Chapter 2
Characterization Methods for BTI Degradation and Associated Gate Insulator Defects

Souvik Mahapatra, Nilesh Goel, Ankush Chaudhary, Kaustubh Joshi and Subhadeep Mukhopadhyay

Abstract In this chapter, different characterization methods are discussed to determine BTI degradation of MOSFET parameters and to directly estimate the pre-existing and generated gate insulator defects responsible for BTI. V_T shift is obtained from full I_D–V_G sweeps and also from spot I_D measurements at fixed V_G; one spot measurements are performed either on-the-fly at stress V_G or by dropping down to a lower V_G from stress. Impact of measurement delay and mobility degradation on V_T extracted from different methods is discussed. Flicker noise method is used to access the density of pre-existing defects for different gate insulator processes. Gated diode or DCIV, charge pumping (CP) and low voltage SILC methods are used to determine trap generation at or near the interface between Si channel and gate insulator. Conventional SILC is used to estimate generation of bulk gate insulator defects. Different artifacts related to improper choice of stress bias and measurement delay are discussed.

2.1 Introduction

As discussed in Chap. 1, Bias Temperature Instability (BTI) in Metal Oxide Semiconductor Field Effect Transistor (MOSFET) is normally assessed by stressing the device at higher than nominal gate bias ($V_G = V_{G\text{-}STR}$), and recording the shift in device parameters in logarithmic intervals of stress time (t_{STR}). Degradation in device parameters, such as threshold voltage shift (ΔV_T), is extracted from drain current (I_D) measured before and after stress. I_D is usually measured in the linear operating regime (I_{DLIN}), with a low drain bias (V_D) to keep oxide field (E_{OX}) almost uniform across the channel during stress.

S. Mahapatra (✉) · N. Goel · A. Chaudhary · K. Joshi · S. Mukhopadhyay
Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India
e-mail: souvik@ee.iitb.ac.in; mahapatra.souvik@gmail.com

S. Mahapatra (ed.), *Fundamentals of Bias Temperature Instability in MOS Transistors*, Springer Series in Advanced Microelectronics 139,
DOI 10.1007/978-81-322-2508-9_2

In earlier days, complete I_D–V_G sweeps were taken to extract V_T before and during BTI stress, the later was achieved by interrupting the stress for few seconds for measurements [1]. It was soon realized that BTI degradation recovers substantially after stress is paused for measurement [2, 3], and for certain situations, as much as 50 % of degradation recovers in just 1 s after stoppage of stress [4]. This makes the conventional, slow I_D–V_G sweep based Measure-Stress-Measure (MSM) technique unsuitable for BTI characterization, since it cannot capture the correct degradation magnitude. To circumvent this recovery issue, three different ultra-fast measurement techniques have been proposed in the literature, i.e., On-The-Fly (OTF) I_{DLIN}, MSM and One Spot Drop Down (OSDD), and are described in this chapter.

All three classes of measurements described in this chapter can be implemented using the setup illustrated in Fig. 2.1 [5–7]. A pulse generator is connected to the gate terminal of a MOSFET and is used to apply the desired stress and measurement biases. Source and substrate terminals are grounded, and the drain terminal is connected to a current–voltage converter (IVC) and subsequently to a digital storage oscilloscope (DSO). The combination of IVC and DSO helps in ultra-fast I_{DLIN} measurement in $\sim$microseconds measurement time (t_M). All methods are described in this chapter are for Negative BTI (NBTI) stress in p-MOSFETs, although these can be readily adopted for Positive BTI (PBTI) stress in n-MOSFETs.

As described in detail later in this book, BTI is due to uncorrelated contribution from charging of pre-existing defects as well as generation of new defects during stress in the gate insulator of a MOSFET, and measured ΔV_T can be decomposed into these underlying sub-components. Therefore, it is important to independently and directly access the pre-existing and generated gate insulator traps for different gate insulator processes, to verify the accuracy of the decomposition method. Such assessment can be done using different trap characterization methods, i.e., Flicker (1/*f*) Noise, Charge Pumping (CP), Gated Diode (DCIV), Stress Induced Leakage Current (SILC) and Low Voltage (LV) SILC, and are described in this chapter.

Finally, it is important to remark that several stress and measurement artifacts can corrupt BTI experiments and are also discussed in this chapter. Since BTI and TDDB stress regimes are essentially the same, shown in Chap. 1, Fig. 1.2, and since

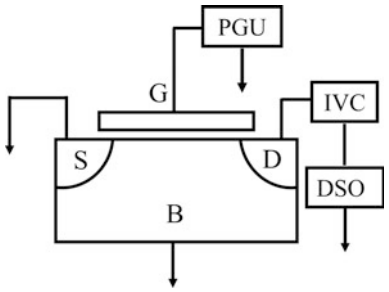

Fig. 2.1 Schematic of a representative setup for ultra-fast BTI measurements

TDDB defects have significantly different t_{STR} and V_G dependence, caution must be used to avoid TDDB corruption of BTI experiments. Several artifacts related to BTI recovery are also discussed.

2.2 Ultra-Fast On-the-Fly (UF-OTF) Method

Figure 2.2 shows the schematic of OTF measurement method. BTI induced I_{DLIN} degradation is measured on-the-fly at stress ($V_G = V_{G\text{-}STR}$) without reduction in V_G, and hence the technique does not suffer from recovery issues [3, 5]. However, the first data point (I_{DLIN0}) measured immediately after the application of $V_{G\text{-}STR}$ is assumed to be unstressed, and extracted ΔI_{DLIN} ($=I_{DLIN} - I_{DLIN0}$) therefore depends on the time-zero delay (t_0 delay) between application of $V_{G\text{-}STR}$ and measurement of I_{DLIN0} [5, 8]. I_{DLIN0} has been measured by using $t_0 = 1$ ms in conventional OTF [3, 9] and $t_0 = 1$ µs in Ultra-Fast OTF (UF-OTF) [5, 8] setups. Furthermore, the OTF technique needs to be modified to account for mobility degradation ($\Delta\mu_{eff}$), as-measured ΔI_{DLIN} is influenced by both ΔV_T and $\Delta\mu_{eff}$, as shown in Chap. 1, especially for NBTI stress in p-channel MOSFETs [10]. In [11], the OTF method is modified presumably to account for mobility degradation, where a small gate pulse is superimposed on $V_{G\text{-}STR}$ to measure transconductance (g_m), and ΔV_T is calculated using $\Delta V_T = \Delta I_{DLIN}/g_m$. However, it will be shown later that this method is not effective in providing correct ΔV_T, and also suffers from complications associated with application of a complex gate waveform. An alternative approach is to compute $\Delta V = -\Delta I_{DLIN}/I_{DLIN0} * (V_{G\text{-}STR} - V_{T0})$ by applying only a DC stress bias and measuring I_{DLIN} at $V_G = V_{G\text{-}STR}$, where V_{T0} is pre-stress V_T of the device. A post-processing correction procedure involving additional I_D–V_G sweep measurements is used to correct for the impact of $\Delta\mu_{eff}$ and convert ΔV to ΔV_T [12], which is discussed later in this chapter. OTF and UF-OTF methods have been

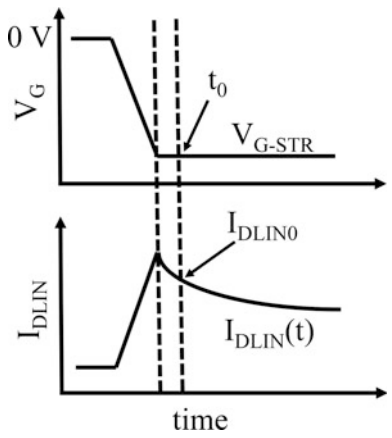

Fig. 2.2 Schematic of OTF measurement of BTI degradation during stress. Default time-zero delay is 1 µs unless mentioned otherwise

extensively used to characterize the gate insulator process dependence of NBTI degradation especially in Silicon Oxynitride (SiON) p-MOSFETs [9, 13], which will be discussed and analyzed later in this book.

2.2.1 Impact of Measurement Delay for DC Stress

Accuracy of the OTF technique depends on time-zero or t_0 delay as mentioned before, as the first data point measured immediately after the application of $V_{G\text{-}STR}$, i.e., I_{DLIN0}, is assumed as unstressed. As shown in Fig. 2.2, t_0 delay is the time lag between application of $V_{G\text{-}STR}$ and measurement of I_{DLIN0}. For a particular t_0 delay, a device that degrades more rapidly soon after the initiation of stress would show larger t_0 delay impact on measured ΔI_{DLIN} and vice versa. OTF measurements with t_0 delay down to 1 μs have been made using the setup shown in Fig. 2.1 [5, 8], and relevant results are discussed in this section. Although the implementation and results are shown for the case of NBTI in SiON p-MOSFETs, OTF method can be used to study NBTI and PBTI in HKMG devices as well.

Figure 2.3a plots I_{DLIN} measured before, during and immediately after application of $V_{G\text{-}STR}$ for OTF measurement in SiON p-MOSFET having Plasma Nitrided Oxide (PNO) gate insulator; refer to Chap. 1, Table 1.1 for SiON device details. In the beginning, the pulse generator shown in Fig. 2.1 supplies a low DC V_G to the gate terminal of the MOSFET, and the DSO is triggered to start sampling of I_{DLIN} in 1 μs intervals. The pulse generator is triggered next and V_G ramps up to $V_{G\text{-}STR}$, and DSO sampling captures the associated rise in I_{DLIN} as the MOSFET is driven to strong inversion. The first data point, i.e., I_{DLIN0}, is captured within 1 μs of the time of V_G becoming equal to $V_{G\text{-}STR}$. Measured I_{DLIN} then starts to decrease with increase in stress time due to BTI degradation.

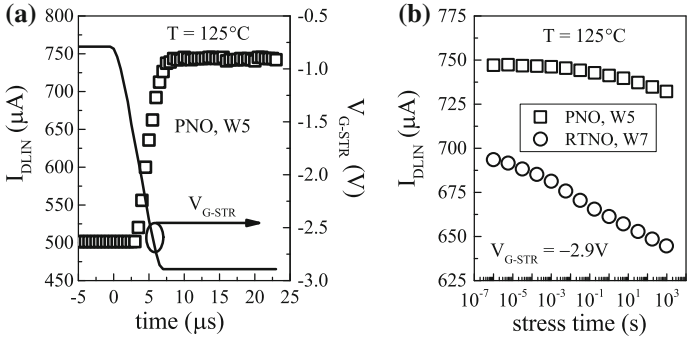

Fig. 2.3 OTF stress measurement: **a** application of gate stress voltage and corresponding measured I_{DLIN}, and **b** time evolution of I_{DLIN} degradation for PNO and RTNO SiON p-MOSFETs under NBTI stress

Figure 2.3b shows measured I_{DLIN} from short to longer stress time for PNO and Rapid Thermal Nitrided Oxide (RTNO) SiON p-MOSFETs. Measured I_{DLIN} at $V_{G\text{-}STR}$ starts to degrade due to BTI degradation. Although the devices have similar Equivalent Oxide Thickness (EOT), RTNO device shows lower starting I_{DLIN} and larger rate of I_{DLIN} degradation with increase in stress time, t_{STR}. As mentioned before, ΔI_{DLIN} is calculated from measured I_{DLIN} degradation using $I_{DLIN} - I_{DLIN0}$, and although the first data point captured within 1 µs can be used as I_{DLIN0}, it is possible to choose measured I_{DLIN} at different t_{STR} as I_{DLIN0} to study the impact of different t_0 delay on ΔI_{DLIN}. BTI degradation is estimated using $\Delta V = -\Delta I_{DLIN}/I_{DLIN0} * (V_{G\text{-}STR} - V_{T0})$ as mentioned above, where V_{T0} is pre-stress V_T of the device, and mobility correction has been used to convert ΔV to ΔV_T using the method described in [12] and discussed later in this chapter.

Figure 2.4 plots ΔV time evolution measured in (a) PNO and (b) RTNO devices using different t_0 delay. Higher degradation is obtained using lower t_0 delay as it captures a less degraded I_{DLIN0} and vice versa. Significantly, larger impact of t_0 delay is observed for RTNO compared to PNO device, as RTNO device has larger rate of I_{DLIN} degradation just after the initiation of stress as shown in Fig. 2.3b.

In spite of having similar EOT and much lower Nitrogen (N) content as mentioned in Chap. 1, Table 1.1, the RTNO device has much larger BTI magnitude compared to the PNO device when stressed at identical electric field, E_{OX}, and temperature (T), and measured using identical t_0 delay; E_{OX} is calculated using $(V_{G\text{-}STR} - V_{T0})$/EOT. Interestingly, the RTNO device shows large degradation in the sub-1 ms time scale when measured using t_0 delay of 1 µs, which is not observed in the PNO device, or with OTF measurements with higher t_0 delay. This observation has significant implication in understanding the underlying physical mechanism of BTI, and will be discussed in detail in Chap. 6. Therefore, Fig. 2.4 clearly demonstrates the importance of ultra-fast OTF measurements to accurately capture device degradation during BTI stress. Time evolution of NBTI degradation

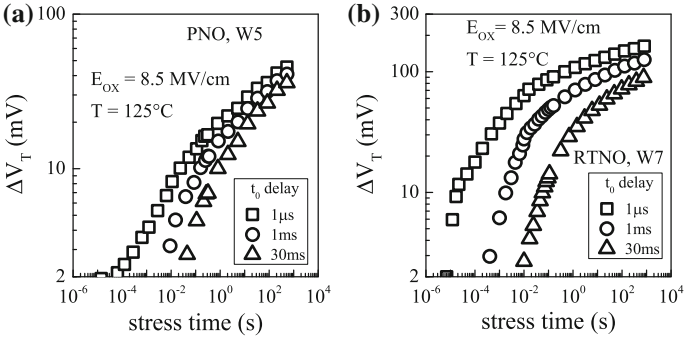

Fig. 2.4 Time evolution of OTF measured and mobility corrected ΔV_T for different time-zero delay, for **a** PNO and **b** RTNO SiON p-MOSFETs under NBTI stress. See Sect. 2.3 for mobility correction

at different stress E_{OX} and T measured using the ultra-fast OTF method in differently processed SiON p-MOSFETs has been shown in Chap. 1 and further analyzed later in this book.

As mentioned in Chap. 1, Fig. 1.16, BTI degradation shows power-law time dependence at longer stress time. Measured data are plotted in a log-log plot and the power-law time exponent (n) is obtained using linear regression in a specified time range. To verify the robustness of power-law time dependence, Fig. 2.5a plots n calculated from ΔV_T time evolution data as a function of $V_{G\text{-}STR}$ for NBTI stress in PNO-SiON p-MOSFET. Extraction is done for different ranges of t_{STR} over which regression has been performed. Note that within measurement error, similar values of n are observed for different t_{STR} range and across different stress V_G, as long as ΔV_T is measured using a small t_0 delay such that recovery artifacts remain negligible. It is important to note that the V_G independence of n shown in Fig. 2.5a does not hold all the time. As mentioned in Chap. 1, measured n can reduce at higher $V_{G\text{-}STR}$ due to saturation effects, which is caused by reduction in stress E_{OX} at longer t_{STR} when ΔV_T becomes large. On the other hand, n can also increase at higher $V_{G\text{-}STR}$ especially for thicker gate insulator devices, due to additional contribution from TDDB like bulk insulator trap generation [14], which is discussed later in this chapter and also in Chap. 4.

Figure 2.5b illustrates the impact of t_0 delay on extracted n for p-MOSFETs having different EOT and N content of PNO and RTNO based SiON gate stacks. Note that n is extracted over a fixed range of t_{STR} and reduces with reduction in t_0 delay, although PNO devices having different EOT and low to moderate N content show higher magnitude of n when compared to the RTNO device. It is important to note that a smaller t_0 captures I_{DLIN0} that is relatively less degraded, and results in larger BTI degradation and lower time exponent n. Similarly for a particular t_0 delay, I_{DLIN0} is less degraded for PNO compared to RTNO device, refer to Fig. 2.3b, and results in smaller degradation magnitude and higher n for the former device. Therefore, it is important to use ultra-fast OTF measurements to properly

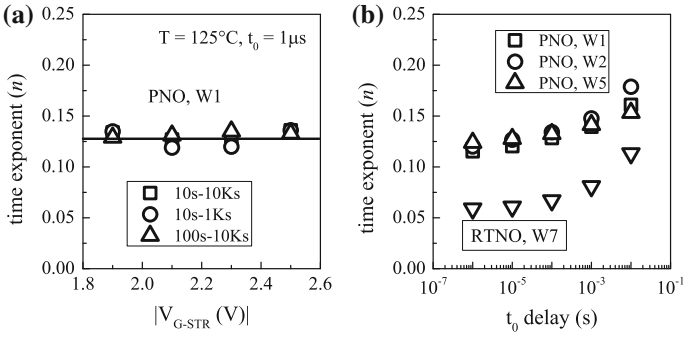

Fig. 2.5 Power-law time exponent n of OTF measured ΔV_T, **a** plotted versus $V_{G\text{-}STR}$ and extracted for different stress time ranges for linear regression, and **b** versus time-zero delay, for PNO and RTNO SiON p-MOSFETs under NBTI stress

measure BTI degradation in wide variety of gate insulator devices. More results of OTF measured time exponent n for different devices and stress conditions are presented and analyzed in other chapters of this book.

2.2.2 Recovery After DC Stress

As mentioned before, it is now well known that BTI degradation recovers substantially after the reduction of stress V_G. Figure 2.6a illustrates the gate bias sequence for OTF measurement of BTI recovery at $V_G = V_{G\text{-}REC}$ following BTI stress at $V_G = V_{G\text{-}STR}$ for a time t_{STR}. The pulse generator applies $V_G = V_{G\text{-}REC}$ to the gate of the MOSFET and the DSO is triggered to measure pre-stress I_{DLIN}, which is recorded as I_{DLIN0}. The pulse generator is then triggered to output $V_G = V_{G\text{-}STR}$ for a duration t_{STR}, during which the device degrades due to BTI, after which V_G drops back to $V_{G\text{-}REC}$. The DSO is synchronously triggered again with the falling edge of the gate pulse, and recovery of degraded I_{DLIN} is measured. Once again, the voltage shift for BTI recovery can be calculated using $\Delta V = -\Delta I_{DLIN}/I_{DLIN0} * (V_{G\text{-}REC} - V_{T0})$, where ΔI_{DLIN} is $I_{DLIN} - I_{DLIN0}$ and I_{DLIN0} is recorded at $V_G = V_{G\text{-}REC}$ before stress. Mobility correction is used to convert ΔV to ΔV_T using the method of [12] as discussed in the following section.

Figure 2.6b plots time evolution of ΔV_T recovery measured using the OTF technique after NBTI stress in PNO-SiON p-MOSFETs for different t_{STR} values. It is important to note that ΔV_T starts to recover as soon as V_G is reduced from $V_{G\text{-}STR}$

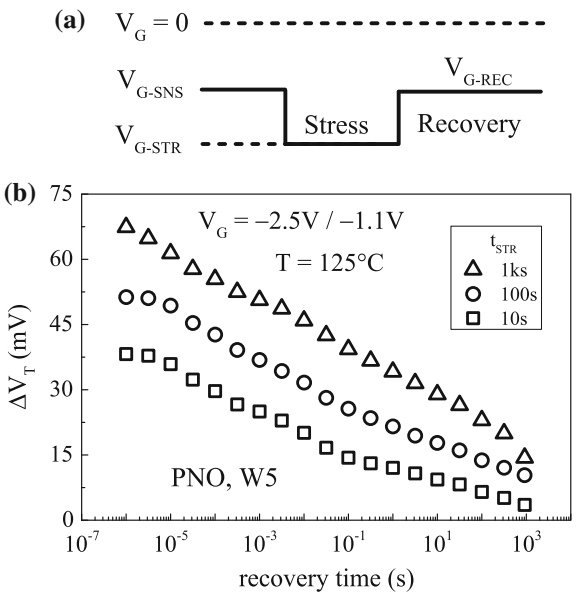

Fig. 2.6 OTF recovery measurement: **a** applied gate voltage sequence, **b** time evolution of measured and mobility corrected ΔV_T recovery after NBTI stress for different stress time in PNO SiON p-MOSFET

to $V_{G\text{-}REC}$. This suggests the necessity of using ultra-fast setup for BTI characterization using MSM or OSDD methods, and is discussed later in this chapter. More BTI recovery results measured using OTF method for different devices and stress conditions are presented and analyzed in Chap. 6 of this book.

2.3 Mobility Degradation

As discussed in Chap. 1, I_{DLIN} degradation especially during NBTI stress is caused by degradation in both V_T and μ_{eff}, the later results in transconductance (g_m) degradation, refer to Chap. 1, Fig. 1.16. Therefore, the contribution of $\Delta\mu_{eff}$ on ΔI_{DLIN} needs to be determined for accurate determination of time evolution of ΔV_T from OTF technique.

The drain current of a MOSFET in above threshold and linear regime of operation, I_{DLIN}, is given by the following equation [15]:

$$I_{DLIN} = \mu_{eff} C_{OX} \frac{W}{L}\left(V_G - V_T - m\frac{V_D}{2}\right)V_D \tag{2.1}$$

In (2.1), μ_{eff} is effective mobility, C_{OX} is gate capacitance, W and L are device width and channel length respectively, V_G and V_D are gate and drain biases respectively, and m is calculated using the subthreshold I_D versus V_G characteristics. For higher gate overdrive ($V_G - V_T$), the effective mobility can be expressed using a simple first order relation as follows [10, 12]:

$$\mu_{eff} = \frac{\mu_0}{1 + \theta(V_G - V_T)} \tag{2.2}$$

In (2.2), μ_0 is low field mobility and θ is high field reduction factor. The variation in I_{DLIN} due to BTI can be obtained by differentiating (2.1) as follows [12]:

$$-\left|\frac{\Delta I_{DLIN}}{I_{DLIN0}}\right| = \frac{\Delta\mu_{eff}}{\mu_{eff0}}\frac{V_G - V_T - m\frac{V_D}{2}}{V_G - V_{T0} - m_0\frac{V_D}{2}} + \frac{\mu_{eff}}{\mu_{eff,0}}\frac{-|\Delta V_T| - \frac{V_D}{2}|\Delta m|}{V_G - V_{T0} - m_0\frac{V_D}{2}} \tag{2.3}$$

Note that besides ΔV_T, ΔI_{DLIN} is also impacted by $\Delta\mu_{eff}$ ($=\mu_{eff} - \mu_{eff,0}$) and therefore by $\Delta\mu_0$ ($=\mu_0 - \mu_{0,0}$) and $\Delta\theta$ ($=\theta - \theta_0$), as well as by Δm ($=m - m_0$), where $\mu_{eff,0}$, θ_0 and m_0 are corresponding pre-stress values. The simplified expression used in Sect. 2.2, $\Delta V = -\Delta I_{DLIN}/I_{DLIN0} * (V_{G\text{-}STR} - V_{T0})$, is derived from (2.3) by ignoring $\Delta\mu_{eff}$ and Δm, and therefore, ΔV needs to be corrected, especially for NBTI stress, to obtain proper, "mobility corrected" ΔV_T [12].

The mobility and m factor degradation can be assessed by measuring I_{DLIN} versus V_G sweeps before and during logarithmically spaced intervals of NBTI stress. Pre- and post-stress I_{DLIN} versus V_G data in above threshold at high gate overdrive can be used to obtain μ_{eff} versus ($V_G - V_T$) data, where the mobility

expression of (2.2) remains valid, and can be fitted to obtain μ_0 and θ, and hence, time evolution of $\Delta\mu_0$ and $\Delta\theta$ can be determined. Time evolution of Δm can be obtained from subthreshold slope degradation and that of ΔV_T from V_T extracted using conventional "peak g_m" method.

Figure 2.7 shows normalized $\Delta\mu_0$, $\Delta\theta$ and Δm as a function of ΔV_T for NBTI stress in SiON p-MOSFETs. The magnitude of μ_0 and θ reduce and that of m increases after NBTI stress. Moreover, $\Delta\mu_0$, $\Delta\theta$ and Δm are linearly correlated to ΔV_T, and this correlation can be used to construct a "mobility correction" algorithm as discussed later. Note that for a particular ΔV_T, much larger variation is usually observed for $\Delta\mu_0$ and $\Delta\theta$ than that for Δm, and therefore, correction due to Δm has negligible impact and can be ignored as discussed in [12].

It is important to remark that the correlation of $\Delta\mu_{eff}$ to ΔV_T depends on θ_0, the pre-stress slope of μ_{eff} versus $(V_G - V_T)$ data, refer to (2.2). Note that when compared at fixed $(V_G - V_T)$, obtained μ_{eff} always reduces after NBTI stress, and therefore, $\Delta\mu_{eff} < 0$. However, actual NBTI measurements are done at fixed V_G, and for devices having high θ_0, reduction in μ_0 can be overcompensated by reduction in θ_0 and $(V_G - V_T)$ after NBTI stress, which can result in $\Delta\mu_{eff} > 0$, refer to (2.2) [12].

As an example, Fig. 2.8 shows normalized $\Delta\mu_{eff}$ as a function of ΔV_T for two different SiON p-MOSFETs having (a) low and (b) high pre-stress θ_0. Note that μ_{eff} is extracted using (2.1) from pre- and post-stress I_{DLIN} versus V_G characteristics at fixed sense bias of $V_G = V_{G\text{-}SNS}$, while V_T is extracted by using the peak g_m method. The magnitude of $|\Delta\mu_{eff}|$ increases with increase in ΔV_T due to NBTI stress for both devices. However, the device with low θ_0 shows $\Delta\mu_{eff} < 0$ as expected, while the device with high θ_0 shows $\Delta\mu_{eff} > 0$ due to overcompensation effect discussed above. It is also important to note that for a particular ΔV_T, $|\Delta\mu_{eff}|$ reduces with

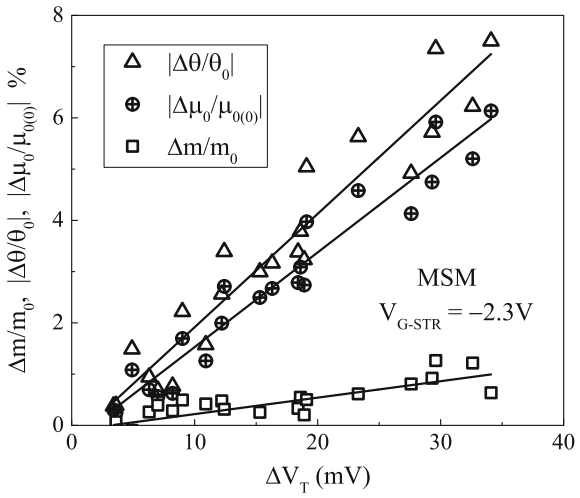

Fig. 2.7 Correlation of degradation in parameters m, μ_0 and θ shown in (2.1) and (2.2) to ΔV_T for NBTI stress in PNO SiON p-MOSFET. Data obtained using I–V sweeps in slow MSM method

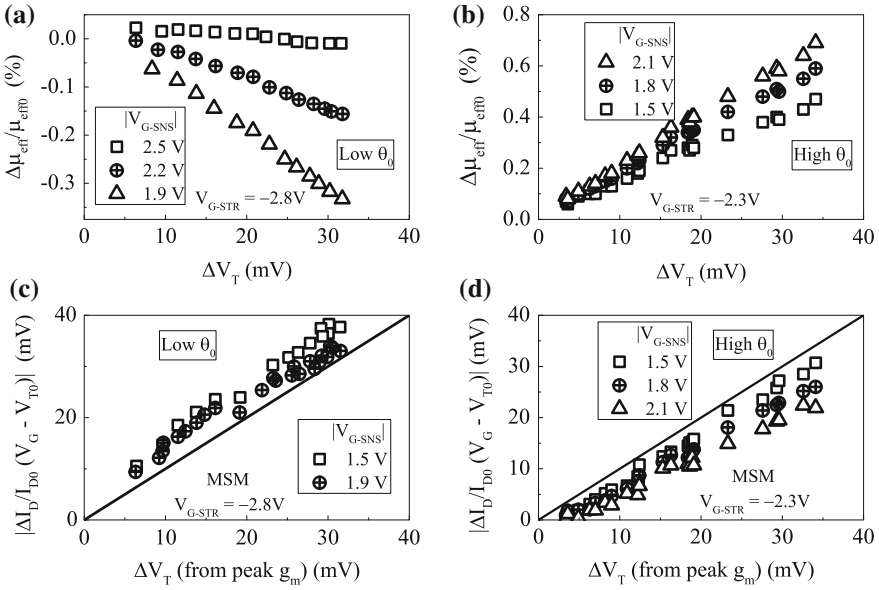

Fig. 2.8 Correlation of normalized mobility degradation measured using different sense V_G to ΔV_T for NBTI stress in PNO SiON p-MOSFET having **a** low θ_0 and **b** high θ_0. Correlation of corresponding mobility uncorrected voltage shift using simple I_{DLIN} expression to ΔV_T for **c** low θ_0 and **d** high θ_0 devices. ΔV_T is obtained using peak g_m method

increase in $|V_{G\text{-}SNS}|$ for low θ_0 device, while for high θ_0 device, $|\Delta\mu_{eff}|$ increases with increase in $|V_{G\text{-}SNS}|$ as shown in Fig. 2.8a, b.

Note that the magnitude and sign of $\Delta\mu_{eff}$ versus ΔV_T correlation would impact ΔI_{DLIN} versus ΔV_T correlation as per (2.3), since ΔI_{DLIN} is also obtained from I_{DLIN} measurement before and after stress at a fixed V_G. As an illustration, Fig. 2.8 plots the correlation of $\Delta V = -\Delta I_{DLIN}/I_{DLIN0} * (V_{G\text{-}SNS} - V_{T0})$ versus ΔV_T at different sense bias $V_{G\text{-}SNS}$ for SiON p-MOSFETs with (c) low and (d) high pre-stress θ_0; the $\Delta\mu_{eff}$ versus ΔV_T correlation of these devices are shown, respectively in Fig. 2.8a, b. Pre- and post-stress V_T and hence ΔV_T are obtained using the peak g_m method, and 1:1 correlation is shown by a solid line. As expected, ΔV is not equal to ΔV_T due to mobility degradation effects. For a particular ΔV_T, the device with lower θ_0 shows higher ΔV than 1:1 correlation line, see Fig. 2.8c, and the error in ΔV increases with reduction in $|V_{G\text{-}SNS}|$. On the other hand, the device with higher θ_0 has lower ΔV than 1:1 correlation line, see Fig. 2.8d, while the error in ΔV increases with increase in $|V_{G\text{-}SNS}|$. Therefore, if mobility correction is not taken into consideration, OTF measurements can result in higher or lower than actual NBTI degradation, and the error will be dependent on $V_{G\text{-}SNS}$, which is $V_{G\text{-}STR}$ for OTF method. Therefore, not only the NBTI magnitude but also the $V_{G\text{-}STR}$ or E_{OX} dependence of measured degradation will get corrupted, which will result in incorrect extrapolated degradation at use condition.

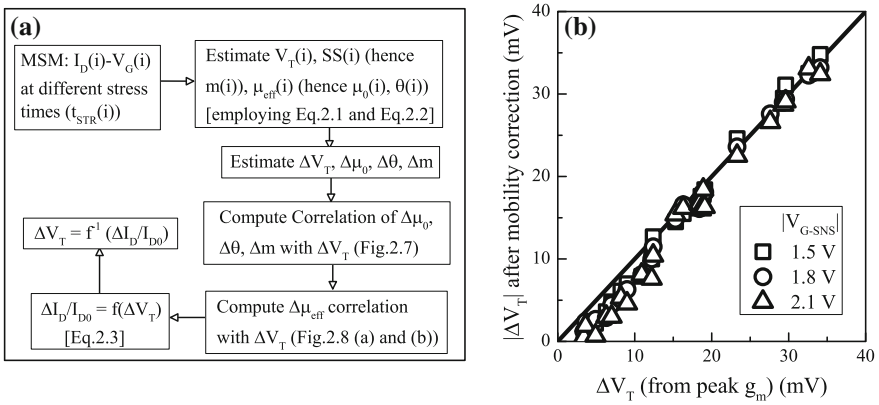

Fig. 2.9 **a** Algorithm for mobility correction of OTF data. **b** Correlation of mobility corrected calculated ΔV_T from measured I_{DLIN} at different sense V_G to measured ΔV_T using peak g_m method for NBTI stress in PNO SiON p-MOSFET

I_{DLIN} versus V_G measurements before and during NBTI stress interruptions can be used to calculate pre-stress $\mu_{0,0}$, θ_0 and m_0 values, as well as stress induced $\Delta\mu_0$, $\Delta\theta$ and Δm versus ΔV_T correlation as shown in Fig. 2.7. Moreover, (2.2) can be used to calculate $\Delta\mu_{eff}$ versus ΔV_T correlation as shown in Fig. 2.8a, b, and finally, (2.3) can be used to calculate "mobility corrected" ΔV_T [12]. Figure 2.9a illustrates this correction algorithm, and Fig. 2.9b plots "mobility corrected" ΔV_T from one spot ΔI_{DLIN} measurements at different $V_{G\text{-}SNS}$ as a function of ΔV_T obtained using the conventional peak g_m method. Note that once mobility degradation is taken into account using (2.3), 1:1 correlation is observed between "peak g_m" ΔV_T and ΔV_T obtained from ΔI_{DLIN} measured at different $V_{G\text{-}SNS}$. All UF-OTF NBTI data presented elsewhere in this book are corrected for mobility degradation using the method proposed in this section.

2.4 Ultra-Fast Measure-Stress-Measure (UF-MSM) Method

Ultra-Fast MSM (UF-MSM) method, illustrated in Fig. 2.10, relies on performing complete I_{DLIN} versus V_G sweeps in few microseconds before and during logarithmically spaced interruptions of BTI stress [6, 7]. V_T is extracted before and after stress from measured I_D–V_G data by using the conventional "peak g_m" method [15], and hence time evolution of ΔV_T can be obtained. Note that similar to conventional MSM method, the UF-MSM method also suffers from potential recovery issues since V_G is lowered from stress for measurement, although the recovery would be substantially lower, if not almost negligible, as stress interruption time is drastically minimized. Note that a tradeoff exists between measurement speed and accuracy,

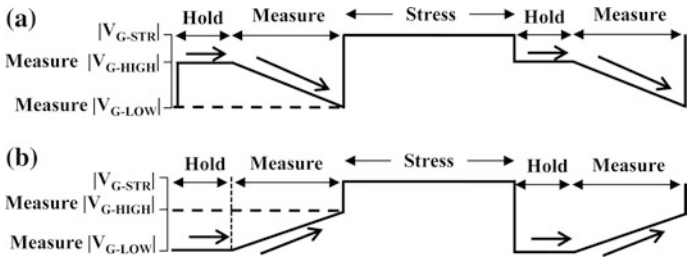

Fig. 2.10 Schematic of UF-MSM method for DC NBTI measurements. Arrows represent directions of V_G sweep for *I–V* measurement. Measurement delay is sum of hold time and sweep time. Default hold time is 50 ns and sweep time is 10 μs unless mentioned otherwise

and commercial setups are now available to reliably perform both DC and AC BTI experiments in MSM mode with ~10 μs delay [16]. The UF-MSM method has been extensively used to characterize DC and AC degradation during NBTI and PBTI stress in p- and n-channel High-K Metal Gate (HKMG) MOSFETs, respectively [16–19]; results are shown in Chap. 1 and further analyzed in later chapters of this book. This section describes the implementation details of the UF-MSM method. Although early implementation of UF-MSM method was made using custom setups [6], all results presented in this book are measured using commercially available instrument [16]. Although UF-MSM method has been used to characterize NBTI stress in both SiON and HKMG devices and PBTI stress in HKMG devices, results presented in this section are obtained from NBTI stress in p-channel HKMG MOSFETs.

2.4.1 Measurements During DC Stress

As shown in Fig. 2.10, V_G sweeps for I_{DLIN} measurements can be made either from high to low (H to L) or from low to high (L to H) values,[1] however, note that it is important to keep identical sweep direction for pre- and post-stress measurements. Figure 2.11 plots pre- and post-stress (a) I_{DLIN} and (b) g_m versus V_G data obtained using H to L and L to H sweeps. Note that due to recovery issues, the H to L and L to H sweep measurements are done on identical but separate devices, stressed at identical V_G and T. Identical pre-stress measured values for different sweep directions (small discrepancy is due to device to device variation) verify that the measurement is free from hysteresis issues, which can be either due to inaccurate setup or due to bad gate insulator quality of the device under test. The degradation, i.e., the difference between pre- and post-stress data is different for different sweep directions and is discussed below.

[1]For NBTI stress, low and high values correspond to the magnitude of gate pulse.

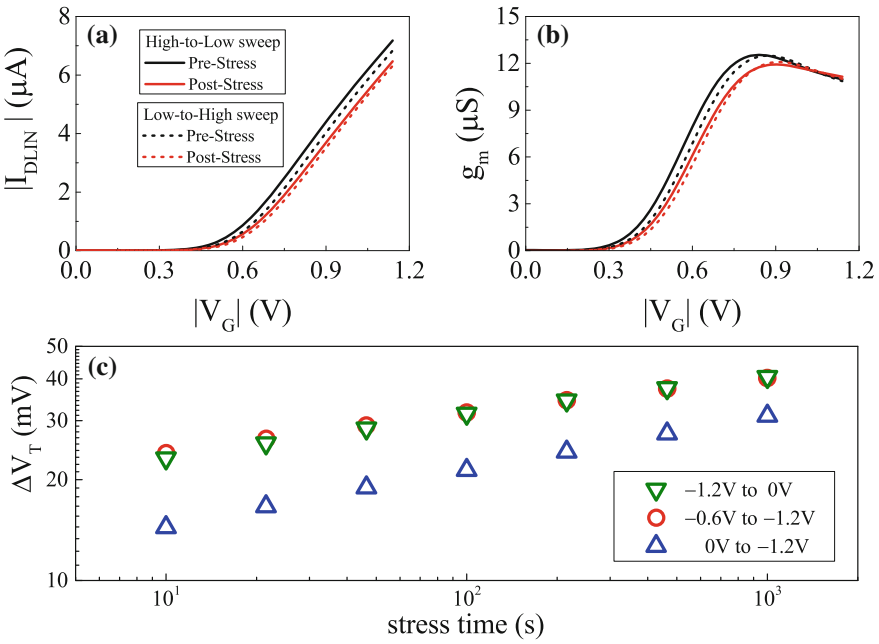

Fig. 2.11 UF-MSM measured **a** I_{DLIN} versus V_G and **b** g_m versus V_G characteristics before and after NBTI stress in HKMG p-MOSFET. Data shown for different sweep directions for *I–V* measurements. **c** Time evolution of UF-MSM measured ΔV_T for different sweep directions and ranges of sweep bias

As mentioned before, V_T is extracted using the conventional "peak g_m" method, where a tangent in drawn on the I_{DLIN} versus V_G curve at the V_G value corresponding to the peak g_m point. *I–V* characteristics are measured before and during logarithmically spaced interruptions in BTI stress, and ΔV_T time evolution is obtained. Figure 2.11c plots time evolution of ΔV_T obtained at identical NBTI stress conditions but using different sweep directions. Note that ΔV_T shows power-law time dependence with exponent n at longer t_{STR}. However, the magnitude of ΔV_T and its exponent n show different values for H to L and L to H sweeps for identical hold and sweep time. Higher ΔV_T and lower n are observed for H to L when compared to L to H sweep as shown; n being obtained using linear regression of measured data in t_{STR} range to 10 s to 1 Ks.

Although H to L and L to H sweeps are taken using identical voltage range (0 to −1.2 V) and sweep time, the hold before sweep is at higher $|V_G|$ for H to L and at 0 V for L to H sweep, see Fig. 2.10. Since recovery accelerates at lower $|V_G|$, refer to Chap. 1, Fig. 1.22, hold at 0 V for L to H sweep has larger recovery as compared to hold at higher $|V_G|$ for H to L sweep, and therefore, the former results in lower ΔV_T magnitude and higher exponent n as shown. Note that for L to H sweep, it is sufficient to record I_{DLIN} from V_G values just below the peak g_m point for extraction

of V_T using the tangent method, and it is not necessary to record I_{DLIN} all the way from $V_G = 0$ V. Figure 2.11c also plots the time evolution of ΔV_T for L to H sweep from −0.6 to −1.2 V, which, for identical $V_{G\text{-}STR}$ and T, shows larger ΔV_T and lower n compared to L to H sweep from 0 to −1.2 V, due to relatively higher hold V_G and correspondingly lower recovery. Note that the H to L sweep can also be done from −1.2 V up to −0.6 V instead of going all the way down to 0 V to reduce measurement time, t_M. It should be remarked that the non-zero lower limit of V_G for L to H and/or H to L sweep should be suitably chosen such that it crosses the V_G value corresponding to the peak g_m point, through which a tangent needs to be drawn to estimate V_T.

Figure 2.12 shows the impact of hold time on time evolution of ΔV_T for (a) H to L sweep from −1.2 to −0.6 V, as well as L to H sweep from (b) −0.6 to −1.2 V and (c) 0 to −1.2 V. The power-law time exponents extracted using linear regression of measured data in t_{STR} range to 10 s to 1 Ks are shown as a function of hold delay in Fig. 2.12d for different sweep directions. All measurements are done on identical but separate devices, stressed at identical $V_{G\text{-}STR}$ and T. Note that H to L sweep shows largest ΔV_T, smallest n and negligible impact of hold time, while for L to H sweep, ΔV_T reduces and n increases with increase in hold time, larger ΔV_T and smaller n are obtained for hold at −0.6 V compared to 0 V. This is due to negligible

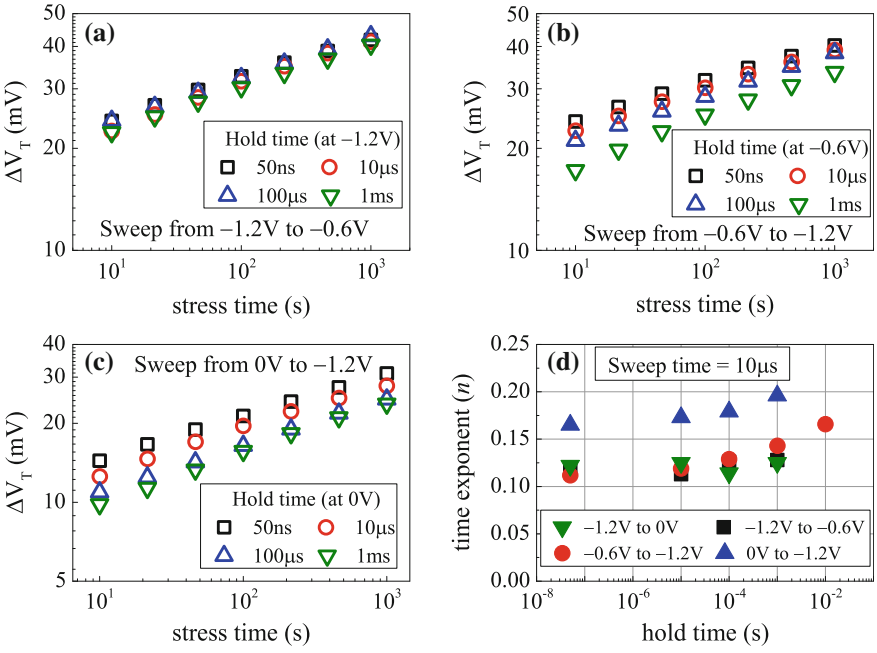

Fig. 2.12 **a–c** Impact of hold time on time evolution of UF-MSM measured ΔV_T for NBTI stress in HKMG p-MOSFETs for different sweep conditions. **d** Measured power-law time exponent n versus hold time for different sweep conditions

ΔV_T recovery at −1.2 V, and much larger recovery at 0 V compared to −0.6 V. Therefore, a non-zero lower limit of V_G sweep reduces the impact of recovery especially for L to H sweep, and the overall measurement time can also be reduced due to reduction in range of sweep V_G.

Besides hold time, the sweep time also impacts time evolution of measured ΔV_T. Higher recovery at lower $|V_G|$ influences both H to L and L to H sweep based *I*–*V* measurements, even for sweeps with relatively higher, non-zero lower limit of V_G. Figure 2.13 plots ΔV_T time evolution for (a) H to L sweep from −1.2 to −0.6 V, as well as L to H sweep from (b) −0.6 to −1.2 V and (c) 0 to −1.2 V. The power-law time exponents obtained using linear regression of measured data in t_{STR} range to 10 s to 1k s are shown as a function of sweep delay in Fig. 2.13d for different sweep directions. Higher sweep time results in lower ΔV_T and higher n for both H to L as well as L to H sweep. However contrary to hold time, the sweep time has relatively larger impact on H to L compared to L to H sweep. This is due to larger recovery at larger sweep delay for H to L compared to L to H sweep, as the later part of the sweep goes towards lower $|V_G|$ for H to L and towards higher $|V_G|$ for L to H sweep. Therefore, the sweep direction, hold time and sweep time affect ΔV_T time evolution and need to be carefully chosen for UF-MSM measurements.

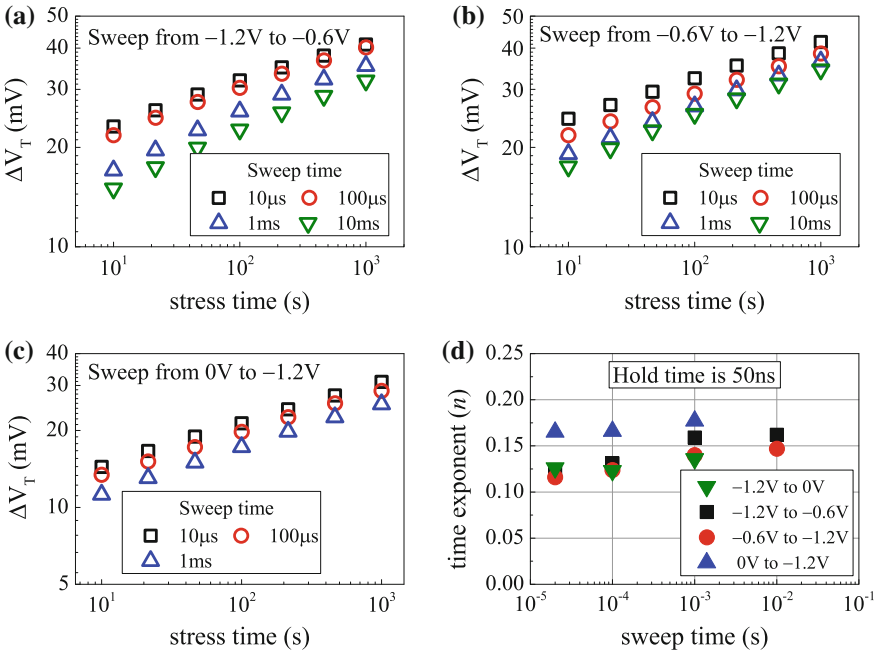

Fig. 2.13 **a–c** Impact of sweep time and sweep direction on UF-MSM measured ΔV_T time evolution for NBTI stress in HKMG p-MOSFETs. **d** Measured power-law time exponent n versus sweep time for different sweep conditions

2.4.2 Measurements During AC Stress

Figure 2.14 shows the gate voltage sequence for AC stress and measurement. AC pulse having a particular frequency (f) and duty cycle (PDC) is applied to the gate for stress, pulse $V_{G\text{-}HIGH}$ and pulse $V_{G\text{-}LOW}$ being the stress high and low values, respectively. The device gets stressed during pulse high phase, and recovers during pulse low phase. The pulse low is usually kept at 0 V, although non-zero $V_{G\text{-}LOW}$ can also be used to study the impact of V_G on recovery for AC stress. I–V sweeps are measured before and during logarithmically spaced interruptions in BTI stress to extract time evolution of ΔV_T. Measure $V_{G\text{-}HIGH}$ and measure $V_{G\text{-}LOW}$ denote the range of V_G sweep for I–V measurements. Note that AC stress can be interrupted for I–V measurements after the completion of last half cycle or last full cycle, defined respectively as Mode-A and Mode-B AC stress. It is important to note that AC BTI results shown in Chap. 1 are from Mode-B stress. Similar to DC stress, the I–V curves can be measured using H to L or L to H directions of V_G sweep for AC stress, although H to L direction is used in this book unless mentioned otherwise. It is important to use identical sweep parameters, such as sweep direction, hold time and sweep time for accurate estimation of the ratio of AC to DC degradation.

Figure 2.15a plots time evolution of ΔV_T for Mode-A and Mode-B AC stress at identical PDC, f, $V_{G\text{-}HIGH}$ (=$V_{G\text{-}STR}$) and $V_{G\text{-}LOW}$ (=$V_{G\text{-}REC}$) and measured using H to L gate sweep from −1.2 and −0.6 V. Time evolution of ΔV_T for AC stress also shows power-law dependence at longer t_{STR}. However, Mode-A stress shows higher ΔV_T magnitude and lower time exponent n when compared to Mode-B stress, which is due to higher recovery for the later as measurements are done after the pulse low phase. The difference between Mode-A and Mode-B AC stress is analyzed in detail later in this book.

Figure 2.15 also shows the impact of (b) PDC and (c) f of the AC gate pulse on measured ΔV_T at fixed t_{STR} for Mode-A and Mode-B stress at fixed $V_{G\text{-}HIGH}$ but

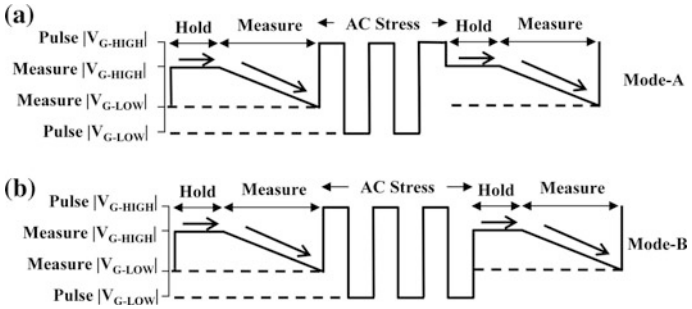

Fig. 2.14 Schematic of UF-MSM method for AC NBTI measurements. Mode-A and Mode-B represents start of I–V sweep after last half and full cycle, respectively. Identical sweep directions should be used for both modes, which should also be identical to that used for DC stress

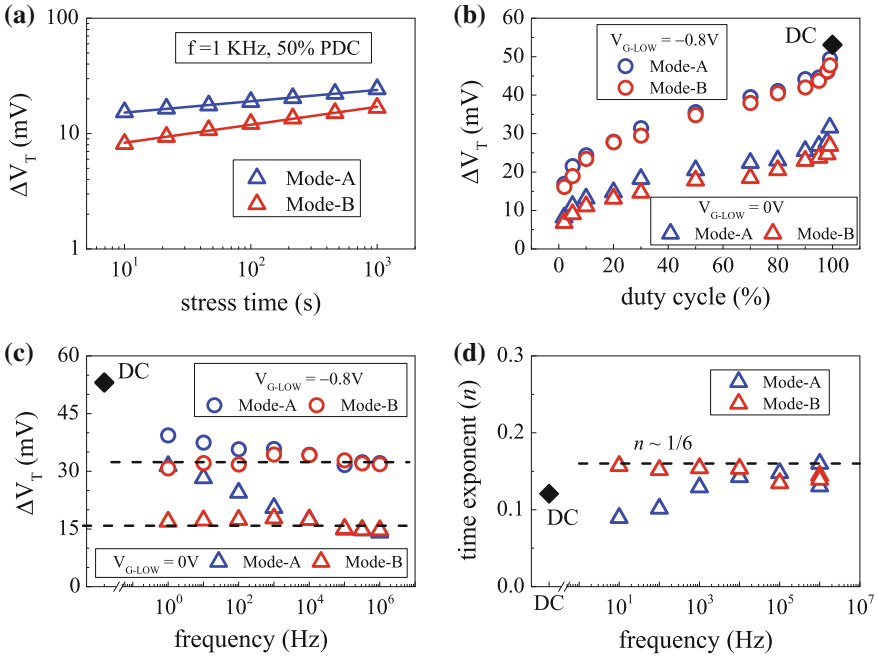

Fig. 2.15 AC NBTI measurements in HKMG p-MOSFETs using UF-MSM method: **a** Time evolution of ΔV_T for Mode-A and Mode-B stress. Fixed time ΔV_T for Mode-A and Mode-B stress using AC pulse having different pulse low bias, as a function of **b** duty cycle and **c** frequency. **d** Power-law time exponent n versus frequency for Mode-A and Mode-B AC stress. DC data are shown as reference

different $V_{G\text{-}LOW}$, the corresponding DC stress value is shown as reference. Measurements are done using H to L V_G sweep from −1.2 to −0.6 V. Note that ΔV_T for DC and AC stress are obtained at fixed t_{STR}, which, for AC stress includes both pulse on and off time, and the net stress or pulse on time for AC stress would depend on PDC of the gate pulse. Higher ΔV_T is obtained at higher $|V_{G\text{-}LOW}|$ for all PDC and f, which is consistent with lower recovery at higher $|V_G|$. ΔV_T increases with increase in PDC due to increase in net pulse high or stress time. The PDC dependence shows a typical "S" shaped characteristics, and a kink or jump is observed between high PDC AC and DC stress. Lower kink in ΔV_T is observed for higher $|V_{G\text{-}LOW}|$ and vice versa, which will be explained later in this book. Measured ΔV_T for Mode-B stress is independent of f for different $V_{G\text{-}LOW}$, although larger AC to DC ratio is observed for higher $|V_{G\text{-}LOW}|$ and vice versa. On the other hand, ΔV_T for Mode-A stress is higher when compared to Mode-B ΔV_T especially at lower f, however, the former reduces with increase in f and merges with the later at higher f and also shows f independence.

The impact of AC pulse f on measured power-law time exponent n for Mode-A and Mode-B stress is shown in Fig. 2.15d. As mentioned before, n is extracted by

using linear regression of measured data in t_{STR} range of 10 s to 1k s. Note that Mode-B stress shows universal $n \sim 1/6$ value and f independence of n. On the other hand, Mode-A stress results in lower n compared to Mode-B especially at lower f, however, n for the former increases with increase in f and merges with the later at higher f and shows f independence. It is important to note the f independence of ΔV_T and n for Mode-B stress. Moreover, the observed f dependence of ΔV_T and n for Mode-A stress especially at lower f is also of interest; ΔV_T reduces but n increases with increase in f. These aspects will be explained in Chap. 6.

The impact of hold and sweep delay on AC stress is also of interest. Figure 2.16 shows ΔV_T time evolution measured using L to H V_G sweep from −0.6 to −1.2 V for (a, c) Mode-A and (b, d) Mode-B AC stress, for different (a, b) hold time and (c, d) sweep time. Similar to DC, Mode-A AC stress is also affected by both hold and sweep time; lower ΔV_T magnitude and higher power-law time exponent n are observed for higher measurement delay. However, the sweep and hold delay have negligible impact on Mode-B AC stress as shown. This can be explained by noting that I–V sweeps for Mode-B stress are done after the pulse off phase where recovery takes place, and therefore, additional delay has negligible impact on measured ΔV_T time evolution. However, this is not the case for Mode-A stress as I–V sweeps are done after pulse high phase, and hence, delay influences measured time evolution of ΔV_T as shown. Therefore, it is necessary to use UF-MSM for accurate estimation of DC to AC ratio for Mode-A and Mode-B AC stress.

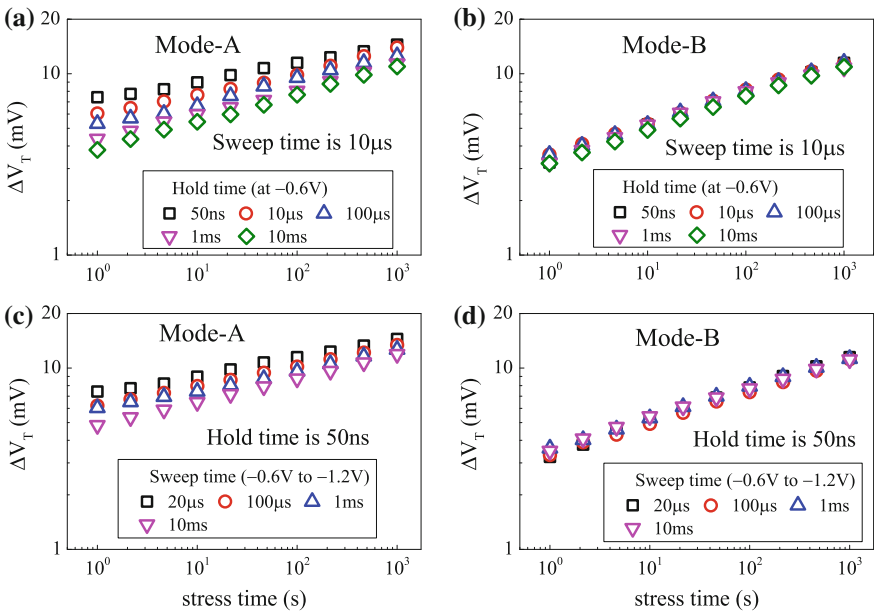

Fig. 2.16 Impact of (**a, b**) hold time and (**c**, **d**) sweep time for UF-MSM measured ΔV_T time evolution during AC (**a**, **c**) Mode-A and (**b**, **d**) Mode-B NBTI stress in HKMG p-MOSFETs

2.5 One Spot Drop Down (OSDD) Method

In OSDD technique, illustrated in Fig. 2.17, BTI stress is interrupted and V_G is reduced from $V_{G\text{-}STR}$ to a suitable sense bias ($V_{G\text{-}SNS}$) to measure I_{DLIN} in time t_M. Similar to MSM, the OSDD technique would also suffer from recovery issues, however, it takes much shorter time to measure a single spot I_D than full I_D–V_G sweep and hence recovery can be minimized. Although specialized commercial setups are now available to reliably perform full I_D–V_G sweeps in few microseconds, OSDD has been particularly useful in the past when such specialized instruments were not available. As shown in Fig. 2.17, once post-stress I_{DLIN} is measured at $V_{G\text{-}SNS}$, it can be compared to pre-stress I_D–V_G sweep to determine BTI degradation. In the vertical shift method, $\Delta V = -\Delta I_{DLIN}/I_{DLIN0} * (V_{G\text{-}SNS} - V_{T0})$ can be estimated by noting the difference in I_{DLIN} between pre- and post-stress at $V_{G\text{-}SNS}$, and in the absence of mobility variation, ΔV equals ΔV_T. Note that the absence of mobility degradation implies parallel I_D–V_G curves before and after stress. In the lateral shift method, the voltage corresponding to post-stress I_{DLIN} is noted from the pre-stress I_D–V_G curve, which is denoted as $V_{G\text{-}PST}$ as shown in Fig. 2.17, and the difference between $V_{G\text{-}PST}$ and $V_{G\text{-}SNS}$ is used to calculate ΔV_{LS}, which becomes equal to ΔV_T in the absence of mobility degradation.

In the technique proposed in [20], I_{DLIN} is measured at $V_{G\text{-}SNS} = V_{T0}$, where V_{T0} is pre-stress threshold voltage of the device and lateral shift method is used. Note that the lateral shift method assumes parallel I_D–V_G curves before and after stress and hence no mobility degradation, which is a fair assumption for PBTI stress as discussed in Chap. 1, and also at lower $V_{G\text{-}SNS}$ close to V_{T0} for NBTI stress, which will be shown later in this section. However, as recovery magnitude increases with increase in the difference between $V_{G\text{-}STR}$ and $V_{G\text{-}SNS}$ [16, 21], a low $V_{G\text{-}SNS}$ readout would suffer from recovery artifacts, shown in [22] and discussed later in this section, unless I_{DLIN} is recorded in few microseconds. There exists a tradeoff

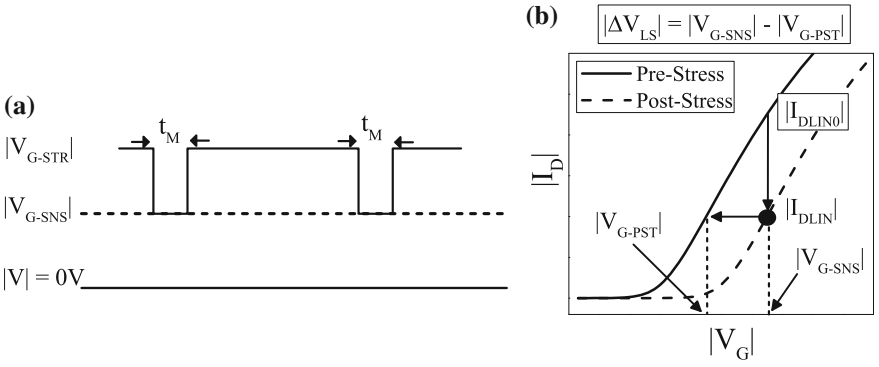

Fig. 2.17 **a** Schematic of OSDD method. **b** Measured I_{DLIN} versus V_G sweep before stress and one spot I_{DLIN} measurement after stress. The lateral and vertical shift methods for ΔV_T extraction are illustrated

between $V_{G\text{-}SNS}$ and drop down delay to minimize the impact of recovery, refer to Chap. 1, Fig. 1.29 for details. Alternatively, especially for relatively slower measurements, the recovery can be reduced by dropping down to $V_{G\text{-}SNS} = V_{DD}$ for I_{DLIN} measurement, where V_{DD} is the operating voltage of the device [23, 24]. ΔV_T can be calculated either using the lateral shift or vertical shift methods. However, measured ΔI_{DLIN} at high $V_{G\text{-}SNS}$ is affected by both ΔV_T and $\Delta \mu_{eff}$, especially for NBTI stress, and the impact of $\Delta \mu_{eff}$ needs to be corrected as described later in this section. Although OSDD method has been used to characterize NBTI in SiON and HKMG p-MOSFETs and PBTI in HKMG n-MOSFETs [20, 24], in this chapter, results are only shown for NBTI in HKMG p-MOSFETs.

Figure 2.18 plots the time evolution of measured ΔV_{LS} during NBTI stress in HKMG p-MOSFETs by using the lateral shift method, for (a) fixed t_M and different $V_{G\text{-}SNS}$ and (b) fixed $V_{G\text{-}SNS}$ and different t_M. The same I_{DLIN} data can be used to calculated ΔV using the vertical shift method using the equation shown above, and obtained time evolution is shown in Fig. 2.18 for (c) fixed t_M and different $V_{G\text{-}SNS}$ and (d) fixed $V_{G\text{-}SNS}$ and different t_M. The impact of $V_{G\text{-}SNS}$ on power-law time exponent n, extracted by linear regression of measured data in t_{STR} range of 10 s to 1 Ks, is plotted in Fig. 2.18 for (e) lateral and (f) vertical shift methods, for different measurement delay t_M. Finally, the time evolution of ΔV_T obtained using 10 μs UF-MSM method is shown in Fig. 2.18a, c, and power-law time exponent n extracted from MSM data for different measurement delay is shown in Fig. 2.18e, f as reference.

Note that when compared to the UF-MSM method for a particular t_M, the magnitude of ΔV_{LS} from the lateral shift and ΔV from the vertical shift OSDD method increases with increase in $V_{G\text{-}SNS}$ as shown respectively in Fig. 2.18a, c, and the corresponding time exponent n reduces as shown in Fig. 2.18e, f. Since recovery is negligible due to the use of small t_M, this is an artifact of mobility degradation at higher $V_{G\text{-}SNS}$ and is discussed below. However, for a particular $V_{G\text{-}SNS}$, the magnitude of ΔV_{LS} and ΔV reduces and corresponding time exponent n increases with increase in t_M, as shown in Fig. 2.18b–f, which can be attributed to recovery artifacts as discussed in the previous section.

To illustrate the impact of mobility degradation, Fig. 2.19a shows I_{DLIN} versus V_G characteristics measured using UF-MSM method in HKMG p-MOSFETs before and after NBTI stress. Pre- and post-stress V_T is determined using the peak g_m method, and the post-stress I–V curve is shifted by ΔV_T to align with the pre-stress curve. Note that although the shifted post-stress curve aligns with the pre-stress curve at lower V_G, the curves do not match at higher V_G, and the post-stress curve is below the pre-stress curve due to additional mobility degradation. The mobility impact for vertical shift method is assessed by $\Delta V^* = -\Delta I_{DLIN}/I_{DLIN0} * (V_{G\text{-}SNS} - V_{T0})$, where ΔI_{DLIN} is difference between pre-stress and ΔV_T shifted post-stress curve. In a similar manner, the mobility impact for lateral shift method is assessed by noting ΔV_{LS}^*, the lateral difference between pre-stress and ΔV_T shifted post-stress curves as shown in Fig. 2.19a. Similar analysis can be done at different $V_{G\text{-}SNS}$ and for post-stress I–V measured at different t_{STR}.

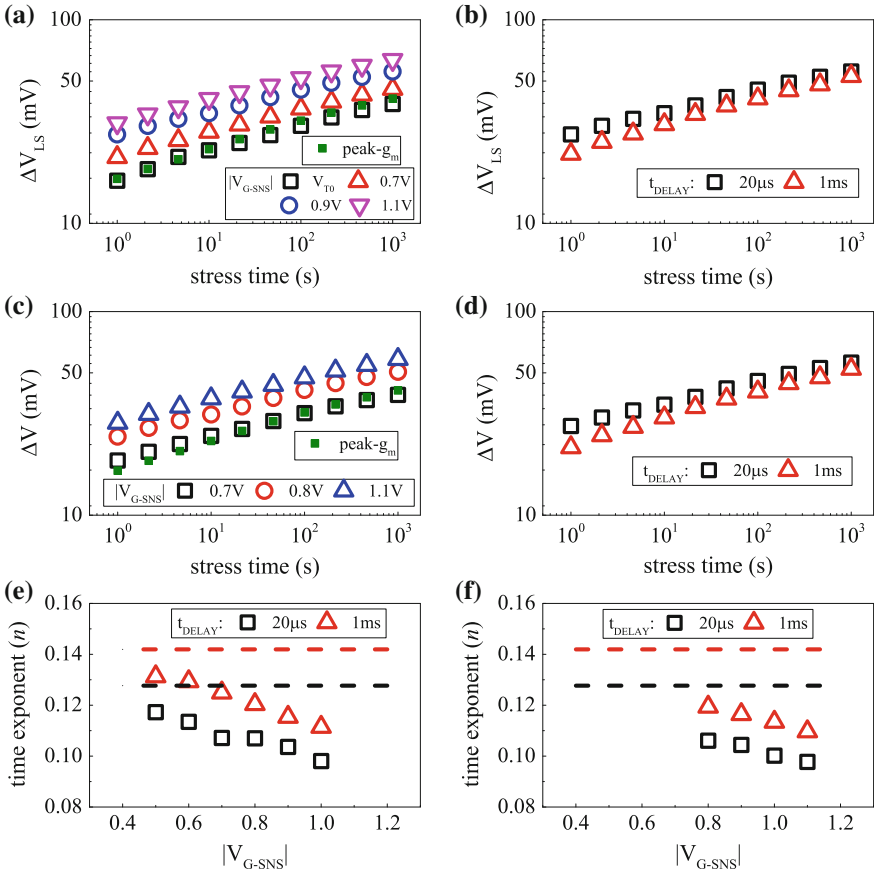

Fig. 2.18 OSDD measurements: Time evolution of (**a**, **b**) ΔV_{LS} using lateral shift and (**c**, **d**) ΔV using vertical shift methods, for different (**a**, **c**) sense bias and (**b**, **d**) measurement delay, for NBTI stress in HKMG p-MOSFETs. Impact of sense V_G on power-law time exponent n for **e** lateral shift and **f** vertical shift method for different measurement delay. UF-MSM data obtained using peak g_m method are shown as reference in (**a**, **c**). *Lines* in (**e**, **f**) represent UF-MSM measured time exponent n for different delay

The resulting time evolution of ΔV_{LS}* and ΔV* at different $V_{G\text{-}SNS}$ is shown respectively in Fig. 2.19b, c. It is evident that post-stress I_{DLIN} is affected by both ΔV_T and $\Delta\mu_{eff}$ especially at higher $V_{G\text{-}SNS}$. The contribution due to $\Delta\mu_{eff}$, manifested as ΔV_{LS}* or ΔV* for lateral or vertical shift methods respectively, increases with increase in $V_{G\text{-}SNS}$. Therefore, ΔV_{LS} and ΔV, extracted from lateral or vertical shift in I_D–V_G curve after stress, increases at higher $V_{G\text{-}SNS}$ due to additional contribution respectively from ΔV_{LS}* and ΔV*. Although time evolution of ΔV_{LS}* and ΔV* shows power-law time dependence, they have lower time exponent n compared to

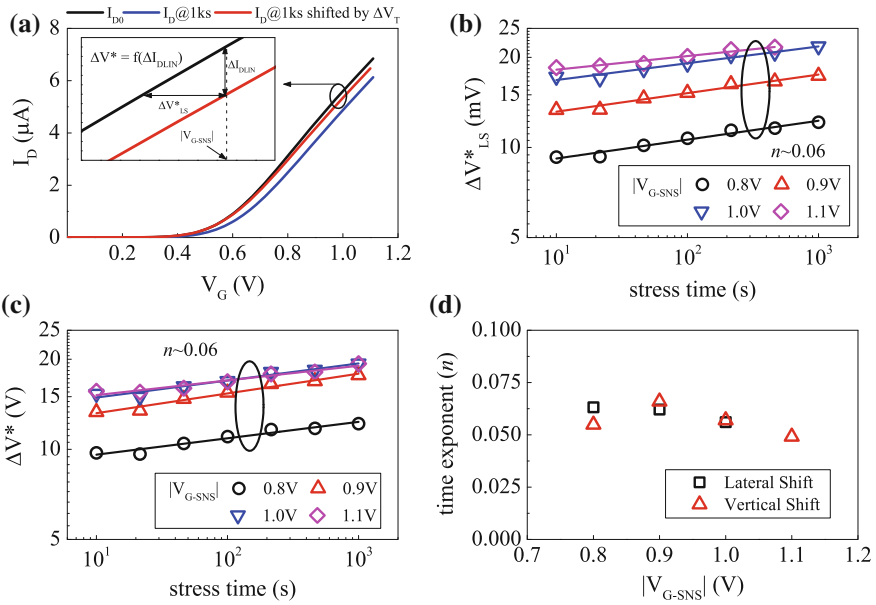

Fig. 2.19 Isolation of mobility impact: **a** Measured I_{DLIN} versus V_G curves before and after stress, and shift of post-stress *I–V* curve by ΔV_T. Calculated degradation using **b** lateral and **c** vertical shift methods, between pre-stress and ΔV_T shifted post-stress data for different sense V_G. **d** Power-law time exponent of mobility degradation induced degradation obtained at different sense V_G from (**b**) and (**c**). Data from NBTI stress in HKMG p-MOSFETs

that for ΔV_T, which is evident from Fig. 2.19b, c. The impact of $V_{G\text{-}SNS}$ on time exponent n, extracted using linear regression of time evolution of ΔV_{LS}* and ΔV* data in t_{STR} range of 10 s to 1 Ks is shown in Fig. 2.19d. Due to lower n for ΔV_{LS}* and ΔV*, increased ΔV_{LS} and ΔV magnitude is always associated with reduced time exponent n when OSDD measurements are done at higher $V_{G\text{-}SNS}$, as shown in Fig. 2.18e, f.

Figure 2.20a, b plot the time evolution of $\Delta V_{LS} - \Delta V_{LS}$* and $\Delta V - \Delta V$* obtained at different $V_{G\text{-}SNS}$ respectively from lateral and vertical shift methods for different measurement delay. It is important to remark that once the voltage shift ΔV_{LS}* and ΔV* corresponding to mobility degradation is subtracted, time evolution of $\Delta V_{LS} - \Delta V_{LS}$* and $\Delta V - \Delta V$* would correspond to that of ΔV_T. Figure 2.20c, d shows the impact of $V_{G\text{-}SNS}$ on corresponding time exponent n, obtained using linear regression of the mobility corrected measured data in t_{STR} range of 10 s to 1 Ks, for lateral and vertical shift methods for different measurement delay. UF-MSM measured ΔV_T time evolution for different measurement delay and corresponding time exponents are also shown as reference. It is important to note that once the impact of mobility is corrected, the lateral and vertical shift methods provide identical ΔV_T and n values at different $V_{G\text{-}SNS}$, which matches well with UF-MSM measured data, and this holds for different measurement delay as shown.

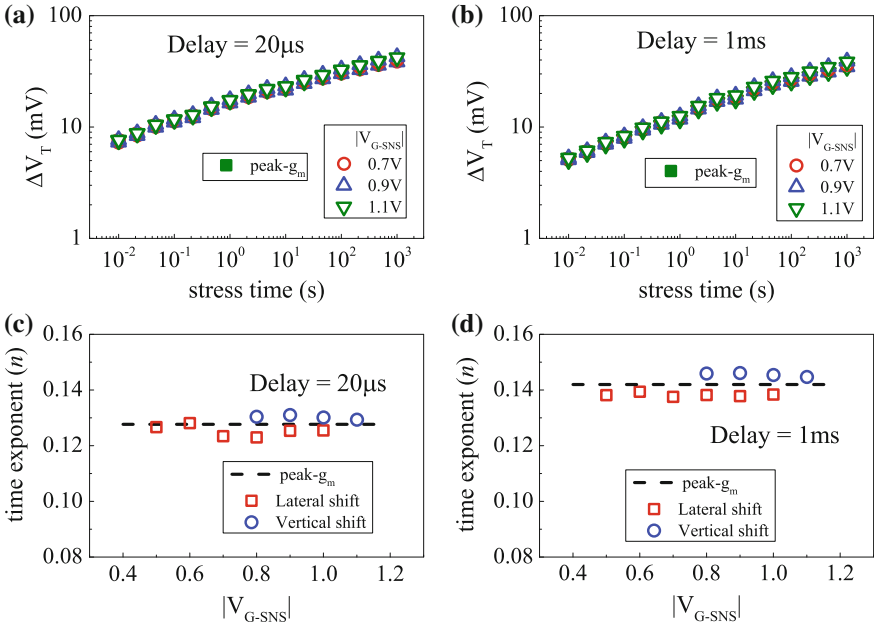

Fig. 2.20 **a**, **b** Time evolution of mobility corrected ΔV_T from lateral and vertical shift OSDD methods at different sense V_G for different measurement delay, for NBTI stress in HKMG p-MOSFETs. **c**, **d** Extracted long-time power-law time exponent n versus sense V_G, obtained from data in (**a**, **b**). UF-MSM measured data obtained using peak g_m method are shown as reference

2.6 Flicker Noise

Figure 2.21 shows the schematic of a flicker noise setup. Gate of the MOSFET is connected to a power supply via a low pass filter; drain is connected to a digital spectrum analyzer via a low noise amplifier. Power spectral density of drain current noise (S_{ID}) is measured in frequency (f) domain for different values of gate overdrive ($V_G - V_{T0}$), and input referred noise (S_{VG}) is obtained from the relation $S_{VG} = S_{ID}/g_m^2$, where V_{T0} and g_m are threshold voltage and transconductance respectively of the device under test [25].

Several mechanisms have been proposed in the past to explain flicker noise in MOSFETs. Some reports suggest noise is due to fluctuation in inversion layer carrier density [26–28], while others relate noise to bulk mobility fluctuation [29, 30]. The number fluctuation model [26–28] is based on McWhorter's theory of random trapping and detrapping of inversion layer carriers in the gate insulator traps [31], which in turn causes surface potential fluctuation and hence variation in inversion layer carrier density. The model suggests S_{VG} to be independent of gate overdrive, and strong correlation of S_{VG} to density of gate insulator traps [32, 33]. The mobility fluctuation model on the other hand is based on Hooge's empirical

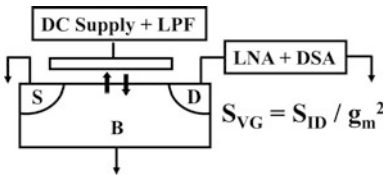

Fig. 2.21 Schematic of flicker noise measurement setup

formulation derived for bulk semiconductors [34], and in MOSFETs it is associated to fluctuations in bulk mobility as a result of fluctuation in phonon population due to phonon scattering [35]. In contrast to the number fluctuation theory, the mobility fluctuation model suggests S_{VG} to be linearly dependent on $(V_G - V_{T0})$.

Flicker noise method has been used in the past to directly estimate gate insulator defects in Silicon Dioxide (SiO_2) MOSFETs [32, 33]. More recently, the method has been used to determine pre-existing trap density in MOSFETs having differently processed SiON and HKMG gate insulators [25, 36–38]. Although flicker noise can also be utilized to determine trap generation after BTI stress [25], accurate flicker noise measurement is a time consuming process and hence the method would suffer from recovery issues and cannot estimate correct magnitude of generated defects.

Figure 2.22 plots S_{VG} as a function of (a) frequency at fixed $(V_G - V_{T0})$ and (b) gate overdrive at fixed f, measured in SiON p-MOSFETs before and after NBTI stress [25]. Note that S_{VG} increases after stress due to generation of new gate insulator traps as discussed in detail later in this chapter, which is consistent with previous reports [32, 33] and number fluctuation theory [26–28].

As discussed in [39] and shown in Fig. 2.23, noise contribution due to trapping and detrapping of inversion carrier for a single gate insulator trap has Lorentzian f dependence of the form $1/[1 + (ff_0)^2]$, where f_0 is the corner frequency of the trap;

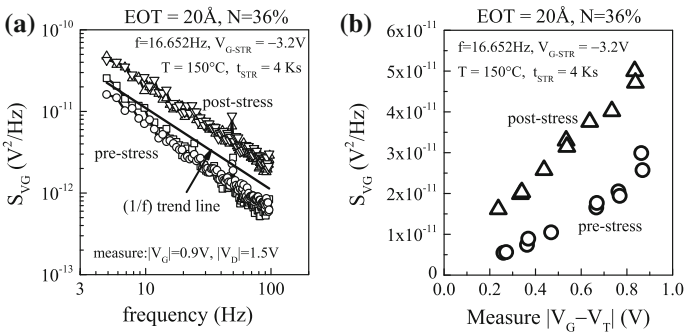

Fig. 2.22 Measured input referred noise versus **a** frequency and **b** gate voltage overdrive, before and after NBTI stress in SiON p-MOSFETs. Multiple measurements are shown in pre- and post-stress, obtained from different devices

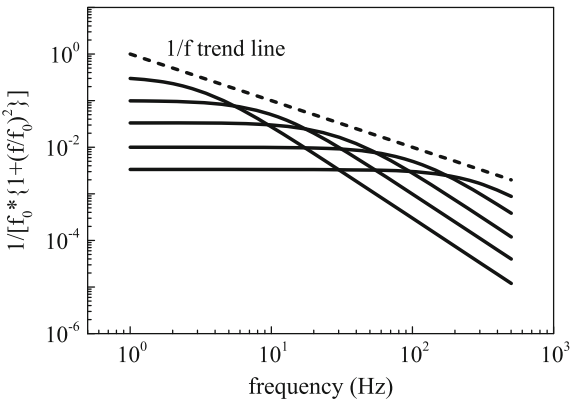

Fig. 2.23 Schematic representation of flicker noise as weighted sum of Lorentzian functions with different corner frequencies, resulting in 1/*f* frequency response

a weighted summation of contribution from uniform spatial distribution of gate insulator traps with different f_0 results in 1/*f* dependence. As evident from Fig. 2.22, measured frequency dependence of S_{VG} shows f^{-k} dependence, with $k \sim 0.7$ before stress, which increases to $k \sim 1$ after stress. A plausible explanation has been provided in [25], which relates noise in SiON devices to Nitrogen (N) related gate insulator traps. Since PNO devices have been used having N concentration that peaks at the SiON/poly-Si interface and exponentially reduces towards the Si/SiON interface, a suitably weighted sum of Lorentzian response from such non-uniform trap distribution results in $k < 1$ for f dependence of S_{VG} before stress. As discussed later in this chapter and in Chap. 3, NBTI stress results in generation of traps at and near the Si/SiON interface and hence the overall spatial distribution of gate insulator traps become relatively more uniform. As a consequence, the f dependence of S_{VG} after stress increases and shows $k \sim 1$.

Although the dependence of S_{VG} on gate insulator traps agrees with the number fluctuation theory, the variation of S_{VG} with gate overdrive shown in Fig. 2.22b cannot be explained in this framework. It is now well known that charges in gate insulator traps not only impacts the inversion carrier density but also influences their mobility by Coulomb scattering. The number fluctuation model has been enhanced to incorporate the effect of surface mobility fluctuation caused by gate insulator charges [40]. The combined number and mobility fluctuation model predicts the following relation for S_{VG} [40]:

$$S_{VG} = \frac{kT\,q^2}{\gamma f WLC_{OX}^2}(1 + \alpha\mu N_C)^2 N_T(E_{FN}) \tag{2.4}$$

In (2.4), q is electronic charge, kT is thermal energy, f is frequency, C_{OX}, W and L are gate oxide capacitance, device width and length respectively; μ is inversion layer mobility, $N_C = C_{OX}/q * (V_G - V_{T0})$ is inversion carrier density; γ is the attenuation factor of electron or hole wave function into the gate insulator and can be calculated using WKB tunneling framework; α is scattering coefficient and N_T is the gate insulator trap density near the Fermi level (E_{FN}). Although missing from

the conventional number fluctuation model, the combined number-mobility fluctuation model has linear dependence on gate overdrive via the term N_C and therefore is consistent with measured data.

As discussed later in this book, gate insulator processes impact the density of pre-existing traps and hence the trapping component of BTI degradation. Of particular interest is the impact of Nitrogen, which is discussed in detail in different chapters of this book. Figure 2.24 plots as processed, pre-existing trap density in (a) SiON p-MOSFETs as well as (b) HKMG p-MOSFETs and n-MOSFETs extracted using the flicker noise method [25, 38]. As discussed in Chap. 1, Table 1.1, PNO SiON devices with different N dose have been used. On the other hand, the HKMG devices were fabricated without (D2) and with (D3) nitridation after ALD High-K deposition and with nitrided IL (D4), refer to Chap. 3, Fig. 3.2 for details. Note that flicker noise in n- and p-channel MOSFETs is respectively due to trapping and detrapping of inversion layer electrons and holes in gate insulator traps. Therefore, it can be remarked that presence of N in gate insulator increases pre-existing hole trap density in SiON and HKMG p-MOSFETs and electron trap density in HKMG n-MOSFETs, and hence impacts both NBTI and PBTI degradation as discussed in detail in Chaps. 3 and 4.

In spite of strong experimental evidence in the literature of trap generation during BTI stress in p- and n-channel MOSFETs, shown later in this chapter and also in Chap. 3 , some reports have suggested NBTI and PBTI to be exclusively due to hole and electron trapping respectively, in pre-existing gate insulator traps [41–43]. Such a framework has to associate strong T activation to the charge trapping process to explain T activation of BTI, which can be independently assessed using flicker noise technique. Figure 2.25 plots T dependence of S_{VG} measured at fixed f and $(V_G - V_{T0})$ in SiON p-MOSFETs [44]. Note that contrary to the above proposition, S_{VG} has negligible T dependence and therefore would suggest negligible T activation of the hole trapping and detrapping process in thin gate insulator stacks used in modern MOSFETs. It is important to remark

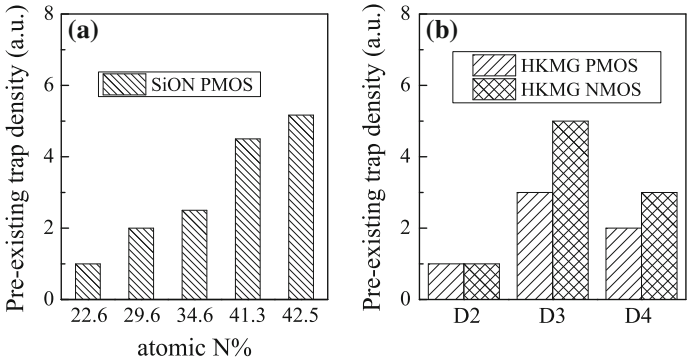

Fig. 2.24 Pre-stress trap density measured using flicker noise method in **a** SiON p-MOSFETs having different N% and **b** HKMG p- and n-MOSFETs having different gate insulator processes (*D2* non-nitrided, *D3* post High-K nitridation, *D4* nitrided IL, refer to Chap. 3, Fig. 3.2) leading to different N%

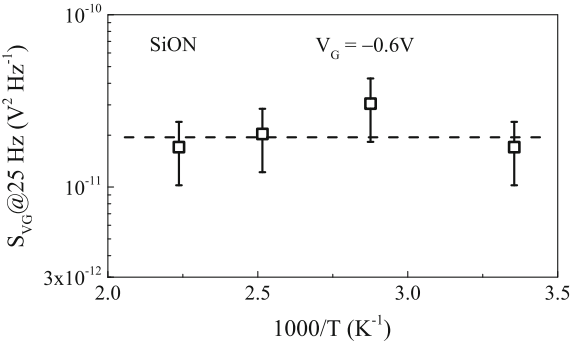

Fig. 2.25 Impact of temperature on measured input referred noise. Data from [44]

that weak T activation of the trapping–detrapping process is significant observation, and when combined with relatively stronger T activation of the trap generation process, shown in Chap. 3, it can explain T activation of BTI for different gate insulator processes as discussed in detail in Chap. 4.

Note that the T activation of NBTI measured in SiON and HKMG p-MOSFETs and of PBTI in HKMG n-MOSFETs reduces for devices having N in the gate insulator stack, as discussed in detail in Chaps. 1 and 4. This can be explained by invoking (a) both trap generation and charge trapping processes for BTI that are mutually uncorrelated, (b) relatively lower T activation for the charge trapping process, shown in Fig. 2.25, when compared to the trap generation process, shown in Chap. 3, and (c) higher relative contribution of charge trapping in pre-existing gate insulator traps for devices having N in gate insulator, as shown in Fig. 2.24; refer to Chap. 4 for further details. Other manifestations of uncorrelated trap generation and trapping BTI sub-components are discussed in Chaps. 3 and 4 for different gate insulator processes, and flicker noise method has been used for independent verification of pre-existing gate insulator traps.

2.7 Charge Pumping (CP)

CP technique [45] estimates trap density at and near the Silicon (Si) channel and gate insulator interface of a MOSFET by measuring trap assisted electron-hole recombination current. In CP method, illustrated in Fig. 2.26, a large signal gate pulse is applied to drive the MOSFET repetitively between inversion and accumulation, the drain and source terminals are shorted and grounded, and DC current due to recombination of electrons and holes in gate insulator traps is measured at the substrate. Although the method can measure pre-stress gate insulator traps, increase in CP current (ΔI_{CP}) after BTI stress can be used as a direct estimation of generation of new gate insulator traps.

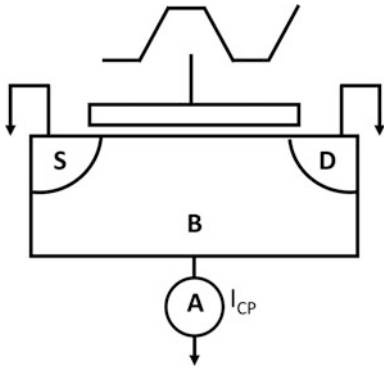

Fig. 2.26 Schematic of charge pumping measurement setup

However, CP is a slow measurement technique and is usually implemented in measure-stress-measure or MSM configuration to estimate trap generation due to BTI stress, and hence suffer from recovery issues. Moreover, CP scans trap generation in an energy range of the Si band gap that is much smaller compared to that scanned by I_{DLIN} and V_T measurement methods. Therefore, measurement delay and band gap correction methods have been proposed [9] for accurate estimation of trap generation contribution to overall BTI degradation and are discussed later in this section. CP technique has been extensively used to characterize trap generation during BTI stress [9, 14, 22, 36, 46–48].

The basic CP method has been modified to estimate the energetic and spatial distribution of traps in the gate insulator [49, 50]. However, such spectroscopic CP techniques are extremely time consuming, and although they can be used to estimate as processed, pre-existing traps, these methods suffer from recovery related issues and hence are not effective to determine the distribution of generated traps. Furthermore, some reports have proposed ultra-fast CP techniques to characterize on-the-fly trap generation without any measurement delay [51, 52]. These methods rely on applying a gate pulse having inversion level equal to the BTI stress bias, $V_{G\text{-}STR}$. The application of such a large gate bias makes these methods prone to gate leakage [53] and unsuitable for thinner gate insulator MOSFETs. More importantly, since CP method involves pulsing the channel from inversion to accumulation, the application of such large bi-polar pulses alters the conventional BTI stress regime, and makes the MOSFET vulnerable to significant bulk trap generation in addition to generation of traps at or near the Si/SiO_2 interface. Owing to the application of large gate pulse, these methods scan traps deep into the gate insulator bulk [49, 50], and therefore, the contribution due to additionally generated bulk traps must be corrected for proper estimation of BTI generated defects. Therefore, although recovery issue can be avoided, the accuracy of these ultra-fast CP methods gets impacted by gate leakage and bulk trap issues, and the methods are of not much use to characterize BTI trap generation.

Figure 2.27 illustrates the dynamics of electron and hole capture-emission processes in gate insulator traps of an n-channel MOSFET under repetitive gate pulses [45]. The channel goes into inversion when V_G goes above V_T, and traps that

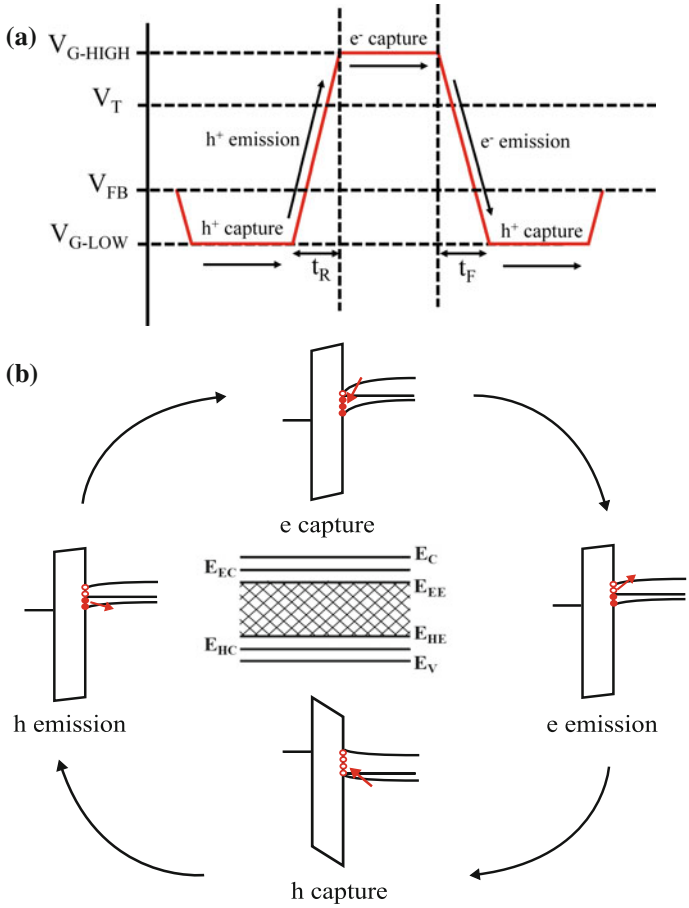

Fig. 2.27 Schematic of charge pumping process: **a** Different regimes of trap assisted capture and emission of electrons and holes as gate is pulsed between inversion and accumulation. **b** Corresponding energy band diagrams showing capture and emission of electrons and holes, and the energy zone for recombination. Example is shown for *n*-channel MOSFET

remain below the Fermi level capture inversion layer electrons from the Si conduction band. The channel transitions from inversion via depletion towards accumulation during the falling edge of the gate pulse, traps start to move above the Fermi level and some traps emit electrons back to the conduction band. Once the channel goes into accumulation as V_G goes below the flatband voltage (V_{FB}), traps above Fermi level capture holes from the valence band of Si. As the channel is pulsed back via depletion to inversion during the rising edge of the gate pulse, traps start to go below the Fermi level and some traps emit holes back to the valence band. Finally, as the channel goes back to inversion as V_G goes above V_T, the electron capture process starts, and subsequent processes get repeated.

Note that time constant of traps for electron emission is smallest near the conduction band edge and increases towards the valence band, while time constant for hole emission is smallest near the valence band edge and increases towards the conduction band. Figure 2.27 also depicts the energy level of traps associated with capture-emission process. During inversion, traps up to the level E_{EC} gets filled with electrons, and electron emission occurs in the levels between E_{EC} and E_{EE} during the falling edge of the gate pulse, as V_G transitions between V_T and V_{FB}. Similarly during accumulation, traps up to the level E_{HC} get filled with holes, and hole emission occurs in the levels between E_{HC} and E_{HE} during the rising edge of the gate pulse, as V_G transitions between V_{FB} and V_T. Therefore, electron-hole recombination occurs in traps between the range E_{EE} and E_{HE} that results in CP current, I_{CP}. The energy zone (ΔE) of traps for electron-hole recombination is given by [45]:

$$\Delta E = -2kT \ln\left[n_i v_{th} \left(\sigma_n \sigma_p\right)^{0.5} \left(t_R t_F\right)^{0.5} \left\{\frac{V_T - V_{FB}}{V_{GH} - V_{GL}}\right\}\right] \tag{2.5}$$

In (2.5), V_{GH} and V_{GL} respectively are pulse high and low levels, t_R and t_F are pulse rise and fall times for transition between V_{GL} to V_{GH} and V_{GH} to V_{GL} respectively, refer to Fig. 2.27; kT and n_i are thermal energy and intrinsic carrier density respectively, v_{th} is thermal velocity, while σ_n and σ_p are trap capture cross sections associated respectively with electron and hole trapping. For gate pulse having frequency f, the CP current I_{CP} is given by [45]:

$$I_{CP} = qfWLN_{IT} \tag{2.6}$$

In (2.6), q is the electronic charge, W and L respectively are width and length of the MOSFET; $N_{IT} = \langle D_{IT} \rangle \Delta E$, where $\langle D_{IT} \rangle$ is the average density of traps per energy in the energy zone ΔE scanned by CP, refer to (2.5) and Fig. 2.27, and N_{IT} is the total trap density probed by CP method.

CP measurements are done in slow MSM mode before and during logarithmically spaced interruptions of BTI stress, and by assuming no change in σ_n and σ_p and therefore in ΔE, changes in CP current (ΔI_{CP}) after stress can be attributed to changes in trap density (ΔN_{IT}). Note that to reduce measurement time, CP measurements can be done at fixed gate pulse amplitude, and the V_{GH} and V_{GL} levels of the gate pulse need to be chosen above V_T and below V_{FB} respectively, to drive the MOSFET between strong inversion and accumulation. Pre-stress N_{IT} can be independently estimated from subthreshold slope measurements, and can be used with pre-stress I_{CP} to determine the geometric mean of σ_n and σ_p. Although CP method has been used to quantify trap generation during NBTI stress in SiON and HKMG p-MOSFETs as well as PBTI stress in HKMG n-MOSFETs, in this section, results are shown for NBTI stress in SiON p-MOSFETs.

As mentioned earlier in this chapter and discussed in Chap. 3, NBTI parametric shift is caused by mutually independent contribution from trapping in pre-existing defects and generation of new traps; the later can be independently verified using

CP measurements. However, caution must be applied before the time evolution of ΔN_{IT} obtained from CP measurements is converted to the trap generation component ($\Delta V_{IT} = q \cdot \Delta N_{IT}/C_{OX}$) of overall V_T shift [9]. First, note that CP method scans trap generation in energy range ΔE in the band gap centered on the midgap, refer to (2.5), while V_T shift estimated from inversion I_{DLIN} degradation is impacted by generated traps throughout the band gap. Moreover, while I_{DLIN} is usually measured using ∼milliseconds or ∼microseconds delay, CP method is implemented in slow MSM mode with measurement delay of ∼seconds, and hence, the later would suffer from recovery related artifacts. Therefore, time evolution of ΔN_{IT} from CP measurements must be corrected for band gap and delay differences, which is discussed below.

Figure 2.28a shows ΔN_{IT} time evolution obtained from CP measurements for NBTI stress in SiON p-MOSFETs [9]. Measurements were done using different t_R and t_F values of gate pulse that result in different scanned energy zone ΔE, refer to (2.5) [45]. Note that measured ΔN_{IT} increases with reduction in t_R and t_F as ΔE is increased, and the ΔN_{IT} versus ΔE correlation can be used to determine corrected ΔN_{IT} for traps corresponding to the full band gap, which is also plotted in Fig. 2.28a. Uniform trap generation in the entire band gap is assumed for this correction. Figure 2.28b plots ΔN_{IT} time evolution after band gap correction, obtained from CP measurements with different measurement delay, t_M. Note that NBTI stress induced generated traps do recover after stress is stopped for measurement. Therefore, the magnitude of ΔN_{IT} reduces and the power-law time exponent n increases with increase in t_M as recovery is increased [2, 22].

Time evolution of ΔN_{IT} at different t_M can be modeled using empirical universal recovery expression [54]:

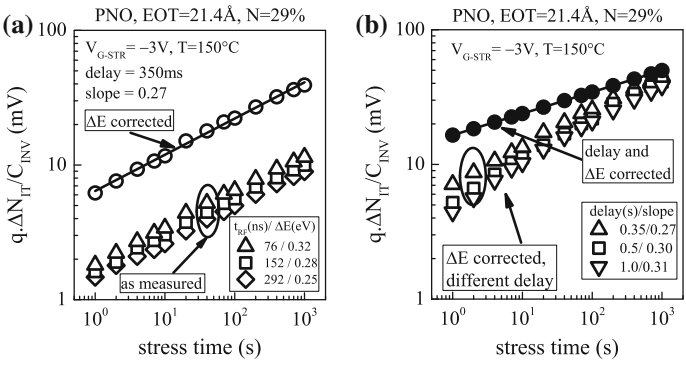

Fig. 2.28 CP measurements: Time evolution of **a** measured ΔN_{IT} for different energy zone of electron-hole recombination determined by different rise and fall time of gate pulse, and estimated ΔN_{IT} after band gap correction. Time evolution of **b** band gap corrected ΔN_{IT} for different measurement delay, and ΔN_{IT} after both band gap and delay correction. Data from NBTI stress in SiON p-MOSFETs

$$\Delta N_{IT}(t_{STR}, t_M) = \frac{\Delta N_{IT}(t_{STR})}{1 + B\left(\frac{t_M}{t_{STR}}\right)^{\beta}} \tag{2.7}$$

The band gap corrected ΔN_{IT} time evolution data can be fitted using (2.7) to determine model parameters, and hence, both band gap and delay corrected ΔN_{IT} can be obtained for any t_M, which is also shown in Fig. 2.28b. It is important to remark that the magnitude of ΔN_{IT} increases and power-law time exponent n reduces after corrections are performed on as-measured CP data. Failure to perform these corrections would severely underestimate the trap generation component and overestimate the trapping component of NBTI, as done in [55]. Since trapping is attributed to the as-processed pre-existing defects, this would severely underestimate quality of the gate insulator; refer to [9, 25] for further details. Once the band gap and delay corrections are performed, time evolution of directly measured ΔN_{IT} from CP would provide correct ΔV_{IT} subcomponent of overall ΔV_T, which is further discussed in later chapters of this book.

Recovery of NBTI stress-generated traps is of interest and is modeled in detail in Chap. 6. An interesting artifact of N_{IT} recovery is shown in Fig. 2.29. CP measurements have been performed in SiON p-MOSFETs before and during logarithmic interruptions of NBTI stress; the stress has been performed at different T, and different measurement delay t_M has been used. Figure 2.29 shows (a) time evolution of ΔN_{IT} for different stress T at fixed t_M, and (b) power-law time exponent n of ΔN_{IT} time evolution data as a function of T for different t_M [22]. The exponent is calculated using linear regression of measured data in t_{STR} range of 10 s to 1 Ks. It is interesting to note that measured n increases at higher T for a given t_M, which has been explained by dispersive Hydrogen (H) transport related trap generation mechanism [20]. However, the above observation is simply a measurement artifact

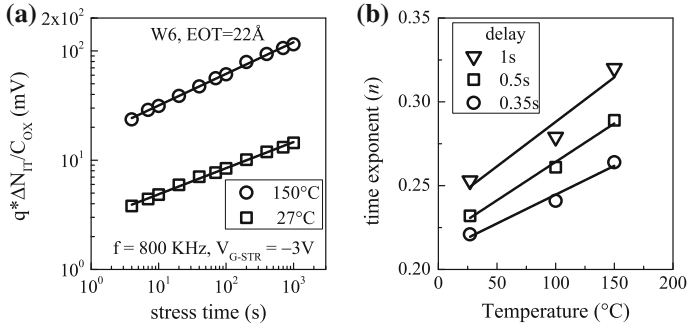

Fig. 2.29 CP measurements: **a** Time evolution of ΔN_{IT} for NBTI stress in SiON p-MOSFETs at different stress T. **b** Extracted power-law time exponent n from ΔN_{IT} time evolution data as a function of stress T for different measurement delay

associated with N_{IT} recovery, since the T dependent increase in n is strongly impacted by t_M as shown in Fig. 2.29b, and larger n is observed at higher t_M for a given T [22]. Therefore, proper corrections to CP measured data must be done before deriving conclusions regarding NBTI physical mechanism.

2.8 Gated Diode (DCIV)

Although CP is a very useful technique to determine trap generation, it requires the application of large pulses to the gate terminal of the MOSFET to drive the channel from strong inversion to accumulation, and hence is not suitable for gate insulators having thin EOT due to gate leakage issues [53]. Alternatively, direct estimation of traps at and near the Si/SiO_2 interface can be obtained using the DCIV method [56, 57], illustrated in Fig. 2.30. The source-drain terminals are shorted and forward biased just below the junction cut-in voltage, the gate bias is swept from accumulation to inversion in the vicinity of $V_G \sim 0$ V, and trap assisted electron hole recombination current I_{DCIV} is measured at substrate terminal. DCIV method has been used to characterize trap generation during NBTI stress in SiON p-MOSFETs [48, 57–59], as well as during NBTI and PBTI stress respectively in HKMG p- and n-MOSFETs [17, 18, 37, 38, 60]; HKMG results are discussed in detail in Chap. 3. In this section, implementation of DCIV method is discussed using the example of NBTI stress in HKMG p-MOSFETs.

Figure 2.31 plots measured I_{DCIV} versus V_G characteristics (a) before and during different intervals of NBTI stress, as well as (b) immediately after and at different intervals following stoppage of NBTI stress in HKMG p-MOSFET. Note that I_{DCIV}

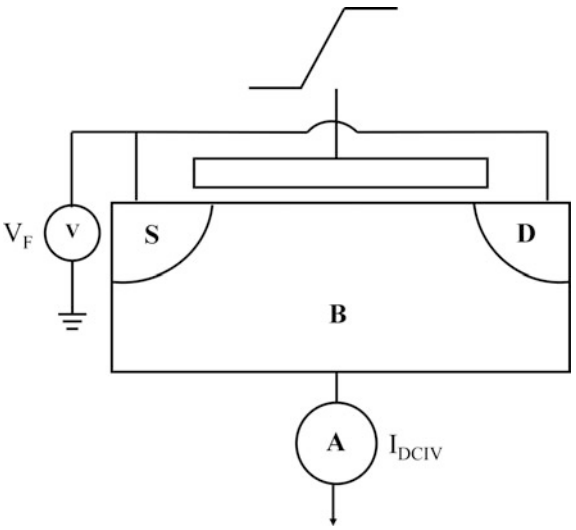

Fig. 2.30 Schematic of gate diode (DCIV) measurement setup

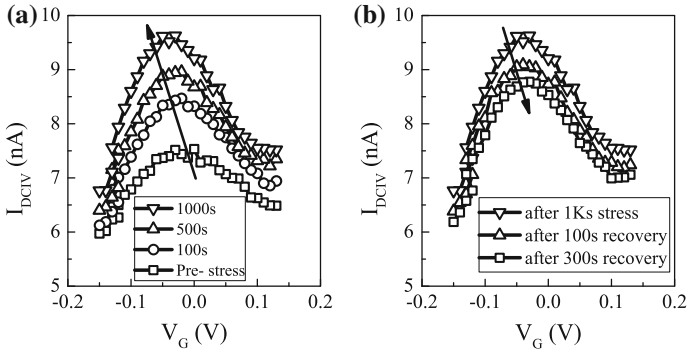

Fig. 2.31 *I–V* characteristics from DCIV measurements **a** before and after different intervals of stress, and **b** immediately after stress and after different recovery intervals, for NBTI stress in HKMG p-MOSFETs

peaks at a particular V_G, and the peak magnitude increases and reduces respectively during and after NBTI stress. The V_G value corresponding to the peak I_{DCIV} also varies during stress and recovery. Since I_{DCIV} is due to electron-hole recombination via the traps at or near the Si/SiO_2 interface, increase and reduction in I_{DCIV} signify generation and recovery of traps during and after NBTI stress. Trap density can be estimated from peak I_{DCIV} using the Shockley–Read–Hall (SRH) formalism of carrier capture and emission [61].

For a single defect situated at energy level E_T having density N_{IT}, the difference between base and peak values of I_{DCIV} versus V_G characteristics is given by the following relation [56, 57, 62]:

$$I_{DCIV,peak} - I_{DCIV,base} = \frac{qn_i v_{th}\left(\sigma_n \sigma_p\right)^{0.5} N_{IT} \frac{WL}{2}\left[\exp\left(\frac{qV_F}{kT} - 1\right)\right]}{\exp\left(\frac{qV_F}{2kT}\right) + \cos h(U_T)} \tag{2.8}$$

In (2.8), q is electronic charge, W and L are device width and length respectively, kT and n_i are thermal energy and intrinsic carrier density respectively, v_{th} is thermal velocity, σ_n and σ_p are trap capture cross sections associated with electron and hole trapping respectively and V_F is forward bias applied to the shorted source and drain terminals. $U_T = (E_T - E_i)/kT + \ln(\sigma_n/\sigma_p)$, where E_i is intrinsic level. The equation shown is for a single trap situated at a particular energy level, and has to be integrated over different trap energy levels for distribution of traps. The energy zone scanned by DCIV technique is given by $\Delta E = q \cdot |V_F|$, and for relatively larger V_F value, which should be still lower than the junction cut-in voltage, (2.8) can be approximated as follows [59, 62]:

$$I_{DCIV,peak} - I_{DCIV,base} = \frac{1}{2} qn_i \left(\sigma_n \sigma_p\right)^{0.5} v_{th} N_{IT} (WL) \exp\left(\frac{qV_F}{2kT}\right) \tag{2.9}$$

In (2.9), $N_{IT} = \langle D_{IT} \rangle \Delta E$ is average density of traps in energy zone ΔE, and similar values of σ_n and σ_p has been assumed. DCIV measurements are done before and during logarithmically spaced interruptions of BTI stress, and by assuming no change in σ_n and σ_p due to stress, time evolution of ΔN_{IT} can be estimated from time evolution of measured peak ΔI_{DCIV} using (2.9). As before, pre-stress subthreshold slope is used to estimate pre-stress N_{IT} and hence the geometric mean of σ_n and σ_p can be found by using pre-stress I_{DCIV}.

Note that similar to CP, corrections should be done on as-measured DCIV data before directly estimated ΔN_{IT} is compared to the ΔV_{IT} component of ΔV_T. Similar to CP, DCIV is also a slow measurement method and is implemented in the MSM mode. A typical V_G sweep takes $\sim$seconds, and the time evolution of as-measured stress-generated defects should be corrected for measurement delay using the universal recovery expression shown in (2.7). Moreover, DCIV method scans traps in energy zone $\Delta E = q \cdot |V_F|$ in the Si band gap centered on the midgap, while ΔV_T is impacted by traps generated throughout the band gap. Therefore, obtained trap density should also be corrected for band gap difference before comparing to the trap generation subcomponent of overall V_T shift [60]. Failure to do these corrections would result in gross underestimation of the trap generation subcomponent.

DCIV measurements were performed in different HKMG p-MOSFETs, before and immediately after NBTI stress as well as after a certain delay following NBTI stress. Devices having different HKMG gate insulator processes have been used [38]; refer to Chap. 3, Fig. 3.2 for details. Different E_{OX} values have been used for stress, and all devices were stressed and recovered for identical stress time t_{STR} and recovery time t_{REC}, respectively. Figure 2.32a correlates ΔI_{DCIV} measured just after NBTI stress to ΔI_{DCIV} after recovery following NBTI stress. Note that universal correlation between generation and recovery of traps has been observed across different devices. Such correlation can be used to determine the parameters of universal recovery expression shown in (2.7), which can further be used to perform delay correction of as-measured DCIV data. Figure 2.32b plots the time evolution of as-measured ΔN_{IT} using the DCIV method for different measurement delay, t_M, for NBTI stress in HKMG p-MOSFET. Time evolution of ΔN_{IT} after delay correction and both delay as well as band gap correction are also shown. Note that the magnitude of ΔN_{IT} increases as well as power-law time exponent n reduces after delay and band gap corrections.

DCIV measurements have been done before and during logarithmically spaced interruptions in NBTI stress, to determine time evolution of generated traps in different HKMG p-MOSFETs listed in Chap. 3, Fig. 3.2. Measured data are corrected for delay and band gap using the procedure discussed above. Figure 2.32c, d plot power-law time exponent n for these devices obtained from corrected trap generation data, as a function of $V_{G\text{-}STR}$ and T, respectively, obtained using linear regression in t_{STR} range of 10 s to 1 Ks. It is important to remark that universal time exponent of $n \sim 1/6$ is always observed, for different devices and across different $V_{G\text{-}STR}$ and T. The physical mechanism behind this universality will be discussed later in this book. Unless mentioned otherwise, all DCIV data presented in Chap. 3 and elsewhere in this book are delay and band gap corrected.

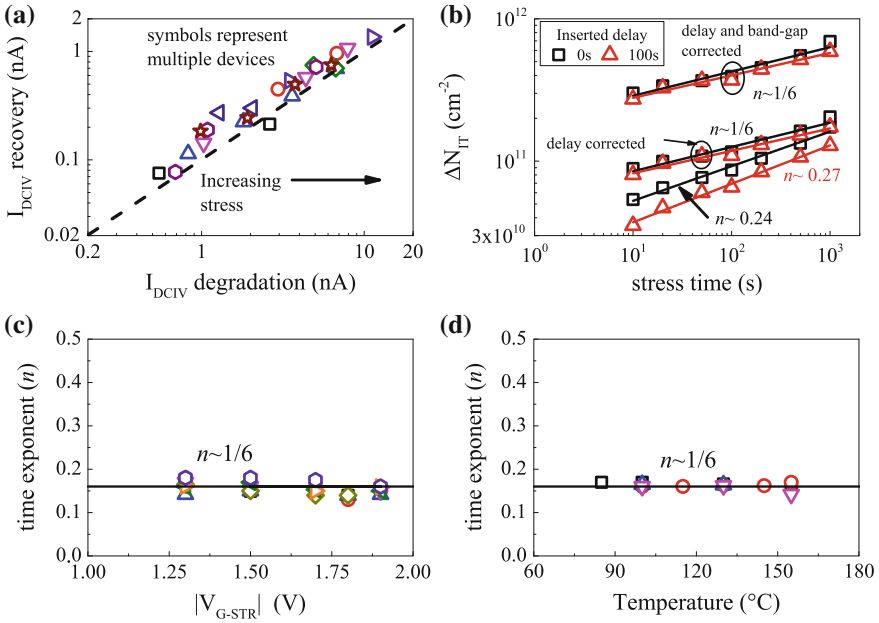

Fig. 2.32 DCIV measurements for NBTI stress in HKMG p-MOSFETs: **a** Correlation of peak DCIV current degradation and recovery for different HKMG devices. **b** Time evolution of measured ΔN_{IT} for different measurement delay, after delay correction and after both delay and band gap correction. Extracted power-law time exponent n versus **c** stress V_G and **d** stress T, obtained from corrected DCIV data for different HKMG devices. Refer to Chap. 3, Fig. 3.2, for details of different HKMG devices

2.9 Stress Induced Leakage Current (SILC)

Increase in gate leakage current (ΔI_G) measured in inversion, refered to as SILC, is a routine characterization technique for Time Dependent Dielectric Breakdown (TDDB) stress in MOSFETs [63–65]. SILC has been attributed to inelastic Trap Assisted Tunneling (TAT) via bulk gate insulator traps generated during TDDB stress, and is illustrated in Fig. 2.33 [63, 66]. Since BTI and TDDB stress regimes are essentially the same, as mentioned in Chap. 1, Fig. 1.2, SILC has been used during NBTI stress in SiON p-MOSFETs [14] and PBTI stress in HKMG n-MOSFETs [24, 37, 38, 67, 68] to access contribution due to generated bulk insulator traps. However, due to band alignment issues, SILC is usually not observed in HKMG p-MOSFETs as discussed in [65]. In this section, the implementation details of SILC measurement are shown using PBTI stress in HKMG n-MOSFETs, and more SILC data are shown later in Chaps. 3 and 4. Independently measured DCIV data are also shown to aid comparison between the two characterization methods.

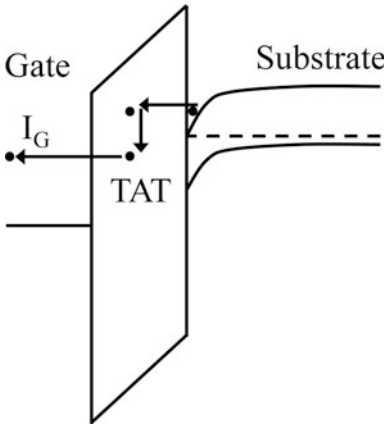

Fig. 2.33 Energy band diagram representing SILC via trap assisted tunneling and structural relaxation of traps. Example shown for *n*-channel MOSFET

Figure 2.34a shows measured I_G versus V_G characteristics before and during logarithmically spaced interruptions of PBTI stress in HKMG n-MOSFET. Time evolution of SILC, manifested in the form of increase in I_G at a particular V_G after stress, can be estimated by noting ΔI_G at a fixed V_G from measured I_G–V_G characteristics. However, note that SILC recovers after stoppage of stress [24] and

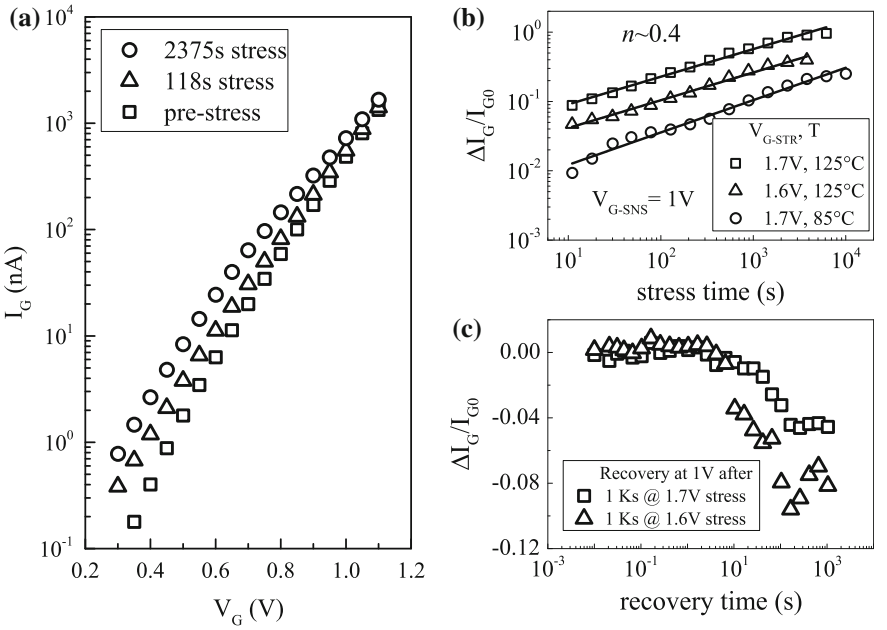

Fig. 2.34 SILC for PBTI stress in HKMG n-MOSFETs: **a** Measured I_G versus V_G characteristics before and after stress. **b** Time evolution ΔI_G obtained at fixed sense V_G for different $V_{G\text{-}STR}$ and T. **c** Time evolution of reduction in ΔI_G sensed at fixed V_G, after removal of stress

I_G–V_G sweep is usually time consuming for accurate I_G measurement. Therefore, it is desirable to avoid full I_G–V_G sweep to minimize recovery, and instead, drop V_G down from $V_{G\text{-}STR}$ to $V_{G\text{-}SNS}$, usually equals to V_{DD}, for one-point I_G measurement. Figure 2.34b shows time evolution of ΔI_G at different $V_{G\text{-}STR}$ and T for PBTI stress in HKMG n-MOSFET. A power-law time dependence is observed with time exponent n ($\sim$0.4), which is much larger than that obtained from DCIV measurements and shown earlier in this chapter. Therefore, trap generation probed by SILC has very different physical origin than that probed by DCIV, which is discussed in detail in Chap. 3. Figure 2.34c shows the reduction of ΔI_G after stoppage of PBTI stress, which signifies the recovery of generated traps as measured using SILC.

Figure 2.35a correlates generation of SILC measured immediately after stress to SILC recovery measured after certain delay following stress in different HKMG n-MOSFETs listed in Chap. 3, Fig. 3.2 [38]. All devices were stressed for fixed t_{STR} duration at different $V_{G\text{-}STR}$, and then allowed to recover for a fixed duration t_{REC}. It is important to note that universal generation to recovery correlation has been observed for different devices, and such data can be used to model SILC recovery using the universal recovery expression of (2.7). As a comparison, Fig. 2.35b correlates trap generation and recovery obtained from measured changes in DCIV current, ΔI_{DCIV}, during and after PBTI stress in these HKMG n-MOSFETs. Identical $V_{G\text{-}STR}$, T, t_{STR} and t_{REC} have been used for SILC and DCIV measurements for fair comparison. Universal correlation has been observed for DCIV measured data across different devices. However, the slope of the correlation line for SILC and DCIV are different, refer to Fig. 2.35a, b. Interestingly, the DCIV measured trap generation to recovery correlation slope is similar between NBTI and PBTI stress, refer to Figs. 2.32a and 2.35b. This aspect is further discussed in Chap. 3.

Figure 2.36 plots the time evolution of as measured and delay corrected (a) ΔI_G and (b) ΔI_{DCIV} respectively from SILC and DCIV measurements, for PBTI stress in

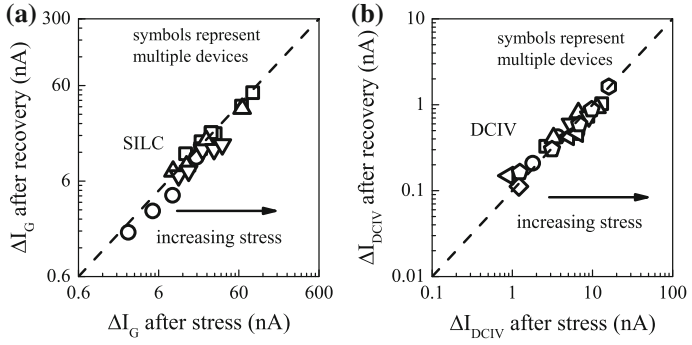

Fig. 2.35 Correlation of trap generation and recovery obtained using **a** SILC and **b** DCIV measurements for PBTI stress in HKMG n-MOSFETs having different gate insulator processes. Refer to Chap. 3, Fig. 3.2, for details of different HKMG devices

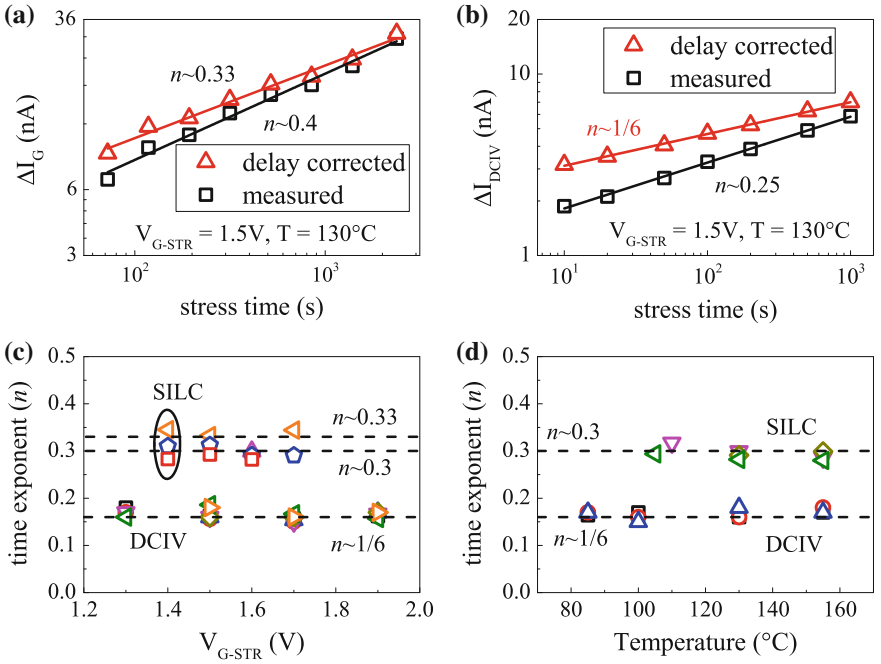

Fig. 2.36 Time evolution of as measured and delay corrected **a** ΔI_G from SILC and **b** ΔI_{DCIV} from DCIV measurements for PBTI stress in HKMG n-MOSFETs. Extracted time exponent n from delay corrected SILC and DCIV data as a function of **c** $V_{G\text{-}STR}$ and **d** stress T, for PBTI stress in different HKMG devices

HKMG n-MOSFET [38]. Delay correction is done using (2.7), and the parameters of the universal recovery expression are calibrated using SILC and DCIV data shown respectively in Fig. 2.35a, b. Note that power-law time evolution of generated traps is observed for both measurements, although n from SILC is much larger than that from DCIV measurements. As expected, magnitude of degradation increases and power-law time exponent n reduces after recovery correction. Figure 2.36 also plots delay corrected n from SILC and DCIV measurements, as a function of (c) $V_{G\text{-}STR}$ and (d) T for PBTI stress in different HKMG n-MOSFETs. The exponent n is obtained using linear regression of measured and recovery corrected data in t_{STR} range of 10 s to 1 Ks. Obtained time exponent n is independent of $V_{G\text{-}STR}$ and T for both SILC and DCIV measurements. SILC measurements show $n \sim 1/3$ after delay correction, which is much larger than $n \sim 1/6$ observed for DCIV measurements. It is important to remark that delay corrected DCIV data show universal $n \sim 1/6$ for both NBTI and PBTI stress, which will be further discussed in Chap. 3.

For a particular bulk trap density N_T (x, E) at spatial location x and energy E in the band gap of the gate insulator, increase in inversion gate leakage due to TAT of electrons is given by the following relation [63, 66]:

$$\Delta I_G = \iint q\sigma_n v_{th} N_C S_N (f_C - f_A) WLN_T(x,E) \frac{T_{C-T}(E)T_{T-A}(E-E_R)}{T_{C-T}(E)+T_{T-A}(E-E_R)} dx\, dE \tag{2.10}$$

In (2.10), $T_{C\text{-}T}$ and $T_{T\text{-}A}$ are transition probabilities of electron tunneling from cathode, i.e., conduction band of Si substrate to the trap, and from the trap to anode, i.e., gate, respectively, see Fig. 2.33. E_R is the energy relaxation of a trap after capture of an electron, q is electronic charge, σ_n is electron capture cross section, v_{th} is thermal velocity, N_C is density of electrons in the conduction band edge, and S_N is energy suppression factor in conduction band, and f_C and f_A are occupation probabilities of the cathode and anode respectively; W and L respectively are width and length of the MOSFET. The integration is performed over all possible spatial and energy values of traps in the gate insulator. Time evolution of delay corrected ΔI_G can be fitted with (2.10) to obtain time evolution of bulk trap density, ΔN_{OT}, and is discussed in detail later in Chap. 4.

2.10 Low Voltage Stress Induced Leakage Current (LV-SILC)

Increase in MOSFET gate leakage current in accumulation or Low Voltage SILC (LV-SILC) has been used as a monitor of interface trap generation for NBTI stress in p-MOSFETs [1, 58, 69] and also TDDB stress in n- and p-MOSFETs [70]. As of now, LV-SILC has been used to characterize trap generation in MOSFETs having SiO_2 and SiON gate insulators, and to the best of our knowledge, this method has not been used in devices having scaled HKMG gate insulators. In this section, LV-SILC method will be briefly reviewed to provide additional evidence of interface trap generation during NBTI stress in SiON p-MOSFETs.

Accumulation I_G versus V_G sweeps are measured before and during interruptions of NBTI stress, and Fig. 2.37 plots the increase in gate leakage, normalized to pre-stress data, $\Delta I_G/I_{G0}$, as a function of measurement V_G for different t_{STR} [69]. Two different $\Delta I_G/I_{G0}$ peaks are observed that evolve with increase in t_{STR}, one near $V_G \sim 0$ V and the other near $V_G \sim V_{FB}$. As also illustrated in Fig. 2.37, LV-SILC is associated with elastic or small energy loss tunneling between traps generated at the Si/SiO_2 and SiO_2/poly-Si interfaces [58, 69, 70]. The peak near $V_G \sim 0$ V is due to electron tunneling from the valence band of poly-Si to traps at Si/SiO_2 interface as well as hole tunneling via traps at Si/SiO_2 interface. On the other hand, the peak near $V_G \sim V_{FB}$ is due to electron tunneling from Si channel to poly-Si gate via traps at both Si/SiO_2 and SiO_2/poly-Si interfaces. The increase in peak ΔI_G magnitude with stress implies generation of new traps at these interfaces during NBTI stress. Although LV-SILC provides crucial evidence of interface trap generation during NBTI stress, this method is not used in this book.

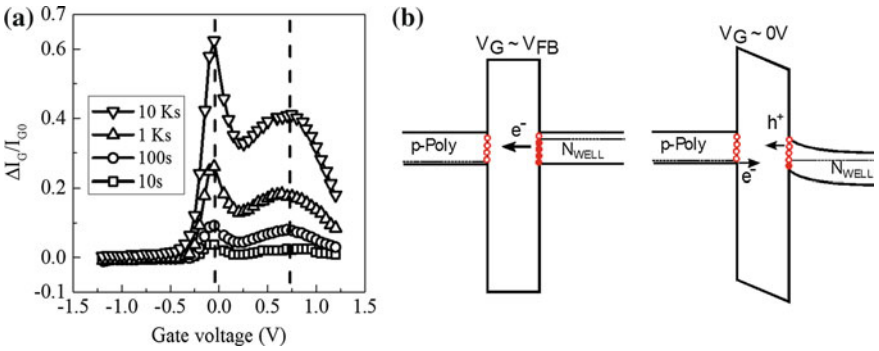

Fig. 2.37 LV-SILC measurements: **a** Increase in I_G after NBTI stress in SiON p-MOSFETs. **b** Energy band diagrams showing trap assisted tunneling corresponding to peaks in *I–V* characteristics. Data from [69]

2.11 Experimental Artifacts

During BTI test, MOSFET is stressed at $V_G = V_{G\text{-}STR}$ that is higher than the nominal operating bias, and the resultant degradation is measured either on-the-fly or by using logarithmically spaced interruptions in stress. Time evolution of degradation, measured under accelerated stress condition but for short duration, is then extrapolated to end-of-life under use condition to determine device lifetime. It is important to choose proper stress and measurement conditions so that measured data remain free from different extraneous artifacts discussed in this section.

Measurement speed is a very important parameter and has very different impact on OTF and MSM or OSDD techniques. Since OTF measures I_{DLIN} without reducing stress V_G, refer to Sect. 2.2, it does not suffer from the usual recovery artifacts. However, since the device starts to degrade as soon as $V_{G\text{-}STR}$ is applied, a faster measurement or smaller t_0 delay would capture a less degraded I_{DLIN0}, which is important as I_{DLIN0} is assumed to be undegraded in this method. A larger t_0 delay results in lower than actual degradation magnitude and higher than actual power-law time exponent as discussed earlier in this chapter. For a particular t_0 delay, the accuracy of OTF measurement would depend on how fast the device degrades after the application of $V_{G\text{-}STR}$. Devices with significant fast trapping component are relatively more sensitive to variations in t_0 delay than devices having negligible trapping, as shown using the example of RTNO and PNO SiON p-MOSFETs.

On the other hand, pre-stress measurement is used as a reference for MSM and OSDD methods, while the gate bias is reduced from $V_{G\text{-}STR}$ for measurements during stress, refer to Sects. 2.3 and 2.4. Therefore, these methods suffer from recovery issues, and a lower than actual degradation magnitude and higher than actual power-law time exponent are captured depending on the measurement speed, t_M. For a particular t_M, the accuracy of MSM and OSDD measurements would

depend on how fast the device recovers after the removal of stress. Note that devices having significant fast trapping would also recover fast, and would be more sensitive to variations in t_M. Therefore, while OTF is impacted by dynamics of degradation at the initiation of stress, MSM and OSDD methods are impacted by dynamics of recovery after removal of stress. Impact of measurement speed on OTF, MSM, and OSDD measurements has been discussed earlier in this chapter.

Recovery plays an important role when BTI degradation is estimated using different measurement methods and compared against each other. For example, since BTI degradation is due to contribution from trap generation and trapping that are mutually uncorrelated, refer to Chap. 4 for details, trap generation is often estimated directly using CP or DCIV techniques and compared to overall V_T shift to determine the trapping component. Multiple measurements are often performed sequentially after stress is interrupted, and depending on measurement delay and the exact sequence, they would produce different results owing to BTI recovery. As an illustration, Fig. 2.38 plots time evolution of ΔV_T and ΔN_{IT} obtained respectively from I_{DLIN}-V_G and CP measurements during NBTI stress in SiON p-MOSFETs [71]. Two different measurement sequences have been used, i.e., *I–V* first CP last and CP first *I–V* last, and produce different results as shown. Note that CP requires larger measurement time than *I–V* measurements, and since NBTI recovers quickly in time as soon as stress is stopped, ΔV_T obtained from *I–V* first CP last is much larger than that obtained from CP first *I–V* last sequence, as much of the degradation recovers during CP measurements for the latter sequence. Therefore, it is important to avoid sequential measurements and perform different measurements independently on identical but different devices.

Even when independently measured, *I–V* and CP measurements result in different time evolution of degradation under identical measurement time. Time evolution of ΔV_T and ΔN_{IT} due to NBTI stress in SiON p-MOSFETs is measured using OSDD and CP methods, respectively. Experiments were performed under

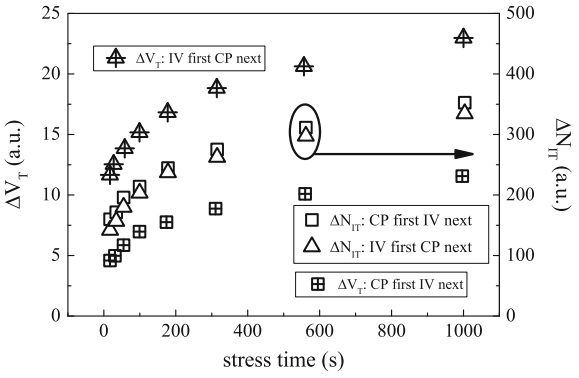

Fig. 2.38 Impact of sequential measurements using multiple methods: Time evolution of ΔV_T and voltage shift corresponding to ΔN_{IT} obtained respectively from *I–V* and CP measurements, for *I–V* first and CP next and CP first and *I–V* next sequence. Data from [71]

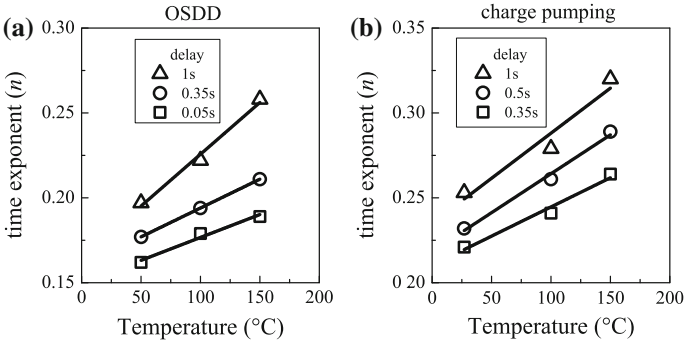

Fig. 2.39 Impact of measurement delay: Measured power-law time exponent n as a function of stress T, extracted from **a** ΔV_T time evolution using OSDD and **b** ΔN_{IT} time evolution using CP methods, for NBTI stress in SiON p-MOSFETs

identical $V_{G\text{-}STR}$ but different stress T, and for different measurement delay t_M. Figure 2.39 plots extracted power-law time exponent n for (a) OSDD and (b) CP measurements as a function of stress T for different t_M, obtained using linear regression of measured data in t_{STR} range of 10 s to 1k s [22]; the CP data are reproduced from Fig. 2.29b. Note that as-measured values of n are different between OSDD and CP measurements at identical stress T and t_M, as these methods suffer from different magnitudes of recovery. For a particular t_M, n increases with increase in stress T, and for both measurement methods, the rate of increase in n is higher for larger t_M due to higher recovery. Therefore, just ensuring identical t_M is not sufficient for error-free comparison of different measurement methods.

As discussed earlier in this chapter, as-measured time evolution of trap generation obtained directly from CP or DCIV must be corrected for measurement delay and band gap differences before comparing with the trap generation component of ΔV_T obtained from ultra-fast OTF, MSM or OSDD I_{DLIN} measurements. It is important to remark that failure to do these corrections would severely underestimate the trap generation and over estimate the trapping component of BTI.

BTI stress is usually performed at different $V_{G\text{-}STR}$, which aids in projecting the accelerated stress data to use condition. BTI and TDDB stress regimes are essentially same, as discussed in Chap. 1, Fig. 1.2, and therefore, it is important to carefully choose $V_{G\text{-}STR}$ so that TDDB effects do not significantly corrupt BTI degradation. As an example, Fig. 2.40a plots the time evolution of ΔI_{DLIN} measured in SiON p-MOSFETs under NBTI stress at different $V_{G\text{-}STR}$ [72]. Note that a power-law time dependence is observed along with the normally observed time exponent of $n \sim 1/6$ for stress at lower $V_{G\text{-}STR}$. However, ΔI_{DLIN} breaks off from simple power-law time dependence and increases at longer t_{STR} for stress at higher $V_{G\text{-}STR}$; the break-off happens earlier in time as $V_{G\text{-}STR}$ is increased. Increased degradation at longer t_{STR} can be modeled using the sum of two independent degradation, having power-law dependence with $n \sim 1/6$ and $n \sim 1/3$ corresponding to NBTI and TDDB respectively, also shown in Fig. 2.40a, refer to

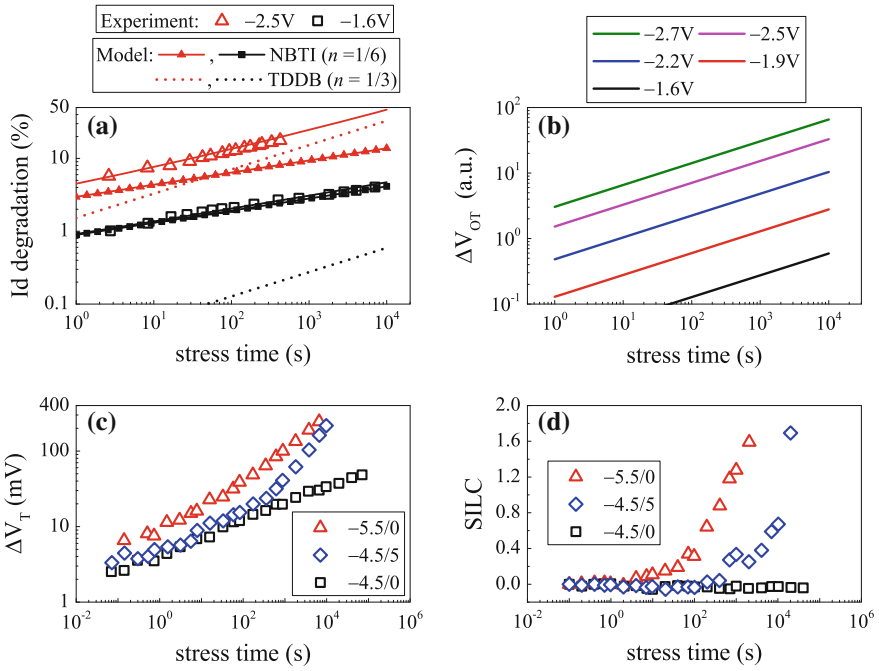

Fig. 2.40 Impact of TDDB defects at high $V_{G\text{-}STR}$: **a** Time evolution of measured ΔI_{DLIN} at different $V_{G\text{-}STR}$, and its decomposition into NBTI (n = 1/6) and TDDB (n = 1/3) components, and **b** time evolution extracted TDDB component for different $V_{G\text{-}STR}$, for NBTI stress in SiON p-MOSFETs. Data obtained from [72]. Time evolution of **c** ΔV_T from I–V and **d** ΔI_G from SILC measurements, for NBTI stress at different $V_{G\text{-}STR}$ and reverse substrate bias V_B in SiO_2 p-MOSFETs

Chap. 4 for further details. Figure 2.40b shows the time evolution of extracted additional degradation for different $V_{G\text{-}STR}$. Identical $n \sim 1/3$ power-law time dependence is observed across all $V_{G\text{-}STR}$, with a V_G acceleration factor of $\Gamma_V \sim 30$ corresponding to time to reach a fixed degradation, which is similar to that observed for TDDB stress [48, 63, 68].

As an additional proof, Fig. 2.40c plots the time evolution of ΔV_T measured in thicker SiO_2 p-MOSFETs for NBTI stress at low and high $V_{G\text{-}STR}$, as well as at low $V_{G\text{-}STR}$ and high reverse substrate bias (V_B) [14]. Note that ΔV_T shows power-law time dependence with a single time exponent of $n \sim 0.25$ for the entire duration of stress, when stress is performed at lower V_G and at V_B = 0 V; higher n is observed due to recovery issues as data were measured using slow MSM method. However, ΔV_T breaks off from the $n \sim 0.25$ power-law trend and increases at longer t_{STR} for stress at higher $V_{G\text{-}STR}$ and V_B = 0 V or at low $V_{G\text{-}STR}$ but $V_B > 0$ V. To understand the mechanism responsible for enhanced degradation at longer t_{STR}, independent SILC measurements were performed in these devices at identical stress conditions. Figure 2.40d plots time the evolution of $\Delta I_G/I_{G0}$ measured in inversion, showing

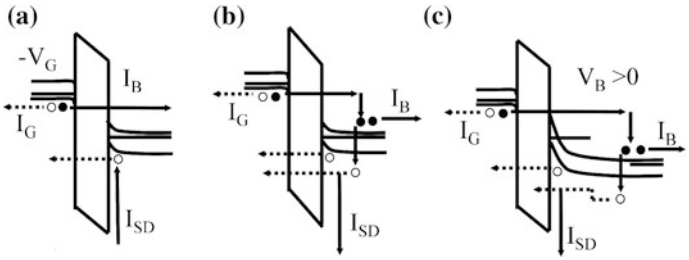

Fig. 2.41 Energy band diagrams for NBTI stress in p-MOSFETs using **a** low V_G and $V_B = 0$, **b** high V_G and $V_B = 0$, and **c** low V_G and high reverse V_B

absence of SILC for low V_G and $V_B = 0$ V stress. However, SILC is observed at longer t_{STR} for stress at higher $V_{G\text{-}STR}$ and $V_B = 0$ V or at lower $V_{G\text{-}STR}$ but $V_B > 0$ V. As SILC is a measure of bulk trap generation, refer to Sect. 2.9, increased ΔV_T at longer t_{STR} can be attributed to the charging of generated TDDB like bulk traps for stress at higher $V_{G\text{-}STR}$ and $V_B = 0$ V or for lower $V_{G\text{-}STR}$ but $V_B > 0$ V.

The energy band diagrams corresponding to NBTI stress in p-MOSFETs, illustrated in Fig. 2.41, can help explain the physical mechanism responsible for TDDB like bulk oxide trap generation under certain stress conditions [14]. For low stress V_G, inversion holes tunnel from Si substrate to gate and electrons tunnel from gate to substrate, refer to Fig. 2.41a, which is normal NBTI condition. At higher V_G, refer to Fig. 2.41b, electrons tunneling from gate to substrate gain enough energy and undergo impact ionization at the substrate and create hot holes, which can get injected back into the oxide by Anode Hole Injection (AHI) mechanism and create oxide defects [63]. Note that high stress V_G results in high E_{OX} in the gate insulator as well as generation of hot holes, and it is difficult to distinguish and identify the main reason behind bulk trap generation. Therefore to further verify the role of hot holes, Fig. 2.41c shows energy bands corresponding to stress at low V_G but high reverse V_B. In spite of low E_{OX}, hot holes can be generated via impacted ionization in the substrate due to high V_B, and can also generate bulk oxide traps as shown. Therefore, TDDB can affect the time evolution of NBTI degradation at long stress time, unless the upper limit of stress gate bias is carefully chosen.

2.12 Summary

To summarize, three different ultra-fast characterization methods are discussed to determine V_T shift of a MOSFET due to BTI stress. The MSM method involves I_D versus V_G sweep measurements before and during periodic interruptions of stress and provides direct estimation of ΔV_T; pre- and post-stress V_T being calculated using peak g_m method. The MSM method is implemented for DC and AC stress and the impact of pulse hold and sweep delay is discussed. The OTF method senses I_D on-the-fly without reducing stress V_G, while the OSDD method measures I_D by

reducing V_G from stress to a fixed sense bias. The impact of mobility degradation at different sense bias is discussed, and a post-processing mobility correction procedure is discussed for these one spot methods to properly estimate ΔV_T from measured I_D degradation. The impact of time-zero delay for OTF and drop down delay for OSDD is discussed. It is shown that OTF measurement accuracy is dependent on how fast a device degrades after stress, while that for MSM or OSDD depends on how fast a device recovers after the removal of stress.

Several characterization techniques are discussed that directly estimate the pre-existing and newly generated defects in MOSFET gate insulator and are responsible for BTI degradation. Pre-existing process related traps are estimated in different SiON and HKMG devices using flicker noise method; a correlated number and surface mobility fluctuation model is used to determine trap density from measured drain current noise. The implementation details of CP and gated diode (or DCIV) techniques are discussed to determine density of generated traps during BTI stress. CP and DCIV are relatively slow methods and scan generated traps only in a limited energy range corresponding to the center of the Si band gap. Delay and band gap correction methods are discussed, which can be used to correct as-measured data for proper comparison with the trap generation component obtained from ultra-fast *I–V* methods. Increase in gate leakage in inversion or SILC is used to characterize generation of bulk traps in the gate insulator. The importance of proper choice of stress V_G is discussed, to minimize corruption of BTI degradation by TDDB like bulk trap generation. Finally, the LV-SILC method that also probes interface trap generation during BTI is briefly discussed.

In later chapters, NBTI in SiON p-MOSFETs is characterized using the mobility corrected OTF method, while flicker noise and CP methods are used respectively to characterize the pre-existing and generated traps. NBTI and PBTI in HKMG MOSFETs are characterized using the MSM method; flicker noise and DCIV are used to characterize the pre-existing and generated traps, respectively. When applicable, SILC is used to characterize generated bulk insulator traps.

Acknowledgments The authors would like to acknowledge E. Naresh Kumar and Vrajesh Maheta for OTF measurements, Gautam Kapila and Bijesh Rajamohanan for flicker noise measurements, Dhanoop Varghese for CP measurements and Applied Materials for providing devices used in this work.

References

1. N. Kimizuka, T. Yamamoto, T. Mogami, K. Yamaguchi, K. Imai, T. Horiuchi, The impact of bias temperature instability for direct-tunneling ultra-thin gate oxide on MOSFET scaling, in *Symposium on VLSI Technology: Digest of Technical Papers* (1999), p. 73
2. M. Ershov, S. Saxena, H. Karbasi, S. Winters, S. Minehane, J. Babcock, R. Lindley, P. Clifton, M. Redford, A. Shibkov, Dynamic recovery of negative bias temperature instability in p-type metal–oxide–semiconductor field-effect transistors. Appl. Phys. Lett. **83**, 1647 (2003)

3. S. Rangan, N. Mielke, E.C.C. Yeh, Universal recovery behavior of negative bias temperature instability [PMOSFETs], in *IEEE International Electron Devices Meeting Technical Digest* (2003), p. 14.3.1
4. H. Reisinger, O. Blank, W. Heinrigs, A. Muhlhoff, W. Gustin, C. Schlunder, Analysis of NBTI degradation- and recovery-behavior based on ultra fast VT-measurements, in *IEEE International Reliability Physics Symposium Proceedings* (2006), p. 448
5. E.N. Kumar, V.D. Maheta, S. Purawat, A.E. Islam, C. Olsen, K. Ahmed, M.A. Alam, S. Mahapatra, Material dependence of NBTI physical mechanism in silicon oxynitride (SiON) p-MOSFETs: a comprehensive study by ultra-fast on-the-fly (UF-OTF) IDLIN technique, in *IEEE International Electron Devices Meeting Technical Digest* (2007), p. 809
6. C. Shen, M.-F. Li, C.E. Foo, T. Yang, D.M. Huang, A. Yap, G.S. Samudra, Y.-C. Yeo, Characterization and physical origin of fast Vth transient in NBTI of pMOSFETs with SiON dielectric, in *IEEE International Electron Devices Meeting Technical Digest* (2006). doi:10.1109/IEDM.2006.346776
7. A. Kerber, M. Kerber, Fast wafer level data acquisition for reliability characterization of sub-100 nm CMOS technologies. IEEE International Integrated Reliability Workshop Final Report (2004), p. 41
8. V.D. Maheta, E.N. Kumar, S. Purawat, C. Olsen, K. Ahmed, S. Mahapatra, Development of an ultrafast on-the-fly I_{DLIN} technique to study NBTI in plasma and thermal oxynitride p-MOSFETs. IEEE Trans. Electron Devices **55**, 2614 (2008)
9. S. Mahapatra, K. Ahmed, D. Varghese, A.E. Islam, G. Gupta, L. Madhav, D. Saha, M.A. Alam, On the physical mechanism of NBTI in silicon oxynitride p-MOSFETs: can differences in insulator processing conditions resolve the interface trap generation versus hole trapping controversy?, in *IEEE International Reliability Physics Symposium Proceedings* (2007), p. 1
10. A. Chaudhary, S. Mahapatra, A physical and SPICE mobility degradation analysis for NBTI. IEEE Trans. Electron Devices **60**, 2096 (2013)
11. M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, Y. Rey-Tauriac, N. Revil, On-the-fly characterization of NBTI in ultra-thin gate oxide PMOSFET's, in *IEEE International Electron Devices Meeting Technical Digest* (2004), p. 109
12. A.E. Islam, V.D. Maheta, H. Das, S. Mahapatra, M. A. Alam, Mobility degradation due to interface traps in plasma oxynitride PMOS devices, in *IEEE International Reliability Physics Symposium Proceedings* (2008), p. 87
13. V.D. Maheta, C. Olsen, K. Ahmed, S. Mahapatra, The impact of nitrogen engineering in silicon oxynitride gate dielectric on negative-bias temperature instability of p-MOSFETs: a study by ultrafast on-the-fly I_{DLIN} technique. IEEE Trans. Electron Devices **55**, 1630 (2008)
14. S. Mahapatra, P. Bharath Kumar, M.A. Alam, Investigation and modeling of interface and bulk trap generation during negative bias temperature instability of p-MOSFETs". IEEE Trans. Electron Devices **51**, 1371 (2004)
15. Y. Taur, T.H. Ning, *Fundamentals of Modern VLSI Devices* (Cambridge University Press, Cambridge)
16. N. Goel, N. Nanaware, S. Mahapatra, Ultrafast AC–DC NBTI characterization of deep IL scaled HKMG p-MOSFETs. IEEE Electron Device Lett. **34**, 1476 (2013)
17. N. Goel, K. Joshi, S. Mukhopadhyay, N. Nanaware, S. Mahapatra, A comprehensive modeling framework for gate stack process dependence of DC and AC NBTI in SiON and HKMG p-MOSFETs. Microelectron. Reliab. **54**, 491 (2014)
18. N. Goel, S. Mukhopadhyay, N. Nanaware, S. De, R. K. Pandey, K.V.R.M. Murali, S. Mahapatra, A comprehensive DC/AC model for ultra-fast NBTI in deep EOT scaled HKMG p-MOSFETs, in *IEEE International Reliability Physics Symposium Proceedings* (2014), p. 6A.4.1
19. S. Mukhopadhyay, Private communication
20. B. Kaczer, V. Arkbipov, R. Degraeve, N. Collaert, G. Groeseneken, M. Goodwin, Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification, in *IEEE International Reliability Physics Symposium Proceedings* (2005), p. 381

21. B. Kaczer, T. Grasser, J. Roussel, J. Martin-Martinez, R. O'Connor, B.J. O'Sullivan, G. Groeseneken, Ubiquitous relaxation in BTI stressing—new evaluation and insights, in *IEEE International Reliability Physics Symposium Proceedings* (2008), p. 20
22. D. Varghese, D. Saha, S. Mahapatra, K. Ahmed, F. Nouri, M. Alam, On the dispersive versus Arrhenius temperature activation of NBTI time evolution in plasma nitrided gate oxides: measurements, theory, and implications, in *IEEE International Electron Devices Meeting Technical Digest* (2005), p. 684
23. S. Deora, P. Narayanasetti, M. Thakkar, S. Mahapatra, Development of a novel ultrafast direct threshold voltage (UF-DVT) technique to study NBTI stress and recovery. IEEE Trans. Electron Devices **58**, 3506 (2011)
24. K. Joshi, S. Mukhopadhyay, N. Goel, S. Mahapatra, A consistent physical framework for N and P BTI in HKMG MOSFETs, in *IEEE International Reliability Physics Symposium Proceedings* (2012), p. 5A.3.1
25. G. Kapila, N. Goyal, V.D. Maheta, C. Olsen, K. Ahmed, S. Mahapatra, A comprehensive study of flicker noise in plasma nitrided SiON p-MOSFETs: process dependence of pre-existing and NBTI stress generated trap distribution profiles. , in *IEEE International Electron Devices Meeting Technical Digest* (2008). doi:10.1109/IEDM.2008.4796625
26. S. Christensson, I. Lundström, C. Svensson, Low frequency noise in MOS transistors—I theory. Solid State Electron. **11**, 797 (1968)
27. F. Horng-Sen, C.-T. Sah, Theory and experiments on surface 1/f noise. IEEE Trans. Electron Devices **19**, 273 (1972)
28. Z. Celik, T.Y. Hsiang, Study of 1/f noise in N-MOSFET's: Linear region. IEEE Trans. Electron Devices **32**, 2797 (1985)
29. L.K.J. Vandamme, Model for 1/f; noise in MOS transistors biased in the linear region. Solid State Electron. **23**, 317 (1980)
30. L.K.J. Vandamme, H.M.M. de Werd, 1/f; noise model for MOSTs biased in nonohmic region. Solid State Electron. **23**, 325 (1980)
31. A.L. McWhorter, 1/f noise and germanium surface properties, in *Semiconductor Surface Physics* (University of Pennsylvania Press, Philadelphia, 1957), p. 207
32. G. Abowitz, E. Arnold, E.A. Leventhal, Surface states and 1/f noise in MOS transistors. IEEE Trans. Electron Devices **14**, 775 (1967)
33. H.E. Maes, S.H. Usmani, G. Groeseneken, Correlation between 1/f noise and interface state density at the Fermi level in field-effect transistors. J. Appl. Phys. **57**, 4811 (1985)
34. F.N. Hooge, 1/f noise. Phys. B + C **83**, 14 (1976)
35. R.P. Jindal, Phonon fluctuation model for flicker noise in elemental semiconductors. J. Appl. Phys. **52**, 2884 (1981)
36. V. Huard, Two independent components modeling for negative bias temperature instability, in *IEEE International Reliability Physics Symposium Proceedings* (2010), p. 33
37. K. Joshi, S. Hung, S. Mukhopadhyay, V. Chaudhary, N. Nanaware, B. Rajamohnan, T. Sato, M. Bevan, A. Wei, A. Noori, B. McDougal, C. Ni, G. Saheli, C. Lazik, P. Liu, D. Chu, L. Date, S. Datta, A. Brand, J. Swenberg, S. Mahapatra, HKMG process impact on N, P BTI: Role of thermal IL scaling, IL/HK integration and post HK nitridation, in *IEEE International Reliability Physics Symposium Proceedings* (2013), p. 4C.2.1
38. S. Mukhopadhyay, K. Joshi, V. Chaudhary, N. Goel, S. De, R.K. Pandey, K.V.R.M. Murali, S. Mahapatra, Trap generation in IL and HK layers during BTI / TDDB stress in scaled HKMG N and P MOSFETs, in *IEEE International Reliability Physics Symposium Proceedings* (2014), p. GD.3.1
39. S. Machlup, Noise in semiconductors: spectrum of a two-parameter random signal. J. Appl. Phys. **25**, 341 (1954)
40. K.K. Hung, P.K. Ko, C. Hu, Y.C. Cheng, A unified model for the flicker noise in metal-oxide-semiconductor field-effect transistors. IEEE Trans. Electron Devices **37**, 654 (1990)

41. T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, M. Nelhiebel, Understanding negative bias temperature instability in the context of hole trapping (Invited Paper). Microelectron. Eng. **86**, 1876 (2009)
42. H. Reisinger, T. Grasser, W. Gustin, C. Schlunder, The statistical analysis of individual defects constituting NBTI and its implications for modeling DC- and AC-stress, in *IEEE International Reliability Physics Symposium Proceedings* (2010), p. 7
43. S. Zafar, A. Kerber, R. Muralidhar, Physics based PBTI model for accelerated estimation of 10 year lifetime, in *Symposium on VLSI Technology: Digest of Technical Papers* (2014). doi:10.1109/VLSIT.2014.6894388
44. B. Kaczer, T. Grasser, J. Martin-Martinez, E. Simoen, M. Aoulaiche, P.J. Roussel, G. Groeseneken, NBTI from the perspective of defect states with widely distributed time scales, in *IEEE International Reliability Physics Symposium Proceedings* (2009), p. 55
45. G. Groeseneken, H.E. Maes, N. Beltran, R.F. De Keersmaecker, A reliable approach to charge-pumping measurements in MOS transistors. IEEE Trans. Electron Devices **31**, 42 (1984)
46. Y. Mitani, "Influence of nitrogen in ultra-thin SiON on negative bias temperature instability under AC stress, in *IEEE International Electron Devices Meeting Technical Digest* (2004), p. 117
47. S. Mahapatra, V.D. Maheta, A.E. Islam, M.A. Alam, Isolation of NBTI stress generated interface trap and hole-trapping components in PNO p-MOSFETs. IEEE Trans. Electron Devices **56**, 236 (2009)
48. S. Mahapatra, A.E. Islam, S. Deora, V.D. Maheta, K. Joshi, A. Jain, M.A. Alam, A critical re-evaluation of the usefulness of R-D framework in predicting NBTI stress and recovery, in *IEEE International Reliability Physics Symposium Proceedings* (2011), p. 6A.3.1
49. M. Zahid, R. Degraeve, M. Cho, L. Pantisano, D.R. Aguado, J. Van. Houdt, G. Groeseneken, M. Jurczak, Deffect profiling in the SiO_2/Al_2O_3 interface using Variable *T* charge-*T* discharge amplitude charge pumping (VT2ACP), in *IEEE International Reliability Physics Symposium Proceedings* (2009), p. 21
50. M. Masuduzzaman, A.E. Islam, M.A. Alam, Exploring the capability of multi frequency charge pumping in resolving location and energy levels of traps within dielectric. IEEE Trans. Electron Devices **55**, 3421 (2008)
51. W.J. Liu, Z.Y. Liu, D. Huang, C.C. Liao, L.F. Zhang, Z.H. Gan, W. Wong, C. Shen, M.-F. Li, On-the-fly interface trap measurement and its impact on the understanding of NBTI mechanism for p-MOSFETs with SiON gate dielectric, in *IEEE International Electron Devices Meeting Technical Digest* (2007), p. 813
52. D.S. Ang, Z.Q. Teo, C.M. Ng, Reassessing NBTI mechanisms by ultrafast charge pumping measurement. ieee international integrated reliability workshop final report (2009), p. 25
53. S.S. Chung, S.-J. Chen, C.-K. Yang, S.-M. Cheng, S.-H. Lin, Y.-C. Sheng, H.-S. Lin, K.-T. Hung, D.-Y. Wu, T.-R. Yew, S.-C. Chien, F.-T. Liou, F. Wen, A novel and direct determination of the interface traps in sub-100 nm CMOS devices with direct tunneling regime (12 $\sim$ 16 A) gate oxide, in *Symposium on VLSI Technology. Digest of Technical Papers* (2002), p. 74
54. T. Grasser, W. Gos, V. Sverdlov, B. Kaczer, The universality of NBTI relaxation and its implications for modeling and characterization, in *IEEE International Reliability Physics Symposium Proceedings* (2007), p. 268
55. J. Franco, B. Kaczer, J. Mitard, M. Toledano-Luque, P.J. Roussel, L. Witters, T. Grasser, G. Groeseneken, NBTI Reliability of SiGe and Ge channel pMOSFETs With SiO_2/HfO_2 dielectric stack. IEEE Trans. Device Mater. Reliab. **13**, 497 (2013)
56. J. Cai, R.-Y. Sah, Monitoring interface traps by DCIV method. IEEE Electron Device Lett. **20**, 60 (1999)
57. A. Neugroschel, G. Bersuker, R. Choi, Applications of DCIV method to NBTI characterization. Microelectron. Reliab. **47**, 1366 (2007)

58. J.H. Stathis, G. LaRosa, A. Chou, Broad energy distribution of NBTI-induced interface states in p-MOSFETs with ultra-thin nitrided oxide, in *IEEE International Reliability Physics Symposium Proceedings* (2004). doi:10.1109/RELPHY.2004.1315292
59. J.P. Campbell, P.M. Lenahan, A.T. Krishnan, S. Krishnan, NBTI: an atomic-scale defect perspective, in *IEEE International Reliability Physics Symposium Proceedings* (2006), p. 447
60. S. Mahapatra, N. Goel, S. Desai, S. Gupta, B. Jose, S. Mukhopadhyay, K. Joshi, A. Jain, A.E. Islam, M.A. Alam, A comparative study of different physics-based NBTI models. IEEE Trans. Electron Devices **60**, 901 (2013)
61. W. Shockley, W. Read, Statistics of the recombinations of holes and electrons. Phys. Rev. **87**, 835 (1952)
62. D.J. Fitzgerald, A.S. Grove, Surface recombination in semiconductors. Surf. Sci. **9**, 347 (1968)
63. M.A. Alam, SILC as a measure of trap generation and predictor of TBD in ultrathin oxides. IEEE Trans. Electron Devices **49**, 226 (2002)
64. S. Pae, T. Ghani, M. Hattendorf, J. Hicks, J. Jopling, J. Maiz, K. Mistry, J. O'Donnell, C. Prasad, J. Wiedemer, J. Xu, Characterization of SILC and its end-of-life reliability assessment on 45nm high-K and metal-gate technology, in *IEEE International Reliability Physics Symposium Proceedings* (2009), p. 499
65. S. Ramey, A. Ashutosh, C. Auth, J. Clifford, M. Hattendorf, J. Hicks, R. James, A. Rahman, V. Sharma, A. St. Amour, C. Wiegand, Intrinsic transistor reliability improvements from 22 nm tri-gate technology, in *IEEE International Reliability Physics Symposium Proceedings* (2013), p. 4C.5.1
66. S. Takagi, M. Takayanagi, A. Toriumi, Experimental examination of physical model for direct tunneling current in unstressed/stressed ultrathin gate oxides, in *IEEE International Electron Devices Meeting Technical Digest* (1999), p. 461
67. E. Cartier, A. Kerber, Stress-induced leakage current and defect generation in nFETs with HfO2/TiN gate stacks during positive-bias temperature stress, in *IEEE International Reliability Physics Symposium Proceedings* (2009), p. 486
68. J. Yang, M. Masuduzzaman, K. Joshi, S. Mukhopadhyay, J. Kang, S. Mahapatra, M.A. Alam, Intrinsic correlation between PBTI and TDDB degradations in nMOS HK/MG dielectrics, in *IEEE International Reliability Physics Symposium Proceedings* (2012), p. 5D.4.1
69. A.T. Krishnan, C. Chancellor, S. Chakravarthi, P.E. Nicollian, V. Reddy, A. Varghese, R.B. Khamankar, S. Krishnan, Material dependence of hydrogen diffusion: implications for NBTI degradation, in *IEEE International Electron Devices Meeting Technical Digest* (2005), p. 688
70. P.E. Nicollian, Insights on trap generation and breakdown in ultra thin SiO_2 and SiON dielectrics from low voltage stress-induced leakage current measurements. Microelectron. Reliab. **48**, 1171 (2008)
71. T. Yang, M.F. Li, C. Shen, C.H. Ang, C. Zhu, Y.C. Yeo, G. Samudra, S.C. Rustagi, M.B. Yu, D.L. Kwong, Fast and slow dynamic NBTI components in p-MOSFET with sion dielectric and their impact on device life-time and circuit application, in *Symposium on VLSI Technology: Digest of Technical Papers* (2005), p. 92
72. Y.M. Lin, C.J. Wang, K. Wu, A new finding on NBTI lifetime model and an investigation on NBTI degradation characteristic for 1.2 nm ultra thin oxide, in *IEEE International Reliability Physics Symposium Proceedings* (2005), p. 704

http://www.springer.com/978-81-322-2507-2

Fundamentals of Bias Temperature Instability in MOS Transistors
Characterization Methods, Process and Materials Impact, DC and AC Modeling
Mahapatra, S. (Ed.)
2016, XVI, 269 p. 201 illus., 67 illus. in color., Hardcover
ISBN: 978-81-322-2507-2