aging effect experiment is done by changing the threshold voltage of the PMOS transistors in the TSPC circuit. We considered the effect of aging on the flip-flop after three years of operation. The increase in the threshold voltage due to NBTI effect is calculated using the model provided in [3].

The characteristic values of TSPC flip-flop for \vec{w}^{1*} are shown in Table 2. It can be seen that the aged setup time and t_{c2q} are increased by 15% and 21%, respectively. This amount of NBTI-induced increase in the timing characteristics of flip-flops is not acceptable and causes timing failure in the VLSI circuits.

Table 2: TSPC FF characteristics for \bar{w}^{1*}

State	E.D (fJ.ns)	setup time (ps)	t_{c2q} (ps)	Power (μW)	Area (fm^2)
fresh	0.371	20	70	1.423	234
aged	0.399	23	85	1.269	234
percentage	7	15	21	-10	0

To overcome this undesirable outcome, transistors in the TSPC flip-flop should be sized up. We use the algorithm given in 5.4 to size the transistors. Here the optimization is just for one voltage corner (1.0V). We consider that up to 5% increase (compared to the corresponding fresh state values) in the setup time and t_{c2q} after a three-year aging process is acceptable. Hence, values of $\tau_{s,max}$ and $t_{c2q,max}$ in the mathematical program (7) are 21ps and 73.5ps, respectively. We denote the sizing solution of this optimization problem as \bar{w}^{2*} . The results of single corner optimization problem for TSPC flip-flop are reported in Table 3. The degradation percentages are calculated with respect to the fresh state values of for sizing vector \bar{w}^{1*} , i.e., the original energy-delay optimized TSPC flip-flop.

Notice that the reason for 9% power consumption decrease in the aged state is the reduction in the leakage power. The leakage power decreases because of the increase in the threshold voltage of the PMOS transistors due to NBTI effect.

Table 3: Single corner optimization results for TSPC with sizing vector \bar{w}^{2*}

	E.D (fJ.ns)	setup time (ps)	t_{c2q} (ps)	Power aged state (μW)	Area (fm^2)
Value	0.393	21	73	1.288	248
Percentage	6	5	5	-9	6

6.2.2 Multi-corner optimization experiments

There are four different corners (A through D) in our experiment corresponding to two voltage levels (1.0 and 1.2V) and two temperature values (25 and 85°C). The first step is to optimize each corner individually; therefore, the sizing vector solution for each corner is different from the others. The single corner optimization results for each corner are shown in Table 4.

We consider the results for fresh state of the flip-flop with sizing vector \bar{w}^{1*} as the baseline. From now on, all the comparisons are with respect to this baseline.

The values of constraints for 25°C corners are 5% increase in the fresh values of setup time and t_{c2q} , which are the same as before. However, for 85°C corners there are no feasible solutions with these constraint values. So, we must relax the constraints (allow 20% increase in the fresh values of setup time and t_{c2q}).

Table 4 entries for corner B show the importance of supply voltage level. When higher voltage value is used, the timing characteristics improve but power consumption becomes much larger. The magnitude of the power

consumption is so high that the solution for the optimization problem becomes the same as \bar{w}^{1*} , i.e., transistor sizes are not increased so as to keep the power consumption as low as possible. However, the objective function (energy-delay product) value for this corner is much higher than that for corner A. Results of corners C and D underline the big influence of temperature on the NBTI effect.

Table 4: Single corner optimization results for TSPC on four different corners

Corner	Temp (°C)	Voltage (V)	E.D (%)	setup time (%)	t_{c2q} (%)	Power aged state (%)	Area (%)
A	25	1.0	6	5	5	-9	6
B	25	1.2	32	2	4	+25	-3
C	85	1.0	127	20	20	+64	65
D	85	1.2	133	20	20	+72	12

The idea of multi-corner optimization is to find a single solution which produces good results in all corners of interest. So far, we found the optimum solution for each corner independent of the other corners. Now, we check the results of using the optimum solution for one corner under the parameter setting of another corner.

We use the following notation to illustrate our results: AB means the optimum sizing vector of corner A (\bar{w}_A^*) is applied to the parameter setting of corner B. Table 5 demonstrates the results of this experiment. It can be seen that the results are worse than the results in Table 4 since the optimum solution of one corner is applied to the setting of another corner.

Table 5: Results of optimum solution of one corner in the environment of other corners

Case	E.D (%)	setup time (%)	t_{c2q} (%)	Power aged state (%)	Area (%)
AB	38	-14	-10	+45	6
BA	12	20	33	-15	-3
CD	164	7	9	+95	65
DC	152	56	39	+37	12

Now, we use mathematical programming (5) to find the multi-corner optimization solution on two voltage corners 1.0V and 1.2V for 25°C, i.e., corners A and B. We consider that both corners are equally important ($r_A = r_B = 0.5$). Table 6 shows the results of this experiment. It is clear that the results of multi-corner optimization are worse than each corner's results for its own optimum sizing vector in Table 4 but they are better than the results in Table 5. Figure 8 illustrates the comparison between the results of Table 4, Table 5 and Table 6 (single corner vs. multi-corner optimizations). In Figure 8, multi-corner A (B) means evaluating the results of multi-corner optimization at corner A (B).

Table 6: Multi-corner optimization results for TSPC evaluated at different corners

Corner	E.D (%)	setup time (%)	t_{c2q} (%)	Power aged state (%)	Area (%)
A	9	7	8	-12	4
B	35	1	2	+40	4

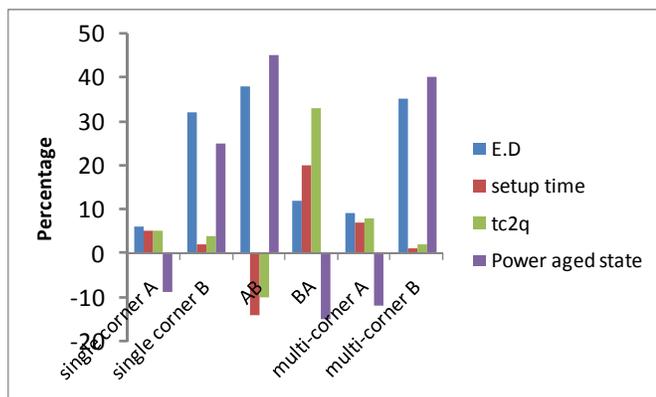


Figure 8: Comparison between single corner and multi-corner optimizations

7. Conclusion

In this paper, we studied the NBTI effect on the codelpendency of setup and hold times of flip-flops. We introduced BES-Algorithm as an efficient method to characterize CSHT contour and find MSPH point on it. We also used polynomial modeling technique and showed that its error is negligible. Consequently, we introduce our multi-corner optimization algorithm to minimize the energy delay product of flip-flops with aging effects on timing characteristics as constraints. We used nonlinear programming to solve the optimization problem to find the best transistor sizes. Finally, experimental results used to show our modeling accuracy and new transistor sizes for TSPC flip-flop.

References

[1] B.C. Paul, K. Kang, H. Kuflouglu, M. A. Alam and K. Roy, "Impact of NBTI on the temporal performance degradation of digital circuits," *Electron Device Letters*, vol. 26, no. 8, pp. 560-562, 2005.

[2] D.K. Schroder and J.A. Babock "Negative bias temperature instability: Road to Cross in Deep

Submicron Silicon Semiconductor Manufacturing," *Journal of Applied Physics*, 2003.

[3] S. Bhardwaj, W. Wang, R. Vattikonda, Y. Cao, and S. Vrudhula, "Predictive modeling of the NBTI effect for reliable design," *Custom Integrated Circuits Conference*, 2006.

[4] B.C. Paul, K. Kang, H. Kufluoglu, M. A. Alam, and K. Roy, "Negative bias temperature instability: estimation and design for improved reliability of nanoscale circuits," *Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, No. 4, pp. 743-751, Apr. 2007.

[5] K. Duraisami, E. Macii, and M. Poncino, "Using soft-edge flip-flops to compensate NBTI-induced delay degradation," *Great Lakes Symposium on VLSI*, 2009.

[6] W. Wang, S. Yang, S. Bhardwaj, R. Vattikonda, S. Vrudhula, F. Liu, and Y. Cao, "The impact of NBTI on the performance of combinational and sequential circuits," *Design Automation Conference*, 2007.

[7] K. Ramakrishnan, X. Wu, N. Vijaykrishnan, and Y. Xie, "Comparative analysis of NBTI effects on low power and high performance flip-flops," *International Conference on Computer Design*, 2008.

[8] H. Abrishami, S. Hatami, B. Amelifard, and M. Pedram, "NBTI-aware flip-flop characterization and design," *Great Lakes Symposium on VLSI*, 2008.

[9] E. Salman, A. Dasdan, F. Taraporevala, K. Kucukcakar, and E.G. Friedman, "Exploiting setup-hold-time interdependence in static timing analysis," *Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, no. 6, Jun. 2007.

[10] *International technology roadmap for semiconductors*. Semiconductor Industry Association, 2005, <http://www.itrs.net/>

[11] R. Vattikonda, W. Wang, and Y. Cao, "Modeling and minimization of PMOS NBTI effect for robust nanometer design," *Design Automation Conference*, 2006.

[12] <http://www.eas.asu.edu/~ptm/>

[13] S. Srivastava and J. Roychowdhury, "Interdependent latch setup/hold time characterization via Euler-Newton curve tracing on state-transition equations," *Design Automation Conference*, 2007.

[14] S. Hatami, H. Abrishami, and M. Pedram, "Statistical timing analysis of flip-flops considering codependent setup and hold times," *Great Lakes Symposium on VLSI*, 2008.