Adaptive Techniques for Overcoming Performance Degradation due to Aging in CMOS Circuits

Sanjay V. Kumar, Chris H. Kim, and Sachin S. Sapatnekar
University of Minnesota, Minneapolis MN 55455

***Abstract*— Negative bias temperature instability (NBTI) in PMOS transistors has become a major reliability concern in present-day digital circuit design. Further, with the recent introduction of Hf-based high-k dielectrics for gate leakage reduction, positive bias temperature instability (PBTI), the dual effect in NMOS transistors, has also reached significant levels. Consequently, designs are required to build in substantial guardbands in order to guarantee reliable operation over the lifetime of a chip, and these involve large area and power overheads. In this paper, we begin by proposing the use of adaptive body bias (ABB) and adaptive supply voltage (ASV) to maintain optimal performance of an aged circuit, and demonstrate its advantages over a guardbanding technique such as synthesis. We then present a hybrid approach, utilizing the merits of both ABB and synthesis, to ensure that the resultant circuit meets the performance constraints over its lifetime, and has a minimal area and power overhead, as compared with a nominally designed circuit.**

I. Introduction

NBTI (Negative Bias Temperature Instability) in PMOS devices has become a major reliability issue in sub-130nm technologies. NBTI manifests itself as a temporal increase in the threshold voltage, V_{th}, of a PMOS transistor, thereby causing circuit delays to degrade over time and exceed their specifications. A corresponding and dual effect, known as Positive Bias Temperature Instability (PBTI) [1]–[3], is seen for NMOS devices, when a positive bias stress is applied across the gate oxide of the NMOS device. Although the impact of PBTI is lower than NBTI [2], it is increasingly becoming important in its own right, particularly with the use of Hf-based high-k gate oxides for leakage reduction [1], [3]. Collectively, NBTI and PBTI are referred to as bias temperature instability (BTI) effects.

Previous approaches to guardbanding a circuit and ensuring optimal performance over its lifetime, such as sizing [4], [5], and synthesis [6][1] can be classified as "one-time" solutions that add appropriate guardbands at design time. These methods are generally formulated as optimization problems of the type:

$$\begin{aligned} &\text{Minimize } \alpha \text{ Area } + \beta \text{ Power} \\ &\text{s.t. } D(0 \le t \le t_{\text{life}}) \le D_{\text{spec}}, \end{aligned} \tag{1}$$

where α and β are weights associated with the area and the power objectives, respectively, while D_{spec} is the specified target delay that must be met at all times, up to the lifetime of the circuit, t_{life}.

Under the framework of (1), both the synthesis and sizing optimizations lead to an increase in area and power, as compared with a nominally designed circuit that is built to meet the specification only at birth, and not necessarily over its entire life. The work in [6] argues that synthesis can lead to area and power savings, as compared with sizing optimizations. However, guardbanding (through sizing or synthesis) is performed during design time, and is a one-time fixed amount of padding added into the circuit in the form of gates with a higher drive strength. Inevitably, this results in large positive slacks during the initial stages of operation of the circuit, and therefore, larger-than-necessary area and power overheads, in comparison with a circuit designed to exactly meet the specifications throughout its lifetime.

We also note that while BTI effects cause the transistor threshold voltages to increase, resulting in larger delays, higher V_{th} also implies lower subthreshold leakage ($I_{\text{sub}} \propto e^{\frac{-V_{th}}{mkT}}$). Therefore, both NBTI and PBTI cause the leakage of the circuit to decrease with time, thereby providing the opportunity to trade off this slack in leakage to restore the lost performance. Adaptive Body Bias (ABB) [7] provides an attractive solution to explore leakage-performance trade-offs. Forward body bias (FBB) can be used to speed up a circuit [8], by reducing the V_{th}, thereby using up the available slack in the leakage budget. Further, the amount of FBB can be determined adaptively, based on the exact temporal degradation of the circuit, and requisite amounts of body bias can be applied to exactly meet the target specifications under all conditions.

The main advantage of a body bias scheme is that the performance can be recovered with a minimal increase in the area overhead as compared with "one-time" approaches such as sizing and synthesis. While [9] demonstrated that ABB could be used to allow the circuit to recover from voltage and temperature variations as well as aging, we believe our work is the first solution to take advantage of the reduction in leakage due to bias temperature instability (BTI). We demonstrate how ABB can be used to maintain the performance of the circuit over its lifetime, by determining the appropriate PMOS and NMOS body bias values (and supply voltages) at all times. We use a look-up table whose entries consist of the optimal body bias and supply voltages, indexed by the cumulative time of BTI stress on the circuit.

Accordingly, we first propose an optimization algorithm to compute the entries of the look-up table, such that the delay specifications of the circuit are met throughout its lifetime

[1]The work in [6] uses a BTI-aware delay model during the technology mapping phase of synthesis, and the circuit is mapped to a library such that its delay at the end of its lifetime, t_{life}, meets the specifications. In the context of this paper, we will refer to this approach as the "synthesis approach."

and the power overhead is minimized. In contrast with the significant area cost for the synthesis-based method, the area overhead using this approach is limited to the look-up tables, body bias generation, and body bias routing networks and associated control circuitry, and is therefore minimal, while the power overhead is similar to that incurred by synthesis. Thus, we show that the adaptive compensation of circuit delay degradation due to aging provides a viable alternative to "one-time" fix techniques such as BTI-aware synthesis.

In the second approach, we propose an alternative hybrid formulation that combines adaptive techniques with synthesis. This iterative method first performs a power-constrained delay minimization through the application of FBB. This optimization recovers some amount of the performance degradation caused by aging by using the power slack that is created as the circuit ages. However, since this power-constrained optimization is not guaranteed to meet the delay specifications, technology mapping is used next to resynthesize the circuit to meet tighter timing specifications at birth. Using a new power specification, the iteration continues through alternate steps of FBB optimization and resynthesis until the timing specification is met. Our simulation results indicate that by combining the merits of the adaptive and synthesis-based approaches, the resulting circuit meets the performance constraints at all times, with only a minimal expense in the area and power.

The paper is organized as follows. Section II outlines the impact of BTI on the delay and leakage of circuits, motivating a scheme for ensuring reliable operation. The effectiveness of FBB in maintaining optimal performance subject to power constraints is explored, and two optimization schemes are outlined, considering some combination of adaptive body bias (ABB), adaptive supply voltage (ASV) and resynthesis. Section III focuses on the control system implementation, while algorithms for circuit optimization are presented in Section IV and V. Simulation results are presented in Section VI where we compare the area, delay and power numbers, as a function of time, for the various approaches, followed by concluding remarks in Section VII.

II. Background and Problem Formulation

We begin this section by determining the impact of NBTI on the delay and leakage of digital circuits. We then explore the potential of FBB to achieve power-performance trade-offs, and accordingly formulate an optimization problem.

A. Impact of BTI on Delay and Leakage

At the transistor level, the reaction-diffusion (R-D) framework [10], [11] has widely been used to determine the long-term impact of NBTI on the threshold voltage degradation of a PMOS device. Accordingly, the V_{th} degradation for a PMOS transistor under DC stress increases asymptotically with time, t, as $\Delta V_{th}(t) \propto t^{\frac{1}{6}}$ [12]–[15]. We also use a PBTI model where the degradation mechanism is similar to NBTI, but the magnitude of V_{th} degradation is lower. Specifically, in our simulations, the ΔV_{th} for a PMOS device after 10^8seconds ($\approx$ 3 years) of DC stress is $\approx$ 50mV, while that for an NMOS device is $\approx$ 30mV. The corresponding nominal values of the threshold voltages, based on PTM 45nm model files [16], are -411.8mV for a PMOS device and 466mV for an NMOS device. Since NBTI affects the V_{th} of PMOS devices, it alters the rising delay of a gate. Similarly, PBTI, which affects NMOS transistors, changes the falling delay of a gate.

At the gate level, we derive models for the delay and the leakage as functions of the transistor threshold voltages. We assume the worst-case degradation [4] model for all gates in the circuit, for reasons that will become apparent in Section III. The delay and leakage numbers for the degraded circuit are computed through SPICE simulations, at T = 105°C, at different times. Since BTI is enhanced with temperature, the library gates are characterized at the maximum operating temperature of the chip, assumed to be T = 105°C.

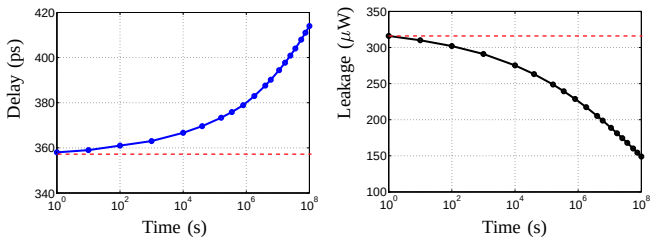
The results from the above SPICE simulations are curve-fitted to obtain models for the delay and leakage as a function of the transistor threshold voltages. The gate delay, D, is modeled as:

$$D(t) = D_0 + \sum_{\text{all transistors } i} \frac{\partial D}{\partial V_{th_i}} \Delta V_{th_i}(t) \qquad (2)$$

where the sensitivity terms, $\partial D \partial V_{th}$, for each of the transistors in the gate, along the input-output path, are determined through a linear least-squares curve-fit. This first order sensitivity-based model is accurate, and has an average error of 1% in comparison with the simulation results, within the ranges of V_{th} degradation caused by BTI. Similarly, a model for leakage, L, can be developed as:

$$\log L(t) = \log L_0 + \sum_{\text{all transistors } i} \frac{\partial L}{\partial V_{th_i}} \Delta V_{th_i}(t) \qquad (3)$$

Note that the $\Delta V_{th}(t)$, D_0, and L_0 values are functions of the supply voltage, V_{dd}. The leakage numbers are experimentally verified to have an average error of 5% with respect to the SPICE simulated values.



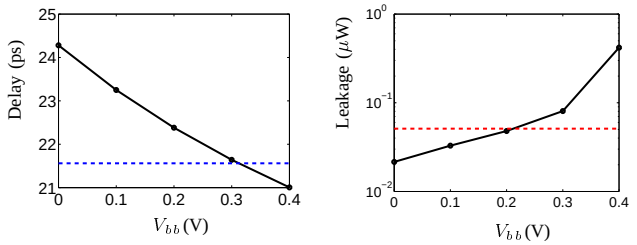
(a) (b)

Fig. 1. The impact of BTI on (a) the delay and (b) the leakage of the LGSYNTH93 benchmark, "des," as a function of time.

At the circuit level, Fig. 1 shows the impact of BTI on the delay and leakage of an LGSYNTH93 benchmark "des" as a function of time. The delay and leakage of the uncompensated circuit at $t = 0$, are shown by flat dotted lines on each plot. The results indicate that the delay degrades by around 14%, whereas the subthreshold leakage reduces by around 50%, after three years of operation. We ignore the contribution of gate leakage current here, since neither BTI nor FBB impacts

the gate leakage. Further, with the use of high-k dielectrics, gate leakage has been reduced by several orders of magnitude, making it negligible in comparison with the subthreshold and junction leakages.

B. Recovery in Performance using FBB



(a) (b)

Fig. 2. The impact of applying FBB to a degraded inverter at $t = t_{\text{life}}$ on its (a) delay and (b) its leakage.

At first glance, one may imagine that by returning the threshold voltage to its original value, FBB could be used to fully recover any degradation in the PMOS/NMOS transistor threshold voltage, bringing the I_{on} and I_{off} values of the device to their original levels, thereby completely restoring its performance and leakage characteristics. However, on closer examination, it is apparent that this is not the case, due to the effect of the substrate junction leakage. The results of applying FBB on a temporally degraded inverter (after three years of constant continuous stress on all the transistors) are shown in Fig. 2. Fig. 2(a) shows the average delay[2] of the inverter, measured as $\frac{1}{2}(\tau_R + \tau_F)$, where τ_R and τ_F are the output rise and fall times, respectively, plotted against the body bias voltage V_{bb}. Here, we apply an equal V_{bb} to all devices. The value of the delay at zero body bias represents the delay of the aged circuit. The horizontal dotted line represents the delay specification, and after three years of maximal aging, the circuit clearly violates this requirement. At this point, the application of a V_{bb} of $\approx$ 0.3V can restore the delay of the inverter to its original value.

Fig. 2(b) plots the corresponding total leakage power, consisting of the sum of the subthreshold leakage and the substrate junction leakage, under maximal V_{th} degradation of both the NMOS and the PMOS transistors, The leakage computed at $t = 0$, i.e., with ΔV_{th} = 0, shown by the horizontal dotted line, is chosen as the leakage budget: after three years of aging, the leakage value (shown at zero body bias) falls below this budget. The figure shows that with the application of V_{bb}, the leakage rises, and exceeds the budget at around 0.2V. In particular, the exponential increase in substrate junction leakages with FBB leads to a sharp increase in the leakage beyond a certain point.

[2]The trend of average of rise and fall delays mimics that of a path delay, since alternate stages of logic in a path (consisting of all inverting gates) undergo rising and falling transitions, respectively.

From Figs. 2(a) and 2(b), it can be inferred that a complete recovery in the delay degradation of the circuit could cause the leakage current to exceed its nominal value. Simulation results indicate that our benchmark circuits require FBB of the order of (0.3-0.4V), which leads to a large increase in the power dissipation, and can potentially exceed the available budget.

In other words, the sole use of ABB (FBB) to restore fully the performance of the circuit results in a substantial power overhead, particularly as we approach the lifetime of the circuit, where large values of FBB are necessary. The use of ASV in combination with ABB has been demonstrated to be more effective than using ABB individually [17]. Hence, we propose our first method, termed the "adaptive approach," that applies ASV in conjunction with ABB to minimize the total power overhead, while meeting the delay constraints throughout the circuit lifetime.

As we will see from the results in Section VI, while the adaptive approach provides area savings in comparison with the synthesis approach, the maximal power dissipation overhead is significant. Although ASV, when used in combination with ABB, tempers the exponential increase in junction leakages with FBB, the corresponding increases in V_{dd} cause the subthreshold leakage to increase exponentially, while the active power increases quadratically. Further, the amount of threshold voltage degradation has a second order dependence on the supply voltage, with larger V_{dd} leading to higher ΔV_{th} [18]. Our second approach further reduces the power dissipation by combining the merits of the adaptive and synthesis approaches, thereby trading off area with power. In particular, it supplements the use of ABB with synthesis, instead of ASV as in the adaptive approach, yielding improved trade-offs. We refer to this as the "hybrid approach."

C. Adaptive Approach

Under the adaptive approach, the optimal choice of the values of the NMOS body bias voltage, v_{bn}, the PMOS body bias voltage, v_{bp}, and the supply voltage, V_{dd}, to meet the performance constraint is such that the total power dissipation at all times is minimized. An optimization problem may be formulated as follows:

$$\begin{aligned} \text{Minimize } P_{\text{act}}(t, V_{dd}) &+ P_{\text{lkg}}(t, v_{bn}, v_{bp}, V_{dd}) \\ \text{s.t. } D(t, v_{bn}, v_{bp}, V_{dd}) &\leq D_{\text{spec}} \\ 0 \leq t &\leq t_{\text{life}}, \end{aligned} \tag{4}$$

where P_{act} and P_{lkg} are the weighted active and leakage (subthreshold + junction leakage) power values, respectively, while D_{spec} is the timing specification that must be met at all times. It can be intuitively seen that a solution to the optimization problem in (4) attempts to maintain the circuit delays to be as close to (but still lower than) the specification as possible, since any further reduction in delay using ABB/ASV is accompanied by a corresponding increase in the active and leakage power dissipation.

D. Hybrid Approach

The hybrid approach uses a combination of adaptive methods and presilicon synthesis to optimize the circuit for aging

effects. The use of ASV results in a quadratic increase in the active power; in contrast, at reasonable design points, synthesis can provide delay improvements with subquadratic increases in the power dissipation. Therefore, the hybrid adaptive approach is restricted to the use of ABB only, at the nominal V_{dd} value.

The hybrid approach employs synthesis and ABB in an iterative loop, tightly controlling the power increase in each step. For the ABB assignment step of the loop, the optimization formulation in (4) is recast within a power envelope, as a problem of delay minimization subject to power constraints.

$$\begin{aligned} \text{Minimize} \quad & \max_{t\in[0,t_{\text{life}}]} D(t, v_{bn}, v_{bp}) \\ \text{s.t. } & P_{\text{lkg}}(t, v_{bn}, v_{bp}) \leq P_{\text{lkg}}(t=0) \\ & 0 \leq t \leq t_{\text{life}}, \end{aligned} \tag{5}$$

where $P_{\text{lkg}}(t=0)$ denotes the leakage power budget. This budget is taken to be the peak leakage of the uncompensated circuit, i.e., its leakage at $t = 0$. Note that this effectively bounds the total power dissipation of the circuit to its value at $t = 0$, since the above optimization has a negligible effect on the active power dissipation.

The solution to the above optimization problem reduces the delay of the circuit under power constraints, but does not guarantee that the delays will be lower than D_{spec}. If this is the case, in a second step, the circuit is resynthesized to meet a heuristically chosen delay specification, tighter than D_{spec}, at $t = 0$. The iteration continues until the optimization in (5) can guarantee that the compensated circuit meets D_{spec} over its entire lifetime.

III. Control System for Adaptive Compensation

In this section, we investigate how an adaptive control system can be implemented to guardband circuits against aging. Prior work in this area can be summarized as follows. A look-up-table-based approach that precomputes and stores the optimal ABB/ASV/frequency values, to compensate for droop and temperature variations, is presented in [9]. An alternative approach [7], [8] uses a replica of the critical path to measure and counter the effects of within-die and die-to-die variations. Techniques for sensor design have been addressed in [19], [20], which propose high-resolution on-chip sensors for capturing the effects of aging.

However, with increasing levels of intra-die variations, critical path replica-based test circuits require a large number of critical paths to provide an f_{max} distribution that is identical to the original circuit, leading to an area overhead. Further, the critical paths in a circuit can dynamically change, based on the relative temporal degradation of the potentially critical paths. Adding every potentially critical path from the original circuit into the critical path replica may cause the test circuit to become extremely large. Apart from a high area overhead, such a large test circuit may incur its own variations that may be different from those in the original circuit.

Owing to these drawbacks, we propose the use of a look-up-table-based implementation to determine the actual v_{bn}, v_{bp}, and V_{dd} values that must be applied to the circuit to compensate for aging. The entries in the look-up table are indexed by the total time for which the circuit has been in operation. This time can be tracked by a software routine, with $t = 0$ representing the beginning of the lifetime of the circuit after burn-in, testing, and binning. The degradation in delays due to accelerated stresses at high temperature during burn-in are accounted for in determining $D(t = 0)$, by adding an additional timing guardband. This software control enables the system to determine the total time for which the circuit has been operational.

The look-up table method requires the critical paths and the temporal delay degradation of the circuit to be known beforehand, to determine the entries of the table. It is impossible to determine, *a priori*, the exact temporal degradation of a circuit, since this depends on the stress patterns, which in turn depend on the percentage of time various circuit nodes are at logic levels 0 and 1. This percentage depends on the profile of computations executed by the circuit, and cannot be captured accurately by, for example, an average probabilistic analysis. The only guaranteed-pessimistic measure for BTI stress uses the worst-case degradation of the circuit. The method in [4] presents such a method, considering the impact of NBTI only, and determines the worst-case scenario by assuming maximal DC stress on every PMOS transistor. The idea can be extended to include maximal impact of PBTI on the NMOS transistors, as well, to compute the maximal degradation of the most critical path in the circuit. The worst-case method to estimate the maximal delay degradation after t seconds of aging is computationally efficient, is input-vector-independent, and requires a single timing analysis run based on the degraded NMOS and PMOS V_{th} values at t. Due to the fact that this is guaranteed-pessimistic over all modes of circuit operation, the set of v_{bn}, v_{bp}, and V_{dd} values in (4), determined using this number as a measure of $D(t)$ in this formulation, is guaranteed to ensure that the circuit meets the delay specification under all operating conditions.

The next sections describe the algorithms for the adaptive and the hybrid approaches to counter the impact of BTI. In Section IV, we first outline an algorithm for the adaptive approach to compute the optimal tuple entries in the look-up table at different times. We then investigate how further area-power trade-offs can be achieved using the hybrid approach in Section V, and describe the implementation.

IV. Optimal ABB/ASV Computation for the Adaptive Approach

We will begin by pictorially illustrating the idea of the adaptive approach. Fig. 3 shows the temporally degraded delay, $D(t)$, of the original circuit without ABB/ASV, where the delay monotonically increases with time, and violates D_{spec} for some $t \geq t_{i-1}$. The figure shows how ABB/ASV may be applied at a time t_{i-1}, to ensure that the delay degradation during the interval $[t_{i-1}, t_i]$ does not cause the circuit delay to exceed the specifications. The delay of the circuit immediately after applying ABB/ASV, based on the look-up table values at t_{i-1}, is denoted as $D(t_{i-1})$, and is guaranteed to always be less than D_{spec}. Similarly, $D(t_i-)$ is the delay of the circuit

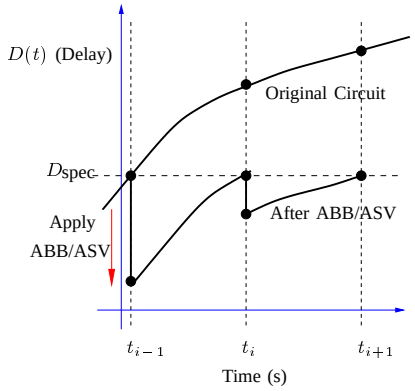


Fig. 3. A plot of the delay of the original circuit, without adaptation, as a function of time, showing degradation due to BTI effects, and a schematic of our compensation scheme using ABB/ASV at three consecutive compensation time points, t_{i-1}, t_i, and t_{i+1}, showing the delay of the compensated circuit as a function of time.

just before applying ABB/ASV at t_i, and this typically touches D_{spec}. Considering the cumulative temporal degradation at $t = t_{i-1}$, the impact of ABB/ASV applied at that time point, and the temporal degradation due to BTI over $[t_{i-1}, t_i]$, we have

$$\begin{aligned} D(t_{i-1}) &< D(t_{i-1}-) \leq D_{\text{spec}} \\ D(t_{i-1}) &< D(t_i-) \leq D_{\text{spec}} \end{aligned} \tag{6}$$

At every compensation time point t_{i-1}, the amount of adaptation required is dependent on the delay degradation up to the next compensation time point t_i, and follows the shape of the figure in Fig. 3. In this figure, if no compensation is applied to the circuit, the delay during the interval $[t_{i-1}, t_i]$ will be above D_{spec}. To ensure that the delay meets specifications during this period, we apply a compensation at time t_{i-1}, whose magnitude is determined by the following result.

Theorem 1: Under small perturbations to the threshold voltage due to aging, let $D(t)$ be the delay of the aged circuit at any time t, and assume that under a specific compensation, $D(t_i-) > D_{\text{spec}}$ just prior to compensation time t_i. To bring $D(t_i-)$ to be under the specification, the value of $D(t_{i-1})$ can be adjusted, through compensation, to

$$D'(t_{i-1}) = D_{\text{spec}} \left(\frac{D(t_{i-1})}{D(t_i-)} \right) \tag{7}$$

Proof: For a MOS device, $V_{th} \propto t^{\frac{1}{6}}$, where the proportionality constant is different for NMOS and PMOS transistors. If we consider the effect of aging from time t_{i-1} to t_i, for a specific transistor type,

$$\frac{V_{th}(t_i)}{V_{th}(t_{i-1})} = \left(\frac{t_i}{t_{i-1}} \right)^{\frac{1}{6}} \tag{8}$$

Since the perturbations to V_{th} over this interval are small (by assumption), the delay of each gate can reasonably be assumed to vary linearly with V_{th}, as defined by a first-order Taylor series approximation. Therefore, the delay of each gate changes by a multiplicative factor, given by the right hand side of (8), implying that the delay of the circuit also changes by the same multiplicative factor. In other words, if the delay at time t_{i-1} is changed to $D'(t_{i-1})$

$$\frac{D'(t_i-)}{D'(t_{i-1})} = \frac{D(t_i-)}{D(t_{i-1})} \tag{9}$$

Since our goal is to set $D'(t_i-) = D_{\text{spec}}$, the result follows immediately. □

Algorithm 1 Adaptive Approach: Enumeration

1: Determine the nominal ($t_0 = 0$) delay and power (active and leakage). By assumption, $D(t_0-) \leq D_{\text{spec}}$.
2: **for** $t_i = t_0, \ldots, t_{n-1}$ **do**
3: Set $V_{dd}(t_i) = V_{dd}(t_{i-1})$ if $i > 0$; else set $V_{dd}(t_i)$ to the nominal V_{dd} value.
4: **repeat**
5: Set $V = V_{dd}(t_i)$.
6: Compute $\Delta V_{th}[t_i, t_{i+1}]$ due to BTI assuming that in this interval, $V_{dd} = V$, and determine $V_{th}(t_{i+1})$.
7: Using static timing analysis (STA), determine $D(t_{i+1}-)$, the delay due to BTI just prior to time t_{i+1}.
8: {*Use $D(t_{i+1}-)$ to determine the target delay at t_i, upon the application of ABB/ASV.*}
9: Set $D(t_i) = D_{\text{spec}} \left(\frac{D(t_i)}{D(t_{i+1}-)} \right)$.
10: {*Determine ABB/ASV values to be applied at time t_i to meet $D(t_i)$.*}
11: Use an enumeration scheme, similar to [21], to solve the optimization problem in (4), i.e., determine (v_{bn}, v_{bp}, V_{dd}) for the interval $[t_i, t_{i+1}]$, such that the delay requirement $D(t_i)$ is met, and power is minimized.
12: **until** $(V_{dd}(t_i) == V)$
13: {*Compute the delay at the end of the interval.*}
14: Compute $D(t_{i+1}-) = D(t_i)$ + temporal degradation over [t_i,t_{i+1}]. At this point, $D(t_{i+1}-) \leq D_{\text{spec}}$.
15: **end for**
16: Return the (v_{bn}, v_{bp}, V_{dd}) tuples at all times $t_0, \ldots, t_{n-1}$.

The adaptive strategy is developed at design time using the scheme shown in Algorithm 1, which shows the pseudocode for computing the optimal ABB/ASV values as the circuit ages. The algorithm begins by determining the amount of ABB/ASV that must be applied at the beginning of the lifetime of the circuit (after burn-in, testing, and binning), denoted by $t_0 = 0$, to compensate for aging until the first time t_1. This can be computed by determining the amount of change in the threshold voltage until t_1 (denoted as $\Delta V_{th}[t_0, t_1]$), and performing an STA run, to determine $D(t_1)$, as shown on lines 6-7 of the algorithm. The target delay after applying ABB/ASV is then computed, as shown on line 9, by applying the scaling factor from Theorem 1 to D_{spec}. As expected, $D(t_0) < D_{\text{spec}}$. Line 11 uses an enumeration scheme, based on the method described in [21], to determine the optimal ABB/ASV that must be applied at time t_0. Line 14 computes the delay of the circuit just prior to time t_1, i.e., $D(t_1-)$, which is less than D_{spec}. The method is repeated for successive values of t_i, and the look-up table entries are computed.

It should be pointed out that there is a second-order dependence between the level of V_{th} degradation and V_{dd} [18]. The value of V_{dd} in the solution at t_i depends on the delay degradation over $[t_i, t_{i+1}]$, which in turn depends on the degradation in V_{th} during this interval, which is a function of the V_{dd} value at time t_i. Hence, an iterative approach is employed, as illustrated by the repeat loop.

The choice of the compensation time points depends on several factors. While we would like to continuously apply the requisite amount of compensation at all times, so as to just meet the performance constraints while minimizing the power overhead, in practice, the circuit can only be compensated at a finite number of time points, n. The number of compensating times chosen, (i.e., the size of the look-up table) and their specific values is limited by the following factors:

- *The resolution in generating the body-bias and supply voltages*: A large number of body bias and supply voltages require a sophisticated network of voltage generators and dividers, adding to the area and power overheads.
- *The minimum change in delay over* $[t_i, t_{i+1}]$*, subject to modeling errors*: Since the delay model has some inaccuracies, a control system with a large number of compensatory points, where the delay over a pair of such successive times changes very marginally, may lead to inaccurate computations, due to modeling errors[3].
- *The resolution of mapping each delay to a unique* (v_{bn}, v_{bp}, V_{dd}) *tuple*: Since there is a fixed discretization in the values of each element of this tuple, each compensation step will reduce the delay by a quantum, and finer-grained delay compensation is not possible.

Section VI-C explores the impact of the number of compensating points chosen on the temporal profiles of the delay and power of the circuit.

V. Implementation of the Hybrid Approach

While the adaptive framework provides considerable savings in area as compared with synthesis, the power overhead over the original circuit can still be appreciably large, as will be shown in Section VI. This is due to the fact that the reduction in delay through FBB is obtained at the expense of an exponential increase in leakage power, as seen in Fig. 2, while an improvement in performance through ASV also results in an exponential increase in leakage power ($I_{\text{off}} \propto e^{V_{dd}}$), as well as a quadratic increase in active power.

On the other hand, technology mapping can map the circuit to use gates with different functionalities and/or drive strengths. The use of this technique has empirically been seen to provide significant performance gains with low area and power overheads, for reasonable delay specifications. Hence, a combination of this synthesis technique with ABB has the potential to provide improved results.

Accordingly, we propose a hybrid approach to design reliable circuits. An iterative approach is followed during design, alternating between the ABB assignment and technology mapping phases, to ensure that the final design is reliable, and has minimal power and area overheads. The algorithm consists of two distinct phases, namely the adaptive compensation phase involving an optimization formulation subject to power constraints, and the resynthesis phase, involving technology mapping to meet a tighter design specification. Algorithm 2 describes the steps involved in this approach.

Algorithm 2 Hybrid Approach - Iterative Adaptive Compensation and Technology Mapping

1: Determine the original delay and leakage power at $t_0 = 0$. By assumption, $D(t_0-) \leq D_{\text{spec}}$.
2: Assume the leakage power at $t = 0$ to be the leakage budget during adaptive compensation phase of optimization.
3: {***Adaptive Phase***}
4: **for** $t_i = t_0, \ldots, t_{n-1}$ **do**
5: Compute $\Delta V_{th}[t_i, t_{i+1}]$ due to BTI at the nominal V_{dd}, and determine $V_{th}(t_{i+1})$.
6: Perform STA to determine the delay, $D(t_{i+1}-)$, due to BTI, just prior to time t_{i+1}, and determine the leakage power, $P_{\text{lkg}}(t_{i+1})$.
7: Use an enumeration scheme, similar to [21], to solve the optimization formulation in (5), i.e., to determine (v_{bn}, v_{bp}) so as to minimize the delay, $D(t_i)$, while staying within the leakage budget from line 2.
8: Determine the delay before applying FBB at the next time point, i.e., $D(t_{i+1}-)$.
9: **end for**
10: **if** all delays are $\leq D_{\text{spec}}$ **then**
11: The optimization has converged; output the computed FBB values to the look-up table.
12: **else**
13: {***Resynthesis Phase***}
14: Identify the highest $D(t_i-)$ and set $D_{\text{spec}} = D_{\text{spec}} \times \left[\frac{D_{\text{spec}}}{\max_{i=1}^{n} D(t_i-)}\right]$ to reduce D_{spec}.
15: {*Tighten the delay specification for synthesis to ensure that after aging and subsequent adaptive compensation,* $D(0 \leq t \leq t_{life}) \leq D_{spec}$.}
16: Perform technology mapping to resynthesize the circuit under the tighter delay specification, at $t = 0$.
17: If leakage power of this new circuit at $t = 0$ is greater than the original leakage budget computed in line 2, increase the budget accordingly.
18: Repeat from line 2.
19: **end if**

The algorithm begins with the adaptive compensation phase, where the ABB optimization formulation from (5) is solved. Lines 3-9 modify the framework of Algorithm 1 to compute the optimal (v_{bn}, v_{bp}) values at different time points, instead of the optimal (v_{bn}, v_{bp}, V_{dd}) tuple, such that the delay is minimized without violating the leakage power constraints. If the delay of the circuit throughout its lifetime[4] is less than the specification D_{spec}, then the optimization ends and the optimal (v_{bn}, v_{bp}) entries are used to populate the look-up table, as shown in lines 10-11.

[3]In this work, we select our times such that the delay changes by at least 1% in each interval.

[4]Practically, this involves checking the values of $D(t_i)$ only at each of the compensation times, t_i.

TABLE I

LOOK-UP TABLE ENTRIES FOR THE LGSYNTH93 BENCHMARK, "DES," USING THE ADAPTIVE AND HYBRID APPROACHES

	Adaptive Approach						Hybrid Approach				
Time $\times 10^8$s	v_{bn} (mV)	v_{bp} (mV)	V_{dd} (V)	Delay (ps)	P_{act} (μW)	P_{lkg} (μW)	v_{bn} (mV)	v_{bp} (mV)	Delay (ps)	P_{act} (μW)	P_{lkg} (μW)
Nominal	0	0	1.00	355	641	327	0	0	355	641	327
0.0000	**0**	**50**	**1.03**	341	680	416	**0**	**0**	330	643	333
0.0001	**0**	**50**	**1.03**	341	680	346	**50**	**50**	334	643	332
0.0004	**0**	**100**	**1.03**	351	680	362	**0**	**100**	337	643	320
0.0016	**50**	**100**	**1.03**	351	680	369	**0**	**150**	338	643	333
0.0035	**0**	**50**	**1.06**	352	721	344	**0**	**150**	340	643	320
0.0080	**50**	**50**	**1.06**	351	721	357	**50**	**150**	339	643	329
0.0180	**50**	**100**	**1.06**	351	721	368	**50**	**150**	342	643	312
0.0400	**100**	**100**	**1.06**	352	721	377	**50**	**200**	343	643	328
0.0600	**0**	**100**	**1.09**	351	762	353	**50**	**200**	345	643	318
0.1100	**50**	**100**	**1.09**	351	762	360	**100**	**200**	343	643	326
0.1700	**100**	**200**	**1.06**	352	720	398	**100**	**200**	345	643	322
0.2500	**50**	**150**	**1.09**	352	762	362	**150**	**200**	343	643	328
0.3600	**50**	**200**	**1.09**	351	762	388	**150**	**200**	346	643	316
0.5500	**100**	**200**	**1.09**	351	762	396	**100**	**250**	351	643	325
0.7500	**50**	**150**	**1.12**	352	804	359	**100**	**250**	353	643	314
1.0000				355	804	350			355	643	305

However, if the delays are higher than D_{spec}, the circuit is technology-mapped to tighter design constraints, as shown in line 16. As a first order measure, the specification of the circuit is lowered from D_{spec} to $D_{\text{spec}} \times \frac{D_{\text{spec}}}{\max_{i=1}^{n} D(t_i-)}$, where $\max_{i=1}^{n} D(t_i-)$ is the maximum delay of the circuit over its lifetime, under the adaptive compensation scheme. If the leakage power of the circuit exceeds its budget value, the nominal value of the leakage power is updated, and this new value is used in (5) for $P_{\text{lkg}}(t = 0)$, as shown in line 17, and adaptive compensation is now repeated on this modified circuit. The process of adaptive compensation (lines 3-9) and technology-mapping for a tighter target delay (lines 13-16) is performed in an iterative manner, until the circuit delays converge, and the timing specifications are met at all times. In practice, only a few iterations are necessary before the delay converges, as seen from our experiments.

As we will demonstrate shortly, our experimental results indicate that this approach provides savings in area as compared with the synthesis approach, and dissipates lower power in comparison with the adaptive approach.

VI. EXPERIMENTAL RESULTS

We now present the results of applying our compensation scheme to circuits in the ISCAS85, LGSYNTH93, and ITC99 benchmark suites, synthesized on a 45nm [16]-based library. The body bias voltage is altered in increments of 50mV, while increments of 30mV are used for the supply voltage.

A. Results on a Sample Benchmark Circuit

We present detailed experimental results on a representative LGSYNTH93 benchmark, "des," whose delay and leakage variations under BTI, without ABB/ASV compensation, were shown in Fig. 1.

1) Look-up Table Entries: Table I shows the entries of the look-up table that encodes the compensation scheme, and the delay, active, and leakage power numbers for the adaptive and the hybrid approaches. The circuit is compensated at different times, as shown in the first column of Table I, up to its t_{life} of 10^8s. The time-entries in the look-up table are chosen such that the increase in delay over any successive time-interval is uniform, and that the circuit is uniformly compensated for degradation, over its entire lifetime. A large starting value of t = 10^4s, is chosen for the adaptive approaches, since the BTI model for estimating the delay degradation of the circuit in Algorithm 1 is asymptotically accurate. Further discussion on the optimality of the selection of the number of time-stamps (n) to compensate the circuit, and its impact on the temporal delay-power curves is deferred to Section VI-C.

The remaining columns of Table I show the details of the compensation scheme. Columns 2-7 correspond to the adaptive approach, and show, for each compensation time, the (v_{bn}, v_{bp}, V_{dd}) tuples computed by Algorithm 1, the final delay after applying ABB/ASV, and the active and leakage power values. Columns 8-12 show the results for the hybrid approach and display, respectively, the optimal (v_{bn}, v_{bp}) pair, and the delay, active power, and leakage power at each compensation time. The first four columns of the table, (bold-faced, with a gray background), denote the actual entries that would be encoded into the look-up table for the adaptive approach, while the first, eighth, and the ninth columns denote the entries of the look-up table for the hybrid approach. The column, "Delay," denotes the delay of the circuit $D(t_i)$, at the given compensation time t_i, immediately after applying ABB/ASV values from the table.

The results indicate that the target delay is met at all time points, up to t_{life} = 10^8s, using both the approaches. The amount of compensation increases with time, as the circuit degrades due to BTI. With the adaptive approach, which optimizes the power under fixed delay constraints, a combination of ABB and ASV is used to counter the effects of aging, on the original design, whose delay and power values are shown in the row labeled "Nominal." The active and leakage power values vary as a function of time, depending on the optimal solution chosen at each time point. As explained in Fig. 3, the circuit is compensated for aging right from the

first time period [0,t_1], by applying ABB/ASV at time $t = 0$. Hence, the delay of the circuit at $t = 0$ in the look-up table is less than D_{spec}. The leakage power decreases temporally due to increase in V_{th} caused by BTI, but increases with ABB/ASV, and in our scheme, it is seen to exceed the nominal leakage. For the hybrid approach, which uses a combination of ABB and synthesis, the circuit at $t = 0$ achieves its delay reduction purely through synthesis. It can be seen that the area overhead of synthesis in this case is low: as compared to the nominal case, the active power increases by 0.3% and the leakage power by 1.8%. The results indicate that the power numbers using the hybrid approach are significantly lower than that using the adaptive approach.

2) Comparison of Transient Power and Delay Numbers: The temporal variation in the delay of "des" is shown in Fig. 4. The delay of the circuit, as a function of time, is shown for

- the adaptive method from Section II-C, where the delay can be seen to always be close to D_{spec}.
- the synthesis-based method from [6], where worst-case BTI-based library gate delays were used during technology mapping to synthesize the circuit: in this case, the delay increases monotonically with time.
- the fixed power case, corresponding to the results of solving the optimization problem in (5), where the delay is minimized through ABB under a power budget, set to the power at $t = 0$: this curve does not satisfy the delay specification.
- the hybrid method from Section II-D, which satisfies the delay specification throughout its lifetime, and essentially corresponds to finding a power specification for a fixed power curve that meets D_{spec} at the end of the circuit lifetime. In this case, the power specification implies that the circuit is mapped to meet a delay specification of 330ps at $t = 0$.

All methods were targeted to meet the same delay specification, $D_{\text{spec}} = 355$ps, throughout the circuit lifetime and this value is shown by a horizontal line in the figure. This delay corresponds to the nominal delay of the original circuit at $t = 0$.

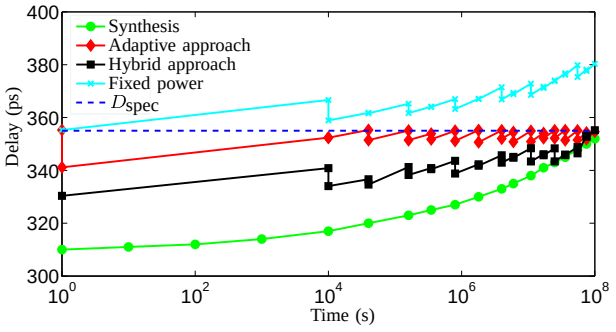
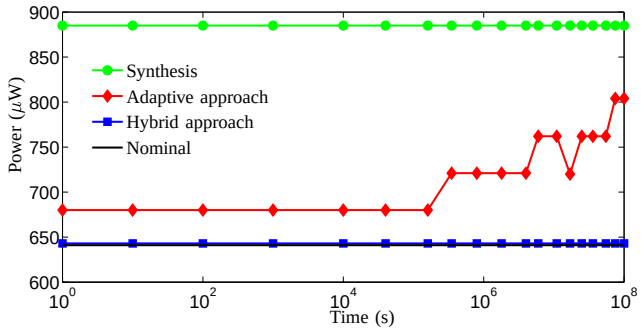
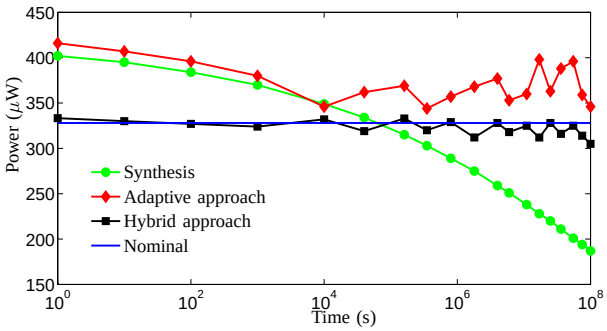


Fig. 4. Temporal delay of benchmark "des" using different approaches.

Figs. 5(a) and (b), respectively, compare the values of the active and leakage power for the three approaches (adaptive, hybrid, and synthesis). The horizontal line marked "Nominal" represents the power dissipation of the original circuit at $t = 0$. Since the synthesis approach performs technology mapping for a tighter delay specification at birth, leading to a large area, as compared with the nominal design, the active power for the synthesis approach is constant over the lifetime of the circuit. For the adaptive approach, the supply voltage generally (but not always) increases gradually with time, as shown in Table I. Correspondingly, the active power increases almost monotonically, as shown in Fig. 5(a). One exception to the monotonicity of V_{dd}, as seen from Table I, is at $t = 0.17\times10^8$s, where the optimal (v_{bn}, v_{bp}, V_{dd}) tuple leads to a decrease in V_{dd} accompanied by a larger increase in (v_{bn}, v_{bp}), with respect to the solution at the previous time point, hence causing the active power to decrease temporally, as seen in the figure. The figure indicates that the maximum active power dissipated using the adaptive approach is less than that for the synthesis-based design.



(a) Active power



(b) Leakage power

Fig. 5. Temporal active and leakage power values of "des" using the various approaches.

Similarly, Fig. 5(b) compares the leakage of the various approaches over the lifetime of the circuit, with respect to its nominal value. The leakage of the synthesis-based circuit is highest at $t = 0$ (when there is no BTI), but monotonically decreases with time. In contrast, the adaptive approach tries to adaptively recover performance, at the expense of increased power. The corresponding overhead implies that the leakage power for this method increases beyond its nominal value. Note that the leakage for the adaptive circuit at $t = 0$

TABLE II

AREA AND POWER OVERHEAD COMPARISON FOR ADAPTIVE AND HYBRID APPROACHES

	Original design				Adaptive approach		Hybrid approach					Synthesis		
Bench-mark	D_{spec}	$\frac{\Delta D(t_{\text{life}})}{D_{\text{spec}}}$	Nominal ($t = 0$s)		Overhead		Delay		Overhead			Overhead		
	(ps)	%	P_{lkg_0} (μW)	P_{act_0} (μW)	Max P_{lkg} %	Max P_{act} %	$(t = 0)$s (ps)	Reduction	Area %	Max P_{lkg} %	Max P_{act} %	Area %	Max P_{lkg} %	Max P_{act} %
b14	1078	14%	426	775	14%	26%	1027	0.95	2%	2%	1%	19%	16%	17%
b15	902	13%	781	1384	26%	19%	839	0.93	4%	4%	4%	16%	15%	18%
b17	1255	15%	3242	1790	26%	23%	1177	0.94	1%	1%	1%	22%	26%	22%
b20	1125	16%	1745	919	19%	31%	1058	0.94	2%	2%	3%	17%	18%	17%
C2670	510	15%	52	94	29%	26%	487	0.96	2%	3%	2%	32%	15%	19%
C3540	769	14%	74	136	30%	25%	724	0.95	2%	3%	4%	32%	38%	37%
C5315	729	15%	114	208	29%	19%	697	0.96	1%	1%	1%	14%	25%	18%
C6288	2190	14%	264	182	13%	28%	2110	0.96	7%	8%	5%	57%	63%	48%
C7552	616	15%	190	337	29%	19%	592	0.96	1%	1%	1%	18%	15%	19%
dalu	560	12%	227	127	12%	28%	535	0.96	1%	1%	1%	19%	19%	27%
des	355	15%	327	641	27%	25%	332	0.93	1%	3%	2%	35%	28%	38%
i8	840	17%	157	305	26%	25%	789	0.94	1%	3%	3%	18%	44%	71%
i10	830	14%	152	307	32%	25%	787	0.95	2%	2%	3%	21%	28%	26%
t481	368	14%	201	572	16%	26%	345	0.94	2%	3%	3%	38%	30%	39%
Average		15%			23%	25%		0.95	2%	3%	2%	26%	27%	30%

is also greater than the nominal value, since some amount of ABB/ASV is applied to the circuit to guardband against temporal degradation during $[0, t_1]$, as shown in Fig. 3. The maximum leakage power (at $t = 0$) at any time point using our approach is almost identical to that using the synthesis method, as seen from Fig. 5(b).

For the hybrid approach, which uses a combination of synthesis and adaptive compensation, the results provide improvements over these two methods, used separately. As shown in Fig. 5(a) and (b), respectively, the active and leakage power at $t = 0$ increase very minimally (by less than 2%), as compared with the corresponding values for the original circuit, due to an increase in the area of the circuit during resynthesis. Subsequent adaptive compensation over the lifetime of the circuit is performed under fixed power constraints to ensure that the power never exceeds its value at $t = 0$. Hence, the curves for the overall leakage and active power, as functions of time, are closest to their corresponding budgets.

B. Area and Power Trade-offs

In this section, we compare the trade-offs in area and power for various approaches proposed in this paper, for the five largest benchmark circuits from ISCAS85 and LGSYNTH93 suites, as well as some large ITC99 benchmarks. Table II presents the area savings and the maximal power overhead of the adaptive and hybrid approaches, in comparison with the synthesis method. The column D_{spec}, is the delay of the original circuit at $t = 0$. The active (denoted as P_{act_0}) and leakage (tabulated as P_{lkg_0}) power values of the uncompensated circuit shown in the table denote their maximal numbers over the lifetime operation of the circuit. The column $\frac{\Delta D(t_{\text{life}})}{D_{\text{spec}}}$ denotes the percentage increase in delay due to maximal BTI after t_{life} (10^8s) seconds of stress. The percentage increase in the **maximum** leakage and active power values dissipated over the time interval [0,t_{life}], and the overhead in area, over the original design, are shown in the table, for the three approaches.

Table II indicates that the synthesis approach has a large average area overhead of 26%. However, the area overhead of the adaptive approach is restricted to the look-up tables, voltage generators for the additional supply voltages, and the body-bias voltages, and is therefore significantly smaller. The work in [7] has shown that this overhead is within 2-3% of the area of the original design. Thus, the adaptive approach provides significant area savings as compared with synthesis.

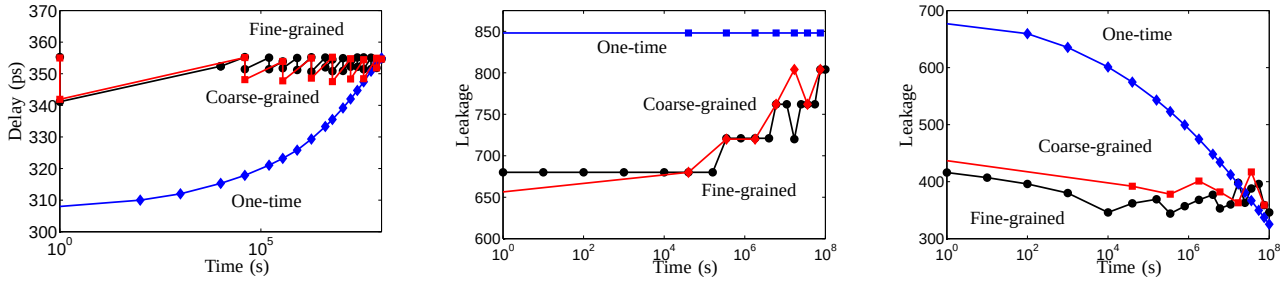
During optimization using the hybrid approach, the resynthesis (technology mapping) phase causes an increase in the area of the circuit, since the circuit is remapped to tighter specifications. The column "Reduction" in Table II indicates that using the hybrid approach, the target delay (at $t = 0$) during the technology mapping phase is only 5% lower than the nominal delay of the circuit, whereas the target delay (at $t = 0$) using BTI-aware synthesis is $\approx$ 15% less than the nominal delay. Expectedly, this small decrease in delay of $\approx$ 5% can be obtained with a marginal penalty in area (average value of the order of around 2%) for most circuits[5]. Hence, this overhead in area is extremely small, particularly when compared with that using synthesis.

The power numbers shown in the table indicate that while the adaptive and synthesis approaches have large power overheads, the power overhead using the hybrid approach is extremely small, with an average increase in active and leakage powers of the order of around 2-3%, over the wide range of benchmarks tested. Thus, by combining the advantages of adaptive compensation and BTI-aware synthesis, we obtain an optimal final design whose area overhead is lower than that of the synthesis based approach, while the power overhead is lower than that of the adaptive approach, for the same delay specifications.

C. Optimal Selection of Look-up Table Entries

For the adaptive approach, the size of the look-up table (i.e., the number of entries) can be chosen according to

[5]In reality, the area overhead is slightly higher, since the overhead in creating wells for body biasing, the look-up tables, and additional control circuitry must be considered. Nevertheless, the area overhead is still lower than that using synthesis.



(a) Delay (b) Active power (c) Leakage power

Fig. 6. Temporal delay, active and leakage power of "des" for the adaptive approach, showing the impact of the number of entries in the look-up table.

various criteria, as discussed in Section IV. In this section, we investigate the impact of the size of the look-up table on the power and delay of the compensated circuit, using the adaptive approach. Accordingly, we perform simulations where the circuit is compensated at eight time points, instead of the 15 times chosen in Table I. The compensation time points correspond to alternate entries from the look-up table in Table I, and the corresponding (v_{bn}, v_{bp}, V_{dd}) tuples, found using Algorithm 1, are shown in Table III.

TABLE III
LOOK-UP TABLE ENTRIES FOR "DES" USING A COARSE-GRAINED ADAPTIVE COMPENSATION WITH FEWER TIME ENTRIES

Time $\times 10^8$s	v_{bn} (mV)	v_{bp} (mV)	V_{dd} (V)	Delay (ps)	P_{act} (μW)	P_{lkg} (μW)
0.0000	**50**	**150**	**1.00**	342	641	466
0.0004	**50**	**100**	**1.03**	348	680	391
0.0035	**0**	**100**	**1.06**	348	720	378
0.0180	**100**	**100**	**1.06**	349	720	401
0.0600	**50**	**100**	**1.09**	348	762	381
0.1700	**0**	**100**	**1.12**	348	804	363
0.3600	**100**	**200**	**1.09**	348	762	417
0.7500	**50**	**150**	**1.12**	352	804	358
1.0000				355	804	340

As expected, the results indicate that the delay of the circuit is still met at all times, but the optimal (v_{bn}, v_{bp}, V_{dd}) tuples, and the corresponding delay and power values, are different from the corresponding values in Table I. We compare these values by plotting the delay and power as functions of time in Fig. 6. We refer to the adaptive approach with 15 entries in the look-up table as the "Fine-grained" method, and that with eight entries as the "Coarse-grained" method. We also consider an extreme coarse-grained approach, where ABB/ASV is only applied at $t = 0$, to ensure that the circuit meets its delay specifications over its lifetime: this can be considered as a look-up table with only one entry, at $t = 0$, and is referred to as the "One-time" approach. Fig. 6(a) shows the delays for all three of these approaches as a function of time.

By design, all methods meet the delay specification over the circuit lifetime, but as the granularity becomes coarser, the variation in circuit delay over time becomes larger, since the incremental delay degradation in each interval is higher, requiring larger changes to the (v_{bn}, v_{bp}, V_{dd}) tuple at each compensation time point, leading to larger swings for $D(t_i)$ below D_{spec}.

The active and leakage power profiles for the three cases are shown in Fig. 6(b), and Fig. 6(c), respectively. These trends show that the peak power dissipation of the circuit over its lifetime, for both the active and leakage power, increase as the granularity becomes coarser. The fine-grained approach used in our work, with 15 compensation time points, therefore satisfies the requirements laid out in Section IV, while maintaining a small overhead in terms of the circuitry required for its implementation.

VII. CONCLUSION

BTI has become an important reliability concern in circuit design. Previous solutions in the presilicon design stage aimed at guaranteeing reliable circuit performance can lead to large area and associated power overheads. An adaptive approach that determines the temporal degradation of the circuit, and compensates for it, through adaptive body biasing (ABB) and adaptive supply voltage (ASV) has been proposed in this work. The results indicate that by combining the adaptive and synthesis approaches, circuits can be efficiently guardbanded over their lifetime, with a minimal overhead in area, and a small increase in power, as compared with a circuit designed only to meet the nominal specifications. Further, techniques such as those in [21] may be used to apply ABB/ASV to simultaneously counter the impact of aging, as well as process and temperature variations.

REFERENCES

[1] S. Zafar, Y. Kim, V. Narayanan, C. Cabral, V. Paruchuri, B. Doris, J. Stathis, A. Callegari, and M. Chudzik, "A Comparative Study of NBTI and PBTI Charge Trapping in SiO_2-HfO_2 Stacks with FUSI, TiN, Re Gates," in *Proc. Symposium on VLSI Technology*, pp. 23–25, October 2006.

[2] M. F. Li, G. Chen, C. Shen, X. P. Wang, H. Y. Yu, Y.-C. Yeo, and D. L. Kwong, "Dynamic Bias Temperature Instability in Ulthathin SiO_2 and HfO_2 Metal Oxide Semiconductor Field Effect Transistors and its Impact on Device Lifetime," *Japanese Journal of Applied Physics*, vol. 43, pp. 7807–7814, November 2004.

[3] F. Crupi, C. Pace, G. Cocorullo, G. Groeseneken, M. Aoulaiche, and M. Houssa, "Positive Bias Temperature Instability in nMOSFETs with Ultra-Thin Hf-silicate Gate Dielectrics," *Journal of Microelectronic Engineering*, vol. 80, pp. 130–133, June 2005.

[4] B. C. Paul, K. Kang, H. Kufluoglu, M. A. Alam, and K. Roy, "Temporal Performance Degradation under NBTI: Estimation and Design for Improved Reliability of Nanoscale Circuits," in *Proceedings of the Design, Automation & Test in Europe*, pp. 1–6, March 2006.

[5] K. Kang, S. Gangwal, and K. Roy, "NBTI Induced Performance Degradation in Logic and Memory Circuits: How Effectively can we Approach a Reliability Solution," in *Proceedings of the Asia-South Pacific Design Automation Conference*, pp. 726–731, January 2008.

[6] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "NBTI-Aware Synthesis of Digital Circuits," in *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 370–375, June 2007.

[7] J. W. Tschanz, J. T. Kao, S. G. Narendra, R. Nair, D. A. Antoniadis, A. P. Chandrakasan, and V. De, "Adaptive Body Bias for Reducing Impacts of Die-to-Die and Within-Die Parameter Variations on Microprocessor Frequency and Leakage," *IEEE Journal of Solid-State Circuits*, vol. 37, pp. 1396–1402, November 2002.

[8] A. Keshavarazi, S. Narendra, B. Bloechel, S. Borkar, and V. De, "Forward Body Bias for Microprocessors in 130-nm Technology Generation and Beyond," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 696–701, May 2003.

[9] J. W. Tschanz, N. S. Kim, S. Dighe, J. Howard, G. Ruhl, S. Vanga, S. Narendra, Y. Hoskote, H. Wilson, C. Lam, M. Shuman, C. Tokunga, S. Somasekhar, S. Tang, D. Finan, T. K. an N. Borkar, N. Kurd, and V. De, "Adaptive Frequency and Biasing Techniques for Tolerance to Dynamic Temperature-Voltage Variations and Aging," in *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 292–294, February 2007.

[10] K. O. Jeppson and C. M. Svensson, "Negative Bias Stress of MOS Devices at High Electric Fields and Degradation of MNOS Devices," *Journal of Applied Physics*, vol. 48, pp. 2004–2014, May 1977.

[11] M. A. Alam, "A Critical Examination of the Mechanics of Dynamic NBTI for pMOSFETs," in *IEEE International Electronic Devices Meeting*, pp. 14.4.1–14.4.4, December 2003.

[12] S. Bhardwaj, W. Wang, R. Vattikonda, Y. Cao, and S. Vrudhula, "Predictive Modeling of the NBTI Effect for Reliable Design," in *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 189–192, September 2006.

[13] A. E. Islam, H. Kufluoglu, D. Varghese, S. Mahapatra, and M. A. Alam, "Recent Issues in Negative Bias Temperature Instability: Initial Degradation, Field Dependence of Interface Trap Generation, Hole Trapping Effects, and Relaxation," *IEEE Transactions on Electron Devices*, vol. 54, pp. 2143–2154, September 2007.

[14] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "An Analytical Model for Negative Bias Temperature Instability (NBTI)," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 493–496, November 2006.

[15] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "A Finite Oxide Thickness Based Model for Negative Bias Temperature Instability," *Submitted to IEEE Transactions on Devices and Materials Reliability*.

[16] "Predictive Technology Model." Available at `http://www.eas.asu.edu/~ptm`.

[17] J. W. Tschanz, S. G. Narendra, R. Nair, and V. De, "Effectiveness of Adaptive Supply Voltage and Body Bias for Reducing Impact of Parameter Variations in Low Power and High Performance Microprocessors," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 826–829, May 2003.

[18] R. Vattikonda, W. Wang, and Y. Cao, "Modeling and Minimization of PMOS NBTI Effect for Robust Nanometer Design," in *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 1047–1052, July 2006.

[19] J. Keane, T.-H. Kim, and C. H. Kim, "An On-Chip NBTI Sensor for Measuring PMOS Threshold Voltage Degradation," in *Proceedings of the ACM International Symposium on Low Power Electronics and Design*, pp. 189–194, August 2007.

[20] T.-H. Kim, J. Liu, R. Persaud, and C. H. Kim, "Silicon Odometer: An On-Chip Reliability Monitor for Measuring Frequency Degradation of Digital Circuits," *IEEE Journal of Solid-State Circuits*, vol. 43, pp. 874–880, April 2008.

[21] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "Body Bias Voltage Computations for Process and Temperature Compensation," *IEEE Transactions on VLSI Systems*, vol. 16, pp. 249–262, March 2008.