Three-Dimensional Mechanistic Modeling of Gate Leakage Current in High-κ MOSFETs

Feilong Liu,[1,2] Yue-Yang Liu,[1] Ling Li,[3] Guofu Zhou,[2,4,*] Xiangwei Jiang,[1,†] and Jun-Wei Luo[1]

[1] *Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, People's Republic of China*

[2] *Guangdong Provincial Key Laboratory of Optical Information Materials and Technology & Institute of Electronic Paper Displays, South China Academy of Advanced Optoelectronics, South China Normal University, Guangzhou 510006, People's Republic of China*

[3] *Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, People's Republic of China*

[4] *Shenzhen Guohua Optoelectronics Tech. Co. Ltd., Shenzhen 518110, People's Republic of China*

(Received 9 August 2019; revised manuscript received 25 November 2019; accepted 10 January 2020; published 10 February 2020)

Gate leakage current is an issue for the reliability of modern high-κ MOSFETs. Although various physical models describing both direct tunneling and trap-assisted contribution of leakage current have been presented in the literature, many of them treats traps in the dielectric as a continuum distribution in energy and position, and trap-to-trap transport of electrons has so far been mostly neglected or not treated three dimensionally. In this work, we present a mechanistic model for calculation of gate leakage current in high-κ MOSFET multilayer stacks based on multiphonon trap-assisted tunneling theory, taking into account the intrinsic three-dimensional (3D) discreteness of traps in the dielectric. Our model can to a good approximation reproduce the experimental results at different dielectric thicknesses, gate voltages, temperatures, and different gate materials. We find that in realistic devices, the 3D trap-to-trap transport of electrons contributes a non-negligible part to the gate leakage current. This contribution is more pronounced at low-voltage device operations, which is important for low-power applications. We calculate the intrinsic fluctuation of gate leakage current due to positional and energetic disorder of traps in the dielectric, and conclude that positional disorder is more important than energetic disorder for realistic material parameters. The calculated gate leakage current depends sensitively on temperature, trap energy, and trap density. We provide a computationally efficient 3D master equation approach that enables 3D mechanistic simulation of 10^3 traps on the order of minutes on a standard desktop computer.

DOI: 10.1103/PhysRevApplied.13.024020

I. INTRODUCTION

MOSFETs gated with high dielectric constant (high-κ) materials have become the mainstream technology for transistors in integrated circuits. Compared with the conventional SiO_2 dielectric, high-κ dielectric materials enable further scaling down of the equivalent oxide thickness (EOT) while significantly suppressing the gate leakage current due to direct tunneling of charge carriers through the ultra-thin oxide layer. However, different from high-quality thermally grown SiO_2, deposited high-κ dielectric materials usually suffer more from defects and trap states, causing device reliability issues that have been studied extensively [1–8]. One of the key reliability issues is the gate leakage current, which directly hampers the transistor performance. It is crucial to understand in depth its physical mechanism for rational device optimization. For high-κ devices, although direct tunneling has been suppressed, the electrons can alternatively penetrate through the dielectric layer via trap states. The trap-induced current may actually be the dominant leakage mechanism.

Research on gate leakage current has already started since the early 2000s. At first, physical models for direct tunneling current through a single-layer dielectric were proposed [9–13]. It was then realized that an interlayer of ultra-thin SiO_2 usually exists in between the Si substrate and the deposited high-κ layer, composing a multilayer gate stack. Physical models of multilayer tunneling based on empirical analytical equations, WKB, or transfer matrix methods were developed [14–20].

In addition to direct tunneling, trap-assisted transport models through the gate dielectric were also proposed, in the framework of analytical or semianalytical "compact modeling" [21–27]. In these models, usually charge transport between the gate or substrate and a single trap is first considered. Then the model is extended (usually integrated numerically) for a continuous spatial and energetic trap

*guofu.zhou@m.scnu.edu.cn
†xwjiang@semi.ac.cn

distribution [described as $N_T(x_T, E_T)$] in the dielectric. By tuning fitting parameters, experimental results can be explained.

However, to our understanding, these models are not adequate in describing the underlying physics in the scenario due to the following reasons. (1) The process of a charge carrier transferring in between two or more traps is not taken into account; all charge carriers are assumed to penetrate through the dielectric via at maximum one trap state. As a rough estimate, considering a device with a 7-nm channel length, a width:length (W:L) ratio of 10, and a high-κ dielectric thickness of 4 nm, the trap density extracted from experiments is typically in the range of $10^{18} - 5 \times 10^{20}$ cm^{-3} [28–32]. Then the number of traps in each transistor can vary from 2 to 1000. In the high trap-density limit, it is very likely that charge carriers can transfer between Si conduction channel and gate via several traps, contributing significantly to the gate leakage current. (2) In the low trap-density limit, although trap-to-trap conduction might not be relevant, the low trap density can lead to significant fluctuations in leakage current among different transistors due to positional and energetic disorder of traps in the intrinsically amorphous dielectric thin film [28,32–35]. (3) Percolation and current filaments. It has been shown both experimentally and theoretically in the oxide breakdown process that the trap-induced current is highly nonuniform throughout the dielectric layers [36–41]. (4) Random telegraph noise and stochastic charge trapping, i.e., the transport of charge carriers associated with traps observed in experiments are stochastic (like a Markov chain process) [42–45]. Due to these reasons, a mechanistic physical model that takes into account the intrinsic three-dimensional (3D) discreteness of trap sites is in high demand for both scientific community and semiconductor industry.

To our best knowledge, the first physical model that tries to include these 3D features in the dielectric was proposed by Larcher in 2003 [33]. In that model, a 3D spatial distribution of traps with transport and percolation through multiple traps is taken into account. However, perhaps due to numerical reasons, in that model the trap-assisted transport is simplified to a quasi-1D percolation path rather than calculating the full 3D percolation network.

In recent years, an advanced 3D mechanistic modeling based on kinetic Monte Carlo (KMC) simulation for trap-assisted leakage current has been developed by Lugli *et al.* [46–50]. This model not only well describes the experimental current-voltage characteristics for various dielectric materials at different temperatures, but also automatically incorporates the mesoscopic physics described above. We also note that the 3D KMC technique has been applied to high-κ MOSFETs for device reliability studies [51,52]. However, in this paper, instead of using KMC, which relies on the statistics of a huge number of charge carrier dynamics, we present a much more computationally efficient algorithm—the 3D master equation approach, making 3D mechanistic simulations of 10^3 traps possible on the order of minutes on a standard desktop computer [53–55]. While Lugli *et al.* mainly applied the model to metal-insulator-metal (M-I-M) capacitors, in this work, we apply his basic ideas to investigation of leakage current in MOSFETs, with charge injection into traps of the dielectric from a semiconductor rather than a metal. The physics of charge injection from a metal and from a semiconductor is fundamentally different, as shown in Sec. II. Because there is abundant density of states (DOS) at the Fermi level of a metal while the DOS is very limited at the conduction- or valence-band edge of a semiconductor. This also leads to significantly different temperature dependence of the direct tunneling current. While in M-I-M capacitors the direct tunneling current is temperature insensitive, in MOSFETs the direct tunneling current is very sensitive to temperature due to the state filling of the conduction band with increasing temperature. We also extend the model from a single-layer dielectric to a realistic multilayer dielectric stack, because a SiO_2 interlayer is usually inevitable between Si substrate and high-κ material. Because the interlayer is very thin (<1 nm), atomistic level simulation has shown that the conduction- and valence-band energies do not change abruptly but gradually. A Gaussian distribution of trap energy levels is included in the model, which is extracted from bias temperature instability (BTI) studies [28,32,56,57].

In addition to device-level models, many researchers have already studied trap-assisted charge transport from the material level using first-principles density-functional-theory (DFT) calculations [5,49,58–70]. These atomistically resolved calculations are good at revealing the origin of a trap by determining its energy levels, especially at different charged states, and its coupling to lattice vibrations (phonons). The DFT calculations can also give the accurate DOS at the Si conduction and valence band, gate, and the dielectric, as well as the charge-transfer rates between a trap and substrate or gate. However, these calculations are very computationally expensive. Typical calculations can only include one (or up to a few) trap(s) and are very difficult to perform an entire device calculation. Ideally, the charge-transfer rates and the trap-energy distributions calculated from first-principles calculations could, in principle, be combined with the mesoscopic 3D device simulation presented in this work, together to achieve a parameter-free predictive simulation of high-κ MOSFETs from atom to device [71–73]. However, that is beyond the scope of this paper. In this work, we focus on using the mesoscopic 3D device simulation to investigate the leakage current mechanism. The traps in the dielectric (mostly oxygen vacancies) are treated as points on which localized charge carriers can reside. The charge-transfer rates are modeled within the framework of an effective-mass-based semiclassical theory.

This paper is organized as follows. In Sec. II, we present the main framework of the 3D mechanistic modeling of gate leakage current of high-κ MOSFETs. More technical details of the model are presented in the Supplemental Material [74]. In Sec. III, we use the model presented to explain experimental data in the well-known literature, and study in depth the physical mechanism of leakage current. In Sec. IV, we conclude this paper and envision possible next steps in future work.

II. MODEL

The device structure under study is shown in Fig. 1. In this work, we take *n*-type MOSFETs as an example and only consider the case that electrons transfer from Si substrate to gate (under a positive gate voltage V_{gate}). *p*-type MOSFETs or *n*-type MOSFETs under a negative gate bias can, in principle, be studied following a similar approach. HfO_2 is taken as the example high-κ material. The band diagrams at flat band are shown in Figs. 1(a) and 1(b). In addition to the HfO_2 layer and the SiO_2 layer in between Si and HfO_2, two transition layers are also considered at Si/SiO_2 and SiO_2/HfO_2 interfaces. The conduction- and valence-band energies are assumed to change linearly in these transition regions [28,70]. Following Ref. [28], we only consider traps in the Si/SiO_2 transition layer and the HfO_2 layer. For the gate part, we consider two different cases: (1) a heavily doped *n*-type poly-Si; (2) a Pt metal gate. Due to the amorphous nature, the trap energies in the dielectric are assumed to follow a Gaussian distribution [28,32]. In Fig. 1(c), the 3D spatial distribution of traps is shown. Most traps in the dielectric are from the HfO_2 layer. All material and device parameters used in simulation are shown in Table I, with values obtained from the literature. For the two devices shown, the trap densities and trap energies are different in simulation. The optimized values are obtained by fitting experimental leakage current data, which is discussed in detail in Sec. III. We note that in the literature, there is no consensus on trap density and energy, depending significantly on the processing condition, characterizing method, and methods of parameter extraction. For (shallow) electron trap density in HfO_2, values from 1×10^{18} cm^{-3} to 2×10^{20} cm^{-3} have been reported [28,41,75,76]. For (shallow) electron trap energy in HfO_2, values range from -0.85 to -2.2 eV with respect to HfO_2 conduction-band edge [28,34,41,70,75,77–79]. The values extracted from our model (given in Sec. III) are well within the literature range.

The charge-transfer rates are modeled within an effective-mass-based semiclassical approach. Following the work of previous researchers [46,50,75,86,87], we consider the following processes. (1) Charge transfer between Si conduction channel and a trap in the dielectric. (2) Charge transfer between gate and a trap in the dielectric. (3) Charge transfer between two traps in the dielectric layer. (4) Poole-Frenkel emission, i.e., charge transfer from a trap to the conduction band of the dielectric.

Charge transfer from Si to a trap in the dielectric can either be an elastic (coherent) [50,88] or an inelastic (incoherent) process [27,33,50,75,89]. For an elastic process, the energy of an injecting electron in the Si conduction band is equal to the trap energy. For an inelastic process, the initial and final energies of the electron are different, and the electron transfer is assisted by either absorbing or emitting phonons. The inelastic process is described in the framework of a multiphonon trap-assisted tunneling model. The rates of charge-transfer processes ω are determined by the DOS g of Si conduction band, the Fermi-Dirac distribution f, the tunneling transmission coefficient $\overline{T}$, and (for the inelastic process only) the electron-phonon coupling strength L_m, given by Refs. [50,88,89]

$$\omega^{\text{inela}}_{\text{Si/gate}\to\text{trap},j} = c_{0,j}\left[\sum_{m<0}^{-\infty} g_{\text{Si/gate}}(E_{m,j}) f\left(E_{m,j},\Phi_{\text{Si/gate}}\right)\overline{T}(E_{m,j},L_x,x_{T,j})L_m \exp\left(\frac{m\hbar\omega_0}{k_BT}\right) + \sum_{m>0}^{+\infty} g_{\text{Si/gate}}(E_{m,j}) f\left(E_{m,j},\Phi_{\text{Si/gate}}\right)\overline{T}(E_{m,j},L_x,x_{T,j})L_m\right], \tag{1}$$

$$\omega^{\text{inela}}_{\text{trap}\to\text{Si/gate},j} = c_{0,j}\left\{\sum_{m<0}^{-\infty} g_{\text{Si/gate}}(E_{m,j})\left[1-f\left(E_{m,j},\Phi_{\text{Si/gate}}\right)\right]\overline{T}(E_{m,j},x_{T,j},L_x)L_m + \sum_{m>0}^{+\infty} g_{\text{Si/gate}}(E_{m,j})\left[1-f\left(E_{m,j},\Phi_{\text{Si/gate}}\right)\right]\overline{T}(E_{m,j},x_{T,j},L_x)L_m \exp\left(-\frac{m\hbar\omega_0}{k_BT}\right)\right\}. \tag{2}$$

$$\omega^{\text{ela}}_{\text{Si/gate}\to\text{trap},j} = \left(\frac{m^*_{\text{Si/gate}}}{m^*_i}\right)^{2.5}\frac{8\left(E_{T,j}-\Phi_{\text{Si/gate}}+\Sigma_{\text{Si/gate}}\right)^{1.5}}{3\hbar\sqrt{E_{D,j}}} f\left(E_{T,j},\Phi_{\text{Si/gate}}\right)\overline{T}(E_{T,j},L_x,x_{T,j}), \tag{3}$$

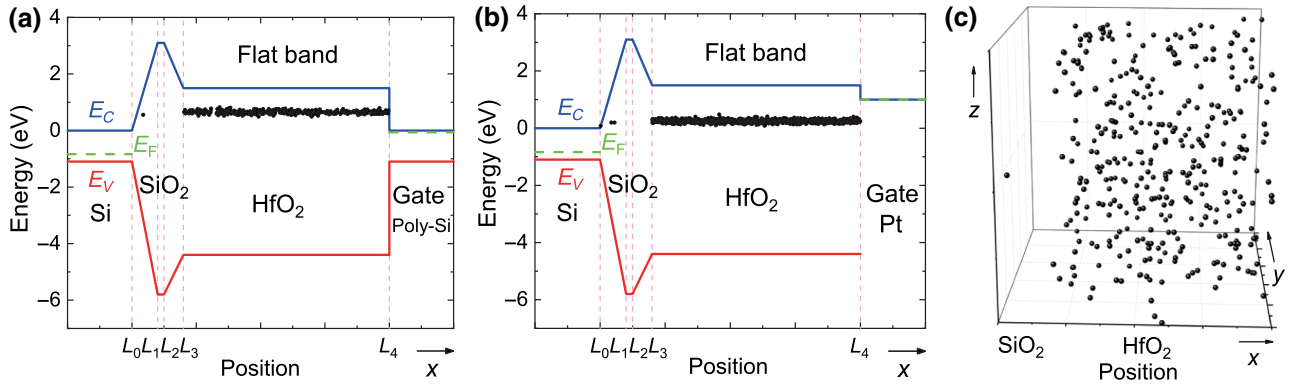

FIG. 1. Schematic band diagrams of the $Si/SiO_2/HfO_2$/gate system under study at flat band, with a (a) poly-Si or (b) Pt gate. Electron traps are located in the Si/SiO_2 transition region and the HfO_2 layer. (c) An example of 3D spatial distribution of traps in the dielectric.

and

$$\omega^{\text{ela}}_{\text{trap}\to\text{Si/gate},j} = \left(\frac{m^*_{\text{Si/gate}}}{m^*_i}\right)^{2.5} \frac{8(E_{T,j} - \Phi_{\text{Si/gate}} + \Sigma_{\text{Si/gate}})^{1.5}}{3\hbar\sqrt{E_{D,j}}} \left[1 - f\left(E_{T,j}, \Phi_{\text{Si/gate}}\right)\right] \overline{T}(E_{T,j}, x_{T,j}, L_x). \tag{4}$$

k_B is the Boltzmann constant and T is temperature. L_x is the total dielectric thickness and $x_{T,j}$ is the x coordinate of trap j. m^*_i is the effective mass in the ith layer of dielectric stack. $m^*_{\text{Si/gate}}$ is the effective mass at Si/gate side. The prefactor $c_{0,j}$ is given by $c_{0,j} = [(4\pi)^2 r^3_{t,j}/E_{g,i}](e^2\hbar F^2_j/2m^*_i)$ [87]. $E_{g,i}$ is the band gap of the ith layer. F_j is the electric

TABLE I. List of parameters.

Parameter	Value	Remark
Si/SiO_2 transition region width d_1	0.4 nm	Refs. [28,70]
SiO_2/HfO_2 transition region width d_3	0.3 nm	Refs. [28,70]
Si conduction band $E_{c,\text{Si}}$	0	Reference energy[a]
Si valence band $E_{v,\text{Si}}$	−1.1 eV	Ref. [80]
SiO_2 conduction band E_{c,SiO_2}	3.1 eV	Ref. [9]
SiO_2 valence band E_{v,SiO_2}	−5.9 eV	Ref. [80]
HfO_2 conduction band E_{c,HfO_2}	1.5 eV	Ref. [16]
HfO_2 valence band E_{v,HfO_2}	−4.3 eV	Ref. [8]
Pt work function	1.0 eV	Ref. [81]
Pt Fermi energy Σ_{Pt}	9.74 eV	Ref. [82]
Si substrate p-type doping concentration N_{sub}	10^{15} cm^{-3}	
Poly-Si gate n-type doping concentration N_{poly}	10^{20} cm^{-3}	See footnote[b]
Si intrinsic carrier concentration n_i	10^{10} cm^{-3}	Ref. [80]
Si dielectric constant $\varepsilon_{r,\text{Si}}$	11.9	Ref. [80]
SiO_2 dielectric constant ε_{r,SiO_2}	3.9	Ref. [8]
HfO_2 dielectric constant ε_{r,HfO_2}	25	Ref. [8]
SiO_2 and HfO_2 trap energy distribution width σ_T	0.05 eV	fitting parameter
Si electron conductivity effective mass m^*_{Si}	0.3 m_0[c]	Refs. [11,80]
SiO_2 electron tunneling effective mass $m^*_{SiO_2}$	0.5 m_0	Refs. [9,16,83,84]
HfO_2 electron tunneling effective mass $m^*_{HfO_2}$	0.2 m_0	Refs. [9,16,83]
Huang-Rhys factor S in dielectric	10	Ref. [50]
Phonon energy $\hbar\omega_0$ in dielectric	0.04 eV	Ref. [50]

[a]All energy values are given at flat band.
[b]N_{sub} and N_{poly} together leads to a flat band voltage $V_{\text{FB}} \approx -0.77$ V given in Ref. [85].
[c]m_0 is the free electron mass.

field at the trap j. $\Phi_{\mathrm{Si/gate}}$ is the Fermi level at Si/gate side. $E_{m,j} = E_{T,j} + m\hbar\omega_0$ is the trap energy $E_{T,j}$ plus ($m > 0$) or minus ($m < 0$) the phonon energy associated with the transition. m is the number of phonons involved in the charge transfer. $\hbar\omega_0 = 40$ meV is the assumed phonon energy [50]. $r_{t,j}$ is the localization radius of the trapped electron in the j th trap. $E_{D,j}$ is the trap energy with respect to the conduction-band energy. Σ_{gate} is the Fermi energy of metal and is assumed to be zero for Si.

The charge-transfer rate from Si to a trap as a function of the energy difference between final and initial states are shown as solid curves in Fig. 2(a). For the elastic process, the rate is only nonzero for $\Delta E > 0$, because there are no states available in the Si band gap. For the inelastic process, the rate is continuous in the whole ΔE range due to possible coupling to phonons. If both elastic and inelastic processes are considered, the total charge-transfer rate results in a discontinuity, leading to an unrealistic shape of calculated current density-voltage curves. Due to this reason, we only consider the inelastic process from Si to a trap, similar to other previous models [27,75]. We note that the physics of charge-transfer process from a semiconductor to a trap is different from the process from a metal to a trap [50]. As shown as the dashed curves in Fig. 2(a), because there are plenty of states available at the Fermi level in metal, there is no discontinuity in the charge-transfer rates as a function of ΔE. In physics, the discontinuity and the choices of inelastic and elastic processes made by us and other researchers are due to the limitation of an effective-mass-based model. In reality, the band tail of Si prevents the discontinuity from occuring [90]. A next-level charge-transfer rate model, such as the DFT calculations, solves the problem due to effective-mass approximation and will be incorporated in the future [70].

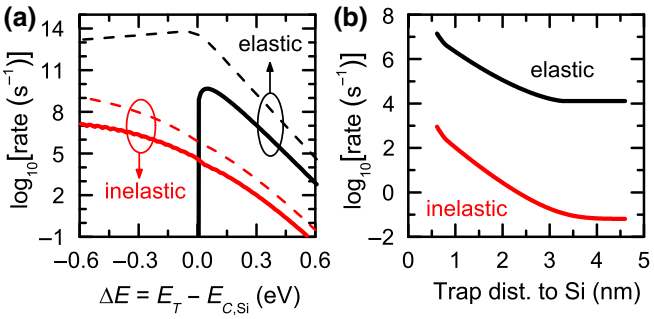

FIG. 2. Charge-transfer rate from Si conduction channel to a trap in the HfO_2 layer. (a) Solid lines: rates of elastic and inelastic processes as a function of the energy difference between the electron trap and Si conduction-band edge. Dashed lines: rates of elastic and inelastic processes from a metal to a trap in the dielectric, with abundant DOS at the Fermi level. (b) Rates of elastic and inelastic processes as a function of the distance between the electron trap and Si/SiO_2 interface.

In Fig. 2(b), it is shown that the distance dependence of charge-transfer rate from Si to a trap is not simply an exponential decay. That is because at a negative electric field, the conduction band decreases with x. At a large distance, the wave function of the incoming electron might be close to or even arrive at an unbound state. The tunneling transmission coefficient thus does not simply decay with distance. The actual shape of the decay also depends on the electric field in the dielectric. This nonexponential decay with distance has also been confirmed recently in more advanced atomistic level DFT calculations [70].

For charge transfer between a trap and metal gate, both elastic and inelastic processes are included. For simplicity, the DOS shape of Si and metal are both assumed to be the free electron parabolic DOS. The transmission coefficient is approximated using the WKB method for charge transfer between Si/gate and a trap. For a multilayer stack, the transmission coefficient T is approximated as the product of transmission through each individual layer. This approximation is valid when $T \ll 1$ [14].

The charge transfer between two traps in the dielectric layer is described using the Miller-Abrahams formula [50,91],

$$\omega_{a\to b} = \nu_0 \exp\left(-\frac{r_{ab}}{r_{D,a}}\right)\left(-\frac{r_{ab}}{r_{D,b}}\right) \times \begin{cases} e^{-(E_b-E_a)/(k_BT)}, & E_b - E_a > 0, \\ 1, & E_b - E_a \leqslant 0, \end{cases} \quad (5)$$

where $r_{D,a} = \hbar/\sqrt{2m_i^* E_{D,a}}$, $r_{D,b} = \hbar/\sqrt{2m_j^* E_{D,b}}$, and $r_{ab} = \sqrt{(z_{T,b} - z_{T,a})^2 + (y_{T,b} - y_{T,a})^2 + (x_{T,b} - x_{T,a})^2}$. The hopping attempt frequency ν_0 is assumed to be 10^{13} Hz [50].

This formula assumes that the charge carrier is localized and the transition is incoherent. The wave function of the trapped electron is localized not only due the intrinsic property of a trap (i.e., oxygen vacancy), but also due to the Anderson localization effect caused by the randomness of trap distribution in the dielectric [92].

Without further information, we assume the traps to be randomly distributed in space and the trap number and positions are fixed throughout the entire calculation. We note that in time-dependent dielectric breakdown studies the situation is more complex [41,93], however, we decide not to include these advanced effects in this work.

In addition, the Poole-Frenkel emission is also included in the model, during which the trapped electron can be excited to the conduction band of the dielectric [46,50]. This process is usually negligible compared to the others, except when the trap is very shallow (trap energy very close to the conduction band).

The trap-assisted charge transport and associated leakage current density J_{trap} are calculated by the 3D master equation approach, originally developed for disordered

organic semiconductor devices [53,54,94,95]. Instead of calculating the dynamics of every electron in KMC, the 3D master equation approach calculates the property of electron occupation probability of each trap, thus greatly reducing computational cost while maintaining the full 3D physics. The 3D master equation is written as

$$\sum_j \left[\omega_{ij} p_i (1 - p_j) - \omega_{ji} p_j (1 - p_i)\right] = 0, \tag{6}$$

where p_i is the time-averaged electron occupancy of trap i. ω_{ij} is the charge-transfer rate described in Eqs. (1)–(5). The term $1 - p_i$ avoids two electrons occupying the same trap simultaneously. For Si and gate, they are specially treated in the model so that they can always supply or drain charge carriers in the dielectric [94,95].

The contribution of a charge-transfer process to J_{trap} is proportional to the charge-transfer rate, the probability that an electron is available in the initial state, and the probability that the final state is unoccupied by an electron, given by

$$J_{\text{trap}} = \frac{e}{L_x L_y L_z} \sum_{i,j} \omega_{ij} p_i (1 - p_j)(x_{T,j} - x_{T,i}), \tag{7}$$

where L_y and L_z are the lateral dimensions of the simulation box. All the traps in the dielectric are taken into account and the master equations of the system are solved

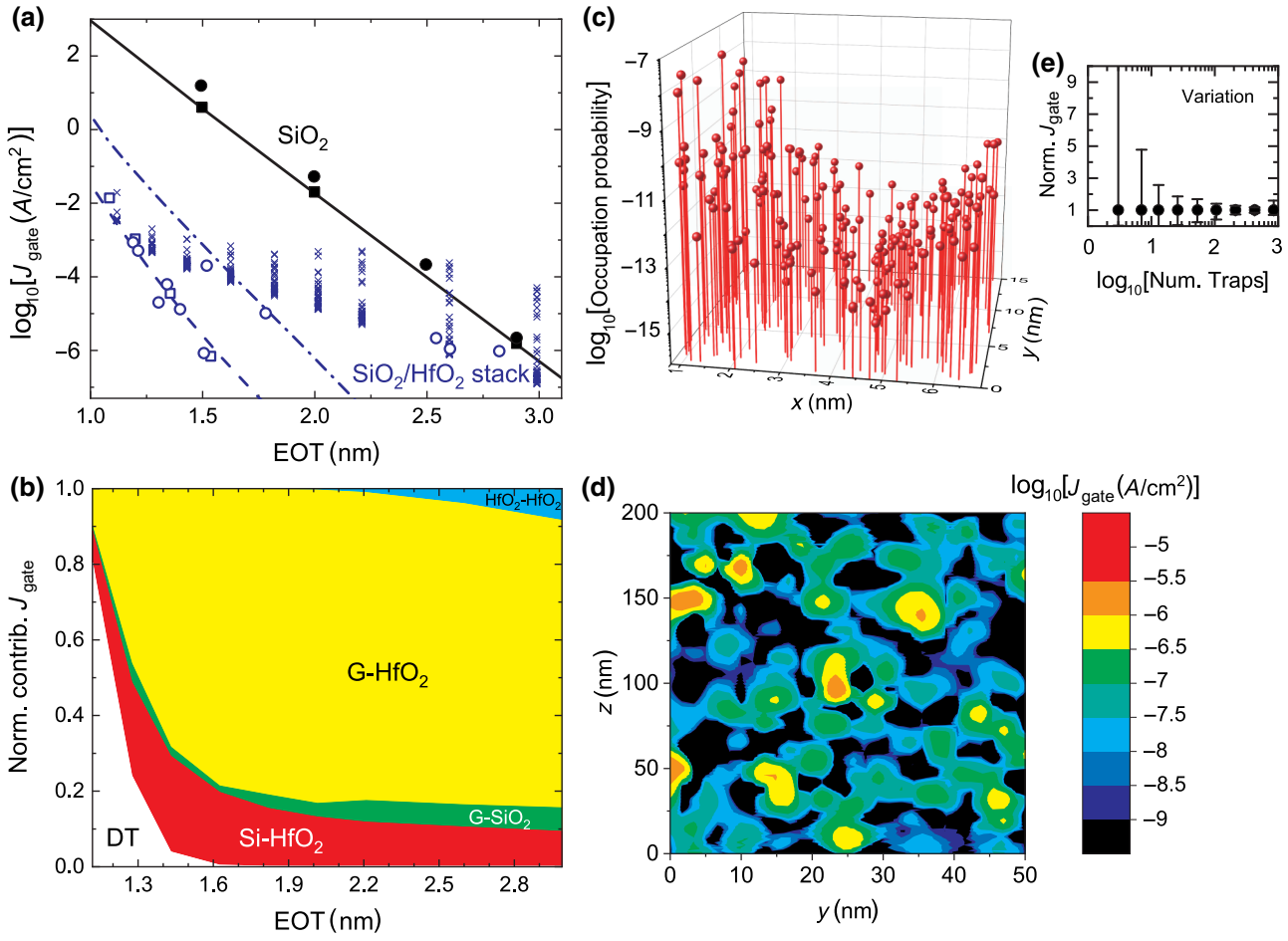

FIG. 3. J_{gate} for poly-Si-gated MOSFETs at $V_{\text{gate}} = 1$ V. The device structure is shown in Fig. 1(a). (a) J_{gate} as a function of EOT. Filled circles (squares) are experimental data of SiO_2-only devices from Ref. [9] ([8]). Open circles (squares) are experimental data of SiO_2/HfO_2-stack-gated devices from Ref. [9] ([8]). Solid curve: calculated direct tunneling current density J_{tun} of SiO_2-only devices. Dashed curve: calculated J_{tun} of high-κ devices with a 0.65-nm-thick SiO_2 layer. Dash-dotted curve: calculated J_{tun} of high-κ devices with a 1-nm-thick SiO_2 layer. Crosses: calculated total gate current density $J_{\text{gate}} = J_{\text{tun}} + J_{\text{trap}}$ of high-κ devices with a 0.65-nm-thick SiO_2 layer. (b) Calculated normalized contributions to J_{gate}. G, gate; DT, direct tunneling. (c) 2D (x-y) distribution of electron occupation probability of traps. In this example, Si conduction channel is located at $x = 0$ and gate is located at $x = 6.9$ nm. (d) 2D (y-z) distribution of J_{gate}. To better visualize the nonuniformity, an MOSFET with a ultra-large area (50×200 nm^2) is shown as an example. (e) Calculated variation of J_{gate} as a function of MOSFET size (proportional to the number of traps in each device). The variation is due to positional and energetic disorder of traps.

under the constraint of the voltage drop across the dielectric layer V_{ox}. This is ensured by coupling the 3D master equation to the 1D discrete Poisson equation

$$\varepsilon_{r,i}\frac{V_{i+1}-V_i}{x_{T,i+1}-x_{T,i}}-\varepsilon_{r,i-1}\frac{V_i-V_{i-1}}{x_{T,i}-x_{T,i-1}}=\frac{ep_i}{\varepsilon_0 L_y L_z}, \tag{8}$$

with the boundary conditions of the electrostatic potential $V_{Si}=0$ and $V_{gate}=V_{ox}+(\Phi_{gate}-\Phi_{Si})/e$. These equations are solved using Newton's iteration method with a proper choice of Jacobian matrices [55]. Once the steady-state solution is found, J_{trap} is calculated by summing up all charge-transfer processes in the x direction. The source-drain voltage is assumed to be zero throughout this paper.

The direct tunneling current density through the dielectric J_{tun} is due to electrons tunneling from Si to gate. Holes tunneling from gate to Si is negligible because of a much higher tunneling barrier. The direct tunneling current is calculated using the Tsu-Esaki equation [96],

$$J_{tun}=\frac{m^*_{Si}ek_BT}{2\pi^2\hbar^3}\int_{E_{c,Si}}^{+\infty}\overline{T}(E)N(E)\,dE,$$
$$N(E)=\ln\left[\frac{1+\exp\left(\frac{\Phi_{Si}-E}{k_BT}\right)}{1+\exp\left(\frac{\Phi_{gate}-E}{k_BT}\right)}\right], \tag{9}$$

with the transmission coefficient calculated from the transfer matrix (TM) method [19]. The TM method is more accurate than WKB because it also includes the resonance effect during the tunneling process. The total gate leakage current density is the sum of direct tunneling and trap-assisted contributions $J_{gate}=J_{tun}+J_{trap}$.

The image charge effect is also important because it lowers the tunneling barrier height of the SiO_2 layer and thus enhances tunneling current. This effect is included in the model following Refs. [11,54,97].

Finally, we need to relate the voltage across the dielectric stack V_{ox} to the applied gate voltage V_{gate}. For the two device examples studied in this work, they are both close to channel inversion even at zero gate voltage. The applied voltage mostly drops across the dielectric stack for a metal gate. For a poly-Si gate, upon an applied gate bias, the poly-Si gate becomes partly depleted. We include the relation between V_{gate} and V_{ox} following the model in Ref. [13].

All technical details of the model are presented in the Supplemental Material [74].

III. RESULTS AND DISCUSSION

The calculated and experimental results of J_{gate} for poly-Si-gated MOSFETs are shown in Fig. 3 at various

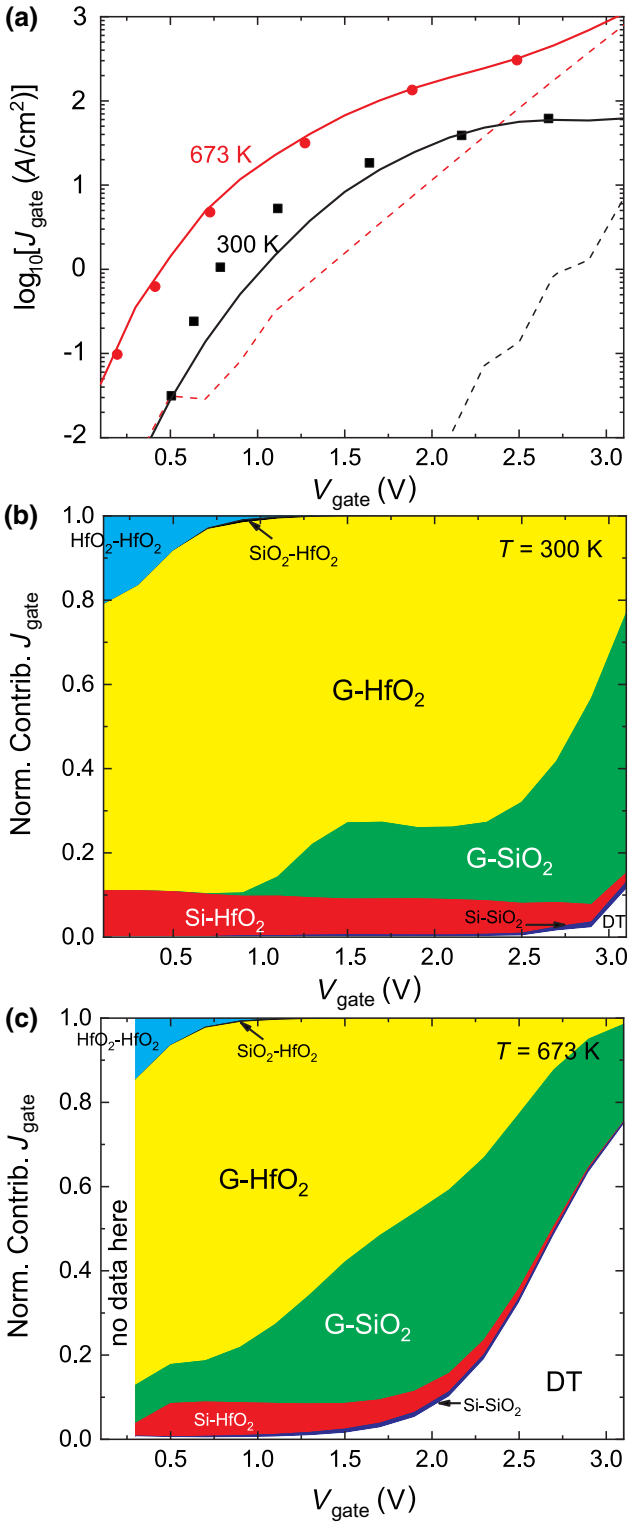

FIG. 4. J_{gate} for a Pt-gated MOSFET. The device structure is shown in Fig. 1(b). (a) $J_{gate}-V_{gate}$ characteristics at high (673 K) and room (300 K) temperatures with EOT = 1.63 nm. Symbols: experimental data from Ref. [81]. Solid curves: calculated $J_{gate}=J_{tun}+J_{trap}$. Dashed curves: calculated J_{tun} only. (b),(c) Calculated normalized contributions to J_{gate} at 673 and 300 K, respectively. G, gate; DT, direct tunneling.

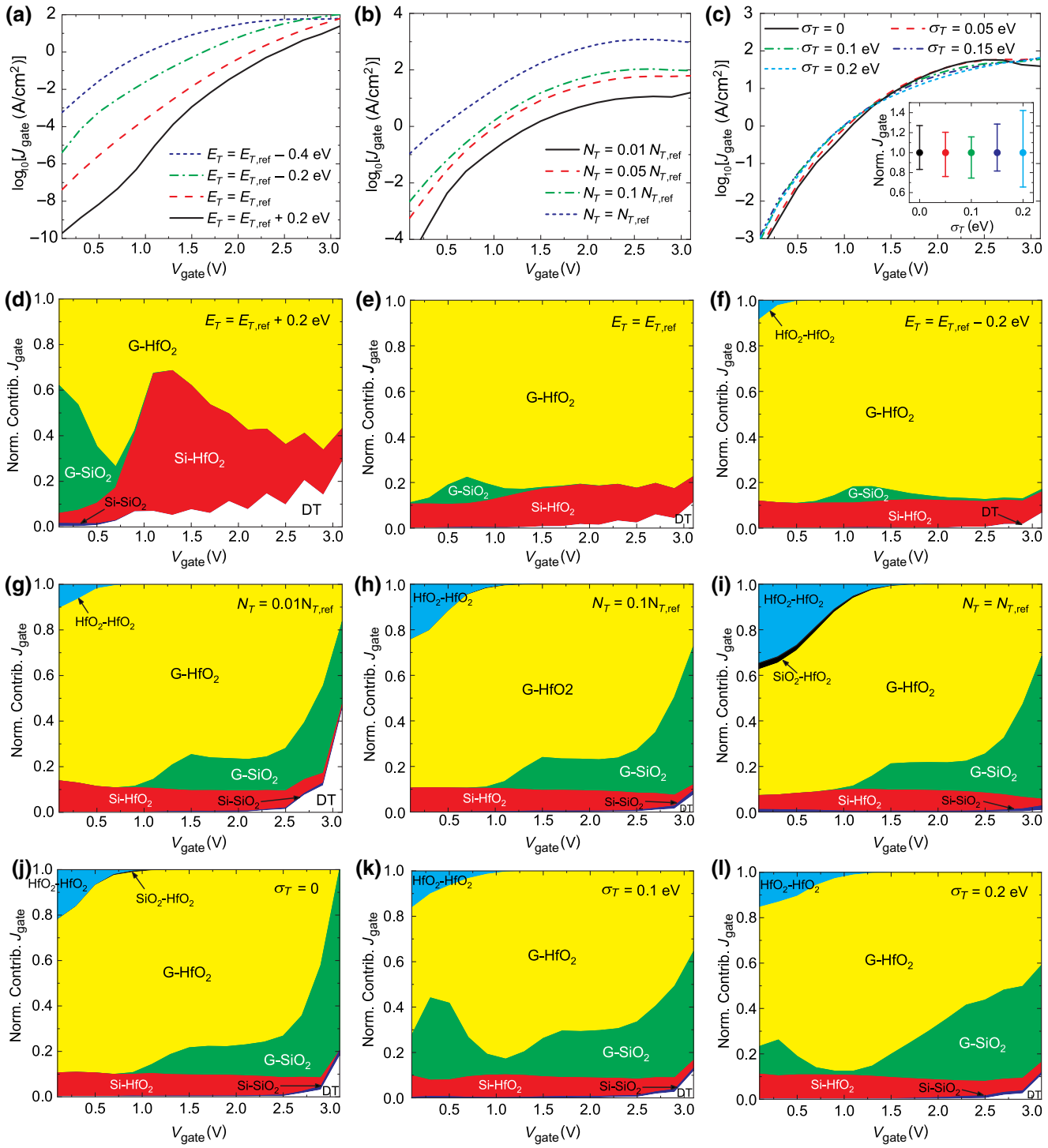

FIG. 5. $J_{\text{gate}} - V_{\text{gate}}$ characteristics for a Pt-gated MOSFET at (a) various trap energies E_T, (b) trap densities N_T, and (c) trap energy distribution widths σ_T. Inset of (c): variation of J_{gate} at $V_{\text{gate}} = 1.1$ V as a function of σ_T. (d)–(f) Calculated normalized contributions to J_{gate} at different E_T. (g)–(i) Calculated normalized contributions to J_{gate} at different N_T. (j)–(l) Calculated normalized contributions to J_{gate} at different σ_T. G, gate; DT, direct tunneling.

EOTs. Our model can excellently reproduce the experimental results for the cases when direct tunneling is dominant. When trap-assisted transport is dominant, in order to fit the experimental data, the trap energies are taken to be $E_{\text{TSiO}_2} = E_{c,\text{SiO}_2} - 2.52$ eV and $E_{\text{THfO}_2} = E_{c,\text{HfO}_2} - 0.92$ eV; the trap density are taken to be $N_{\text{TSiO}_2} = 3 \times 10^{17}$ cm^{-3} and $N_{\text{THfO}_2} = 1.1 \times 10^{19}$ cm^{-3}. The agreement with experimental data is not perfect because the

trap-assisted gate current is sensitive to trap energy level and density (see discussions below), which is difficult to control accurately when fabricating samples with different thicknesses.

For SiO_2-gated device, J_{gate} is always dominated by direct tunneling. For a high-κ device, J_{gate} is dominated by direct tunneling only at very small EOTs. At larger EOTs, J_{gate} is dominated by trap-assisted transport. J_{gate} is very sensitive to the SiO_2 interlayer thickness. Most contribution to J_{gate} comes from the charge transfer between gate and traps in HfO_2. The charge transfer between two HfO_2 traps starts to play a role at EOT > 2.2 nm for the case of $V_{gate} = 1$ V shown in this figure.

The electron occupation probability of traps are always much smaller than 1, meaning that the traps are unoccupied most of the time. The minimum of the occupation probability occurs near the middle of the dielectric and the maximum occurs near the Si substrate from which electrons are injected into the traps. We note that for p-type HfO_2-gated MOSFETs this might be totally different because the hole trap energy there lies within the band gap of Si [28,70]. The gate leakage is highly nonuniform, with a distribution of "hot spots" through which electrons can penetrate easily. Significant variation of J_{gate} is found in calculation due to positional and energetic disorder of electron traps. When the number of traps in a single MOSFET is less than 100, variation of J_{gate} increases rapidly with decreasing device area. This leads to an intrinsic fluctuation of leakage current as the device size scales down. This effect is to some extent similar to the random dopant fluctuation effect [98], which is also caused by atomistic level randomness.

The calculated and experimental results of $J_{gate} - V_{gate}$ characteristics for a Pt-gated MOSFET are shown in Fig. 4 at high (673 K) and room (300 K) temperatures. The model calculation can to a good approximation agree with experimental results, provided that the trap energy and trap density are both adjusted from those in the poly-Si-gated devices by 0.4 eV downwards and 2.5 times larger. The oscillation in the calculated J_{tun} is due to the transition of tunneling barrier from a trapezoidal-like shape to triangularlike shape (Fowler-Nordheim tunneling) [12]. Although J_{tun} is much smaller than J_{trap} at low and medium voltages, J_{tun} has a much stronger voltage dependence and dominates at high voltages. J_{tun} is also shown to have a high sensitivity on temperature. This is fundamentally different from the temperature-insensitive tunneling process with metal or metal-like electrodes [27]. Because for a metal, there is plenty of DOS at the Fermi level, which contributes mostly to J_{tun}; while for a semiconductor, the DOS at the conduction-band edge is zero, and the electrons in the conduction band for tunneling are filled by Fermi-Dirac distribution strongly affected by temperature.

The contribution of charge transfer between two HfO_2 traps to J_{gate} is significant at low voltages, and decreases with V_{gate}. Because when the electric field in the dielectric is lower, electrons can stay in the dielectric for a longer time and are prone to visit more trap sites. This means realistically that for low-power applications, it is crucial to take into account trap-to-trap transport in order to correctly describe the physics of gate leakage current.

While at low and medium V_{gate}, J_{gate} is dominated by traps in HfO_2; at high V_{gate}, J_{gate} is dominated by traps in SiO_2. This is because the SiO_2 trap energy and thus the energy barrier between Si conduction band and a SiO_2 trap decreases with V_{gate}. This decrease continues until to a point that the conduction path from Si to the gate via a SiO_2 trap opens up.

In order to understand the key parameters affecting J_{gate}, we systematically vary the trap energy E_T, trap density N_T, and trap energy distribution width σ_T, and study their effect on J_{gate} in Figs. 5(a)–5(c). It is obvious that E_T and N_T can sensitively affect J_{gate}. However, we observe that σ_T has only a minor effect on J_{gate}. To better understand this, we examine the variation of J_{gate} at different σ_T, and found that the fluctuation of J_{gate} is almost unaffected by σ_T

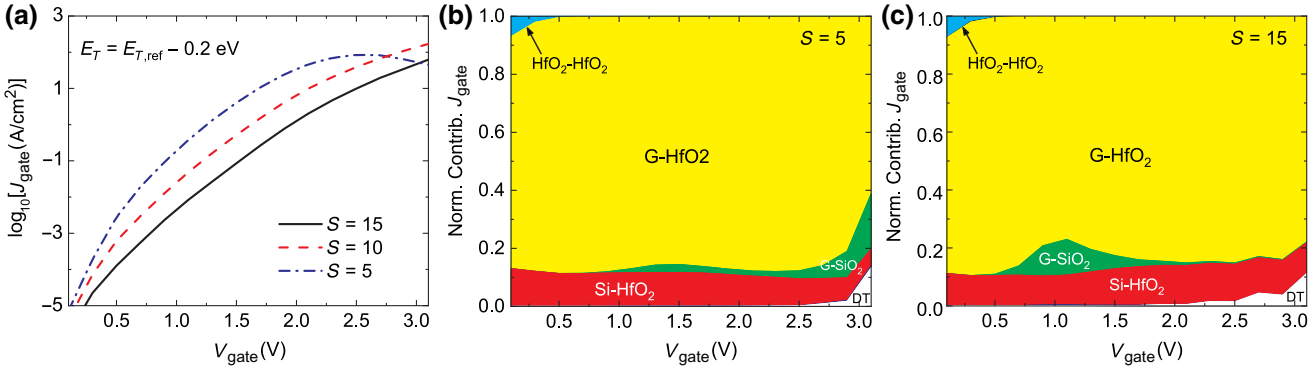

FIG. 6. (a) $J_{gate} - V_{gate}$ characteristics for a Pt-gated MOSFET at various electron-phonon coupling strength (Huang-Rhys factor S). (b),(c) Calculated normalized contributions to J_{gate} at low ($S = 5$) and high ($S = 15$) electron-phonon coupling strengths, respectively. G, gate; DT, direct tunneling.

unless $\sigma_T \geq 0.2$ eV. We thus conclude that for $\sigma_T < 0.2$ eV the positional disorder rather than energetic disorder is the main cause of the leakage current fluctuation. We also note that σ_T values of 0.15–0.25 eV have been used for electron traps in SiO_2 and HfO_2 in BTI studies [28,32].

These key parameters also strongly affect the detailed mechanism of leakage current, as shown in Figs. 5(d)–5(l). When E_T is high, trap-assisted current is low, and direct tunneling has a larger contribution to J_{gate} [Fig. 5(d)]. The oscillation is consistent with J_{tun} shown in Fig. 4(a). When E_T is high, charge transfer from Si to a HfO_2 trap close by is energetically more unfavorable, however, the charge transfer from Si to a HfO_2 trap farther away is energetically more favorable although the transfer distance is larger. This is the reason for a higher contribution of J_{gate} from Si to a HfO_2 trap in Fig. 5(d). A significant difference between Figs. 5(d)–5(f) and Fig. 4(b) is that the contribution from SiO_2 traps to gate no longer plays a role at high V_{gate} here. Because SiO_2 traps are very close to Si, V_{gate} has very limited control over the barrier height between a SiO_2 trap and Si conduction band. The conduction path via a SiO_2 trap can only form when E_T is low enough. We also note that a peak is observed for the contribution between SiO_2 traps and gate. We suspect this to be a phonon-assisted transport mechanism, which is studied in more detail in Fig. 6.

In Figs. 5(g)–5(i), the leakage current mechanism is investigated as a function of N_T. Even at a low trap density $N_T = 0.01 N_{T,ref}$ (with $N_{T,HfO_2} = 5 \times 10^{18}$ cm^{-3}), the trap-to-trap conduction is still not negligible at low voltages. With increasing N_T, the trap-to-trap conduction becomes more important. In Figs. 5(j)–5(l), the leakage current mechanism is investigated as a function of σ_T. σ_T seems to mostly affect the contribution of transport between SiO_2 traps and gate. Compared to the $\sigma_T = 0$ case, on the one hand, it is more likely to find a low-energy SiO_2 trap at $\sigma_T > 0$, which opens up the conduction path even at low V_{gate}; on the other hand, it is also more likely to find a high-energy SiO_2 trap at $\sigma_T > 0$, which lowers the contribution of transport between SiO_2 traps and gate at high V_{gate}.

In Fig. 6, the $J_{gate} - V_{gate}$ characteristics and leakage current mechanism are investigated as a function of the Huang-Rhys factor S representing the electron-phonon coupling strength. In most of the voltage range shown, J_{gate} decreases when S increases. The results here are consistent with the results in Ref. [73] that the charge-transfer rate decreases with S when a small number of phonons are coupled to the process. We also observe that at a smaller S, the contribution to J_{gate} from the charge transfer between SiO_2 traps and gate increases at high V_{gate} and decreases at low V_{gate}.

Until now we assume that the trap distribution in the dielectric is spatially uniform. However, there has been experimental evidence showing that the trap density in HfO_2 may increase with distance away from the Si/SiO_2 layers [29,76]. In Fig. 7, the $J_{gate} - V_{gate}$ characteristics and leakage current mechanism are investigated with a HfO_2 trap distribution that linearly increases with x, while the total trap density is equal to the case of a uniform HfO_2 trap distribution. Compared to the uniform trap distribution case, the leakage current is significantly reduced for the case with linearly increasing trap distribution. This indicates that J_{gate} is limited by the first step of an electron transferring from Si to a HfO_2 trap. The decrease of J_{gate} is due to the larger average distance and thus the lower electron transfer rate from Si to HfO_2 traps. As the electron transfer via HfO_2 traps is suppressed, SiO_2 traps play a significantly larger role in the contribution to J_{gate}, as shown in Fig. 7(b).

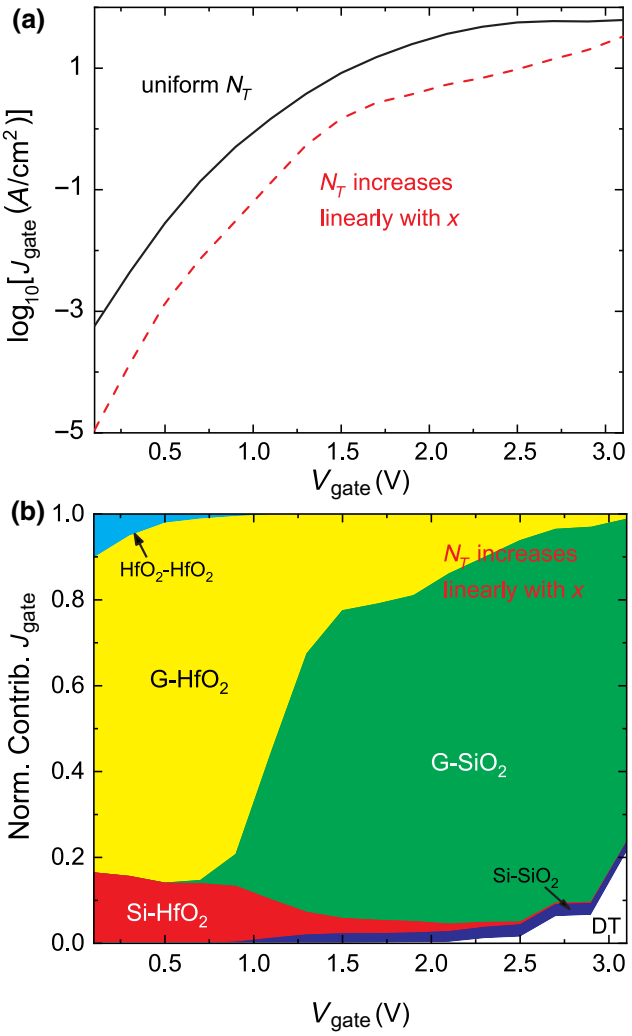

FIG. 7. (a) $J_{gate} - V_{gate}$ characteristics for a Pt-gated MOSFET with a spatially uniform HfO_2 trap distribution and with a HfO_2 trap distribution that linearly increases with x. (b) Calculated normalized contributions to J_{gate} for the HfO_2 trap distribution that linearly increases with x. G, gate; DT, direct tunneling.

IV. CONCLUSION AND OUTLOOK

We present a 3D mechanistic model of gate leakage current in high-κ MOSFETs. Experimental data are reproduced to a good approximation. Physical mechanism of leakage current is studied in detail. Different from other models, we have quantitatively identified the contribution of 3D trap-to-trap conduction to the leakage current. The contribution is more important for low-voltage and low-power applications. Our calculation shows the intrinsic fluctuation in the gate leakage current due to positional and energetic disorder of traps, and concludes that the positional disorder is more important than the energetic disorder. The leakage current is sensitively dependent on temperature, trap energy, and trap density. We also provide a computationally efficient 3D master equation approach that can help to gain fruitful physical insight of the gate leakage scenario without the need of expensive computing power.

The next step of model development could be to link the mesoscopic 3D device model presented in this work to first-principles calculations. The semiclassical effective-mass-based charge-transfer rates could be improved with rates calculated from atomistic level DFT, following the two-state and four-state nonradiative multiphonon theory [59] or Marcus theory [70], or GW calculations [49]. A more challenging task would be to accurately determine the trap energy distribution in the dielectric from DFT. We for the moment use in this model a random spatial distribution of traps, however, it is a valid question to ask whether in reality the trap distribution is truly random. For example, trap aggregation occurs if two traps are very close to each other [99]. Moreover, the number and position of traps in the model are assumed to be fixed, however, at a large leakage current stress, new oxygen vacancies may be formed dynamically due to hot electron injection [51,52,100].

ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (Grants No. 61927901, No. 11774338, andNo. 11574304), the Youth Innovation Promotion Association CAS (Grant No. 2016109), project Grant No. 6J6011000, the leading talents of Guangdong Province Program (No. 00201504), and Guangdong Innovative Research Team Program (Grant No. 2013C102). F. Liu acknowledges R. Coehoorn and P.A. Bobbert from Eindhoven University of Technology, the Netherlands, for the idea of the 3D master equation methodology.

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

[2] J. C. Ranuárez, M. J. Deen, and C.-H. Chen, A review of gate tunneling current in MOS devices, Microelectron. Reliab. **46**, 1939 (2006).

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

[4] B. Kaczer, J. Franco, P. Weckx, Ph. J. Roussel, V. Putcha, E. Bury, M. Simicic, A. Chasin, D. Linten, B. Parvais, F. Catthoor, G. Rzepa, M. Waltl, and T. Grasser, A brief overview of gate oxide defect properties and their relation to MOSFET instabilities and device and circuit time-dependent variability, Microelectron. Reliab. **81**, 186 (2018).

[5] A. L. Shluger, K. P. McKenna, P. V. Sushko, D. M. Ramo, and A. V. Kimmel, Modelling of electron and hole trapping in oxides, Modell. Simul. Mater. Sci. Eng. **17**, 084004 (2009).

[6] J. Strand, M. Kaviani, D. Gao, A.-M. El-Sayed, V. V. Afanas'ev, and A. L. Shluger, Intrinsic charge trapping in amorphous oxide films: Status and challenges, J. Phys.: Condens. Matter **30**, 233001 (2018).

[7] 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 Trans. Device Mater. Reliab. **5**, 5 (2005).

[8] J. Robertson and R. M. Wallace, High-K materials and metal gates for CMOS applications, Mater. Sci. Eng. R **88**, 1 (2015).

[9] Y.-C. Yeo, T.-J. King, and C. Hu, MOSFET gate leakage modeling and selection guide for alternative gate dielectrics based on leakage considerations, IEEE. Trans. Electron Devices **50**, 1027 (2003).

[10] H. Wu, Y. Zhao, and M. H. White, Quantum mechanical modeling of MOSFET gate leakage for high-k gate dielectrtics, Solid-State Electron. **50**, 1164 (2006).

[11] A. Schenk and G. Heiser, Modeling and simulation of tunneling through ultra-thin gate dielectrics, J. Appl. Phys. **81**, 7900 (1997).

[12] E. I. Goldman, N. F. Kukharskaya, and A. G. Zhdan, The effect of imaging forces in ultra thin gate insulator on the tunneling current and its oscillations at the region of transition from the direct tunneling to the Fowler-Nordheim tunneling, Solid-State Electron. **48**, 831 (2004).

[13] W.-C. Lee and C. Hu, Modeling CMOS tunneling currents through ultrathin gate oxide due to conduction- and valence-band electron and hole tunneling, IEEE Trans. Electron Devices **48**, 1366 (2001).

[14] J. Zhang, J. S. Yuan, Y. Ma, and A. S. Oates, Design optimization of stacked layer dielectrics for minimum gate leakage currents, Solid-State Electron. **44**, 2165 (2000).

[15] W. Wang, N. Gu, J. P. Sun, and P. Mazumder, Gate current modeling of high-k stack nanoscale MOSFETs, Solid-State Electron. **50**, 1489 (2006).

[16] T. Ando, N. D. Sathaye, K. V. R. M. Murali, and E. A. Cartier, On the electron and hole tunneling in a HfO_2 gate stack with extreme interfacial-layer scaling, IEEE Electron Device Lett. **32**, 865 (2011).

[17] Y. T. Hou, M. F. Li, H. Y. Yu, and D.-L. Kwong, Modeling of tunneling currents through HfO_2 and $(HfO_2)_x(Al_2O_3)_{1-x}$ gate stacks, IEEE Electron Device Lett. **24**, 96 (2003).

[18] H. Yu, Y.-T. Hou, M.-F. Li, and D.-L. Kwong, Investigation of hole-tunneling current through ultrathin oxynitride/oxide stack gate dielectrics in p-MOSFETs, IEEE Trans. Electron Devices **49**, 1158 (2002).

[19] B. Govoreanu, P. Blomme, M. Rosmeulen, J. V. Houdt, and K. De Meyer, A model for tunneling current in multilayer tunnel dielectrics, Solid-State Electron. **47**, 1045 (2003).

[20] Y. Zhao and M. H. White, Modeling of direct tunneling current through interfacial oxide and high-K gate stacks, Solid-State Electron. **48**, 1801 (2004).

[21] T.-E. Chang, C. Huang, and T. Wang, Mechanisms of interface trap-induced drain leakage current in off-state n-MOSFET's, IEEE Trans. Electron Devices **42**, 738 (1995).

[22] S. Takagi, N. Yasuda, and A. Toriumi, A new I–V model for stress-induced leakage current including inelastic tunneling, IEEE Trans. Electron Devices **46**, 348 (1999).

[23] J. Lee, G. Bosman, K. R. Green, and D. Ladwig, Model and analysis of gate leakage current in ultrathin Nitrided oxide MOSFETs, IEEE Trans. Electron Devices **49**, 1232 (2002).

[24] F. Jiménez-Molinos, A. Palma, F. Gámiz, J. Banqueri, and J. A. López-Villanueva, Physical model for trap-assisted inelastic tunneling in metal-oxide-semiconductor structures, J. Appl. Phys. **90**, 3396 (2001).

[25] G. Bersuker, J. H. Sim, C. S. Park, C. D. Young, S. V. Nadkarni, R. Choi, and B. H. Lee, Mechanism of electron trapping and characteristics of traps in HfO_2 gate stacks, IEEE Trans. Device Mater. Reliab. **7**, 138 (2007).

[26] F. Jiménez-Molinos, F. Gámiz, A. Palma, P. Cartujo, and J. A. López-Villanueva, Direct and trap-assisted elastic tunneling through ultrathin gate oxides, J. Appl. Phys. **91**, 5116 (2002).

[27] A. Gehring and S. Selberherr, Modeling of tunneling current and gate dielectric reliability for nonvolatile memory devices, IEEE Trans. Device Mater. Reliab. **4**, 306 (2004).

[28] G. Rzepa, J. Franco, B. O'Sullivan, A. Subirats, M. Simicic, G. Hellings, P. Weckx, M. Jech, T. Knobloch, M. Waltl, P. J. Roussel, D. Linten, B. Kaczer, and T. Grasser, Comphy – A compact-physics framework for unified modeling of BTI, Microelectron. Reliab. **85**, 49 (2018).

[29] 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**, 152907 (2006).

[30] C. D. Young, D. Heh, S. V. Nadkarni, R. Choi, J. J. Peterson, J. Barnett, B. H. Lee, and G. Bersuker, Electron trap generation in high-κ gate stacks by constant voltage stress, IEEE Trans. Device Mater. Reliab. **6**, 123 (2006).

[31] Y. Son, Y. Kim, and M. Kang, Characterization of oxide trap density with the charge pumping technique in dual-layer gate oxide, IEICE Electron. Express **14**, 1 (2017).

[32] G. Rzepa, J. Franco, A. Subirats, M. Jech, A. Chasin, A. Grill, M. Waltl, T. Knobloch, B. Stampfer, T. Chiarella, N. Horiguchi, L. Å. Ragnarsson, D. Linten, B. Kaczer, and T. Grasser, in *2017 IEEE International Reliability Physics Symposium (IRPS)*, (IEEE, Monterey, CA, USA, 2017), p. XT-11.1.

[33] L. Larcher, Statistical simulation of leakage currents in MOS and flash memory devices with a new multiphonon trap-assisted tunneling model, IEEE Trans. Electron Devices **50**, 1246 (2003).

[34] J.-P. Han, E. M. Vogel, E. P. Gusev, C. D'Emic, C. A. Richter, D. W. Heh, and J. S. Suehle, Asymmetric energy distribution of interface traps in n- and p-MOSFETs with HfO_2 gate dielectric on ultrathin SiON buffer layer, IEEE Electron Device Lett. **25**, 126 (2004).

[35] R. D. Long, C. M. Jackson, J. Yang, A. Hazeghi, C. Hitzman, S. Majety, A. R. Arehart, Y. Nishi, T. P. Ma, S. A. Ringle, and P. C. Mclntyre, Interface trap evaluation of $Pd/Al_2O_3/GaN$ metal oxide semiconductor capacitors and the influence of near-interface hydrogen, Appl. Phys. Lett. **103**, 201607 (2013).

[36] M. A. Alam, B. E. Weir, P. J. Silverman, Y. Ma, and D. Hwang, in *International Electron Devices Meeting 2000*, Technical Digest, IEDM (Cat. No.00CH37138), (IEEE, San Francisco, CA, USA, 2000), p. 529.

[37] R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel, and H. E. Maes, New insights in the relation between electron trap generation and the statistical properties of oxide breakdown, IEEE Trans. Electron Devices **45**, 904 (1998).

[38] S. Choi and Y. J. Park, in *2015 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, (IEEE, Washington, DC, USA, 2015), p. 405.

[39] J. H. Stathis, Percolation models for gate oxide breakdown, J. Appl. Phys. **86**, 5757 (1999).

[40] H. Nakamura and Y. Asai, Competitive effects of oxygen vacancy formation and interfacial oxidation on an ultrathin HfO_2-based resistive switching memory: Beyond filament and charge hopping models, Phys. Chem. Chem. Phys. **18**, 8820 (2016).

[41] L. Vandelli, A. Padovani, L. Larcher, and G. Bersuker, Microscopic modeling of electrical stress-induced breakdown in poly-crystalline Hafnium oxide dielectrics, IEEE Trans. Electron Devices **60**, 1754 (2013).

[42] T. Grasser, Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities, Microelectron. Reliab. **52**, 39 (2012).

[43] T. Grasser, K. Rott, H. Reisinger, M. Waltl, J. Franco, and B. Kaczer, in *2014 IEEE International Reliability Physics Symposium*, (IEEE, Waikoloa, HI, USA, 2014), p. 4A.5.1.

[44] J. Lee, G. Bosman, K. R. Green, and D. Ladwig, Noise model of gate-leakage current in ultrathin oxide MOSFETs, IEEE Trans. Electron Devices **50**, 2499 (2003).

[45] C. M. Chang, S. S. Chung, Y. S. Hsieh, L. W. Cheng, C. T. Tsai, G. H. Ma, S. C. Chien, and S. W. Sun, in *2008 IEEE International Electron Devices Meeting*, (IEEE, San Francisco, CA, USA, 2008), p. 1.

[46] G. Jegert, A. Kersch, W. Weinreich, U. Schröder, and P. Lugli, Modeling of leakage currents in high-κ dielectrics: Three-dimensional approach via kinetic Monte Carlo, Appl. Phys. Lett. **96**, 062113 (2010).

[47] G. Jegert, A. Kersch, W. Weinreich, and P. Lugli, Ultimate scaling of $TiN/ZrO_2/TiN$ capacitors: Leakage currents and limitations due to electrode roughness, J. Appl. Phys. **109**, 014504 (2011).

[48] G. Jegert, A. Kersch, W. Weinreich, and P. Lugli, Monte Carlo simulation of leakage currents in TiN/ZrO_2/TiN capacitors, IEEE Trans. Electron Devices **58**, 327 (2011).

[49] G. Jegert, D. Popescu, P. Lugli, M. J. Häufel, W. Weinreich, and A. Kersch, Role of defect relaxation for trap-assisted tunneling in high-κ thin films: A first-principles kinetic Monte Carlo study, Phys. Rev. B **85**, 045303 (2012).

[50] B. Weiler, T. Haeberle, A. Gagliardi, and P. Lugli, Kinetic Monte Carlo of transport processes in Al/AlO_x/Au-layers: Impact of defects, AIP Adv. **6**, 095112 (2016).

[51] Y. Li, P. Huang, S. Di, X. Zhang, G. Du, and X. Liu, Comprehensive investigation of multitraps-induced degradation in HfO_2-based nMOSFETs with interfacial layer by three-dimensional KMC method, IEEE Trans. Nanotechnol. **17**, 198 (2018).

[52] Y. Li, H. Jiang, Z. Lun, Y. Wang, P. Huang, H. Hao, G. Du, X. Zhang, and X. Liu, Reliability investigation of high-k/metal gate in nMOSFETs by three-dimensional kinetic Monte-Carlo simulation with multiple trap interactions, Jpn. J. Appl. Phys. **55**, 04ED15 (2016).

[53] F. Liu, H. van Eersel, P. A. Bobbert, and R. Coehoorn, Three-dimensional Modeling of Bipolar Charge-Carrier Transport and Recombination in Disordered Organic Semiconductor Devices at low Voltages, Phys. Rev. Appl. **10**, 054007 (2018).

[54] J. J. M. van der Holst, M. A. Uijttewaal, B. Ramachandhran, R. Coehoorn, P. A. Bobbert, G. A. de Wijs, and R. A. de Groot, Modeling and analysis of the three-dimensional current density in sandwich-type single-carrier devices of disordered organic semiconductors, Phys. Rev. B **79**, 085203 (2009).

[55] M. Z. Szymański, B. Łuszczyńska, and D. Djurado, Modeling the transient space-charge-limited current response of organic semiconductor diodes using the master equation approach, IEEE J. Sel. Top. Quantum Electron. **19**, 7800107 (2013).

[56] J. Franco, V. Putcha, A. Vais, S. Sioncke, N. Waldron, D. Zhou, G. Rzepa, Ph. J. Roussel, G. Groeseneken, M. Heyns, N. Collaert, D. Linten, T. Grasser, and B. Kaczer, in *2017 IEEE International Electron Devices Meeting (IEDM),* (IEEE, San Francisco, CA, USA, 2017), p. 7.5.1.

[57] J. Franco, B. Kaczer, P. J. Roussel, J. Mitard, M. Cho, L. Witters, T. Grasser, and G. Groeseneken, SiGe channel technology: Superior reliability toward ultrathin EOT devices – Part I: NBTI, IEEE Trans. Electron Devices **60**, 396 (2013).

[58] N. Umezawa, K. Shiraishi, T. Ohno, H. Watanabe, T. Chikyow, K. Torii, K. Yamabe, K. Yamada, H. Kitajima, and T. Arikado, First-principles studies of the intrinsic effect of nitrogen atoms on reduction in gate leakage current through Hf-based high-k dielectrics, Appl. Phys. Lett. **86**, 143507 (2005).

[59] W. Goes, Y. Wimmer, A.-M. El-Sayed, G. Rzepa, M. Jech, A. L. Shluger, and T. Grasser, Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence, Microelectron. Reliab. **87**, 286 (2018).

[60] T. Grasser, B. Kaczer, W. Goes, Th. Aichinger, Ph. Hehenberger, and M. Nelhiebel, Understanding negative bias temperature instability in the context of hole trapping, Microelectron. Eng. **86**, 1876 (2009).

[61] P. C. McIntyre, Bulk and interfacial oxygen defects in HfO2 gate dielectric stacks: A critical assessment, ECS Trans. **11**, 235 (2007).

[62] K. Takagi and T. Ono, First-principles study on leakage current caused by oxygen vacancies at HfO_2/SiO_2/Si interface, Jpn. J. Appl. Phys. **57**, 066501 (2018).

[63] E. Cho, B. Lee, C.-K. Lee, S. Han, S. H. Jeon, B. H. Park, and Y.-S. Kim, Segregation of oxygen vacancy at metal-HfO_2 interfaces, Appl. Phys. Lett. **92**, 233118 (2008).

[64] K. Y. Tse, D. Liu, and J. Robertson, Electronic and atomic structure of metal-HfO_2 interfaces, Phys. Rev. B **81**, 035325 (2010).

[65] H. H. Pham and L.-W. Wang, Oxygen vacancy and hole conduction in amorphous TiO_2, Phys. Chem. Chem. Phys. **17**, 541 (2015).

[66] A. L. Shluger and K. P. McKenna, in *2013 IEEE International Reliability Physics Symposium (IRPS),* (IEEE, Anaheim, CA, USA, 2013), p. 5A.1.1.

[67] P. Broqvist, A. Alkauskas, J. Godet, and A. Pasquarello, First principles investigation of defect energy levels at semiconductor-oxide interfaces: Oxygen vacancies and hydrogen interstitials in the Si-SiO_2-HfO_2 stack, J. Appl. Phys. **105**, 061603 (2009).

[68] P. Broqvist and A. Pasquarello, Oxygen vacancy in monoclinic HfO_2: A consistent interpretation of trap assisted conduction, direct electron injection, and optical absorption experiments, Appl. Phys. Lett. **89**, 262904 (2006).

[69] E.-A. Choi and K. J. Chang, Charge-transition levels of oxygen vacancy as the origin of device instability in HfO_2 gate stacks through quasiparticle energy calculations, Appl. Phys. Lett. **94**, 122901 (2009).

[70] Y.-Y. Liu, F. Liu, R. Wang, J.-W. Luo, X. Jiang, R. Huang, S.-S. Li, and L.-W. Wang, Characterizing the Charge Trapping across Crystalline and Amorphous Si/SiO_2/HfO_2 Stacks from First-Principle Calculations, Phys. Rev. Appl. **12**, 064012 (2019).

[71] A. Massé, P. Friederich, F. Symalla, F. Liu, R. Nitsche, R. Coehoorn, W. Wenzel, and P. A. Bobbert, Ab initio charge-carrier mobility model for amorphous molecular semiconductors, Phys. Rev. B **93**, 195209 (2016).

[72] F. Liu, A. Massé, P. Friederich, F. Symalla, R. Nitsche, W. Wenzel, R. Coehoorn, and P. A. Bobbert, Ab initio modeling of steady-state and time-dependent charge transport in hole-only α-NPD devices, Appl. Phys. Lett. **109**, 243301 (2016).

[73] P. Kordt, J. J. M. van der Holst, M. Al Helwi, W. Kowalsky, F. May, A. Badinski, C. Lennartz, and D. Andrienko, Modeling of organic light emitting diodes: From molecular to device properties, Adv. Funct. Mater. **25**, 1955 (2015).

[74] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevApplied.13.024020 for more details of the model presented in this work.

[75] L. Vandelli, A. PadovanFi, L. Larcher, R. G. Southwick, W. B. Knowlton, and G. Bersuker, A physical model of the temperature dependence of the current through SiO_2/HfO_2 stacks, IEEE Trans. Electron Devices **58**, 2878 (2011).

[76] D. Heh, C. D. Young, G. A. Brown, P. Y. Hung, A. Diebold, E. M. Vogel, J. B. Bernstein, and G. Bersuker, Spatial distributions of trapping centers in HfO_2/SiO_2 gate stack, IEEE Trans. Electron Devices **54**, 1338 (2007).

[77] X.-F. Zheng, S. Fan, D. Kang, J.-K. Zhang, Y.-R. Cao, X.-H. Ma, and Y. Hao, Electron trap energy distribution in HfO_2 by the discharge-based pulse *I–V* technique, Chin. Phys. Lett. **31**, 127701 (2014).

[78] J. L. Gavartin, D. Muñoz Ramo, A. L. Shluger, G. Bersuker, and B. H. Lee, Negative oxygen vacancies in HfO_2 as charge traps in high-*k* stacks, Appl. Phys. Lett. **89**, 082908 (2006).

[79] S. Cimino, A. Padovani, L. Larcher, V. V. Afanas'ev, H. J. Hwang, Y. G. Lee, M. Jurczac, D. Wouters, B. H. Lee, H. Hwang, and L. Pantisano, A study of the leakage current in $TiN/HfO_2/TiN$ capacitors, Microelectron. Eng. **95**, 71 (2012).

[80] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices* (John Wiley & Sons, Hoboken, New Jersey, 2007), 3rd ed.

[81] W. J. Zhu, T.-P. Ma, T. Tamagawa, J. Kim, and Y. Di, Current transport in Metal/Hafnium Oxide/Silicon structure, IEEE Electron Device Lett. **23**, 97 (2002).

[82] E. Gharibshahi and E. Saion, Influence of dose on particle size and optical properties of colloidal Platinum nanoparticles, Int. J. Mol. Sci. **13**, 14723 (2012).

[83] C. L. Hinkle, C. Fulton, R. J. Nemanich, and G. Lucovsky, A novel approach for determining the effective tunneling mass of electrons in HfO_2 and other high-*K* alternative gate dielectrics for advanced CMOS devices, Microelectron. Eng. **72**, 257 (2004).

[84] B. Brar, G. D. Wilk, and A. C. Seabaugh, Direct extraction of the electron tunneling effective mass in ultrathin SiO_2, Appl. Phys. Lett. **69**, 2728 (1996).

[85] S. J. Lee, H. F. Luan, W. P. Bai, C. H. Lee, T. S. Jeon, Y. Senzaki, D. Roberts, and D. L. Kwong, in *International Electron Devices Meeting 2000*, Technical Digest (IEDM, San Francisco, CA, USA, 2000), p. 31.

[86] J. E. Medvedeva, Averaging of the electron effective mass in multicomponent transparent conducting oxides, EPL **78**, 57004 (2007).

[87] The optical dielectric constant is simply the square of the refractive index, obtained at https://refractiveindex.info.

[88] C. Svensson and I. Lundstroem, Trap-assisted charge injection in MNOS structures, J. Appl. Phys. **44**, 4657 (1973).

[89] M. Herrmann and A. Schenk, Field and high-temperature dependence of the long term charge loss in erasable programmable read only memories: Measurements and modeling, J. Appl. Phys. **77**, 4522 (1995).

[90] P. Long, J. Z. Huang, M. Povolotskyi, P. Sarangapani, G. A. Valencia-Zapata, T. Kubis, M. J. W. Rodwell, and G. Klimeck, Atomistic modeling trap-assisted tunneling in hole tunnel field effect transistors, J. Appl. Phys. **123**, 174504 (2018).

[91] A. Miller and E. Abrahams, Impurity conduction at low concentrations, Phys. Rev. **120**, 745 (1960).

[92] G. Blaise, Charge localization and transport in disordered dielectric materials, J. Electrostat. **50**, 69 (2001).

[93] A. Padovani, D. Z. Gao, A. L. Shluger, and L. Larcher, A microscopic mechanism of dielectric breakdown in SiO2 films: An insight from multi-scale modeling, J. Appl. Phys. **121**, 155101 (2017).

[94] Z. G. Yu, D. L. Smith, A. Saxena, R. L. Martin, and A. R. Bishop, Molecular Geometry Fluctuation Model for the Mobility of Conjugated Polymers, Phys. Rev. Lett. **84**, 721 (2000).

[95] J. Cottaar, PhD Thesis, Eindhoven University of Technology, (2012).

[96] R. Tsu and L. Esaki, Tunneling in a finite superlattice, Appl. Phys. Lett. **22**, 562 (1973).

[97] M. Kleefstra and G. C. Herman, Influence of the image force on the band gap in semiconductors and insulators, J. Appl. Phys. **51**, 4923 (1980).

[98] A. Asenov, Random dopant induced threshold voltage lowering and fluctuations in sub-0.1/spl mu/m MOSFETs: A 3-D "atomistic" simulation study, IEEE Trans. Electron Devices **45**, 2505 (1998).

[99] S. R. Bradley, G. Bersuker, and A. L. Shluger, Modelling of oxygen vacancy aggregates in monoclinic HfO_2: Can they contribute to conductive filament formation? J. Phys.: Condens. Matter **27**, 415401 (2015).

[100] C. Mannequin, P. Gonon, C. Vallée, L. Latu-Romain, A. Bsiesy, H. Grampeix, A. Salaün, and V. Jousseaume, Stress-induced leakage current and trap generation in HfO_2 thin films, J. Appl. Phys. **112**, 074103 (2012).