

Layout-Dependent Aging Mitigation for Critical Path Timing

Che-Lun Hsu¹, Shaofeng Guo², Yibo Lin¹, Xiaoqing Xu¹, Meng Li¹, Runsheng Wang², Ru Huang², and David Z. Pan¹

¹ECE Department, University of Texas at Austin, Austin, TX, USA

²Institute of Microelectronics, Peking University, Haidian, Beijing, China

{chsu1, yibolin, xiaoqingxu, mengli, dpan}@cerc.utexas.edu, {r.wang, ruhuang}@pku.edu.cn

Abstract— Layout-dependent effects (LDEs) are becoming increasingly important as technology node continues to shrink into the regime of FinFET transistors. Prior LDE studies mainly focus on accurate transistor modeling and fast circuit performance evaluations at the early lifetime of a design. Few studies have been performed on the layout dependency of the circuit aging towards the end of life (EOL). This study demonstrates that, due to transistor-level layout-dependent aging (LDA) behaviors, circuit-level timing degradations are greatly impacted by layout configurations, including length of diffusion and oxide spacing. In this paper, we propose the first circuit-level aging mitigation framework to improve the critical-path timing towards the EOL. Our framework features comprehensive LDA evaluations for standard cell timing, which shows that multiple-row height cells lead to worse EOL timing than single-row height cells due to length-of-diffusion effects. We further propose a min-cost-flow-based placement approach to concurrently allocate the oxide spacing among neighboring standard cells, which generates much better EOL timing than a conventional greedy approach. Experimental results demonstrate that under the concurrent approach in the proposed aging mitigation framework, the total and worst negative slacks for EOL timing are on average reduced by 42% and 25%, respectively.

I INTRODUCTION

Layout-dependent effects (LDEs), including stress and aging on CMOS transistors, are becoming increasingly important as technology node continues to shrink into the domain of FinFET transistors [1]. As transistor scaling continues, LDEs greatly contribute to the variations of transistor parameters, such as threshold voltage (V_{th}) and mobility (μ), which eventually penalizes circuit performance. The semiconductor industry is adapting to this trend and incorporating LDE parameters into transistor modeling [2]. Moreover, for digital circuit design, LDEs not only degrade the transistor parameters (V_{th} and μ) at the early lifetime [3], but also introduce complex layout-dependent aging (LDA) behaviors towards the end of lifetime (EOL) [4]. Ren et al. [4] and Liu et al. [5] have demonstrated that transistor aging behaviors, including Biased Temperature Instability (BTI) and Hot Carrier Injection (HCI), highly depend on the local layout configurations, which further induces layout-dependent aging (LDA) effects.

Prior LDE and aging studies mainly focus on accurate transistor modeling and circuit timing evaluations. Ndiaye et al. and Wang et al. [2, 6] demonstrate EOL reliability studies for V_{th} degradation at transistor level by applying stress over a period of time. Similar studies from transistor level to the standard cell

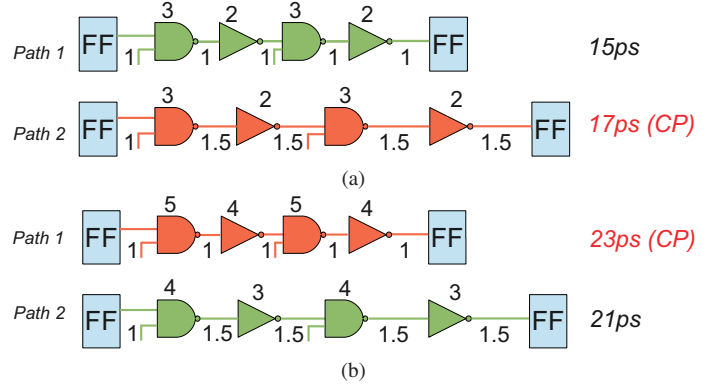


Fig. 1. (a) Early-life timing, (b) EOL timing. The numbers above gates and below interconnects are for gate and wire delay, respectively.

level can be found in [7, 8]. Firouzi [9] further presents a machine learning technique to predict timing paths that are prone to failure due to aging and variation. Faraji et al. [10] focuses on novel techniques to overcome aging performance degradation in SRAM cells. Fang et al. [11, 12] study efficient and accurate circuit-level timing analysis leveraging complex transistor aging mechanisms, such as BTI and HCI.

At the circuit level, LDA makes the reliability analysis and optimization much more complex than dealing with traditional LDEs or aging behaviors. LDA severely impacts the EOL circuit performance and is highly sensitive to local layout configurations. For circuit design, this further means EOL timing is hard to predict and optimize because a typical reliability analysis tool only guarantees timing closure at early lifetime, while assuming pessimistic guard-banding for EOL timing [13]. Fig. 1 compares early-life timing and EOL timing of one design. Fig. 1(a) illustrates two timing paths, where “Path 2” has larger total delay (17ps) than “Path 1” (15ps). This makes “Path 2” the critical path (CP) for early-life timing. However, for EOL timing in Fig. 1(b), “Path 1” becomes the timing CP (23ps) due to distinct path aging rates. The timing aging rate of “Path 1” is much larger than that of “Path 2”. The major reason is that the local layout configurations of the gates on “Path 1” are much more susceptible to aging than those on “Path 2” due to complex LDA behaviors [4, 5]. Although this significantly complicates EOL timing analysis and optimization, it also implies conventional guard-banding method could be overly pessimistic. Therefore, to obtain much better EOL timing, it becomes pivotal to invent novel optimization techniques (beyond pessimistic guard-banding) to miti-

gate the EOL CP timing while leveraging complex LDA behaviors.

In this study, we first embed the transistor LDA into standard cell timing evaluations, which generates comprehensive EOL timing data for each standard cell. Our study reveals that, due to the length of diffusion effects, the multiple-row height cells are more susceptible to LDA behaviors. We further identify that circuit placement plays an important role in the local layout configurations for LDA. Therefore, we propose novel placement techniques to appropriately allocate the whitespace among neighboring standard cells, which significantly improves the EOL timing. More importantly, the transistor-level modeling and validation have demonstrated that LDA has a similar dependency on local layout parameters as traditional LDEs [4]. As traditional LDEs only affect early-life timing, this similarity benefits the early-life timing closure after applying the proposed aging optimizations. While maintaining early life timing closure, we propose mitigation schemes to improve EOL timing on CPs of a design. Our major contributions are summarized as follows.

1. To the best of our knowledge, this is the first circuit-level mitigation framework using intelligent standard cell placement techniques for EOL timing optimization.
2. The LDA effects on heterogeneous-height cells are evaluated for the first time, which demonstrates inferior EOL timing compared to the single-row height cells due to the length-of-diffusion effect.
3. A min-cost flow-based cell placement technique is proposed to concurrently allocate spacing among cells on critical paths for the EOL timing optimization.

The rest of this paper is organized as follows: Section II explains different types of LDAs and problem definitions. Section III explains how we characterize the aging behaviors at the standard cell level and our analysis for aging impacts on multiple-row height cells. Section IV presents the optimization methods and the proposed framework for aging mitigation. Section V shows the experimental results. Section VI concludes the paper.

II PRELIMINARIES

A Layout Dependent Aging

Circuit-level aging is mainly affected by layout-dependent parameters, including *length of diffusion* (LOD), *metal boundary effect* (MBE), and *oxide-to-oxide spacing effect* (OSE) [4]. Fig. 2 illustrates the aforementioned parameters. LOD is defined as the distance between the edge of the gate and active area. OSE indicates the spacing between neighboring active areas. In Fig. 2, LOD_l is the LOD on the left and LOD_r is the LOD on the right. A similar naming convention applies to OSE. MBE is quantified by the proximity of Nwell/Pwell (N/P) boundary (SPM), i.e., the distance between active area and N/P boundary. This study focuses on the analysis and optimization of LOD and OSE while assuming SPM as a fixed value. This is because MBE is a fixed parameter under modern standard cell architecture, where the standard cell height is fixed and N/P boundary is set by design rule check (DRC). LOD and OSE can be adjusted more freely as long as the cell is on a legal placement site.

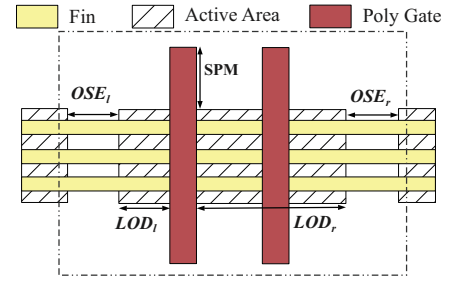


Fig. 2. LOD and OSE parameters.

LOD and OSE-related transistor degradation is due to the stress effects from *shallow trench isolation* (STI) [14, 15]. Based on the model in [4] and experimental silicon data in [5], we create a predictive LDA model for 7nm FinFET. The predictive LDA model considers LDA behaviors (BTI and HCI) for EOL timing and traditional LDEs for early-life timing. Fig. 3 shows the aging degradation in ΔV_{th} v.s layout parameters. When LOD and OSE decrease, the corresponding LOD and OSE aging effects increase, i.e., the amount of V_{th} degradation (ΔV_{th}) increases. PMOS and NMOS show the same trend as $|\Delta V_{th}|$ becomes larger when LOD/OSE is smaller. $|\Delta V_{th}|$ increases drastically especially in extremely-scaled transistor parameters ($LOD, OSE < 100nm$). Note that in Fig. 3, both LOD and OSE are the average of left and right side values of a transistor. For example, $LOD = (LOD_l + LOD_r)/2$ for the transistor in Fig. 2.

B End of Life Timing

LDA penalizes transistor performance continuously, where the induced degradation increases over the circuit lifetime. Aging phenomenon has to be addressed in circuit design stage in order to obtain a reliable system that functions across the lifetime. EOL timing optimization is challenging because LDA will not only degrade the early-life CPs, but paths that are not critical at early-life may become timing critical towards the EOL. This is because different timing paths experience different magnitude of LDA based on the LDA parameter profiles on the path as shown in Fig. 1. This study aims at EOL timing optimization (beyond the conventional pessimistic guard-banding scheme) while maintaining early-life timing closure.

C Problem Definition

The aging mitigation aims at optimizing EOL CP timing by remedying LDA impacts across the standard cell design and placement stages, which includes cell-level aging evaluation and circuit-level aging mitigation. Related problems are defined as follows.

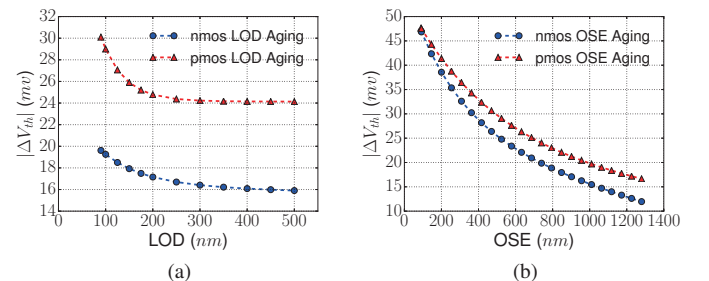


Fig. 3. Aging effects on EOL V_{th} , (a) LOD, (b) OSE.

Problem 1 (Cell-level Aging Evaluation).

Given a standard cell library, for each cell, extracts and enumerates LDA parameters for EOL timing characterization.

EOL timing characterization for each cell can be composed to an EOL timing library. With EOL timing library, we can obtain accurate EOL timing information for a design. This further enables circuit-level aging mitigation studies at a placement stage.

Problem 2 (Circuit-level Aging Mitigation).

Given a legal placement, early-life and EOL timing, improve EOL timing while maintaining early-life timing closure.

III LAYOUT DEPENDENT AGING EVALUATION

A Cell-Level Aging Simulation

Cell-level aging evaluation in Problem 1 requires a series of SPICE simulations. Given a standard cell layout, we first perform cell extraction and obtain LDA parameters. State-of-the-art aging model deriving from [4, 5, 16] enables aging SPICE simulations, which modified the device parameters such as V_{th} and μ based on extracted LDA parameters. The simulation with aging model outputs the EOL cell delay data for each cells which allows us to characterize cells and build the aging timing library.

From extensive SPICE simulations with various OSE values, we obtain a relationship between OSE and delay. Fig. 4 illustrates an example, where INV_X1 delay is obtained with different OSE values and linearly fitted to get a piecewise linear relationship between OSE and cell delay. We refer to the linear line fitted with OSE smaller than $600nm$ as *region 1* and the other as *region 2* where the OSE range starts from $600nm$ to $1.28\mu m$. We assume the impacts to delay can be ignored when the OSE goes beyond *region 2* where delay is flattened out. The threshold that divides *region 1* and *region 2* is derived based on approximation error. Currently, we limit the maximum difference between the piecewise linear fit and the simulation data to be less than $1ps$ and we are able to improve the accuracy of the piecewise linear approximation by increasing the number of linear regions.

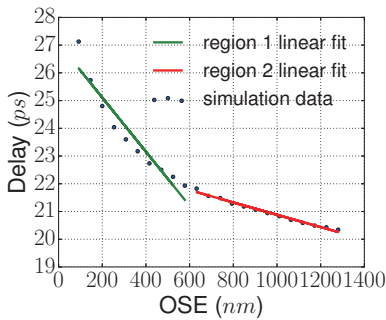


Fig. 4. The OSE dependence of the INV_X1 fall delay.

B Aging Impact on Multiple-Row Height Cells

Multiple-row height cells have gained importance due to standard cell tracks shrinkage [17–19]. Therefore, it is essential to incorporate multiple-row height cells in $7nm$ process studies. Multiple-row height cells have different layout configurations which translate to different LDA profiles. To understand the aging impact of multiple-row height cells, a set of dual-row height cells are created for cell-level LDA evaluations and comparisons. From the

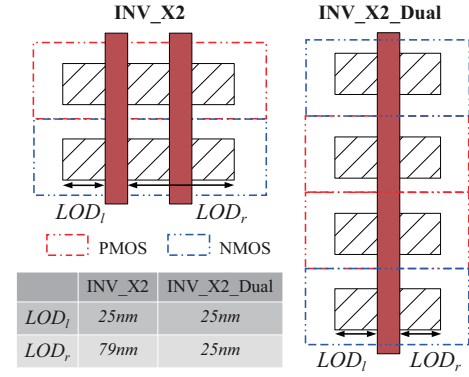


Fig. 5. Multiple-row height cell comparisons.

aging evaluation, we discover that multiple-row height cells are more susceptible to aging degradations.

Fig. 5 illustrates the layout for INV_X2 in single-row height and dual-row height configuration. The effective width in terms of fin count for one-finger transistor is 3 fins. Single-row height cell in Fig. 5 (INV_X2) is a two-finger transistor so the effective width is 6 fins. For the LOD values, the single-row height cell has $LOD_l = 25nm$ and $LOD_r = 79nm$. Dual-row height cell in Fig. 5 (INV_X2_Dual) is an one-finger transistor where the additional driving strength comes from the vertical extension of the cell height so that the effective width is also 6 fins. The LOD_l and LOD_r are both $25nm$ for dual-row height cell which means the average LOD is smaller than the LOD in single-row height configuration. This makes LOD-induced aging for the dual-row height cell larger than that for the single-row height cell as previously shown in Fig. 3a.

After evaluating cells with different LDA parameters, we achieve accurate information for LDA parameters vs cell EOL timing. The cell-level EOL timing can be used to guide the circuit-level aging mitigation.

IV AGING MITIGATION FOR CRITICAL PATH TIMING

This section presents the aging mitigation framework to solve Problem 2 in Section II-C. It has been mentioned in Section II-B, EOL CP timing needs to be emphasized since conventional timing library and static timing analysis (STA) engine handle only early-life timing. The aging timing library from Section III-A enables us to perform STA for EOL timing. In a given design with circuit placement, we identify a set of timing CPs and associated standard cells at the EOL. To mitigate LDA, we propose two strategies including cell replacement and cell spreading to optimize LDA parameters, i.e. LOD and OSE.

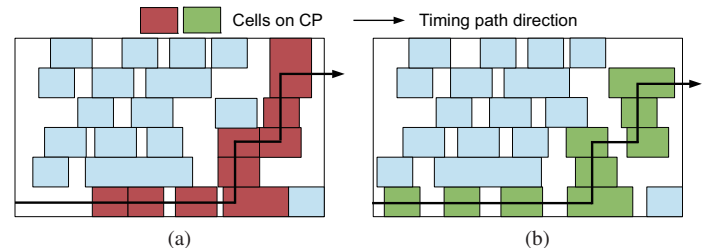


Fig. 6. EOL aging mitigation for CP timing, (a) initial placement, (b) aging mitigation with cell replacement and spreading.

A Cell Replacement

We assume an initial design is implemented with multiple-row height cells. The cell replacement strategy directly targets the multiple-row height cells. As shown in Section III-B, compared with single-row height cells of the same driving strength, multiple-row height cells are more vulnerable to aging. Therefore, replacing multiple-row height cells with single-row height cells on CP can mitigate the LDA effect on timing. For instance, in Fig. 6(a), one dual-row height cell exists on the CP of the initial placement. To improve EOL timing, the dual-row height cell is replaced with the single-row height cell as shown in Fig. 6(b). We adopt a greedy approach for the cell replacement if the width of the single-row cell can fit into the horizontal spaces surrounding the original dual-row height cell.

B Cell Spreading

The other aspect of LDA is OSE where dense oxide spacing causes significant aging degradations. Oxide spacing can be directly translated to cell spacing; therefore, cell spreading can achieve sparser oxide spacing and thereby reduce the OSE-induced aging. Using the piecewise linear relationship in Fig. 4, we propose two approaches for cell spreading.

Greedy Approach: This greedy method always optimizes the cell with the steepest OSE delay slope ($|\Delta\text{delay}/\Delta\text{OSE}|$). We first sort cells on CP according to their OSE delay slope in *region 1* in a descending order. Then for each cell, we spread out its left and right neighboring cells to increase its OSE as much as possible until the OSE goes beyond *region 2*. This approach is straightforward but the drawback is that spreading out adjacent cells will affect the OSE values of other cells as shown in Fig. 7, which might degrade the overall EOL timing.

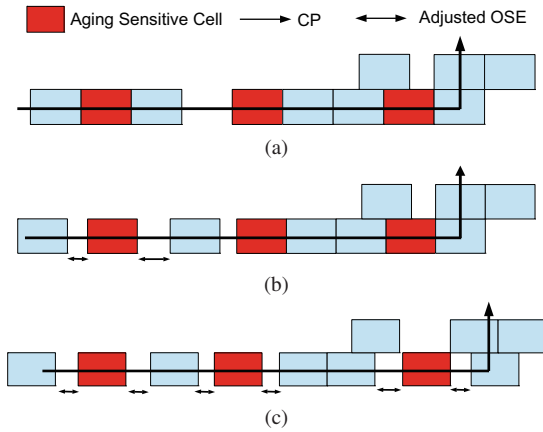


Fig. 7. (a) initial placement, (b) greedy cell spreading - one aging-sensitive cell movement affects other cells, (c) concurrent cell spreading - all aging-sensitive cells are moved concurrently.

Concurrent Approach: The greedy approach suffers from low quality of results when neighboring cells compete for OSE resources. Therefore, we propose a concurrent approach to optimize the EOL timing of cells within a row simultaneously. The mathematical formulation is shown in Eq. (1), where we optimize

cell locations in a row while keeping the order of cells.

$$\min \sum_i -k_i(x_{i+1} - x_{i-1}) + \sum_i (d_i^r - d_i^l), \quad (1a)$$

$$\text{s.t. } x_{i+1} - x_i \geq w_i, \quad (1b)$$

$$d_i^r \geq x_i^0, \quad d_i^r - x_i \geq 0, \quad (1c)$$

$$d_i^l \leq x_i^0, \quad d_i^l - x_i \leq 0, \quad (1d)$$

where x_i denotes the position of i^{th} cell in the row, x_i^0 denotes the initial position of the cell and w_i denotes the width of the cell. Variables d_i^l and d_i^r are introduced to compute displacement of cells in the objective. Eq. (1b) ensures no overlaps between cells. Eq. (1c) ensures that $d_i^r \geq \max(x_i, x_i^0)$. Eq. (1d) ensures that $d_i^l \leq \min(x_i, x_i^0)$. Thus minimizing $d_i^r - d_i^l$ in the objective is equivalent to minimize the displacement $|x_i - x_i^0|$. The objective consists of two parts: OSE and displacement. The OSE terms are the weighted summation of $(x_{i+1} - x_{i-1})$ which correlates to the overall left and right OSE of the i^{th} cell. The weight k_i is the OSE delay slope ($|\Delta\text{delay}/\Delta\text{OSE}|$) mentioned in the greedy approach and $-k_i$ guarantees to maximize the weighted overall OSE. The displacement terms ($d_i^r - d_i^l$) minimize the perturbation on the circuit placement to avoid potential timing degradations.

While Formulation (1) can be solved by linear programming (LP), more efficient algorithm based on dual min-cost flow is proposed [20–22]. Fig. 8 is an example for min-cost flow implementation on optimizing the OSE of cell 2 and cell 3 while involving 4 cells. The formulation can be written as follows and we omit the displacement terms for simplicity.

$$\min \quad -k_2(x_3 - x_1) - k_3(x_4 - x_2), \quad (2a)$$

$$\text{s.t. } x_2 - x_1 \geq w_1, x_3 - x_2 \geq w_2, x_4 - x_3 \geq w_3, \quad (2b)$$

$$x_1 \geq l, -x_4 \geq w_4 - r, \quad (2c)$$

We can construct its network flow graph in Fig. 8. Each cell corresponds to one node in the graph and each differential constraint $x_{i+1} - x_{i-1} \geq w_i$ corresponds to an edge from node $i+1$ to node $i-1$ with an edge cost of $-w_i$ and a capacity of infinity. Additional nodes s and t are introduced to handle bound constraint in Eq. (2c). The weights in the objective correspond to node supplies shown as the labels next to nodes. The final position of i^{th} cell can be computed from the difference between the potentials of node i and node t after solving the network flow problem.

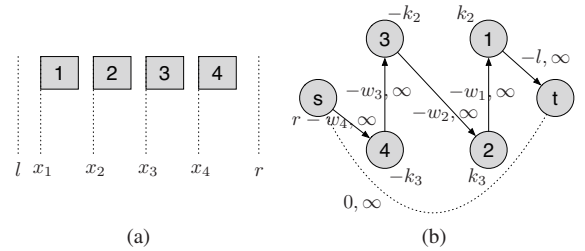


Fig. 8. (a) Cell locations and (b) an example network flow graph.

C Proposed Aging Mitigation Framework

Fig. 9 is an overview of our proposed aging mitigation framework. Given an initial design with gate-level netlist and placement file, we add aging information based on cell locations in placement

file. Then, commercial auto place & route (APR) tool routes the design and performs RC extraction. Next step is updating aging timing using STA engine. STA engine takes in aging library and timing constraints for EOL timing analysis. Section III-A explains how the modeling work is done. After getting the EOL timing, aging mitigation process will identify a set of EOL CPs and optimize EOL timing by cell replacement and cell spreading, which is explained in Section IV-A&B. After the optimization, we get the aging-optimized placement for LDA mitigation.

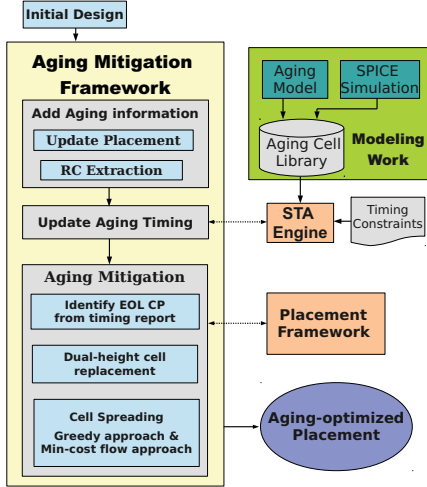


Fig. 9. LDA mitigation framework.

V EXPERIMENTAL RESULTS

This section presents the experimental results of the cell-level aging evaluation and circuit-level aging mitigation framework. In cell-level aging evaluation, we adopt the *ASAP 7nm PDK* [23, 24] for layout creation. CalibreXRC [25] is used to extract layout parameters including aging parameters. EOL cell delay is simulated with HSPICE based on the aging model [4]. An aging timing library is further built based on the HSPICE results. To evaluate our timing mitigation framework, both the greedy method and concurrent method are implemented in C++ for comparisons. Benchmarks are selected from IWLS2005 benchmark suite [26]. The RTLs are synthesized with Synopsys Design Compiler [27], and placed/routed with Cadence Innovus [28]. We analyze the timing results with Synopsys Primetime [29]. All the experiments are performed on a Linux machine with 8-core 3.4GHz Intel processor and 32 GB memory.

A Cell-Level Aging Evaluation

The cell-level aging evaluation focuses on the EOL delay degradation considering different LDA parameters. Table I compares the delay degradation for cells with different OSEs. We select cells with 92nm OSE and 740nm OSE as examples. 92nm OSE is selected because it is the minimum OSE to pass DRC in this technology. 740nm is selected because it falls into region 2 in Fig. 4 where the delay degradation due to OSE is significantly less. Table I shows that, with 92nm OSE, EOL delay increases

TABLE I
Cell delay comparison for different OSE

cell	Early-life delay (ps)	Δ Delay (92nm)	Δ Delay (740nm)
INV_X1	9.01	37.96%	11.30%
NAND2_X1	8.64	46.16%	11.97%
NOR2_X1	13.42	24.63%	12.07%

TABLE II
Cell delay comparison for different cell height.

cell	Early Life			EOL		
	Single (ps)	Dual (ps)	Δ Delay (%)	Single (ps)	Dual (ps)	Δ Delay (%)
INV_X2	6.18	7.31	18.27%	9.24	10.94	18.35%
NAND2_X1	8.64	9.50	9.88%	12.63	13.72	8.56%

more than 36% on average and can be as large as 46.16% for NAND2_X1 cell. When OSE is increased to 740nm, cell delay reduces by 25% on average, which demonstrates the importance to consider OSE-induced aging impacts.

We compare the early-life and EOL delay for cells with different heights in Table II. INV_X2 and NAND2_X1 are used as examples. For early life timing, dual-row height cells have 18.27% and 9.88% additional delay in comparison to the single-row height cells for INV_X2 and NAND2_X1. This is because our predictive model includes the conventional LDEs for early-life timing. Similarly, EOL cases have 18.35% and 8.56% additional delay in dual-row height cells for INV_X2 and NAND2_X1. It has been discussed that the difference between early-life and EOL timing is resulted from the change of stress and aging induced by LOD. Replacing dual-row height cells with single-row height cells benefits both early life and EOL delay.

B Compare Different Aging Mitigation Approaches

We demonstrate the effectiveness of the concurrent aging optimization by comparing EOL timing results with the cases when no optimization is applied and when only greedy optimization is applied. Experiments are performed on IWLS2005 benchmarks shown in Table III, including the gate counts (Gate #) and clock period (T). We first compare the early-life timing for the three schemes aforementioned. In Table IV early-life timing column, D_o , D_g , and D_c are denoted as the CP delay for original placement, CP delay for placement after greedy optimization, and CP delay for placement after concurrent optimization, respectively. Note that both greedy optimization and concurrent optimization include cell replacement for better EOL timing. The change of the early-life timing is less than 0.2% on average after both optimizations and all benchmarks still achieve the clock periods in Table III. Some benchmarks, e.g., *tv80s*, experience slight CP delay increase due to cell spreading. This is because cell spreading causes larger wire RC delay on CPs. Other benchmarks, e.g., *aes_core*, have smaller CP delay due to better LOD and OSE stress effects on early-life timing. The reason is that early-life LDEs, such as LOD and OSE stress effects, have similar dependency on layout parameters as LDA as discussed in Section I. For benchmark *tv80s*, we further show the detailed timing slack distribution in Fig. 10(a). Timing slack ranges after greedy and concurrent optimization are controlled between 70ps to 150ps, which is the same as the slack range of the original placement. We set the benchmark initial area utilization rate to be 0.7 and the increased area due to spreading is accommodated by the whitespace in the original placement so that overall area is not impacted.

We further compare the EOL timing for all the three schemes. Timing violations are represented using total negative slack

TABLE III
Benchmark description

ckt	ss_pcm	simple_spi	sasc	tv80s	ac97_ctrl	usb	aes_core
Gate #	767	1493	1209	14664	25807	29710	49557
T (ps)	357	345	303	833	500	667	588

TABLE IV
Comparisons on solution qualities of different aging mitigation approaches

ckt [26]	Early-life Timing			EOL Timing w/o Opt.		EOL Timing w/ Greedy Opt.					EOL Timing w/ Concurrent Opt.				
	D_o (ps)	D_g (ps)	D_c (ps)	TNS (ps)	WNS (ps)	TNS (ps)	ΔTNS (%)	WNS (ps)	ΔWNS (%)	cpu (s)	TNS (ps)	ΔTNS (%)	WNS (ps)	ΔWNS (%)	cpu (s)
ss_pcm	336.42	336.76	337.18	182.79	37.15	153.14	18.48%	33.69	9.31%	0.01	144.09	23.27%	32.29	13.08%	0.17
simple_spi	342.03	342.02	342.37	202.59	29.27	71.69	64.61%	15.54	46.90%	0.03	58.31	71.22%	14.06	51.96%	0.34
sasc	290.96	291.78	292.82	211.40	30.35	149.81	29.13%	27.36	9.85%	0.02	128.91	39.02%	25.61	15.62%	0.34
tv80s	749.93	753.79	753.48	2436.16	77.55	1860.37	23.64%	65.67	15.32%	0.18	1598.48	34.39%	59.61	23.13%	2.99
ac97_ctrl	440.41	441.33	439.06	401.18	28.12	267.06	33.43%	22.31	20.66%	0.43	172.22	57.07%	16.59	41.00%	6.60
usb	626.84	627.53	637.91	352.80	61.76	331.46	6.05%	61.66	0.16%	0.49	235.34	33.29%	51.41	16.76%	7.51
aes_core	575.48	561.67	563.10	220.81	47.69	155.29	29.67%	42.53	10.82%	0.58	136.41	38.22%	40.65	14.76%	8.87
Average Reduction				—	—	—	29.28%	—	16.15%	—	—	42.35%	—	25.19%	—

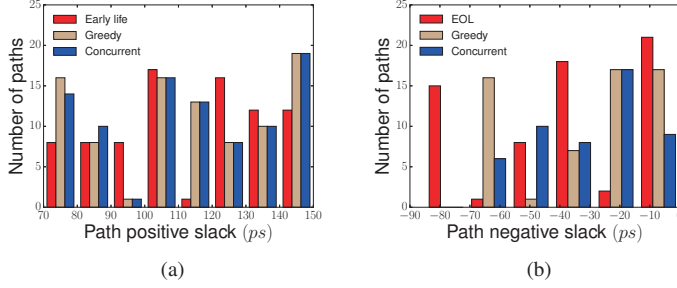


Fig. 10. tv80s slack histogram for, (a) early-life (b) EOL.

(TNS) and worst negative slack (WNS) in ps. The EOL timing w/ optimization columns show the reduced TNS/WNS and the reduction percentage ($\Delta TNS/\Delta WNS$) for both greedy approach and concurrent approach. The optimization runtime is reported as “cpu (s)”. The greedy cell spreading with cell replacement method achieves 29.28%, 16.15% average reduction in TNS and WNS, respectively. With the concurrent optimization approach, TNS and WNS can be further reduced to 42.35% and 25.19%, respectively. Runtimes are less than 1s and 10s across all benchmarks for the greedy and concurrent approach, respectively. We also show the detailed timing slack distribution for benchmark tv80s in Fig. 10(b). In Fig. 10(b), EOL has 15 paths near -80ps slack range for the original placement while there is no path at < -80ps range for greedy and concurrent optimization results. The WNS values for tv80s are -77.55ps, -65.67ps, and -59.61ps for EOL timing w/o optimization, EOL timing w/greedy optimization, and EOL timing w/ concurrent optimization from Table IV. After greedy and concurrent optimizations, negative slack distributions are shifting to the right in the histogram, which means TNS is decreasing and the design is closer to timing closure.

VI CONCLUSION

An LDA mitigation framework for critical path timing is proposed for the first time. We first perform comprehensive standard cell timing evaluations leveraging complex LDA behaviors from LOD and OSE, which reveals that multiple-row height cells are more susceptible to aging than single-row height cells due to LOD effects. The OSEs among neighboring cells are concurrently allocated based on a min-cost flow model, which greatly improves the EOL timing and reduces the pessimistic guard-banding. The min-cost-flow-based concurrent aging mitigation reduce the TNS and WNS by 42% and 25%, respectively, for the EOL timing.

REFERENCES

- [1] B. Yu, X. Xu, S. Roy, Y. Lin, J. Ou, and D. Z. Pan, “Design for manufacturability and reliability in extreme-scaling VLSI,” *Science China Information Sciences*, vol. 59, no. 6, p. 061406, 2016.
- [2] C. Ndiaye, R. Berthelon, V. Huard, A. Bravaix *et al.*, “Reliability compact modeling approach for layout dependent effects in advanced CMOS nodes,” in *IEEE International Reliability Physics Symposium (IRPS)*, April 2017, pp. 4C-4.1-4C-4.7.
- [3] H. C. Ou, K.-H. Tseng, J. Y. Liu, I.-P. Wu, and Y. W. Chang, “Layout-dependent-effects-aware analytical analog placement,” in *ACM/IEEE Design Automation Conference (DAC)*, June 2015, pp. 1-6.
- [4] P. Ren, X. Xu, P. Hao, J. Wang, R. Wang *et al.*, “Adding the missing time-dependent layout dependency into device-circuit-layout co-optimization - New findings on the layout dependent aging effects,” in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2015, pp. 11.7.1-11.7.4.
- [5] C. Liu, M. Jin, T. Uemura *et al.*, “New insights into 10nm FinFET BTI and its variation considering the local layout effects,” in *IEEE International Reliability Physics Symposium (IRPS)*, April 2017, pp. XT-2.1-XT-2.4.
- [6] Y. Wang, S. Cotoana, and L. Fang, “A unified aging model of NBTI and HCI degradation towards lifetime reliability management for nanoscale MOSFET circuits,” in *IEEE/ACM International Symposium on Nanoscale Architectures*, June 2011, pp. 175-180.
- [7] M. G. Bardon, V. Moroz, G. Eneman, P. Schuddinck *et al.*, “Layout-induced stress effects in 14nm & 10nm FinFETs and their impact on performance,” in *Symposium on VLSI Technology (VLSIT)*, 2013, pp. T114-T115.
- [8] S. K. Marella, A. R. Trivedi, S. Mukhopadhyay, and S. S. Sapatnekar, “Optimization of FinFET-based circuits using a dual gate pitch technique,” in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Nov 2015, pp. 758-763.
- [9] F. Firouzi, F. Ye, K. Chakrabarty, and M. B. Tahoori, “Aging- and Variation-Aware Delay Monitoring Using Representative Critical Path Selection,” *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, vol. 20, no. 3, pp. 39:1-39:23, Jun. 2015.
- [10] R. Faraji and H. R. Naji, “Adaptive Technique for Overcoming Performance Degradation Due to Aging on 6T SRAM Cells,” *IEEE Transactions on Device and Materials Reliability*, vol. 14, no. 4, pp. 1031-1040, Dec 2014.
- [11] J. Fang and S. S. Sapatnekar, “Incorporating hot-carrier injection effects into timing analysis for large circuits,” *IEEE Transactions on Very Large Scale Integration Systems (TVLSI)*, vol. 22, no. 12, pp. 2738-2751, 2014.
- [12] —, “The impact of BTI variations on timing in digital logic circuits,” *IEEE Transactions on Device and Materials Reliability*, vol. 13, no. 1, pp. 277-286, 2013.
- [13] “Cadence RelXpert Manual,” ---.
- [14] R. Li, L. Yu, H. Xin, Y. Dong, K. Tao, and C. Wang, “A comprehensive study of reducing the STI mechanical stress effect on channel-width-dependent Idsat,” *Semiconductor Science and Technology*, vol. 22, no. 12, p. 1292, 2007.
- [15] G. Scott, J. Lutze, M. Rubin, F. Nouri, and M. Manley, “NMOS drive current reduction caused by transistor layout and trench isolation induced stress,” in *IEEE International Electron Devices Meeting (IEDM)*, Dec 1999, pp. 827-830.
- [16] “BSIM-CMG Technical Manual and Code, 2012,” <http://bsim.berkeley.edu/?page=BSIMCMG>.
- [17] S.-H. Baek, H.-Y. Kim, Y.-K. Lee, D.-Y. Jin, S.-C. Park, and J.-D. Cho, “Ultra-high density standard cell library using multi-height cell structure,” pp. 72 680C-72 680C-8, 2008.
- [18] Y. Lin, B. Yu, and D. Z. Pan, “Detailed placement in advanced technology nodes: a survey,” in *Solid-State and Integrated Circuit Technology (ICSIT), 2016 13th IEEE International Conference on*. IEEE, 2016, pp. 836-839.
- [19] Y. Lin, B. Yu, X. Xu, J.-R. Gao, N. Viswanathan, W.-H. Liu, Z. Li, C. J. Alpert, and D. Z. Pan, “Mrdp: Multiple-row detailed placement of heterogeneous-sized cells for advanced nodes,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2017.
- [20] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, 1993.
- [21] X. Tang, R. Tian, and M. D. F. Wong, “Optimal redistribution of white space for wire length minimization,” in *IEEE/ACM Asia and South Pacific Design Automation Conference (ASPDAC)*, 2005, pp. 412-417.
- [22] Y. Lin, B. Yu, and D. Z. Pan, “High performance dummy fill insertion with coupling and uniformity constraints,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. PP, no. 99, pp. 1-1, 2016.
- [23] L. T. Clark, V. Vashishtha, L. Shifren, A. Gujja, S. Sinha, B. Cline, C. Ramamurthy, and G. Yeric, “ASAP7: A 7-nm finFET predictive process design kit,” *Microelectronics Journal*, vol. 53, pp. 105 - 115, 2016.
- [24] X. Xu, N. Shah, A. Evans, S. Sinha, B. Cline, and G. Yeric, “Standard Cell Library Design and Optimization Methodology for ASAP7 PDK,” in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Nov 2017.
- [25] M. Graphics, “Calibre verification user’s manual,” 2008.
- [26] “IWLs 2005 Benchmarks,” <http://www.iwls.org/iwls2005/Benchmarks.html>.
- [27] “Synopsys Design Compiler,” <http://www.synopsys.com>.
- [28] “Cadence Innovus,” <http://www.cadence.com>.
- [29] “Synopsys PrimeTime,” <http://www.synopsys.com>.