Energetic mapping of oxide traps in MoS₂ field-effect transistors

Yury Yu Illarionov, Theresia Knobloch, Michael Waltl, Gerhard Rzepa, Andreas Pospischil, Dmitry K Polyushkin, Marco M Furchi, Thomas Mueller and Tibor Grasser

1 Institute for Microelectronics (TU Wien), Gusshausstrasse 27–29, 1040 Vienna, Austria
2 Ioffe Physical-Technical Institute, Polytechnicheskaya 26, 194021 St-Petersburg, Russia
3 Institute for Photonics (TU Wien), Gusshausstrasse 27–29, 1040 Vienna, Austria

E-mail: illarionov@iue.tuwien.ac.at and grasser@iue.tuwien.ac.at

Keywords: MoS₂, Al₂O₃, SiO₂, transistor, oxide traps, charge trapping, hysteresis sweep

Supplementary material for this article is available online

Abstract
The performance of MoS₂ transistors is strongly affected by charge trapping in oxide traps with very broad distributions of time constants. These defects degrade the mobility and additionally lead to the hysteresis of the gate transfer characteristics, which presents a crucial performance and reliability issue for these new technologies. Here we perform a detailed study of the hysteresis in double-gated MoS₂ FETs and show that this issue is nothing else than a combination of threshold voltage shifts resulting from positive and negative bias-temperature instabilities. While these instabilities are well known from silicon devices, they are even more important in 2D devices given the considerably larger defect densities. Most importantly, the magnitudes of these threshold voltage shifts depend strongly on the density and energetic alignment of the active oxide traps. Based on this, we introduce the incremental hysteresis sweep method which allows for an accurate mapping of these defects and extract their energy distributions from simulations. By applying our method to analyze the impact of oxide traps situated in the Al₂O₃ top gate of several devices, we confirm its versatility. Since all 2D devices investigated so far suffer from a similar hysteresis behavior, the incremental hysteresis sweep method provides a unique and powerful way for the detailed characterization of their defect bands.

Introduction
Molybdenum disulfide (MoS₂) is a next-generation semiconductor from a wide range of transition metal dichalcogenides which is now considered for applications in beyond-CMOS electronic devices. Owing to a direct electronic bandgap of around 2.6 eV in the single-layer limit [1, 2], MoS₂ has attracted considerable attention for digital device applications. In particular, numerous successful attempts at fabricating MoS₂ field-effect transistors (FETs) have been undertaken recently [3–10]. However, these studies mostly deal with the analysis of the performance characteristics of these devices, such as mobilities and on/off current ratios, as well as the exploration of fabrication techniques allowing them to realize their theoretical performance potential predicted by simulations [11]. In addition, there is some understanding that MoS₂ FETs are suitable for circuit integration [12–14] and high-frequency applications [15]. However, one of the most important performance limitations is due to defects. Even though at the moment these defects severely reduce the potential for industrial integration of these new technologies, they have not yet received the attention they deserve.

The most critical component in terms of hysteresis and reliability is the insulator materials itself and in particular its interface to the semiconducting channel. Many of these materials and interfaces have already received a lot of attention in Si transistors [16] and other electronic devices, such as flash memory cells [17]. These studies have shown that every insulator studied up to now contains some preexisting defects [18] with widely distributed time constants [19], which can act as trapping sites for the charge carriers in the channel, and thus affect the device performance and reliability. In particular, the most obvious consequence of oxide traps is the ubiquitous charge trapping, which can lead to long-term drifts of the transfer characteristics during device operation or under the presence of gate bias stress. These issues are known as bias-temperature instabilities (BTI) and are a serious reliability concern in conventional Si technologies [20–25]. During the last decade BTI in Si FETs has been...
thoroughly characterized [22, 24, 26–28]. In particular, charge trapping was identified to be a crucial contributor to the degradation [19, 29]. At the same time, the reliability of next-generation 2D devices, which is known to be far below the standards of modern industrial FETs [30–33], is still poorly understood. As such, commercialization of these new technologies requires systematic reliability assessment studies, which would allow to understand and minimize charge trapping in 2D FETs.

Although the reliability of MoS2 FETs and other 2D devices is not properly understood yet, the similarities to available data on Si technologies clearly suggest that charge trapping in oxide traps is at the heart of the problem [34–39]. As the dynamics appear to be very similar to Si FETs [40], we expect that they can be described using the models previously developed for Si technologies [25, 29, 41]. However, the most ubiquitous issue, which is typically observed in addition to partially recoverable BTI shifts [30, 31, 42], is the hysteresis of the gate transfer characteristics [34–37]. While in mature Si technologies the hysteresis is negligible, in MoS2 devices this issue can still lead to considerable instabilities of the device characteristics. In our previous work [40] we have performed a combined study of the hysteresis and BTI in MoS2/SiO2 FETs and speculated that both the hysteresis and BTI are due to the same type of oxide traps, with the faster ones leading to the hysteresis and slower ones to BTI. However, the oxide traps are known to be localized within certain defect bands which are unique for every insulator [18]. As such, information about the density and energetic alignment of the defects would allow for a considerable advancement in our general understanding of the performance and reliability of MoS2 FETs.

Here we perform a detailed study of the hysteresis behavior in double-gated MoS2 FETs and show that the observed instability of the threshold voltage is a consequence of what is typically referred to as positive and negative BTI (PBTI and NBTI, respectively). Based on this, we introduce the incremental hysteresis sweep method which allows us to perform an accurate mapping of oxide traps with different time constants. The main ingredients of this method are the experimental technique, which gives the density of active oxide traps at different gate voltages, and the technology computer aided design (TCAD) simulations which allow to link the applied gate voltage with the trap level alignment in a gate insulator. Thus, the results of our method are the energy distributions of the density of oxide traps with different time constants. In order to demonstrate the versatility of our approach, we apply the method to extract the distributions of oxide traps in the Al2O3 top gate insulator of our MoS2 FETs. Although we demonstrate that the density of oxide traps in Al2O3 layers is relatively large, our results are instrumental for understanding the charge trapping dynamics in this high-k oxide, which is typically used either as a top gate insulator [7, 32, 43] or encapsulation layer [33, 44] in many 2D devices.

Note that our method targets exactly those oxide traps which are responsible for the hysteresis and BTI in MoS2 FETs. These traps are typically localized within certain defect bands which are a fundamental property of any insulator. This is in contrast to some other techniques based on capacitive measurements [45] or on the analysis of the subthreshold swing [46]. For example, the authors of [45] extract the density of band tail trapping states in MoS2 FETs, which result from the impact of the defects in MoS2 and interface states. At the same time, the method of [46] allows to extract the density of localized states in the semiconductor bandgap using the subthreshold swing model. However, this approach does not suggest any distinction between semiconductor defects and oxide traps. Furthermore, the analysis is done in a simplistic manner ignoring the impact of oxide traps, which is not suitable for 2D FETs.

**Devices**

Our devices are single-layer and bilayer double-gated MoS2 FETs. After chemical vapor deposition (CVD) of the MoS2 film, it was transferred onto a 280 nm thick oxide coated silicon wafer, which was used as a backgate dielectric. Following fabrication of source and drain contacts, the devices were coated with a 23.5 nm thick Al2O3 top-gate dielectric using atomic layer deposition (ALD). The channel length \( L \) of the devices varies from 500 nm to 3 \( \mu \)m, while the channel width \( W \) is kept constant at 3 \( \mu \)m. Top gate electrodes were designed to be slightly smaller than \( L \) (350 nm to 2 \( \mu \)m) to avoid electrical break-through. A schematic layout and microscope image of the devices are shown in figures 1((a) and (b)). Electrical characterization shows that the devices exhibit a relatively high on/off current ratio, which can reach \( 10^6–10^7 \) for single-layer devices in both top gate and back gate operation modes (figures 1((c) and (d))). Single-layer devices with \( L = 500 \text{ nm} \) were selected for the detailed study.

**Charge trapping in MoS2 FETs: basics**

Charge trapping in preexisting oxide traps is one of the major issues known to affect the reliability of Si technologies [20–25]. These defects are energetically localized within certain defect bands and present a fundamental property of every insulator [18]. During device operation, traps within several nanometers from the oxide/channel interface, known as border traps [47], can exchange charge with the channel by means of carrier capture or emission via tunneling process. These charge trapping events are well described by non-radiative multiphonon processes [25, 48, 49] and their dynamics depend on the capture and emission time constants \( \tau_c \) and \( \tau_e \), respectively. The essential aspect of these processes is that the time constants are dominated by structural relaxation at the defect sites rather than the tunneling probabilities. Due to the amorphous nature of most oxides, these time constants are widely
distributed for different defects and present the time which is required for each particular defect to capture or emit a carrier under favorable bias conditions. Note that the most important characteristic of a non-radiative multiphonon process is its strong temperature- and bias-dependence.

As has been found recently, next-generation 2D devices [32, 33, 50, 51], in particular MoS₂ FETs [30, 31, 34–37, 40, 42], also suffer from the charge trapping at preexisting oxide defects. The dynamics of the underlying processes are surprisingly similar to Si technologies [32, 33, 40]. Depending on the microscopic structure of the defect, studied in detail for SiO₂ based on Si/SiO₂ FETs, one usually speaks of electron or hole traps respectively. The charge transfer process itself is the same in both cases. The two processes only differ in the change of the charge of the trapping defect in the oxide. This defect either goes from positive to neutral (hole trap) or from neutral to negative (electron trap). Therefore, the difference between electron and hole traps is only visible in an offset of the transfer characteristic, as it only changes the balance of fixed charges. From the overall charge balance required for our TCAD simulations, we conclude that the defect band in Al₂O₃, which dominates the charge capture and emission processes causing the hysteresis, is an electron trapping band.

In figure 2 we schematically illustrate the charge trapping in our devices operated in the top gate mode. In equilibrium, which corresponds to the flat-band voltage \( V_{fb} \), the defects localized below the Fermi level \( E_F \) are negatively charged, while the ones above \( E_F \) are neutral. At the same time, the threshold voltage of the device strongly depends on the concentration of charged defects and can be given as

\[
V_{th} \approx V_{th}^{eq} + \frac{q(N_{ot} - N_{ot}^{eq})}{C_{tg}} \tag{1}
\]

with \( q \) being the elementary charge, \( C_{tg} \) the top gate oxide capacitance, \( N_{ot} \) the concentration of charged defects, \( V_{th}^{eq} \) the equilibrium threshold voltage and \( N_{ot}^{eq} \) the concentration of charged defects. In particular, if \( V_{tg} < V_{fb} \) is applied, band-bending shifts most defects above \( E_F \). As such, charged defects, except those with very large emission time constants, can emit an electron into the channel and become neutralized (i.e. discharged). Thus \( N_{ot} \) becomes smaller, which makes the threshold voltage \( V_{th} \) of the device more negative. This issue is known as NBTI. Conversely, if \( V_{tg} > V_{fb} \) is applied, a considerable number of defects is below \( E_F \), which is close to the conduction band of MoS₂. Thus, neutral defects, except those with very large capture time constants, can capture an electron from the channel and become charged. As a result, \( N_{ot} \) becomes larger and \( V_{th} \) is more positive, which is known as PBTI.

The most obvious consequence of both issues on the performance of MoS₂ FETs is the hysteresis of the gate.

![Figure 1](image-url). (a) Schematic layout of our double-gated MoS₂ FETs. The top gate insulator is 23.5 nm thick Al₂O₃ and the back gate is 280 nm thick SiO₂. (b) A microscope image of our CVD MoS₂ devices. (c) Back gate and (d) top gate transfer characteristics measured using the autorange mode.
transfer characteristics [34–37]. Namely, \( V_{th} \) measured using a \( V^+ \) sweep (from \( V_{g} \) to \( V_{tgmax} \)) is typically more negative than the one measured using \( V^- \) sweep (from \( V_{g} \) to \( V_{tgmin} \)), which is because \( N_{tot} \) at \( V_{g} \) is smaller than at \( V_{tg} \). Furthermore, those oxide traps which are too slow to follow the hysteresis sweeps remain charged and can cause long-term NBTI and PBTI drifts which appear if a constant gate bias stress is applied for a considerable time and can be recovered if the device is returned back to the equilibrium (for more details see our previous work [40]). Below we will show that the charge trapping behavior can be used as an efficient instrument for accurate mapping of oxide traps in MoS\(_2\) FETs.

**Results and discussions**

We performed all our measurements in complete darkness and in a vacuum (\( 5 \times 10^{-6} \rightarrow 10^{-5} \) torr, \( T = 27 \) °C). The latter was necessary to avoid the detrimental impact of the ambient [36]. The hysteresis was investigated by measuring the gate transfer characteristics at \( V_g = 0.5 \) V using both \( V^+ \) and \( V^- \) sweep directions. In order to capture the impact of oxide traps with widely distributed time constants, we varied the sweep rate \( S = \frac{V_{step}}{t_{step}} \) between 0.02 and 5000 V s\(^{-1}\) by adjusting the step voltage \( V_{step} \) and the sampling time \( t_{step} \). While using smaller \( V_{step} \) allowed us to access more oxide traps, a larger \( t_{step} \) allowed to increase the amount of slower traps which are able to contribute to the hysteresis. Also, we mostly focused on the analysis of the hysteresis on the top gate transfer \( (I_d-V_g) \) characteristics and used different sweep ranges \( V_{gmin} \) to \( V_{gmax} \). The hysteresis width \( \Delta V_{tg} \) was measured around \( V_{tg} \), which was extracted using a constant current method at \( I_d = 10 \) nA.

In figure 3 we compare the \( I_d-V_g \) characteristics measured using both sweep directions and different sweep rates for the \( V_{tg} \) sweep range from \(-10 \) to \(-3 \) V. The threshold voltages measured using the \( V^+ \) sweep mode become more negative as \( S \) is decreased. As shown by the schematic band diagrams (figure 3(b)), this is because around \( V_{gmin} \), most defects are above the Fermi level \( E_F \), which allows their efficient discharging by means of electron emission. As a result, an NBTI degradation is observed, which is more pronounced for slower sweeps, i.e. larger stress times. Conversely, around \( V_{gmax} \), we are dealing with PBTI degradation, which is associated with charging of some defects. Thus, the threshold voltages measured using \( V^- \) sweeps are more positive compared to their \( V^+ \) sweep counterparts, and a clockwise, i.e. PBTI-like, hysteresis is observed. Since the magnitude of PBTI degradation is strongly dependent on \( S \), which determines the stress time, for slower sweeps the hysteresis becomes larger. However, if the sweep rate is as fast as 5000 V s\(^{-1}\), we do not see any considerable hysteresis, while \( V_{tg} \) is more positive than even for \( S = 100 \) V s\(^{-1}\). This means...
that the time constants of most oxide traps accessible within this narrow sweep range are larger than the corresponding sweep time, which is around 1 ms. As such, we assume that the $I_d-V_g$ characteristic measured using $S = 5000 \text{ V s}^{-1}$ is weakly affected by charge trapping and take it as a reference curve (see more details in figure S1 in the supporting information (SI) (stacks.iop.org/TDM/4/025108/mmedia)). The latter allows us to split the total hysteresis width $\Delta V_{th}$ into the threshold voltage shifts $\Delta V_{th}^+$ and $\Delta V_{th}^-$ obtained for the $I_d-V_g$ characteristics measured using the $V^+$ and $V^−$ sweep modes, respectively. As shown in figure 3(c), both shifts are of NBTI-like nature, while being larger for smaller measurement frequencies $f = 1/N \text{step}$ with $N$ being the number of $V_g$ steps of duration $f_{\text{step}}$ [40]. This is because for the narrow sweep range the major fraction of the total sweep time is spent at $V_g$ corresponding to an NBTI bias condition, i.e. $V_g$ is below the equilibrium voltage. Nevertheless, $\Delta V_{th}^+$ is smaller than $\Delta V_{th}^−$ due to PBTI degradation, which occurs more close to $V_{g\text{max}}$ and becomes stronger for smaller $f$. The latter leads to the observed hysteresis.

The results corresponding to the $V_g$ sweep range from $−10$ to $14 \text{ V}$ are shown in figure 4. While the $I_d-V_g$ curve measured using the $V^+$ sweep mode is shifted in an NBTI-like manner with respect to $S = 5000 \text{ V s}^{-1}$ curve, a larger $V_{g\text{max}}$ leads to a PBTI-like shift of the $V^−$ sweep characteristics. Thus, a considerable hysteresis is observed, which is a result of both NBTI and PBTI degradation. While the NBTI magnitude remains comparable to the case of narrow sweep ranges, the PBTI contribution is dramatically increased. The latter is because the larger $V_{g\text{max}}$ increases both the number of oxide traps which can be charged (figure 4(b)) and the total time spent at a PBTI bias condition, when most of the defects are shifted below $E_F$ due to the band bending. Since both NBTI and PBTI shifts become larger for slower sweeps, the hysteresis width also increases. However, for very slow sweep frequencies $\Delta V_{th}$ tends to saturate, which is consistent with the universal behavior reported in our previous work [40]. The results obtained on the back gate can be found in figure S2 in the SI.

The results above show that the number of oxide traps which are able to contribute to the hysteresis depends strongly on $V_{g\text{max}}$, which determines the magnitude of the PBTI contribution and hence the total hysteresis width. In addition, the number of traps able to contribute strongly depends on the sweep rate, as the traps with capture/emission times larger than the sweep time will not be able to react. Thus, aiming to map these defects with their widely distributed time

---

Figure 3. (a) The $I_d-V_g$ characteristics of our single-layer MoS$_2$ FETs measured using the $V_g$ sweep range from $−10$ to $−3 \text{ V}$ and different sweep rates. Both the $V^+$ and $V^−$ curves are shifted in an NBTI-like manner with respect to the $S = 5000 \text{ V s}^{-1}$ curve. (b) This is due to discharging of oxide traps around $V_{g\text{max}}$, which becomes more efficient for slower sweeps and presents nothing else than an NBTI degradation. Conversely, when approaching $V_{g\text{max}}$ there is some charging of oxide traps, i.e. PBTI degradation. (c) The hysteresis width $\Delta V_{th}$ can be treated as a difference between the $V_{th}$ shifts of the $I_d-V_g$ characteristics measured using the $V^+$ and $V^−$ sweep modes.

Figure 4. (a) The $I_d-V_g$ characteristics of our single-layer MoS$_2$ FETs measured using the $V_g$ sweep range from $−10$ to $14 \text{ V}$ and different sweep rates. Contrary to figure 3, the transfer characteristics measured using the $V^−$ sweep mode are shifted in a PBTI-like manner with respect to the $S = 5000 \text{ V s}^{-1}$ curve. (b) The reason for this is a larger $V_{g\text{max}}$, which leads to a larger fraction of sweep time spent at $V_g$ corresponding to PBTI condition and also increases the number of accessible defects. (c) Owing to a dramatically increased PBTI contribution, the hysteresis width is considerably larger than it was for the narrower sweep range.
constants and different energy levels, we employed the following experimental technique which presents the main ingredient of our incremental hysteresis sweep method. An elementary loop consists of measurements of the $I_g-V_g$ characteristics using both $V^+$ and $V^-$ sweep directions with a fixed sweep range and different $V_{\text{step}}$ and $t_{\text{step}}$ [40]. As shown in figure 5(a), the full measurement procedure consists of repeated loops of this kind for different sweep ranges using a fixed $V_{\text{tgmin}} = -10 \text{ V}$ and $V_{\text{tgmax}}$ varied from $-3$ to $14 \text{ V}$ in $1 \text{ V}$ steps. This allows us to obtain a set of $\Delta V_{\text{tg}}(f)$ characteristics which strongly depend on $V_{\text{tgmax}}$ and thus contain the information about the energy distribution of the density of charged oxide traps with different time constants (see figure S3 in the SI). Next we follow the approach of figures 3–4 and evaluate the PBTI and NBTI contributions into the total hysteresis widths by splitting the $\Delta V_{\text{tg}}(f)$ characteristics into $\Delta V_{\text{tg}}^\text{th}(f)$ and $\Delta V_{\text{tg}}^\text{th}(f)$ parts. The results measured using $V_{\text{tgmin}} = -10 \text{ V}$ and $V_{\text{tgmax}}$ between $-3$ and $14 \text{ V}$ for single-layer devices are shown in figure 5(b). Since we have used the same $V_{\text{tgmin}}$ for all these measurements, the $\Delta V_{\text{tg}}^\text{th}(f)$ characteristics, which are associated with the NBTI contribution into the total $\Delta V_{\text{tg}}$, are nicely reproducible. Some negligible variations of $\Delta V_{\text{tg}}^\text{th}(f)$ originate from a slight drift of the device in between the measurement loops with different $V_{\text{tgmax}}$. Conversely, the $\Delta V_{\text{tg}}^\text{th}(f)$ curves, which contain the fingerprint of the PBTI contribution, follow an increase of $V_{\text{tgmax}}$. The latter is because the defect band of Al$_2$O$_3$ is bent by applying a top gate voltage, which shifts the traps below the Fermi level. Thus, an increase in the maximum of the sweep range by an interval $\Delta V_{\text{tgmax}}$ bends the defect band downwards a bit stronger. In this way, traps which formerly have been situated above the Fermi level during the whole sweep can now contribute as well. In other words, if higher top gate voltages are applied, defects with a higher energy level can be accessed. At the same time, the high carrier density in the accumulation regime assures that the band bending affects first and foremost the Al$_2$O$_3$ layer. As such, after polynomial smoothening of the obtained $\Delta V_{\text{tg}}^\text{th}(f)$ characteristics (see figure S4 in the SI) we calculate the concentration of oxide traps which come into play between $V_{\text{tgmax}}^i$ and $V_{\text{tgmax}}^{i+1}$ as

$$\Delta N_{\text{tg}}^i(f) = \left( \Delta V_{\text{tg}}^i(f) - V_{\text{tgmax}}^{i+1} - V_{\text{tgmax}}^i \right) \frac{C_{\text{tg}}}{q}$$

with $C_{\text{tg}}$ being the top gate oxide capacitance and $q$ the elementary charge. In order to be able to contribute to the charge trapping processes, these traps should be able to capture the electrons which are tunneling from the MoS$_2$ channel through the top gate dielectric. As such, the active oxide traps should be situated not farther than a maximum distance $d_{\text{max}} \approx 2.6 \text{ nm}$ from the MoS$_2$/Al$_2$O$_3$ interface (see the detailed evaluation in the SI). Therefore, the oxide trap density within the device operation range can be estimated as

$$D_{\text{tg}} \left( V_{\text{tg}} = \frac{V_{\text{tgmax}}^i + V_{\text{tgmax}}^{i+1}}{2}, f \right) = \left| \frac{\Delta N_{\text{tg}}^i(f)}{d_{\text{max}} V_{\text{tgmax}}^{i+1} - V_{\text{tgmax}}^i} \right|$$

Processing of the results shown in figure 5(b) using the equations (2)–(3) allows us to obtain the distribution of the oxide trap density for $V_{\text{tg}}$ between $-2.5$ and $13.5 \text{ V}$. In order to link $V_{\text{tg}}$ with the Al$_2$O$_3$ trap level $E_T$, we next perform TCAD simulations with the parameters taken from the literature [1, 2, 52, 53]. Namely, $E_{\text{MoS}_2} = 2.6 \text{ eV}$, $E_{\text{Al}_2\text{O}_3}^\text{Al} = 6.7 \text{ eV}$, and $E_{\text{MoS}_2} = 3.74 \text{ eV}$, $E_{\text{Al}_2\text{O}_3}^\text{Mo} = 1.35 \text{ eV}$ have been used for the electronic bandgap and the electron affinity of single-layer MoS$_2$ and ALD-grown Al$_2$O$_3$, respectively. Furthermore, the four-state non-radiative multphonon (NMP) model known from Si technologies [25] has been implemented to describe the charge trapping by oxide traps which leads to the hysteresis. In figures 6(a) and (b) we show that both the shape of the $I_g-V_g$ characteristics and the hysteresis dynamics measured using different sweep rates and sweep ranges can be reasonably well matched by our TCAD simulations. Furthermore, simulations of a large number of $I_g-V_g$ characteristics with different sweep parameters allowed us to extract the hysteresis widths and reproduce the experimental set of the $\Delta V_{\text{tg}}(f)$ characteristics for different $V_{\text{tgmax}}$ (figure 6(c)).
allowed us to validate our experimental approach and to determine the position of the trap level $E_T$ in $\text{Al}_2\text{O}_3$.

The band diagrams for the $\text{SiO}_2$/MoS$_2$/$\text{Al}_2\text{O}_3$ system underlying the results of figure 6 are shown in figure 7(a). For $\text{SiO}_2$ we are using two distinct defect bands which have been identified in our previous works [33, 54]. The upper defect band is located at $E_T = -2.75 \pm 0.4 \text{ eV}$ below the $\text{SiO}_2$ conduction band edge [33], which almost exactly matches the value previously reported for Si technologies ($\sim 2.6 \text{ eV}$) [55]. The lower defect band, also known from Si technologies [54], is at $E_T = -4.56 \pm 0.35 \text{ eV}$ below the $\text{SiO}_2$ conduction band edge.

[Figure 6 and Figure 7]
in reasonable agreement with the one previously identified for Si technologies (~2.0 eV) [55], which further confirms that every insulator has its unique defect bands. The distributions of oxide traps for the defect bands in SiO2 and Al2O3 used in our TCAD simulations are shown in figure 7(b). While our mapping range from Vg = −3 V to 14 V corresponds to accumulation, the Fermi level is pinned close to the conduction band of MoS2, which agrees with previous literature reports [56–58]. Thus, in agreement with our qualitative interpretation above, for more positive Vg the number of Al2O3 defects in the active region, i.e. below the Fermi level EF, increases due to band-bending which changes the shape of the Al2O3 defect band and brings more defects downwards. This leads to their efficient charging and, consequently, a larger hysteresis for wider sweep ranges.

The results of figures 7(a) and (b) allow us to recalculate the applied Vg into the trap level E[τ] by considering the band-bending of the Al2O3 defect band within dmax ≈ 2.6 mm from the interface. In figure 7(c) we show the obtained E[τ](Vg) dependence with Eτ given in the units of electronic energy used in the band diagrams. We notice that there is a saturation at larger Vg which is an artifact related to the discrete defect bands used in our TCAD simulations (see figure S5 in the SI). Therefore, in order to obtain a more physical Eτ(Vg) dependence, we use a linear approximation. With this mapping we can convert our experimental D[τ](Vg) dependences into the differential energy distributions D[τ](E). In figure 7(d) we show the D[τ](E) curves obtained for the measurement frequency spaced logarithmically between 10 and 1 × 10−3 Hz, which corresponds to capture times τc ∼ 1/τ from 10−1 to 103 s. The typical defect densities determined for the Al2O3 defect band are similar to those previously obtained for the same oxide in Si devices (~1020 cm−3 eV−1) [55]. At the same time, the D[τ](E) distributions are broadly consistent with the simple Gaussian shape used in our TCAD simulations (see figure S6 in the SI), although additional peaks, likely associated with some imperfections of real devices, are present. Furthermore, the dependence of D[τ] versus the measurement frequency along the defect band is also consistent with the shift of the Gaussian peak which follows from our TCAD simulations (figure S6). In particular, we can experimentally resolve three regions with different D[τ](f) behavior. In the lower (L) region, i.e. below the D[τ](E) peak, we observe a monotonic increase of D[τ] as f is decreased. This means that the capture times are distributed within the whole interval [10−5 s...102 s], while the number of slower traps is larger (figure 7(e)). In the middle (M) region, i.e. around the D[τ](E) peak, the D[τ](f) dependence is weak, which suggests that the fractions of slower and faster traps are comparable. Finally, in the upper (U) region D[τ] decreases for smaller f, i.e. faster traps dominate. Thus, at very slow sweeps most defects have enough time to become charged, which leads to some saturation in the magnitude of the PBTI contribution and, consequently, in the total hysteresis width. The latter is fully consistent with the results of our TCAD simulations and some experimental observations reported in the previous work [40], which suggest that the maximum of ΔVh can be reached at a comparably small measurement frequency.

Finally, we have performed a detailed verification of our method by comparing the D[τ](E) distributions extracted from the simulated ΔVh(f) curves (figure 6(c)) with the total trap density which has been originally used in the TCAD simulator (see the details in figures S7–8 in the SI). While a reasonable agreement has been achieved, we found that the peak D[τ] values extracted using our approach are smaller than those for the input trap density (e.g. figure 7(b)). This observation confirms that our incremental hysteresis sweep method is sensitive exactly to those oxide traps which contribute to the hysteresis, while the number of traps which can be captured depends on the ratio between 1/f and the time constants. At the same time, the typical D[τ] ∼ 1020 cm−3 eV−1 obtained for MoS2/Al2O3 system are comparable to those reported previously for MoS2/SiO2 FETs [59], while being larger than for SiGe/Al2O3 (∼ 5 × 1018 cm−3 eV−1) [60] devices, not to mention the Si/SiO2 FETs (< 1017 cm−3 eV−1) [61].

Conclusions

In summary, we have performed a detailed study of the hysteresis dynamics in double-gated MoS2 FETs. We found that this issue is a consequence of device degradation due to positive and negative bias-temperature instabilities, which are related to charging and discharging of oxide traps, respectively. Based on this finding, we have developed the incremental hysteresis sweep method which allows to perform an accurate mapping of oxide traps with widely distributed time constants. By using the experimental technique of our method and TCAD simulations to convert the applied top gate voltage into the electronic energy, we have extracted the differential energy distributions of oxide traps in single-layer MoS2 FETs and confirmed the validity of our approach. Taking into account that the hysteresis appears to have the same origin in all new 2D technologies, we are confident that the reported method is universal.

Methods

Device fabrication

Single-layer MoS2 was grown by chemical vapor deposition (CVD) on c-plane sapphire similarly to the method suggested in [62]. After cleaving the sapphire substrates and consecutive cleaning by ultrasonication in acetone and 2-propanol, they have been placed face-down over an alumina crucible containing ~5 mg MoO3 (99.998%, Alfa Aesar) and loaded into a quartz tube (diameter 20 mm) of a three-zone split-tube CVD furnace. A second alumina crucible containing
30 mg of sulfur (99.9%, Sigma-Aldrich) was loaded upstream of the substrate in a colder region of the furnace. The tube was flushed several times with ultra-high purity argon at room temperature. After heating up the furnace (ramp rate: 50°C/min, Ar flow: 10 sscm, atmospheric pressure), the temperature was kept constant at 700°C for 15 minutes to grow the MoS2 film. Subsequently, the furnace was left to cool down to ambient temperature. The synthesized MoS2 film was transferred from the sapphire substrate onto a 280 nm thick oxide coated silicon wafer for electrical characterization similarly to the method suggested in [63]. To perform the transfer, the MoS2 film was spin-coated with a thin polysiloxane film, which was lifted-off in water. Subsequently the carrier film was performed onto the target wafer and dissolved in toluene. Transistor devices were fabricated by e-beam lithography and dry etching of the MoS2 film in rectangular shape (20 × 3 µm2) to create single devices. Source and drain contacts were defined by e-beam lithography, e-beam evaporation of Ti/Au (5 nm/40 nm) and lift-off. The channel length varied from 500 nm to 3 µm, while the channel width was kept constant at 3 µm. For electrical isolation of the top-gate, a thin layer (23.5 nm) of Al2O3 was deposited on the whole wafer by atomic-layer deposition. Top-gate electrodes were fabricated by e-beam lithography (gate length from 350 nm to 2 µm, depending on the channel length), e-beam evaporation of Ti/Au (5 nm/40 nm) and lift-off in acetone.

Experimental technique
All our measurements have been performed using a Keithley-2636A in a chamber of a Lakeshore vacuum probe station (5 × 10^-6–10^-3 torr). We measured the I_d–V_g characteristics of our double-gated MoS2 FETs in both sweep directions using step voltages V_step in the range [1 V…0.01 V] and a sampling time t_step varied between 0.2 ms and 500 ms. This allowed us to vary the sweep rate S = V_step/t_step between 0.02 and 5000 V s^-1. An elementary loop of our experimental technique consists of measurements using a fixed sweep range V_gmin to V_gmax and different V_step and t_step. By loops using a fixed V_gmin = −10 V and V_gmax between −3 to 14 V in 1 V steps, we obtain a set of ΔV(t) characteristics which contain the information about the density of charged oxide traps with different time constants. The measurement frequency is given as f = 1/(N/t_step) with N = 2×(V_gmax−V_gmin)/V_step + 1) being the number of voltage step points.

Modeling
The modeling has been done using the drift–diffusion based TCAD simulator Minimos-NT [64]. First a two-dimensional model of the device cross-section was implemented using parameters taken from the literature [1, 2, 53] and validated against measured I_d–V_g and I_d–V_bg characteristics. Then the modeling of oxide traps was performed based on our previously developed four-state non-radiative multiphonon (NMP) model [25]. This model has already been successfully applied to capture various aspects of charge trapping by oxide traps in Si technologies [54, 65, 66] and back-gated FETs with MoS2 [40] and black phosphorus [33]. To simulate the hysteresis widths and offsets using the four-state NMP model, a set of microscopic defects was generated while assuming normally distributed model parameters. Finally, the model parameters were calibrated to the comprehensive experimental data set, which covers different sweep rates and sweep ranges corresponding to different time constants and energy level ranges of the traps involved.

Acknowledgments
The authors thank for the financial support through the FWF grant n° I2606-N30 and the European Union (grant agreement No. 696656 Graphene Flagship). We thank Benedikt Gottschieder and Ole Bethge for assistance with CVD and ALD setups, respectively, Alois Lugstein and Emmerich Bertagnolli for providing access to CVD and ALD systems. Y.Y.I. is a member of the Mediterranean Institute of Fundamental Physics (MIFP).

References
[1] Klots A et al 2014 Sci. Rep. 4 6608
[2] Rasmussen F and Thygensen K 2015 J. Phys. Chem. C 119 13169–83
[3] Radisavljevic B, Radenovic A, Berivio J, Giacometti V and Kis A 2011 Nat. Nanotechnol. 6 147–50
[4] Das S, Chen H, Penumatcha A and Appenzeller J 2012 Nano Lett. 13 106–5
[5] Fiori G, Szafaranc B, Iannaccone G and Neumaier D 2013 Appl. Phys. Lett. 103 233509
[6] Lee G H et al 2013 ACS Nano 7 7793–6
[7] Choi K et al 2015 Nanoscale 7 5617–23
[8] English C, Gine G, Dorgan V, Saraswat K and Pop E 2016 Nano Lett. 16 3824–30
[9] Kim T Y, Amanti M, Ahn G, Song Y, Javey A, Chung S and Lee T 2016 ACS Nano 10 2819–26
[10] Nourbakhsh A et al 2016 Nano Lett. 16 7798–806
[11] Yoon Y, Ganapathi K and Salahuddin S 2011 Nano Lett. 11 3768–73
[12] Radisavljevic B, Whitwick M and Kis A 2011 ACS Nano 5 9934–8
[13] Wang H, Lii Y, Lee Y H, Shi Y, Hsu A, Chin M, Li L J, Dubey M, Kong J and Palacios T 2012 Nano Lett. 12 4674–80
[14] Wachtler S, Polyushkin D, Bethge O and Mueller T 2016 Nat. Commun. 8 14948
[15] Krasnozhto D, Lembke D, Nyfelter C, Leblebici Y and Kis A 2014 Nano Lett. 14 5905–11
[16] Lenahan P 2003 Microelectron. Eng. 69 173–81
[17] Ghidini G 2012 Microelectron. Reliab. 52 1876–82
[18] Franco J et al 2011 Superior NBTI reliability of SiGe channel pMOSFETs: replacement Gate, finFETS, and impact of body bias IEEE Int. Electron Devices Meeting https://doi.org/10.1109/IEDM.2011.6131580
[19] Grasser T et al 2014 On the microscopic structure of hole traps in pMOSFETs IEEE Int. Electron Devices Meeting pp 21.1.1–1.4
[20] Schroder D and Babcock J 2003 J. Appl. Phys. 94 1–18
[21] Huard V, Denis M and Parthasarathy C 2006 Microelectron. Reliab. 46 1–23
[22] Huard V 2010 Two independent components modeling for negative bias temperature instability IEEE Int. Reliability Physics Symp. pp. 53–52
[23] Ang D, Yeo Z, Ho T and Ng C 2011 IEEE Trans. Device Mater. Reliab. 11 19–34
[24] Grasser T, Kaczer B, Gös W, Reisinger H, Aichinger T, Hehenberger P, Wagner P J, Franco J, Toledano-Luque M and Nelhiebel M 2011 IEEE Trans. Electron Devices 58 3652–66
[25] Grasser T 2012 Microelectron. Reliab. 52 39–70
[26] Heinrigs W, Reisinger H, Gustin W and Schlunder C 2007 Consideration of recovery effects during NBTI measurements for accurate lifetime predictions of state-of-the-art pMOSFETs IEEE Int. Reliability Physics Symp. pp. 288–92
[27] Grasser T, Rott K, Reisinger H, Walli M, Schanovsky F and Kaczer B 2014 IEEE Trans. Electron Devices 61 3586–93
[28] Grasser T, Walli M, Rzega G, Goes W, Wimmer Y, El-Sayed A, Shluger A, Reisinger H and Kaczer B 2016 The permanent component of NBTI revisited: saturation, degradation-reversal, and annealing IEEE Int. Reliability Physics Symp. pp. 5A–5
[29] Grasser T, Reisinger H, Wagner P J, Goes W, Schanovsky F and Kaczer B 2010 The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability IEEE Int. Reliability Physics Symp. pp. 16–25
[30] Cho K, Park W, Park J, Jeong H, Jang I, Kim T Y, Hong K W, Hong S and Lee T 2013 ACS Nano 7 7751–8
[31] Yang S, Park S, Jang S, Kim H and Kwon J Y 2014 Phys. Status Solidi RRL 8 714–8
[32] Iliarionov Y, Smith A, Vaziri S, Ostling M, Mueller T, Lemme M and Grasser T 2014 Appl. Phys. Lett. 105 143507
[33] Iliarionov Y, Walli M, Rzega G, Kim J S, Kim S, Dodabalapur A, Akinwande D and Grasser T 2016 ACS Nano 10 9543–9
[34] Lee Y, Kang C, Jung U, Kim J, Hwang H, Chung H J, Seo S, Choi R and Lee B 2011 Appl. Phys. Lett. 98 183508
[35] Qin H, Pan L, Yao Z, Li J, Shi Y and Wang X 2012 Appl. Phys. Lett. 100 123104
[36] Late D, Liu B, Matte H, Dravid V and Rao C 2012 ACS Nano 6 5635–41
[37] Cho A, Yang S, Park K, Namgung S, Kim H and Kwon J Y 2014 ECS Solid State Lett. 3 Q67–9
[38] Guo Y, Wei X, Shu J, Liu B, Yin J, Guan C, Han Y, Gao S and Chen Q 2015 Appl. Phys. Lett. 106 103109
[39] Park Y, Isaac H, Heo J and Yoo G 2016 Appl. Phys. Lett. 108 083102
[40] Iliarionov Y, Rzega G, Walli M, Knobloch T, Grill A, Furchi M, Mueller T and Grasser T 2016 2D Mater. 3 035004
[41] Grasser T, Wagner P J, Reisinger H, Aichinger T, Pobegein G, Nelhiebel M and Kaczer B 2011 Analytic modeling of the bias temperature instability using capture/emission time maps IEEE Int. Electron Devices Meeting pp. 27.4–1–4
[42] Park W, Lee Y, Kim J, Lee S, Kang C, Cho C, Lim S, Jung U, Hong W and Lee B 2013 Reliability characteristics of MoS2 FETs Extended Abstracts of the 2013 Int. Conf. on Solid State Devices and Materials pp. 684–5
[43] Sundaram R, Engel M, Lombardo A, Krupke R, Ferrari A, Avouris P and Steiner M 2013 Nano Lett. 13 1416–21