Comphy v3.0—A compact-physics framework for modeling charge trapping related reliability phenomena in MOS devices

Dominic Waldhoer [a,*], Christian Schleich [a,d], Jakob Michl [a,d], Alexander Grill [b], Dieter Claes [b], Alexander Karl [a], Theresia Knobloch [a], Gerhard Rzepa [c], Jacopo Franco [b], Ben Kaczer [b], Michael Waltl [a,d], Tibor Grasser [a,*]

[a] *Institute for Microelectronics, TU Wien, Austria*
[b] *imec, Leuven, Belgium*
[c] *Global TCAD Solutions, Wien, Austria*
[d] *Christian Doppler Laboratory for SDS at the Institute for Microelectronics, TU Wien, Austria*

ARTICLE INFO

ABSTRACT

Charge trapping plays an important role for the reliability of electronic devices and manifests itself in various phenomena like bias temperature instability (BTI), random telegraph noise (RTN), hysteresis or trap-assisted tunneling (TAT). In this work we present *Comphy v3.0*, an open source physical framework for modeling these effects in a unified fashion using nonradiative multiphonon theory on a one-dimensional device geometry. Here we give an overview about the underlying theory, discuss newly introduced features compared to the original *Comphy* framework and also review recent advances in reliability physics enabled by these new features. The usefulness of *Comphy v3.0* for the reliability community is highlighted by several practical examples including automatic extraction of defect distributions, modeling of TAT in high-κ capacitors and BTI/RTN modeling at cryogenic temperatures.

1. Introduction

Charge trapping at defects in the gate dielectric or at the oxide/semiconductor interface of a MOSFET device causes various reliability challenges in different applications. For instance, charge trapping due to temperature and/or bias stress induces a drift in the device characteristic, commonly referred to as bias temperature instability (BTI) [1–3], leading to a potential failure in both digital and analog circuits over the course of the device lifetime. In memory applications on the other hand, trap-assisted tunneling (TAT) [4] through oxide defects can limit the data retention times of the stored information [5]. In non-standard technologies, on the other hand, the defect densities are typically higher, making charge trapping already relevant during initial operation and can lead to a clearly visible hysteresis in the transfer characteristics [6,7].

While the aforementioned phenomena typically exhibit a pronounced temperature activation [8–10], which might lead to the conclusion that charge trapping becomes irrelevant at lower temperatures, it has been demonstrated that charge trapping processes also occur at cryogenic temperatures due to nuclear tunneling. One example is the prominence of positive BTI at 4 K observed in high-κ gate stacks [11]. Another manifestation of nuclear tunneling is the occurrence of random telegraph noise (RTN) [12] in control circuits for solid-state qubits, which causes a loss of quantum coherence [13,14].

Besides these detrimental effects, charge trapping can also be leveraged for certain applications, e.g. to store information in charge trap flash devices [15] or as physical unclonable function by providing a unique device-specific noise signature [16].

These examples emphasize the importance of understanding the microscopic nature of defects as well as the need to accurately describe charge trapping in devices. Modern TCAD simulation packages like Minimos-NT [17,18] or Sentaurus Device [19] include a multi-state nonradiative multiphonon (NMP) model [20] to incorporate the effects of charge trapping on the device performance and lifetime. However, such models are often too complex to be well calibrated with the available experimental data. Furthermore, an overly elaborate modeling approach can distract from the essential physics of charge trapping, which, in many relevant cases, can be well described within a simplified 2-state NMP model [21]. In this spirit, we recently introduced

---

* Corresponding authors.
  E-mail addresses: waldhoer@iue.tuwien.ac.at (D. Waldhoer), grasser@iue.tuwien.ac.at (T. Grasser).

the *Comphy* (compact-physics) framework [22], a light-weight Python package, which is designed to simulate various charge-trapping related reliability issues based on the 2-state NMP model within a compact 1D device simulator.

While *Comphy* in its original form was intended as a proof-of-concept to demonstrate the feasibility of accurate and efficient reliability modeling of core logic FETs, it is now frequently employed to assess and understand the degradation in experimental devices, ranging from novel memory applications [23] to power electronics [24] in silicon but also other semiconductors such as SiC. Although accurate reliability models are valuable in their own right for device lifetime predictions, the physics-based modeling approach used in *Comphy* also allows to draw conclusions about the atomistic nature of defects responsible for charge trapping by comparison of model parameters to theoretical predictions based on density functional theory (DFT) [25, 26]. Furthermore, a physical model can provide clues for possible design improvements. For example, based on *Comphy* predictions, a dipole layer introduced at the SiO_2/HfO_2 interface of high-κ gate stacks was demonstrated to improve BTI by suppressing the defect-carrier interaction [27,28].

In this paper we summarize our recent efforts to further extend the original *Comphy* framework towards a powerful and user-friendly reliability code, culminating in the publically available *Comphy v3.0* package [29].

2. Models and features

Comphy v1.0 was designed to provide a simple physical description of BTI within an easy-to-use compact model implemented in Python. However its capabilities were limited to simulating the BTI response of preset defect bands for a certain temperature and gate bias profile in a given gate-stack. While this is still a core use case of *Comphy v3.0*, several newly developed models and features have been added in order to meet the current demands for nanoscale device modeling. In particular, we added a new scheme for automatic defect parameter extraction from experimental data, a unified description of BTI and trap-assisted tunneling (TAT), as well as the capability of modeling charge trapping at cryogenic temperatures which is highly relevant for emerging quantum technologies. Furthermore, a lot of effort was undertaken to rewrite the code and optimize its run-time by using techniques like caching. In order to provide a code which can be easily extended by the community, we also redesigned the code with a modular class-based architecture. In the following we give a brief overview of these new features, a detailed description is then provided in later sections of this paper.

2.1. Parameter extraction methods

One of the main motivations behind a framework like *Comphy* is to obtain a physically-motivated degradation model based on experimental data points at accelerated stress conditions. A well-calibrated model can then be used to estimate the device lifetime by extrapolating the degradation to typical use conditions. As most gate oxides are amorphous thin films, they show a wide distribution of defect parameters due to the varying local chemical environment of the defects. Furthermore, most degradation experiments are conducted on large-area devices where only the collective response of a whole defect ensemble can be observed. Hence, the underlying defect parameter distributions have to be inferred from these observations. Previous studies mostly assumed Gaussian distributions for the parameters [22,30] and their mean and sigma values were obtained from a non-linear optimization procedure by fitting experimental data. However, such a procedure can lead to extraction artifacts because it enforces a certain shape of the distribution. Furthermore, the optimization requires a good initial guess, becomes more tedious and requires frequent manual intervention particularly when multiple defect bands are involved, e.g. interacting

electron and hole traps as observed in SiC/SiO_2 devices [24,31]. In order to circumvent these issues, *Comphy v3.0* offers a novel method of parameter extraction named *Effective Single Defect Decomposition* (ESiD) [32], which allows for a semi-automatic extraction of defect parameters from experimental measure-stress-measure traces without the aforementioned assumptions about their distribution and, in particular, without requiring a good initial guess.

2.2. Trap-Assisted Tunneling (TAT)

While the most pronounced effects of charge trapping at defects in the oxide are electrostatic shifts manifesting as BTI, RTN and hysteresis, the same mechanism can also facilitate a parasitic gate leakage current by conductance ("hopping") over defects [4]. Contrary to direct (DT) or Fowler–Nordheim (FN) tunneling [33], the resulting currents exhibit a strong temperature dependence, indicating a charge transfer mediated by a multiphonon process similar to charge trapping. Although several different models have been proposed in literature to describe this phenomenon [34–36], it is usually treated separately from charge trapping in the context of BTI. *Comphy v3.0* includes our recently developed unified approach [37] for TAT and BTI, where both are treated on the same footing within the NMP framework. This approach allows to obtain a consistent defect parameter set for the defect–channel interaction which is then transformed to a corresponding parameter set describing the defect–defect interaction.

2.3. Charge trapping at cryogenic temperatures

For applications at room temperature and above, the full quantum mechanical NMP model can be reasonably well approximated by its classical analog, the celebrated Marcus theory for charge transfer reactions [38]. This model is preferable from a computational point of view, since it only requires calculating the classical transition barrier instead of all vibrational overlaps in NMP theory. However, the classical model predicts a complete freeze-out of charge transfer towards cryogenic temperatures, whereas it is well known that charge trapping and its resulting effects on the device, i.e. BTI and RTN, can be observed at lower temperatures [11,39–41]. In this regime, the defect reconfiguration upon charge transfer is no longer temperature activated but rather dominated by *nuclear tunneling*. In order to efficiently model charge trapping under these conditions, e.g. for studying RTN in emerging quantum information applications, a Wentzel–Kramers–Brillouin (WKB) approximation to the full quantum mechanical model has been developed [42] and incorporated into *Comphy v3.0*. This approach remedies the computational overhead associated with the quantum mechanical description while remaining sufficiently accurate, hence allowing to simulate a whole ensemble of defects in the nuclear tunneling regime.

2.4. Coupling to TCAD

While *Comphy* provides a fast way to assess device reliability, an inherent limitation of the employed 1D geometry is related to the charge sheet approximation (CSA) which is used to model the impact of oxide charges on device electrostatics [43]. Even for ideal planar devices, this approximation fails to describe the distribution of threshold voltage shifts caused by the charges of individual traps, furthermore the deviations worsen in the subthreshold regime, where current percolation paths are formed, for example through random discrete dopants (RDD) and random trapped charges [44]. In the case of more complex device geometries such as FinFETs or gate-all-around (GAA) FETs, the modeling of transmission coefficients and the strongly inhomogeneous electric fields and carrier concentrations lead to further inaccuracies in the 1D approximation.

An efficient simulation method to go beyond this 1D approach and to accurately describe the variability-aware reliability of 3D structures

is to couple TCAD simulations with *Comphy*. In this approach, the 3D electrostatics are simulated with TCAD for all bias and temperature conditions which are relevant for a given stress scenario to obtain the quantities required by *Comphy* to compute the defect occupancies: The local lattice temperature, the electrostatic potential, the transmission coefficients to all charge reservoirs, as well as the carrier concentrations, energies, and effective masses of the semiconductor and the metal gate. Since these parameters can be passed to *Comphy* in the form of single values or generic value tables, this feature does not depend on the availability of a particular TCAD tool. In addition, the perturbative impedance field method (IFM) [45] can be applied on top of the quasistationary TCAD solutions to accurately model the impact of the defect charges on the device characteristics instead of using the CSA.

With this method, the efficient *Comphy* simulations can leverage full 3D TCAD accuracy as recently demonstrated for a reliability and variability aware design technology co-optimization (DTCO) study of FinFET and nanosheet devices, including BTI degradation of ringoscillators for 10 years of AC operation [46].

3. Theory

Due to the strong impact on electronic devices, charge trapping at oxide defects is the subject of numerous experimental and theoretical investigations to this day. While RTN is evidently caused by charge trapping [12,47], a connection between oxide defects and BTI has long been suspected [48,49] but has only recently been firmly established [50,51]. Early theoretical modeling attempts for BTI assumed a reaction–diffusion (RD) transport of hydroxyl groups [52] as underlying cause. Irrespective of the fact that hydrogen was ruled out as diffusing species in the pioneering paper [52], primarily because its diffusivity was known to be very high, the RD model was nonetheless later modified to be based on the diffusion of H_2 [53], dispersive H or H_2 transport [54] or reaction-dispersive proton transport [55,56]. However, more recent elaborate measurements show that such mechanisms are incompatible with the experimental evidence [57,58]. In addition, the parameters used in the reaction–diffusion model, in particular those related to the properties of hydrogen, are in contradiction to both independent experiments and theoretical expectations [3].

Instead, a reaction-limited process involving the phonon-mediated charge transfer to oxide defects is now widely accepted as the dominant cause of BTI [50]. Interestingly, this possibility was already suggested in the very first papers on BTI [48], including the pioneering paper on RD theory by Jeppson and Svensson [52]. This conclusion is mainly based on the pronounced bias- and temperature dependence observed in nanoscale devices [58] as well as the observation that BTI and RTN are two manifestations of the same microscopic mechanism at shorter time scales [59]. At longer time scales additional interface defects are created triggered by a gate-sided hydrogen release mechanism [51]. Furthermore, recent atomistic studies based on density functional theory of defects for SiO_2 [60] and HfO_2 [61], the most widely used gate dielectrics, are fully consistent with the parameters extracted from experimental data [60].

Since charge trapping is an essential part of reliability physics, the fundamentals of the underlying nonradiative multiphonon (NMP) theory, its limitations for reliability modeling as well as approximations necessary for its implementation in *Comphy* are discussed in the following.

3.1. Defect states

In order to model the dynamics of a defect in a device, details about the atomic configurations and the pathways between them are often too complicated to be dealt with explicitly. Following a certain transition, a defect typically relaxes into its new equilibrium configuration in a matter of picoseconds [62]. This implies that the properties of a defect and its possible future pathways only depend on the current
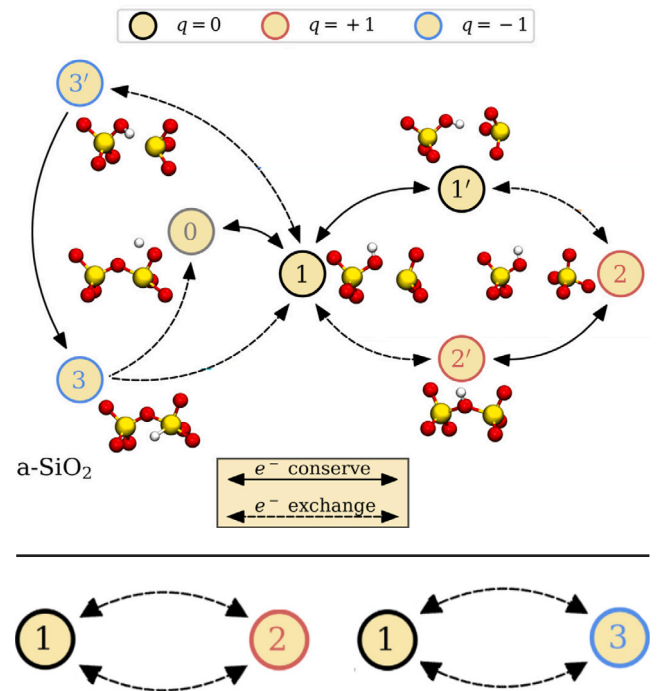

Fig. 1. Top: State diagram of the amphoteric hydroxyl E′ defect in SiO₂. Notice how a single amphotheric defect can capture either electrons or holes by going through its various states. Reprinted from [63]. **Bottom:** Simplified 2-state models for hole and electron trapping used in *Comphy*. Note that in this approximation electron and hole traps become independent.

defect state. In other words, the defect is treated as a memory-less system. Using this assumption, the defect states and the transition rates between different states form a Markov chain. All the defect physics, i.e. temperature and bias dependence, is encoded in the transition rates k_{ij} from an initial state i to a final state j. The probabilities P_j of finding a defect in a certain state at a given time evolve according to the Master equations

$$\frac{dP_j}{dt} = \sum_{i \neq j} k_{ij} P_i(t) - k_{ji} P_j(t) \quad \text{with} \quad \sum_i P_i(t) = 1. \tag{1}$$

This set of coupled linear differential equations can be solved analytically for an arbitrary number of interacting states using the matrix exponential approach.

Time-dependent defect spectroscopy (TDDS) [64] and RTN experiments have unambiguously shown that defects can not only undergo charge transitions, but also relaxations to meta-stable configurations within the same charge state. Prime examples for the existence of such meta-stable states are observations of anomalous RTN [65] and socalled switching traps [66], which show an unusual bias-dependence in their emission time constants, as explained in Section 3.2.1. The existence of these states was also established using paramagnetic resonance studies [66,67] and formed the basis for the development of the 4-state NMP model [20] currently implemented in commercial TCAD software like Minimos-NT or Sentaurus Device. While this model can give a comprehensive description of hole trapping in Si/SiO₂ devices during negative BTI (NBTI) stress, treatment of the less pronounced electron trapping during positive BTI (PBTI) was still lacking. As recently has been demonstrated [63], defects in amorphous SiO₂ are amphoteric, i.e. they can act as both hole and electron traps. The resulting 7-state diagram is depicted in Fig. 1 for the case of the hydroxyl E′ (HE) defect. Although the defect dynamics in such a model are quite complex, its behavior would still be described by (1). However, for a proper parameterization this model would require extensive experimental data

based on single-defect characterizations, which is exceedingly hard to obtain.

For this reason, *Comphy* condenses the defect dynamics into effective 2-state models for electron and hole traps separately, as shown in Fig. 1(bottom). Such a simplified model accurately captures the essential physics of charge trapping in most scenarios [21]. Furthermore, this model can be easily parameterized with only a few extended measure-stress-measure (eMSM) [68] sequences using the novel ESiD approach implemented in *Comphy v3.0*, as will be discussed in Section 5.3. Besides these advantages, (1) has a simple analytical solution

$$P_i(t) = P_i(\infty) + (P_i(0) - P_i(\infty)) \cdot \exp\left(-\frac{t}{\tau}\right) \qquad (2)$$

with the time constant $\tau = 1/(k_{ij} + k_{ji})$ and the equilibrium solution $P_i(\infty) = \tau k_{ji}$. This is exploited to efficiently evaluate the occupancies of defects analytically for arbitrary V_G and T profiles. Note, however, that this approach is only valid for non-interacting defects, an assumption that must be dropped in fully self-consistent calculations or to describe phenomena like multi-trap TAT, see Section 6.2 for more details.

3.2. Limitations of the 2-state defect model

While the effective 2-state model is well suited to reproduce the averaged response of a large ensemble of defects, it is worth discussing its limitations. The following subsections summarize the features introduced by these additional states that cannot be captured by the 2-state model. Whether any of these effects are relevant for a certain application can be checked by comparison to the full 4-state model available in TCAD solutions.

3.2.1. Switching traps

First, the transition between states 2 and 1′ as depicted in Fig. 1 leads to the change of the charge state observed in switching traps. Essentially, upon hole capture, the defect goes from the neutral state 1 into the positively charged state 2. This transition occurs via the metastable state 2′ which becomes clear when looking at atomic configurations of the defect states obtained from DFT. However, after becoming positively charged, many defects have a secondary trap level, typically inside the Si bandgap, which allows them to become neutralized while still in the secondary configuration 1′. Note that while a transition to 1′ electrically neutralizes the defect, it does not release the distortion at the defect site as 1′ is separated by a large barrier from the relaxed neutral configuration 1 [25]. Contrary to the time constants linking 1 and 2 over 2′, which are typically large, the time constants linking 2 and 1′ are typically short, meaning that these transitions are fast. This allows for a clear experimental distinction: without a transition to 1′, the defect would appear like a fixed positive charge which does not react to changes of the Fermi-level, at least up to the point where a transition back to state 1 is favored. On the other hand, with a transition to 1′ it appears like the transition from 1 to 2 has created an electrically active defect, a switching trap, which can quickly react to changes in the Fermi-level. As an example, this means that these defects can lead to a subthreshold slope change and show up as defects in charge pumping (CP) and capacitance–voltage (CV) measurements.

Another interesting consequence is the following: during BTI experiments the gate voltage is typically switched from a stress to a read-out or recovery voltage. For example in a pMOS transistor, if this recovery voltage results in a Fermi-level below the secondary trap level, all these defects will be positively charged and contribute to 100% to the threshold voltage shift. However, if the Fermi-level during read-out is positioned at the mean of the distribution of the secondary trap level, only 50% of these traps will be positively charged and thus visible in the threshold voltage shift. This strong read-out bias-dependence can be easily observed experimentally and originally triggered the extension of the 2-state model to a 3-state model (states 1, 2, and 1′), see [69].
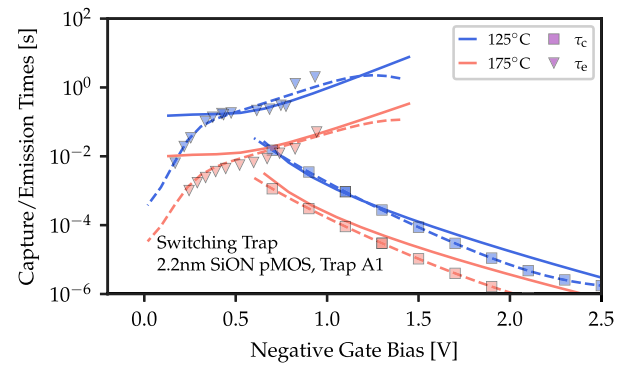

Fig. 2. 2-state **(solid)** and 4-state **(dashed)** model fits to the capture and emission times of a switching trap obtained from TDDS in [60]. The rapidly decreasing emission times of this switching trap below the threshold voltage ($V_G \approx -0.5\,\text{V}$) cannot be explained within the 2-state model.

Additionally, if the transition 2 to 1′ leads to a relaxation of the defect back to state 1 via state 1′, the emission time will be dramatically reduced, as shown in Fig. 2. One consequence of this transition is that such "positive charge" can be easily neutralized and the defects restored back to state 1 by switching the transistor into accumulation. Such an accumulation pulse can be used to experimentally extract the emission time constant under these bias conditions [70].

3.2.2. Metastable state 2′

Although it has long been known that transitions between states 1 to 2 occur via a metastable state 2′ [67], the necessity of including this transition into the model was originally deduced from the stronger temperature activation of the emission time compared to the capture time constant [71]. Later studies demonstrated that state 2′ also leads to a frequency dependence of the capture time constant since at higher frequencies back-transitions from 2′ to 1 slow down the overall charge capture [72].

3.2.3. Precursor activation

Another interesting property of oxide defects observed in TDDS experiments is their volatility: they can disappear without a trace just to randomly reappear weeks or months later with exactly the same properties [73]. This volatility was interpreted as a transition into the precursor state 0, e.g. by releasing hydrogen from the defect site. The hydrogen is then free to move around, potentially activating new defects nearby via a transition from 0 to 1, or even depassivating Si–H bonds at the interface. Note that contrary to the assumptions of the reaction–diffusion model, the diffusivity of H is very high [3]. Thus, H moves around very quickly, not limiting any subsequent reactions. An interesting consequence of the H released in the insulators is that it is available to trigger further reactions. In fact, it has long been known that the Si–H bond is too strong to easily release its H and that the reaction Si–H + H → Si-* + H$_2$ is the only feasible reaction [3,74–76]. Note that the depassivation could also proceed via a proton instead of neutral hydrogen [77]. Assuming that hydrogen is trapped at the gate side in H-E′ centers, they would mostly be positively charged (state 2). Upon application of a bias, they could be neutralized (transition to state 1) or release their H (transition to state 0). This idea led to the gate-sided hydrogen release model [51] and results in a coupling between charge trapping in the oxide and interface trap creation [78]. While providing a physical explanation, this model in its current stage is computationally demanding and suffers from numerical instabilities, hence in *Comphy* the creation of interface states is modeled via an empirical double-well barrier [22].
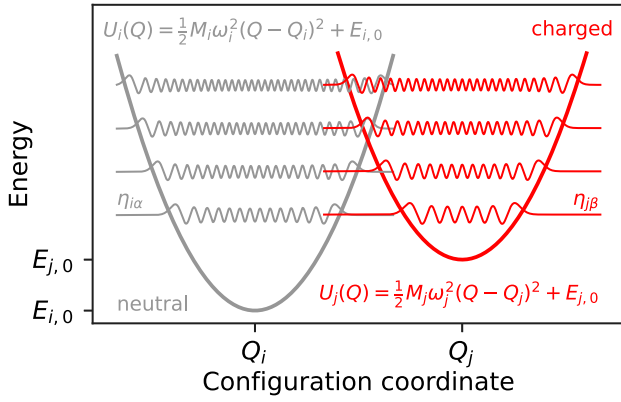
Fig. 3. Potential energy curves of a neutral and charged defect state within the harmonic approximation. The overlap integrals of the vibrational wavefunctions $\eta_{i\alpha}$ and $\eta_{j\beta}$ determine the nonradiative transition rates.
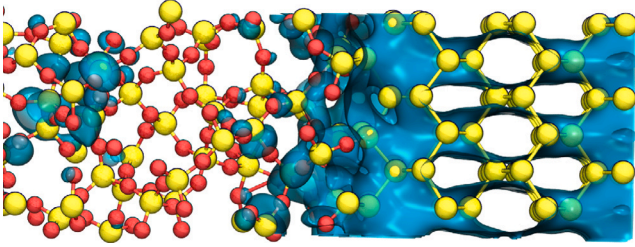

Fig. 4. DFT calculations of an atomistic Si/SiO_2 interface. **Left:** Localized defect state in the SiO_2. **Right:** Delocalized wavefunction in the silicon substrate penetrating into the oxide.

3.3. Nonradiative multiphonon transitions

While mathematically the defects are modeled as Markov chains, the transition rates k_{ij} containing the physical details within this model still have to be determined. As indicated in Fig. 1, there are two fundamentally different types of transitions. First, there are so-called thermal transitions (solid lines) between states with the same charge. In such cases, the transition rates are determined by a reaction barrier and can be evaluated within the classical transition state theory. Since these transitions are not treated explicitly in the simplified Comphy model, we will not discuss them further here and refer the interested reader to the literature [79]. The second type of transitions involves a charge transfer event typically accompanied by the emission or absorption of multiple phonons from the environment. These transitions are governed by the nonradiative multiphonon formalism [80,81] which will be outlined in the following. A detailed description can be found in [25].

During a charge transition, typically the atomic configuration of the defect changes due to structural relaxations. After decoupling the electronic and vibrational degrees of freedom by applying the Born–Oppenheimer approximation, the system is described by potential energy curves (PECs) $U_i(Q)$ and $U_j(Q)$ for both involved charge states along a so-called reaction coordinate Q as depicted in Fig. 3. Usually the PECs are assumed to be parabolic, i.e. the defect states are approximated as harmonic oscillators. Each PEC belongs to a different diabatic electronic state, e.g. the carrier is either delocalized in the semiconductor substrate or localized at the oxide defect. The corresponding electronic wavefunctions Ψ_i are shown in Fig. 4. On the other hand, the phonons involved in the charge transition are described by the vibrational wavefunctions $\eta_{i\alpha}$ and $\eta_{j\beta}$ arising from the PECs.

In order to obtain the total rate for a transition $i \rightarrow j$, all possible transitions between initial and final vibrational states have to be considered. Assuming a weak coupling strength between the defect

and the device substrate, the individual vibrational transition rates are given by Fermi's Golden Rule within time-dependent perturbation theory. By assuming a thermal equilibrium in the initial state and by invoking the Franck–Condon principle [82], the total transition rate can be expressed as [25]

$$k_{ij} = A_{ij} f_{ij} \tag{3}$$

with the electronic matrix element

$$A_{ij} = \frac{2\pi}{\hbar} |\langle \Psi_i | \hat{H}_{el} | \Psi_j \rangle|^2 \tag{4}$$

and the Franck–Condon lineshape function

$$f_{ij} = \frac{1}{Z} \sum_{\alpha,\beta} |\langle \eta_{i\alpha} | \eta_{j\beta} \rangle|^2 \exp\left(-\frac{E_{i\alpha}}{k_B T}\right) \delta(E_{i\alpha} - E_{j\beta}). \tag{5}$$

Here, $\hat{H}_{el}$ is the Hamiltonian of the electronic subsystem, Z is the canonical partition function of the initial state, $E_{i\alpha}$ and $E_{j\beta}$ are the vibrational eigenenergies of the initial and final states respectively, as shown in Fig. 3. While (3) describes the full quantum mechanical NMP transition rates, the usage in device simulations requires further approximations.

First, the electronic matrix element A_{ij} cannot be calculated within a classical device simulator, since the electronic wavefunctions are not available. Even when employing actual atomistic simulations [83] based on DFT, accurately obtaining A_{ij} remains highly challenging. However, due to the localized nature of the defect wavefunction and the exponentially decaying substrate wavefunction into the oxide, A_{ij} can be reasonably approximated by a classical capture cross section σ_0 from kinetic gas theory modified by a WKB based tunneling factor, yielding

$$A_{ij} \propto v_{th} \sigma_0 \vartheta_{WKB}, \tag{6}$$

with $v_{th} = \sqrt{3k_B T / \pi m^*}$ being the thermal velocity of carriers in the substrate. The tunneling factor ϑ_{WKB} is evaluated similarly to (20), as discussed in Section 6.1. All results shown in this work assumed a capture cross section of $10^{-15}\,\mathrm{cm}^2$.

Second, evaluating all vibrational overlap integrals in (5) is computationally very expensive and hence not suitable for an efficient model. Instead, we employ the classical limit ($T \rightarrow \infty$), in which the lineshape function is dominated by the classical barrier ε_{ij}, given by the crossing point of the two PECs, see also Fig. 5. In this case, the lineshape function can be approximated by

$$f_{ij} \approx \xi \exp\left(-\frac{\varepsilon_{ij}}{k_B T}\right), \tag{7}$$

with ξ being a prefactor depending on the exact shape of the PECs. However, this prefactor can usually be ignored due to the dominance of the Arrhenius-like temperature activation. While this approximation gives good results at room temperature and above [25], it is not suitable for cryogenic applications. In such cases, a different approach based on a WKB approximation of the vibrational wavefunctions is used, see Section 7 for details.

Although not explicitly stated, the quantities A_{ij} and f_{ij} both depend on the energy E of the carrier being exchanged between defect and substrate. Hence (3) has to be integrated over the whole conduction and valence band of the substrate semiconductor in order to obtain the full transition rates. However, this integration can only be done numerically and is hence associated with large computational costs. To circumvent this, we employ the so-called band-edge approximation [22], which assumes that all the available carriers are located exactly at the conduction or valence band edge. Under this assumption the classical transition rates can be expressed analytically and, in the case of an electron trap interacting with the semiconductor conduction band edge, are given by

$$k_{ij}^{cl} = n v_{th,n} \sigma_{0,n} \vartheta_n e^{-\beta(\varepsilon_{ij} - E_F + E_{CB})}$$

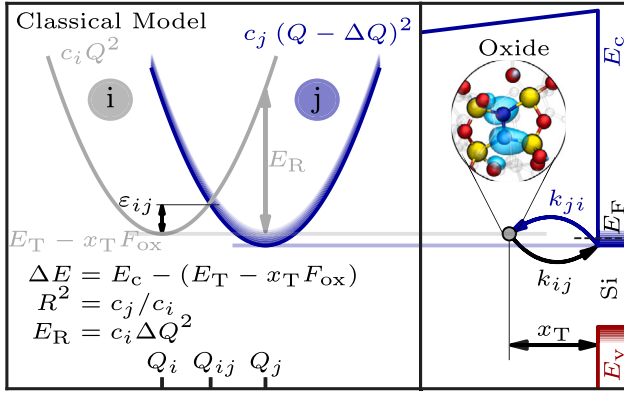$$k_{ji}^{cl} = n v_{th,n} \sigma_{0,n} \vartheta_n e^{-\beta \varepsilon_{ji}} \tag{8}$$

Fig. 5. Classical (high temperature) limit of the NMP model. Here, the transition rate is only determined by the classical barrier ε_{ij}. The PECs and their relative position are defined by the trap level E_T, the relaxation energy E_R and the curvature ratio R. Additionally, the position of the trap (x_T) together with the applied electric field F_{ox} across the oxide introduces a bias-dependent energy offset ΔE.

Table 1
Overview of the most important defect band parameters for Gaussian defect bands in *Comphy v1.0* and *v3.0*. Note that in *v1.0* all energies are referenced to the Si midgap, whereas in *v3.0* the vacuum level is the universal energy reference.

Parameter		v1.0	v3.0
Traplevel E_T	[eV]	Et_mean	Et
		Et_sigma	Et_sigma
Relaxation energy E_R	[eV]	S_mean	Er
		S_sigma	Er_sigma
Trap depth x_T	[m]	xt_min	xt_start
		xt_max	xt_end
Trap density N_T	[m^{-3}]	Nt	Nt

Note however, that for some special cases like the weak electron–phonon coupling regime, the correct value of the barrier ε_{ij} can deviate from the definition given in Fig. 5. Details about this intricacy can be found in the original *Comphy* paper [22], in particular Table 4.

3.4. Defect parameters

In order to simulate the threshold voltage shift ΔV_{th} caused by a (partially) charged defect ensemble, the PECs of the individual defects have to be parameterized. As depicted in Fig. 5, within the harmonic approximation the PECs of the two defect states can be described by

$$U_i(Q) = c_i Q^2 \quad \text{and} \quad U_j(Q) = c_j (Q - \Delta Q)^2 + \Delta E . \tag{9}$$

While the PECs can be directly parameterized by the curvatures $c_{i/j}$ at their respective minima, it is more common to use the *relaxation energy*

$$E_R = U_i(\Delta Q) - U_i(0) = c_i \Delta Q^2 \tag{10}$$

and the *curvature ratio* $R = \sqrt{c_j/c_i}$ instead. Note that the relaxation energy was denoted as S in previous works, however this was changed to E_R to avoid possible confusion with the Huang-Rhys factor, an overview of parameter names in *Comphy v1.0* and *v3.0* is given in Table 1. The energetic offset ΔE between the two parabolas is determined by the intrinsic trap level of the defect E_T as well as the applied electric field F_{ox} in the oxide and is given by

$$\Delta E = E_C - E_T - x_T F_{ox} \tag{11}$$

for a charge-free dielectric. Note that such a non-self-consistent treatment is only valid if the charges accumulating in the oxide defects cause only a small shift in the device electrostatics. Otherwise a self-consistent approach solving the Poisson equation including the oxide charges is necessary.
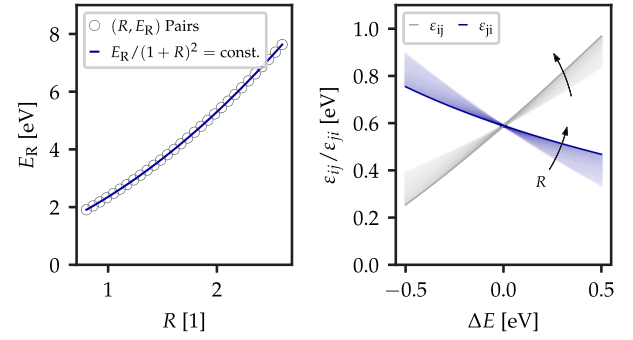

Fig. 6. Cross-correlation between the parameters R and E_R. When modified along the correlation curve (**left**), R and E_R can vary widely, however the resulting classical barriers remain rather similar (**right**).

For the defect PECs and their relative position with respect to each other to be uniquely defined, one has to specify the parameter tuple

$$\boldsymbol{p} = \left(E_T, E_R, R, x_T, \Delta Q \right) . \tag{12}$$

Note that in the classical limit the transition rates are only determined by the energetic barrier defined by the PEC crossing point. Hence, the displacement ΔQ along the configuration coordinate is immaterial in this case. However, in a quantum mechanical treatment, ΔQ is an important defect parameter determining the likelihood of nuclear tunneling.

Parameter correlations

While such a tuple $\boldsymbol{p}$ is sufficient to determine the transition rates of a defect and hence also its capture (τ_c) and emission times (τ_e), for the inverse problem, i.e. extracting defect parameters from a set of experimental ΔV_{th} curves, one has to check for cross-correlations between the model parameters. Framed differently, the question is, whether or not a *unique* set of defect parameters can be extracted from the available experimental data obtained with electrical characterization methods.

Considering only the classical limit of the NMP model, we have previously identified a non-linear cross-correlation between the parameters R and E_R [32]. The classical barrier can be expanded around $\Delta E = 0$, yielding

$$\varepsilon_{ij}(\Delta E) = \frac{E_R}{(1+R)^2} + \frac{R}{1+R}\Delta E + \mathcal{O}(\Delta E^2) . \tag{13}$$

As illustrated in Fig. 6, we observe that keeping the zeroth order term in this expansion fixed, R and E_R can vary over a wide parameter range resulting in very similar barriers. However, since the barrier is the defining feature for the electrically measurable ΔV_{th} response in the classical model, this implies that simultaneously extracting both R and E_R uniquely from experiments is challenging.

We suspect this (R, E_R) correlation to be responsible for the unphysically high relaxation energies of up to $8.0\,\text{eV}$ reported for oxide defects in previous works [22]. In order to obtain a unique, physical parameter set, we therefore recommend to restrict the simulations to the so-called *linear coupling regime* defined by $R = 1$ [79,84]. This choice is supported by extensive theoretical DFT studies on oxide defects in a-SiO$_2$ [25]. Besides this empirical justification, one could also make an argument based on the curvatures of the PECs being related to the phonon frequencies in the material. Since in an amorphous material many different phonon modes will be involved in a charge transition, the phonon frequency changes in different charge states will mostly average out, resulting in an effective curvature ratio of $R = 1$.

Although we demonstrated the (R, E_R) correlation in the 2-state NMP model, we cannot rigorously rule out the existence of other cross-correlations at the current time. For example, we suspect a similar correlation between ΔQ and E_R in the quantum mechanical model, however, this is subject to further investigations.
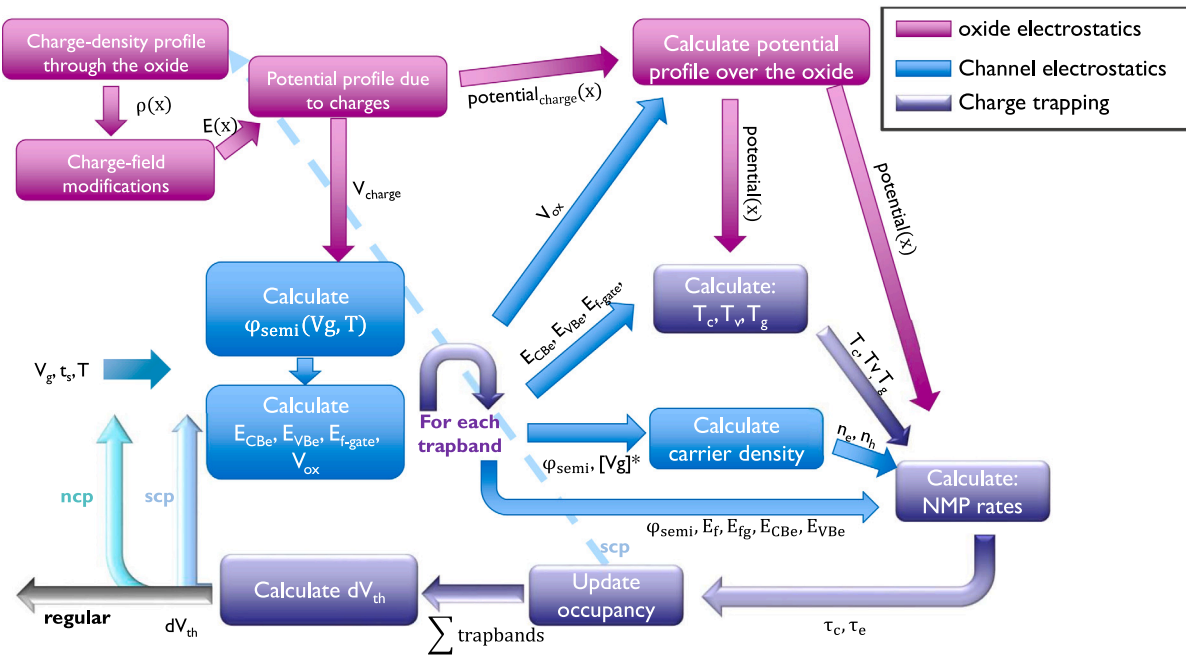
Fig. 7. A schematic representation of the internal structure of Comphy and the flow of information for calculating a ΔV_{th}-response (bottom left) starting from a set of input quantities (middle left). Boxes correspond to different stages in the calculation (each focused on obtaining a specific set of intermediary quantities), colors correspond to larger sections (stages grouped by a common denominator), arrows correspond to quantities being transferred from one stage to another. "Regular", "ncp" and "scp" refer to different ways of coupling charged traps to the electric field. "Regular" implies no coupling, "ncp" implies the charges of the current field will influence the next time step, "scp" implies the charges are coupled self-consistently and for every time step a convergent solution is acquired.

4. Code structure

Fig. 7 illustrates the internal logic of the simulation engine in *Comphy*. Each block represents a set of operations (functions) with the arrows illustrating the data propagation between the different blocks. This flow chart assumes all device parameters have already been loaded. The simulation starts at the left side where the gate bias V_G as a function of the time t at a certain temperature T is given as input quantity and is propagated through the framework, eventually producing a ΔV_{th} prediction (black arrow bottom left). Propagating through the framework, operations are subdivided into three major categories depending on whether the calculated quantity belongs to the channel (blue), the dielectric (purple) or the charge trapping at the defects (violet). Each block calculates one or more quantities using the provided input from earlier blocks up the chain. The flow is not static and depends on the features used/physics requested for a simulation. To improve the overall performance and to reduce redundant calculations, transparent caching is applied at various stages, as e.g. the electrostatic profiles may be requested multiple times by different blocks. In addition, certain quantities remain identical (given the same input) throughout the waveform, especially for cases where the coupling between the electric field profile and the trapped charges is neglected.

Elaborating on this coupling, the black, teal and cyan arrows in the bottom left corner of Fig. 3 represent the various available charge-field coupling schemes in *Comphy* ("regular", "ncp" and "scp" respectively). These allow the user various levels of speed-accuracy trade-offs by either not coupling the trapped charge and the electrostatics ("regular"), accounting for the trapped charge using an equivalent projected charge for the next timepoint (non-self consistent coupling, "ncp") or coupling the trapped charge with the oxide field and iteratively solving for a consistent solution (self-consistent coupling, "scp").

5. Parameter extraction methods

One of the main goals in device reliability physics is to accurately predict the performance and degradation of a device over its lifetime.
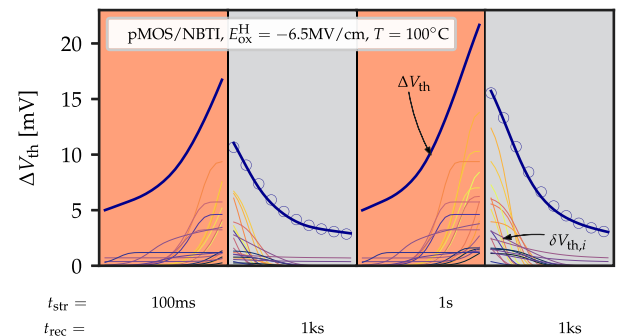

Fig. 8. Assuming non-interacting defects, the transient ΔV_{th} during an eMSM sequence can be linearly decomposed into small contributions $\delta V_{th,i}$ from individual defects. The red and gray regions indicate the stress and recovery phases of the eMSM sequence respectively.

Since the degradation processes happen on the timescale of years under normal operating conditions, accelerated stress conditions are used to observe the degradation in a reasonable time frame of hours to days. The main difficulty is to extrapolate the observations back to operating conditions in a physically meaningful way. Due to their simplicity, empirical power laws are frequently used in industry to assess the device reliability, however their predictive capabilities are limited due to the lack of a physical model. Instead, *Comphy* relies on a physical description of charge trapping based on the rigorous NMP framework. This approach does not only provide a more adequate extrapolation scheme, but also allows to compare the defect parameters of the simulation to theoretical DFT studies in order to identify possible defect candidates responsible for the observed degradation.

5.1. Fitting problem

Due to the amorphous nature of the gate dielectric, the defect parameters are usually distributed, resulting in capture and emission
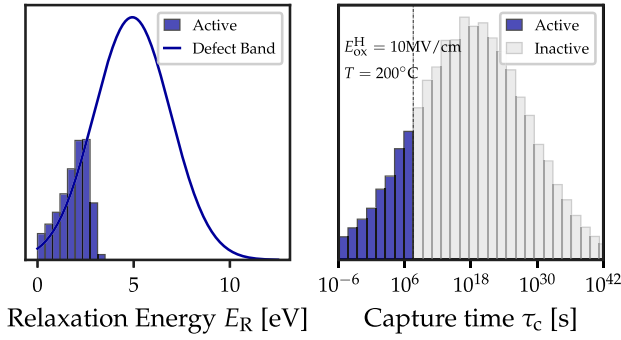
Fig. 9. Extracted Gaussian trap band for electron traps in SiO$_2$ [22] **(left)**. However, most defects in the band have exceedingly large capture time constants, rendering them electrically inactive. Only a small portion at the distribution tail can become charged within the measurement window **(right)**.

time constants spanning many orders of magnitude. Furthermore, most experiments on ΔV_{th} degradation are carried out on large-area devices, where only the collective response of a whole defect ensemble is observable. Typically, eMSM schemes [68] are used in experiments, resulting in recovery curves as depicted in Fig. 8. Using the 2-state NMP model, one can deduce from (2) that these recovery traces have to be composed of multiple exponentially decaying functions. Assuming non-interacting defects, the total ΔV_{th} can be expressed as a linear superposition of individual defect contributions, i.e.

$$\Delta V_{th}(t, V_G, T) = \int_{\Omega} N(\boldsymbol{p}) \cdot \delta V_{th}(t, V_G, T; \boldsymbol{p}) \mathrm{d}\boldsymbol{p}. \tag{14}$$

Here, $\boldsymbol{p}$ is a set of defect parameters similar to (12), δV_{th} is the contribution of a single defect with parameters $\boldsymbol{p}$ and $N(\boldsymbol{p})$ is the distribution function in the parameter space Ω.

Hence, extracting defect parameters from experiments is equivalent to finding a suitable distribution function $N(\boldsymbol{p})$ in order to fulfill (14). However, such an *inverse problem* [85] is mathematically ill-posed and requires to impose further restrictions on $N(\boldsymbol{p})$ to obtain a physical solution. Another relevant example of such a problem would be the *multiexponential analysis* required in deep level transient spectroscopy (DLTS) [86,87].

In the following, we discuss two different approaches implemented in Comphy to deal with this parameter extraction problem.

5.2. Gaussian defect bands

The main challenge in solving (14) is that the unknown $N(\boldsymbol{p})$ is a scalar field and hence has far too many degrees of freedom. A straightforward approach to obtain a solution is to enforce a certain distribution shape. The most common choice found in literature is a normal distribution for the energetic parameters E_T and E_R resulting in a *Gaussian defect band* [22]. The spatial variable x_T, on the other hand, is usually assumed to be uniform, leading to a constant defect density in a certain region of the oxide. Besides this simple model, an exponentially decaying spatial distribution is often chosen to reflect the increasing defect density towards the interface [88]. By fixing the mathematical form of the distribution, only its defining parameters, like the mean and standard deviation in the case of Gaussian defect bands, have to be optimized to fit the experimental ΔV_{th}. This can be achieved by using the simplex-method or other local optimizers.

The concept of Gaussian defect bands is intuitive and has been used successfully to reproduce experimental data across a large variety of different devices [22,23,89,90]. However, artificially constraining the shape of the parameter distribution might lead to some artifacts making it harder to provide physical interpretations and to link the obtained distributions to theoretical DFT studies. One example for such a misrepresentation is shown in Fig. 9 for the case of electron traps

in SiO$_2$ [22]. While the extracted defect band would suggest a mean relaxation energy of 5 eV, closer examination shows that even under severe stress conditions most defects within this band are electrically inactive due to their exceedingly large time constants. Instead, when filtering out too large capture times, only the tails of the Gaussian band contribute to the observed charge trapping, leading to a significantly altered parameter distribution. For this reason, it is advisable to check the corresponding time constants when using Gaussian bands and to truncate the distributions if appropriate. While such long time constants in Gaussian defect bands might seem to explain the more permanent component of BTI, these long time constants are typically associated with unrealistically large relaxation energies and thus cannot provide a physical explanation for "permanent" BTI. As discussed in Section 3.2.3, BTI at longer timescales is probably linked to the generation of new defects in the oxide and hydrogen kinetics instead of simply charging preexisting defects. With that being said, the degradation due to these additional mechanisms also follows a Arrhenius-like behavior and thus can be mathematically expressed in a similar manner as Eq. (8). This can be exploited to model permanent BTI with artificial oxide defects as discussed in more detail in [32].

Another drawback of Gaussian bands is that the optimization procedure often needs manual intervention and a good initial guess. Especially in SiC based devices, where multiple different defect bands are involved as a result of the increased stoichiometric complexity near the interface [31], the optimization becomes increasingly difficult and time consuming.

5.3. Effective Single Defect Decomposition (ESiD)

In order to eliminate the drawbacks of predefined distribution functions, we developed a novel extraction method named *Effective Single Defect Decomposition* (ESiD) [32] which is capable of inferring an unrestricted physical defect distribution from experimental eMSM sequences in a semiautomatic way.

5.3.1. Algorithm

Going back to the original problem (14), we directly exploit the fact that the macroscopic ΔV_{th} is a linear superposition of independent defect responses denoted by δV_{th}. While this assumption is strictly true only in the limit of low defect concentrations, it is still a good approximation for typically observed defect concentrations in the range of $10^{18} - 10^{20} \, \mathrm{cm}^{-3}$. Regardless, the quality of this approximation can always be checked retrospectively by using the resulting defect distributions in a self-consistent Poisson (scp) calculation and comparison to the non-selfconsistent results used for obtaining aforementioned distributions in the first place. Under this assumption, finding an optimal distribution function $N(\boldsymbol{p})$ can be recast into the non-negative linear least square (NNLS) problem

$$N(\boldsymbol{p}) = \arg \min_{\hat{N} \geq 0} \left| \Delta V_{th} - \int_{\Omega} \hat{N}(\boldsymbol{p}) \delta V_{th} \mathrm{d}\boldsymbol{p} \right|^2. \tag{15}$$

Note that the non-negativity constraint is essential here, since negative values for the defect density would constitute an unphysical solution. However, (15) is merely a reformulation of the original problem and hence is still ill-posed.

Naively trying to solve the optimization problem as stated above would lead to unstable solutions which are highly sensitive to noise in the input data. In order to obtain a physically meaningful solution, the problem has to be regularized. The implementation in Comphy uses the *Tikhonov* scheme [91], in which a regularization term depending on the total density

$$\hat{N}_{tot} = \int_{\Omega} \hat{N}(\boldsymbol{p}) \mathrm{d}\boldsymbol{p} \tag{16}$$

is introduced. In this scheme, the term to be minimized is given by

$$\left| \Delta V_{th} - \int_{\Omega} \hat{N}(\boldsymbol{p}) \delta V_{th} \mathrm{d}\boldsymbol{p} \right|^2 + \gamma^2 N_{tot}^2 \tag{17}$$
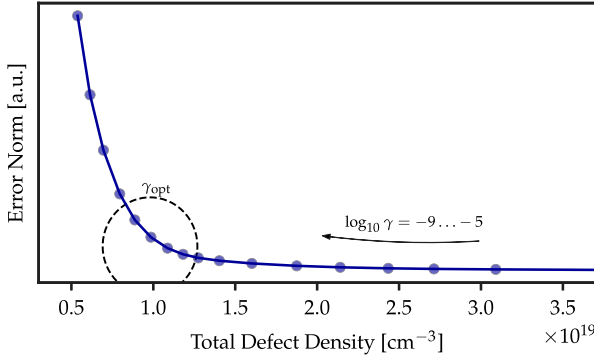
Fig. 10. Impact of the regularization term γ on the obtained solutions. The optimal value for γ is obtained at the "corner" of the L-shaped curve, before a further increase of γ leads to a steep increase in the error norm of the solution.
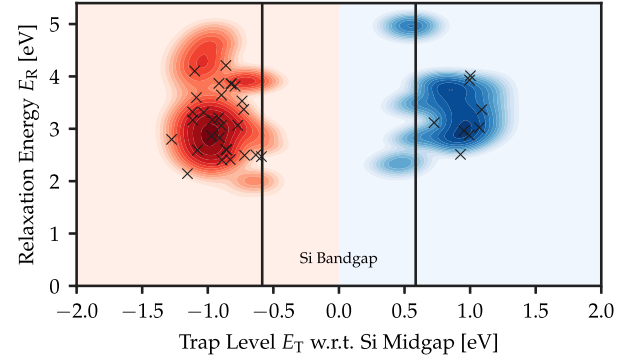

Fig. 11. (E_T, E_R) maps extracted from large-area SiON devices using ESiD [32] (heatmap). The distributions agree well with individual electron and hole traps (crosses) extracted with TDDS from similar small-area devices [59,93].
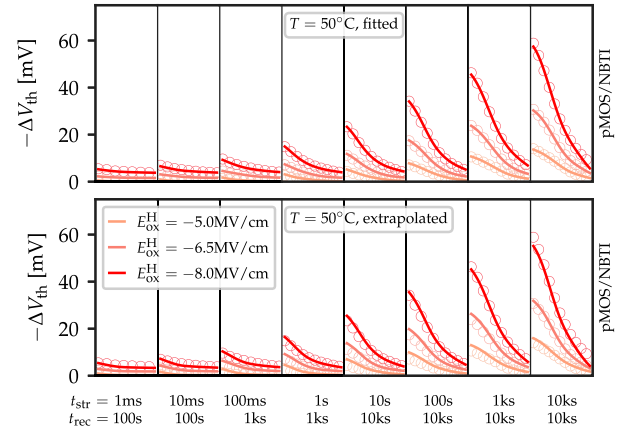

Fig. 12. Top: eMSM traces for SiON device [32] at T = 50 °C fitted with ESiD. **Bottom:** A different ESiD model was calibrated to experimental data obtained at T = 100 °C and T = 150 °C and subsequently extrapolated down to T = 50 °C. As can be seen, the prediction of this model for the ΔV_{th} degradation shows excellent agreement with the actual measurement data.

with γ being an adjustable regularization parameter. The effects of γ on the optimization are twofold. It favors solutions which require only a low defect density to match the experimental data and, although not obvious, also implicitly forces the solutions to be smooth. This can be shown by performing a singular-value decomposition (SVD) on the linear optimization problem and realizing that γ sets a lower bound for the smallest possible singular values, which are responsible for discontinuous solutions. In practice, γ has to be adjusted to the particular problem at hand. However, as illustrated in Fig. 10 a reasonable value can be estimated by plotting the total defect density versus the error norm to the experimental data for different values of γ, resulting in a L-shaped curve. As can be seen, for too small values, the approximation error is very low at the expense of a very high defect concentration. Quite to the contrary, if γ is too large, the problem becomes overregularized, meaning that there is a steep increase in approximation error when further increasing γ. According to the *L-criterion* [92], the optimal value for γ lies at the "corner" of this L-shaped curve, providing a good compromise between accurate representation of the experimental data with a reasonably small defect concentration.

5.3.2. Application

For an efficient implementation in *Comphy*, we only employ ESiD for the energetic parameters E_T and E_R since they are most relevant for identifying possible defect candidates. As mentioned earlier, we fix $R = 1$ in order to eliminate cross-correlations between the model parameters. The distributions along x_T and ΔQ are assumed to be uniform, meaning that the responses along these dimensions can be added up before optimization with ESiD. By doing so, we can reduce the parameter extraction to determining a 2-dimensional (E_T, E_R) distribution function. The NNLS optimization problem can then be easily discretized by defining a search region $[E_{T,min}, E_{T,max}] \times [E_{R,min}, E_{R,max}]$ on a uniform grid.

An example of such an extracted (E_T, E_R) map is shown in Fig. 11 for electron and hole traps in SiON dielectrics [32]. Note that this map extracted from large-area devices with eMSM agrees well with data obtained for individual defects by means of TDDS on similar small-area devices [59,93]. Although both methods give comparable results, the ESiD distribution was extracted automatically from measurement data obtained in a matter of days, whereas the single-defect characterization took several months. This clearly demonstrates the advantages provided by the ESiD extraction scheme. To further show the predictive capabilities of a device model created automatically by ESiD from eMSM data, Fig. 12 shows a comparison between a model which was actually fitted to traces at T = 50 °C (top) and a different model which was extracted from experimental data at T = 100 °C and 150 °C and subsequently extrapolated down to 50 °C (bottom). As can be seen,

the extrapolated degradation from the device model almost perfectly matches the actual measurement data. Note that in order to obtain a good model, the experimental data used for parameter extraction needs to include high temperature traces in order to also probe slower defects responsible for the long-term degradation at lower temperatures.

6. Gate-leakage currents

Besides charge-trapping which leads to considerable V_{th} shifts in MOSFETs, as described in the previous sections, time-zero gate-leakage currents pose a severe threat for power dissipation and to gate oxide reliability. Within *Comphy*, an efficient modeling approach is incorporated to enable the computation of these leakage currents that arise from the same inelastic charge transfer reactions driving BTI, but without additional parameters compared to the BTI simulations [37]. Depending on the gate stack material and defect properties within the insulating layers, either band-to-band tunneling or defects acting as charge transition centers between gate and channel dominate the leakage mechanism. As shown in Fig. 13, the trap-assisted tunneling (TAT) component appears as either a single- or multi-step process.

6.1. Tsu–Esaki model

The most commonly used approximation of the current density due to tunneling of charge carriers between two carrier reservoirs separated
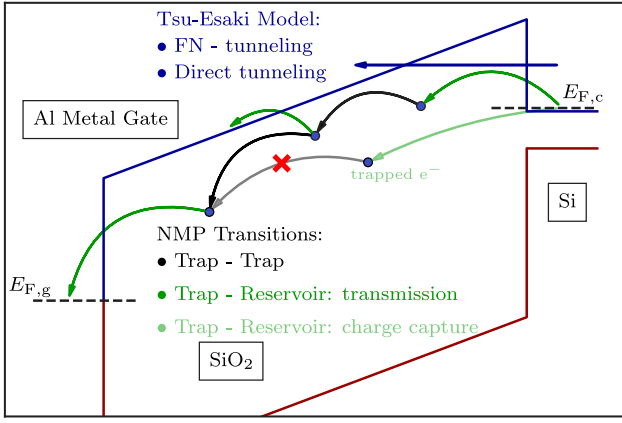
Fig. 13. The different tunneling mechanisms that typically appear in a MOS gate stack are shown schematically. The band to band tunneling mechanisms (green) are included in the well-known Tsu–Esaki model as implemented in Comphy, while all trap-assisted tunneling reactions (blue) are described by two-state NMP transitions.

by an energetic barrier is given by the Tsu–Esaki formalism [33]. For electron tunneling the current density is given as

$$J_{TE,e} = \frac{4\pi m_e q_0}{h^3} \int_{E_{CB}}^{\infty} \vartheta_{WKB}(E) N_e(E) \, dE \tag{18}$$

with the effective electron mass m_e in the semiconductor channel and the supply function

$$N_e(E) = k_B T \ln\left(\frac{1 + \exp\left(-\frac{E-E_{F1}}{k_B T}\right)}{1 + \exp\left(-\frac{E-E_{F2}}{k_B T}\right)} \right). \tag{19}$$

accounting for the Fermi–Dirac distribution of the carriers in both electrodes. The tunneling probability is calculated using a WKB approximation by

$$\vartheta_{WKB}(E) = \exp\left(-\frac{4\pi}{h} \int_{x_1}^{x_2} \sqrt{2m_{e,diel}(W(x) - E)} \, dx \right) \tag{20}$$

with the tunneling effective mass in the dielectric layer $m_{e,diel}$ and the energetic barrier $W(x)$. Depending on the shape of $W(x)$, i.e. triangular or trapezoidal, (18) evolves to the well-known Fowler–Nordheim or direct tunneling formulas.

6.2. Efficient trap-assisted tunneling current computation

Contrary to charge capture at oxide defects in large-area devices, the calculation of charge transfer between individual defects in amorphous materials is not feasible with a simulation on grids with equidistant grid-points. Therefore, for the calculation of the TAT current within an amorphous dielectric, the defects are randomly sampled based on a uniform distribution in a three dimensional volume V_i, as schematically shown in Fig. 14. The defect density N_T and average defect number $\overline{N}$ thereby define $V_i = \overline{N}/N_T$. The actual defect number N is drawn from a Poisson distribution within each simulated V_i. From the three dimensional sampling, the distances d_T between the defects can be calculated, while electrostatic quantities are computed in a one-dimensional simulation space, as described in [22]. Defects sampled with unphysically low distances to each other (below 1 nm) are removed from the sample and redrawn. In order to capture the stochastic variation of the simulation results, the simulation can be performed within M simulation volumes in parallel. Note that the ESiD extraction scheme is not applicable to obtain the model parameters from experiments due to the nonlinearity caused by the defect–defect interaction. Here, a Gaussian distribution is used for the sampling of the energetic defect parameters E_T and E_R which fully determine the

classically approximated PECs of the two charge states of a defect within the NMP model. This representation allows for the calculation of the energetic barriers ε_{ij} that determine the classical charge transition rates k. As described in detail in [37], considering all possible charge transfer paths, the trap-assisted leakage current can be computed based on a generalized Shockley-Ramo theorem [94] by

$$I_{G,TAT} = \underbrace{C_{ox} \frac{dV_G}{dt}}_{\text{displacement current}}$$

$$+ q_0 \underbrace{\sum_i^N k_{e,i,gate} f_i - k_{c,i,gate}(1 - f_i)}_{\text{single-TAT current}}$$

$$+ q_0 \underbrace{\sum_i^N \sum_{j \neq i}^N k_{e,ij} f_i (1 - f_j) \frac{x_i - x_j}{t_{ox}}}_{\text{multi-TAT current}}$$

$$+ q_0 \underbrace{\sum_i^N \left[k_{c,i,channel}(1 - f_i) - k_{e,i,channel} f_i \right] \left(1 - \frac{x_i}{t_{ox}}\right)}_{\text{charge trapping current (channel)}}. \tag{21}$$

Hereby the required NMP parameters to compute the charge transition rates k_{ij} for the multi-TAT current can be readily derived from the PECs for the carrier reservoir to defect charge transfer reactions. This parameter conversion is performed internally in *Comphy* and valid within the limitations presented in [37]. Hence no additional simulation parameters are required for including the multi-TAT current computation. Due to the coupling of the defects in multi-TAT mode, the arising non-linear coupled system of Master equations to compute the defect occupations needs to be solved. The Newton scheme used to calculate the occupations in multi-TAT mode leads to a reduced computational efficiency compared to single-TAT mode, whose performance is comparable to the simple ΔV_{th} calculation. In addition, the single-TAT mode is also accessible by ESiD. As shown in [37] for many applications a multi-step TAT current only becomes relevant at large N_T and for low relaxation energies E_R of the conducting defect bands. Due to these prerequisites, single-step TAT already provides efficient and accurate results for the majority of technologies.

6.3. Application example: Leakage currents in high-κ capacitor

Storage capacitors for RAM application fabricated from a thin ZrO$_2$ layer stacked between two TiN electrodes (TZT) enable further scaling of the memory cells. However, these capacitors show increased thermally activated leakage currents at low to medium electric field strengths [95,96]. By using the Comphy framework extension for leakage current computation as described in the previous section, the measured leakage currents are reproduced accurately for a capacitor employing a 8 nm thick ZrO$_2$ layer, as shown in Fig. 15. Therefore, two defect bands are required to explain the two branches with a shallow and steep current increase over the applied gate bias. These defect bands exhibit relatively large relaxation energies E_R of about 2.6 eV for the charge trapping band on the one hand and low E_R of about 0.8 eV for the defects responsible for the steady-state TAT current. These low E_R show excellent parameter agreement with those calculated with DFT for a statistically relevant number of model structures with polarons in partly recrystallized ZrO$_2$ [97].

7. Cryogenic charge trapping

The classical approximation of the 2-state NMP transition rates (8) is well suited for describing charge transfer kinetics above room temperature, however, it breaks down at cryogenic temperatures. While multiple studies show active charge trapping at cryogenic temperatures [11, 40,41,98] the classical limit

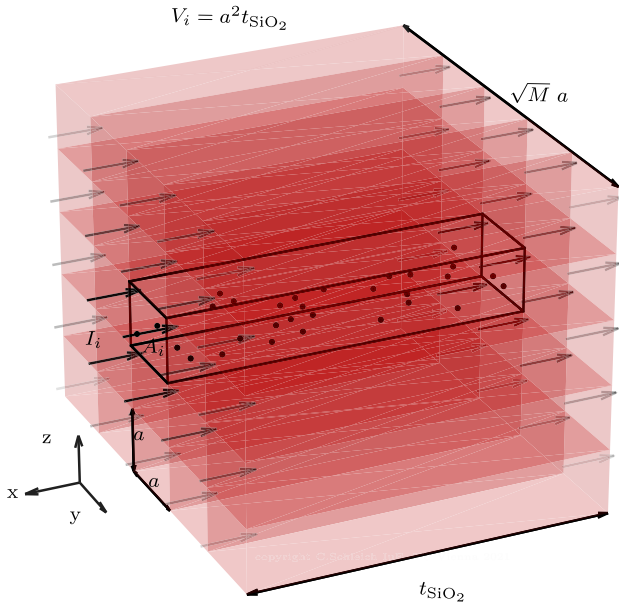$$\lim_{T \to 0K} k_{ij}^{cl}(T) = 0 \tag{22}$$

Fig. 14. The three dimensional spatial simulation domain is divided into M slabs in the dielectric layer. Within each slab a Poisson distributed number of defects N with mean value $\overline{N}$ is sampled throughout the volume $V_i = a^2 t_{\mathrm{SiO_2}}$, with a resulting from the defect density N_T.
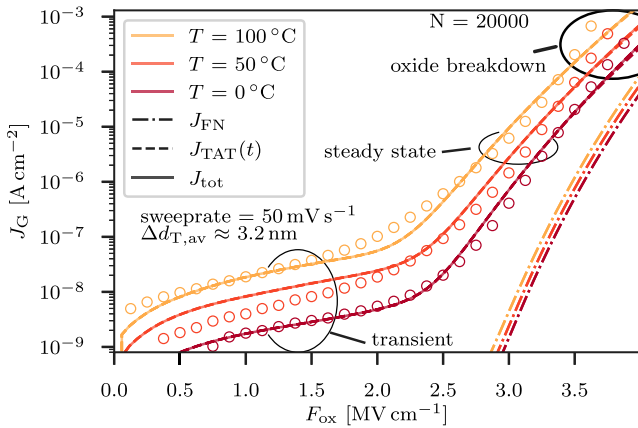

Fig. 16. The lineshape function ζ_{ij} for various configuration coordinate offsets ΔQ for $E_R = 2.5\,\mathrm{eV}$, $R = 1$ and $\Delta E = 0\,\mathrm{eV}$ computed with the full quantum mechanical model (solid), the WKB-based model (dashed) and the undistorted PEC model (dash-dotted) becomes temperature independent towards cryogenic temperatures while the classical model (dotted) freezes out completely.

Fig. 17. Experimental time constants observed in RTN signals at cryogenic temperatures together with fits of the classical and full quantum mechanical model [41]. As can be seen, the temperature independence of the time constants cannot be described within the classical model.

Fig. 15. Gate leakage currents measured (circles) on a TZT MIM capacitor show two branches of oxide field dependence. The simulation (lines) reveals that the shallow branch at low F_{ox} can be explained by a transient charge trapping current. The steep branch at increased F_{ox}, on the other hand, is explained by fast defect transitions between the electrodes via the oxide conduction band.

predicts a total freeze out of charge transitions between neutral and charged defect configurations. Thus, it is necessary to use the full quantum mechanical version of the 2-state NMP model for cryogenic modeling. Here, the transition rates at cryogenic temperatures are not dominated by the height of the classical barrier ε_{ij} but by the overlap of the vibrational wave functions as can be seen in Fig. 3. Towards cryogenic temperatures, the transition rate is dominated by the overlaps of the vibrational ground state, because the low thermal energy restricts access to excited states. As the most obvious consequence, the lineshape function which is proportional to the transition rate becomes temperature independent as shown in Fig. 16. This transition from one atomic configuration to another at energies below the classical barrier is called nuclear tunneling. This temperature independence is also demonstrated for defects causing RTN at cryogenic temperatures in Fig. 17. The computation of the full quantum mechanical transition
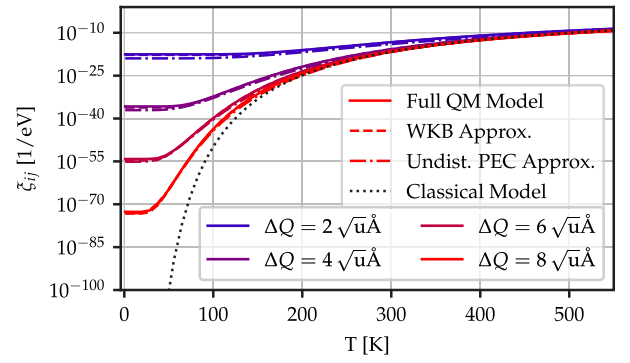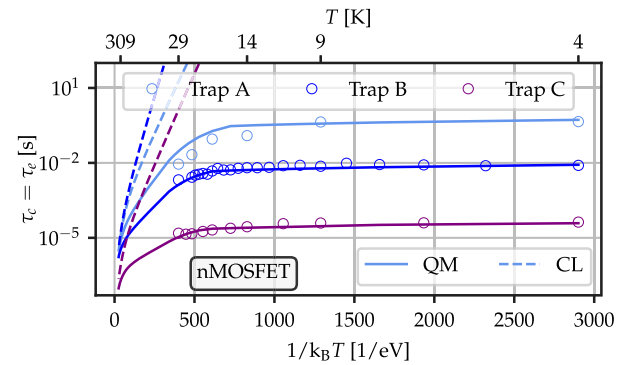
rate is numerically expensive, because it requires to compute the eigenstates of the potential energy surfaces and the corresponding overlap functions which then need to be summed up. This procedure is not suitable for reliability simulations with thousands of defects. Therefore, a numerically effective model based on a WKB approximation of the vibrational wave functions was developed in [42] and implemented in *Comphy v3.0*.

7.1. WKB-based approximation

Within the developed approximation scheme for the 2-state NMP model, the exact vibrational wave functions are replaced by their respective WKB-approximations as shown in Fig. 18.

At cryogenic temperatures, the overlaps of the wave functions which are needed for the computation of (1) are dominated by the exponentially decaying part of the WKB wave function. This allows to simplify the overlap integral (5) to a continuous form and approximate it analytically using the stationary phase method [99]. The transition rate k_{ij} then evaluates to

$$k_{ij}(T) = C(E^*) \exp\left(\frac{-E^*}{k_B T} + \varphi(E^*)\right) \sqrt{\frac{2\pi}{\varphi''(E^*)}}, \tag{23}$$

where $\varphi(E)$ can be expressed analytically as shown in [42]. The energy E^* can be obtained via the saddle-point method using a numerical optimizer for solving

$$\left.\frac{d\varphi(E)}{dE}\right|_{E=E^*} = \frac{1}{k_B T}. \tag{24}$$
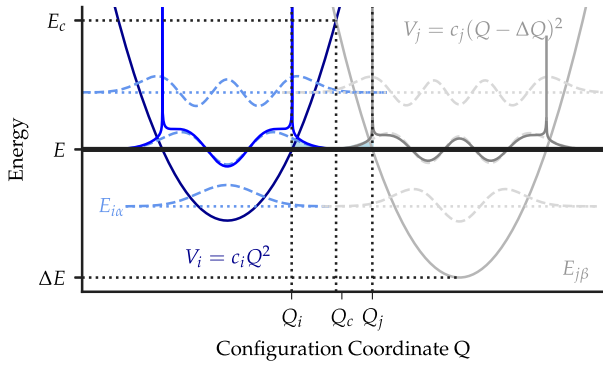
Fig. 18. The vibrational wave functions (dashed) of the potential energy curves representing the neutral and charged states can be approximated using WKB-based wave functions (solid). The overlap of the wavefunctions is dominated by their exponentially decaying parts which allows the formulation of a simplified expression for the charge transition rates.

Fig. 19. eMSM curves recorded between 4 K and room temperature on a 28 nm HKMG technology show that the threshold voltage shift ΔV_{th} of the pMOS device freezes out completely while the nMOS still shows a significant BTI shift.

This model gives very similar results to the full quantum mechanical (FQM) model while also being computationally much more efficient [42]. This is shown exemplary in Fig. 16, where the lineshape function of the WKB-based approximation and the (FQM) model are shown for $E_R = 2.5$ eV, $R = 1$ and $\Delta E = 0$ eV for a wide range of temperatures and configuration coordinate offsets ΔQ. This example shows the ability of describing the non-freeze-out behavior of the quantum mechanical transition rates compared to the classical model and the experimentally observed temperature independent behavior.

However, due to the required optimization step (24) no closed form expression is available which results in a considerably slower computation compared to the classical approximation. For further numerical improvement a model based on the assumption of linear coupling ($R = 1$) is developed.

7.2. Closed form for linear coupling

For $R = 1$, the curvatures of the PECs do not change upon charge trapping, implying that the effective phonon frequencies are also the same in both defect charge states ($\omega_i = \omega_j = \omega$). In this case, it is possible to develop a closed form expression for the quantum mechanical transition rates in the strong electron–phonon coupling regime at cryogenic temperatures [100]. Here, the transition rate k_{ij} can be computed as

$$k_{ij}(T) = \omega \frac{\sqrt{2\pi}}{2D\hbar} \left((1 + \coth\alpha)\, e^{-E_-/2D^2\hbar^2} - \right.$$
$$\left. (1 - \coth\alpha)\, e^{-E_+/2D^2\hbar^2} \right) \tag{25}$$

with $D = \sqrt{\omega E_R(2n+1)/\hbar}$, $n = 1/(\exp(\hbar\omega/k_B T) - 1)$, $\alpha = \hbar\omega/2k_B T$ and $E_\pm = \Delta E \pm \hbar\omega - E_R$. The phonon frequency ω can be obtained for a given configuration coordinate offset ΔQ from the relaxation energy by

$$\omega = \frac{\sqrt{2E_R}}{\Delta Q}. \tag{26}$$

As shown in the example in Fig. 16, the closed form expression for linear coupling gives the same results as the FQM model over a wide temperature range. Furthermore, computing this analytical expression is comparable to the classical model in terms of computational costs and is thus well suited for efficient reliability simulations with thousands of defects in a device.

7.3. Cryogenic modeling and ESiD

Using these cryogenic models for charge transitions in combination with the ESiD framework presented in Section 5.3, it is possible to
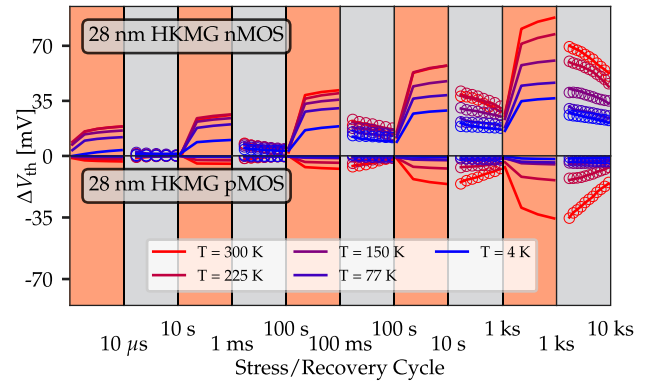
extract trapbands corresponding to recorded eMSM-measurements. As shown for a 28 nm-process of a high-κ metal-gate device, with a HfO$_2$ layer processed on a thin SiO$_2$ layer and dimensions $W \times L = 10\,\mu m \times 10\,\mu m$ in Fig. 19, the threshold voltage shift ΔV_{th} after a stress bias gets smaller towards cryogenic temperatures. However, while for pMOS devices ΔV_{th} freezes out completely below $T \approx 150$ K there is still a significant BTI shift on nMOS devices even at temperatures as low as $T = 4$ K.

Trapbands describing this asymmetry between nMOS and pMOS can be found using the ESiD-algorithm. The optimized trapbands describing the recorded ΔV_{th} curves are shown in Fig. 20. In line with earlier investigations based on CP experiments [101], the trap distribution extracted with ESiD shows large amounts of electron traps in the HfO$_2$ layer. As discussed in [11], the nMOS-pMOS asymmetry at cryogenic temperatures can be explained by these electron traps in the HfO$_2$ layer and fast electronic interface traps located close to the Si conduction band edge.

8. Outlook for future releases

While all these new features already provide a substantial enhancement compared to the original version, future works will also include compact models for emerging 2D devices, refined electrostatic models at cryogenic temperatures, coupling of charge trapping and ferroelectricity for FeFET applications. In order to add more flexibility, a numerical Poisson-solver will be included as well to account for incomplete ionization and movement of possible mobile ions in the oxide. Moreover, it will be possible to efficiently study the impact of charge trapping on CV characteristics. Furthermore, future releases will also include an interface to circuit simulators like PSpice to transfer the obtained reliability models for single devices to whole circuits, e.g. to validate switching timings across a circuit. Although this has been demonstrated already in a recent proof-of-concept work for the case of a simple CMOS inverter [102], further efforts are needed for a seamless link between *Comphy* and PSpice.

9. Conclusions

Device reliability becomes an ever increasing concern with the continued scaling into the atomic realm. Hence, considerable effort is put into developing predictive reliability models in order to assess the non-ideal behavior of nanoscale devices in simulations rather than costly and time-intensive experiments. Many prevalent reliability issues like bias temperature instability (BTI) or random telegraph noise (RTN) are dominated by charge trapping in oxides, demanding a solid understanding and a physical model for such charge transfer processes in devices. While the nonradiative multiphonon (NMP) theory provides a rigorous
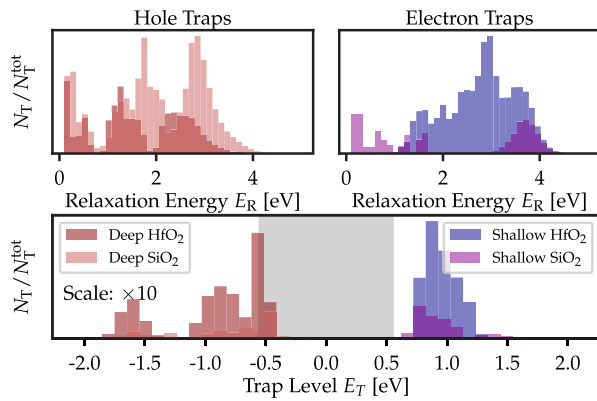
Fig. 20. Distributions of trap levels **(top)** and relaxation energies **(bottom)** for electron and hole traps as extracted with the ESiD scheme. The Si bandgap is indicated by the gray area.

framework for the treatment of charge trapping, its implementation in commercial TCAD software is often slow and cumbersome to calibrate accurately.

Here, we have summarized our recent efforts to provide a lightweight realization of NMP theory coupled to a 1D compact device model, resulting in the release of *Comphy v3.0*, an open source Python package for reliability simulations. While the original *Comphy v1.0* [22] was intended as a proof-of-concept for simple but accurate reliability models, this new release [29] provides a comprehensive framework to meet the current demands in industry and academic device research. Among the key features of *Comphy v3.0* are *(i)* an automated parameter extraction scheme which allows to extract defect parameters from simple extended measure-stress-measure (eMSM) experiments and to build predictive BTI models with ease; *(ii)* a new unified approach to treat trap-assisted tunneling and BTI on an equal footing within NMP theory; and *(iii)* a compact charge trapping model including nuclear tunneling at cryogenic temperatures to study RTN and BTI in emerging fields like quantum computing.

CRediT authorship contribution statement

Dominic Waldhoer: Conceptualization, Methodology, Software, Writing – original draft. **Christian Schleich:** Methodology, Software, Writing – review & editing. **Jakob Michl:** Methodology, Software, Writing – review & editing. **Alexander Grill:** Software, Validation. **Dieter Claes:** Software, Validation. **Alexander Karl:** Software, Validation. **Theresia Knobloch:** Writing – review & editing, Validation. **Gerhard Rzepa:** Conceptualization, Software. **Jacopo Franco:** Investigation, Data curation. **Ben Kaczer:** Project administration, Supervision, Conceptualization. **Michael Waltl:** Funding acquisition. **Tibor Grasser:** Project administration, Supervision, Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

The code is publicly available under https://comphy.eu.

References

[1] Y. Miura, Y. Matukura, Investigation of silicon-silicon dioxide interface using MOS structure, Japan. J. Appl. Phys. 5 (2) (1966) 180, http://dx.doi.org/10.1143/JJAP.5.180.

[2] D.K. Schroder, Negative bias temperature instability: What do we understand? Microelectron. Reliab. 47 (6) (2007) 841–852, http://dx.doi.org/10.1016/j.microrel.2006.10.006.

[3] J.H. Stathis, S. Mahapatra, T. Grasser, Controversial issues in negative bias temperature instability, Microelectron. Reliab. 81 (2018) 244–251, http://dx.doi.org/10.1016/j.microrel.2017.12.035.

[4] A.I. Chou, K. Lai, K. Kumar, P. Chowdhury, J.C. Lee, Modeling of stress-induced leakage current in ultrathin oxides with the trap-assisted tunneling mechanism, Appl. Phys. Lett. 70 (25) (1997) 3407–3409, http://dx.doi.org/10.1063/1.119186.

[5] K. Lee, H. Shin, Investigation of retention characteristics for trap-assisted tunneling mechanism in sub 20-nm NAND flash memory, IEEE Trans. Device Mater. Reliab. 17 (4) (2017) 758–762, http://dx.doi.org/10.1109/TDMR.2017.2772046.

[6] G. Rescher, G. Pobegen, T. Aichinger, T. Grasser, On the subthreshold drain current sweep hysteresis of 4H-SiC nMOSFETs, in: 2016 IEEE International Electron Devices Meeting, IEDM, IEEE, 2016, http://dx.doi.org/10.1109/iedm.2016.7838392.

[7] T. Knobloch, B. Uzlu, Y.Y. Illarionov, Z. Wang, M. Otto, L. Filipovic, M. Waltl, D. Neumaier, M.C. Lemme, T. Grasser, Improving stability in two-dimensional transistors with amorphous gate oxides by Fermi-level tuning, Nat. Electron. 5 (6) (2022) 356–366, http://dx.doi.org/10.1038/s41928-022-00768-0.

[8] T. Aichinger, M. Nelhiebel, T. Grasser, On the temperature dependence of NBTI recovery, Microelectron. Reliab. 48 (3) (2008) 1178–1184, http://dx.doi.org/10.1016/j.microrel.2008.06.018.

[9] T. Aichinger, M. Nelhiebel, T. Grasser, Unambiguous identification of the NBTI recovery mechanism using ultra-fast temperature changes, in: Proc. Intl. Reliability Physics Symposium, IRPS, 2009, pp. 2–7.

[10] M. Toledano-Luque, B. Kaczer, E. Simoen, P.J. Roussel, A. Veloso, T. Grasser, G. Groeseneken, Temperature and voltage dependences of the capture and emission times of individual traps in high-k dielectrics, Microelectron. Eng. 88 (7) (2011) 1243–1246, http://dx.doi.org/10.1016/j.mee.2011.03.097.

[11] J. Michl, A. Grill, D. Waldhör, W. Goes, B. Kaczer, D. Linten, B. Parvais, B. Govoreanu, I. Radu, T. Grasser, M. Waltl, Efficient modeling of charge trapping at cryogenic temperatures-part II: Experimental, IEEE Trans. Electron Devices 68 (12) (2021) 6372–6378, http://dx.doi.org/10.1109/TED.2021.3117740.

[12] K.S. Ralls, W.J. Skocpol, L.D. Jackel, R.E. Howard, L.A. Fetter, R.W. Epworth, D.M. Tennant, Discrete resistance switching in submicrometer silicon inversion layers: Individual interface traps and low-frequency (1/f?) noise, Phys. Rev. Lett. 52 (1984) 228–231, http://dx.doi.org/10.1103/PhysRevLett.52.228.

[13] A. Ayachi, W.B. Chouikha, S. Jaziri, R. Bennaceur, Telegraph noise effects on two charge qubits in double quantum dots, Phys. Rev. A 89 (2014) 012330, http://dx.doi.org/10.1103/PhysRevA.89.012330.

[14] R.E. Throckmorton, S. Das Sarma, Crosstalk- and charge-noise-induced multiqubit decoherence in exchange-coupled quantum dot spin qubit arrays, Phys. Rev. B 105 (2022) 245413, http://dx.doi.org/10.1103/PhysRevB.105.245413.

[15] H.-S. Choi, H. Kim, J.-H. Lee, B.-G. Park, Y. Kim, AND flash array based on charge trap flash for implementation of convolutional neural networks, IEEE Electron Device Lett. 41 (11) (2020) 1653–1656, http://dx.doi.org/10.1109/LED.2020.3025587.

[16] J. Chen, T. Tanamoto, H. Noguchi, Y. Mitani, Further investigations on traps stabilities in random telegraph signal noise and the application to a novel concept physical unclonable function (PUF) with robust reliabilities, in: 2015 Symposium on VLSI Technology, VLSI Technology, 2015, pp. T40–T41, http://dx.doi.org/10.1109/VLSIT.2015.7223695.

[17] IuE, MINIMOS-NT User's Guide, Institut für Mikroelektronik, Technische Universität Wien, Austria, 1998-2008, http://www.iue.tuwien.ac.at/software/minimos-nt.

[18] Global TCAD Solutions, Vienna, Austria 2022 MINIMOS-NT, URL https://www.globaltcad.com/minimos-nt.

[19] Synopsys, Mountain View, CA USA, 2022 Sentaurus Device, URL https://www.synopsys.com/silicon/tcad/device-simulation/sentaurus-device.html.

[20] T. Grasser, Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities, Microelectron. Reliab. 52 (1) (2012) 39–70, http://dx.doi.org/10.1016/j.microrel.2011.09.002, 2011 Reliability of Compound Semiconductors (ROCS) Workshop.

[21] G. Rzepa, J. Franco, A. Subirats, M. Jech, A. Chasin, A. Grill, M. Waltl, T. Knobloch, B. Stampfer, T. Chiarella, N. Horiguchi, L.A. Ragnarsson, D. Linten, B. Kaczer, T. Grasser, Efficient physical defect model applied to PBTI in high-k stacks, in: 2017 IEEE International Reliability Physics Symposium, IRPS, 2017, http://dx.doi.org/10.1109/IRPS.2017.7936425, XT–11.1–XT–11.6.

[22] G. Rzepa, J. Franco, B.O. Sullivan, A. Subirats, M. Simicic, G. Hellings, P. Weckx, M. Jech, T. Knobloch, M. Waltl, P. Roussel, D. Linten, B. Kaczer, T. Grasser, Comphy - A compact-physics framework for unified modeling of BTI, Microelectron. Reliab. 85 (2018) 49–65, http://dx.doi.org/10.1016/j.microrel.2018.04.002.

[23] B.J. O'Sullivan, V. Putcha, R. Izmailov, V. Afanas'ev, E. Simoen, T. Jung, Y. Higashi, R. Degraeve, B. Truijen, B. Kaczer, N. Ronchi, S. McMitchell, K. Banerjee, S. Clima, L. Breuil, G. Van den Bosch, D. Linten, J. Van Houdt, Defect profiling in FEFET Si:HfO2 layers, Appl. Phys. Lett. 117 (20) (2020) 203504, http://dx.doi.org/10.1063/5.0029072.

[24] C. Schleich, J. Berens, G. Rzepa, G. Pobegen, G. Rescher, S. Tyaginov, T. Grasser, M. Waltl, Physical modeling of bias temperature instabilities in SiC MOSFETs, in: 2019 IEEE International Electron Devices Meeting, IEDM, 2019, pp. 20.5.1–20.5.4, http://dx.doi.org/10.1109/IEDM19573.2019.8993446.

[25] W. Goes, Y. Wimmer, A.-M. El-Sayed, G. Rzepa, M. Jech, A. Shluger, T. Grasser, Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence, Microelectron. Reliab. 87 (2018) 286–320, http://dx.doi.org/10.1016/j.microrel.2017.12.021.

[26] D. Waldhoer, A.-M.B. El-Sayed, Y. Wimmer, M. Waltl, T. Grasser, Atomistic modeling of oxide defects, in: T. Grasser (Ed.), Noise in Nanoscale Semiconductor Devices, Springer International Publishing, Cham, 2020, pp. 609–648, http://dx.doi.org/10.1007/978-3-030-37500-3_18.

[27] J. Franco, B. Kaczer, P.J. Roussel, J. Mitard, S. Sioncke, L. Witters, H. Mertens, T. Grasser, G. Groeseneken, Understanding the suppressed charge trapping in relaxed- and strained-Ge/SiO$_2$/HfO$_2$ pMOSFETs and implications for the screening of alternative high-mobility substrate/dielectric CMOS gate stacks, in: 2013 IEEE International Electron Devices Meeting, IEEE, 2013, http://dx.doi.org/10.1109/iedm.2013.6724634.

[28] J. Franco, Z. Wu, G. Rzepa, A. Vandooren, H. Arimura, L.Å. Ragnarsson, G. Hellings, S. Brus, D. Cott, V. De Heyn, G. Groeseneken, N. Horiguchi, J. Ryckaert, N. Collaert, D. Linten, T. Grasser, B. Kaczer, BTI reliability improvement strategies in low thermal budget gate stacks for 3D sequential integration, in: 2018 IEEE International Electron Devices Meeting, IEDM, 2018, pp. 34.2.1–34.2.4, http://dx.doi.org/10.1109/IEDM.2018.8614559.

[29] Comphy v3.0, TU Wien, Austria, URL https://www.comphy.eu.

[30] G. Pobegen, T. Grasser, On the distribution of NBTI time constants on a long, temperature-accelerated time scale, IEEE Trans. Electron Devices 60 (7) (2013) 2148–2155, http://dx.doi.org/10.1109/TED.2013.2264816.

[31] V.V. Afanas'ev, M. Bassler, G. Pensl, M. Schulz, Intrinsic SiC/SiO$_2$ interface states, Phys. Status Solidi a 162 (1) (1997) 321–337, http://dx.doi.org/10.1002/1521-396X(199707)162:1<321::AID-PSSA321>3.0.CO;2-F.

[32] D. Waldhoer, C. Schleich, J. Michl, B. Stampfer, K. Tselios, E.G. Ioannidis, H. Enichlmair, M. Waltl, T. Grasser, Toward automated defect extraction from bias temperature instability measurements, IEEE Trans. Electron Devices 68 (8) (2021) 4057–4063, http://dx.doi.org/10.1109/TED.2021.3091966.

[33] R. Tsu, L. Esaki, Tunneling in a finite superlattice, Appl. Phys. Lett. 22 (11) (1973) 562–564, http://dx.doi.org/10.1063/1.1654509.

[34] M. Herrmann, 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 (9) (1995) 4522–4540, http://dx.doi.org/10.1063/1.359414.

[35] 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 (5) (2003) 1246–1253, http://dx.doi.org/10.1109/TED.2003.813236.

[36] M. Zhang, Z. Huo, Z. Yu, J. Liu, M. Liu, Unification of three multiphonon trap-assisted tunneling mechanisms, J. Appl. Phys. 110 (11) (2011) 114108, http://dx.doi.org/10.1063/1.3662195.

[37] C. Schleich, D. Waldhör, T. Knobloch, W. Zhou, B. Stampfer, J. Michl, M. Waltl, T. Grasser, Single- versus multi-step trap assisted tunneling currents—Part I: Theory, IEEE Trans. Electron Devices 69 (8) (2022) 4479–4485, http://dx.doi.org/10.1109/TED.2022.3185966.

[38] R.A. Marcus, Electron transfer reactions in chemistry. Theory and experiment, Rev. Modern Phys. 65 (1993) 599–610, http://dx.doi.org/10.1103/RevModPhys.65.599.

[39] J.H. Scofield, N. Borland, D.M. Fleetwood, Temperature-independent switching rates for a random telegraph signal in a silicon metal–oxide–semiconductor field-effect transistor at low temperatures, Appl. Phys. Lett. 76 (22) (2000) 3248–3250, http://dx.doi.org/10.1063/1.126596.

[40] T. Knobloch, J. Michl, D. Waldhör, Y. Illarionov, B. Stampfer, A. Grill, R. Zhou, P. Wu, M. Waltl, J. Appenzeller, T. Grasser, Analysis of single electron traps in nano-scaled MoS$_2$ FETs at cryogenic temperatures, in: Proceedings of the Device Research Conference, DRC, 2020, pp. 52–53.

[41] J. Michl, A. Grill, B. Stampfer, D. Waldhör, C. Schleich, T. Knobloch, E. Ioannidis, H. Enichlmair, R. Minixhofer, B. Kaczer, B. Parvais, B. Govoreanu, I. Radu, T. Grasser, M. Waltl, Evidence of tunneling driven random telegraph noise in Cryo-CMOS, in: 2021 IEEE International Electron Devices Meeting, IEDM, 2021, pp. 31.3.1–31.3.4, http://dx.doi.org/10.1109/IEDM19574.2021.9720501.

[42] J. Michl, A. Grill, D. Waldhör, W. Goes, B. Kaczer, D. Linten, B. Parvais, B. Govoreanu, I. Radu, M. Waltl, T. Grasser, Efficient modeling of charge trapping at cryogenic temperatures-Part I: Theory, IEEE Trans. Electron Devices 68 (12) (2021) 6365–6371, http://dx.doi.org/10.1109/TED.2021.3116931.

[43] O. Roux dit Buisson, G. Ghibaudo, J. Brini, Model for drain current RTS amplitude in small-area MOS transistors, Solid-State Electron. 35 (9) (1992) 1273–1276, http://dx.doi.org/10.1016/0038-1101(92)90161-5.

[44] B. Kaczer, S.M. Amoroso, R. Hussin, A. Asenov, J. Franco, P. Weckx, P.J. Roussel, G. Rzepa, T. Grasser, N. Horiguchi, On the distribution of the FET threshold voltage shifts due to individual charged gate oxide defects, in: Proc. Intl. Integrated Reliability Workshop, IIRW, 2016, pp. 18–20, http://dx.doi.org/10.1109/IIRW.2016.7904892.

[45] E. Karim, E. Lyumkis, A. Wettstein, Modeling statistical variability with the impedance field method, in: Proc. Simulation of Semiconductor Processes and Devices, SISPAD, 2012, pp. 205–208.

[46] G. Rzepa, M. Karner, O. Baumgartner, G. Strof, F. Schanovsky, F. Mitterbauer, C. Kernstock, H.-W. Karner, P. Weckx, G. Hellings, et al., Reliability and variability-aware DTCO flow: Demonstration of projections to N3 FinFET and nanosheet technologies, in: Proc. Intl. Reliability Physics Symposium, IRPS, 2021, pp. 1–6, http://dx.doi.org/10.1109/IRPS46558.2021.9405172.

[47] D. Fleetwood, H. Xiong, Z.-Y. Lu, C. Nicklaw, J. Felix, R. Schrimpf, S. Pantelides, Unified model of hole trapping, 1/f noise, and thermally stimulated current in MOS devices, IEEE Trans. Nucl. Sci. 49 (6) (2002) 2674–2683, http://dx.doi.org/10.1109/TNS.2002.805407.

[48] D.J. Breed, A new model for the negative voltage instability in MOS devices, Appl. Phys. Lett. 26 (3) (1975) 116–118, http://dx.doi.org/10.1063/1.88085.

[49] V. Huard, M. Denais, C. Parthasarathy, NBTI degradation: From physical mechanisms to modelling, Microelectron. Reliab. 46 (1) (2006) 1–23, http://dx.doi.org/10.1016/j.microrel.2005.02.001.

[50] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, M.T. Toledano Luque, M. Nelhiebel, The paradigm shift in understanding the bias temperature instability: From reaction–diffusion to switching oxide traps, IEEE Trans. Electron Devices 58 (11) (2011) 3652–3666, http://dx.doi.org/10.1109/TED.2011.2164543.

[51] T. Grasser, M. Waltl, Y. Wimmer, W. Goes, R. Kosik, G. Rzepa, H. Reisinger, G. Pobegen, A. El-Sayed, A. Shluger, B. Kaczer, Gate-sided hydrogen release as the origin of "permanent" NBTI degradation: From single defects to lifetimes, in: Proc. Intl. Electron Devices Meeting, IEDM, 2015, http://dx.doi.org/10.1109/iedm.2015.7409739.

[52] K.O. Jeppson, C.M. Svensson, Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices, J. Appl. Phys. 48 (5) (1977) 2004–2014, http://dx.doi.org/10.1063/1.323909.

[53] M.A. Alam, A critical examination of the mechanics of dynamic NBTI for pMOSFETs, in: IEEE International Electron Devices Meeting 2003, IEEE, 2003, pp. 4.14.1–4.14.4, http://dx.doi.org/10.1109/IEDM.2003.1269295.

[54] T. Grasser, W. Gos, B. Kaczer, Dispersive transport and negative bias temperature instability: Boundary conditions, initial conditions, and transport models, IEEE Trans. Device Mater. Reliab. 8 (1) (2008) 79–97.

[55] S. Chakravarthi, A. Krishnan, V. Reddy, C. Machala, S. Krishnan, A comprehensive framework for predictive modeling of negative bias temperature instability, in: 2004 IEEE International Reliability Physics Symposium. Proceedings, 2004, pp. 273–282, http://dx.doi.org/10.1109/RELPHY.2004.1315337.

[56] M. Houssa, M. Aoulaiche, S. De Gendt, G. Groeseneken, M.M. Heyns, A. Stesmans, Reaction-dispersive proton transport model for negative bias temperature instabilities, Appl. Phys. Lett. 86 (9) (2005) 093506, http://dx.doi.org/10.1063/1.1871357.

[57] F. Schanovsky, T. Grasser, On the microscopic limit of the RD model, in: T. Grasser (Ed.), Bias Temperature Instability for Devices and Circuits, Springer New York, New York, NY, 2014, pp. 379–408, http://dx.doi.org/10.1007/978-1-4614-7909-3_15.

[58] T. Grasser, K. Rott, H. Reisinger, M. Waltl, F. Schanovsky, B. Kaczer, NBTI in nanoscale MOSFETs – The ultimate modeling benchmark, IEEE Trans. Electron Devices 61 (11) (2014) 3586–3593, http://dx.doi.org/10.1109/TED.2014.2353578.

[59] T. Grasser, K. Rott, H. Reisinger, M. Waltl, J. Franco, B. Kaczer, A unified perspective of RTN and BTI, in: 2014 IEEE International Reliability Physics Symposium, 2014, pp. 4A.5.1–4A.5.7, http://dx.doi.org/10.1109/IRPS.2014.6860643.

[60] T. Grasser, W. Goes, Y. Wimmer, F. Schanovsky, G. Rzepa, M. Waltl, K. Rott, H. Reisinger, V. Afanas'ev, A. Stesmans, A.-M. El-Sayed, A. Shluger, On the microscopic structure of hole traps in pMOSFETs, in: 2014 IEEE International Electron Devices Meeting, 2014, pp. 21.1.1–21.1.4, http://dx.doi.org/10.1109/IEDM.2014.7047093.

[61] M.N.K. Alam, S. Clima, B.J. O'Sullivan, B. Kaczer, G. Pourtois, M. Heyns, J. Van Houdt, First principles investigation of charge transition levels in monoclinic, orthorhombic, tetragonal, and cubic crystallographic phases of HfO$_2$, J. Appl. Phys. 129 (8) (2021) 084102, http://dx.doi.org/10.1063/5.0033957.

[62] F. Schanovsky, Atomistic Modeling in the Context of the Bias Temperature Instability (Ph.D. thesis), TU Wien, 2013, http://dx.doi.org/10.34726/hss.2013.28781.

[63] C. Wilhelmer, D. Waldhoer, M. Jech, A.-M.B. El-Sayed, L. Cvitkovich, M. Waltl, T. Grasser, Ab initio investigations in amorphous silicon dioxide: Proposing a multi-state defect model for electron and hole capture, Microelectron. Reliab. 139 (2022) 114801, http://dx.doi.org/10.1016/j.microrel.2022.114801.

[64] T. Grasser, H. Reisinger, P.-J. Wagner, B. Kaczer, Time-dependent defect spectroscopy for characterization of border traps in metal-oxide-semiconductor transistors, Phys. Rev. B 82 (2010) 245318, http://dx.doi.org/10.1103/PhysRevB.82.245318.

[65] M. Kirton, M. Uren, Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency (1/f) noise, Adv. Phys. 38 (4) (1989) 367–468, http://dx.doi.org/10.1080/00018738900101122.

[66] A. Lelis, T. Oldham, Time dependence of switching oxide traps, IEEE Trans. Nucl. Sci. 41 (6) (1994) 1835–1843, http://dx.doi.org/10.1109/23.340515.

[67] E. Poindexter, W. Warren, Paramagnetic point defects in amorphous thin films of SiO_2 and Si_3N_4: Updates and additions, J. Electrochem. Soc. 142 (7) (1995) 2508–2516, http://dx.doi.org/10.1149/1.2044326.

[68] B. Kaczer, T. Grasser, J. Roussel, J. Martin-Martinez, R. O'Connor, B.J. O'Sullivan, G. Groeseneken, Ubiquitous relaxation in BTI stressing—New evaluation and insights, in: 2008 IEEE International Reliability Physics Symposium, 2008, pp. 20–27, http://dx.doi.org/10.1109/RELPHY.2008.4558858.

[69] T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, M. Nelhiebel, A two-stage model for negative bias temperature instability, in: 2009 IEEE International Reliability Physics Symposium, 2009, pp. 33–44, http://dx.doi.org/10.1109/IRPS.2009.5173221.

[70] T. Grasser, K. Rott, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, M. Waltl, M. Toledano-Luque, B. Kaczer, Advanced characterization of oxide traps: The dynamic time-dependent defect spectroscopy, in: Proc. Intl. Reliability Physics Symposium, IRPS, 2013, pp. 2D.2.1–2D.2.7, http://dx.doi.org/10.1109/IRPS.2013.6531957.

[71] T. Grasser, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, B. Kaczer, The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability, in: Proc. Intl. Reliability Physics Symposium, IRPS, 2010, pp. 16–25, http://dx.doi.org/10.1109/IRPS.2010.5488859.

[72] T. Grasser, H. Reisinger, K. Rott, M. Toledano-Luque, B. Kaczer, On the microscopic origin of the frequency dependence of hole trapping in pMOSFETs, in: Proc. Intl. Electron Devices Meeting, IEDM, 2012, pp. 19.6.1–19.6.4, http://dx.doi.org/10.1109/IEDM.2012.6479076.

[73] T. Grasser, K. Rott, H. Reisinger, M. Waltl, P. Wagner, F. Schanovsky, W. Goes, G. Pobegen, B. Kaczer, Hydrogen-related volatile defects as the possible cause for the recoverable component of NBTI, in: Proc. Intl. Electron Devices Meeting, IEDM, 2013, http://dx.doi.org/10.1109/IEDM.2013.6724637.

[74] E. Cartier, J.H. Stathis, D.A. Buchanan, Passivation and depassivation of silicon dangling bonds at the Si/SiO_2 interface by atomic hydrogen, Appl. Phys. Lett. 63 (11) (1993) 1510–1512, http://dx.doi.org/10.1063/1.110758.

[75] J.H. Stathis, E. Cartier, Atomic hydrogen reactions with P_b centers at the (100) Si/SiO_2 interface, Phys. Rev. Lett. 72 (1994) 2745–2748, http://dx.doi.org/10.1103/PhysRevLett.72.2745.

[76] M. Houssa, V. Afanas'ev, A. Stesmans, M. Aoulaiche, G. Groeseneken, M. Heyns, Insights on the physical mechanism behind negative bias temperature instabilities, Appl. Phys. Lett. 90 (4) (2007) 043505, http://dx.doi.org/10.1063/1.2434176.

[77] S.N. Rashkeev, D.M. Fleetwood, R.D. Schrimpf, S.T. Pantelides, Defect generation by hydrogen at the Si- SiO_2 interface, Phys. Rev. Lett. 87 (2001) 165506, http://dx.doi.org/10.1103/PhysRevLett.87.165506.

[78] T. Grasser, B. Kaczer, Evidence that two tightly coupled mechanism are responsible for negative bias temperature instability in oxynitride MOSFETs, IEEE Trans. Electron Devices 56 (5) (2009) 1056–1062, http://dx.doi.org/10.1109/TED.2009.2015160.

[79] A. Nitzan, Chemical Dynamics in Condensed Phases: Relaxation, Transfer and Reactions in Condensed Molecular Systems, in: Oxford Graduate Texts, OUP Oxford, 2006, http://dx.doi.org/10.1093/oso/9780198529798.001.0001.

[80] K. Huang, A. Rhys, N.F. Mott, Theory of light absorption and non-radiative transitions in F-centres, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 204 (1078) (1950) 406–423, http://dx.doi.org/10.1098/rspa.1950.0184.

[81] C.H. Henry, D.V. Lang, Nonradiative capture and recombination by multiphonon emission in GaAs and GaP, Phys. Rev. B 15 (1977) 989–1016, http://dx.doi.org/10.1103/PhysRevB.15.989.

[82] E. Condon, A theory of intensity distribution in band systems, Phys. Rev. 28 (1926) 1182–1201, http://dx.doi.org/10.1103/PhysRev.28.1182.

[83] Y.-Y. Liu, F. Zheng, X. Jiang, J.-W. Luo, S.-S. Li, L.-W. Wang, Ab initio investigation of charge trapping across the crystalline-Si–amorphous-SiO_2 interface, Phys. Rev. Appl. 11 (2019) 044058, http://dx.doi.org/10.1103/PhysRevApplied.11.044058.

[84] C.S. Kelley, Moments of semiclassical and classical absorption and emission band shapes of impurities in solids, Phys. Rev. B 20 (1979) 5084–5089, http://dx.doi.org/10.1103/PhysRevB.20.5084.

[85] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, 2005, http://dx.doi.org/10.1137/1.9780898717921.

[86] T.R. Hanak, R.K. Ahrenkiel, D.J. Dunlavy, A.M. Bakry, M.L. Timmons, A new method to analyze multiexponential transients for deep-level transient spectroscopy, J. Appl. Phys. 67 (9) (1990) 4126–4132, http://dx.doi.org/10.1063/1.344973.

[87] A.A. Istratov, O.F. Vyvenko, Exponential analysis in physical phenomena, Rev. Sci. Instrum. 70 (2) (1999) 1233–1257, http://dx.doi.org/10.1063/1.1149581.

[88] J. Franco, J.-F. de Marneffe, A. Vandooren, H. Arimura, L.-Å. Ragnarsson, D. Claes, E.D. Litta, N. Horiguchi, K. Croes, D. Linten, T. Grasser, B. Kaczer, Low temperature atomic hydrogen treatment for superior NBTI reliability—Demonstration and modeling across SiO_2 IL thicknesses from 1.8 to 0.6 nm for I/O and core logic, in: 2021 Symposium on VLSI Technology, 2021, pp. 1–2.

[89] G. Rzepa, F. Schanovsky, M. Karner, From gate oxide characterization to TCAD predictions: Exploring impact of defects across technologies, in: 2019 Silicon Nanoelectronics Workshop, SNW, 2019, pp. 1–2, http://dx.doi.org/10.23919/SNW.2019.8782971.

[90] D. Claes, J. Franco, N. Collaert, D. Linten, M. Heyns, Positive bias temperature instability of HfO$_2$-based gate stacks at reduced thermal budget for future CMOS technologies, J. Appl. Phys. 128 (10) (2020) 104101, http://dx.doi.org/10.1063/5.0006110.

[91] A.N. Tikhonov, A. Goncharsky, V. Stepanov, A.G. Yagola, Numerical Methods for the Solution of Ill-Posed Problems, Vol. 328, Springer Science & Business Media, 1995, http://dx.doi.org/10.1007/978-94-015-8480-7.

[92] D. Calvetti, S. Morigi, L. Reichel, F. Sgallari, Tikhonov regularization and the L-curve for large discrete ill-posed problems, J. Comput. Appl. Math. 123 (1) (2000) 423–446, http://dx.doi.org/10.1016/S0377-0427(00)00414-3, Numerical Analysis 2000. Vol. III: Linear Algebra.

[93] M. Waltl, B. Stampfer, G. Rzepa, B. Kaczer, T. Grasser, Separation of electron and hole trapping components of PBTI in SiON nMOS transistors, Microelectron. Reliab. 114 (2020) 113746, http://dx.doi.org/10.1016/j.microrel.2020.113746, 31st European Symposium on Reliability of Electron Devices, Failure Physics and Analysis, ESREF 2020.

[94] C. Jungemann, C. Zimmermann, DC, AC and noise simulation of organic semiconductor devices based on the master equation, in: 2014 International Conference on Simulation of Semiconductor Processes and Devices, SISPAD, IEEE, 2014, pp. 137–140, http://dx.doi.org/10.1109/SISPAD.2014.6931582.

[95] G. Jegert, A. Kersch, W. Weinreich, P. Lugli, Ultimate scaling of TiN/ZrO$_2$ / TiN capacitors: Leakage currents and limitations due to electrode roughness, J. Appl. Phys. 109 (1) (2011) 014504, http://dx.doi.org/10.1063/1.3531538.

[96] A. Padovani, B. Kaczer, M. Pešić, A. Belmonte, M. Popovici, L. Nyns, D. Linten, V.V. Afanas'ev, I. Shlyakhov, Y. Lee, H. Park, L. Larcher, A sensitivity map-based approach to profile defects in MIM capacitors from $I - V$, $C - V$, and $G - V$ measurements, IEEE Trans. Electron Devices 66 (4) (2019) 1892–1898, http://dx.doi.org/10.1109/TED.2019.2900030.

[97] C. Schleich, D. Waldhör, A.-M. El-Sayed, K. Tselios, B. Kaczer, T. Grasser, M. Waltl, Single- versus multi-step trap assisted tunneling currents—Part II: The role of polarons, IEEE Trans. Electron Devices 69 (8) (2022) 4486–4493, http://dx.doi.org/10.1109/TED.2022.3185965.

[98] A. Grill, E. Bury, J. Michl, S.E. Tyaginov, D. Linten, T. Grasser, B. Parvais, B. Kaczer, M. Waltl, I. Radu, Reliability and variability of advanced CMOS devices at cryogenic temperatures, in: Proceedings of the IEEE International Reliability Physics Symposium, IRPS, 2020, pp. 1–6, http://dx.doi.org/10.1109/IRPS45951.2020.9128316.

[99] T. Markvart, Determination of potential surfaces from multiphonon transition rates, J. Phys. C: Solid State Phys. 14 (15) (1981) L435, http://dx.doi.org/10.1088/0022-3719/14/15/002.

[100] K.F. Freed, J. Jortner, Multiphonon processes in the nonradiative decay of large molecules, J. Chem. Phys. 52 (12) (1970) 6272–6291, http://dx.doi.org/10.1063/1.1672938.

[101] A. Kerber, E. Cartier, L. Pantisano, R. Degraeve, T. Kauerauf, Y. Kim, A. Hou, G. Groeseneken, H. Maes, U. Schwalke, Origin of the threshold voltage instability in SiO_2/HfO_2 dual layer gate dielectrics, IEEE Electron Device Lett. 24 (2) (2003) 87–89, http://dx.doi.org/10.1109/LED.2003.808844.

[102] M. Waltl, D. Waldhoer, K. Tselios, B. Stampfer, C. Schleich, G. Rzepa, H. Enichlmair, E.G. Ioannidis, R. Minixhofer, T. Grasser, Impact of single-defects on the variability of CMOS inverter circuits, Microelectron. Reliab. 126 (2021) 114275.