# Hot carrier aging and its variation under use-bias: kinetics, prediction, impact on Vdd and SRAM

M. Duan[1], J. F. Zhang[1], A. Manut[1], Z. Ji[1], W. Zhang[1], A. Asenov[2], L. Gerrer[2], D. Reid[3], H. Razaidi[2], D. Vigar[4], V. Chandra[5], R. Aitken[5], B. Kaczer[6], and G. Groeseneken[6]

[1]School of Engineering, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK ( j.f.zhang@ljmu.ac.uk )
[2]Dept. Electronics and Electrical Engineering, University of Glasgow, UK. [3]GSS, Glasgow, UK.
[4]CSR, Cambridge CB4 0WZ, UK. [5]Arm R&D, San Jose, USA. [6]IMEC, Leuven B3001, Belgium.

## Abstract

As CMOS scales down, hot carrier aging (HCA) scales up and can be a limiting aging process again. This has motivated re-visiting HCA, but recent works have focused on accelerated HCA by raising stress biases and there is little information on HCA under use-biases. Early works proposed that HCA mechanism under high and low biases are different, questioning if the high-bias data can be used for predicting HCA under use-bias. A key advance of this work is proposing a new methodology for evaluating the HCA-induced variation under use-bias. For the first time, the capability of predicting HCA under use-bias is experimentally verified**.** The importance of separating RTN from HCA is demonstrated. We point out the HCA measured by the commercial Source-Measure-Unit (SMU) gives erroneous power exponent. The proposed methodology minimizes the number of tests and the model requires only 3 fitting parameters, making it readily implementable.

## Introduction

Recent results (**Fig.1**) show Hot Carrier Aging (HCA) can be severe for current/future CMOS nodes [1-7], because: (i) Channel length downscaling enhances HCA (**Fig.2a**). For some sub-30nm processes, HCA can be higher than BTI (**Figs.1b&2b**); (ii) HCA can have larger time exponents (**Figs.1b&2b**) [3-7] and its importance increases with aging. (iii) NBTI recovery [8-10] is higher than HCA (**Fig.2c**); (iv) Conventionally, the worst HCA occurs during switch near Vg~Vd/2 and duty factor (DF) is typically low (1~2%) [7,11]. For modern CMOS, however, more damage occurs under Vg=Vd (**Fig. 3**) [3,6,7] and DF can be high. For example, during 'read 0' in a SRAM cell, one access nMOSFET can suffer HCA for ~50% of time (**Fig. 4**).

The renewed HCA-threat has motivated its re-visit [1-7,12]. It is reported aging mechanisms and time exponent, 'n' (**Eq.1** in **Table 1**), are different under different stress biases [1,6,7]. 'n' can also vary with time (e.g. **Fig.5**) [2,4,5,7], challenging the lifetime, τ, prediction based on **Eq.1** that requires a constant 'n' [11,13,14]. The recent works have focused on bias-accelerated HCA [1-7,12] and there is little data on HCA under **use-bias**. For test engineers, *two pressing questions* are: can τ under use-bias still be predicted by the established JEDEC method based on **Eq.1** and how to evaluate 'n' correctly for HCA? *A key advance of this work is answering them and finding the pitfalls for extracting 'n'. For the first time, the capability of predicting HCA under use-bias is experimentally verified (**Fig.6**).*

## Devices and Experiments

nMOSFETs of MG/HK were made by an industrial process with L×W of 27×(90~900)nm and use-Vdd of 0.9V. Vd=Vg is chosen to represent stress, as Isub/Id has a device-to-device variation (DDV) at stress-0 for nm-devices (**Fig.7a**) and it does not correlate with HCA (**Fig.7b**). All tests were at 125ºC.

## Methodology

*A. Selecting parameter for extracting power exponent, 'n'*

HCA was widely monitored by forward saturation current shift under Vg=Vd=Vdd, ΔId/Id_F, although reverse saturation current shift, ΔId/Id_R, and ΔVth(Vd≤0.1V) also were used [1-7,11,12]. The problem is 'n' for ΔId/Id_F is larger than 'n' for ΔId/Id_R, leading to their incorrect cross-over and errors in prediction at 10 years (**Fig.8**), highlighting the importance of 'n'-accuracy. Under Vg=Vd, ΔId/Id_F does not sense the HCA-defects above space charge region (**Fig.8**), resulting in an apparent larger 'n', as simulated by subtracting a constant from real power law (inset of **Fig.8**). The 'n' extracted from the forward ΔId/Id is erroneous. To capture all defects, ΔVth(Vd=0.1V) should be used for extracting 'n', as ΔVth_F=ΔVth_R (**Fig.9**). Once ΔVth is predicted, we propose evaluating ΔId/Id_F and ΔId/Id_R by using their measured relation with ΔVth (**Fig.10**).

*B. HCA acceleration*

SRAM often is used for qualifying new processes [15], where the access nMOSFETs suffer the worst HCA under Vg≈Vd (**Figs.3&4**). HCA under use-Vg=Vd must be predicted and we focus on it here. Under use-bias, **Fig.11** shows that HCA is too low to establish its kinetics reliably within a practical time and acceleration is needed. One may accelerate HCA by raising Vd only [12] or both Vg and Vd with Vg=Vd [1,6]. **Fig.12a** confirms 'n' is larger under Vg<Vd than under Vg=Vd [1,7], so that Vg<Vd must not be used to predict HCA under use-Vg=Vd. When accelerated by Vg=Vd, 'n' is bias-independent (**Fig.12b**) and should be used.

*C. DC versus AC*

Unlike NBTI(AC)<NBTI(DC) [8-10], the AC and DC HCAs agree well, regardless of frequency and duty-factor (DF) for the same equivalent stress time, i.e. DF×time (**Fig.13**), confirming the frequency-independence [3]. DC will be used.

*D. Voltage-Step-Stress (VSS)*

The VSS technique recently developed for NBTI [16] allows extracting both 'n' and 'm' (**Eq.1**) from just one large

device, reducing test numbers by ~80%, and will be adopted for HCA here. For an L×W=27×900nm, stress under each Vg=Vd lasted for T=1ks and biases were then stepped up (**Fig.14**), lifting HCA up from the power law (**Fig.14c**). Based on **Eqs.1-3**, HCA under a high Vg=Vd was converted into a longer equivalent stress time under a low Vg=Vd (**Fig.14c**) and ΔVth follows a power law well even when ΔVth>150mV, corresponding to ΔId/Id>30% (**Fig.10**), well beyond the typical 10% HCA lifetime criterion and allowing reliable extraction of 'n' and 'm' (**Fig.14c**).

## Prediction

A model is useful only if it can predict aging under use-bias. The HCA predicted by the model extracted from the VSS data in **Fig.14** agrees well with the test data in **Figs.6b-g**. The highest ΔVth in **Figs.6a&14** is ~2-orders above ΔVth under 0.9V (**Fig.6b**), verifying its prediction capability. We emphasize the model was extracted from the data in **Fig.6a** only and the test data in **Figs.6b-g** were not used for fitting. The extracted model (**Eq.1**) can be used for evaluating HCA under any bias and time and for predicting lifetime and operation Vdd (**Fig. 15**).

## HCA in nm-width devices

Unlike L×W=27×900nm, 27×90nm devices suffer from RTN-like within-a-device fluctuation (WDF) and large device-to-device variation (DDV) (**Fig.16**). To extract HCA kinetics, one has to use the smooth mean of 50 devices, but 'n' depends on how data is taken (**Fig.17**). After a stress, ΔVth fluctuates and one can use its up-envelope (UE), lower-envelope (LE) [17], or average over a period of time, e.g. ~10ms (**Fig.17b**), as a typical quasi-DC Source-Measure-Unit (SMU) does. The 'n' from UE and DC (inset of **Fig. 17a**) is smaller than the 'n=0.29' from a device of W=900nm (**Fig.14c**), but the 'n' from LE agrees well with it. The smaller 'n' for UE incorrectly takes it below LE when extrapolating (see the 'cross-over' in **Fig. 17a**).

To explain the difference in 'n', **Figs.18a&b** show that WDF=(UE-LE) does not increase with aging. It must originate from as-grown defects and should be excluded from aging kinetics, so that LE must be used for extracting 'n'. LE_F and LE_R correlates (**Fig.19a**), but WDF_F and WDF_R does not (**Fig.19b**), supporting their different origins.

Since HCA-recovery is insignificant (**Fig.2c**), one may think it can be measured by a quasi-DC SMU [7,18]. This, however, gives an erroneous lower 'n' (**Fig.17a**) by including some as-grown WDF. Adding a constant to a power law leads to an apparent lower 'n' at short time and a variation of 'n' with time (inset, **Fig.17a**) [13].

## Statistic HCA

The DDV of LE at different time (**Figs.20a&b**) and voltage (**Figs.20c&d**) follows the defect-centric distribution (**Eqs.4&5**) well [19]. LE_mean of 50 W=90nm devices agrees well with ΔVth of one W=900nm device (**Fig.21a**) and can be predicted by the same method (**Figs.6&14**). After knowing LE_mean, the standard deviation, σ, can be evaluated from its power law relation with the mean (**Fig.21b**).

### *A. Impact on use-Vdd*

To have a yield corresponding to i×σ, ΔId/Id=10% is required at i×σ, resulting in smaller mean value (**Fig.22a**) and in turn lower use-Vdd (**Fig.22b**) for higher i. For a yield of 3×σ (99.7%), HCA-only and HCA+WDF (**Eqs.6,7**) reduces Vdd from its zero-spread value by 75mV and 100mV, respectively.

### *B. Impact on 6T-SRAM*

Assuming only one access nMOSFET suffered HCA and using the predicted HCA distribution at 10 years under 0.9V for simulation [20], static write/read noise margin reduces/rises, respectively (**Fig.23**), as a weakened access nMOSFET is not in favor of write. Both the dynamic read (**Figs.24a-c**) and write (**Figs.24d-f**) access time deteriorates, since longer time is required through a weakened access nMOSFET. This demonstrates that the extracted HCA model can be incorporated into a compact simulator to evaluate the required margin for a specified yield.

## Conclusions

As CMOS scales down, HCA scales up. For the first time, this work experimentally verifies that the HCA under use-Vdd can be predicted by the power law extracted from VSS-method, provided that correct acceleration and 'n'-evaluation are made. We point out the forward saturation ΔId/Id and HCA measured by SMU gives erroneous 'n' for nm-width devices. The model requires only 3 fitting parameters (**Eq.1**), making it readily implementable.

## Acknowledgement

The authors thank Plamen Asenov of GSS for his critical comments. This work was supported by the EPSRC of U.K. under the Grant No. EP/L010607/1.

## References

[1] A. Bravaix et al, IRPS 2013, 2D.6.1-2D6.9.
[2] J. H. Stathis et al, IEDM 2014, p. 522.
[3] G. T. Sasse, IEEE Trans.Elec.Dev., 55, pp.3167-3174, 2008.
[4] A. J. Scholten et al, IEEE Trans.Elec.Dev., 58, pp.1-8, 2011.
[5] F. Cacho et al. IRPS 2014, 5D.4.1.
[6] A. Bravaix et al, IEDM 2011, p.622.
[7] M. Cho et al, IEEE Trans.Elec.Dev., 60, pp.4002-4007, 2013.
[8] S. Ramey et al, IRPS 2014, XT.2.1.
[9] M. H. Chang and J. F. Zhang, J.Appl.Phys., 101, art. no. 024516, 2007.
[10] S. F. W. M. Hatta et al, IEEE Trans.Elec.Dev., 60, pp.1745-1753, 2013.
[11] JEDEC 2011 p.15.
[12] C. Liu et al, IEDM 2014, p 836.
[13] Z. Ji et al, IEDM 2013, p.413
[14] Z. Ji et al, IEEE Trans.Elec.Dev., 57, pp.228-237, 2010.
[15] U. Bhattacharya, Intel Tech. J. 2008, p.111.
[16] Z. Ji et al, IRPS 2014, GD-2.
[17] M. Duan et al, IEDM 2013, p.774.
[18] N. H. H. Hsu et al, IRPS 2012.
[19] L. M. Procel et al, IEEE Elec.Dev.Lett., 35, p. 1167-1169, 2014.
[20] http://ptm.asu.edu/.

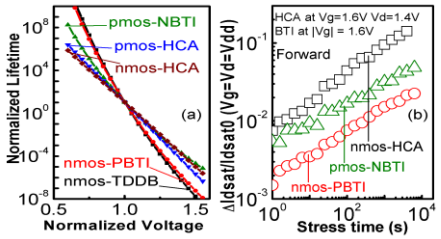



Fig.1 A comparison of Hot Carrier Aging (HCA) with BTIs reported by early works. (a) and (b) are re-plots of data from refs. [2] and [1], respectively.

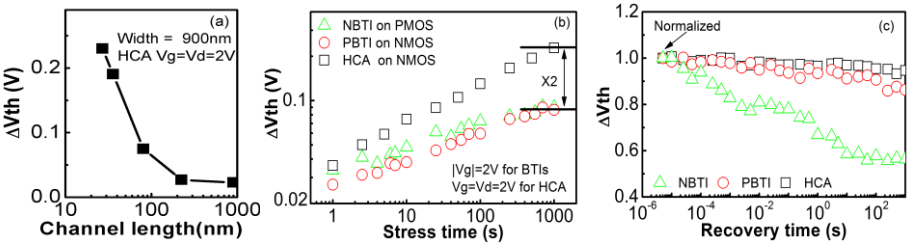



Fig.2 (a) Downscaling L increases HCA. The stress was at 125ºC for 1000sec. (b) A comparison of HCA and BTIs for L=27nm used in this work. Stresses were under the same |Vg|, with Vd=Vg for HCA and Vd=0 for PBTI and NBTI. (c) A comparison of their recovery under Vg=Vd=0.

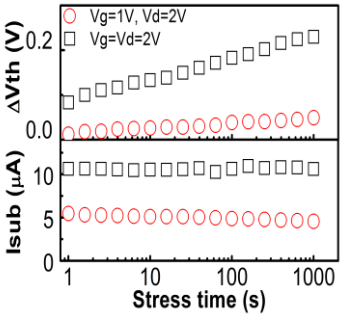



Fig.3 HCA under Vg=Vd is more than HCA under Vg=Vd/2 for L=27nm.

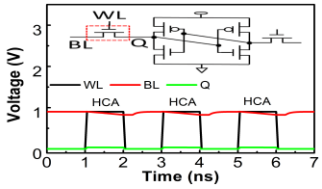



Fig.4 The access nMOS in a SRAM during read-0 has a HCA duty factor of ~50% (simulation).

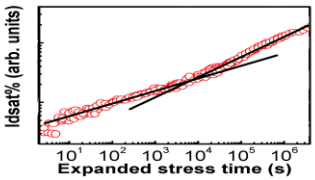



Fig.5 Variation of time exponent, 'n' (the line slope), with HCA time. A re-plot of data from ref. [2].

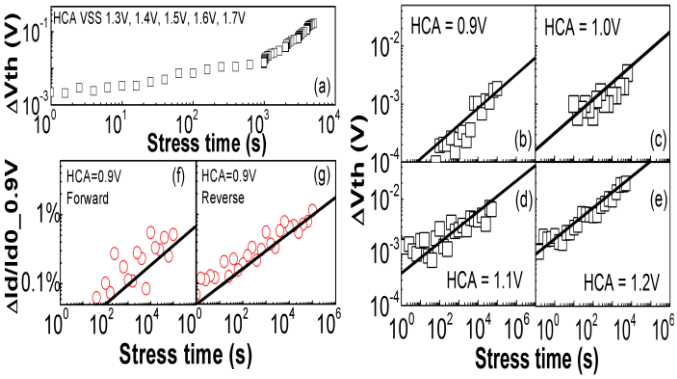



Fig.6 Verification of predicting HCA under use-Vdd. The model was extracted from accelerated VSS test data given in (a) with Vg=Vd rising from 1.3 to 1.7V (For details, see **Fig. 14**). The symbols in (b)-(e) were measured from 4 devices and not used for fitting. The lines in (b)-(e) are the predicted HCA. The lines in (f) and (g) were obtained from **Fig.10** by converting ΔVth to ΔId/Id(Vg=Vd=0.9V).

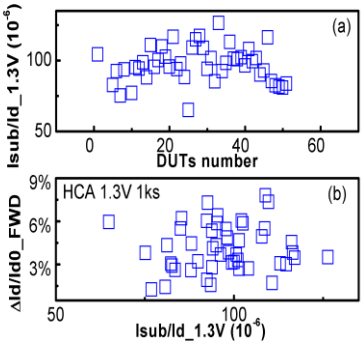
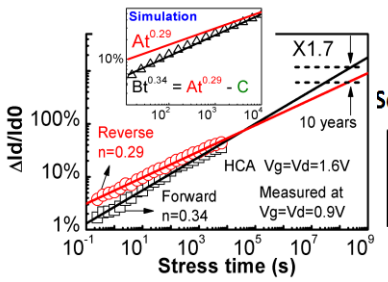



Fig. 7 Isub/Id does not represent HCA-stress well for nm-devices, as it has a device-to-device variation (DDV) at stress=0 (a) and its DDV does not correlate with that of HCA-induced ΔId/Id.



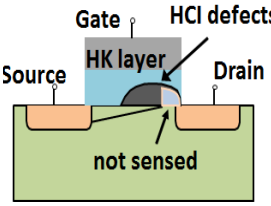




Fig.8 Although test data (**o** and □) show (ΔId/Id_F)<(ΔId/Id_R), higher 'n' for ΔId/Id_F leads to incorrect (ΔId/Id_F)>(ΔId/Id_R) when extrapolating. ΔId/Id_F does not sense the defects above space charges. The '$\bar{\Delta}$' in inset is calculated from ($\bar{At}^{0.29}$-**Constant**), which fits well with $Bt^{0.34}$ (black line). Subtracting a constant from a real power law (red line) can give an 'apparent' higher 'n'. Id was measured under Vg=Vd=0.9V.

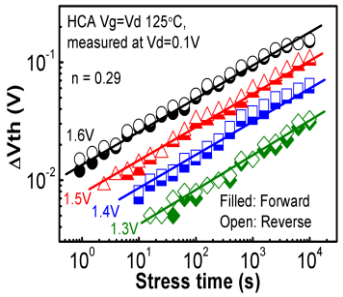



Fig.9 The forward and reverse ΔVth measured under Vd=0.1V agrees well.

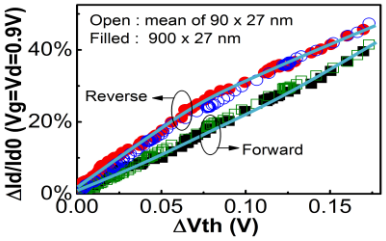



Fig.10 Relation of HCA-induced ΔVth under Vd=0.1V with ΔId/Id under Vg=Vd=0.9V. The open symbols are the mean of 50 small devices (90×27nm)

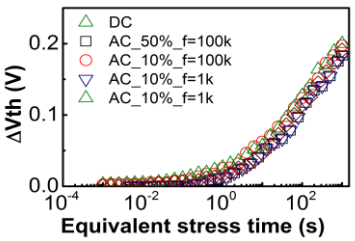



Fig.13 The AC and DC HCAs agree well when using equivalent stress time, i.e. time×Duty Factor (DF). The AC stress conditions are given in the format of 'AC_DF_Frequency' and $\bar{V}g$=$\bar{V}d$=2V.

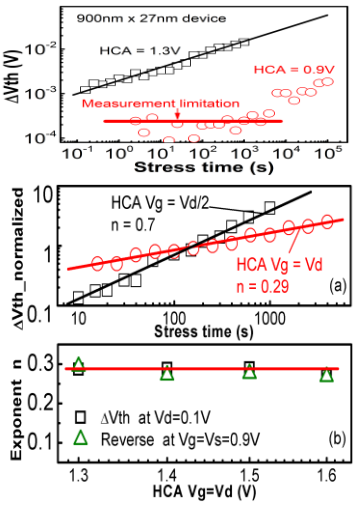



Fig. 11 HCA is too low to establish kinetics reliably under use-Vdd=0.9V and acceleration (e.g. 1.3V) is needed.

*Table* 1:

(1) $\Delta Vth = A \cdot t^n \cdot {V_g}^m$

(2) $T_{eff} = T \cdot (\frac{V_2}{V_1})^{m/n}$

(3) $A \cdot T^n \cdot (V_2)^m = A \cdot {T_{eff}}^n \cdot {V_1}^m \quad (V_2 > V_1)$

(4) $F_N(\Delta V_{th}, \eta) = \sum_{k=0}^{\infty} \frac{e^{-N} N^k}{k!} F_k(\Delta V_{th}, \eta)$

(5) $N = \frac{2\mu^2}{\sigma^2} \quad \eta = \frac{\sigma^2}{2\mu}$

(6) $\mu\_\Delta Vth = \mu\_LE\_ + \mu\_WDF\_$

(7) $\sigma\_\Delta Vth = (\sigma\_LE^2 + \sigma\_WDF^2)^{0.5}$

Fig. 12 (a) The time exponents under Vg=Vd are smaller than that under Vg=Vd/2. (b) The time exponent is insensitive to stress biases under Vg=Vd

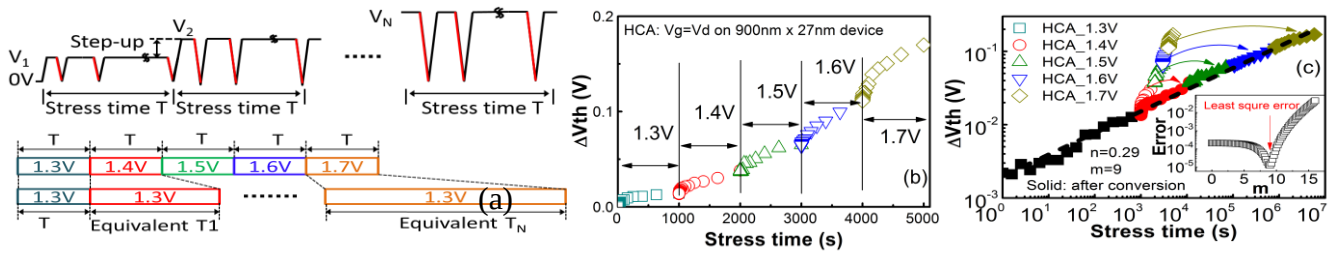



Fig.14 Voltage-Step-Stress (VSS) technique for HCA. (a) One device was stressed for a time T and the stress Vg=Vd was then stepped up. ΔVth is plotted against linear (b) and log (c) stress time. The stress time under high bias is converted to an equivalent longer time at low bias by fitting the voltage exponent 'm' (inset of (c)), based on **Eqs.1-3** in **Table 1**. The dashed line has n=0.29 and m=9.

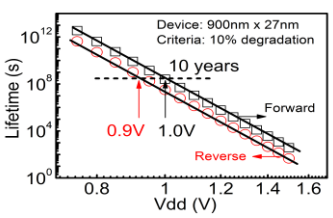



Fig.15 Evaluation of lifetime versus Vdd based on the model extracted from VSS tests in **Fig.14**.

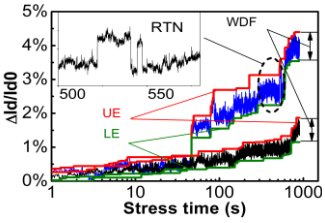



Fig.16 HCA of two W=90nm devices shows large DDV. WDF, UE, and LE is 'within-a-device-fluctuation', the upper- and the lower- envelope.

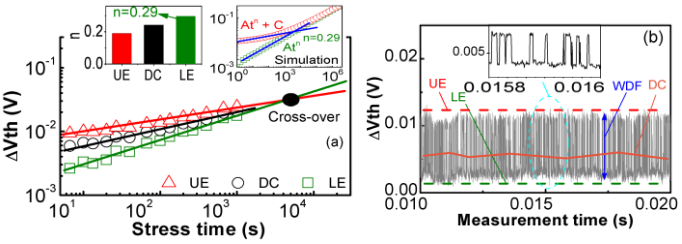



Fig. 17 (a) HCA kinetics for the mean of 50 W=90nm devices. UE, DC, and LE have different 'n' (inset). Incorrect inclusion of an as-grown component, '**C**', gives an apparent lower 'n' at short time. (b) The definition of UE, DC, and LE. DC is the average over 10ms.

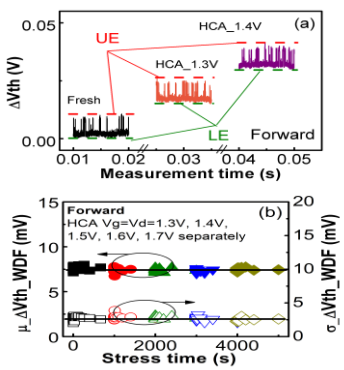



Fig.18 (a) For L×W=27×90nm, LE increases with HCA, but WDF=UE-LE does not. (b) The WDF_mean of 50 devices and its sigma do not increase with stress time.

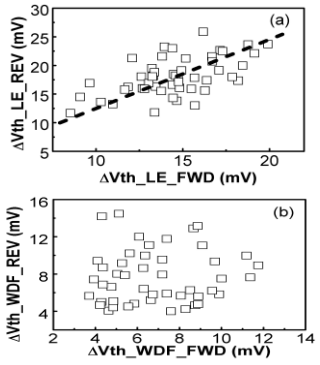



Fig.19 (a) The forward LE_F correlates with LE_R. (b) WDF_F does not correlate with WDF_R.

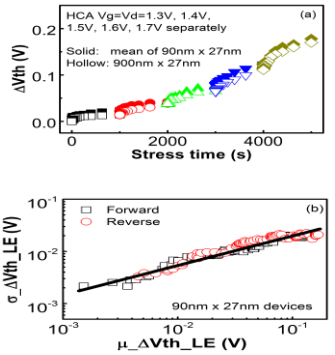



Fig.21 (a) The mean of 50 90×27nm agrees well with one 900×27nm for VSS stresses. (b) Sigma versus mean. The fitted exponent is 0.55.

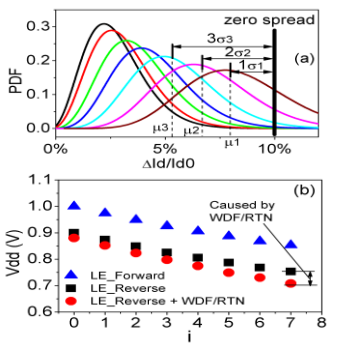



Fig.22 Impact of DDV on use-Vdd. When ΔId/Id reaches 10% at i×σ, the mean ΔId/Id, μ, of defect-centric distributions reduces for higher i (a). This in turn requires a lower use-Vdd (b). For the reverse: '■'--- HCA only and '●' --- HCA and RTN/WDF (**Eqs.6&7**).

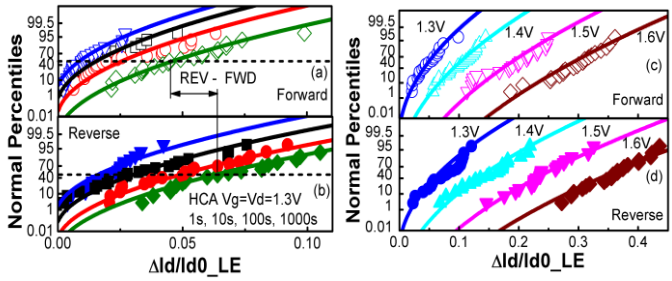
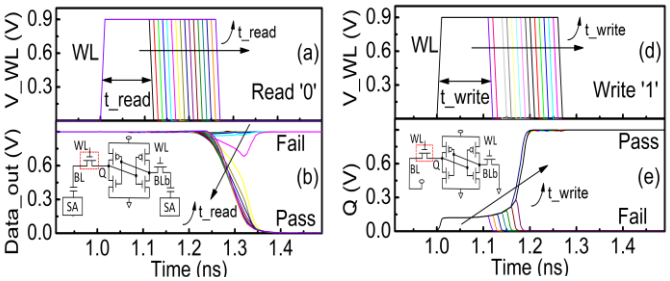



Fig.20 Statistics of LE DDV after different stress time (a&b) and voltage (c&d). The lines are fitted with the defect-centric distribution (**Eqs.4&5**).



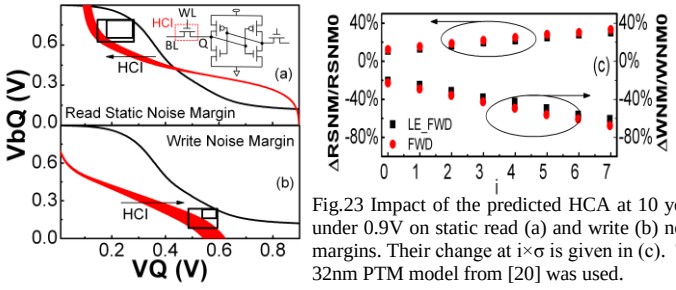




Fig.23 Impact of the predicted HCA at 10 years under 0.9V on static read (a) and write (b) noise margins. Their change at i×σ is given in (c). The 32nm PTM model from [20] was used.

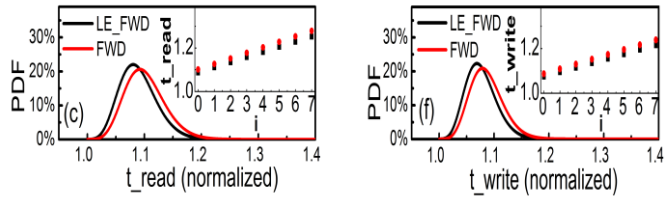



Fig.24 Impact of HCA of access nMOS on dynamic read (a-c) and write (d-f) assess time, normalized against their fresh value. The mean and sigma of HCA under Vdd=0.9V at 10 years were predicted based on test data and then used to compute the defect-centric PDF vs HCA. For a given PDF, the corresponding HCA were used to compute the failure time, as illustrated in (a,b) and (d,e). The (t_Fail, PDF) pairs were plotted in (c) & (f). The insets of (c)&(f) shows the normalized margins against i×σ.