Chapter 16
Advanced Modeling of Oxide Defects

Wolfgang Goes, Franz Schanovsky, and Tibor Grasser

Abstract During the last couple of years, there is growing experimental evidence which confirms charge trapping as the recoverable component of BTI. The trapping process is believed to be a non-radiative multiphonon (NMP) process, which is also encountered in numerous physically related problems. Therefore, the underlying NMP theory is frequently found as an important ingredient in the youngest BTI reliability models. While several different descriptions of the NMP transitions are available in literature, most of them are not suitable for the application to device simulation. In this chapter, we will present a rigorous derivation that starts out from the microscopic Franck–Condon theory and yields generalized trapping rates accounting for all possible NMP transitions with the conduction and the valence band in the substrate as well as in the poly-gate. Most importantly, this derivation considers the more general quadratic electron–phonon coupling contrary to several previous charge trapping models. However, the pure NMP transitions do not suffice to describe the charge trapping behavior seen in time-dependent defect spectroscopy (TDDS). Inspired by these measurements, we introduced metastable states, which have a strong impact on the trapping dynamics of the investigated defect. It is found that these states provide an explanation for plenty of experimental features observed in TDDS measurements. In particular, they can explain the behavior of fixed as well as switching oxide hole traps, both regularly observed in TDDS measurements.

16.1 Introduction

For a long time, the research in bias temperature instability (BTI) was dominated by variants of the reaction–diffusion (RD) model [1–8], discussed in [9]. In the course of the last decade it was realized that the concept of the RD model cannot

W. Goes (✉) • F. Schanovsky • T. Grasser
Institute for Microelectronics, TU Vienna, Gusshausstrasse 27-19, 1040 Vienna, Austria
e-mail: goes@iue.tuwien.ac.at; schanovsky@iue.tuwien.ac.at; grasser@iue.tuwien.ac.at

T. Grasser (ed.), *Bias Temperature Instability for Devices and Circuits*,
DOI 10.1007/978-1-4614-7909-3_16,

explain BTI [8, 10, 11]. At the same time, a new measurement technique called time-dependent defect spectroscopy (TDDS) emerged, which indicated that some sort of charge trapping is involved in BTI. This method is capable of detecting single charge emission events from individual defects [12–18] in recovery traces that last up to a few hundred seconds. Thus TDDS allows for the analysis of the recoverable component of BTI and opened the doors toward in-depth investigations of the physical trapping mechanism underlying BTI. For a detailed description of this measurement method see [19].

First variants of charge trapping models relied on elastic hole tunneling of holes between the substrate and oxide defects [20–24]. However, these models show a negligible temperature dependence, which is in contrast to what has been observed experimentally. Other variants were based on the famous Shockley–Read–Hall (SRH) model [25] and modified to account for the tunneling effect [26] and the thermal activation of BTI [27, 28]. For the latter, transition barriers were phenomenologically introduced to reproduce the observed temperature dependence. They were reasoned by non-radiative multiphonon (NMP) transitions but were not rigorously derived from a microscopic theory [29–39]. The underlying theory provides a rigorous framework for the description of the charge transfer process between the substrate and the oxide defects in BTI. Hence, this theory forms the basis of our multi-state model and will be discussed in detail at first. Subsequently they will be simplified to make them applicable for analytical calculations.

Furthermore, TDDS studies demonstrated that the trapping dynamics must involve metastable states as well as thermal but field-insensitive transitions. This observation suggested a bistable BTI defect, which features quite complicated trapping dynamics including two-step capture and emission processes. This kind of defect also allows for different transition paths, which can explain the dual trap behavior seen in TDDS. For validation of this new model, we will evaluate the simulation results to the experimental data obtained from TDDS.

16.2 Benchmarks for a BTI Model

As a result of the continuous downscaling of the device geometries, single charge detrapping events have become visible as discrete steps in the BTI recovery curves. These steps came into the focus of scientific interest so that measurement techniques, such as TDDS, have become frequently employed. The TDDS relates these steps to several single charging or discharging events [12, 13, 18] of defects and therefore allows for the analysis of individual oxide defects and their trapping behavior. As such, the findings from TDDS [14–17] are used as criteria for the development of an atomistic BTI model and are listed in the following:

(i) The plot in Fig. 16.1 reveals that the defects exhibit a strong, nearly exponential stress voltage dependence of τ_c. Empirically, this dependence can be described by $\exp(-c_1 F_{ox} + c_2 F_{ox}^2)$. However, it differs from defect to defect, implying that it is related to certain defect properties.

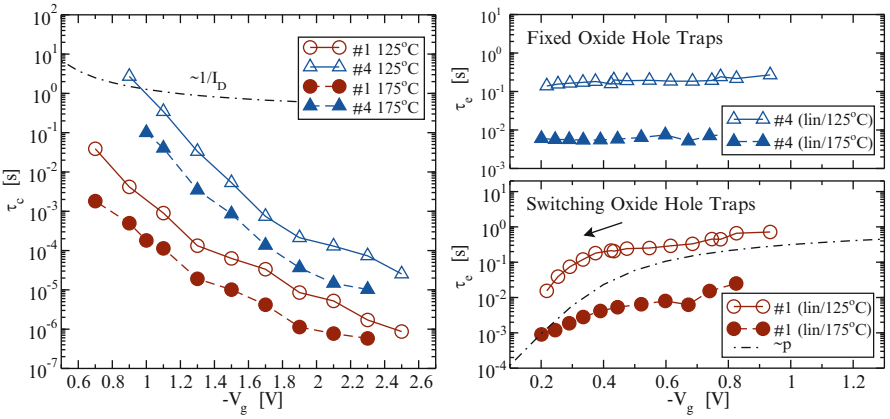

Fig. 16.1 *Left*: The capture time constants τ_c as a function of V_g for two defects at different temperatures extracted from a single device. *Open and closed symbols* mark measurements carried out at 125 °C and 175 °C, respectively. The τ_c curves show a strong field acceleration and temperature activation. However, the observed field acceleration does not follow the $1/I_d \approx 1/p$ dependence (*dot-dashed line*) as predicted by the conventional SRH model. *Right*: The emission time constants τ_e for single defects gathered from the TDDS for varying recovery gate voltages. The two distinct field dependences (*upper and lower panel*) suggest the existence of two types of defects present in the oxide. The defect #1 shows different field behaviors depending on whether the device is operated in the linear or the saturation regime during the measurement (not shown here). This suggests that the electrostatics within the device are responsible for the two distinct field dependences. It is noteworthy that the drop in τ_e goes hand in hand with the decrease in the interfacial hole concentration p (*dot-dashed line*)

(ii) The time constant plots show a marked temperature dependence, which becomes obvious by the downward shift of the τ_c curves at higher temperatures. The activation energies extracted from Arrhenius plots are about 0.6 eV.
(iii) One type of the oxide defects ("fixed oxide hole traps") has a τ_e that remains unaffected by changes in V_g [40, 41].
(iv) The other type ("switching oxide hole traps") shows a drop in τ_e toward lower V_g [40–42].
(v) The τ_e of both types shows a temperature activation with a large spread (0.6–1.4 eV).

Furthermore, it was found that several TDDS recovery traces display random telegraph noise (RTN) when studying a device at certain bias conditions [14]. After a while, this RTN signal vanishes and does not reoccur during the remaining measurement time. The termination of the noise signal is ascribed to hole traps which change to their neutral charge state and remain therein. This kind of noise is termed temporary RTN [14] (tRTN) since it occurs only for a limited amount of time. A similar phenomenon called anomalous RTN (aRTN) was discovered earlier by Kirton and Uren [27]. Therein, electron traps were observed, which repeatedly

produce noise for random time intervals. During the interruptions of this RTN signal, the defects dwell in their negative charge state generating no RTN noise signal. The behavior of these traps was interpreted by the existence of a metastable defect state. Unfortunately, there exist only a quite limited amount of noise data so that no reliable statistics can be generated. Nevertheless, it is viewed as a stringent requirement that the sought BTI model can also capture these noise phenomena in principle.

16.3 Previous Modeling Attempts

Early BTI modeling attempts relied on the classical reaction–diffusion model [1, 2, 5–7] or variants thereof [3, 5, 8] accounting for dispersive diffusion [3, 8] and three-dimensional effects [10, 11]. Even though these models are still popular, it has been demonstrated that the underlying concept cannot describe the basic feature of BTI (see [9]). As an alternative explanation for BTI, charge trapping based on elastic electron tunneling was previously suggested. However, this process exhibits a far too weak temperature dependence as compared to measurements. The next evolution of trapping models rested upon SRH theory combined with elastic tunneling, thereby mimicking an inelastic and thus temperature-activated trapping process. To its disadvantage, the underlying trapping process is not specified within the general SRH framework and can therefore not be linked to simulations based on well-founded atomistic theories. A prototype version of this SRH model was proposed by McWhorter [26], who extended the SRH equations by the factor $\exp(x_\mathrm{t}/x_0)$ in order to account for the effect of electron tunneling. Since this model suffers from a weak temperature dependence of τ_c and small time constants, Kirton and Uren [27] incorporated a term with field-independent energy barriers ΔE_b. This "ad hoc" introduction of barriers has been motivated by the theory of non-radiative multi-phonon transitions (NMP) process [38]. However, Kirton and Uren did not provide a detailed theoretical derivation based on NMP theory. Nevertheless, their work must be regarded as a substantial improvement in the interpretation of charge trapping at semiconductor–oxide interfaces. In this variant, the capture and emission time constants read

$$\tau_\mathrm{c} = \tau_0 \exp\left(\frac{x_\mathrm{t}}{x_0}\right) \exp(\beta \Delta E_\mathrm{b}) \frac{N_\mathrm{v}}{p} \begin{cases} 1, & E_\mathrm{t} > E_\mathrm{v} \\ \exp(-\beta \Delta E_\mathrm{t}) \exp(\beta q_0 F_\mathrm{ox} x_\mathrm{t}), & E_\mathrm{t} < E_\mathrm{v} \end{cases} \tag{16.1}$$

$$\tau_\mathrm{e} = \tau_0 \exp\left(\frac{x_\mathrm{t}}{x_0}\right) \exp(\beta \Delta E_\mathrm{b}) \begin{cases} \exp(\beta \Delta E_\mathrm{t}) \exp(-\beta q_0 F_\mathrm{ox} x_\mathrm{t}), & E_\mathrm{t} > E_\mathrm{v} \\ 1, & E_\mathrm{t} < E_\mathrm{v} \end{cases} \tag{16.2}$$

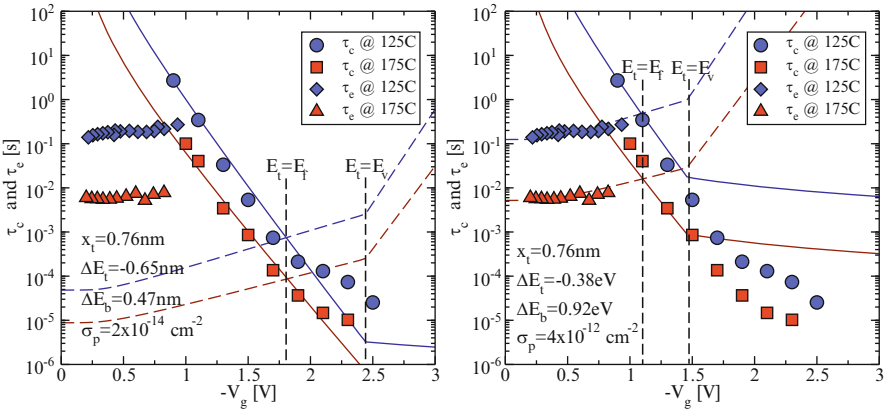

Fig. 16.2 Two fits of the Kirton model to the TDDS data. The symbols stand for the measurement data and the lines represent the simulated time constants. *Left*: When the Kirton model is optimized to the hole capture times τ_c, reasonable fits can be achieved but τ_e is predicted three orders of magnitudes too low. *Right*: Alternatively, a good agreement can be obtained for the hole emission times τ_e but with a strong mismatch of the capture times τ_c for $E_t > E_v$. From this it is concluded that the Kirton model is not capable of fitting τ_c and τ_e at the same time

where the trap level E_t is defined as

$$E_t(x_t) = E_v + \underbrace{E_{t,0} - E_{v,0}}_{=\Delta E_t} - q_0 x_t F_{ox} \tag{16.3}$$

with $E_{t,0}$ and $E_{v,0}$ denoting the trap level and the valence band edge in the absence of an electric field.

The behavior of the model with respect to the temperature and the oxide field is illustrated in the left plot of Fig. 16.2 (left). When the trap level lies below the valence band edge ($E_t < E_v$), τ_c shows an exponential field dependence. At low gate biases, the breakdown of the inversion layer gives rise to a drop in the hole concentration and in consequence to a strong increase in τ_c. Comparing the model to the experimental TDDS data, this exponential behavior allows for reasonably good, approximative fits of τ_c but is still incompatible with the observed curvature in τ_c (see the left fit in Fig. 16.2). τ_e is experimentally observed to be field insensitive, which goes hand in hand with (16.2) based on Boltzmann statistics. However, when accurate Fermi–Dirac statistics (as implemented in device simulators) are employed, the emission times exhibit a weak field dependence that agrees reasonably well with the behavior of fixed oxide hole traps (constant emission times) but is incompatible with the behavior of switching oxide hole traps (a drop at weak oxide fields). Alternatively, when τ_e is optimized in the Kirton model (see right fit in Fig. 16.2), a reasonable fit can be achieved but at the same time a strong mismatch arises for τ_c in the range $E_t > E_v$. Furthermore, Fig. 16.2 reveals that the introduction of ΔE_b yields the required temperature activation and larger time constants in agreement with the

points (ii) and (v) of the TDDS findings. Even though the model can reproduce several features seen in the TDDS data separately—except for the curvature in τ_c—no reasonable agreement with the whole set of measurement data can be achieved.

16.4 NMP Transitions Between Single States

Contrary to the previously discussed charge trapping models, the non-radiative multiphonon (NMP) theory [37–39] relies on a solid physical foundation. Its understanding requires the knowledge of fundamental microscopic theories, which are briefly discussed in the following. In the Huang–Born approximation, a certain atomic configuration is split into a system of electrons and nuclei, which are described by two separated Schrödinger equations.

$$\left\{\hat{T}_\mathrm{e} + \hat{V}_\mathrm{ee}(\mathbf{r}) + \hat{V}_\mathrm{en}(\mathbf{r};\mathbf{R}) + \hat{V}_\mathrm{nn}(\mathbf{R})\right\}\varphi_i(\mathbf{r};\mathbf{R}) = V_i(\mathbf{R})\varphi_i(\mathbf{r};\mathbf{R}) \tag{16.4}$$

$$\left\{\hat{T}_\mathrm{n} + V_i(\mathbf{R})\right\}\eta_{i\alpha}(\mathbf{R}) = E_{i\alpha}\eta_{i\alpha}(\mathbf{R}) \tag{16.5}$$

These equations contain Coulomb contributions from the electron–electron ($\hat{V}_\mathrm{ee}$), electron–nucleus ($\hat{V}_\mathrm{en}$), and nucleus–nucleus ($\hat{V}_\mathrm{nn}$) interactions as well as the kinetic energies of the electrons ($\hat{T}_\mathrm{e}$) and the nuclei ($\hat{T}_\mathrm{n}$). The electronic Hamiltonian in (16.4) depends on the electronic ($\mathbf{r}$) and the nuclear ($\mathbf{R}$) degrees of freedom, where the latter only enter parametrically. The solution $V_i(\mathbf{R})$ of the electronic Schrödinger equation (16.4) corresponds to the energy of a certain atomic configuration and acts as a potential for the nuclei in the Schrödinger equation (16.5). Therefore, $V_i(\mathbf{R})$ is usually referred to as the adiabatic potential energy. In the Huang–Born approximation, the nuclei of the atoms are treated as a system of quantum mechanical particles with quantized states $\eta_{i\alpha}$ and discrete energies $E_{i\alpha}$. Also the wavefunction of the composite electron–nucleus system is split into an electronic $\varphi_i(\mathbf{r};\mathbf{R})$ and nuclear $\eta_{i\alpha}(\mathbf{R})$ part, denoted the electronic and the vibrational wavefunction, respectively.

In the case of charge trapping in BTI, one deals with a process that is frequently termed "charge transfer reaction" in the theoretical literature. Such a kind of process must be described by a system consisting of all atoms involved. Since the trapped charge carrier is exchanged between the defect and the substrate, the system includes the atoms surrounding the BTI defect as well as the atoms in the substrate. Altogether, these atoms span a $3N$-dimensional space with N being the number of considered atoms. The adiabatic potential energy surface in this configurational space is usually visualized in a configuration coordinate diagram (see Fig. 16.3). Therein, the atomic positions are reduced to a one-dimensional quantity called configuration coordinate, which allows to describe the correlated motion of atoms, such as lattice relaxation. In these plots, the adiabatic potential energy surfaces assume an almost parabolic shape for small atomic displacements and are thus usually approximated by harmonic quantum oscillators in solid state theory.

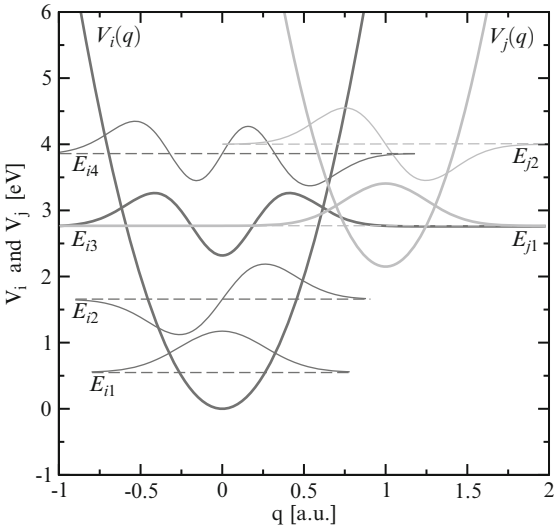

Fig. 16.3 Adiabatic potentials involved in a charge transfer reaction. Each of the two parabola corresponds to one charge state of the defect where the left (V_i) one represents the initial and the right (V_j) one the final charge state. Their corresponding wavefunctions and eigenenergies are depicted as *solid* and *dashed lines*, respectively. An NMP transition only occurs when the initial and the final energies coincide as it is the case for E_{i3} and E_{j1}. Then the overlap of their corresponding vibrational wavefunctions enters the calculation of the lineshape function f_{ij} and consequently determines the NMP transition probability

During a charge trapping process, the defect changes from the charge state i to j, where each of the charge states is represented by its own adiabatic potential in the configuration coordinate diagram (see Fig. 16.3). The NMP transition rate k_{ij} is then derived from first-order time-dependent perturbation theory using the Franck–Condon approximation [37, 39, 43, 44].

$$k_{ij} = A_{ij} f_{ij} \tag{16.6}$$

$$A_{ij} = \frac{2\pi}{\hbar} |\langle \varphi_i | V' | \varphi_j \rangle|^2 \tag{16.7}$$

$$f_{ij} = \underset{\alpha}{\text{ave}} \sum_{\beta} |\langle \eta_{i\alpha} | \eta_{j\beta} \rangle|^2 \tag{16.8}$$

Here, "ave" stands for the thermal average over all initial states "α" and the sum runs over the final states "β". A_{ij} is the electronic matrix element with the adiabatic operator as a perturbation V' and is associated with a simple electronic transition. The Franck–Condon factor $|\langle \eta_{i\alpha} | \eta_{j\beta} \rangle|^2$ in (16.8) only gives a contribution when the initial and the final state have the same energy. If this is the case, this factor is calculated as the overlap integral of the two vibrational wavefunctions "$i\alpha$" and "$j\beta$" and corresponds to the respective transition probability (cf. Fig. 16.3). Calculating the thermal average over the initial states α and summing over the final

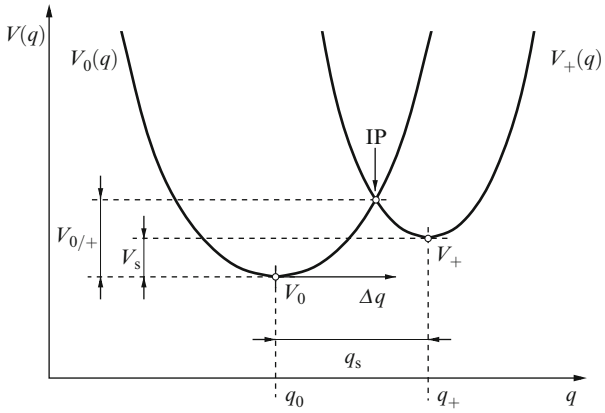

Fig. 16.4 The configuration coordinate diagram for an NMP transition. The adiabatic potentials for the initial and the final states are denoted as $V_0(q)$ and $V_+(q)$, respectively. They are defined by their corresponding minima V_0 and V_+ located at their equilibrium configurations q_0 and q_+, respectively. To simplify the mathematical calculations, the axis origin is shifted into the energy minimum V_0

states β yield the lineshape function f_{ij} that will be found to govern the gate bias and temperature dependence of the NMP transition rate. In solids the eigenspectrum $E_{i\alpha}$ is usually densely spaced so that there are numerous possible transitions from the initial charge i to the final charge state j. This lineshape function has its largest contributions from those energies that lie close to the intersection point (IP) of the adiabatic potentials. Around this point, the lineshape function is assumed to have a Dirac peak in the classical limit [45]. This assumption allows for simple analytical expressions that can be conveniently employed for device simulation.

In the following, the NMP transition rates will be derived for a defect which changes between its neutral (0) and its positive (+) charge state upon hole trapping or detrapping. The corresponding initial ($i = 0$) and final ($j = +$) potential energy surface can be expressed as

$$V_0(q) = c_0(q - q_0)^2 + V_0 = c_0 \Delta q^2 + V_0 \tag{16.9}$$

$$V_+(q) = c_+(q - q_+)^2 + V_+ = c_+(\Delta q - q_s)^2 + V_0 + V_s \tag{16.10}$$

using the quantities defined in Fig. 16.4 and the shorthands $\Delta q = q - q_0$, $q_s = q_+ - q_0$, and $V_s = V_+ - V_0$. c_0 and c_+ denote the curvature of the adiabatic potentials for the neutral and the positively charged defect, respectively. Without loss of generality, V_0 can be chosen to be zero and will thus be neglected from now on. Note that the two parabolas are characterized by different curvatures ($c_0 \neq c_+$), implying that there exist two intersection points given by

$$\Delta q_{1,2} = \frac{c_+ q_s \pm \sqrt{c_0 c_+ {q_s}^2 + V_s(c_0 - c_+)}}{c_+ - c_0}. \tag{16.11}$$

In the literature, this case is usually referred to as quadratic electron–phonon coupling. For equal curvatures ($c_0 = c_+ = c$), linear electron–phonon coupling is obtained, which yields only one intersection point located at

$$\Delta q_1 = \frac{V_s/c + q_s^2}{2q_s}. \tag{16.12}$$

The classical lineshape function for hole capture is obtained from

$$f_{0/+}(c_0, c_+, q_s, V_0, V_+) = Z^{-1} \int_q \mathrm{e}^{-\beta V_0(q')} \delta\left(V_0(q') - V_+(q')\right) \mathrm{d}q' \tag{16.13}$$

with the partition function

$$Z = \int_q \mathrm{e}^{-\beta V_0(q')} \mathrm{d}q'. \tag{16.14}$$

In accordance with the classical limit, the Dirac delta function in (16.13) ensures that the integral is only evaluated at the intersection point of the two parabolas. Using the integration rule for Dirac delta functions, this integral evaluates to

$$\begin{aligned} &\int_{\Delta q} \mathrm{e}^{-\beta V_0(\Delta q')} \delta\left(V_0(\Delta q') - V_+(\Delta q')\right) \mathrm{d}(\Delta q') \\ &= \frac{\mathrm{e}^{-\beta c_0 \Delta {q_1}^2}}{|2c_0 \Delta q_1 - 2c_+(\Delta q_1 - q_s)|} + \frac{\mathrm{e}^{-\beta c_0 \Delta {q_2}^2}}{|2c_0 \Delta q_2 - 2c_+(\Delta q_2 - q_s)|} \end{aligned} \tag{16.15}$$

and the partition function simplifies to

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-\beta c_0 \Delta {q'}^2} \mathrm{d}(\Delta q') = \sqrt{\frac{\pi}{c_0 \beta}}. \tag{16.16}$$

Inserting (16.15) and (16.16) into the definition of the lineshape function (16.13) leads to [45]

$$\begin{aligned} &f_{0/+}(c_0, c_+, q_s, V_0, V_+) = f_{0/+}(c_0, c_+, q_s, V_s) \\ &= \frac{1}{2}\sqrt{\frac{c_0\beta}{\pi}} \left(\frac{\mathrm{e}^{-\beta c_0 \Delta {q_1}^2}}{|c_0 \Delta q_1 - c_+(\Delta q_1 - q_s)|} + \frac{\mathrm{e}^{-\beta c_0 \Delta {q_2}^2}}{|c_0 \Delta q_2 - c_+(\Delta q_2 - q_s)|} \right). \end{aligned} \tag{16.17}$$

Keep in mind that the lineshape function may also vanish ($f_{0/+} = 0$) when the two parabolas do not share a common intersection point. For linear electron–phonon coupling ($c = c_0 = c_+$), the above expression reduces to

$$f_{0/+}(c, q_s, V_0, V_+) = f_{0/+}(c, q_s, V_s) = \frac{1}{2}\sqrt{\frac{c\beta}{\pi}} \frac{\mathrm{e}^{-\beta c \Delta {q_1}^2}}{c|q_s|} \tag{16.18}$$

with

$$\Delta q_1 = \frac{V_s/c + {q_s}^2}{2q_s}. \tag{16.19}$$

It is emphasized here that the lineshape function is most strongly affected by the exponential term, where the expression $c_0 {\Delta q_{1,2}}^2$ can be identified with the energy barrier from the minimum V_0 to the saddle point IP (cf. Fig. 16.4). This NMP transition barrier can be expressed as

$$\begin{aligned} V_{0/+} &= V_0(\Delta q_{1,2}) \\ &= \frac{c_0 q_s^2}{(\frac{c_0}{c_+} - 1)^2} \left(1 \pm \sqrt{\frac{c_0}{c_+} + \frac{V_s(\frac{c_0}{c_+} - 1)}{c_+ q_s^2}} \right)^2, \end{aligned} \tag{16.20}$$

or

$$V_{0/+} = \left(\frac{V_s + c q_s^2}{2\sqrt{c} q_s} \right)^2 \tag{16.21}$$

for linear electron–phonon coupling. For hole emission the roles of the initial and the final states are reversed. The corresponding lineshape function $f_{+/0}$ and the NMP barrier $V_{+/0}$ are of the same form as in (16.17) and (16.20), respectively, but have their subscripts "0" and "+" exchanged.

As will be demonstrated in Sect. 16.7, the NMP transition barrier varies strongly with the temperature and the gate bias and therefore governs the trapping behavior of BTI defects. In the following calculations, the above analytical expressions for the lineshape function are preferred to the Franck–Condon overlap factors since they can be easily implemented in simple device simulators at computational feasible costs.

Next, the NMP theory has to be specified for the situation of charge capture and emission in MOSFETs. Therefore, the energy minima V_0 and V_+ at the potential energy surfaces must be linked to the energy of the transferred electron in the band energy diagram before and after an NMP transition. In a simplified picture, it can be envisioned that only the energy of the transferred electron changes while the energy of the other electrons ($\tilde{V}_0$) remains unaffected. In the following, we discuss a hole capture[1] process, during which an electron is emitted from the energy level E_t of a trap into an energy level E in the substrate valance band state. Then the energy minima V_0 and V_+ can be expressed as

$$V_0 = \tilde{V}_0 + E_t \tag{16.22}$$

$$V_+ = \tilde{V}_0 + E \tag{16.23}$$

[1]It is stressed that the term "hole capture" refers to either a capture of hole from the valence band into a trap or an emission of an electron from the trap into the valence band. Keep in mind that both of these processes are equivalent from a physical point of view.

with $\tilde{V}_0$ being the energy of the system minus the energy of the transferred electron. The NMP transition rate is then written as

$$k_{0/+} = A_{0/+}(E) f_{0/+}(c_0, c_+, q_s, E - E_t). \tag{16.24}$$

The unknown auxiliary quantity $\tilde{V}_0$ cancels out in the lineshape function, which only depends on the energy difference

$$V_s = V_+ - V_0 = E - E_t\,. \tag{16.25}$$

The trap wavefunction in the electronic matrix element $A_{0/+}(E)$ is strongly localized around the defect so that the integrand in (16.7) has its largest contribution at the defect site and $A_{0/+}(E, x_t)$ can be approximated by

$$\begin{aligned} A_{0/+}(E, x_t) &= A_0 |\langle x_t | \varphi \rangle|^2 = A_0 |\varphi(x_t)|^2 \\ &= A_1 \lambda(E, x_t)\,. \end{aligned} \tag{16.26}$$

Here, A_0 is a not further specified prefactor and $\varphi(E)$ stands for the channel wavefunction with an energy E. The electronic matrix element is governed by the exponential decay of the channel wavefunction and can be approximated using a WKB factor $\lambda(E, x_t)$ for the implementation in simple device simulators.

16.5 NMP Transition with a Whole Band of States

So far, the theoretical foundation for NMP transitions between two certain states has been discussed. In BTI, however, the oxide defects interact with the whole conduction or valence band of the substrate so that the current formulation of the NMP processes must be extended to account for transitions with a multitude of band states at different energies E. For this reason, one has to introduce a summation over all possible valence band states n in (16.24). Since the valence band states form a continuous spectrum, this summation can also be transformed to an integral over a density of states [46].

$$\sum_n \to \Omega \int_{-\infty}^{E_v} D_p(E)\mathrm{d}E \tag{16.27}$$

Using the above transformation, the NMP hole capture rate can be expressed as

$$k_{0/+}^{pc} = \Omega \int_{-\infty}^{E_v} D_p(E) A_{0/+}(E, x_t) f_{0/+}(c_0, c_+, q_s, E - E_t)\mathrm{d}E\,. \tag{16.28}$$

The density of states $D_\mathrm{p}(E)$ can be calculated using a simple expression based on the parabolic band approximation

$$D_\mathrm{p}(E) = \sum_\nu \frac{g_\nu m_{\mathrm{p}\nu}}{\hbar^3 \pi^2} \sqrt{2 m_{\mathrm{p}\nu}(E - E_\mathrm{c})}\,, \tag{16.29}$$

where g_ν is the degeneracy of the νth valence band valley and m_p its corresponding effective hole mass. Alternatively, the density of states may originate from a more sophisticated Schrödinger–Poisson solver that allows for quantized states $E_{\nu k}$ arising from the one-dimensional confinement of the charge carriers in the inversion layer.

$$D_\mathrm{p}(E) = \sum_\nu \frac{g_\nu m_{\mathrm{p}\nu}}{\hbar^2 \pi} \sum_k \Theta(E - E_{\nu k}) \tag{16.30}$$

Next, the hole occupancy of the band states (f_p for E) and electron occupancy of the trap state (f_t for E_t) have to be taken into account. Then, the resulting NMP transition rates read

$$k^{\mathrm{pc}}_{0/+} = \Omega \int_{-\infty}^{E_\mathrm{v}} D_\mathrm{p}(E) f_\mathrm{p}(E, E_\mathrm{f}) A_{0/+}(E, x_\mathrm{t}) f_{0/+}(c_0, c_+, q_\mathrm{s}, E - E_\mathrm{t}) f_\mathrm{t} \mathrm{d}E\,. \tag{16.31}$$

For the case of electron emission,[2] the electron is emitted into the substrate conduction band and thus $D_\mathrm{p}(E)$ must be replaced by $D_\mathrm{n}(E)$.

$$k^{\mathrm{ne}}_{0/+} = \Omega \int_{E_\mathrm{c}}^{+\infty} D_\mathrm{n}(E) f_\mathrm{p}(E, E_\mathrm{f}) A_{0/+}(E, x_\mathrm{t}) f_{0/+}(c_0, c_+, q_\mathrm{s}, E - E_\mathrm{t}) f_\mathrm{t} \mathrm{d}E \tag{16.32}$$

The configuration coordinate diagrams of both processes are combined in Fig. 16.5, which now covers all electron or hole transitions from the defect into the substrate. Interestingly, the final states span an energy spectrum V_+ that can be identified with band energy diagram including the conduction as well as the valence band. Each of these states is associated with a distinct position of its adiabatic potential $V_+(q)$ and thus has a different NMP barrier height along with a different transition probability according to the lineshape function in the transition rates (16.31) and (16.32) (cf. Fig. 16.6). For hole capture (case A), the defect has to undergo an NMP transition from the parabola $V_0(q)$ to the parabola $V_+(q)$. This transition occurs the fastest when $V_+(q)$ cuts the minimum of $V_0(q)$. Then the corresponding transition barrier $V_{0/+}$ is negligible and the lineshape function $f_{0/+}(E)$ reaches its maximum value. When hole emission is considered (case B), the roles of the initial and the final states

[2]Note that electron emission corresponds to hole capture into the substrate conduction band.

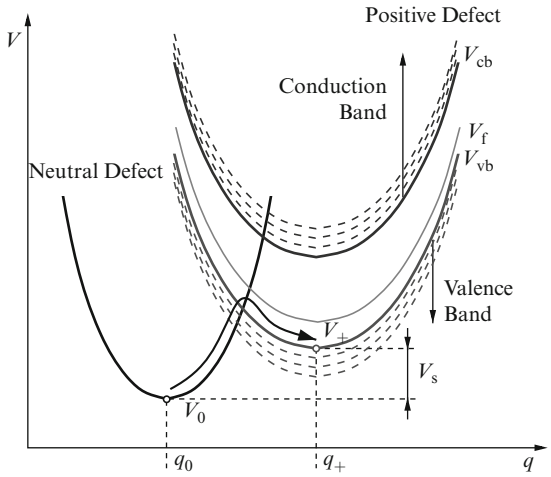

Fig. 16.5 A combined configuration coordinate diagram for hole capture and electron emission. According to the relation $V_+ = \tilde{V}_0 + E$, an electron located in an energetically higher band state E is represented by higher adiabatic potential $V_+(q)$. As a consequence, the upper and the lower family of curves constitute the set of adiabatic potentials V_+ associated with the conduction and valence band, respectively. It is noted that this configuration coordinate diagram remains unchanged for hole emission and electron capture and can therefore be used for both processes. As such, this diagram covers all possible NMP transitions of the considered defect with the substrate

are reversed so that the NMP transition proceeds from the adiabatic potential $V_+(q)$ to $V_0(q)$. Then the corresponding lineshape function $f_{+/0}(E)$ peaks when $V_+(q)$ is cut in its minimum. Note that the maximal transition rates for hole capture and emission are associated with different energy levels E, which are frequently referred to as the switching trap levels[3] in literature [47–54]. However, they should not be confused with the thermodynamic trap levels E_t that enter SRH-like formulations of the charge transfer process used here. The thermodynamic trap level (case C) is associated with the energy level E, at which the hole capture and emission are balanced and the two lineshape functions $f_{0/+}(E)$ and $f_{+/0}(E)$ assume the same value (cf. Fig. 16.6). In the configuration coordinate diagram, this is the case for the situation when the minima of adiabatic potentials $V_0(q)$ and $V_+(q)$ are at the same height. Note that special importance is attached to this energy level with respect to the equilibrium occupancy of the defect. If the Fermi level is located above the thermodynamic level, the dominating trapping process is hole emission and the defect becomes neutral. However, when the Fermi level falls below the thermodynamic level, the hole capture rate exceeds the hole emission rate and the defect becomes occupied by a hole.

[3]Note that the same term "switching trap level" is also used for the thermodynamic trap level for a switching oxide hole trap introduced in Fig. 16.1.

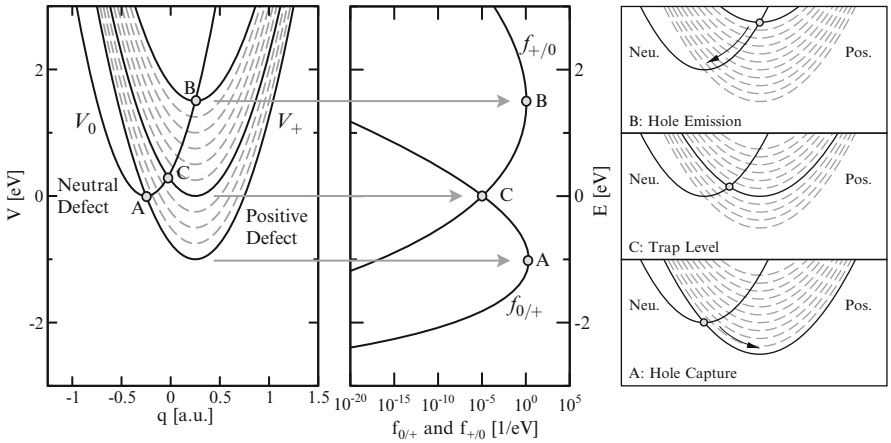

Fig. 16.6 Configuration coordinate diagram (*left*) for a continuum of adiabatic potentials $V_+(q)$, the corresponding lineshape functions (*middle*), and sketches of the cases A, B, and C (*right*). For hole capture, the lineshape function $f_{0/+}(E)$ reaches its maximal value when $V_+(q)$ intersects the minimum of $V_0(q)$ and thus the NMP transition has a vanishing barrier $V_{0/+}$ (case A). When changing from the configuration coordinate diagram (*left*) to the lineshape function (*middle*), the adiabatic potentials are converted to electron energies according to (16.23). For the hole emission, the analogous considerations apply as for hole capture. Now the intersection point must lie in the minimum of $V_+(q)$, giving rise to the peak of the lineshape function $f_{+/0}(E)$ (case B). If the minima of both parabolas coincide, the barriers for both directions have the same heights, which leads to equaling NMP transition rates (case C)

In semiconductor theory—especially when NBTI in pMOS transistors is considered—the trapping dynamics is preferentially described in the "hole picture." In this case the hole is emitted from a continuum of states where its energy in the initial state is undefined. By contrast, the hole energy is exactly specified by the trap level E_t in the final state (cf. Fig. 16.7). As a consequence, the trap level E_t and the band states E change their roles. Furthermore, the energy axis of the charge carriers is inverted so that the energy spectrum of V_+ in Fig. 16.5 is flipped in the hole picture in Fig. 16.7.

$$V_0 = \tilde{V}_0 - E \tag{16.33}$$

$$V_+ = \tilde{V}_0 - E_\mathrm{t} \tag{16.34}$$

The energy difference of the adiabatic potentials is then given by

$$V_\mathrm{s} = V_+ - V_0 = E - E_\mathrm{t}\,, \tag{16.35}$$

implying that the same activation energy is required as in the electron picture. Following the same derivation as for the electron picture, the NMP transition rate for hole capture reads

$$k_{0/+}^{\mathrm{pc}} = \Omega \int_{-\infty}^{E_\mathrm{v}} D_\mathrm{p}(E) f_\mathrm{p}(E, E_\mathrm{f}) A_{0/+}(E, x_\mathrm{t}) f_{0/+}(c_0, c_+, q_\mathrm{s}, E - E_\mathrm{t}) f_\mathrm{t} \mathrm{d}E\,. \tag{16.36}$$

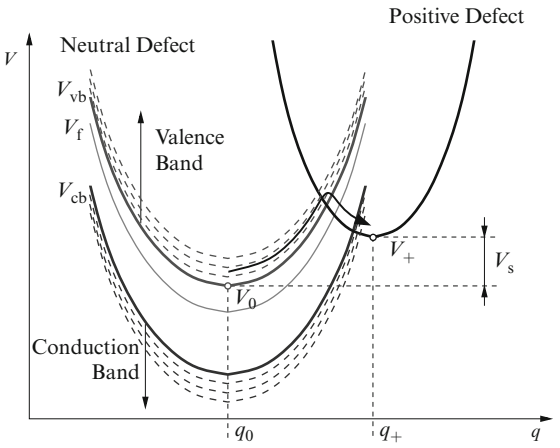

Fig. 16.7 The same configuration coordinate diagram as in Fig. 16.5 but in the "hole picture." Note that the energy scale of the charge carriers and thus the band diagram is inverted compared to the "electron picture." Furthermore, the energy of the transferred charge carrier is now undefined for the initial state since the hole is in one of the valence band states. By contrast, it can be specified by E_t when the hole is trapped

It is remarked that the electronic matrix elements $A_{0/+}(E,x_t)$ in the hole picture and in the electron picture equal since they are determined by the same channel wavefunction. Using the approximation (16.26), the whole set of NMP trapping rates can be written as

$$k^{nc} = k_0^n \int_{E_c}^{+\infty} D_n(E) f_n(E,E_f) \lambda(E,x_t) f_{+/0}(c_+,c_0,q_s,E_t-E)\mathrm{d}E \tag{16.37}$$

$$k^{ne} = k_0^n \int_{E_c}^{+\infty} D_n(E) f_p(E,E_f) \lambda(E,x_t) f_{0/+}(c_0,c_+,q_s,E-E_t)\mathrm{d}E \tag{16.38}$$

$$k^{pc} = k_0^p \int_{-\infty}^{E_v} D_p(E) f_p(E,E_f) \lambda(E,x_t) f_{0/+}(c_0,c_+,q_s,E-E_t)\mathrm{d}E \tag{16.39}$$

$$k^{pe} = k_0^p \int_{-\infty}^{E_v} D_p(E) f_n(E,E_f) \lambda(E,x_t) f_{+/0}(c_+,c_0,q_s,E_t-E)\mathrm{d}E\,, \tag{16.40}$$

where the quantities $k_0^{n/p}$ are used as shorthands for the product of the prefactors Ω and A_1. "n" and "p" refer to electrons or holes while "c" and "e" stand for capture and emission processes, respectively. It has to be noted that the integrands of the above rate equations are usually sharply peaked due to the strong exponential dependences of the occupancies $f_p(E)$ and $f_n(E)$ as well as the lineshape functions $f_{0/+}(E)$ and $f_{+/0}(E)$. Hence, these integrals are solved numerically using adaptive integration schemes in order to keep the computation costs low and to ensure a sufficient accuracy of the computed rates.

The above set of rate equations can also be modified to the case where the defect exchanges charge carriers with the poly-gate by replacing the band edges

and the Fermi level with their respective values at the poly-gate. They can also be adapted for an electron trap, whose charge state switches between neutral and negative. As such, these rate equations form the basis for charge trapping involving the substrate as well as the gate and could consequently also cover trap-assisted tunneling occurring via NMP transitions.

16.6 Huang–Rhys Parameter

The employed NMP theory was initially derived for the fluorescence and absorption spectra of gases and solids, where the Huang–Rhys factor S was introduced to obtain compact analytical solutions [37]. This quantity corresponds to the number of absorbed or emitted phonons during an optical transition and thereby characterizes the shape of two adiabatic potentials $V_0(q)$ and $V_+(q)$. For quadratic electron–phonon coupling, the adiabatic potentials are represented by two parabolas that are shifted against each other and have different curvatures. To define them, we introduce the quantities S and R (see Fig. 16.8), which are defined as follows:

$$S\hbar\omega = c_0 q_s^2 \tag{16.41}$$

$$R^2 = \frac{c_0}{c_+}. \tag{16.42}$$

Using the above substitutions, the NMP transition barrier in (16.20) can be rewritten as

$$V_{0/+}(V_s) = \frac{S\hbar\omega}{(R^2-1)^2}\left(1 \pm R\sqrt{\frac{S\hbar\omega + V_s(R^2-1)}{S\hbar\omega}}\right)^2. \tag{16.43}$$

The prefactor $\xi_{0/+}(\Delta q)$ of the exponential term in (16.17) is of the form

$$\xi_{0/+}(\Delta q_{1,2}) = \sqrt{\frac{\beta c_0}{4\pi}}\frac{1}{|c_0\Delta q - c_+(\Delta q - q_s)|} \tag{16.44}$$

and can be expressed as

$$\xi_{0/+}(V_s) = \sqrt{\frac{\beta}{4\pi}}\frac{R}{\sqrt{S\hbar\omega + V_s(R^2-1)}}. \tag{16.45}$$

For linear electron–phonon coupling, one obtains the frequently applied result

$$V_{0/+}(V_s) = \frac{(V_s + S\hbar\omega)^2}{4S\hbar\omega} \tag{16.46}$$

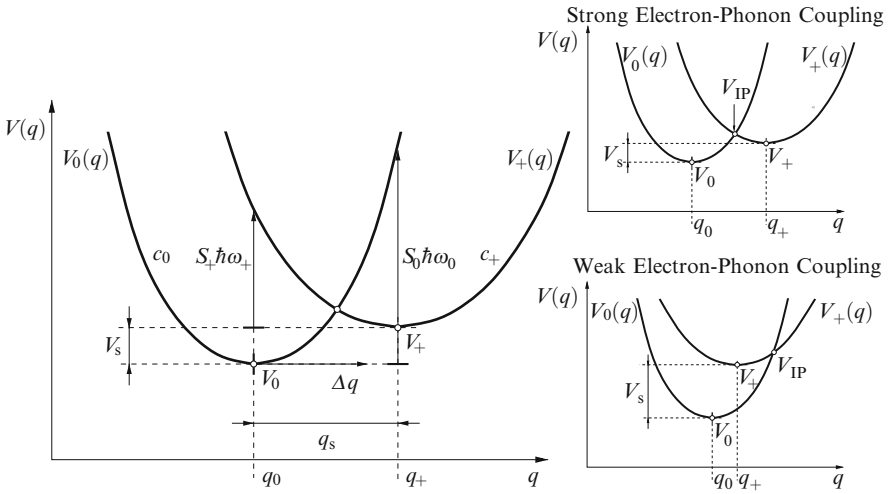

Fig. 16.8 *Left*: The configuration coordinate diagram including the Huang–Rhys factors S_0 and S_+. The adiabatic potentials are often defined as harmonic oscillators of the form $V_0(q) = {}^1\!/_2 M\omega_0^2(q-q_0)^2 + V_0$ and $V_+(q) = {}^1\!/_2 M\omega_+^2(q-q_+)^2 + V_+$, where ω_0 and ω_+ are their respective oscillator frequencies. For an optical transition, the energy delivered by the photon must equal the energy difference $V_0(q_+) - V_+$, which is indicated by the upwards arrow and can be expressed as an integral multiple S_0 of $\hbar\omega_0$. In analogy, $S_+\hbar\omega_+$ equals the energy difference $V_+(q_0) - V_0$. In the remainder of this chapter, $S_0\hbar\omega_0$ and $S_+\hbar\omega_+$ will be replaced by $S\hbar\omega$ and $R^2 S\hbar\omega$, respectively. *Right*: Strong (*top*) and weak (*bottom*) electron–phonon coupling. In the first case the parabolas are positioned such that the intersection point is situated in between their minima while in the second case one parabola lies inside the other and the intersection point is located beside the two minima

for the NMP transition barrier with the prefactor

$$\xi_{0/+}(\Delta q) = \sqrt{\frac{\beta c_0}{4\pi}} \frac{1}{|c_0 \Delta q - c_+(\Delta q - q_s)|} = \sqrt{\frac{\beta}{4\pi}} \frac{1}{\sqrt{S\hbar\omega}} . \tag{16.47}$$

16.6.1 Analytical Expressions for the NMP Rates

A second order expansion of the expression (16.43) delivers

$$V_{0/+}(V_s) \approx \frac{S\hbar\omega}{(1+R)^2} + \frac{R}{1+R} V_s + \frac{R}{4S\hbar\omega} V_s^2 . \tag{16.48}$$

If the curvatures c_0 and c_+ differ, the quantity R deviates from unity. Since R also enters the above expression for the barrier height, the ratio of the curvatures has a

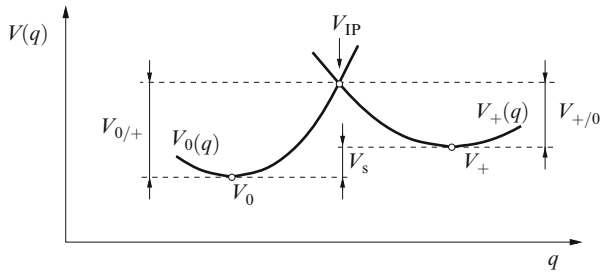

Fig. 16.9 The hole capture ($V_{0/+}$) and emission ($V_{+/0}$) barrier for an NMP transition. The barrier heights are calculated as the energy differences between the corresponding minimum and the intersection point in the hole picture, which yields $V_{0/+} = V_{\mathrm{IP}} - V_0 = V_{\mathrm{IP}} - \tilde{V}_0 + E$ and $\Delta V_{+/0} = V_{\mathrm{IP}} - V_+ = V_{\mathrm{IP}} - \tilde{V}_0 + E_\mathrm{t}$ for the capture and the emission barrier, respectively

strong impact on the NMP transition rates (cf. Fig. 16.9). As in the previous section, V_s can be expressed as

$$V_\mathrm{s} = V_+ - V_0 = E - E_\mathrm{t} = \underbrace{E - E_\mathrm{v}}_{=-\Delta E} + E_\mathrm{v} - E_\mathrm{t} \tag{16.49}$$

so that (16.48) can be rewritten as

$$V_{0/+}(\Delta E) \approx \frac{S\hbar\omega}{(1+R)^2} + \frac{R}{1+R}\left(E_\mathrm{v} - E_\mathrm{t} - \Delta E\right) + \frac{R}{4S\hbar\omega}\left(E_\mathrm{v} - E_\mathrm{t} - \Delta E\right)^2. \tag{16.50}$$

In the case of strong electron–phonon coupling (see Fig. 16.8) $S\hbar\omega \gg |E_\mathrm{v} - E_\mathrm{t} - \Delta E|$ holds and the third term of (16.50) can be neglected. In order to evaluate the integral in the hole capture rate (16.39), the following assumptions must be made:

- Assuming the parabolic-band approximation, the valence band density of states (16.29) is given by $D_\mathrm{p}(E) = D_{\mathrm{p},0}\sqrt{\Delta E}$ with $D_{\mathrm{p},0}$ being an energy-independent prefactor.
- The occupancy $f_\mathrm{p}(E, E_\mathrm{f})$ follows Boltzmann statistics.
- The WKB factor is approximated by the factor $\exp(-x_\mathrm{t}/x_0)$ with the tunneling length x_0.
- The lineshape function is dominated by the exponential barrier term so that the prefactor $\xi_{0/+}$ can be neglected to first order.

With the above simplifications, the hole capture rate (16.39) evaluates to

$$\begin{aligned} k^\mathrm{pc} &= k_0^\mathrm{p} \int_{-\infty}^{E_\mathrm{v}} D_\mathrm{p}(E) f_\mathrm{p}(E, E_\mathrm{f}) \lambda(E, x_\mathrm{t}) \exp(-\beta V_{0/+}(\Delta E)) \mathrm{d}E \\ &= k_0^\mathrm{p} (1+R)^{3/2} p \exp(-x_\mathrm{t}/x_0) \exp\left(-\beta\left(\frac{S\hbar\omega}{(1+R)^2} - \frac{R}{1+R}\Delta E_\mathrm{t}\right)\right), \end{aligned} \tag{16.51}$$

where the hole density p is given by the expression

$$p = D_{\mathrm{p},0}\,\exp\big(\beta(E_\mathrm{v}-E_\mathrm{f})\big)\;\beta^{-3/2}\,\Gamma(3/2) \tag{16.52}$$

with $\Gamma(x)$ being the Gamma function. Motivated by the similarity to the rate equations in the standard SRH theory, the prefactor k_0^p has been identified with the hole thermal velocity $v_\mathrm{th,p}$ times a hole capture cross-section σ_p. The lengthy expression in the exponent of the last term of (16.51) can be related to the hole capture barrier ε^pc, which is evaluated for $E = E_\mathrm{v}$.

$$\frac{S\hbar\omega}{(1+R)^2} + \frac{R}{1+R}E_\mathrm{v} - E_\mathrm{t} = V_{0/+}\Big|_{\Delta E=0} = \varepsilon^\mathrm{pc} \tag{16.53}$$

This is actually surprising since the NMP transition barrier $V_{0/+}(\Delta E)$ is a function of the hole energy E per definition. However, for strong electron–phonon coupling, the rate integral (16.39) delivers its largest contribution close to the valence band edge ($\Delta E = 0$) so that the barrier $V_{0/+}(E)$ can be approximated by $V_{0/+}(E_\mathrm{v})$. As a consequence, the hole capture rate simplifies to

$$k^\mathrm{pc} = v_\mathrm{th,p}\sigma_\mathrm{p}(1+R)^{3/2}\exp(-x_\mathrm{t}/x_0)p\exp\left(-\beta\varepsilon^\mathrm{pc}\right). \tag{16.54}$$

The hole emission rate is derived from (16.40) using the two relations: First, the electron occupation function can be replaced by

$$f_\mathrm{n}(E,E_\mathrm{f}) = f_\mathrm{p}(E,E_\mathrm{f})\exp\big(-\beta(E-E_\mathrm{f})\big). \tag{16.55}$$

Second, the ratio of the exponential barrier terms (see Fig. 16.9) gives

$$\exp(-\beta V_{+/0})/\exp(-\beta V_{0/+}) = \exp\left(-\beta\left(E_\mathrm{t}-E\right)\right) \tag{16.56}$$

for each band state E. Inserting both relations in (16.40) and using the same assumptions as before yields the hole emission rate

$$\begin{aligned} k^\mathrm{pe} &= v_\mathrm{th,p}\sigma_\mathrm{p}\int_{+\infty}^{E_\mathrm{v}} D_\mathrm{p}(E)f_\mathrm{n}(E,E_\mathrm{f})\lambda(E,x_\mathrm{t})\exp(-\beta V_{0/+})\mathrm{d}E \\ &= v_\mathrm{th,p}\sigma_\mathrm{p}(1+R)^{3/2}\exp(-x_\mathrm{t}/x_0)p\exp\left(-\beta\varepsilon^\mathrm{pc}\right)\exp\left(-\beta\left(E_\mathrm{t}-E_\mathrm{f}\right)\right). \end{aligned} \tag{16.57}$$

Interestingly, (16.54) and (16.57) closely resemble the rates obtained from the standard SRH theory except from the exponential barrier terms and even have the same shape as those of Kirton and Uren. However, the NMP transition barriers derived above are calculated from the intersection point of two adiabatic potentials—in this case parabolas—and thus reflect their gate bias dependence governed by the energy separation between the trap level and the valence band edge according to (16.53). Even though they rely on a series of approximations, they contain the main physics

involved in charge trapping. As such, they promote the understanding of the gate bias and temperature tendencies in charge trapping and allow compact analytical expressions for the assumption of strong electron–phonon coupling.

16.7 State Diagram of the Multi-State Model

The NMP transition rates derived in the previous sections describe charge transfer reactions, i.e., the pure charge trapping or detrapping processes. However, the TDDS studies revealed that some defects are found to disappear on the spectral maps. This observation can only be reasoned by the existence of metastable states, in which the oxide defects dwell for a certain amount of time. Furthermore, the TDDS also reveals gate bias-independent transitions that cannot be related to charge transfer reactions. These transitions are associated with an activation over thermal barriers, leaving the charge state of the defect unchanged. Both observations suggest a bistable defect, which has an additional metastable configuration (marked by primes) that appears in two charge states (cf. Fig. 16.10). This means that the defect features two neutral (1, $1'$) and two positive (2, $2'$) charge states (cf. Fig. 16.10), where thermal transitions allow for transitions between same charge states ($1 \leftrightarrow 1'$ and $2 \leftrightarrow 2'$) and NMP transitions between opposite charge states ($1 \leftrightarrow 2'$ and $2 \leftrightarrow 1'$). The bistable defect described above is the heart of the "multi-state model" and will be discussed in detail in the following.

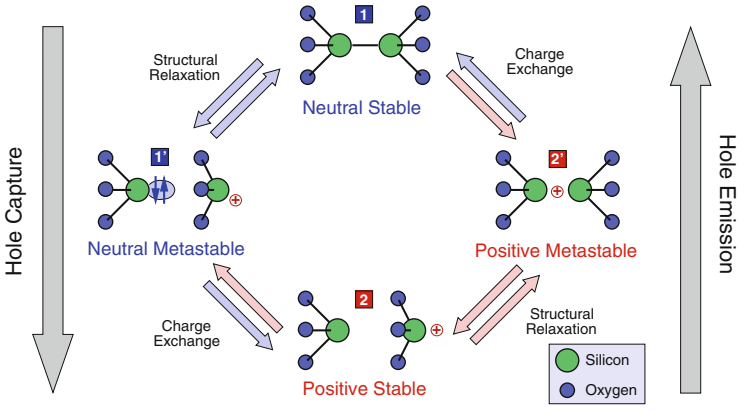

Fig. 16.10 State diagram of the multi-state model. The defect is present in a stable neutral (1) and a stable positive (2) charge state, where each of them has a second metastable state marked by a prime ($1'$, $2'$). The NMP transitions $1 \leftrightarrow 2'$ and $2 \leftrightarrow 1'$ occur between different charge states while the thermal transitions $1 \leftrightarrow 1'$ and $2 \leftrightarrow 2'$ between same charge states. Note that the transitions between the stable states are of main interest since they correspond to the experimentally measured capture and emission times in BTI. However, they involve intermediate states, which are metastable and important for the gate-bias and temperature dependence of the overall transition. The stick-and-ball models correspond to the configurations of a possible defect candidate, i.e., the oxygen vacancy, which is only shown for illustration purpose

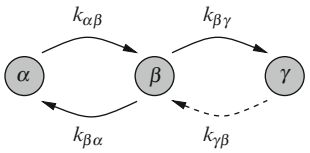

Fig. 16.11 The state diagram for a two-step process from the state α to γ. The first passage time of such a process is calculated by (16.59). Consider that the transition rate $k_{\gamma\beta}$, indicated by the *dashed arrow*, does not enter this equation

Such defects [55] show complex dynamics between those four states and must be correctly treated using homogeneous continuous-time Markov chain theory [56]. This theory rests upon the assumption that the future transitions between the states do not depend on the past of the investigated system. This assumption is justified as long as the defect relaxes after each transition by interacting with its environment, thereby losing the memory of its past. In fact, this is the case for both pure thermal and NMP transitions disregarding special theories, such as recombination-enhanced defect reaction. The time evolution of such a defect system is described by a first-order differential equation termed the Master equation.

$$\partial_t \pi_i(t) = \sum_{j \neq i} \pi_j(t) k_{ji} - \sum_{i \neq j} \pi_i(t) k_{ij} \tag{16.58}$$

Here, $\pi_i(t)$ is the time-dependent occupation probability that the defect is in state i and k_{ij} denotes the transition rate from state i to state j. When going from a single to a multitude of defects, the occupation probabilities must be averaged and become occupancies. The resulting rate equations, which are of the same form as the above Master equation, are usually solved in device simulators in order to predict the degradation for large area devices. Those kind of simulations can also account for the fact that the defect properties vary from trap to trap. The wide distributions of the defect properties arise from the amorphous defect environments but also come from the random dopant fluctuations, which have increasingly attracted scientific interest during the last several years [18, 57–62]. (For a detailed discussion of this topic, the interested reader is referred to [63].) For a comparison to the TDDS data, one is primarily interested in the transition times between stable states. The metastable states will only be occupied temporarily and are not observable in experiments. However, they gain their relevance for the overall gate bias and temperature dependence of two-step processes. The transitions between stable states are obtained from first-passage times. For a two-step process, the transition time from a state α to a state γ over a state β (cf. Fig. 16.11) reads

$$\tau_{\alpha\gamma} = \frac{k_{\alpha\beta} + k_{\beta\gamma} + k_{\beta\alpha}}{k_{\alpha\beta} k_{\beta\gamma}} = \frac{1}{k_{\alpha\beta}} + \frac{1}{k_{\beta\gamma}} + \frac{1}{k_{\beta\gamma}} \frac{k_{\beta\alpha}}{k_{\alpha\beta}}. \tag{16.59}$$

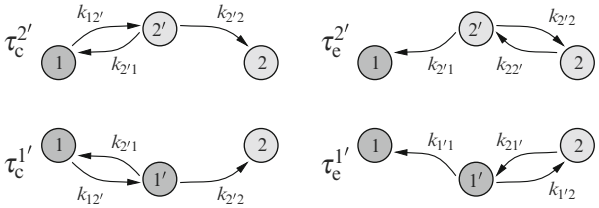

Fig. 16.12 Simplified state diagrams of hole capture and emission over the metastable states $1'$ and $2'$. The superscripts of τ denote the intermediate state, which has been passed through during a complete capture or emission event. Note that there exist two competing pathways for a hole capture event, namely one over the intermediate state $1'$ and one over $2'$. Of course, the same holds true for a hole emission event

The multi-state model with its four states allows for four distinct transition pathways (see Fig. 16.12), whose first-passage times are listed below:

$$\tau_c^{2'} = \frac{1}{k_{12'}} + \frac{1}{k_{2'2}} + \frac{1}{k_{2'2}}\frac{k_{2'1}}{k_{12'}} \tag{16.60}$$

$$\tau_c^{1'} = \frac{1}{k_{11'}} + \frac{1}{k_{1'2}} + \frac{1}{k_{1'2}}\frac{k_{1'1}}{k_{11'}} \tag{16.61}$$

$$\tau_e^{2'} = \frac{1}{k_{22'}} + \frac{1}{k_{2'1}} + \frac{1}{k_{2'1}}\frac{k_{2'2}}{k_{22'}} \tag{16.62}$$

$$\tau_e^{1'} = \frac{1}{k_{21'}} + \frac{1}{k_{1'1}} + \frac{1}{k_{1'1}}\frac{k_{1'2}}{k_{21'}} \tag{16.63}$$

The transition barriers for the partial rates can be extracted from the configuration coordinate diagram of the bistable defect (see Fig. 16.13). The bistability of the defect is reflected in the double-well shape of the adiabatic potentials. The transitions $T_{1\leftrightarrow 1'}$ and $T_{2\leftrightarrow 2'}$ are thermally activated and do not vary with the applied gate bias. According to transition state theory, they can be expressed as

$$k_{11'} = \nu_0 \exp(-\beta\varepsilon_{11'}) \tag{16.64}$$

$$k_{1'1} = \nu_0 \exp(-\beta\varepsilon_{1'1}) \tag{16.65}$$

$$k_{22'} = \nu_0 \exp(-\beta\varepsilon_{22'}) \tag{16.66}$$

$$k_{2'2} = \nu_0 \exp(-\beta\varepsilon_{2'2}) \tag{16.67}$$

where the barriers ε_{ij} are defined in Fig. 16.13 and ν_0 is the attempt frequency, which is typically of the order $10^{13}\,\mathrm{s}^{-1}$. The NMP transition rates are evaluated using (16.37)–(16.40), which contain lineshape functions and thus depend on V_s. The energy minima in the configuration coordinate diagram of Fig. 16.13 are given by

$$V_1 = \tilde{V}_0 - E \tag{16.68}$$

$$V_{2'} = \tilde{V}_0 + \varepsilon_{T2'} - E_t \tag{16.69}$$

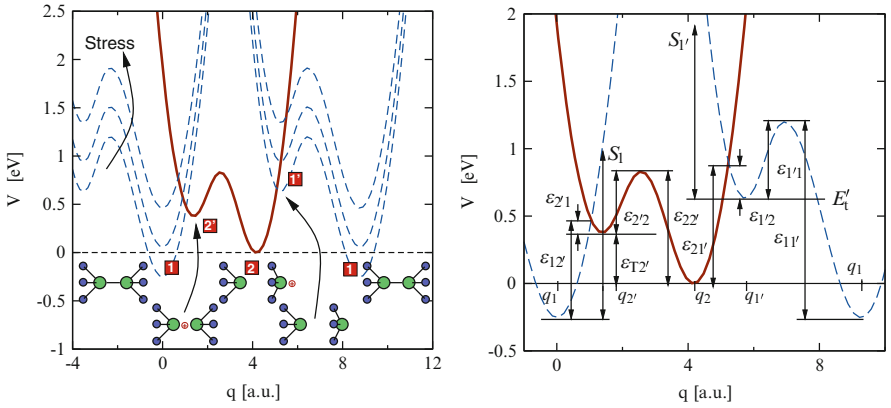

Fig. 16.13 *Left*: A schematic of the configuration coordinate diagram for a bistable defect. The *solid* and the *dashed lines* represent the adiabatic potentials for a defect in its positive and neutral charge state, respectively. The energy minima correspond to the stable or metastable defect configurations, labeled 1, 1′, 2, and 2′. The present configuration coordinate diagram describes the exchange of holes with the valence band and thus is associated with a hole capture or emission process. The stick-and-ball models display a defect in its various stable and metastable configurations for illustration purpose. *Right*: Definitions of the used energies and barriers in the multi-state model. Recall that two adiabatic potentials must be shown for one transition. It is assumed that an alternative transition pathway with an additional crossing point exists in the multidimensional atomic configuration space. In order to show both intersections (related to the transitions $1 \leftrightarrow 2'$ and $2 \leftrightarrow 1'$) in one configuration coordinate diagram, the "neutral" potential must be plotted twice. Obviously, $\varepsilon_{22'} = \varepsilon_{2'2} + \varepsilon_{T2'}$ holds

$$V_2 = \tilde{V}_0 - E'_t \tag{16.70}$$

$$V_{1'} = \tilde{V}_0 - E \tag{16.71}$$

in the hole picture. Here, the V_i stands for the adiabatic potentials with i being one of the states in Fig. 16.10. Furthermore, the hole is assumed to be energetically located at the valence band edge. It is emphasized that the energy $\varepsilon_{T2'}$ must be added to $\tilde{V}_0$ to obtain the correct energy minimum of state $2'$.

$$E_t \rightarrow E_t - \varepsilon_{T2'} \tag{16.72}$$

As a consequence, $\varepsilon_{T2'}$ modifies the energy differences V_s extracted from the configuration coordinate diagram

$$V_{12'} = V_{2'} - V_1 = E - E_t + \varepsilon_{T2'} \tag{16.73}$$

$$V_{1'2} = V_2 - V_{1'} = E - E'_t \tag{16.74}$$

and enters the NMP rates

$$k_{12'} = v_{\mathrm{th,n}}\sigma_{\mathrm{n}} \int_{E_{\mathrm{c}}}^{+\infty} D_{\mathrm{n}}(E) f_{\mathrm{p}}(E,E_{\mathrm{f}}) \lambda(E) f_{0/+}(c_0,c_+,q_{\mathrm{s}}, \underbrace{E-E_{\mathrm{t}}+\varepsilon_{\mathrm{T2'}}}_{=V_{12'}}) \mathrm{d}E$$
$$+ v_{\mathrm{th,p}}\sigma_{\mathrm{p}} \int_{-\infty}^{E_{\mathrm{v}}} D_{\mathrm{p}}(E) f_{\mathrm{p}}(E,E_{\mathrm{f}}) \lambda(E) f_{0/+}(c_0,c_+,q_{\mathrm{s}}, \underbrace{E-E_{\mathrm{t}}+\varepsilon_{\mathrm{T2'}}}_{=V_{12'}}) \mathrm{d}E \qquad (16.75)$$

$$k_{2'1} = v_{\mathrm{th,n}}\sigma_{\mathrm{n}} \int_{E_{\mathrm{c}}}^{+\infty} D_{\mathrm{n}}(E) f_{\mathrm{n}}(E,E_{\mathrm{f}}) \lambda(E) f_{+/0}(c_+,c_0,q_{\mathrm{s}}, \underbrace{E_{\mathrm{t}}-\varepsilon_{\mathrm{T2'}}-E}_{=-V_{12'}}) \mathrm{d}E$$
$$+ v_{\mathrm{th,p}}\sigma_{\mathrm{p}} \int_{-\infty}^{E_{\mathrm{v}}} D_{\mathrm{p}}(E) f_{\mathrm{n}}(E,E_{\mathrm{f}}) \lambda(E) f_{+/0}(c_+,c_0,q_{\mathrm{s}}, \underbrace{E_{\mathrm{t}}-\varepsilon_{\mathrm{T2'}}-E}_{=-V_{12'}}) \mathrm{d}E \qquad (16.76)$$

$$k_{1'2} = v_{\mathrm{th,n}}\sigma_{\mathrm{n}} \int_{E_{\mathrm{c}}}^{+\infty} D_{\mathrm{n}}(E) f_{\mathrm{p}}(E,E_{\mathrm{f}}) \lambda(E) f_{0/+}(c_0,c_+,q_{\mathrm{s}}, \underbrace{E-E'_{\mathrm{t}}}_{=V_{1'2}}) \mathrm{d}E$$
$$+ v_{\mathrm{th,p}}\sigma_{\mathrm{p}} \int_{-\infty}^{E_{\mathrm{v}}} D_{\mathrm{p}}(E) f_{\mathrm{p}}(E,E_{\mathrm{f}}) \lambda(E) f_{0/+}(c_0,c_+,q_{\mathrm{s}}, \underbrace{E-E'_{\mathrm{t}}}_{=V_{1'2}}) \mathrm{d}E \qquad (16.77)$$

$$k_{21'} = v_{\mathrm{th,n}}\sigma_{\mathrm{n}} \int_{E_{\mathrm{c}}}^{+\infty} D_{\mathrm{n}}(E) f_{\mathrm{n}}(E,E_{\mathrm{f}}) \lambda(E) f_{+/0}(c_+,c_0,q_{\mathrm{s}}, \underbrace{E'_{\mathrm{t}}-E}_{=-V_{1'2}}) \mathrm{d}E$$
$$+ v_{\mathrm{th,p}}\sigma_{\mathrm{p}} \int_{-\infty}^{E_{\mathrm{v}}} D_{\mathrm{p}}(E) f_{\mathrm{n}}(E,E_{\mathrm{f}}) \lambda(E) f_{+/0}(c_+,c_0,q_{\mathrm{s}}, \underbrace{E'_{\mathrm{t}}-E}_{=-V_{1'2}}) \mathrm{d}E\,. \qquad (16.78)$$

The above NMP transition rates along with the thermal transition rates (16.64)–(16.67) enter the expressions of the capture and emission times (16.60)–(16.63) that are comparable to time constants observed in the TDDS data. In the next section, they will be used to evaluate the multi-state model against the TDDS data and allow a verification of this model.

16.8 Model Evaluation

As outlined in Sect. 16.2, TDDS experiments measure the response of single defects to different gate voltage or temperature conditions. Based on these data, they give insight into the behavior of single defects and can thus reveal whether a BTI trapping

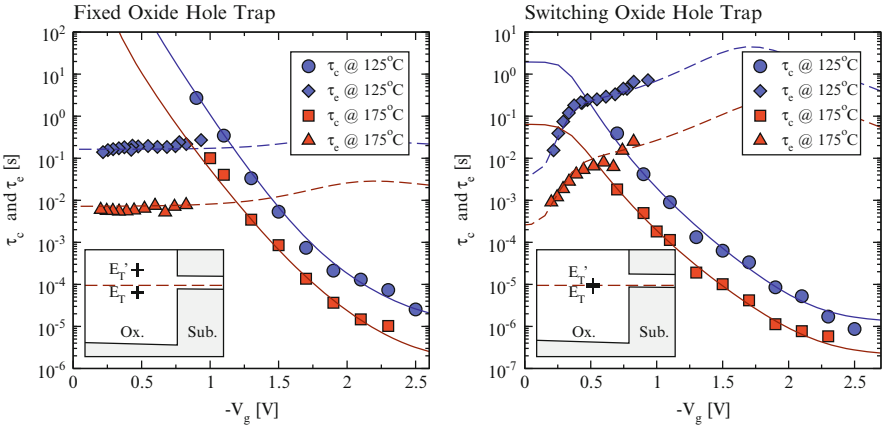

Fig. 16.14 *Left*: The capture (*solid lines*) and emission (*dashed lines*) times of a fixed oxide hole trap as a function of the gate bias. The symbols stand for the measurement data and the lines represent the simulation results of the multi-state model. The latter are shown to be in remarkable agreement with the experimental data. The inset (*bottom left*) depicts the band diagram of a MOSFET with the trap levels E_t and E_t' for the case when no bias is applied to the gate. Under these conditions the trap level E_t' is located far above the substrate Fermi level and the emission time remains unaffected by the gate bias. This fact eventually characterizes this defect as a fixed oxide hole trap *Right*: The same but for a switching oxide hole trap as presented in the Sect. 16.2. Compared to the fixed oxide hole trap, it shows a strong gate bias dependence of τ_e at small gate biases. In contrast to a fixed oxide hole trap, the Fermi level and the trap level E_t' coincide there, resulting in the strong sensitivity of τ_e to V_g

model reflects the physics of real defects. The time constant plots in Fig. 16.14 depict a fit of the multi-state model against the time constants extracted from the TDDS measurement data. The following calculations are carried out on a device simulator that delivers the band energy diagram for the devices used in the TDDS measurements. With these data, the thermal and the NMP transition rates were evaluated, which were subsequently used to calculate the capture and emission times. In these simulation, we accounted for the exchange of charge carriers with the substrate as well as the gate from the conduction and the valence band. An evaluation of the TDDS checklist is given below:

(i) The curvature in τ_c is reproduced by the multi-state model.
(ii) τ_c shows a marked temperature activation over the whole range of V_g, visible as a parallel upward shift.
(iii) In general, the multi-state model yields field-insensitive τ_e as displayed in Fig. 16.14 (left). It is important to note here that at larger oxide fields this model also predicts an exponential dependence, which has also been observed for some defects in RTN measurements [31].
(iv) The multi-state model also allows for a field-dependent τ_e provided that the substrate Fermi level and the trap level E_t' are separated by only a few hundredth of an electron Volt at small V_g (cf. Fig. 16.14, right).
(v) In both cases, τ_e is thermally activated.

The above checklist demonstrates that the multi-state model can reproduce the key features of the hole capture and emission process correctly, strongly indicating that the multi-state model can describe the physics of the defects seen in TDDS.

16.9 Discussion of the Multi-State Model

In Sect. 16.7, we derived a full set of rate equations that can accurately describe charge trapping within the multi-state model. However, they rely on complicated integrals which obscure the gate bias and temperature-dependent behavior of defects. For this reason, we also provide analytical expressions that promote understanding of the essential physical behind the mathematical framework.

Following the derivation in Sect. 16.6, the NMP transition rates can be written as

$$k_{12'} = \nu_{\mathrm{th,p}} \sigma_{\mathrm{p}} (1+R)^{3/2} \lambda(E_{\mathrm{v}}) p \exp(-\beta \varepsilon_{12'}) \tag{16.79}$$

$$k_{2'1} = \nu_{\mathrm{th,p}} \sigma_{\mathrm{p}} (1+R)^{3/2} \lambda(E_{\mathrm{v}}) p \exp(-\beta \varepsilon_{12'}) \exp(-\beta (E_{\mathrm{t}} - \varepsilon_{\mathrm{T}2'} - E_{\mathrm{f}})) \tag{16.80}$$

$$k_{1'2} = \nu_{\mathrm{th,p}} \sigma_{\mathrm{p}} (1+R')^{3/2} \lambda(E_{\mathrm{v}}) p \exp(-\beta \varepsilon_{1'2}) \tag{16.81}$$

$$k_{21'} = \nu_{\mathrm{th,p}} \sigma_{\mathrm{p}} (1+R')^{3/2} \lambda(E_{\mathrm{v}}) p \exp(-\beta \varepsilon_{1'2}) \exp(-\beta (E_{\mathrm{t}} - E_{\mathrm{f}})) \tag{16.82}$$

with

$$\varepsilon_{12'} = \frac{S_1 \hbar \omega_1}{(1+R_1)^2} + \frac{R_1}{1+R_1} (E_{\mathrm{v}} - E_{\mathrm{t}} + \varepsilon_{\mathrm{T}2'}) \tag{16.83}$$

$$= \frac{S_1 \hbar \omega_1}{(1+R_1)^2} - \frac{R_1}{1+R_1} (\Delta E_{\mathrm{t}} - \varepsilon_{\mathrm{T}2'}) + \frac{R_1}{1+R_1} q_0 x_{\mathrm{t}} F_{\mathrm{ox}} \tag{16.84}$$

$$\varepsilon_{1'2} = \frac{S_{1'} \hbar \omega_{1'}}{(1+R_{1'})^2} + \frac{R_{1'}}{1+R_{1'}} (E_{\mathrm{v}} - E'_{\mathrm{t}}) \tag{16.85}$$

$$= \frac{S_{1'} \hbar \omega_{1'}}{(1+R_{1'})^2} - \frac{R_{1'}}{1+R_{1'}} \Delta E'_{\mathrm{t}} + \frac{R_{1'}}{1+R_{1'}} q_0 x_{\mathrm{t}} F_{\mathrm{ox}} \tag{16.86}$$

using (16.3). In analogy to the derivation of the exact NMP transition rates (16.68)–(16.78), the trap level E_{t} must again be referenced to the minimum $2'$ according to (16.69). This reference of E_{t} is required in the calculation of the NMP barriers (16.84) and (16.86) as well as the last term of (16.80) following the concept outlined in Fig. 16.9. With the thermal transitions (16.64)–(16.67) and the above expression of the NMP rates (16.79)–(16.82), the capture and emission times (16.60)–(16.63) read

$$\tau_{\mathrm{c}}^{2'} = \tau_{\mathrm{c,min}}^{2'} + \tau_{\mathrm{p0}} \frac{N_2}{p} \exp\left(\beta \frac{R_1 q_0 x_{\mathrm{t}} F_{\mathrm{ox}}}{1+R_1}\right) + \tau_{\mathrm{c,min}}^{2'} \frac{N_1}{p} \exp(\beta q_0 x_{\mathrm{t}} F_{\mathrm{ox}}) \tag{16.87}$$

$$\tau_{\mathrm{c}}^{1'} = \tau_{\mathrm{c,min}}^{1'} + \tau_{\mathrm{p0}} \frac{N_3}{p} \exp\left(\beta \frac{R_{1'} q_0 x_{\mathrm{t}} F_{\mathrm{ox}}}{1+R_{1'}}\right) \tag{16.88}$$

$$\tau_{\mathrm{e}}^{2'} = \tau_{\mathrm{e,min}}^{2'} + \tau_{2'} \exp\left(-\beta \frac{q_0 x_{\mathrm{t}} F_{\mathrm{ox}}}{1+R_1}\right) \tag{16.89}$$

$$\tau_{\mathrm{e}}^{1'} = \tau_{1'} \exp\left(-\beta \frac{q_0 x_{\mathrm{t}} F_{\mathrm{ox}}}{1+R_{1'}}\right) + \tau_{\mathrm{e,min}}^{1'} \left(1+\exp\left(\beta(E_{\mathrm{t}}' - E_{\mathrm{f}})\right)\right) \tag{16.90}$$

using the definitions

$$N_1 = N_{\mathrm{v}} \exp\left(\beta(\varepsilon_{\mathrm{T2'}} - \Delta E_{\mathrm{t}})\right) \tag{16.91}$$

$$N_2 = \frac{N_{\mathrm{v}}}{(1+R_1)^{3/2}} \exp\left(\beta \frac{S_1 \hbar \omega_1}{(1+R_1)^2}\right) \exp\left(-\beta \frac{R_1(\Delta E_{\mathrm{t}} - \varepsilon_{\mathrm{T2'}})}{1+R_1}\right) \tag{16.92}$$

$$N_3 = \frac{N_{\mathrm{v}}}{(1+R_{1'})^{3/2}} \exp\left(\beta \frac{S_{1'} \hbar \omega_{1'}}{(1+R_{1'})^2}\right) \exp\left(-\beta \frac{R_{1'}}{1+R_{1'}} \Delta E_{\mathrm{t}}'\right) \times \left(1+\exp\left(\beta(\Delta E_{\mathrm{t}}' - \Delta E_{\mathrm{t}})\right)\right) \tag{16.93}$$

$$\tau_{2'} = \frac{\tau_{\mathrm{p0}}}{(1+R_1)^{3/2}} \exp\left(\beta \frac{S_1 \hbar \omega_1}{(1+R_1)^2}\right) \exp\left(\beta \frac{\Delta E_{\mathrm{t}} - \varepsilon_{\mathrm{T2'}}}{1+R_1}\right) \times \left(1+\exp(\beta \varepsilon_{\mathrm{T2'}})\right) \tag{16.94}$$

$$\tau_{1'} = \frac{\tau_{\mathrm{p0}}}{(1+R_{1'})^{3/2}} \exp\left(\beta \frac{S_{1'} \hbar \omega_{1'}}{(1+R_{1'})^2}\right) \exp\left(\beta \frac{\Delta E_{\mathrm{t}}'}{1+R_{1'}}\right) \tag{16.95}$$

$$\tau_{\mathrm{c,min}}^{2'} = 1/k_{2'2} \tag{16.96}$$

$$\tau_{\mathrm{e,min}}^{2'} = 1/k_{22'} \tag{16.97}$$

$$\tau_{\mathrm{c,min}}^{1'} = 1/k_{11'} \tag{16.98}$$

$$\tau_{\mathrm{e,min}}^{1'} = 1/k_{1'1} \tag{16.99}$$

$$\tau_{\mathrm{p0}} = \frac{1}{\sigma_{\mathrm{p}} v_{\mathrm{th,p}} N_{\mathrm{v}} \lambda(E_{\mathrm{v}})}\,. \tag{16.100}$$

Recall that the hole capture process can proceed from state 1 over one of the metastable states $2'$ or $1'$ to the final state 2 according to the state diagram of Fig. 16.12. The corresponding capture time constants are denoted as $\tau_{\mathrm{c}}^{2'}$ and $\tau_{\mathrm{c}}^{1'}$, respectively, and will be discussed in the following. If the transition pathway $T_{1\to 2'\to 2}$ is preferred, the capture time constant has the same shape as (16.59).

$$\tau_{\mathrm{c}}^{2'} = \frac{k_{12'} + k_{2'1} + k_{2'2}}{k_{12'} k_{2'2}} \tag{16.101}$$

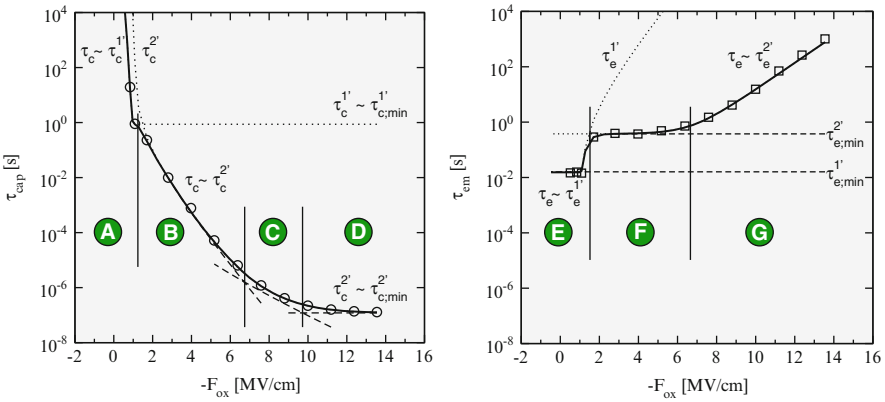

Fig. 16.15 *Left*: The calculated hole capture time constants as a function of the oxide field. The different regimes of τ_c (A, B, C, and D) are separated by the *thin vertical lines* and labeled by the *circles* with the capital letters. The *dotted curves* τ_c^i show the capture processes over a metastable state *i*. The field dependence of τ_c within a certain regime is shown by the *dashed curve*. *Right*: The same but for the hole emission time constants with the regimes (E, F, and G)

$$= \underbrace{\frac{1}{k_{12'}}}_{D} + \underbrace{\frac{1}{k_{2'2}}}_{C} + \underbrace{\frac{1}{k_{2'2}}\frac{k_{2'1}}{k_{12'}}}_{B} \tag{16.102}$$

Each summand in the nominator can be dominant, leading to (16.60), which is characterized by three distinct regimes, namely B, C, and D in Fig. 16.15. At extremely high negative oxide fields (regime D), $k_{12'}$ is the dominant rate meaning that the transition[4] $T_{1\to 2'}$ proceeds much faster than $T_{2'\to 2}$ (cf. Fig. 16.16). Thus complete capture process ($T_{1\to 2'\to 2}$) is controlled by the second transition $T_{2'\to 2}$, which is much slower and has a time constant of $\tau_{c,\min}^{2'}$. Since this second step is only thermally activated, $\tau_c^{2'}$ does not depend on the oxide field. This is consistent with (16.87), in which both exponential terms become negligible at extremely high negative oxide fields. At moderate negative oxide fields (regime C), the rate $k_{12'}$ approaches the order of $k_{2'1}$ and even falls below $k_{2'2}$. In this case the thermal transition $T_{2'\to 2}$ immediately follows the hole capture process from the state 1 to $2'$. As a result, the trapping kinetics are governed by the forward rate of the NMP process $T_{1\to 2'}$. Then $\tau_c^{2'}$ shows an exponential oxide field dependence, which is reflected in the second term of (16.87). At low negative oxide fields (regime B), $k_{12'}$ is already outbalanced by its reverse rate $k_{2'1}$ (see Fig. 16.16) and the ratio of both rates determines the oxide field dependence. This gives an increased exponential

[4]Keep in mind that the term "transition" does not refer to the duration of the physical process itself, such as the time it takes an electron to tunnel through an energy barrier. It rather denotes the mean time until the physical process takes place and the defect change its state.

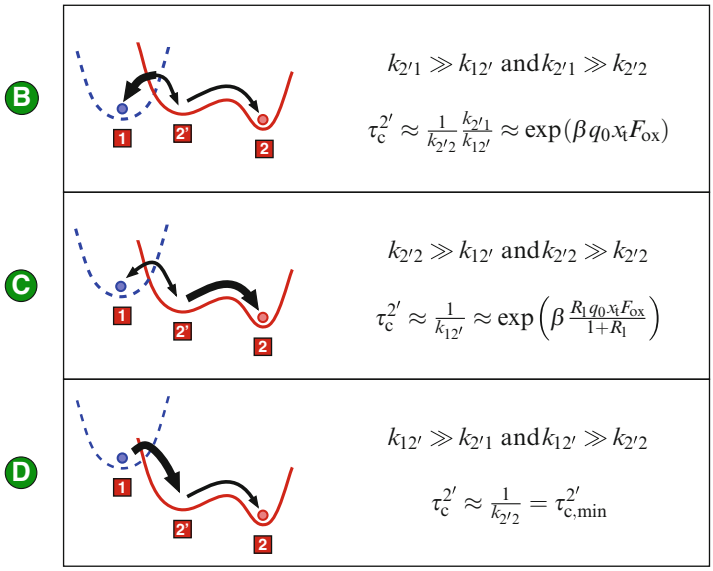

Fig. 16.16 A schematic representation of adiabatic potentials in the regimes B, C, and D. The *arrows* show the transitions involved in the hole capture process. Their thicknesses indicate the magnitude of their rates. This means that the *thinner arrows* are associated with larger transitions times and thus governs the oxide field and temperature dependence of the complete capture process $T_{1\to 2}$. With higher oxide fields (B $\to$ D) the potential of the neutral defect (*dashed line*) is raised relative to that of the positive defect (*solid line*). This is associated with an increase of $k_{12'}$ and a decrease of the reverse rate $k_{2'1}$. In contrast to the charge transfer reactions $T_{1\to 2'}$ and $T_{2\to' 1}$, the thermal transition $T_{2'\to 2}$ is not affected by the oxide field

slope originating from the third term of (16.87). *The transitions between these three regimes are smooth so that the capture time becomes curved in its time constant plots (cf. Fig. 16.16)*. It emphasized here that the curvature in the capture times are one of the most obstinate feature for BTI modeling and has only been reproduced by the multi-state model so far.

However, if the transition over the metastable state $1'$ is favored (regime A), the capture time constant can be formulated using first-passage times:

$$\tau_c^{1'} = \frac{k_{11'} + k_{1'1} + k_{1'2}}{k_{11'} k_{1'2}} \tag{16.103}$$

Since the metastable state $1'$ is situated above the state 1 by definition, $k_{1'1} \gg k_{11'}$ holds. Therefore, the expression (16.103) can be approximated by

$$\tau_c^{1'} \approx \underbrace{\frac{k_{1'1}}{k_{11'} k_{1'2}}}_{A''} + \underbrace{\frac{1}{k_{11'}}}_{A'}, \tag{16.104}$$

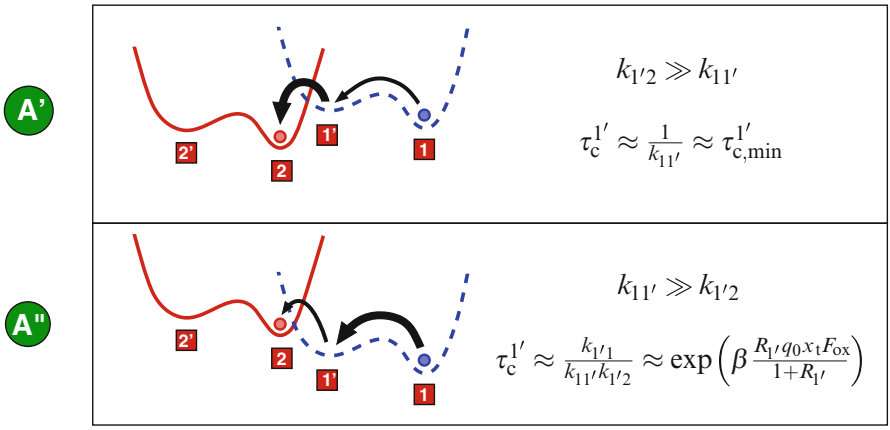

Fig. 16.17 The same as in Fig. 16.16 but for the regimes A' and A" of the oxide field dependence of $\tau_c^{1'}$

which is characterized by only two regimes (A' and A") now. At negative oxide fields (regime A'), the state $1'$ is located high (see Fig. 16.17) so that the transition rate $k_{1'2}$ is large compared to $k_{11'}$. Then the first term of expression (16.104) vanishes and the field-dependent transition $T_{1\to1'}$ with a time constant of $\tau_{c,min}^{1'}$ dominates $\tau_c^{1'}$ in (16.104). When reducing the oxide field, the state $1'$ is shifted downwards in the configuration coordinate diagram, thereby decreasing the transition rate $k_{1'2}$. At a certain oxide field, $k_{1'2}$ falls below $k_{1'1}$ and the first term of the expression (16.104) becomes dominant (regime A"). As a consequence, $\tau_c^{1'}$ is governed by the field-dependent transition $T_{1'\to2}$, which is reflected in the exponential term of the expression (16.88). The transition between A' and A" yields a kink, which is visible in $\tau_c^{1'}$ (cf. Fig. 16.15) but not in the overall hole capture τ_c time given by

$$\frac{1}{\tau_c} \approx \frac{1}{\tau_c^{1'}} + \frac{1}{\tau_c^{2'}} \,. \tag{16.105}$$

So far, this transition has not been observed in TDDS experiments, which is why the regimes A' and A" are not differentiated in Fig. 16.16.

Also the hole emission process has the possibility to proceed over either the state $1'$ or $2'$, with $\tau_e^{1'}$ and $\tau_e^{2'}$ being the corresponding emission time constants (see Fig. 16.18). For the transition pathway over $2'$, the emission time constant can be expressed as:

$$\tau_e^{2'} = \frac{k_{22'} + k_{2'2} + k_{2'1}}{k_{22'}k_{2'1}} \tag{16.106}$$

Since $k_{2'2} \gg k_{22'}$ applies, $\tau_e^{2'}$ has only two regimes, labeled with the capital letters F and G in Fig. 16.15.

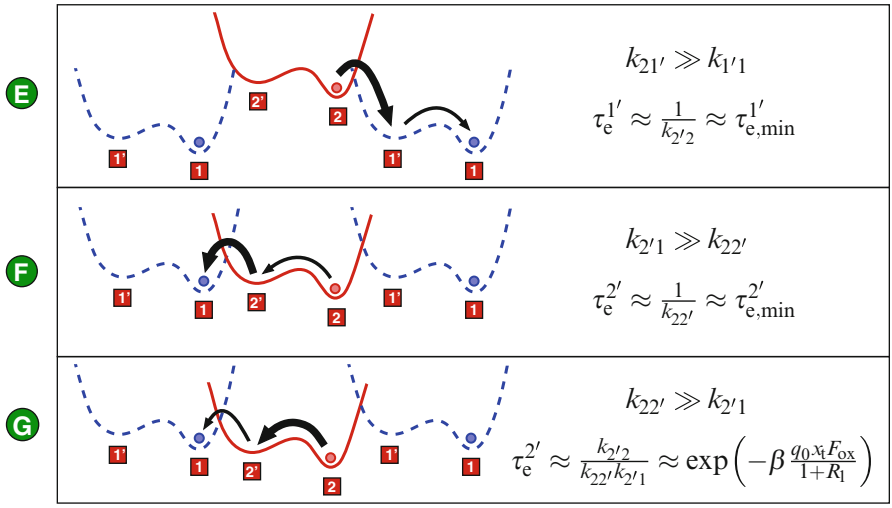

Fig. 16.18 The same as in Fig. 16.16 but for the regimes E, F, and G of the oxide field dependence of τ_e

$$\tau_e^{2'} \approx \underbrace{\frac{k_{2'2}}{k_{22'}k_{2'1}}}_{G} + \underbrace{\frac{1}{k_{22'}}}_{F} \tag{16.107}$$

At high negative oxide fields (regime G), the state 1 is shifted upwards so that $k_{22'}$ is the dominant rate and the field-dependent NMP transition $T_{2'\to 1}$ controls the transition $T_{1\to 2'\to 2}$. The oxide field dependence $T_{2'\to 1}$ is reflected in the second term of (16.89). At moderate negative oxide fields (regime F), the transition $T_{2'\to 1}$ proceeds much faster than $T_{2\to 2'}$. Thus, $\tau_e^{2'}$ is determined by the field-insensitive transition $T_{2\to 2'}$ with a time constant of $\tau_{e,min}^{2'}$. It is pointed out that *the regime F can give an explanation for the field-independent emission time constants observed for fixed oxide hole traps (cf. Fig. 16.14 left)*. This is a direct consequence of the assumed bistability of the defect in the multi-state model.

At a low oxide field (regime E), the state $1'$ is further shifted down, which speeds up the transition $T_{2\to 1'}$ and allows the pathway over the metastable state $1'$. The corresponding emission time constant $\tau_e^{1'}$ is then given by

$$\tau_e^{1'} = \frac{k_{21'} + k_{1'2} + r_{1'1}}{k_{21'}k_{1'1}} \,. \tag{16.108}$$

For a sufficiently large barrier $\varepsilon_{1'1}$, the rate $k_{1'1}$ becomes negligible compared to $k_{21'}$ and $k_{1'2}$ and the above equation simplifies to

$$\tau_e^{1'} = \frac{1}{k_{1'1}} + \frac{k_{1'2}}{k_{21'}k_{1'1}} \,. \tag{16.109}$$

In this case, the state diagram reduces to a subsystem that includes the states $1'$ and 2 and is marginally disturbed by the rate $k_{1'1}$. In this subsystem the states $1'$ and 2 can be assumed to be in quasi-equilibrium

$$f_{1'}k_{1'2} = f_2 k_{21'} \tag{16.110}$$

and the condition $f_{1'} + f_2 = 1$ is met. Then the trap occupancy $f'_t = f_{1'}$ is given by

$$f_{1'} = \frac{1}{1 + \frac{k_{21'}}{k_{1'2}}} = \frac{1}{1 + \exp(\beta(E'_t - E_f))} \,. \tag{16.111}$$

From this equation, it follows that the condition $k_{1'2} = k_{21'}$ is equivalent to $E'_t = E_f$. Furthermore, this equation can also be used to simplify (16.90) to

$$\tau_e^{1'} = \tau_{1'} \exp\left(-\beta \frac{q_0 x_t F_{ox}}{1 + R_{1'}}\right) + \frac{\tau_{e,min}^{1'}}{f_{t'}} \,. \tag{16.112}$$

If E'_t falls below E_f at a certain relaxation voltage, the state $1'$ becomes occupied and the emission time $\tau_e^{1'}$ is determined by the field-independent transition $T_{1'\to 1}$ with the time constant $\tau_{e,min}^{1'}$. By contrast, if E'_t is raised above E_f, the state $1'$ is underpopulated thereby slowing down the hole emission process. This occupancy effect is reflected in the second term, which is sensitive to changes in E_f.

The overall hole emission time τ_e follows approximately from

$$\frac{1}{\tau_e} \approx \frac{1}{\tau_e^{1'}} + \frac{1}{\tau_e^{2'}} \tag{16.113}$$

and is depicted in Fig. 16.15. At a certain oxide field, when the state $1'$ is shifted below state 2, $\tau_e^{1'}$ reaches its minimum value and falls below $\tau_e^{2'}$. *The resulting drop in τ_e is observed as the field dependence characterizing fixed oxide hole traps at weak oxide fields (cf. Fig. 16.14 right).* The drop in τ_e occurs when the minimum of the state $1'$ passes that of state 2, and is thus related to the exact shape of the configuration coordinate diagram. It is emphasized here that in the multi-state model the bistability of the defect allows for fixed as well as switching oxide hole traps while there is no explanation for these two kinds for defects in other models.

In summary, several features observed in the TDDS data have been quantitatively reproduced as shown in Sect. 16.8 and qualitatively understood following the above discussion based on analytical expressions. As such, this model can be regarded as a suited model to describe hole trapping in BTI.

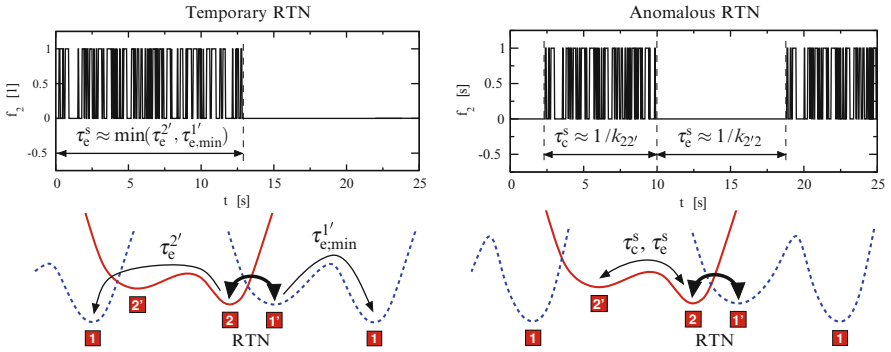

Fig. 16.19 *Top left*: The hole occupancy during tRTN. At $t = 0$ the stress voltage has been removed and the defect is in its positive state 2. After a time τ_e^s the defect ceases to produce noise. *Bottom left*: Configuration coordinate diagram for a tRTN defect. The *thick arrow* indicates the fast switches between the states 2 and $1'$ related to the occurrence of noise. The possibilities to escape from these states are shown by the *thin arrows*. *Top right*: Electron occupancy during aRTN. *Bottom right*: Configuration coordinate diagram for an aRTN defect. Since this defect is an electron trap, the *solid* and the *dashed line* correspond to the negative and the neutral charge state of the defect, respectively. The *double-sided thick arrow* is associated with aRTN while the *thin one* represents the transitions into and out of the metastable state $2'$

16.10 Noise

So far it has been shown that the multi-state model accounts for all features seen in the time constant plots for the fixed as well as the switching oxide hole traps. Beyond that, the model can also give an explanation for tRTN observed in TDDS (see Sect. 16.2). The generated noise stems from defects switching back and forth between states 2 and $1'$. The associated charge transfer reactions $T_{2\leftrightarrow 1'}$ do not involve any intermediate states and are therefore simple NMP processes. It is remarked here that the transitions $T_{2\leftrightarrow 1'}$ require the energy minima 2 and $1'$ in the configuration coordinate diagram to be on approximately the same level at the relaxation voltage. This is only the case for a group of defects whose energy minima 1 and $1'$ are energetically not far separated. In the TDDS measurements, the investigated devices are stressed at a high V_g so that the defects are forced from the state 1 into the state 2 or $1'$. During this step, the defects undergo the transition $T_{1\rightarrow 2'\rightarrow 2}$ into the state 2 or even further into $1'$. The other direct pathway $T_{1\rightarrow 1'}$ into the state $1'$ or 2 is assumed to go over a large barrier $\varepsilon_{11'}$. Therefore, the transition $T_{1\rightarrow 1'}$ proceeds on much larger timescales compared to $T_{1\rightarrow 2'\rightarrow 2}$ and can be neglected. After stressing, the recovery traces are monitored at low V_g or F_{ox}, respectively, at which the energy minima of the states 2 and $1'$ coincide and noise is produced. However, the state 1 is thermodynamically preferred due to its energetically lower position compared to the states 2 and $1'$. When the defect returns to its initial state 1, the RTN signal disappears with a time constant of τ_e^s. The corresponding transition could be either $T_{2\rightarrow 2'\rightarrow 1}$ or $T_{1'\rightarrow 1}$ with a time constant of $\tau_e^{2'}$

or $\tau_{e,min}^{1'}$, respectively (cf. Fig. 16.19). The termination of the noise signal after a time period of τ_e^s is determined by the minimum of these time constants. Consider that the NMP barriers $\varepsilon_{21'}$ and $\varepsilon_{1'2}$ must not be too large since otherwise trapping events will occur too fast and are therefore not detected using a conventional measurement equipment.

Interestingly, there also exists a type of defect which repeatedly produces noise for stochastically distributed time intervals (see Sect. 16.2). This kind of noise was observed for electron traps [27] in nMOSFETS and is referred to as aRTN. Just as in the case of tRTN, the noise signal is generated by charge transfer reactions between the states 2 and $1'$. The recurrent pauses of the noise signal (see Fig. 16.19) originate from transitions into the metastable state $2'$, which is electrically indistinguishable from the state 2. These interruptions correspond to the time during which the defect dwells in this state and no charge transfer reaction can take place. Thereby it has been presumed that the NMP transition $T_{2'\to 1}$ occurs on larger time scales than the return to the state 2 through the transition $T_{2'\to 2}$. The slow capture time constant τ_c^s in Fig. 16.19 defines the mean time interval during which noise is observed. Its value is given by the inverse of the transition rate $1/k_{22'}$. The slow emission time constant $\tau_e^s = 1/k_{2'2}$ corresponds to the mean time interval until the next noise period starts.

One should keep in mind that when adopting the concept of aRTN to hole traps in pMOSFET, it may also explain the tRTN behavior seen in TDDS measurements. During TDDS stress, this sort of defects are forced into one of the states 2 and $1'$ where they produce an RTN signal. As in aRTN, they undergo a transition to the metastable state $2'$ thereby stopping to produce a noise signal. However, this special sort of defects is characterized by a slow emission time constant τ_e^s, which is much larger than the typical measurement time of TDDS. As a consequence, the next transition back to the state 2 and the subsequent noise period are shifted out of the experimental time window of TDDS and will not be recorded during the measurement run. According to this explanation, tRTN can also be explained as a stimulated variant of aRTN.

In summary, the multi-state model can account for the features from the time constant plots and is consistent with the observation of tRTN as well as aRTN. This fact is presented here since it is regarded as an additional support for the validity of this model.

16.11 Conclusion

With the departure from the established reaction–diffusion model, charge trapping in BTI has recently attracted scientific interest. Therefore, the nature of charge trapping has remained vaguely understood for a long time and has been intensively studied within our group. In this chapter we presented a detailed derivation of our charge trapping model, termed multi-state model, in which the focus was on correctly modeling microscopic processes involved in BTI. In order to support understanding

of the tendencies in this model, we have also given analytical expressions, which still capture the main physics underlying charge trapping in BTI.

For the verification of our model, we have chosen the TDDS technique since it allows to analyze the behavior of single defects. The evaluation of our multi-state model was based on five criteria including the curvature in the capture times, the gate bias and temperature dependences, and the fixed as well as the switching oxide hole trap behavior. So far, all these features have only been reproduced by the multi-state model, which strongly indicates that this model is based on correct assumptions. Interestingly, the model gives also an explanation for temporary and anomalous RTN, thereby further corroborating its validity.

Acknowledgements This work has received funding from the Austrian Science Fund (FWF) project n° 23390-N24 and the European Communities FP7 n° 261868 (MORDRED).

References

1. K.O. Jeppson and C.M. Svensson, "Negative Bias Stress of MOS Devices at High Electric Fields and Degradation of MNOS Devices," *J.Appl.Phys.*, vol. 48, no. 5, pp. 2004–2014, 1977.
2. S. Ogawa and N. Shiono, *Phys.Rev.B*, vol. 51, no. 7, pp. 4218–4230, 1995.
3. B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, and M. Goodwin, "Disorder-Controlled-Kinetics Model for Negative Bias Temperature Instability and its Experimental Verification," in *Proc.IRPS*, 2005, pp. 381–387.
4. S. Zafar, "Statistical Mechanics Based Model for Negative Bias Temperature Instability Induced Degradation," *J.Appl.Phys.*, vol. 97, no. 10, pp. 1–9, 2005.
5. M. Houssa, M. Aoulaiche, S. De Gendt, G. Groeseneken, M.M. Heyns, and A. Stesmans, "Reaction-Dispersive Proton Transport Model for Negative Bias Temperature Instabilities," *Appl.Phys.Lett.*, vol. 86, no. 9, pp. 1–3, 2005.
6. M.A. Alam, H. Kufluoglu, D. Varghese, and S. Mahapatra, "A Comprehensive Model for pMOS NBTI Degradation: Recent Progress," *Microelectron.Reliab.*, vol. 47, no. 6, pp. 853–862, 2007.
7. S. Chakravarthi, A.T. Krishnan, V. Reddy, C.F. Machala, and S. Krishnan, "A Comprehensive Framework for Predictive Modeling of Negative Bias Temperature Instability," in *Proc.IRPS*, 2004, pp. 273–282.
8. T. Grasser, W. Goes, and B. Kaczer, "Towards Engineering Modeling of Negative Bias Temperature Instability," in *Defects in Microelectronic Materials and Devices*, D. Fleetwood, R. Schrimpf, and S. Pantelides, Eds., pp. 1–30. Taylor and Francis/CRC Press, 2008.
9. F. Schanovsky and T. Grasser (2013) On the microscopic limit of the RD model. In: T. Grasser (eds) Bias temperature instability for devices and circuits. Springer, New York
10. F. Schanovsky and T. Grasser, "On the Microscopic Limit of the Reaction-Diffusion Model for Negative Bias Temperature Instability," in *Proc.IIRW*, 2011, pp. 17–21.
11. F. Schanovsky and T. Grasser, "On the Microscopic Limit of the Modified Reaction-Diffusion Model for Negative Bias Temperature Instability," in *Proc.IRPS*, 2012, pp. XT.10.1–6.
12. B. Kaczer, T. Grasser, J. Martin-Martinez, E. Simoen, M. Aoulaiche, Ph.J. Roussel, and G. Groeseneken, "NBTI from the Perspective of Defect States with Widely Distributed Time Scales," in *Proc.IRPS*, 2009, pp. 55–60.
13. H. Reisinger, T. Grasser, and C. Schlünder, "A Study of NBTI by the Statistical Analysis of the Properties of Individual Defects in pMOSFETs," in *Proc.IIRW*, 2009, pp. 30–35.

14. T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, *Phys.Rev.B*, vol. 82, no. 24, pp. 245318, 2010.
15. P.-J. Wagner, T. Grasser, H. Reisinger, and B. Kaczer, "Oxide Traps in MOS Transistors: Semi-Automatic Extraction of Trap Parameters from Time Dependent Defect Spectroscopy," in *Proc.IPFA*, 2010, pp. 249–254.
16. T. Grasser, H. Reisinger, P.-J. Wagner, F. Schanovsky, W. Goes, and B. Kaczer, "The Time Dependent Defect Spectroscopy (TDDS) for the Characterization of the Bias Temperature Instability," in *Proc.IRPS*, 2010, pp. 16 –25.
17. H. Reisinger, T. Grasser, W. Gustin, and C. Schlünder, "The Statistical Analysis of Individual Defects Constituting NBTI and its Implications for Modeling DC- and AC-Stress," in *Proc.IRPS*, 2010, pp. 7–15.
18. B. Kaczer, T. Grasser, Ph.J. Roussel, J. Franco, R. Degraeve, L.A. Ragnarsson, E. Simoen, G. Groeseneken, and H. Reisinger, "Origin of NBTI Variability in Deeply Scaled PFETs," in *Proc.IRPS*, 2010, pp. 1095–1098.
19. H. Reisinger (2013) The time dependent defect spectroscopy. In: T. Grasser (eds) Bias temperature instability for devices and circuits. Springer, New York
20. V. Huard, C. Parthasarathy, N. Rallet, C. Guerin, M. Mammase, D. Barge, and C. Ouvrard, "New Characterization and Modeling Approach for NBTI Degradation from Transistor to Product Level," in *Proc.IEDM*, 2007, pp. 797–800.
21. T.L. Tewksbury, *Relaxation Effects in MOS Devices due to Tunnel Exchange with Near-Interface Oxide Traps*, Ph.D. Thesis, MIT, 1992.
22. I. Lundstrom and C. Svensson, "Tunneling to Traps in Insulators," *J.Appl.Phys.*, vol. 43, no. 12, pp. 5045–5047, 1972.
23. F.P. Heiman and G. Warfield, "The Effects of Oxide Traps on the MOS Capacitance," *IEEE Trans.Elect.Dev.*, vol. 12, no. 4, pp. 167–178, 1965.
24. S. Christensson, I. Lundström, and C. Svensson, "Low Frequency Noise in MOS Transistors — I Theory," *Sol.-St.Electr.*, vol. 11, pp. 797–812, 1968.
25. W. Shockley and W.T. Read, *Phys.Rev.*, vol. 87, no. 5, pp. 835–842, 1952.
26. A.L. McWhorter, "1/f Noise and Germanium Surface Properties," in *Sem.Surf.Phys*. RH Kingston (Univ Penn Press), 1957.
27. M.J. Kirton and M.J. Uren, "Noise in Solid-State Microstructures: A New Perspective on Individual Defects, Interface States, and Low-Frequency (1/f) Noise," *Adv.Phys.*, vol. 38, no. 4, pp. 367–486, 1989.
28. M. Masuduzzaman, A.E. Islam, and M.A. Alam, "Exploring the Capability of Multifrequency Charge Pumping in Resolving Location and Energy Levels of Traps Within Dielectric," *IEEE Elect.Dev.Let.*, vol. 55, no. 12, pp. 3421–3431, 2008.
29. A. Avellan, D. Schroeder, and W. Krautschneider, "Modeling Random Telegraph Signals in the Gate Current of Metal-Oxide-Semiconductor Field Effect Transistors after Oxide Breakdown," *J.Appl.Phys.*, vol. 94, no. 1, pp. 703–708, 2003.
30. M. Isler and D. Liebig, *Phys.Rev.B*, vol. 61, no. 11, pp. 7483–7488, 2000.
31. N. Zanolla, D. Siprak, P. Baumgartner, E. Sangiorgi, and C. Fiegna, "Measurement and Simulation of Gate Voltage Dependence of RTS Emission and Capture Time Constants in MOSFETs," in *Ultimate Integration of Silicon*, 2008, pp. 137–140.
32. R.R. Siergiej, M.H. White, and N.S. Saks, "Theory and Measurement of Quantization Effects on $Si - SiO_2$ Interface Trap Modeling," *Sol.-St.Electr.*, vol. 35, no. 6, pp. 843–854, 1992.
33. N.B. Lukyanchikova, M.V. Petrichuk, N.P. Garbar, E. Simoen, and C. Claeys, "Influence of the Substrate Voltage on the Random Telegraph Signal Parameters in Submicron *n*-Channel Metal-Oxide-Semiconductor Field-Effect Transistors under a Constant Inversion Charge Density," *Appl.Phys.A*, vol. 70, no. 3, pp. 345–353, 2000.
34. S. Makram-Ebeid and M. Lannoo, *Phys.Rev.B*, vol. 25, no. 10, pp. 6406–6424, 1982.
35. S.D. Ganichev, W. Prettl, and I.N. Yassievich, "Deep Impurity-Center Ionization by Far-Infrared Radiation," *Phys.Solid State*, vol. 39, no. 1, pp. 1703–1726, 1997.
36. S.D. Ganichev, I.N. Yassievich, V.I. Perel, H. Ketterl, and W. Prettl, *Phys.Rev.B*, vol. 65, pp. 085203, 2002.

37. K. Huang and A. Rhys, "Theory of Light Absorption and Non-Radiative Transitions in F-Centres," *Proceedings of the Royal Society of London. Series A*, vol. 204, pp. 406–423, 1950.
38. C.H. Henry and D.V. Lang, *Phys.Rev.B*, vol. 15, no. 2, pp. 989–1016, 1977.
39. K.V. Mikkelsen and M.A. Ratner, "Electron Tunneling in Solid-State Electron-Transfer Reactions," *Chemical Reviews*, vol. 87, no. 1, pp. 113–153, 1987.
40. Conley and Lenahan, "Electron Spin Resonance Evidence that E'_γ Centers Can Behave as Switching Traps," *IEEE Trans.Nucl.Sci.*, vol. 42, no. 6, pp. 1744–1749, 1995.
41. J.F. Conley Jr., P.M. Lenahan, A.J. Lelis, and T.R. Oldham, "Electron Spin Resonance Evidence for the Structure of a Switching Oxide Trap: Long Term Structural Change at Silicon Dangling Bond Sites in SiO_2," *Appl.Phys.Lett.*, vol. 67, no. 15, pp. 2179–2181, 1995.
42. A.J. Lelis and T.R. Oldham, "Time Dependence of Switching Oxide Traps," *IEEE Trans.Nucl.Sci.*, vol. 41, no. 6, pp. 1835–1843, 1994.
43. M. Lax, "The Franck-Condon Principle and Its Application to Crystals," *Journ.Chem.Phys.*, vol. 20, no. 11, pp. 1752–1760, 1952.
44. T.H. Keil, *Phys.Rev.*, vol. 140, no. 2A, pp. A601–A617, 1965.
45. F. Schanovsky, O. Baumgartner, V. Sverdlov, and T. Grasser, "A Multi Scale Modeling Approach to Non-Radiative Multi Phonon Transitions at Oxide Defects in MOS Structures," *Journ. of Computational Electronics*, vol. 11, no. 3, pp. 218–224, 2012.
46. S. Datta, *Quantum Transport — Atom to Transistor*, Cambridge University Press, 2005.
47. M.O. Andersson, Z. Xiao, S. Norrman, and O. Engström, "Model Based on Trap-Assisted Tunneling for Two-Level Current Fluctuations in Submicrometer Metal-Silicon-Dioxide Diodes," *Phys.Rev.B*, vol. 41, no. 14, pp. 9836–9842, 1990.
48. P.E. Blöchl and J.H. Stathis, "Hydrogen Electrochemistry and Stress-Induced Leakage Current in Silica," *Phys.Rev.Lett.*, vol. 83, no. 2, pp. 372–375, 1999.
49. P.E. Blöchl and J.H. Stathis, "Aspects of Defects in Silica Related to Dielectric Breakdown of Gate Oxides in MOSFETs," *Phys.B*, vol. 273-274, pp. 1022–1026, 1999.
50. W.B. Fowler, J.K. Rudra, M.E. Zvanut, and F.J. Feigl, *Phys.Rev.B*, vol. 41, no. 12, pp. 8313–8317, 1990.
51. W. Goes and T. Grasser, "First-Principles Investigation on Oxide Trapping," in *Proc.SISPAD*, 2007, pp. 157–160.
52. W. Goes and T. Grasser, "Charging and Discharging of Oxide Defects in Reliability Issues," in *Proc.IIRW*, 2007, pp. 27–32.
53. W. Goes, M. Karner, V. Sverdlov, and T. Grasser, "Charging and Discharging of Oxide Defects in Reliability Issues," *IEEE Trans.Dev.Mater.Rel.*, vol. 8, no. 3, pp. 491–500, 2008.
54. A. Alkauskas and A. Pasquarello, "Alignment of Hydrogen-Related Defect Levels at the $Si - SiO_2$ Interface," *Phys.B Condens.Matter*, vol. 401–402, pp. 546–549, 2007.
55. T. Grasser, "Stochastic Charge Trapping in Oxides: From Random Telegraph Noise to Bias Temperature Instabilities," *Microelectron.Reliab.*, vol. 52, no. 1, pp. 39–70, 2012.
56. O.C. Ibe, *Markov Processes for Stochastic Modeling*, Academic Press, 2009.
57. M. Bina, O. Triebl, B. Schwarz, M. Karner, B. Kaczer, and T. Grasser, "Simulation of Reliability on Nanoscale Devices," in *Proc.SISPAD*, 2012, pp. 109–112.
58. B. Kaczer, J. Franco, M. Toledano-Luque, Ph.J. Roussel, M.F. Bukhori, A. Asenov, B. Schwarz, M. Bina, T. Grasser, and G. Groeseneken, "The Relevance of Deeply-Scaled FET Threshold Voltage Shifts for Operation Lifetimes," in *Proc.IRPS*, 2012.
59. A. Asenov, "Random Dopant-Induced Threshold Voltage Lowering and Fluctuations in Sub-0.1 μm MOSFET's: A 3-D "Atomistic" Simulation Study," *IEEE Trans.Elect.Dev.*, vol. 45, no. 12, pp. 2505–2513, 1998.
60. A. Mauri, N. Castellani, C.M. Compagnoni, A. Ghetti, P. Cappelletti, A.S. Spinelli, and A.L. Lacaita, "Impact of Atomistic Doping and 3D Electrostatics on the Variability of RTN Time Constants in Flash Memories," in *Proc.IEDM*, 2011, pp. 17.1.1–17.1.4.
61. N. Sano, K. Matsuzawa, M. Mukai, and N. Nakayama, "On Discrete Random Dopant Modeling in Drift-Diffusion Simulations: Physical Meaning of "Atomistic" Dopants," *Microelectron.Reliab.*, vol. 42, no. 2, pp. 189–199, 2002.

62. A. Asenov, G. Slavcheva, A.R. Brown, J.H. Davies, and S. Saini, “Increase in the Random Dopant-Induced Threshold Fluctuations and Lowering in Sub-100 nm MOSFETs due to Quantum Effects: a 3-D Density-Gradient Simulation Study,” *IEEE Trans.Elect.Dev.*, vol. 48, no. 4, pp. 722–729, 2001.
63. S.M. Amoroso (2013) Statistical study of bias temperature instabilities by means of 3D ‘atomistic’ simulation. In: T. Grasser (eds) Bias temperature instability for devices and circuits. Springer, New York