Investigation of the Temperature Dependence of Hot-Carrier Degradation in Si MOSFETs Using Spectroscopic Charge Pumping

Bernhard Ruch®, Gregor Pobegen®, and Tibor Grasser®, Fellow, IEEE

Abstract—The temperature dependence of defect generation due to hot-carrier stress (HCS) is investigated in long-channel nMOSFETs with a SiO₂ insulator. In order to improve the understanding of hot-carrier degradation, we systematically investigate the impact of temperature on the HCS-induced trap densities and their energetic profiles using standard and spectroscopic charge pumping (CP). With the aid of polyheaters, which can rapidly change the device temperature, recovery of the degradation at higher temperatures is minimized by quickly switching the temperature to the characterization temperature of −60 °C. In order to be able to quantitatively analyze the trap profiles measured with spectroscopic CP, a method for finding the correct capture cross section is suggested. Both CP methods yield comparable results, and it is shown that the same types of defects are created at all temperatures. In agreement with the literature, lower stress temperatures to the characterization temperature of −60 °C minimize the recovery of the degradation at higher temperatures. Consequently, whenever HCD is studied, border traps need to be considered.

Index Terms—Charge pumping (CP), hot carrier degradation (HCD), MOSFET, reliability, silicon, spectroscopic CP (SPCP).

I. INTRODUCTION

The impact of hot-carrier stress (HCS) on the reliability of silicon MOSFETs has been a topic of interest for decades [1] and has led to the investigation of the temperature dependence of this phenomenon [2]–[5]. It was found early that decreasing the device temperature during stress results in a larger number of defects for our long-channel devices. In this context, a transistor with a channel length above 100 nm is typically considered a long-channel device [5], in the sense that an increase in temperature will result in a decrease in HCD.

At the microscopic level, the temperature dependence of HCD can be understood by the investigation of the bond-breakage mechanism of the Si–H bonds at the Si–SiO₂ interface. Theoretical studies [5] have shown that the temperature dependence in long- and short-channel devices depends on the dominating scattering mechanisms. Electron–electron scattering, which has been suggested to be an important contributor to HCD in short-channel devices [5], causes an increase in the number of carriers with higher energies. Other scattering mechanisms, such as scattering at ionized impurities, impact ionization, and electron–phonon interactions, are dominating in long-channel MOSFETs and have the opposite effect. They cause a depopulation of the high energy tails of the carrier density functions. Therefore, an elevation of the temperature enhances the degradation effect in short-channel devices but suppresses HCD in long-channel devices [6], [7].

By desorbing hydrogen from Si–H bonds in the aforementioned reactions, a dangling bond, which is a chargeable defect, is left behind. Such a dangling bond is often referred to as P₆ center [8]. In addition to the P₆ centers that are interface traps, border traps are found after HCS [9]. Their capture and emission processes show oxide field dependencies and additional tunneling time constants due to being located in the oxide. Furthermore, their time constants are strongly temperature-dependent, resulting in a stronger impact on charge pumping (CP) measurements at higher measurement temperatures. Consequently [10], whenever HCD is studied, border traps need to be considered.

When analyzing the temperature dependence of HCD using CP, one faces three fundamental problems that could potentially lead to misleading results.

1) The experimental window in the Si bandgap accessible via CP measurements strongly depends on temperature [9]. For example, while, at −60 °C, already 83% of the bandgap is accessible, at 125 °C, this shrinks down to 63%. Although the size of this window as a function of temperature is well understood, it is difficult to correct for its impact on CP data since the correction factor will...
depend on the shape of the density of trap states in the bandgap, which is clearly not constant [9].

2) It is not fully clear how HCD recovers once the stress condition is removed [10]. In any case, it is reasonable to assume that recovery proceeds faster at higher temperatures, so HCD data taken at higher temperatures are more likely to be more strongly affected by recovery.

3) In order to analyze the impact of temperature on HCD, so far, both stress and measurement have been conducted at the same temperature. Consequently, for a meaningful comparison of the degradation, CP currents of high-temperature HCD must be compared with CP currents measured at much lower temperatures. However, given 1) and 2), these measured values would be dramatically different even if the true HCDs were temperature-independent.

In this study, we attempt to experimentally evaluate the temperature dependence of HCD in long-channel devices as rigorously as possible. Previously, no studies have investigated the impact of the temperature on the energetic distribution of HCS induced traps in these devices. Thus, no information about possible temperature dependence of the defect types has been available. We use spectroscopic CP (SPCP) [11] to show that the energetic spectrum of the defects is affected by increasing stress times in the same manner as by decreasing temperatures. This is a strong indication that HCS generates the same traps at all temperatures. Furthermore, we suggest a method to experimentally identify the effective Shockley–Read–Hall (SRH) capture cross section \( \sigma \) of the traps, which is needed for the correct quantitative analysis of the energetic trap spectra. In order to enable comparison at the same characterization temperature, degradation quenching (DQ) [12] is applied. Furthermore, DQ ensures that no traps generated by HCS are lost to recovery effects after the end of the stress due to high temperatures. This is achieved by leaving the stress conditions applied after the desired stress time has passed. The structure is then cooled down to a temperature where no recovery is expected, and the bias conditions are immediately switched from the stress conditions to measurement conditions. In this subsequent measurement, recovery can be ruled out by monitoring a parameter that is sensitive to the investigated defects, such as the drain current at a fixed gate and drain voltage. This solves all three aforementioned fundamental issues. As for 1), the highest possible bandgap is observed in the CP measurement, and for 2), recovery is suppressed, both due to the low temperature at which the measurement is performed. Also, for 3), all measurements are conducted at the same temperature.

II. STRUCTURES AND MEASUREMENT PROCEDURE

We investigate long-channel nMOSFETs with an \( n^+ \) doped polysilicon electrode and a 30-nm-thick thermal SiO\(_2\) (see Fig. 1). Their 6–\( \mu \)m long-channel clearly exceeds the threshold to be considered a long-channel device. All devices are surrounded by a polysilicon heater (polyheater, see Fig. 1) that allows for a wide range of device temperature adjustments [13] by Joule heating after a calibration process described in [9]. Measurements are performed on the wafer level with a needle probe. Unless stated otherwise, the chuck, on which the wafer is placed, is set to a temperature of \(-60\) °C, and the temperature adjustments are done with the polyheater, resulting in a dynamic range of \(-60\) °C–200°C. Note that the device temperature after a change in the polyheater quickly reaches the target temperature within about a second (error margin of 1%) [13]. To be on the safe side, large jumps in the temperature are allowed to settle for 5 (on) or 3 s (off).

HCS experiments are performed. In order to investigate the temperature-dependent defect formation caused by the stress, they are preceded and followed by SPCP and standard CP. The HCS experiment is structured into five steps (see Fig. 2):

1) **Heating Step of 5 s:** Voltage is applied to the polyheater that heats the device to the desired stress temperature \( T_{\text{stress}} \). During these 5 s, \( V_G = 2 \) V and \( V_D = 0.1 \) V are applied. This operating point results in no HCD but causes a strongly temperature-dependent drain current. Thus, the heating process can be monitored and the convergence of the temperature to its target value verified (see Fig. 3).

2) **HCS:** The device is stressed at its worst case conditions at \( V_D = 8 \) V and \( V_G = 4 \) V at the target temperature.

3) **DQ:** The stress voltages stay applied for 3 s, while the polyheater is switched off [12]. This prevents the fast recovery of defects after stress due to a high device temperature. During this short period, the devices are stressed more severely due to the decreasing temperature. This induces a negligible error that will be discussed in Section IV.

4) **\( I_D \) Recovery Measurement:** \( I_D \) is measured for 10 s at \( V_D = V_G = 1.5 \) V. This current is very sensitive to traps and a qualitative indicator for a possible trap recovery.

5) **Standard CP** [14] for 10 s: A quantitative indicator of possible trap recovery.

Before and after this HCS measurement, standard and SPCP readouts are performed to characterize the traps created by the stress. While the usual routine to identify HCS damage includes the measurement of a threshold voltage shift or change in the drain current [2], [5], we rely on CP to characterize the trap densities quantitatively.

The underlying theory of CP is the SRH theory [15] that describes four possible interactions of carriers with defects: hole and electron capture as well as hole and electron emission. In the standard CP method, the electrode on the oxide under investigation is pulsed between deep inversion and
deep accumulation. Source and drain are tied together and grounded. The bulk is grounded separately. When switching from accumulation to inversion (or vice versa), the captured electrons/holes from traps with sufficiently long emission time constants recombine with carriers of the opposite type that is flushed into the channel. Thus, only traps sufficiently far away from the valence and conduction band edges contribute to $I_{\text{CP}}$. Their emission times are sufficiently long to prevent them from being reemitted before recombining with carriers of the opposite type. This causes a net current flow $I_{\text{CP}}$ from drain/source to bulk that can be measured and is proportional to the trap density $N$ and frequency $f$ of the CP signal.

If the emission time constants of the traps are too short, carriers are reemitted during the slopes of the CP pulse and do not contribute to the CP current. The SPCP method [11] relies on exactly this parasitic effect. By varying the device temperature and then sweeping rise and fall times of the CP pulse, parts of the bandgap can be selectively deactivated. This allows for quantitative energetic scanning of a large fraction of the bandgap and yields the energy-dependent density of trap states $D(E_t)$. The energy axis is obtained by transformation of the slopes, i.e., rise and fall times, of the CP pulse

$$E_t = E_V + k_B T \ln \left( \frac{V_{\text{TH}} - V_{\text{FB}}}{\Delta V_G} \right) t_r$$

(1a)

$$E_t = E_C - k_B T \ln \left( \frac{V_{\text{TH}} - V_{\text{FB}}}{\Delta V_G} \right) t_f$$

(1b)

where $E_t$ is the trap energy, $E_V$ and $E_C$ are the valence and conduction band energies, $T$ is the device temperature, $V_{\text{TH}}$ and $V_{\text{FB}}$ are the threshold and flat-band voltages of CP [11], $\Delta V_G$ is the CP pulse amplitude, $t_r$ and $t_f$ are the rise and fall times of the CP pulse, $N_V$ and $N_C$ are the numbers of states in the valence and conduction bands, and $k_B$ is Boltzmann’s constant. The thermal velocity $v_{\text{th}}$ is often assumed to be $10^7$ cm/s for both holes and electrons [16], but we use a more accurate temperature-dependent model [17]. Finally, the capture cross section $\sigma$ is the essential parameter that describes the mapping into the bandgap. Unfortunately, it is an empirical parameter, and no accepted value exists. However, by comparing $I_{\text{CP}}$ and SPCP data, the value can be extracted with good accuracy as will be discussed in Section III-A. While (1a) describes the contribution of hole traps in the lower half of the bandgap, (1b) gives the contribution of electron traps in the upper half of the bandgap. To obtain the density of trap states, the change in the CP current in a rise and fall time sweep is used

$$D(E_t(t_r,t_f)) = (Af qk_B T)^{-1} t_f/\int t_f \frac{dI_{\text{CP}}}{dt_f}$$

(2)

where $A$ is the Si–SiO$_2$ interface area, $f$ is the CP frequency, and $q$ is the elementary charge. The complete spectrum is obtained by the sweeping rise and fall times at various temperatures; per (1), each sweep scans only a small fraction of the bandgap for the respective temperature. This is because the scanned energy level is logarithmically dependent on the transition times. Combining the sweeps at all temperatures yields the full density of states. An advantage of sweeping the bandgap piecwise by changing the device temperature
after each sweep is that those rise and fall time sweeps overlap energetically at the boundaries. This can be used for plausibility checks as the energetic regions with an overlap must give the same results. We remark that the polyheater allows for an exceptionally straightforward implementation of the SPCP method.

III. RESULTS

The analysis of the temperature dependence of HCD is split into three sections. First, a method for determining the correct capture cross section for SPCP is presented. The obtained value is used to analyze the energetic spectra of the traps caused by HCS in the second section. Finally, these densities of trap states are analyzed quantitatively.

A. Selection of the Capture Cross Section for SPCP

The selection of a suitable value for \( \sigma \) in the SPCP evaluation is often a topic of discussion as it is a free parameter in the model. Previous authors recommend values in the range of \( 1 \times 10^{-16} \text{ cm}^2 \) to \( 1 \times 10^{-14} \text{ cm}^2 \) [18] for SPCP. Generally, a very wide range of values for \( \sigma \) can be found in the literature ranging from \( 1.4 \times 10^{-25} \) [19] to \( 4 \times 10^{-13} \text{ cm}^2 \) [20], and they are often reported to be energy-dependent. In practice, this factor stretches or compresses the energy scale logarithmically and shifts it, as evident from (1) (see Fig. 4).

For a quantitative analysis, it is very important to use a correct value. In order to determine \( \sigma \), we suggest a simple method where the determination of \( \sigma \) is based on the comparison of the trap density change, which can be obtained by integrating the density of trap states obtained from SPCP, to the trap density changes measured with standard CP.

\[
\Delta N_{\text{spec}}(\sigma) = \int_{E_1}^{E_2} \Delta D(E_t(\sigma)) \, dE_t \quad (3a)
\]

\[
\Delta N_{\text{std}} = \frac{I_{\text{post}} - I_{\text{pre}}}{Afq}. \quad (3b)
\]

Since the trap density change \( \Delta N_{\text{spec}}(\sigma) \) obtained by numerical integration of the \( D(E_t(\sigma)) \) spectrum over the respective energy values [see (3a)] must be equal to the measured trap density change \( \Delta N \) of standard CP [see (3b)], the \( \sigma \) at which this condition is fulfilled can be extracted. The integration boundaries \( E_1 \) and \( E_2 \) must match the boundaries of the active CP window of the standard CP experiment that is inherently set by the temperature, as well as the rise and fall times of the CP pulse, and can be obtained from (1).

Fig. 5 shows the dependence of \( \Delta N_{\text{spec}}(\sigma) \) on the integrated trap density that is logarithmically dependent on \( \sigma \) because the integration variable \( E \) depends logarithmically on \( \sigma \) [see (1)]. Comparing \( \Delta N_{\text{spec}} \) from (3a) to \( \Delta N_{\text{std}} \) to (3b), which does not depend on \( \sigma \), yields a value of \( 14 \pm 9.0 \times 10^{-16} \text{ cm}^2 \) at the point where both curves intersect, which is in excellent agreement with the values obtained by our first proposed method. It is important to note that the evaluation, which leads to the \( D(E_t(\sigma)) \) spectrum, never uses the absolute value of the CP current but only its derivatives [see (2)]. Therefore, it is independent of the standard CP evaluation, where \( \Delta N \) is directly proportional to the CP current change before and after stress. Thus, the two methods for determining \( \Delta N \) are independent. We note that we only used the data of the HCS experiments at the lowest four temperatures as the integrated defect spectra obtained in the other experiments deviate from the trap densities found with standard CP. This is discussed more thoroughly in Section III-C.

B. Energetic Profiling

We investigate the spectrum of the traps generated by HCS with SPCP. Such a collection of trap energy densities \( D(E_t(\sigma)) \) is shown in Fig. 6 for several devices that are stressed at different stress temperatures for 20 ks. As expected for long-channel devices, increasing the stress temperature...
Fig. 6. Change in the density of states after HCS measured by the SPCP method. The legend refers to the stress temperatures. For a given CP temperature, the color bar indicates the edges of the active energy window in CP for the rise and fall times of 600 ns. All curves are smoothed with a Savitzky–Golay filter.

Fig. 7. Change in the density of states after HCS at the lowest and highest stress temperatures from Fig. 6. For comparison, two devices are stressed for shorter stress times at \(-60\,\degree C\) represented by dashed–dotted lines. A decrease in the impact of HCS is observed as the stress temperature increases, showing no additional peaks or kinks. This indicates that the temperature only acts as a deceleration factor for HCD. To confirm this, we compare the trap profiles generated by HCD after devices are stressed at high and low temperatures to the trap profiles measured after varying stress times at the low temperature. Fig. 7 shows a comparison of HCD after 20 ks of stress at 186 \(\degree C\) to measurements after 1-, 4-, and 20-ks HCSs at \(-60\,\degree C\). The spectra of the long measurements at 186 \(\degree C\) and \(-60\,\degree C\) show the lowest and highest \(D(E_t(\sigma))\), respectively. The cold and shorter measurements fall right between them as if the stress temperature instead of the stress time is varied. If different types of defects were generated depending on the stress temperature, a smooth transition between these spectra would be highly unlikely. It could only be explained if the different hypothetical microscopic defects resulted in the same energetic spectrum. This leads us to the conclusion that the same types of defects are generated at all temperatures albeit at a different rate.

C. Quantitative Analysis of Hot-Carrier Induced Damage

In this section, the spectra obtained with SPCP are quantitatively analyzed and compared with trap densities obtained from standard CP. Considering the previously found capture cross section, the HCD energy distributions in the bandgap and the trap densities found in standard CP can be compared. The parameter used for the comparison is the number of traps after HCS, \(N\), which can be easily calculated for both standard and SPCP [see (3b) and (3a)]. The results of this analysis are shown in Fig. 8. As previously discussed, the temperature has a massive decelerating impact on HCD. The integrated spectra fit the trap densities measured by standard CP very well. A deviation of spectroscopic and standard CP is observed in the measurements at \(-60\,\degree C\) after HCS at high-stress temperatures although the readouts at stress temperature show little deviation. It can, therefore, be concluded that the energetic profiles of low degradation overestimate the density of states near the band edges. This discrepancy of readouts at \(-60\,\degree C\) and stress temperature is caused by the different energy windows of traps captured by CP, which decreases with increasing temperatures. For comparison, measurements are also conducted without the use of the polyheater (black in Fig. 8). Instead, the chuck temperature of the needle prober is adjusted. Quite reassuringly, they do not deviate significantly from the polyheater measurements. It is, therefore, safe to assume that the polyheater temperatures are adjusted correctly. Due to the matching of \(N\) in both CP methods, it can be
concluded that the spectra are not only qualitatively but also quantitatively correct. Also, the possible uncertainty of the value of $\sigma$ found in Section III-A does not introduce large error margins into the quantitative analysis of the data.

IV. Conclusion

The temperature dependence of HCD has been rigorously investigated for long-channel nMOS transistors. With the help of an embedded polysilicon heater setup, DQ to minimize possible sources of error is made possible. In this method, the devices are stressed at arbitrary temperatures above the chuck temperature. Before the measurement of trap densities, the device is cooled down with the stress bias still applied. In this way, the recovery of stress-induced damage due to the high-stress temperatures can be suppressed. In the subsequent measurements, indeed, no recovery is observed. In addition, the spectra of the hot-carrier induced traps are measured with SPCP. It is shown that changing the stress temperature has the same impact on the density of states as varying the stress time. This is evidence that lowering the temperature during HCS only accelerates the degradation process but does not influence the type of defects created. Comparing the results of spectroscopic and standard CP yields a capture cross section $\sigma$ of $14 \pm 9.0 \times 10^{-16}$ cm$^2$ for the underlying SRH model of SPCP that is consistent with the literature values. Furthermore, it is shown that the integrated spectra of SPCP lead to very similar results as those obtained from standard CP. This confirms that the SPCP method yields quantitatively correct trap densities. Deviations are only observed at low-temperature CP measurements that indicate an overestimation of the spectra near the valance and conduction band edges.

Appendix

Possible Errors of the HCS Experiment

The design of the HCS experiment (see Fig. 2) introduces two possible sources of error. First, the stress biases remain applied during the DQ step that opens the possibility of additional defect generation. Second, some of the stress-induced defects may recover directly after the stress.

As for the first issue, the quantitative analysis of the trap densities has to consider the impact of DQ. While allowing the device to cool down after stress, the stress conditions are still applied for 3 s. Thus, the worst possible error due to DQ can be defined as the trap density a device would accumulate in HCD after a 3-s stress at $-60^\circ$C. This overestimates the real error for two reasons. On the one hand, the device cools down during the DQ period and will experience less severe stress for the time during which the temperature is above $-60^\circ$C. On the other hand, HCD is not linearly accumulative, and more traps are generated at the beginning of the stress. Thus, the stress during DQ on top of the previous stress will be less severe than when a new device is stressed for 3 s. As our devices show relatively low degradation even after 20 ks of stress, we do not measure the degradation after 3 s directly on a fresh device to estimate the impact of the DQ. Instead, we extrapolate on data points after various longer stress times back to 3 s to avoid measurement errors of low trap densities due to noise. Fig. 9 provides such a measurement. After 3 s of HCS, a trap density of approximately $\Delta N = 4.5 \times 10^{16}$ cm$^{-2}$ can be expected. The extrapolation is based on a power law that is commonly used for modeling HCD [21], [22].

Following the DQ step, the bias conditions are switched from stress to an operation point where a defect sensitive drain current can be monitored (Step 4). This allows for the detection of possible poststress recovery of traps. Subsequently, CP is applied for quantitative analysis of recovery (Step 5). In our experiments, no recovery of defects is observed neither in the drain current nor in the CP signal, as shown in Fig. 10.

Acknowledgment

The authors would like to thank Florian Peter Prähbsnack for fruitful discussions about the performance of data processing algorithms.

References


