# Chapter 15
# On the Microscopic Limit of the RD Model

**Franz Schanovsky and Tibor Grasser**

**Abstract** The popular reaction–diffusion model for the negative bias temperature instability is discussed from the viewpoint of stochastic chemical kinetics. We present a microscopic formulation of the reaction–diffusion model based on the reaction–diffusion master equation and solve it using the stochastic simulation algorithm. The calculations are compared to the macroscopic version as well as established experimental data. The degradation predicted by the microscopic reaction–diffusion model strongly deviates from the macroscopic version and the experimentally observed behavior. Those deviations are explained as necessary consequences of the physical processes involved. The presented results show the impact of the unphysical assumptions in the reaction–diffusion model. Further, we generally question the suitability of the mathematical framework of reaction rate equations for a reactive-diffusive system at the given particle densities.

## 15.1 Introduction

The first model for NBTI was put forward by Jeppson and Svensson in 1977 [1]. This model was based on the following ideas, which are illustrated in Fig. 15.1. Due to the lattice mismatch between silicon and silicon dioxide, some of the silicon atoms do not have an oxygen neighbor. A silicon atom in this situation has one unpaired valence electron, which is called a dangling bond. This dangling bond is visible in electronic measurements as it gives rise to states within the band-gap [2]. During the manufacturing process the wafer is exposed to a hydrogen-rich atmosphere so that hydrogen atoms can penetrate through the oxide and passivate the silicon dangling bonds, leading to a removal of the band-gap states.

F. Schanovsky (✉) • T. Grasser
Institute for Microelectronics, Gußhausstraße 27-29/E360, 1040 Vienna, Austria
e-mail: 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_15, 

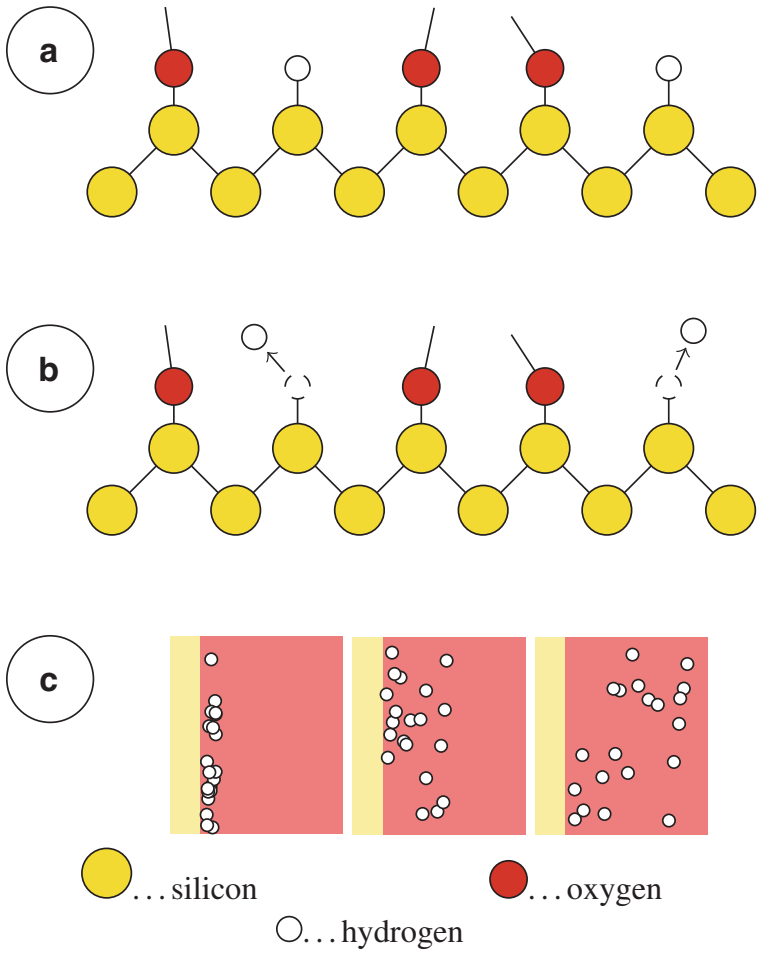


**Fig. 15.1** The basic concept behind the reaction–diffusion model for NBTI. (**a**) Silicon dangling bonds at the Si–$SiO_2$ interface are initially passivated by hydrogen atoms. (**b**) During stress, hydrogen atoms are liberated leaving behind the unpassivated silicon dangling bonds which degrade the device properties. (**c**) The time evolution is determined by the depopulation of the interface due to the flux of hydrogen into the oxide

During stress, the presence of holes at the interface and the increased temperature leads to a liberation of the hydrogen atoms. The remaining silicon dangling bonds become electrically active carrier traps. According to the model, the depassivation and repassivation of dangling bonds at the interface reach an equilibrium in a very short time [3, 4], and it is the constant flux of hydrogen atoms (or some hydrogenic species) away from the interface that determines the temporal evolution of the degradation. Because of the two proposed stages—the electrochemical reaction at the interface and the subsequent diffusion of the hydrogenic species—this model bears the name *reaction–diffusion* (RD) model.

The mathematical framework of the model is based on a macroscopic description using a rate equation for the interface reaction and a Fickian diffusion equation for the motion of the hydrogen in the oxide. Central actors are the density of depassivated silicon dangling bonds at the interface $N_{\mathrm{it}} = [\mathrm{Si}^*]$, and the concentration of hydrogen in the oxide $H = [\mathrm{H}](x,t)$ and at the interface $H_{\mathrm{it}} = [\mathrm{H}](0,t)$.

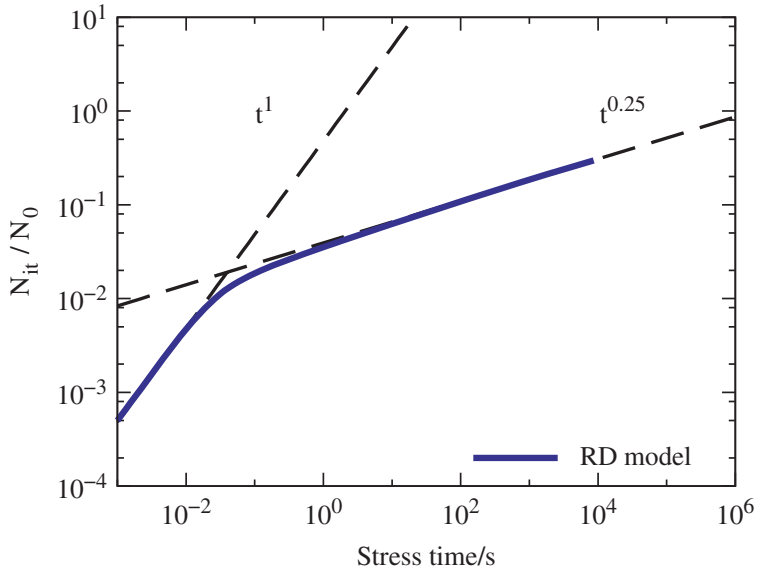


**Fig. 15.2** Basic features of the degradation predicted by the RD model for NBTI. In the initial phase, the depassivation reaction with rate $k_f$ dominates, giving rise to a degradation that increases linearly with time. After the depassivation and repassivation reactions have reached an equilibrium, the degradation is determined by the flux of hydrogen, which gives rise to a power-law with an exponent of $1/4$

During degradation, a fraction $N_{it}$ of the initially passivated silicon dangling bonds $N_0 = [\mathrm{SiH}]_0$ is depassivated according to

$$\frac{\partial N_{it}}{\partial t} = k_f(N_0 - N_{it}) - k_r N_{it} H_{it}, \tag{15.1}$$

with the depassivation (forward) rate $k_f$ and the repassivation (reverse) rate $k_r$. The hydrogen liberated at the interface then diffuses into the oxide as

$$\frac{\partial H}{\partial t} = -D\frac{\partial^2 H}{\partial x^2} \tag{15.2}$$

with the diffusion coefficient $D$. The RD model became popular among reliability engineers as it features a simple mathematical description and a small set of parameters which have a sound physical interpretation. Most importantly, as shown in Fig. 15.2, this model predicts a constant-stress degradation that initially grows linearly with time and then follows a power-law of the form [3, 4]

$$N_{it}(t) = \sqrt{\frac{k_f N_0}{2k_r}}(Dt)^{1/4}. \tag{15.3}$$

This power-law degradation corresponded well with experimental results of the 1970s.

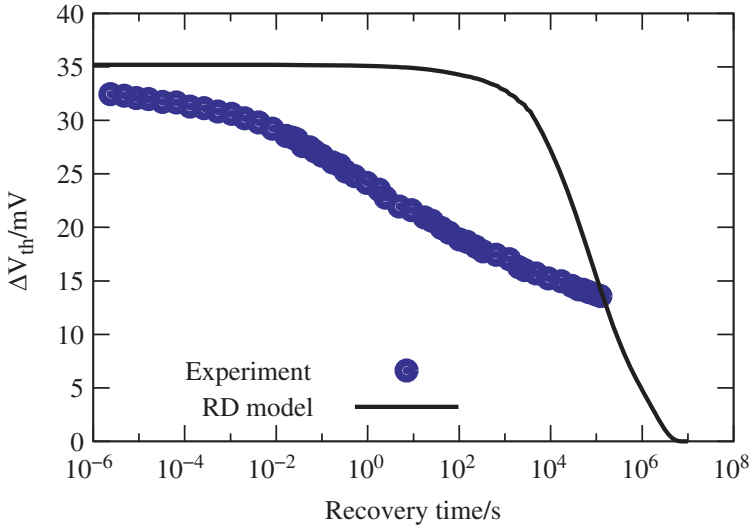


**Fig. 15.3** Typical recovery trace as predicted by the RD model for NBTI using (15.4)–(15.6), which is similar for all variants of the RD mechanism. The comparison with experimental data [7] shows that the RD predicted recovery occurs much too late and proceeds much too fast

In later experiments, power-law exponents were found that differed from the $1/4$ prediction of the model. These findings led to a modification of the original RD model to account for different diffusing species such as $H_2$ [5]. For almost four decades, the reaction–diffusion idea was the unquestioned standard interpretation for NBTI until around 2005 NBT recovery moved into the focus of the scientific attention. The experiments showed that NBTI recovery starts immediately (even before a microsecond) after the removal of stress and extended over several decades, continuing even after more than $10^5$ s [6, 7]. This behavior stands in strong contrast to the predictions of the RD model, which predicts a recovery that proceeds within four decades, centering around the duration of the preceding stress phase [8, 9]. A comparison of a typical experimental NBT recovery trace and the corresponding prediction of the RD model is shown in Fig. 15.3. Several extensions to the RD model have been put forward, such as dispersive transport of the hydrogenic species [4, 6], but none could give the observed experimental behavior. The current state-of-the-art RD-based modeling supplements the RD theory with empirical hole-trapping expressions. It is assumed that short-time (1 s) degradation and recovery is dominated by hole trapping into oxide and interface defects, while the long-term degradation and recovery are determined by the RD mechanism [10–13]. The RD theory employed in these modeling efforts is the *modified* RD model [14–16] that has been developed as an extension of the classical RD models and explicitly considers diffusion of H and $H_2$ and their interconversion reactions. Classical models assume an instantaneous transition between the liberated interfacial hydrogen and the diffusing species, usually $H_2$ [5]. The reactions present in the modified RD model are the interface reaction $\mathrm{SiH} \rightleftharpoons \mathrm{Si}^* + \mathrm{H}$, the dimerization reaction $2\mathrm{H} \rightleftharpoons \mathrm{H}_2$, and the diffusion of both species. The mathematical framework is an extension of (15.1) and (15.2) [15, 16],

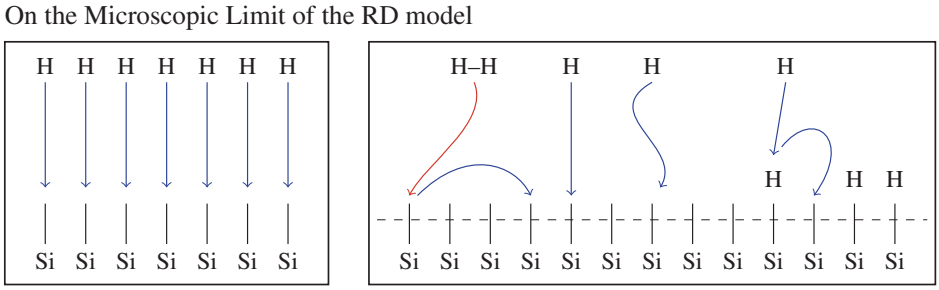


**Fig. 15.4** (*Left*) According to Mahapatra et al. [11], the inability of the macroscopic reaction–diffusion model (15.4)–(15.6) to predict the experimentally observed NBTI recovery is due to the one-dimensional description of the diffusive motion which makes it too easy for the hydrogen to find its dangling bond. (*Right*) A correct description of the three-dimensional atomic motion, so the argument, leads to much richer repassivation kinetics and thus to a distribution of repassivation times

$$\frac{\partial N_{\mathrm{it}}}{\partial t} = k_{\mathrm{f}}(N_0 - N_{\mathrm{it}}) - k_{\mathrm{r}} N_{\mathrm{it}} H_{\mathrm{it}}, \tag{15.4}$$

$$\frac{\partial H}{\partial t} = -D\frac{\partial^2 H}{\partial x^2} - k_H H^2 + k_{H_2} H_2, \tag{15.5}$$

$$\frac{\partial H_2}{\partial t} = -D_2\frac{\partial^2 H_2}{\partial x^2} + \frac{k_H}{2} H^2 - \frac{k_{H_2}}{2} H_2, \tag{15.6}$$

with the additional parameters $k_H$ and $k_{H_2}$ which are the reaction rates for dimerization and atomization, respectively. Again the motion of H and $H_2$ is described by a simple diffusion law with the corresponding diffusion coefficients $D$ and $D_2$ [17]. The combination of this modified reaction–diffusion model with empirical hole-trapping somewhat improves the match with experimental DC and AC stress data. The failure of the RD model to properly describe NBT recovery is shifted out of the time window of some experiments, but essentially remains.

Quite recently it was claimed that the misprediction of recovery is due to the one-dimensional description of the diffusing species in the macroscopic model (15.5) and (15.6) [11]. As illustrated in Fig. 15.4, it was suggested that this formulation makes it too easy for the hydrogen atom to find a dangling bond to passivate because the one-dimensional diffusion considers only two options of motion: forward and backward jumping. In a higher-dimensional description the diffusion and reaction kinetics are much richer:

1. The atoms can move in all three dimensions equally likely, leading to a distribution of arrival times at the interface during recovery.
2. $H_2$-molecules dissociate at a dangling bond, creating a passivated dangling bond and a free hydrogen atom that does not immediately find another dangling bond to passivate.
3. Hydrogen atoms arriving later have to hover along the interface to find an unoccupied dangling bond.

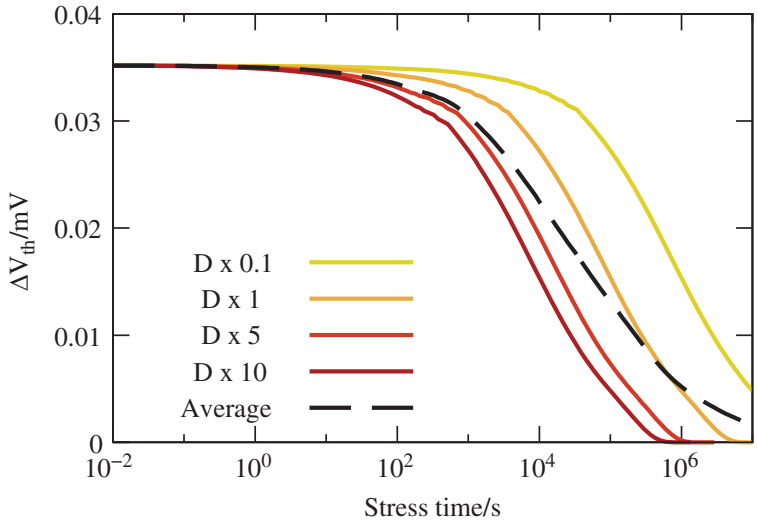


**Fig. 15.5** Numerically calculated recovery traces for different diffusion coefficients during recovery and the average of these traces. In accord with [11], this average trace shows a recovery that proceeds over more time-scales than the individual traces

A simple estimate of the recovery in this hypothetical three-dimensional model is given in [11]. This estimate tries to mimic the different repassivation kinetics arising in the atomic description within the framework of the usual macroscopic RD model. To account for the longer "effective" recovery paths, the diffusion coefficients in the macroscopic model are reduced by different factors during recovery and the resulting recovery traces are averaged. Although this approach gives a recovery that proceeds over more time scales, as shown in Fig. 15.5, no derivation for the quasi-three-dimensional description is given and its physical validity is at least questionable. One of our targets is to test the claims of [11] within a firm theoretical framework.

We have derived and implemented a microscopic formulation of the RD model [18, 19], in order to study the behavior of the RD mechanism on the atomic scale. This effort was made not only to test the claims of [11], but also to investigate general issues of the rate-equation-based description in the context of MOS reliability. As a literature study reveals, reaction–diffusion systems have been studied in numerous scientific communities from both the theoretical and the experimental side for more than a century [22–28]. Although the mathematical framework of the RD model (15.4)–(15.6) seems physically sound and the description using densities and rate equations is commonly considered adequate, it is a well-known and experimentally confirmed result of theoretical chemistry that the partial differential equation-based description of chemical kinetics breaks down for low concentrations [22]. Additionally, in reaction–diffusion systems bimolecular reactions, such as the passivation and the dimerization reaction, require a certain proximity of the reactant species, termed *reaction radius* [23, 28]. Usually the elementary bimolecular reactions happen almost instantaneously and it is the required collision, i.e., the reduction of the distance between two reactants below the reaction radius, which is the rate-limiting step [22]. In chemical kinetics, these reactions are called *diffusion-limited* or *diffusion-controlled* reactions [24].

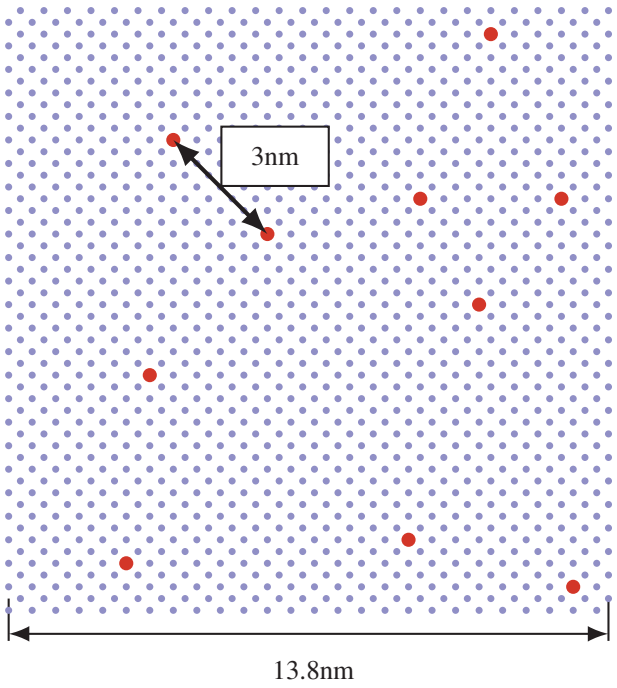


**Fig. 15.6** A random distribution of ten dangling bonds on a silicon (100) surface corresponding to a dangling bond density of $5 \times 10^{12}\,\mathrm{cm}^{-2}$, which is a common assumption for the number of bonding defects at the Si–$SiO_2$ interface [15, 20, 21]. The surface silicon atoms are shown in *blue*, the dangling bonds in *red*. At this density the average distance between two dangling bonds is ∼4.5 nm

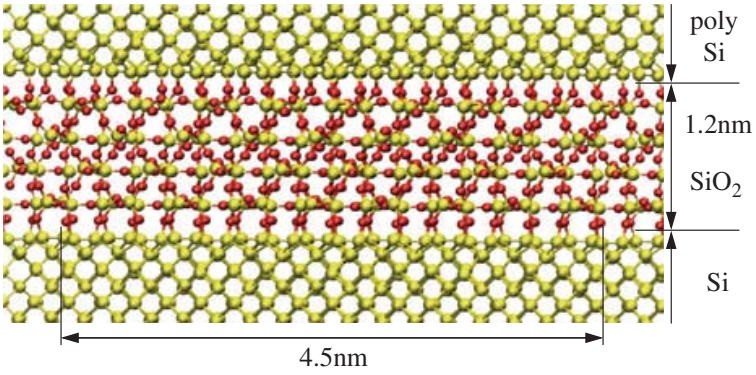


**Fig. 15.7** An idealized atomic model of an MOS structure and the average dangling bond distance of 4.5 nm, which spans several interstitial positions. It is intuitively clear that an elementary reaction between particles separated by this distance is strongly influenced by diffusion

It is easy to show that diffusion must play a dominant role in the bimolecular reactions in the RD model for NBTI. The density of bonding defects on oxidized silicon (100) surfaces is about $1 \times 10^{12}\,\mathrm{cm}^{-2}$ [20]. Figure 15.6 schematically shows a uniform random distribution of dangling bonds on a silicon (100) surface that corresponds to a density of $5 \times 10^{12}\,\mathrm{cm}^{-2}$, which is a usual assumption for $N_0$ [15, 21] in the RD model (15.1) or (15.4). The average distance between two nearest neighbors at this density is $d = {N_0}^{-1/2} \approx 4.5\,\mathrm{nm}$. An atomic model of the Si–$SiO_2$ interface as in Fig. 15.7 shows that two points separated by this distance have a large number of atoms in between. The assumption of an elementary reaction over this distance is clearly inappropriate, so any reaction between particles of this separation must involve a diffusive step.

Once established by the atomic viewpoint above, the diffusive influence on the bimolecular reactions leads to contradictions in the RD model and its physical interpretation. The predicted degradation of the RD model that is compatible with experimental data is only obtained if the hydrogen atoms that are liberated during stress compete for the available dangling bonds and dimerize at a certain rate. Both requirements involve diffusion over distances much larger than the nearest neighbor distance, which takes about 2 s at a commonly assumed diffusion coefficient of $D = 10^{-13}\,\mathrm{cm}^2/\mathrm{s}$ [15, 21]. The reaction radius $\rho_\mathrm{H}$ of the dimerization reaction can be estimated from the Smoluchowski theory for irreversible bimolecular reactions [23, 25, 26]

$$\rho_\mathrm{H} = \frac{k_H}{4\pi D}. \tag{15.7}$$

While a reasonable reaction radius is in the regime of the average radius of the oxide interstitials, which is about 4 Å [29], the application of (15.7) to published dimerization rates gives values ranging from 70 µm for the parametrization of [21] to thousands of kilometers for other parametrizations [15]. Although both values for $\rho_\mathrm{H}$ seem quite unreasonable, they only indicate a limited physical validity of the selected parametrization. An evaluation of the physical validity of the reaction–diffusion model itself requires a more detailed study using a computational model that properly treats the stochastic chemical kinetics involved. For the present study of the microscopic properties of the RD mechanism we have developed an atomistic reaction–diffusion simulator, which is described in the following.

## 15.2 Stochastic Description of Reaction–Diffusion Systems

Our microscopic RD model attempts to mimic the proposed mechanisms of the reaction–diffusion model at a microscopic level. The basic actors are H atoms, $H_2$ molecules, and the silicon dangling bonds at the interface. The investigations are carried out at the stochastic chemistry level. Several approaches have been used in the chemical literature for the stochastic simulation of reaction–diffusion systems [25, 26]. These approaches can roughly be categorized as grid-based methods or grid-less methods [26], owing to the description of the diffusion of the reactants, see Fig. 15.8. Grid-less methods propagate the coordinates of the diffusing species through Newton's equations of motion, quite similarly to molecular dynamics methods. Instead of explicitly treating all atoms of the solvent and their effect on the trajectory of the diffusors, the motion of the diffusing particles is perturbed by an empirical random force to generate a Brownian motion. Bimolecular reactions happen at a certain rate as soon as two reaction partners approach closer than a given radius. Although this technique suffers from its sensitivity to the time-step and the specific choice of the random force, it is a popular choice for the simulation of reaction–diffusion processes in liquid solutions where real molecular-dynamics

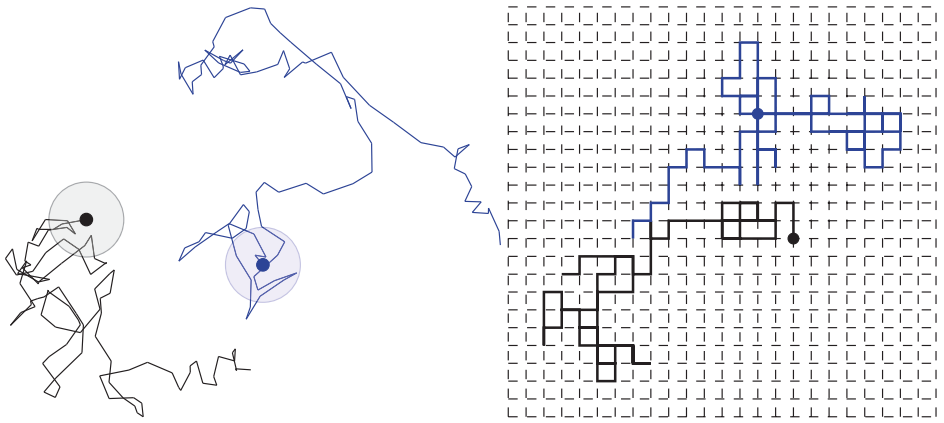

**Fig. 15.8** Schematic illustration of the stochastic modeling approaches to reaction–diffusion systems. *(Left)* In grid-less methods, a molecular trajectory is generated from the transient solution of Newton's equations of motion as in molecular dynamics simulation. The interaction with the solvent is modeled by a random force that acts on the diffusors. Bimolecular reactions occur when two particles approach closer than the reaction radius (indicated as *circles* around the particles) *(Right)* In grid-based methods diffusion proceeds as jumps between the sub-domains defined by the grid. Bimolecular reactions occur when two particles occupy the same grid-point

simulations are not feasible [25, 26]. In grid-based methods the simulated volume is divided into small domains and each diffusing particle is assigned to a specific domain. The motion of the diffusors proceeds as hopping between the grid-points. In these models the bimolecular reactions happen at a certain rate as soon as two reactants occupy the same sub-volume. The advantage of this approach is that it can be formulated on top of the chemical master equation. This equation can be solved without artificial time-stepping, as explained in the following section. A problem of the grid-based method that is repeatedly discussed in chemical literature is the choice of the spatial grid as it induces a more or less unphysical motion in liquid solutions. Additionally, the probability to find two particles on the same grid point and in consequence the rate of bimolecular reactions are quite sensitive to the volume of the sub-domains [26].

In the reaction–diffusion model for NBTI, the diffusion of the particles proceeds inside a solid-state solvent. Contrary to diffusion in gases or liquids, the motion of an impurity in a solid-state host material proceeds via jumps between metastable states as illustrated in Fig. 15.9. This hopping diffusion is understood as a hopping process over energetic barriers. In the case of H or $H_2$, which do not react with the host atoms, these barriers arise from the repelling Coulomb interaction between the electron clouds of the host lattice and the diffusor. The minima of the potential energy surface are thus the interstitial positions of the host lattice [30, 31]. In between the jumps, the motion of the atom is randomly vibrational rather than diffusive. This discreteness of motion not only strongly suggests the use of a grid-based method, where the grid points are interstitial positions of the host lattice, but also induces a natural discretization into the reaction–diffusion equations. As a consequence, the description based on macroscopic diffusion equations in the RD

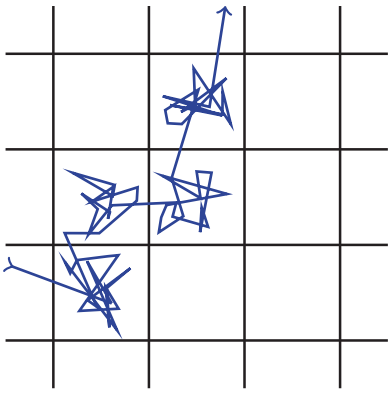

**Fig. 15.9** Schematic trajectory of an inert interstitial atom diffusing in a solid-state host material. The potential energy barriers of the host material are indicated as black grid. The diffusion itself proceeds via jumps between the interstitial positions. In between the jumps, the atom vibrates randomly around an energetic minimum

model (15.5) and (15.6) are only valid at distances that are much larger than the interstitial radius and it has to be assumed that at very short distances a description using hopping diffusion is more accurate.

## 15.3 The Chemical Master Equation

From the considerations of the previous section we conclude that the most appropriate description of the physics considered in the present work is obtained from the reaction–diffusion master equation approach [25–28]. Within the natural lattice of interstitial positions the actors of our RD system exist in well-defined and discrete states. Once the chemical states and reactions that comprise the chemical system under consideration are defined, their dynamics can be described as a random process that switches between the states [32, 33]. Mathematically, the state of the chemical system is described as a vector $\vec{x}$. In addition, a set of reaction channels is established, which cause the transitions between the discrete states of this vector. Due to the unpredictable nature of the dynamics of the microstates, the time at which a reaction takes place is not a deterministic quantity. Instead, if the chemical system is in a given state $\vec{x}_\alpha$ at time $t$, for every reaction channel $\gamma$ a reaction rate $c_\gamma$ can be defined, so that $c_\gamma \mathrm{d}t$ is the probability of the reaction taking place between $t$ and $t + \mathrm{d}t$ [33]. Different chemical states have different reaction rate constants for their reaction channels. These reaction rate constants depend only on the current state of the chemical system irrespective of the previous states of the system. In this case a function can be defined for every reaction channel that assigns a specific rate to a specific state $c_\gamma = a_\gamma(\vec{x}_\alpha)$. These functions are called the *propensity functions* [33]. The change induced by the reaction channel $\gamma$ is described using the state change vector $\vec{v}_\gamma$. The thus formulated model describes a memory-less random process with discrete states, which is usually called a Markov process [34]. The removal

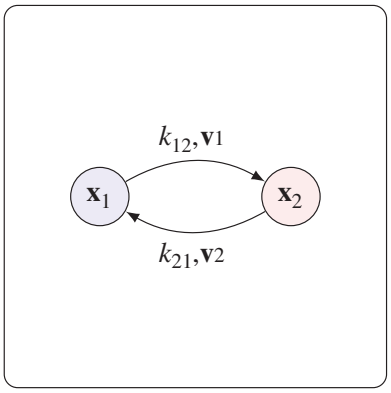


**Fig. 15.10** Illustration of the example system discussed in the text. The system exists in one of the two states $\vec{x}_1$ and $\vec{x}_2$. Transitions between these states are caused by two reaction channels with the state-change vectors $\vec{\nu}_1$ and $\vec{\nu}_2$, and the associated rates $k_{12}$ and $k_{21}$

of memory from the system occurs through the thermal equilibration, which is assumed to happen much faster than the chemical reactions. According to the theory of stochastic chemical kinetics [32, 33], the evolution of this system over time can then be described by a chemical master equation

$$\frac{\partial P(\vec{x},t)}{\partial t} = \sum_{\gamma=1}^{\Gamma} [a_\gamma(\vec{x}-\vec{\nu}_\gamma)P(\vec{x}-\vec{\nu}_\gamma,t) - a_\gamma(\vec{x})P(\vec{x},t)], \tag{15.8}$$

where $P(\vec{x},t) = P(\vec{X}=\vec{x},t|\vec{x}_0,t_0)$ is the probability that the stochastic process $\vec{X}(t)$ equals $\vec{x}$ at time $t$, given that $\vec{X}(t_0)=\vec{x}_0$.

The master equation approach can be illustrated using the simple example of a system with two states $\vec{x}_1$ and $\vec{x}_2$ [34], see Fig. 15.10. The system has two reaction channels 1 and 2, which connect the two states through the state change vectors $\vec{\nu}_1$ and $\vec{\nu}_2$ as

$$\vec{x}_1 + \vec{\nu}_1 = \vec{x}_2 \quad \text{and} \quad \vec{x}_2 + \vec{\nu}_2 = \vec{x}_1 \tag{15.9}$$

The propensity functions $a_1$ and $a_2$ assume the form

$$a_1(\vec{x}_1) = k_{12}, \qquad a_1(\vec{x}_2) = 0, \tag{15.10}$$

$$a_2(\vec{x}_1) = 0, \text{and} \quad a_2(\vec{x}_2) = k_{21}. \tag{15.11}$$

The master equation for this system consequently reads

$$\frac{\partial P(\vec{x}_1,t)}{\partial t} = k_{21}P(\vec{x}_2,t) - k_{12}P(\vec{x}_1,t) \tag{15.12}$$

$$\frac{\partial P(\vec{x}_2,t)}{\partial t} = k_{12}P(\vec{x}_1,t) - k_{21}P(\vec{x}_2,t) \tag{15.13}$$

As the system can only exist in one of the two states at a time, it follows that

$$P(\vec{x}_1,t) = 1 - P(\vec{x}_2,t) = p(t), \tag{15.14}$$

which reduces the master equation of the two-state system to

$$\frac{\partial p(t)}{\partial t} = k_{21}(1 - p(t)) - k_{12}p(t). \tag{15.15}$$

This is the rate-equation of the two-state system, which is equivalent to the master equation for this simple example.

Within the theoretical framework of the chemical master equation, all the microphysical details elaborated in the previous section are now contained in the propensity functions $a_\gamma$ and the state-change vectors $\vec{\nu}_\gamma$ for the $\Gamma$ reaction channels.

## 15.4 States and Reactions in the Microscopic RD Model

The main actors of the microscopic RD model are the H atoms and the $H_2$ molecules. The state vector $\vec{x}$ of the system consequently contains the interstitial positions and bonding states of all actors. The reactions employed in our simulations are the hopping transport between interstitial sites, the passivation/depassivation reaction, and the dimerization/atomization reaction. These reactions are treated as elementary reactions and are formalized in the reaction channels given in Fig. 15.11. The stochastic chemical model is solved using the stochastic simulation algorithm (SSA) explained in Sect. 15.5.

In the microscopic RD model employed in this work the interstitial sites form a regular and orthogonal three-dimensional grid and the hopping rates for the diffusors are assumed to be constant in accord with the isotropic and non-dispersive diffusion underlying the conventional macroscopic RD model [4]. In a real $SiO_2$ of a MOS transistor the amorphous structure will of course lead to a random network of interstitial sites [29] with a variety of hopping rates and a more complex topology. However, as the power-law degradation predicted by the macroscopic RD model requires a constant diffusion coefficient, these variations must be assumed unimportant [17] in order to obtain agreement with the established model. As illustrated in Fig. 15.12, the simulation region in our calculations is a rectangular box which extends to infinity normal to the Si–$SiO_2$ interface and has closed lateral boundaries. The Si–$SiO_2$ interface itself is represented by a special region at the bottom of the simulation box where selected interface sites have the ability to bond or release a diffusing hydrogen atom, see Figs. 15.11 and 15.12. The positions of the dangling bond sites in the interface region are picked randomly, similar to Fig. 15.6.

As mentioned above, the choice of the grid size requires special attention as it determines the probability of the bimolecular reactions. The interstitial size of amorphous silica has been calculated for molecular dynamics generated atomic structures and is about 4 Å [29]. We take this value as the physically most reasonable grid size.

Once the microscopic model is defined, the relation to the macroscopic RD model (15.4)–(15.6) has to be established. Using the number of dangling bonds in the simulation box $n_{\mathrm{DB}}$, the number of hydrogen atoms passivating a dangling bond

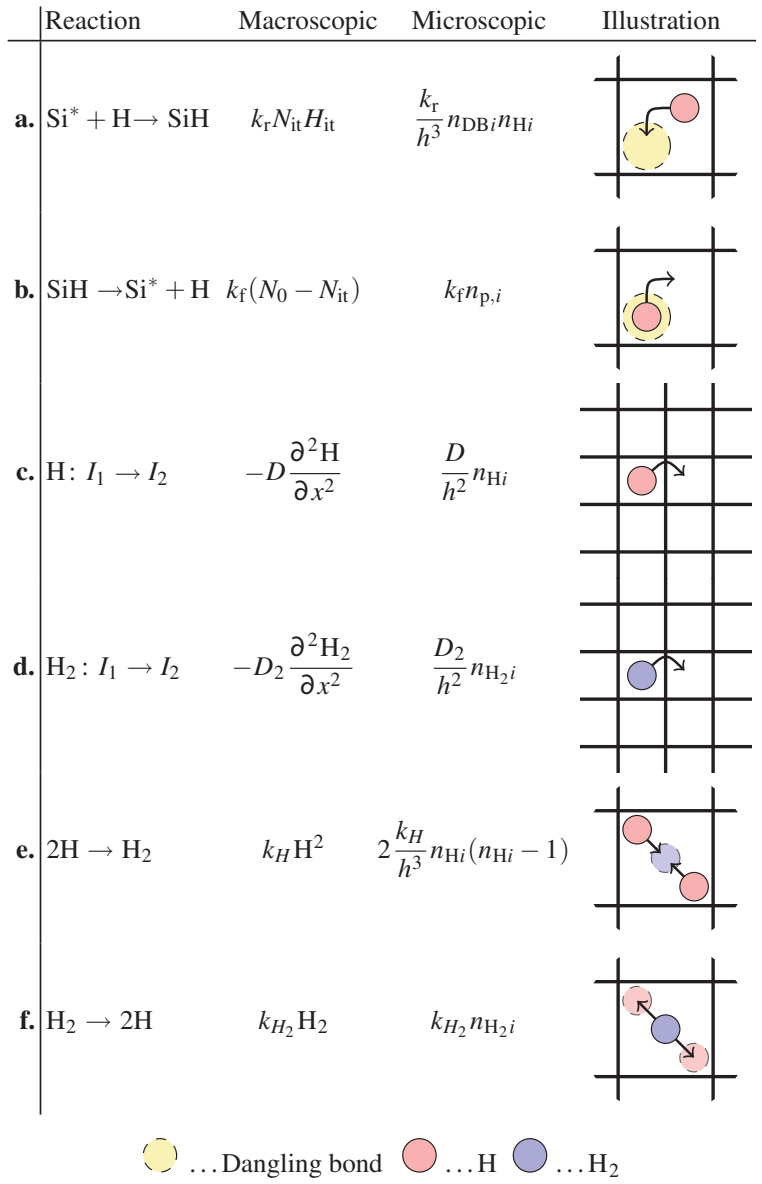


**Fig. 15.11** Reaction channels and propensities in the microscopic RD model along with their macroscopic counterpart. (**a**) The dangling bonds are represented by special sites at the bottom of the simulation box. Empty dangling bond sites can be passivated by a free hydrogen atom. (**b**) Occupied dangling bond sites do not offer a bonding reaction channel, they can only emit their hydrogen atom. (**c**, **d**) Within the bulk $SiO_2$, the atoms or molecules are allowed to jump from an interstitial $I_1$ to any neighboring site $I_2$. (**e**) When two hydrogen atoms occupy the same interstitial position, they can undergo a dimerization at rate $k_H$ and form $H_2$. (**f**) Each hydrogen molecule dissociates at a rate $k_{H_2}$ back into two hydrogen atoms. For interstitial site $i$, $n_{DBi}$ is the number of (depassivated) dangling bonds, $n_{p,i}$ is the number of passivating hydrogen atoms, $n_{Hi}$ is the number of free hydrogen atoms, and $n_{H_2 i}$ is the number of hydrogen molecules. $h$ denotes the step size of the spatial grid

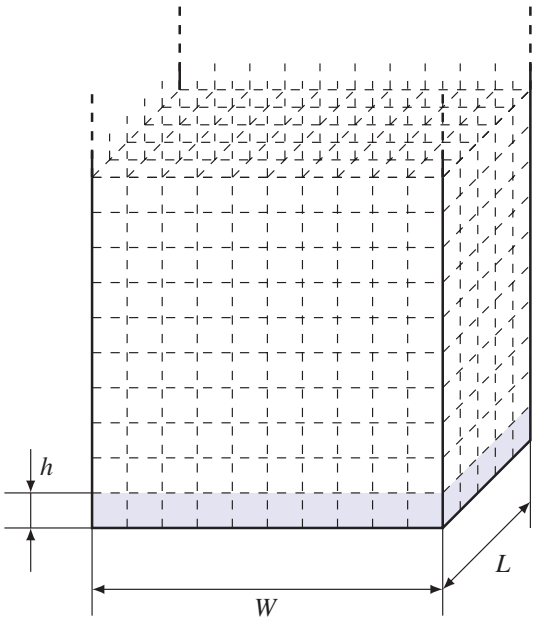


**Fig. 15.12** The simulation structure for the microscopic RD model employed in this work is a bounded region of width $W$ and length $L$ and infinite extension in $z$-direction, i.e., normal to the interface. The silicon dangling bonds are connected to special interstitial positions in the Si–$SiO_2$ interface region at the bottom of the simulation box (*blue*/*gray*). The interstitial positions are assumed to form an orthogonal lattice with constant jump-width and constant diffusion coefficients

$n_\mathrm{p}$ and the numbers $n_{\mathrm{H}i}$ of H and $n_{\mathrm{H}_2 i}$ of $H_2$ at interstitial $i$, this relation is obtained from the discretization induced by the grid [21] as

$$N_0 = \frac{n_\mathrm{DB}}{WL}, \tag{15.16}$$

$$N_\mathrm{it} = \frac{n_\mathrm{DB} - n_\mathrm{p}}{WL}, \tag{15.17}$$

$$\mathrm{H}(x_i) = \frac{n_{\mathrm{H}i}}{V_i}, \tag{15.18}$$

$$\mathrm{H}_2(x_i) = \frac{n_{\mathrm{H}_2 i}}{V_i}, \tag{15.19}$$

where $W$, $L$ and $h$ are illustrated in Fig. 15.12 and $V_i$ is the volume of interstitial $i$ which is $V_i = h^3$ in this work. The relation between the rates of the macroscopic model and the microscopic propensity functions are given in Fig. 15.11. Initially, all hydrogen atoms are passivating silicon dangling bonds

$$n_\mathrm{p}(t = 0) = n_\mathrm{DB}, \tag{15.20}$$

in accordance with the assumptions of the macroscopic RD model.

## 15.5 Solution of the Master Equation

Now that the chemical states and reactions are defined we can calculate the time evolution of the chemical system from the chemical master equation (15.8). As explained above, this equation is a stochastic differential equation which assigns a

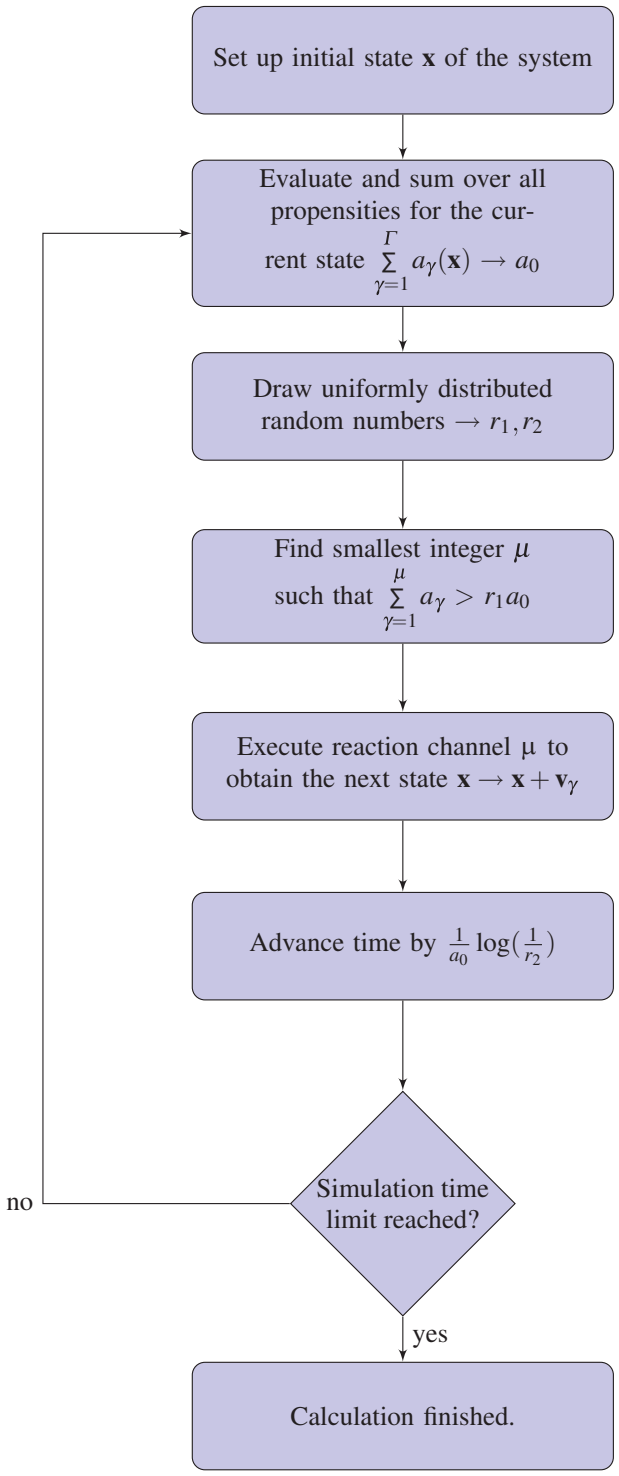


**Fig. 15.13** Sketch of the stochastic simulation algorithm (SSA) [32]. The algorithm generates a realization of the stochastic process described by the chemical master equation (15.8)

probability at time $t$ to any state vector $\vec{x}$, given that $\vec{X} = \vec{x}_0$ at $t = t_0$. As in the simple example above, for a system with a small set of states $\{\vec{x}_1, \ldots, \vec{x}_\Omega\}$, a direct solution can be attempted, which results in a coupled system of $\Omega$ differential equations [34]. However, in many situations the number of states will be large or even infinite, rendering a direct solution of the master equation unfeasible or even impossible. A feasible alternative is the SSA [32] explained in Fig. 15.13, which is also known as

the kinetic Monte Carlo method. Instead of solving the differential equation (15.8), a realization of the stochastic process $\vec{X}$ is generated using pseudo-random numbers. The SSA does not have any algorithmic parameters and is a mathematically exact description of the system defined by the states and reaction channels [32]. Averages of the probability distribution $P(\vec{x},t)$ are trivially calculated over several simulation runs, although care has to be taken to ensure the randomness of the pseudo-random numbers between two runs to avoid correlation effects.

## 15.6 Results and Discussion

Two different systems have been studied in detail: a model system and a "real-world-example." The model system is used to study the general features of the microscopic reaction–diffusion process. It is parametrized in order to clearly show all relevant features at a moderate computational effort. The parametrization of the real-world system is based on a published parametrization of the modified reaction–diffusion model. This system is used to relate our microscopic model to published data.

### *15.6.1 General Behavior of the Microscopic RD Model*

The parametrization that is used to study the general behavior is given in Table 15.1. As the time evolution in the SSA proceeds reaction by reaction, the channels with the fastest rates determine the execution time. The computational effort scales linearly with the number of particles in the system, which is determined by the lateral extent of the simulation box. As the dangling bonds and in consequence also the diffusing particles are uniformly distributed along the interface at any time in our calculations, the reflecting boundary conditions in our calculations are equivalent to periodic boundary conditions. Thus, our calculations correspond to an infinitely extended Si–$SiO_2$ interface and the lateral box size only determines the

**Table 15.1** Parameters of the model system

| Reaction | Propensity ($s^{-1}$) |
|---|---|
| Depassivation | 0.5 |
| Passivation | $4 \times 10^4$ |
| Dimerization | $2 \times 10^5$ |
| Atomization | 5 |
| H-hopping | 100 |
| $H_2$-hopping | 100 |

The parameters have been selected to enable a study of the different regimes of the microscopic RD model at moderate computational expense. The rates are given in terms of the microscopic model as in Fig. 15.11

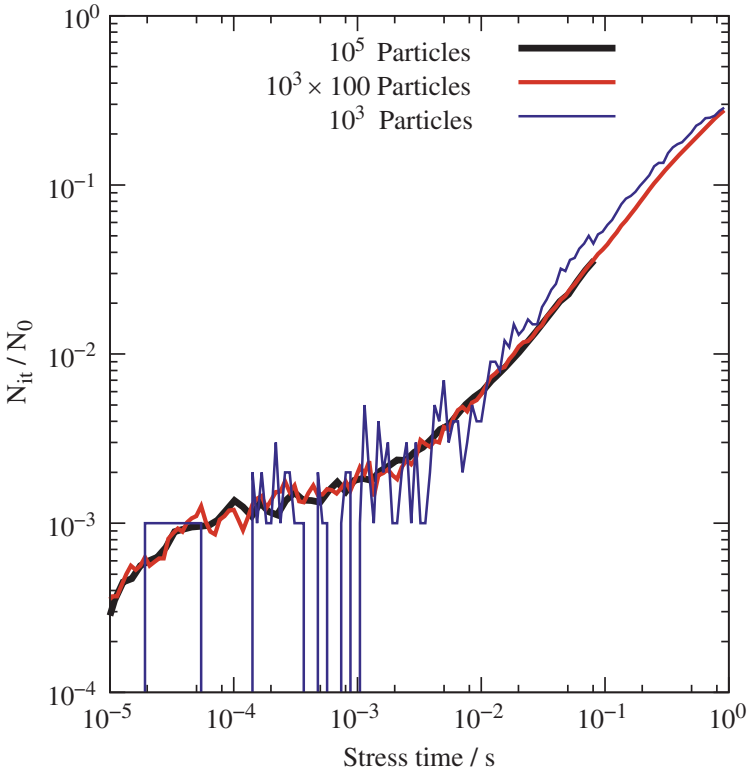


**Fig. 15.14** As the interaction between the diffusing particles is small at early degradation times, the computation can be parallelized by averaging over several simulation runs. The figure shows that a calculation with $10^5$ particles is equivalent to the average of 100 calculation runs with $10^3$ particles. The result of a single $10^3$ particle run is shown for comparison. $k_f$ was increased by a factor of 100 for this calculation, in order to obtain smooth curves from the $10^5$ particle run

resolution, i.e., the noise level, of the degradation curves. The computational effort scales roughly linearly with the simulated time, which means exponential scaling for the logarithmic abscissa that is used for BTI degradation curves. The choice of the lateral extent of the simulation box is thus based on a trade-off between accuracy and computational speed and has to be adapted for the study of the different degradation regimes.

At early degradation times the low degradation level requires a high resolution, i.e., a large number of particles is required to obtain smooth results. Fortunately, as reactions between the hydrogen atoms or between hydrogen atoms and neighboring dangling bonds do not happen in this regime, a good parallelization can be obtained by averaging over separate simulation runs, see Fig. 15.14.

The earliest degradation times are dominated by the depassivation of the silicon dangling bonds leading to a linear increase of the degradation, which is equivalent to the initial "reaction limited" degradation of the macroscopic RD model [3]. However, the degradation predicted by the microscopic RD model quickly saturates as an equilibrium forms between depassivation and repassivation *for each dangling bond separately*. In the absence of any diffusion the time evolution of the number of hydrogen atoms passivating a silicon dangling bond is given by

$$\frac{\partial n_p}{\partial t}(t) = -k_f n_p(t) + \frac{k_r}{h^3}(n_{DB} - n_p(t)) \tag{15.21}$$

$$n_p(t=0) = n_{DB} \tag{15.22}$$

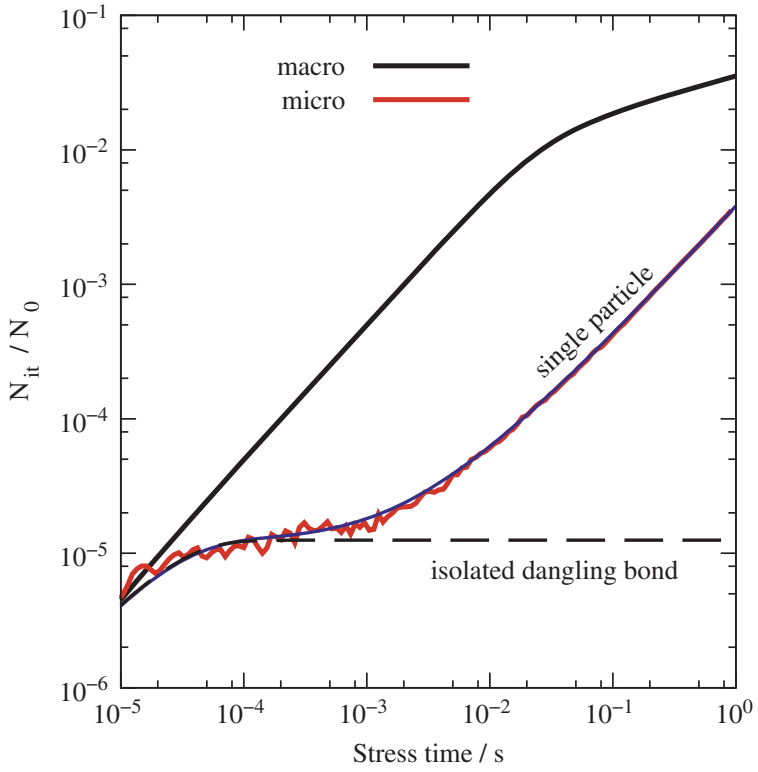


**Fig. 15.15** Comparison of the microscopic RD model with 25,000 particles averaged over 50,000 runs to its macroscopic counterpart, the single-particle expressions (15.24) and (15.25) and the isolated dangling-bond equilibration (15.23). The earliest degradation times are dominated by the equilibration between the depassivation and passivation reaction at every dangling bond. Around 1 ms, the departure of the hydrogen atoms from the dangling bond site begins but the interaction between the diffusors is still negligible

with the solution

$$n_{\mathrm{p}}(t) = n_{\mathrm{DB}} - \frac{n_{\mathrm{DB}} k_{\mathrm{f}}}{k_{\mathrm{f}} + \frac{k_{\mathrm{r}}}{h^3}} \left(1 - e^{-\left(k_{\mathrm{f}} + \frac{k_{\mathrm{r}}}{h^3}\right) t}\right). \tag{15.23}$$

A comparison of the microscopic RD model and (15.23) is shown in Fig. 15.15. The initial behavior of the microscopic RD model stands in stark contrast to the degradation in the macroscopic model where the linear regime continues until a global equilibrium has formed at the interface.

As this initial behavior takes a central position in our further discussion, it requires a deeper analysis. The microscopic single-particle regime can be accurately described using rate equations as it does not contain any second-order reactions. The required equations are basically those of the RD model, but as every hydrogen atom can be assumed to act independently, the expressions for the hydrogen bonding as well as the competition for dangling bonds are neglected. As the kinetic behavior in this regime is strongly determined by the first diffusive steps of the hydrogen atoms, the diffusion part of this approximation must have the same interstitial topology as the microscopic model. As all hydrogen atoms act independently, only one atom and

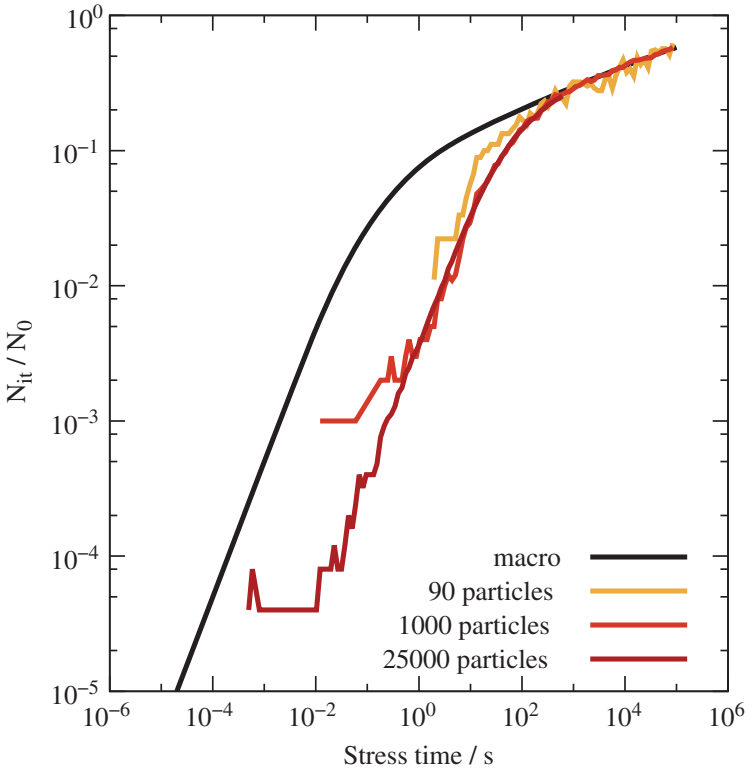


**Fig. 15.16** Due to the larger fraction of depassivated dangling bonds, the number of diffusing particles can be reduced for long-term simulations. Three microscopic calculations are compared to the macroscopic result. The 25,000 particle simulation clearly shows the transition between the single-particle and the macroscopic diffusion-limited regime. The 1,000 particle calculation captures the transition region but is too noisy for $t < 100\,\text{ms}$. The 90 particle simulation captures the macroscopic regime with reasonable accuracy

one dangling bond need to be considered. The interface reaction and the diffusion of the hydrogen atom is thus described as

$$\frac{\partial n_\text{p}}{\partial t} = -k_\text{f} n_\text{p} + \frac{k_\text{r}}{h^3} n_\text{DB} n_\text{H0} \text{ and} \tag{15.24}$$

$$\frac{\partial n_{\text{H}i}}{\partial t} = \sum_{j \in \mathcal{N}(i)} \frac{D}{h^2} (n_{\text{H}j} - n_{\text{H}i}), \tag{15.25}$$

respectively, where $\mathcal{N}$ denotes the set of neighboring interstitials to $i$. Figure 15.15 compares the microscopic RD calculation with the approximations for the different regimes at early degradation times, which shows that the single-particle approximation perfectly matches the behavior of the full model in the initial phase.

After the atoms have traveled sufficiently long distances, the interaction between the particles becomes relevant and the single-particle approximation becomes invalid. In Fig. 15.16 this is visible as a transition away from the single-particle behavior toward the macroscopic solution between 1 s and 1 ks. As the fraction of depassivated dangling bonds in this regime is much higher than during early degradation times, the results are not as strongly influenced by the noise of the SSA calculation. Consequently the number of particles can be reduced for longer simulation times, which makes the prediction of long-term degradation possible.

Finally, Fig. 15.17 compares the microscopic RD model to the macroscopic version over the course of one complete stress cycle, where the microscopic curve was obtained by combining calculations of different accuracy, as explained above.

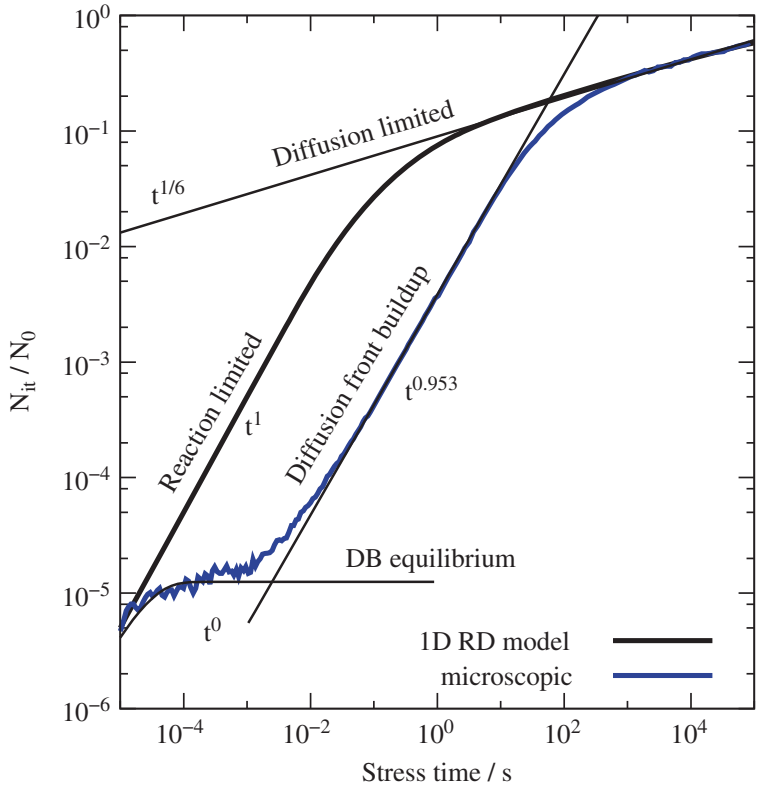


**Fig. 15.17** Comparison of all regimes of the microscopic RD model to the degradation predicted by the macroscopic RD model. Obviously there is a large discrepancy between the two descriptions and the behavior of the physically more reasonable microscopic model is not experimentally observed

Instead of the three regions which arise from the macroscopic RD model—reaction-limited, equilibration, and diffusion-limited—the H–$H_2$ microscopic description has four to five regimes depending on the particular parametrization:

- The earliest degradation times ($t < 20\,\mu s$ in this case) are dominated by the depassivation of dangling bonds. In this regime, the microscopic and the macroscopic model give identical degradation behavior.
- After the passivation and depassivation have reached an equilibrium between $k_f$ and $k_r$ separately for each Si–H bond, the fraction of depassivated dangling bonds remains constant until the diffusion of the hydrogen atoms becomes dominant. This regime only shows when the individual hydrogen atoms are considered and consequently is not obtained from any model based on rate equations.
- As more and more hydrogen atoms leave their initial position, the degradation is determined by the buildup of a diffusion front along the Si–$SiO_2$ interface and the equilibration between the dangling bonds. This regime has a very large power-law exponent of almost one[1] that is not experimentally observed. The stress time

[1]In our earlier studies on two-dimensional systems this exponent was around 0.8 [19], owing to the topology dependence of this regime.

range in which this regime is observed depends on the average distance between two dangling bonds, the diffusion coefficient, and the interstitial size.
- As the bimolecular reactions become relevant, the macroscopic diffusion-limited regime begins to emerge. For some parametrizations we have observed a time window in which the initial diffusion-limited regime has the typical $t^{1/4}$-form that arises from the classical RD model without $H_2$ [19]. In this case the dimerization rate is reduced by the diffusive step and the H diffusion dominates the degradation until a sufficient amount of $H_2$ has formed.

The initial single-particle phase of the degradation is a remarkable feature of the microscopic model. As it is incompatible with experimental data and very sensitive to the parametrization, its relevance for real-world reliability projections has to be investigated. For this purpose we have run calculations based on a published parametrization of the reaction–diffusion model for NBTI, see Sect. 15.6.4.

### *15.6.2 Recovery*

In agreement with our investigations on two-dimensional systems [18, 19], the three-dimensional stochastic motion of the hydrogen atoms *does not influence* the recovery behavior of the system after long-term stress, which contradicts the suggestions of [11]. As shown in Fig. 15.18, a longer relaxation transient is only obtained if the preceding stress phase does not show a power-law regime. As the system comes closer to the macroscopic degradation behavior, the recovery in the microscopic model also approaches the macroscopic version, which is incompatible with experimental data [7, 8, 35]. This behavior is to be expected as the $t^{1/6}$ degradation regime requires an equilibration and thus a quasi-one-dimensional behavior. The recovery proceeds on a timescale that is at least two orders of magnitude longer than the stress time. The lateral search of hydrogen atoms for unoccupied dangling bonds was suggested to dominate at the end of the recovery. However, due to the logarithmic time scale on which recovery is monitored, the equilibration along the interface has negligible impact at the end of the recovery trace if this equilibration proceeds about two orders of magnitude faster. Thus, the hovering of hydrogen atoms along the interface does not influence the shape of the recovery transient.

### *15.6.3 Approximations in the Macroscopic Model*

After the microscopic RD theory Fig. 15.11 has been established and its general behavior has been investigated, one can use this framework to analyze the assumptions that are implicit to the macroscopic RD model (15.4)–(15.6), which is still widely considered to be an adequate approximation.

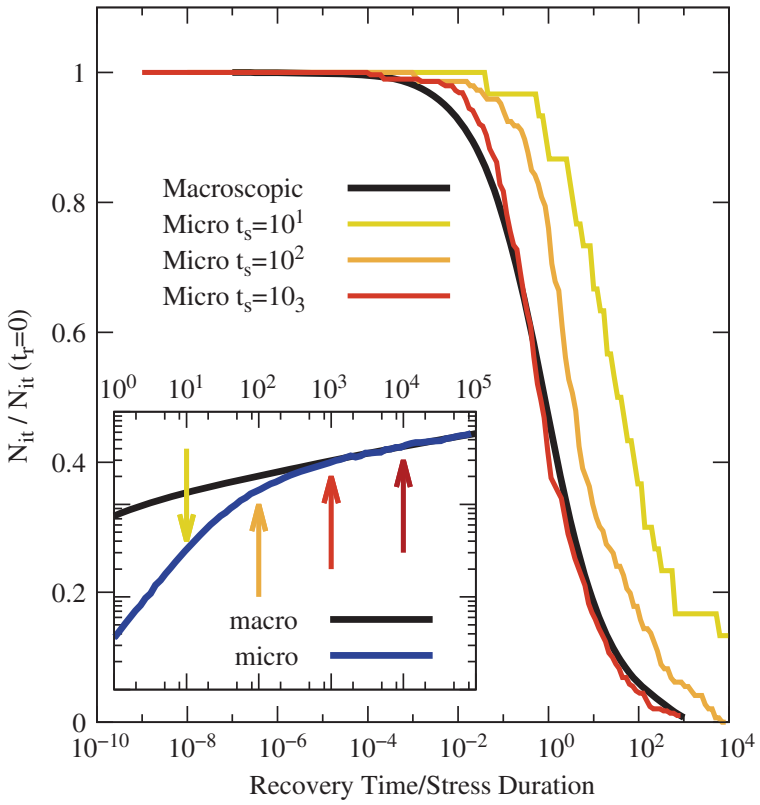


**Fig. 15.18** Recovery transients for different stress times. As the degradation transient approaches the macroscopic diffusion limited regime (see inset), the recovery comes closer to the macroscopic recovery, leading to a perfect match as soon as the degradation assumes the experimentally relevant $t^{1/6}$ form

The most obvious approximation in the macroscopic RD model is the one-dimensional description of diffusion. While this may seem to be appropriate as boundary effects in the diffusion of both H and $H_2$ are negligible, it tacitly introduces the assumption of lateral homogeneity along the interface. This homogeneity includes the following assumptions:

- *All* the liberated hydrogen atoms at the interface ($H_{it}$ in (15.1) and (15.4)) compete instantaneously with *all* the other free interfacial hydrogen atoms for *all* the available dangling bonds.
- *All* the pairs of hydrogen atoms at a certain distance from the Si–$SiO_2$ interface are equally likely to undergo dimerization and form $H_2$, independently of their spatial separation.

As was shown above, a hydrogen atom liberated during stress initially stays in the vicinity of its original dangling bond and thus the lateral homogeneity has to be considered a long-term approximation. It is accurate when the diffusion of hydrogen has led to enough intermixing so that there is no significant variability in the concentration of free hydrogen along the interface. Following [33], this condition can be called "lateral well-stirredness" of the system.

The second and more delicate approximation in the macroscopic RD model is the mathematical description using rate- and diffusion-equations. In the microscopic

**Table 15.2** The parameters employed in the real-world simulations

| | |
|---|---|
| $k_\mathrm{f}$ | $3\,\mathrm{s}^{-1}$ |
| $k_\mathrm{r}$ | $6 \times 10^{-13}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$ |
| $k_H$ | $5.6 \times 10^{-11}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$ |
| $k_{H_2}$ | $95.4\,\mathrm{s}^{-1}$ |
| $D$ | $10^{-13}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ |
| $D_2$ | $1.8126 \times 10^{-14}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ |
| $N_0$ | $5 \times 10^{12}\,\mathrm{cm}^{-2}$ |

The parameter set is based on the values published in [15] but was slightly modified to give the same degradation behavior with physically more reasonable $k_\mathrm{r}$ and $k_H$

RD model, the rate at which an atom at the interface passivates a dangling bond depends not only on the rate $k_\mathrm{r}$ but also on the probability of finding this atom at the position of the dangling bond. In the macroscopic model the precondition of having an unoccupied dangling bond at the interface is described multiplicatively as $k_\mathrm{r} N_\mathrm{it} H_\mathrm{it}$. At early times during degradation, when each hydrogen atom still resides near its dangling bond, this term introduces an unphysical self-interaction where each hydrogen atom competes with *itself* for its dangling bond. As the root of this problem lies in the assumptions implicit to a formulation based on rate-equations, the error is also present in a macroscopic model with three-dimensional diffusion. As explained in [19], this means that a rate-equation-based RD model will not accurately describe the degradation at early times even if higher-dimensional diffusion and discrete dangling bonds are considered.

Similar to the passivation rate, the rate at which $\mathrm{H}_2$ is formed in the microscopic RD model depends on both the dimerization rate $k_H$ and the probability of finding two hydrogen atoms which occupy the same interstitial position. In the macroscopic RD model, this dimerization reaction is modeled as $k_H \mathrm{H}^2$. As thoroughly explained in [22], this approximation is only valid for large numbers of particles, as the number of pairs of hydrogen atoms in an interstitial goes as $N(N-1)$ which can only be approximated as $N^2$ if $N$ is sufficiently large.

All in all, the macroscopic RD model can only be considered a valid approximation of the microscopic RD model for very long stress times and a sufficient amount of liberated hydrogen atoms. The time it takes for the macroscopic approximation to become valid, however, may exceed the time range in which it is usually applied, depending on the parametrization.

### *15.6.4 A Real-World Example*

To study the behavior of the atomistic model for a real-world example, we compare to the measurements of Reisinger et al. [7] using the parametrization of Islam et al. [15] in a modified form, see Table 15.2. Figure 15.19 shows the results of our

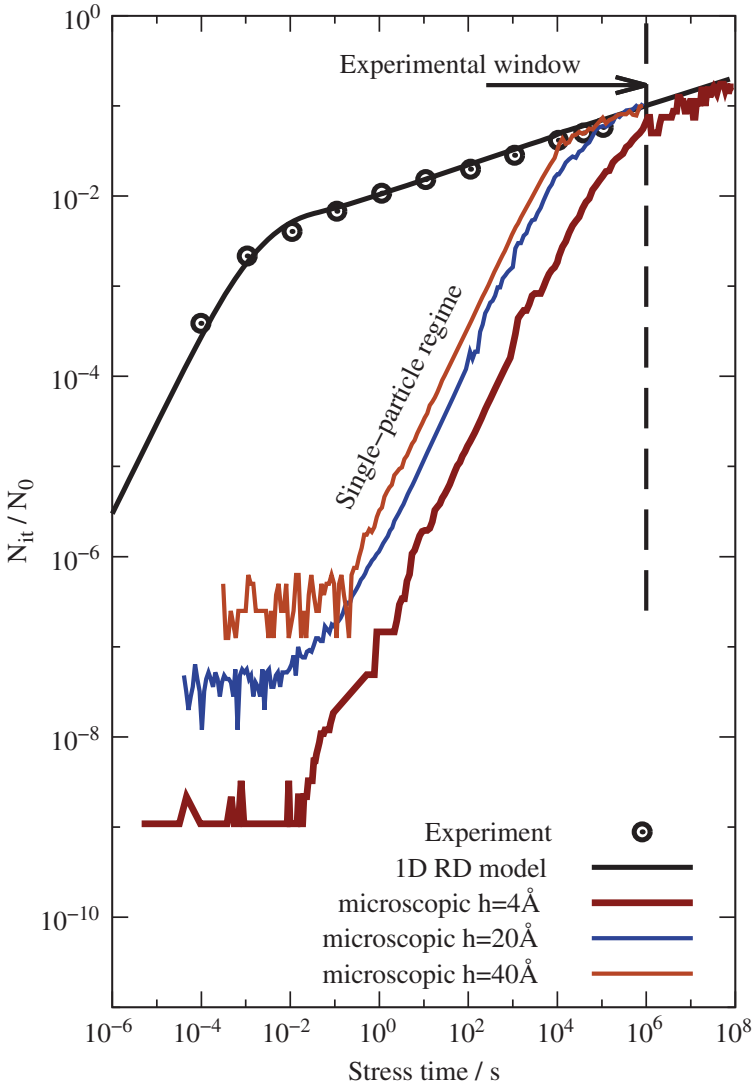


**Fig. 15.19** The degradation transient predicted by the microscopic RD model for four interstitial sizes compared to the macroscopic one-dimensional model and experimental data. Using the parameters in Table 15.2, the prediction of the microscopic RD model is completely incompatible with the experimental data as the onset of the $t^{1/6}$ regime is delayed beyond $10^8$ s (about 3 years) for a reasonable interstitial size of 4 Å. Increasing the interstitial size reduces the effect as it increases the effective reaction radius for the bimolecular reactions. However, even for unphysically large interstitial sizes, the onset of the $t^{1/6}$ regime is delayed to $10^4$ s ($h = 40$ Å) or $10^5$ s ($h = 20$ Å)

calculations for several interstitial sizes. While the macroscopic one-dimensional RD model fits the data very well, the kinetic Monte Carlo data shows a completely different behavior. Again, the single-particle regime is clearly present. However, due to the low density of dangling bonds at the interface, the single-particle regime dominates the degradation for a large part of the stress time. For a realistic interstitial size of 4 Å [29, 36], the onset of the $t^{1/6}$ regime lies far beyond the experimental window of $10^5$ s. When the interstitial size is increased, the onset of the $t^{1/6}$ regime moves to earlier times, which is due to the increase of the reaction radius for the bimolecular reactions as explained above. For the given parameter set, an interstitial

size of $h = 2\,\mathrm{nm}$, which is the total thickness of the oxide of the device under consideration [7], is required to at least have the $t^{1/6}$ regime touch the experimental window.

A shift of the onset of the experimentally observed regime to earlier times at a realistic interstitial size requires a dramatic increase of either the hydrogen diffusion coefficient or the availability of free hydrogen near the interface. An increase of the hydrogen diffusion coefficient, however, breaks the dominance of $H_2$ flux over the flux of atomic hydrogen and changes the predicted degradation away from the experimentally observed $t^{1/6}$ towards $t^{1/4}$. Increasing the availability of hydrogen at the interface by adjusting the ratio $k_\mathrm{f}/k_\mathrm{r}$ causes similar problems, as the $H_2$ diffusion coefficient has to be lowered in order to give the same overall degradation.

This indicates that in the given microscopic model it is impossible to obtain the experimentally observed $t^{1/6}$ degradation within the experimental window at a reasonable interstitial size.

### *15.6.5 Increased Interface Diffusion*

The behavior predicted by the microscopic model is completely incompatible with any experimental data, while the description is much closer to the physical reality than the macroscopic RD model. Only two interpretations are possible to resolve this dilemma. Either the ability of the macroscopic RD model to fit degradation measurements has to be regarded as a mathematical artifact without physical meaning, or the structure of the $Si/SiO_2$ interface somehow accelerates the lateral equilibration considerably. We investigated the second option more closely by considering first-principles calculations that have shown a lowering of diffusion barriers for hydrogen (molecules) along the $Si/SiO_2$-interface as compared to the bulk $SiO_2$ [37]. These findings indicate that the motion of hydrogen might proceed at a much higher rate along the interface. As a higher diffusivity at the interface aids the lateral equilibration, it might be the sought process that makes the one-dimensional RD model physically meaningful. To account for it in our microscopic model, we applied different diffusion coefficients $D_\mathrm{I}$ and $D_\mathrm{B}$ in the interface region and in the bulk, respectively.

As can be seen in Fig. 15.20, the increase of the interface diffusion coefficient accelerates the degradation during the initial phase as it increases the transport of hydrogen atoms away from their dangling bonds. Interestingly, even if the interface diffusion coefficient is increased by four orders of magnitude there is no $t^{1/6}$ behavior visible, but instead the degradation takes on the typical $t^{1/4}$ behavior of a hydrogen-only reaction–diffusion model. While the competition for dangling bonds sets in earlier for increased interface diffusion coefficients, the formation of $H_2$ is not accelerated in the same way. Inspection of the atomic diffusion shows that the acceleration of the dimerization is much less pronounced as the liberated hydrogen atoms constantly leave the interface region into the bulk where the diffusion proceeds slower and the collision rate is reduced. Only in the limit

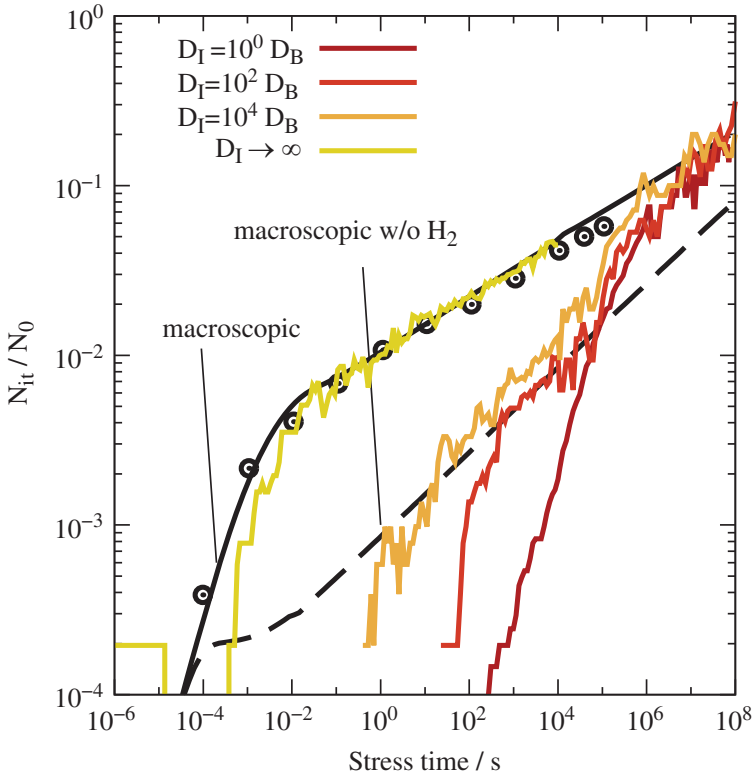


**Fig. 15.20** For increasing $D_I$, the departure of hydrogen atoms from their dangling bond sites starts earlier, leading to an increased degradation at earlier times. Comparison to the classical RD model without $H_2$ formation shows that competition for dangling bonds sets in after about 100 s, leading to a $t^{1/4}$ degradation. The formation of $H_2$ is slowly accelerated by the increased $D_I$ and only for $D_I \to \infty$, the macroscopic behavior is obtained

of $D_I \to \infty$ will the microscopic RD model match the experimentally observed behavior. Although these extremely high interface diffusion coefficients lack any physical justification, this is still closer to the physical reality than the assumption of immediate equilibration along the Si–$SiO_2$-interface at any depth that is inherent to the usually employed one-dimensional macroscopic RD model.

As a side note we remark that in a real wafer, a nearly infinite diffusion coefficient along the Si/$SiO_2$-interface would make the hydrogen spread out through the waver during stress. This would again alter the degradation slope and give rise to cross-talk between neighboring devices that would be measurable, but has never been reported.

## 15.7 Related Work

Four other scientific groups have put forward microscopic RD models recently [21, 38, 39] and interestingly those investigations find a reasonable agreement between their microscopic description and the macroscopic RD model. In the work of Islam et al. [21] the atomic description is basically equivalent to the work presented

here but is built upon a one-dimensional foundation which carries the same implicit assumptions as the macroscopic model. Clearly this model cannot capture the effects discussed in this chapter as those are solely due to higher-dimensional effects. From a physical point of view, however, the one-dimensional approximation lacks justification considering the results presented above.

The work of Choi et al. [39] considers the three-dimensional diffusion of the particles based on a grid-less stochastic formulation. Although the degradation in that work seems to match the macroscopic RD model quite well at first sight, also strong discrepancies arise between the two for longer stress times. Interestingly, for situations where the approach presented above predicts a degradation far below the prediction of the macroscopic model, the degradation predicted by Choi et al. overshoots the macroscopic model considerably. Only for an enormous density of dangling bonds or a very large reaction radius the macroscopic behavior is obtained, in accord with our results. The degradation behavior in [39] initially follows $N_{\mathrm{it}}(t) = k_{\mathrm{f}} t$, which suggests that the depassivated hydrogen atoms instantly leave the reaction radius of their respective dangling bond. The following excessively high power-law exponent suggests that the repassivation of the silicon dangling bonds is somehow inhibited in this formulation. The most likely explanation for this behavior is a too low resolution of the time-stepping, in combination with the physically unjustifiable description of the diffusive motion.

The work of Panagopoulos and Roy [38] uses a grid-based stochastic RD model that seems to be compatible with our description. The surprisingly good agreement between their results and the macroscopic RD model may be an artifact of the employed method which is based on an adaptive time-stepping. Also, the paper states that the passivation reaction occurs if a hydrogen atom is "close" to a dangling bond. This indicates an artificial capture radius, but this is not explicitly stated. Also, the grid spacing is not given in the paper and its physical relevance is not discussed. However, as shown by our calculations, an unphysically large grid spacing strongly promotes bimolecular reactions and thus induces a degradation behavior that is (falsely) compatible with the macroscopic RD model.

Finally, Naphade et al. [40] recently presented a stochastic version of the poly H/$H_2$ RD model. In this approach it is assumed that the diffusion of H is restricted to the oxide, while the diffusion of $H_2$ happens in the gate contact and beyond. The large H diffusion coefficient in these calculations in combination with the large grid size of 1 nm reduces the effect of the diffusion limitation on the bimolecular reactions. Although this model is formulated on a stochastic description, and is in better agreement with the experimental data, its physical validity is again questionable due to the assumed H diffusion coefficient of $10^{-5}\,\mathrm{cm}^2/\mathrm{s}$. At this diffusion coefficient, the hydrogen diffusion front would extend to 1 cm after 100 ks of stress. Even if the assumption that hydrogen is unable to penetrate into the gate contact holds, the isotropic nature of the diffusion process would lead to a lateral diffusion whose front quickly exceeded the dimensions of the MOS device. This would lead to a considerable out-diffusion of hydrogen from the gate area, resulting in a sharp increase of $N_{\mathrm{it}}$ and again a destruction of the $t^{1/6}$ power-law degradation. In the calculations of Naphade et al., however, this effect is not present due to the

reflecting lateral boundary conditions which in this case are not justifiable anymore. Apart from these issues, the poly $H/H_2$ RD model shares the shortcomings of all RD-based models with respect to the prediction of NBTI recovery. Based on our calculations with increased interfacial diffusion coefficients, which corresponds to the increased oxide diffusion coefficient in the model of Naphade et al., we expect this problem to be also present in the stochastic version.

## 15.8 Conclusion

Our work shows that the reaction–diffusion model for the negative bias temperature instability, which has been used for nearly 40 years to interpret experimental data, has a number of inherent assumptions on the underlying physics that lack any physical justification. Those are:

1. *Continuous diffusion in the sub-nm regime.* Diffusion of neutral hydrogen atoms and $H_2$ proceeds via jumps between the interstitial sites of the host material. Positional changes that are smaller than about 4 Å are atomic vibrations around an equilibrium position and thus not diffusive in nature. This is especially relevant as in the macroscopic modified $H–H_2$ RD model, the onset of the power-law regime is quite discretization dependent.
2. *Instantaneous well-stirredness along the interface.* The one-dimensional macroscopic RD model, which gives the experimentally relevant $t^{1/6}$ behavior, inherently assumes that all hydrogen atoms that are liberated during stress instantaneously compete with all other hydrogen atoms at the interface for available dangling bonds or dimerize with each other. However, at typically assumed dangling bond densities of $5 \times 10^{12}\,\mathrm{cm}^{-2}$, the distance between two dangling bonds will be about 4.5 nm. At a depassivation level of 1% this means that the average initial distance between two hydrogen atoms is even in the range of 45 nm. The reduction of this distance to the typical $H_2$ bonding distance of 0.7 Å [30] needs to be overcome by a diffusion step, which takes about 200 s at a diffusion coefficient of $10^{-13}\,\mathrm{cm}^2/\mathrm{s}$.
3. *Rate-equation-based description.* It is well established in chemical literature that bimolecular reactions are not sufficiently described by reaction rate equations if the particle numbers are small. In a reaction rate equation system it is for instance possible for 0.5 H atoms to form 0.25 $H_2$, which is physically meaningless. An accurate description in the limit of small particle numbers is only obtained from an atomistic description.

We have implemented a stochastic three-dimensional modified reaction–diffusion model for NBTI to study the degree to which a more realistic description changes the predicted behavior. The model is theoretically well founded on the theory of stochastic chemical kinetics and is understood as a consequent realization of the physical picture behind the reaction–diffusion theory.

The degradation predicted by the microscopic model features a unique new initial regime in which the motion of each hydrogen atom is completely independent from the others. This regime features a strongly increased power-law exponent that is not observed experimentally, yet it is a necessary consequence of the liberation of hydrogen during stress. Application of the atomic RD model to a real-world example shows that for a realistic jump width it is impossible to obtain the experimentally observed behavior due to the apparent diffusion limitation of the dimerization and passivation rates. The match of the microscopic model with the macroscopic version and experimental data can be improved by using an increased diffusion coefficient at the interface. However, the required diffusion coefficients are many orders of magnitude above $10^{-9}\,\mathrm{cm}^2/\mathrm{s}$, which corresponds to a diffusion length of 100 µm after 100 ks. The lateral diffusion of the hydrogen in this case would reach way beyond the dimensions of individual microelectronic devices, leading to cross talk and a dramatically increased degradation due to the loss of hydrogen.

The recovery predicted by the microscopic model matches the macroscopic counterpart as soon as the previous degradation has entered the classical diffusion-limited regime. This behavior is due to the prerequisite that the system has to be equilibrated along the interface before the $t^{1/6}$ regime can emerge. As the recovery happens on much larger time-scales than the stress duration, lateral equilibration effects are invisible in recovery traces. A distribution of arrival times as predicted by the simple estimate using different diffusion coefficients during recovery as in [11] could not be found.

In summary, our study of the microscopic limit reveals a number of serious problems in the traditional mathematical formulation of the reaction diffusion model for NBTI, rendering all variants that are based on partial differential equations physically meaningless. In a physically meaningful microscopic version of the model, no experimental feature remains that can be accurately predicted. The apparent match of the RD models with experimental data must therefore be considered a mathematical artifact without any physical background.

**Acknowledgements** This work has received funding from the EC's FP7 grant agreement NMP.2010.2.5-1 (MORDRED).

## References

1. K. Jeppson, C. Svensson, J.Appl.Phys. **48**(5), 2004 (1977)
2. D.K. Schroder, Microelectronics Reliability **47**, 841 (2007)
3. H. Kufluoglu, M. Alam, IEEE Trans.Electron Devices **53**(5), 1120 (2006). DOI 10.1109/TED.2006.872098
4. T. Grasser, W. Goes, B. Kaczer, IEEE Trans.Device and Materials Reliability **8**(1), 79 (2008). DOI 10.1109/TDMR.2007.912779
5. S. Ogawa, N. Shiono, Physical Review B **51**(7), 4218 (1995)
6. B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, M. Goodwin, In *Proc. Intl.Rel.Phys.Symp.* (2005), pp. 381–387

7. H. Reisinger, O. Blank, W. Heinrigs, A. Mühlhoff, W. Gustin, C. Schlünder, In *Proc. Intl.Rel.Phys.Symp.* (2006), pp. 448–453
8. T. Grasser, W. Goes, V. Sverdlov, B. Kaczer, In *Proc. Intl.Rel.Phys.Symp.* (2007), pp. 268–280. DOI 10.1109/RELPHY.2007.369904
9. T. Grasser, B. Kaczer, W. Gös, H. Reisinger, T. Aichinger, P. Hehenberger, P.J. Wagner, F. Schanovsky, J. Franco, P.J. Roussel, M. Nelhiebel, In *Proc. Intl.Electron Devices Meeting* (2010), pp. 82–85
10. S. Mahapatra, V.D. Maheta, A.E. Islam, M.A. Alam, IEEE Trans.Electron Devices **56**(2), 236 (2009)
11. S. Mahapatra, A. Islam, S. Deora, V. Maheta, K. Joshi, A. Jain, M. Alam, In *Proc. Intl.Rel.Phys.Symp.* (2011), pp. 6A.3.1 –6A.3.10. DOI 10.1109/IRPS.2011.5784544
12. S. Mahapatra, A. Islam, S. Deora, V. Maheta, K. Joshi, M. Alam, In *Proc. Intl.Symp. on Physical and Failure Analysis of Integrated Circuits* (2011), pp. 1–7. DOI 10.1109/IPFA.2011.5992794
13. K. Joshi, S. Mukhopadhyay, N. Goel, S. Mahapatra, In *Proc. Intl.Rel.Phys.Symp* (2012), pp. 5A.3.1–10
14. A. Islam, H. Kufluoglu, D. Varghese, S. Mahapatra, M. Alam, IEEE Trans.Electron Devices **54**(9), 2143 (2007). DOI 10.1109/TED.2007.902883
15. A.E. Islam, H. Kufluoglu, D. Varghese, M.A. Alam, Appl.Phys.Lett. **90**(8), 083505 (2007). DOI: 10.1063/1.2695998. URL http://dx.doi.org/doi/10.1063/1.2695998
16. H. Kufluoglu, M. Alam, IEEE Trans.Electron Devices **54**(5), 1101 (2007)
17. A. Islam, H. Kufluoglu, D. Varghese, M. Alam, Appl.Phys.Lett. **90**(1), 083505 (2007)
18. F. Schanovsky, T. Grasser, In *Proc. Intl.Integrated Reliability Workshop* (2011), pp. 17–21
19. F. Schanovsky, T. Grasser, In *Proc. Intl.Rel.Phys.Symp* (2012), pp. XT.10.1–6
20. A. Stesmans, B. Nouwen, V.V. Afanas'ev, Phys. Rev. B **58**, 15801 (1998). DOI 10.1103/PhysRevB.58.15801. URL http://link.aps.org/doi/10.1103/PhysRevB.58.15801
21. A. Islam, M. Alam, J.Comp.Elect. pp. 1–11 (2011). URL http://dx.doi.org/10.1007/s10825-011-0369-4. 10.1007/s10825-011-0369-4
22. D.A. McQuarrie, J.Appl.Prob **4**(3), 413 (1967)
23. P. Hänggi, P. Talkner, M. Borkovec, Rev.Mod.Phys **62**(2), 251 (1990)
24. S. Torquato, C.L.Y. Yeong, J.Chem.Phys. **106**, 8814 (1997)
25. S.S. Andrews, D. Bray, Phys.Biol. **1**, 137 (2004)
26. R. Erban, S.J. Chapman, Phys.Biol. **6**, 046001 (2009)
27. S.A. Isaacson, D. Isaacson, Physical Review E **80**, 066106 (2009)
28. D. Fange, O.G. Berg, P. Sjöberg, J. Elf, Proc.Nat.Acad.Sci. **107**(46), 19820 (2010)
29. G. Malavasi, M.C. Menziani, A. Pedone, U. Segre, Journal of Non-Crystalline Solids **352**(3), 285 (2006). DOI 10.1016/j.jnoncrysol.2005.11.022. URL http://www.sciencedirect.com/science/article/pii/S0022309305007994
30. P.E. Blöchl, Physical Review B **62**(10), 6158 (2000)
31. A. Bongiorno, L. Colombo, F. Cargnoni, Chem.Phys.Lett. **264**, 435 (1997)
32. D. Gillespie, J.Comp.Phys. **22**, 403 (1976)
33. D.T. Gillespie, in *Proc. Int. Conf. Form. Meth. Sys. Bio.* (Springer-Verlag, Berlin, Heidelberg, 2008), SFM'08, pp. 125–167. URL http://dl.acm.org/citation.cfm?id=1786698.1786704
34. T. Grasser, Microelectronics Reliability **52**(1), 39 (2012). DOI 10.1016/j.microrel.2011.09.002
35. V. Huard, M. Denais, C. Parthasarathy, Microelectronics Reliability **46**(1), 1 (2006)
36. B. Tuttle, **61**(7), 4417 (2000)
37. S.T. Pantelides, L. Tsetseris, S. Rashkeev, X. Zhou, D. Fleetwood, R. Schrimpf, Microelectronics Reliability **47**(6), 903 (2007). DOI DOI: 10.1016/j.microrel.2006.10.011. URL http://www.sciencedirect.com/science/article/pii/S0026271406003817
38. G. Panagopoulos, K. Roy, IEEE Trans.Electron Devices **58**(8), 2337 (2011). DOI 10.1109/TED.2011.2148720
39. S. Choi, Y. Park, C.K. Baek, S. Park, in *Proc. Simu.Semicond.Proc.Dev.* (2012), pp. 185–188
40. T. Naphade, N. Goel, P.R. Nair, S. Mahapatra, in *Proc. Intl.Rel.Phys.Symp.* (2013)