Combining Electrically Detected Magnetic Resonance Techniques to Study Atomic-Scale Defects Generated by Hot-Carrier Stressing in $HfO_2/SiO_2/Si$ Transistors

S. J. Moxim[1,a)], J.P. Ashton[1*], M.A. Anders[1**], and J.T. Ryan[1]

[1]Alternative Computing Group,
National Institute of Standards and Technology
100 Bureau Drive, Gaithersburg, MD 20899 USA
a) Corresponding Author Email: stephen.moxim@nist.gov

Abstract

This work explores the atomic-scale nature of defects within hafnium dioxide/silicon dioxide/silicon ($HfO_2/SiO_2/Si$) transistors generated by hot-carrier stressing. The defects are studied via electrically detected magnetic resonance (EDMR) through both spin-dependent charge pumping (SDCP) and spin-dependent tunneling (SDT). When combined, these techniques probe defects both at the Si-side interface, and within the oxide-based gate stack. The defects at the Si-side interface are found to strongly resemble P_b-like defects common in the Si/SiO_2 system. The defect within the gate stack has not been positively identified in the literature thus far; this work argues that it is a Si-dangling bond coupled to one or more hafnium atoms. The use of electrically detected magnetic resonance (EDMR) techniques indicates that the defects detected here are relevant to electronic transport, and thus device reliability. This work also highlights the impressive analytical power of combined EDMR techniques when studying complex, modern materials systems.

*Current Address: Keysight Technologies Inc., Santa Rosa, CA 95403
**Current Address: Broadcom, 9999 Hamilton Blvd., Breiningsville, PA 18031

Introduction

Until relatively recently, the question of whether hafnium-based materials would supplant conventional silicon dioxide (SiO_2) based gate dielectrics in metal-oxide-semiconductor field-effect-transistors (MOSFETs) was still very much unanswered.[1–5] At the time, predictions about the end of Moore's law and transistor scaling were abundant and the move towards so called high-k dielectrics was recognized as the biggest challenge in the history of semiconductor technology.[1–7] With the fate of the industry at stake, an enormous effort was undertaken by all sectors of the community including academia, consortia, government, and commercial manufacturers from around the world.[1–7] Now, about twenty years later, we see what a monumental success the effort was; the semiconductor industry continues to grow exponentially, and hafnium-based materials are commonplace.[6,7] In fact, high-k dielectrics are now just one of many revolutionary departures from conventional SiO_2 based planar MOSFET technologies.[8,9] Additionally, hafnium-based materials have found applications beyond MOSFET gate dielectrics and have enabled wholly new technologies, including resistive random-access memory (ReRAM) and ferroelectric random-access memory (FeRAM).[10–13] Additionally, these hafnium-based materials have found applications beyond the sense of traditional data storage, including alternative computing concepts such as neuromorphic or "brain-inspired" computing, in which device operation mimics synaptic coupling between neurons.[10–12]

As such, hafnium-based materials are relevant to several technologies, and understanding their fundamental material properties/limitations is more imperative than ever. While the materials system may seem to have been "figured out" due to its commercialization in MOSFET technologies, information about the electronic properties and physical nature of atomic-scale defects is still a topic of considerable inadequacy. This includes their roles in MOSFET reliability/failure mechanisms.[14–18] Of particular note, studies of stress-induced defects in hafnium-based transistors are limited.

This type of atomic-scale information is also important for developing technologies in which defects are purposely created and/or are intrinsic to the physical operation of the device; a prime example being the creation and motion of oxygen vacancies as the often-cited mechanism responsible for forming, writing, and reset operations of hafnium oxide (HfO_2) based ReRAM devices[10,12,15,16].

Of all the defect characterization tools available, electron spin resonance (ESR), along with a variant known as electrically detected magnetic resonance (EDMR), are among the most powerful due to their unique ability to directly determine specific chemical, physical, and energetic information about paramagnetic defect centers.[19–21]

ESR techniques have been successfully applied to HfO_2 based materials and devices in the past, and helped to provide the understanding needed to bring HfO_2 to its current complementary metal-oxide-semiconductor (CMOS) central commercial potential.[22–27] Note that in almost all cases, the existence of an intentional or unintentional SiO_2 based interfacial layer exists between the Si substrate and any deposited Hf based dielectric, significantly complicating the ability to decipher the defect/material interactions.[22,28–31] Additionally, the mere presence of Hf atoms in the vicinity of paramagnetic defect species leads to significant spin-spin interactions, which can complicate matters significantly (this is due mostly to the transition metal hafnium's d-shell electrons, as well as magnetic ^{177}Hf and ^{179}Hf nuclei)[23,25]. Electrons localized on the Hf atoms experience large spin-orbit coupling interactions. This is very much unlike conventional Si/SiO_2 systems which have well defined regions of crystalline Si and amorphous $SiO_{2,}$[32] and relatively few atoms with magnetic nuclei. The Si/HfO_2 system is much less definitive and often requires the use of vague terms to describe crystallinity, elemental composition, and physical dimensions (such as layer thickness).

Nevertheless, the ESR/EDMR literature generally indicates three classes of defects in Si/HfO_2 based systems: interface, near interface, and bulk.[22–33] Interface defects exist within the very last region of highly ordered crystalline Si substrate. At the conventional Si/SiO_2 interface, the interface defects are dangling bonds known as P_b centers. Although several variants of the P_b center are possible at (100) Si/SiO_2 interfaces, they all consist of a dangling bond orbital hosted on a central Si atom which is back-bonded to three other silicon atoms. The dangling bond orbitals point in specific crystallographic directions leading to an anisotropic g-tensor.[23,24,27,30,31] Comparable defects observed in Si/SiO2/HfO_2 systems have often been called "P_b-like" defects since their g-values are measured to be nearly identical to conventional P_b centers in magnetic resonance studies.[23,24,27]

On the other hand, bulk defects are those atomic-scale imperfections found well into the HfO_2 layer and are akin to what would be found in a large area/volume sample of HfO_2 (this terminology specifically excludes large scale defects such as grain boundaries, cracks, voids, etc.)[23,25,30,31,33]. While equally important to device operation as the interface defects, bulk defects are poorly understood and have proved to be rather difficult to study with ESR/EDMR. This is due to the additional spin interactions mentioned above, as well as the large orbital angular momentum experienced by electrons whose wavefunctions are localized on Hf atoms. Nevertheless, the dominating defects are generally thought to be HfO_2 oxygen vacancies, (Hf^{3+}) and O_2^- centers coupled to Hf ions.[23,25,30,31,33]

Finally, near-interface defects are those centers found in the poorly defined interfacial region. Literature reports indicate that this region is almost always composed of some type of SiO_2-like material and is dominated by defects similar to those found in bulk SiO_2.[22,23,26,28,29,31–33] Such defects are called E' centers; they are silicon dangling bonds with the host Si atom back-bonded to three oxygen atoms, and are well-characterized by a magnetic resonance spectrum with an isotropic g-value of g = 2.0007. While E' centers in amorphous SiO_2 generally display isotropic g-tensors, the "E' like" centers found in the near-interface region of Si/HfO_2 systems often display subtle indications of anisotropic g-tensors, possibly owing to nanoscale crystallinity in the adjacent HfO_2.[22,23,26,28,29,31–33] Additionally, the spectra of these E'-like centers are often influenced by large spin-spin interactions caused by nearby hafnium atoms, but they are not subjected to the large spin-orbit coupling interactions as is the case with the HfO_2 bulk defects discussed above.[22,23,25,28,29,31–33]

Previous ESR and EDMR studies have largely focused on native defects related to processing conditions, and the field of magnetic resonance has largely ignored those generated throughout the lifetime of real transistors. Several short EDMR studies exist concerning radiation damage[34] and the negative bias temperature instability (NBTI)[35,36] in hafnium oxide-based devices. Cochrane et al. observed only defects in the oxide layers of NBTI-stressed, hafnium oxide-based MOSFETs, and observed that the EDMR response changed based on the temperature and duration of stressing.[36] In the only other NBTI study,[35] the P_b-like interface defects were observed along with a response from hafnium oxide traps; this combination of defects was consistent with those generated by gamma irradiation of identical device structures.[34] Magnetic resonance studies of other reliability problems in hafnium-based transistors, such as hot-carrier damage and time-dependent dielectric breakdown, do not appear in the literature.

In this work, we combine two EDMR techniques in order to gain an understanding of the nature of hot-carrier damage-induced defects in HfO_2 based devices. The first is a recently developed technique, known as EDMR via spin dependent charge pumping (SDCP).[37] Due to its versatility, sensitivity, and unique strengths, SDCP is quickly becoming a useful tool to study defects in both silicon carbide[38] (SiC) and (100) Si/SiO_2 MOSFET devices.[39] SDCP is based on the MOSFET interface defect electrical measurement known as charge pumping (CP).[40] On its own, CP is a powerful tool and is easy to implement; it is traditionally used for counting the number of interface defects (also known as traps) in MOSFETs. Furthermore, variations also exist which overcome specific measurement challenges associated with advanced technologies such as nanometer-scale devices which may only contain a single defect (a single broken bond).[41–43]

It's useful to briefly describe both CP and ESR to serve as a foundation for the description of the merged SDCP measurement. In the simplest case of MOSFET CP, the source and drain of the device are grounded and a trapezoidal waveform is applied to the gate contact that acts to cyclically accumulate and invert the gate dielectric/substrate interface. During each cycle, electrons and holes are forced to undergo recombination events at deep level defects (commonly known as interface defects or interface traps). Thus, a subsequent recombination current is generated and measured through the substrate/body contact of the MOSFET, which is held at virtual ground. Any excess carriers diffuse out through the source/drain and body contacts during the rise (t_r) and fall (t_f) times of the gate pulse, while trapping of carriers happens during the high and low times of the voltage pulse. The high and low times correspond to voltage levels V_{High} and V_{Low}, respectively. The CP current, I_{CP}, is at a maximum when $V_{High} > V_{th}$ and $V_{Low} < V_{FB}$, where V_{th} and V_{FB} are the threshold and flat band voltages, respectively. A diagram describing the CP measurement is provided in Figure 1. Here, the specific components of the gate waveform pulse (voltage and time) are mapped to the corresponding energy diagrams of the MOSFET gate dielectric/substrate interface. The current I_{CP} thus depends upon the density of interface traps D_{it} (in units of cm^{-2} eV^{-1}) and the frequency at which the traps are filled and emptied, or the frequency of the accumulation/inversion cycles, known as the CP frequency f_{CP}. The expression for I_{CP} is given by:

$$I_{CP} = qAf_{CP}D_{it}\Delta E. \quad (1)$$

Here, q is electronic charge, A is the gate area, and ΔE is the range of energy within the bandgap explored, which is given by:

$$\Delta E = 2kTln\left(\frac{V_{High} - V_{Low}}{v_{\bar{th}}\bar{\sigma}n_i(V_{th} - V_{FB})\sqrt{t_r t_f}}\right). \quad (2)$$

Here, k is Boltzmann's constant, T is temperature, $\bar{v_{th}}$ is the geometric mean of the electron and hole thermal velocities, $\bar{\sigma}$ is the geometric mean of the electron and hole capture cross sections, and n_i is the intrinsic carrier concentration. Thus, I_{CP} is linear with respect to f_{CP} with a slope dictated by D_{it}. A representative figure showing I_{CP} versus f_{CP} of a Si/SiO_2/HfO_2 MOSFET used in this study is shown in Figure 2. It should be noted that at room temperature, $\Delta E \approx 50$ % of the Si bandgap.

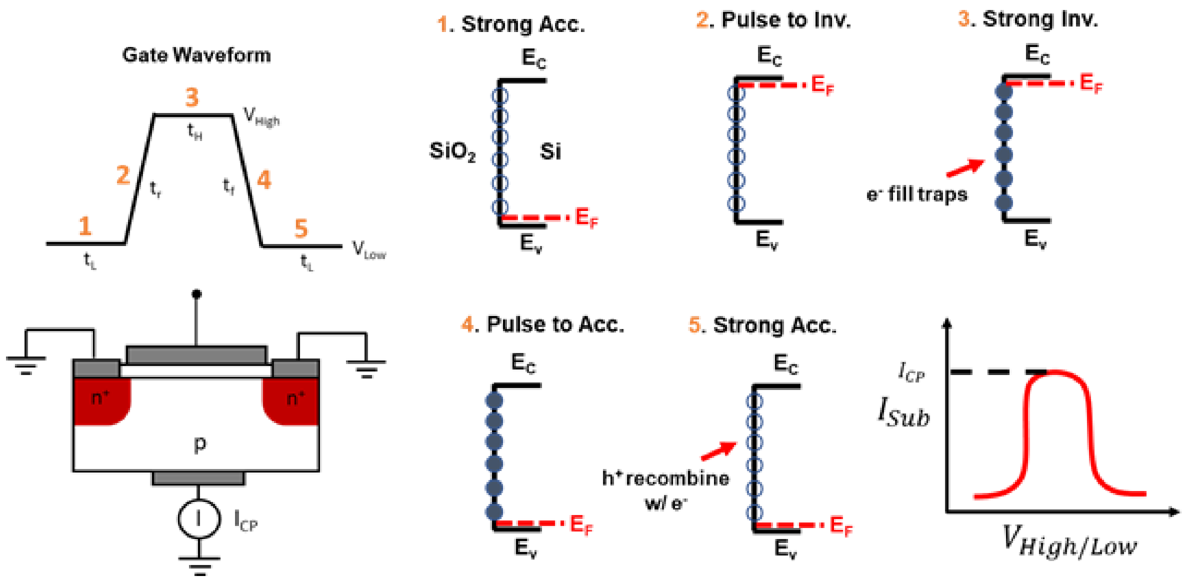

Fig. 1. Step-by-step description of the charge pumping cycle. t_L, t_r, t_H, t_f are the low, rise, high, and fall time of the waveform and V_{High}/V_{low} are the high and low voltages.

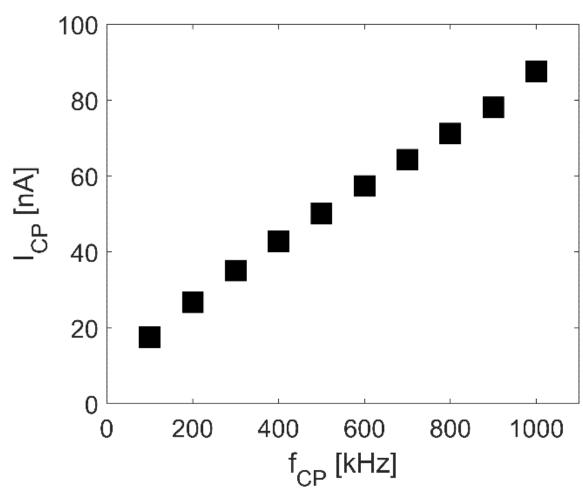

Fig. 2. Experimental I_{CP} vs. f_{CP} characteristics for a $HfO_2/SiO_2/Si$ MOSFET with D_{it} = 1.5 x 10^{12} cm^{-2} eV^{-1}.

We will now give a brief introduction to ESR principles, but more detailed descriptions are available from a variety of textbooks.[19–21,44] In ESR, a material with paramagnetic point defects is placed within an electromagnet that provides a polarizing magnetic field B_0 that acts to split the energy levels of unpaired electrons, with the levels corresponding to the two allowed electron spin quantum numbers, m_s = +1/2 and m_s = -1/2. Simultaneously, the sample is held in a microwave cavity and is exposed to photons of energy *hv*, where *h* Planck's constant and *v* is the frequency. When the incoming photon energy matches the energy difference between the two electron spin states, resonance occurs and transitions can be made between the two spin states (electrons can flip their spin quantum number). Assuming the electrons are completely unperturbed by their local environment, the resonance condition is given by:

$$h\nu = g_e \mu_B B_0 \qquad (3)$$

where, $g_e \approx 2.0023$ is the Landé g factor and μ_B is the Bohr magneton. This experiment is typically (but not necessarily) performed by holding the photon energy constant and recording absorbed energy as a function of magnetic field. Perturbations to (3) shed light on the physical and chemical nature of the defects involved. The two most important perturbations are spin-orbit coupling, which involves the orbital angular momentum of the electron with respect to the positively charged nucleus, and electron-nuclear hyperfine interactions which arise from nearby magnetic nuclei. These two perturbations determine the structure of an ESR spectrum, which can be utilized to identify defects and obtain information about the surrounding lattice. The two mechanisms can be described via a spin Hamiltonian of the form

$$\mathcal{H} = \mu_B \widehat{\boldsymbol{B}} \cdot \boldsymbol{g} \cdot \widehat{\boldsymbol{S}} + \sum_i \widehat{\boldsymbol{I}}_i \cdot \boldsymbol{A}_i \cdot \widehat{\boldsymbol{S}}. \qquad (4)$$

Here, $\widehat{\boldsymbol{B}} = B\hat{k}$ is the applied magnetic field vector, $\boldsymbol{g}$ and $\boldsymbol{A}$ are 2nd rank tensors that describe the spin-orbit coupling and electron-nuclear hyperfine interactions, respectively, $\widehat{\boldsymbol{S}}$ is the electron spin operator, and $\widehat{\boldsymbol{I}}_i$ is the nuclear spin operator for the ith nucleus. The components of $\boldsymbol{g}$ yield deviations from the g_e

of (3) caused by spin-orbit coupling, which depends on the atomic number of the nucleus and the orbital angular momentum of the electron. The components of $\boldsymbol{A}$ yield information regarding nearby magnetic nuclei in the system and the strength of their magnetic interactions with the defect electron spins. Other spin-spin interactions, such as dipolar and exchange interactions, are also observable via ESR and more advanced resonance methods[45].

For studies of micro- and nano-scale technology, EDMR is implemented rather than ESR. Unlike ESR, where the energy absorbed by the sample is monitored as a function of magnetic field, a device *current* is monitored as a function of magnetic field in an EDMR measurement. Several spin-dependent transport phenomena can be invoked as the source of the EDMR-induced current change; in this work we utilize both spin-dependent recombination (SDR) via SDCP, and spin-dependent tunneling (SDT). EDMR is typically about 10^7 times more sensitive than ESR[46] (ESR has an absolute sensitivity of about 10^{11} spins per mT linewidth at X-band frequencies)[44], yet it provides much of the same analytical power. The absolute sensitivity limit of EDMR in fully-processed MOSFETs is difficult to define since it depends on both the defect density and the kinetics of electronic transport, which can vary with biasing. EDMR also has the advantage of being selective to only electrically active defect centers which affect device performance.

One of the most important mechanisms for spin-dependent transport is spin-dependent recombination. It is described in the seminal work by Kaplan, Solomon, and Mott (KSM)[47] and has since been refined.[48] Consider the case when a device containing paramagnetic deep level defects is subjected to B_0 and photons of energy hv. The magnetic field aligns the spin of the defect electron and a nearby conduction level electron (In actuality, the "conduction level" is an intermediate level close to one of the bands, such as a shallow-state donor in Si)[39]. If both the conduction electron and defect electron have the same spin quantum number m_s (a triplet spin pair), the transition of the conduction electron into the defect site is forbidden by the Pauli exclusion principle. However, under ESR resonance conditions, the defect electron can flip spin sates, converting the triplet spin pair to a singlet spin pair, and enabling the previously forbidden capture of a conduction electron. The captured electron is now available for electron-hole recombination. The process produces a measurable change in recombination current which, when plotted vs. magnetic field, is nearly identical to a classical ESR response involving the same defect centers.

SDCP is one method of generating SDR current at MOSFET interfaces; it involves forcing SDR events to occur during CP gate voltage cycles. Thus, SDR can be measured via EDMR through the substrate contact. This SDCP method probes defects over a large percentage of the bandgap, even at room temperature, as determined by equation (2). This is a significant advantage over other MOSFET-based SDR techniques, in which the range of the bandgap explored is usually only about $\frac{1}{2}qV_{Bias}$ centered around the middle of the bandgap, where V_{Bias} is the diode forward bias voltage.[49,50]

Experimental

The MOSFETs used in this study consist of a (100) Si substrate with a 2 nm purposely grown SiO_2 interfacial layer and a 3 nm HfO_2 dielectric layer capped with a TiN gate contact. The channel length and channel widths are 1 µm and 100 µm, respectively. The samples were hot carrier stressed (consisting of drain or source voltage of 3.5 V and gate voltage of 1.3 V for 2000 seconds). An analysis of I_{CP} as a function of f_{CP} for a pre-stress and post-stress device confirms a linear relationship consistent with (1), and yields D_{it} = 1.5 x 10^{12} $cm^{-2}eV^{-1}$. Pre-stress, the devices have an order of magnitude less D_{it}. Measurements were carried out utilizing an arbitrary waveform generator connected to the gate terminal of the device, the source and drain held at ground and the substrate current measured while being held at virtual ground. The substrate was connected to a transimpedance amplifier, the output of which was fed into a data acquisition system. Since the expected EDMR-induced changes in device current are often on the order of pA, we utilize virtual lock-in amplification with magnetic field modulation. EDMR measurements in

MOSFETs are subject to white noise, flicker noise, and shot noise. The use of lock-in detection reduces both white and flicker noise contributions. The lock-in amplifier output is approximately the derivative of the EDMR-induced change in device current. We define the EDMR intensity as the maximum change in current between on-resonance and off-resonance fields. The EDMR intensity is read after numerically integrating the output of the lock-in amplifier with respect to magnetic field and dividing the resulting spectrum by the modulation amplitude. B_0 is provided by a 4-inch electromagnet with power supply and Gaussmeter/Hall probe with magnetic field control achieved utilizing proportional-integral-derivative feedback. The sample is held within a microwave cavity with optimal dimensions for 9 GHz to 10 GHz (X-band) standing waves. The microwaves are provided by a microwave source with maximum output of 33 dBm. The microwaves are channeled through a custom-built microwave bridge that is connected to the cavity plumbing allowing for cavity tuning and conventional ESR detection. Strong pitch was used as a standard to calibrate the magnetic field for accurate g component measurements. The uncertainty in g is $\pm$0.0003 based on the combined uncertainty in magnetic field and microwave frequency. All measurements were performed at room temperature. Signal averaging times of several hours to several days were used to obtain the EDMR results in this work. At these averaging times, the typical detection limit for the EDMR intensity was 4 pA.

Results and Discussion

No EDMR responses were resolved above the detection limit in pre-stress SDCP measurements. After hot-carrier stressing, two differing responses appear in SDCP measurements, depending upon the values of V_{High} and V_{Low}. Since these responses are not present in fresh transistors, we conclude that they are a result of the hot-carrier stressing. Figure 3 compares three SDCP spectra taken at three different voltage ranges corresponding to (a) V_{High} = 0.75 V, V_{Low} = -2.1 V, (b) V_{High} = -1.9 V, V_{Low} = -2.1 V, and (c) V_{High} = -0.1 V, V_{Low} = -0.5 V. In all three cases, f_{CP} was 2 MHz.

The spectra in Figures 3(a) and 3(b) exhibit nearly identical responses. This is surprising, since one would expect that V_{High} = -1.9 V and V_{Low} = -2.1 V *would not* produce a SDCP response since both of these biases are outside of $V_{th} < V < V_{fb}$, while one *would* expect a SDCP response from the voltages used for Figure 4(a) since they span the entirety of V_{fb} to V_{th}. Interestingly, when V_{High} and V_{Low} are set within the range of V_{th} and V_{fb}, respectively, a different SDCP response is observed, as shown in shown in Figure 3(c).

The observed $g \approx 2.0044$ in Figure 3(c) is slightly off from that typically observed for P_b centers at SiO_2/(100)Si MOSFET interfaces (2.0059 for P_{b0} and 2.0031 for P_{b1}).[32,51,52] Pribicko *et al.* report gated diode measurement SDR in similar devices and observe a signal very similar to this.[27] They argue that their spectrum could be P_b-like defects that are a superposition of P_{b0} and P_{b1} within a highly disordered SiO_2/Si interface region. The near perfect 50 % superposition of the $g \approx 2.0044$ reported here also agrees with this interpretation as one would expect to have a near uniform distribution of both P_{b0} and P_{b1} in an interface region of high disorder. Similar scenarios involving combinations of these two defects have been noted in other EDMR studies as well.[53–55]

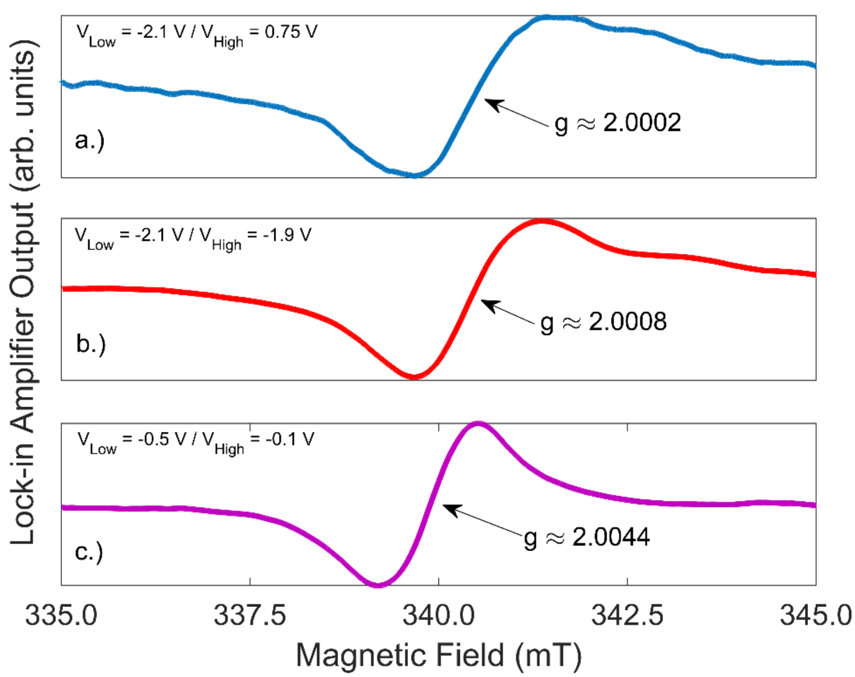

Fig. 3. (a) SDCP spectrum taken with V_{High} = 0.75 V and V_{Low} = -2.1V. (b) SDCP spectrum taken with V_{High} = -1.9 V and V_{Low} = -2.1 V. (c) SDCP spectrum taken with V_{High} = -0.1 V and V_{Low} = -0.5 V. Note that the spectra in (a) and (b) are nearly identical aside from noise and the spectrum in (c) is very different than that of (a) and (b), implying two different voltage parameter-dependent defects.

The $g \approx 2.0002/2.0008$ spectrum is likely not related to an interface defect, but rather exists within the gate stack. This spectrum is notably broader than the $g \approx 2.0044$ spectrum, implying that the spectrum may be broadened by either hyperfine interactions, spin-spin interactions between the defect electron and other unpaired electrons, or saturation. If the defects are within the dielectric stack, it is possible that the defects may be coupled to comparatively large Hf atoms with magnetic nuclei and d-shell electrons. Thus, large, additional spin-spin interactions likely broaden the spectrum. This idea of broadening has been reported in works by Ryan *et al.*[22,26] and Cochrane *et al.*[36] in magnetic resonance-based measurements in similar devices. They also observe the signal $g = 2.0002/2.0008$ and indicate that it is likely an E' variant (likely a E' center coupled to a Hf atom).[22,26] In a paper by Wang *et al*., it was proposed that oxygen deficient HfO_2 will pull oxygen from the SiO_2 interfacial layer, stimulating the generation of oxygen vacancy centers in the SiO_2.[56] If the defect electrons were localized on Hf atoms, we would expect significantly larger spin-orbit interactions, and a g-value which deviates much farther from the free electron g (g_e = 2.0023…) than that of the E' center. Thus, we conclude that the observed responses in Figures 3(a) and 3(b) are an E'-variant in which the defect electron is interacting with (but is not localized on) one or more Hf atoms. Since it would be impossible for V_{High} = -1.9 V to create enough electrons at the interface to facilitate charge pumping, it is possible that the response observed here is caused by an alternate EDMR mechanism.

To better understand the nature of these two different defect spectra, variable V_{Low}, constant V_{High} = 0.75 V SDCP experiments were conducted. The relative EDMR intensity as a function of V_{Low} is shown in Figure 4. The dotted line indicates the V_{Low} corresponding to an approximate point at which the spectrum changes shape and g value (no changes in spectrum shape and center crossing are observed at other voltages). It appears that the spectrum corresponding to $g \approx 2.0002/2.0008$ increases when V_{Low} < -1.0 V (voltages beyond -2.1 V were not explored as the devices underwent additional stressing over the length of the measurement time past this bias). The spectrum corresponding to $g \approx 2.0044$ peaks in amplitude at V_{Low} = -0.5 V. Similar peaks in EDMR intensity from have been observed via dc SDR EDMR measurements in the past[57].

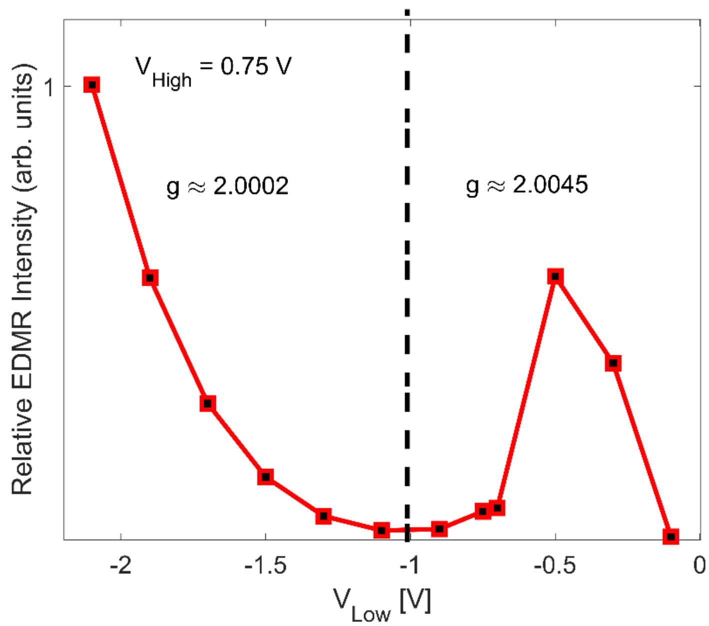

Fig. 4. Relative EDMR intensity vs. V_{Low} of the SDCP results showing the presence of two different defects based upon two different g values. The parameter V_{Low} is varied and two defects become prominent at two different ranges of V_{Low}.

Figures 5 and 6 show variable V_{High}, constant V_{Low} = -0.5 V and V_{Low} = -2.1 V, respectively. Note that in both figures, a peak is also observed but corresponds to V_{High} = -0.1 V. The presence of this peak is consistent with Figure 4, and we thus conclude its validity (it is not an anomaly). These somewhat strange voltage dependencies obviously cannot be attributed to spin-dependent charge pumping alone, and must be at least partially related to another EDMR mechanism.

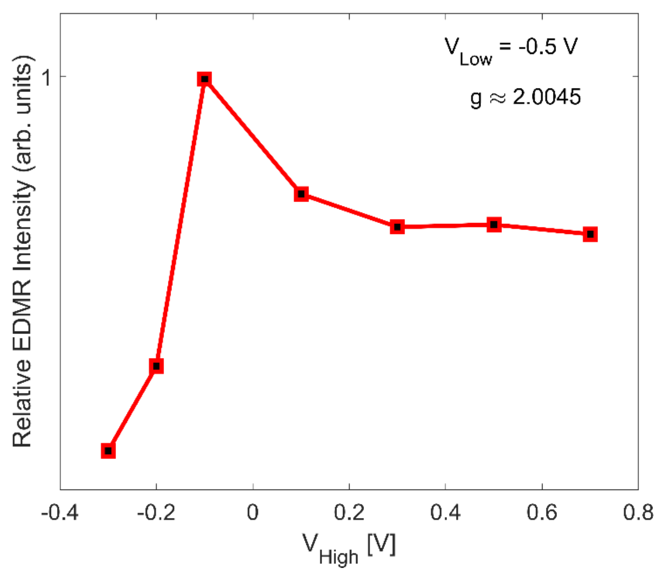

Fig. 5. Relative EDMR intensity vs. V_{High} SDCP results with V_{Low} = -0.5 V. Here, all the spectra obtained are identical aside from relative intensity. Note the peak at -0.1 V.

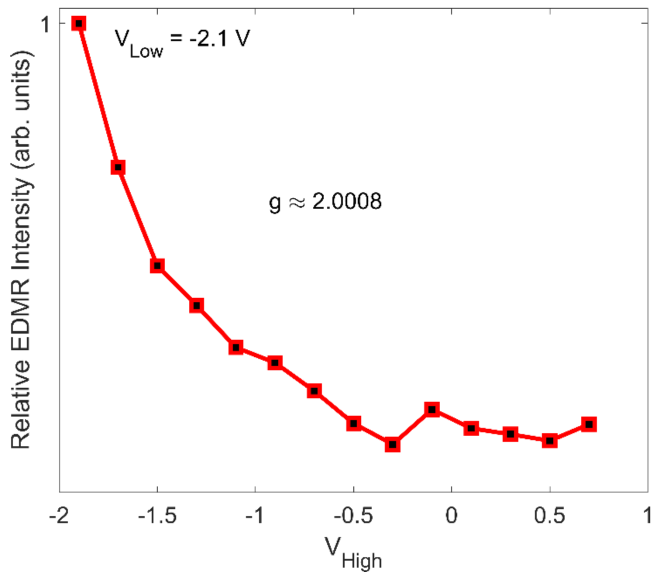

Fig. 6. Relative EDMR intensity vs. V_{High} SDCP results with V_{Low} = -2.1 V. Here, all the spectra obtained are identical aside from relative intensity.

EDMR can also be detected via spin-dependent tunneling (SDT)[58]. This method is used to probe defects that act as tunneling centers in dielectrics. The principle behind SDT is similar to that of the KSM picture of SDR; in order for an electron to use a paramagnetic defect in a dielectric layer as a tunneling center, the tunneling electron and defect electron must be a singlet pair. As discussed above, magnetic resonance can facilitate the conversion of triplet spin pairs, for which a tunneling event would be forbidden, to singlet spin pairs. In the case of SDT, this means that at the resonance condition, an increase in tunneling current is observed. A plot of tunneling current vs magnetic field yields a spectrum nearly identical to that of the ESR spectra of the defect(s) involved. The SDT process is illustrated in Figure 7.

To determine if SDT is responsible for the E'-variant response, SDT measurements were made through the gates of large area, hot-carrier-stressed transistors on the same die. For these SDT

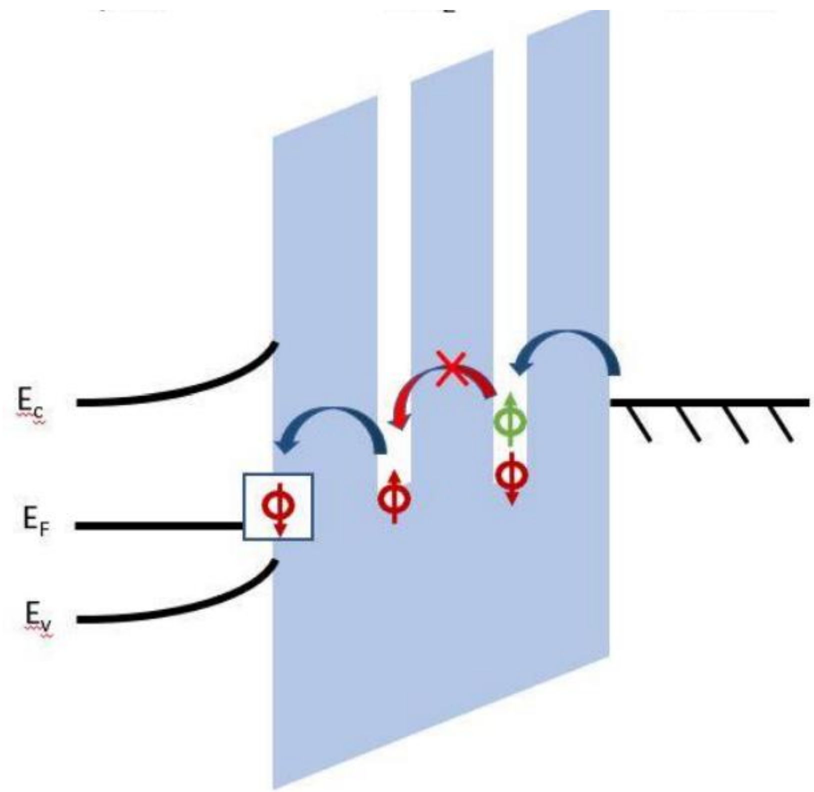

Fig. 7. Simplified illustration of SDT in a metal-oxide-semiconductor structure. Under bias, defects in the dielectric (and sometimes defects at the oxide interface) act as tunneling centers. If the defects are paramagnetic, some tunneling events are forbidden by the Pauli exclusion principle.

measurements, the source and drain were floated, ensuring that contribution from SDCP was impossible. In Figure 8(a), we plot the spin-dependent tunneling (SDT) response from a 100 μm $\times$ 100 μm transistor of the same composition and after the same stressing conditions as those used in the SDCP measurements. Here, the gate voltage was set to -2.1 V with the substrate current measured. The spectrum in Figure 8(a) closely resembles that of Figures 3(a) and 3(b). In Figure 8(b), we plot the amplitude of the SDT spectrum as a function of the gate voltage. Note that the left sides of Figures 4 and 8(b) follow very similar trends in amplitude, and also exhibit the same g-value (within experimental error). Thus, we confirm that this defect response is purely due to SDT in both cases. The $g \approx 2.0044$ signal was not detected in the SDT measurements, indicating that this response must be purely due to a traditional

SDCP process. In work by Mitrovic *et al.*, defect levels near 1.8 eV below the conduction band of HfO_2 were shown to act as effecting trap-assisted tunneling traps.[59] It is possible that holes tunneling from the valence band at $V_G \approx 2$ V into these levels are responsible for the defect spectrum observed for the $g \approx 2.0002/2.0008$ defect.

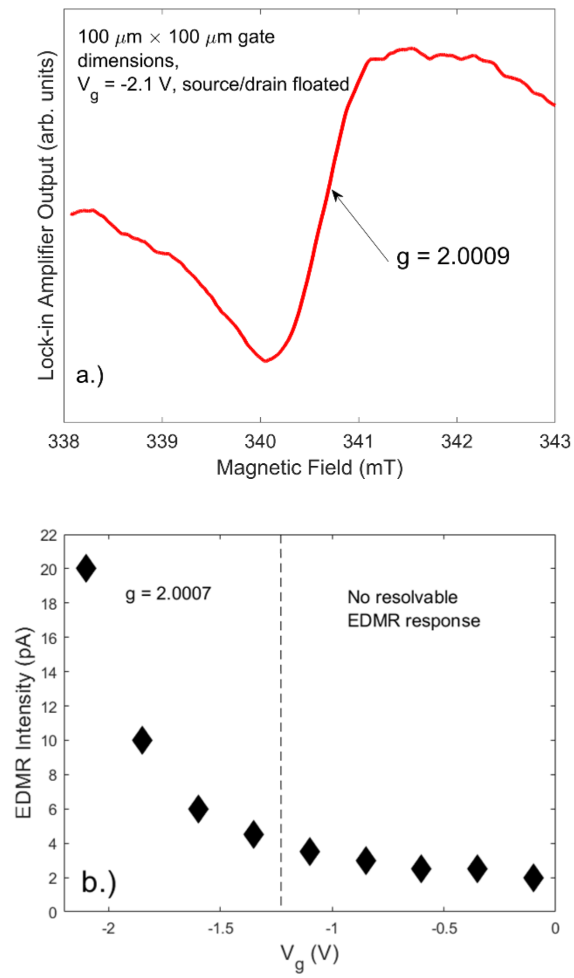

Fig. 8. (a) SDT response measured through the substrate contact with V_g = -2.1 V and the source and drain floated. Note the similarity to Figures 3(a) and 3(b). (b) Integrated SDT amplitude as a function of V_g.

Conclusions

SDCP and SDT have been utilized to study $HfO_2/SiO_2/Si$-based MOSFETs after hot-carrier stressing. SDCP reveals that defects are generated by hot-carrier stressing at the Si/SiO_2 interface and in the bulk oxide when certain CP parameters are used. We have confirmed that the bulk defects are observed via SDT during the SDCP process by making DC SDT measurements on comparable structures. We were able to make a tentative physical and chemical identification of both defects, one being ascribed to a P_b-like interface defect and the other a E'-like defect within the dielectric stack, likely coupled to one or more Hf atoms. The P_b-like defect lies within the energy range explored by our charge pumping measurements (the middle 50% of the Si bandgap). Although more work is needed to get a truly complete atomic scale picture of both defects, it is clear that these two distinctly different defects are generated via hot-carrier stressing and contribute to electronic transport in Hf-based transistors. Obviously, the tunneling currents may not be observed in systems with much thicker dielectric stack regions, and only near-interface SDCP defects could be observed in this case; however, in modern systems utilizing $HfO_2/SiO_2/Si$, field dielectrics are often thin enough to allow for trap-assistant tunneling currents. This work also highlights the power of combining EDMR methods when studying complex systems. Such a strategy allows one to determine not only the identities of the defects, but also where they exist physically in the device (i. e. in the channel region vs. within a dielectric layer).

Future work should involve SDCP measurements at variable temperatures, precise orientation studies, and multi-field/frequency SDCP. Further comparisons to other damage mechanisms, such as time-dependent dielectric breakdown, should also be the subject of future research. The application of EDMR techniques, especially SDT, to Hf-based ReRAM (or other Hf-based memory devices) is also of interest since oxygen vacancies are vital to their operation.

Conflict of Interest:

The authors have no conflict to disclose

Data Availability:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] K.J. Hubbard, and D.G. Schlom, “Thermodynamic stability of binary oxides in contact with silicon,” J Mater Res **11**(11), 2757–2776 (1996).

[2] G.D. Wilk, R.M. Wallace, and J.M. Anthony, “High-κ gate dielectrics: Current status and materials properties considerations,” J Appl Phys **89**(10), 5243–5275 (2001).

[3] G.D. Wilk, M. Verghese, P. Chen, and J.W. Maes, “(Invited) Current Status of High-k and Metal Gates in CMOS,” ECS Trans **50**(4), 207–210 (2013).

[4] S. Stemmer, and D.G. Schlom, in *Nano and Giga Challenges in Microelectronics: Elsevier Science B.V* (2003), pp. 129–150.

[5] J. Robertson, “High dielectric constant oxides,” EPJ Applied Physics **28**(3), 265–291 (2004).

[6] S. Kol, and A.Y. Oral, “Hf-Based High-k Dielectrics: A Review,” Acta Phys Pol A **136**(6), 873–881 (2019).

[7] J. Robertson, and R.M. Wallace, “High-K materials and metal gates for CMOS applications,” Materials Science and Engineering R: Reports **88**, 1–41 (2015).

[8] H.H. Radamson, X. He, Q. Zhang, J. Liu, H. Cui, J. Xiang, Z. Kong, W. Xiong, J. Li, J. Gao, H. Yang, S. Gu, X. Zhao, Y. Du, J. Yu, and G. Wang, “Miniaturization of CMOS,” Micromachines (Basel) **10**(5), (2019).

[9] H.H. Radamson, H. Zhu, Z. Wu, X. He, H. Lin, J. Liu, J. Xiang, Z. Kong, W. Xiong, J. Li, H. Cui, J. Gao, H. Yang, Y. Du, B. Xu, B. Li, X. Zhao, J. Yu, Y. Dong, and G. Wang, “State of the art and future perspectives in advanced CMOS technology,” Nanomaterials **10**(8), 1–86 (2020).

[10] N.K. Upadhyay, H. Jiang, Z. Wang, S. Asapu, Q. Xia, and J. Joshua Yang, “Emerging Memory Devices for Neuromorphic Computing,” Adv Mater Technol **4**(4), (2019).

[11] M.H. Park, Y.H. Lee, T. Mikolajick, U. Schroeder, and C.S. Hwang, “Review and perspective on ferroelectric HfO_2-based thin films for memory applications,” MRS Commun **8**(3), 795–808 (2018).

[12] Y. Li, Z. Wang, R. Midya, Q. Xia, and J. Joshua Yang, “Review of memristor devices in neuromorphic computing: Materials sciences and device challenges,” J Phys D Appl Phys **51**(50), (2018).

[13] W. Banerjee, “Challenges and applications of emerging nonvolatile memory devices,” Electronics (Switzerland) **9**(6), 1–24 (2020).

[14] G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. Parthasarathy, E. Vincent, and G. Ghibaudo, "Review on high-k dielectrics reliability issues," IEEE Transactions on Device and Materials Reliability **5**(1), 5–19 (2005).

[15] A. Padovani, L. Larcher, O. Pirrotta, L. Vandelli, and G. Bersuker, "Microscopic modeling of HfO_x RRAM operations: From forming to switching," IEEE Trans Electron Devices **62**(6), 1998–2006 (2015).

[16] S. Dirkmann, J. Kaiser, C. Wenger, and T. Mussenbrock, "Filament Growth and Resistive Switching in Hafnium Oxide Memristive Devices," ACS Appl Mater Interfaces **10**(17), 14857–14868 (2018).

[17] S. Zafar, A. Kumar, E. Gusev, and E. Cartier, "Threshold voltage instabilities in high-κ gate dielectric stacks," IEEE Transactions on Device and Materials Reliability **5**(1), 45–64 (2005).

[18] T.P. Ma, H.M. Bu, X.W. Wang, L.Y. Song, W. He, M. Wang, H.H. Tseng, and P.J. Tobin, "Special reliability features for Hf-based high-k gate dielectrics," IEEE Transactions on Device and Materials Reliability **5**(1), 36–44 (2005).

[19] J.A. Weil, J.R. Bolton, and J.E. Wertz, *Electron Paramagnetic Resonance* (Wiley, New York, NY, 1994).

[20] Poole C. P., *Electron Spin Resonance: A Comprehensive Treatise on Experimental Techniques*, 2nd ed. (John Wiley and Sons, Inc. , New York, NY, 1983).

[21] W. Gordy, *Theory and Applications of Electron Spin Resonance* (John Wiley and Sons, Inc., New York, NY, 1980).

[22] J.T. Ryan, P.M. Lenahan, J. Robertson, and G. Bersuker, "Direct observation of electrically active interfacial layer defects which may cause threshold voltage instabilities in HfO_2 based metal-oxide-silicon field-effect transistors," Appl Phys Lett **92**(12), (2008).

[23] P.M. Lenahan, and J.F. Conley, "Magnetic resonance studies of trapping centers in high-κ dielectric films on silicon," IEEE Transactions on Device and Materials Reliability **5**(1), 90–101 (2005).

[24] A.Y. Kang, P.M. Lenahan, J.F. Conley, and R. Solanki, "Electron spin resonance study of interface defects in atomic layer deposited hafnium oxide on Si," Appl Phys Lett **81**(6), 1128–1130 (2002).

[25] A.Y. Kang, P.M. Lenahan, and J.F. Conley, "Electron spin resonance observation of trapped electron centers in atomic-layer-deposited hafnium oxide on Si," Appl Phys Lett **83**(16), 3407–3409 (2003).

[26] J.T. Ryan, P.M. Lenahan, G. Bersuker, and P. Lysaght, "Electron spin resonance observations of oxygen deficient silicon atoms in the interfacial layer of hafnium oxide based metal-oxide-silicon structures," Appl Phys Lett **90**(17), (2007).

[27] T.G. Pribicko, J.P. Campbell, P.M. Lenahan, W. Tsai, and A. Kerber, "Interface defects in Si/HfO_2 -based metal-oxide-semiconductor field-effect transistors," Appl Phys Lett **86**(17), 1–3 (2005).

[28] G. Bersuker, J. Barnett, N. Moumen, B. Foran, C.D. Young, P. Lysaght, J. Peterson, B.H. Lee, P.M. Zeitzoff, and H.R. Huff, in *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers* (2004), pp. 7899–7902.

[29] G. Bersuker, C.S. Park, J. Barnett, P.S. Lysaght, P.D. Kirsch, C.D. Young, R. Choi, B.H. Lee, B. Foran, K. van Benthem, S.J. Pennycook, P.M. Lenahan, and J.T. Ryan, "The effect of interfacial layer properties on the performance of Hf-based gate stack devices," J Appl Phys **100**(9), (2006).

[30] K. Xiong, J. Robertson, and S.J. Clark, in *Phys Status Solidi B Basic Res* (2006), pp. 2071–2080.

[31] J. Robertson, "Interfaces and defects of high-K oxides on silicon," Solid State Electron **49**(3), 283–293 (2005).

[32] P.M. Lenahan, "What can electron paramagnetic resonance tell us about the Si/SiO_2 system?," Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures **16**(4), 2134 (1998).

[33] D. Heh, C.D. Young, G.A. Brown, P.Y. Hung, A. Diebold, G. Bersuker, E.M. Vogel, and J.B. Bernstein, "Spatial distributions of trapping centers in HfO_2/SiO_2 gate stacks," Appl Phys Lett **88**(15), (2006).

[34] K.J. Myers, R.J. Waskiewicz, P.M. Lenahan, and C.D. Young, "A Multifield and Frequency Electrically Detected Magnetic Resonance Study of Atomic-Scale Defects in Gamma Irradiated Modern MOS Integrated Circuitry," IEEE Trans Nucl Sci **66**(1), 405–412 (2019).

[35] K.J. Myers, P.M. Lenahan, B.C. Bittle, and I. Meric, "Exploring Negative Bias Temperature Instability in Tri-Gate MOSFETs Through Electrically Detected Magnetic Resonance," Proceedings of the International Integrated Reliability Workshop (IIRW), (2019).

[36] C.J. Cochrane, P.M. Lenahan, J.P. Campbell, G. Bersuker, and A. Neugroschel, "Observation of negative bias stressing interface trapping centers in metal gate hafnium oxide field effect transistors using spin dependent recombination," Appl Phys Lett **90**(12), (2007).

[37] B.C. Bittel, P.M. Lenahan, J.T. Ryan, J. Fronheiser, and A.J. Lelis, "Spin dependent charge pumping in SiC metal-oxide-semiconductor field-effect-transistors," Appl Phys Lett **99**(8), (2011).

[38] M.A. Anders, P.M. Lenahan, and A.J. Lelis, "Multi-resonance frequency spin dependent charge pumping and spin dependent recombination - Applied to the 4H-SiC/SiO_2 interface," J Appl Phys **122**(23), (2017).

[39] M. Hori, and Y. Ono, "Charge Pumping under Spin Resonance in Si (100) Metal-Oxide-Semiconductor Transistors," Phys Rev Appl **11**(6), (2019).

[40] J.S. Brugler, and P.G.A. Jespers, "Charge pumping in MOS devices," IEEE Trans Electron Devices **16**(3), 297–302 (1969).

[41] K.P. Cheung, C. Wang, and J.P. Campbell, "Nanoscale MOSFET as a potential room-temperature quantum current source," Micromachines (Basel) **11**(4), (2020).

[42] T. Tsuchiya, and Y. Ono, "Charge pumping current from single Si/SiO_2 interface traps: Direct observation of P_b centers and fundamental trap-counting by the charge pumping method," Jpn J Appl Phys **54**(4), (2015).

[43] J.P. Ashton, M.A. Anders, and J.T. Ryan, "Detection of individual spin species via frequency-modulated charge pumping," Appl Phys Lett **120**(5), (2022).

[44] G.R. Eaton, S.S. Eaton, D.P. Barr, and R.T. Weber, *Quantitative EPR* (Springer Vienna, Vienna, 2010).

[45] C.J. Cochrane, and P.M. Lenahan, "Spin counting in electrically detected magnetic resonance via low-field defect state mixing," Appl Phys Lett **104**(9), (2014).

[46] G. Kawachi, C.F.O. Graeff, M.S. Brandt, and M. Stutzmann, "Carrier transport in amorphous silicon-based thin-film transistors studied by spin-dependent transport," Phys Rev B **54**(11), (1996).

[47] D. Kaplan, I. Solomon, and N.F. Mott, "Explanation of the Large Spin-Dependent Recombination Effect in Semiconductors," J Phys (Paris) Lett **39**(4), (1978).

[48] C. Boehme, and K. Lips, "Theory of time-domain measurement of spin-dependent recombination with pulsed electrically detected magnetic resonance," Phys Rev B Condensed Matter Phys **68**(24), (2003).

[49] J. Fitzgerald, and A.S. Grove, "Surface Recombination in Semiconductors," Surf Sci **9**, 347–369 (1968).

[50] J.P. Ashton, S.J. Moxim, A.D. Purcell, P.M. Lenahan, and J.T. Ryan, "A quantitative model for the bipolar amplification effect: A new method to determine semiconductor/oxide interface state densities," J Appl Phys **130**(13), (2021).

[51] Y.Y. Kim, and P.M. Lenahan, "Electron-spin-resonance study of radiation-induced paramagnetic defects in oxides grown on (100) silicon substrates," J Appl Phys **64**(7), 3551–3557 (1988).

[52] E.H. Poindexter, P.J. Caplan, B.E. Deal, and R.R. Razouk, "Interface states and electron spin resonance centers in thermally oxidized (111) and (100) silicon wafers," J Appl Phys **52**(2), 879–884 (1981).

[53] S.J. Moxim, F. v. Sharov, D.R. Hughart, G.S. Haase, C.G. McKay, and P.M. Lenahan, "Atomic-scale defects generated in the early/intermediate stages of dielectric breakdown in Si/SiO_2 transistors," Appl Phys Lett **120**(6), (2022).

[54] E.B. Frantz, D.J. Michalak, N.J. Harmon, E.M. Henry, S.J. Moxim, M.E. Flatté, S.W. King, J.S. Clarke, and P.M. Lenahan, "Electrically detected magnetic resonance and near-zero field magnetoresistance in $^{28}Si/^{28}SiO_2$," J Appl Phys **130**(6), (2021).

[55] J.P. Campbell, P.M. Lenahan, C.J. Cochrane, A.T. Krishnan, and S. Krishnan, "Atomic-scale defects involved in the negative-bias temperature instability," IEEE Transactions on Device and Materials Reliability **7**(4), 540–557 (2007).

[56] S.J. Wang, P.C. Lim, A.C.H. Huan, C.L. Lu, J.W. Chai, S.Y. Chow, J.S. Pan, Q. Li, and C.K. Ong, "Reaction of SiO_2 with hafnium oxide in low oxygen pressure," Appl Phys Lett **82**(13), 2047–2049 (2003).

[57] C.J. Cochrane, and P.M. Lenahan, "Detection of interfacial Pb centers in Si/SiO_2 metal-oxide-semiconducting field-effect transistors via zero-field spin dependent recombination with observation of precursor pair spin-spin interactions," Appl Phys Lett **103**(5), (2013).

[58] M.A. Anders, P.M. Lenahan, C.J. Cochrane, and J. van Tol, "Physical nature of electrically detected magnetic resonance through spin dependent trap assisted tunneling in insulators," J Appl Phys **124**(21), (2018).

[59] I.Z. Mitrovic, Y. Lu, O. Buiu, and S. Hall, "Current transport mechanisms in $(HfO_2)x(SiO_2)$ 1-x/SiO_2 gate stacks," Microelectron Eng **84**(9–10), 2306–2309 (2007).