Quantum Modeling of Semiconductors Leakage Currents Induced by Defects

Antoine Jay, C. Mesnard, I. Nicholson, R. Helleboid, G. Mugny, D. Rideau, V. Goiffon, L. Martin Samos, N. Richard, Anne Hémeryck

Quantum Modeling of Semiconductors Leakage Currents Induced by Defects

A. Jay[1], C. Mesnard[1], I. Nicholson[2], R. Helleboid[2], G. Mugny[2], D. Rideau[2], V. Goiffon[3], L. Martin Samos[5], N. Richard[4], A. Hémeryck[1]
[1]*LAAS CNRS, Université de Toulouse*, CNRS Toulouse, France [2]*ST Microelectronics*, Crolles, France, [3]*ISAE Supaéro*, Toulouse, France [4]*CEA, DAM, DIF*, Arpajon, France [5]*CNR IOM*, Trieste, Italie

***Abstract*—We present a full quantum calculation of the electron-hole generation/recombination rate caused by atomic-scale defects, which are known to be at the origin of the noise in semiconductors. While numerous attempts have been made to model this non-radiative multiphonon processes (MPE) using analytical approximations, only a few have reported numerical solutions that consider both the electronic and ionic overlaps. In this study, we tackle this issue within the framework of a complete *ab initio* calculation, which involves computing the dynamical matrix in large supercells. To address the challenge of calculation the ionic integral, we propose a semi-classical method.**

***Index Terms*—Dark current, electronic cross section, multi-phonon processes, deep center, semiconductors, ionic wave functions**

I. Origin of the problem

Electronic devices are degraded when they are used in aggressive radiating environments such as nuclear, spatial, medical, or during the implantation processes. The incoming particles destroy the crystalline structure of the semi-conductor by crossing it, leaving in its wake some atomic scale defects. These defects induce new electronic states in the semi-conductor band gap, which can act as bridges for the electrons to thermally cross the band gap, or as traps for free carriers. The so created current generates a noise known as the Shockley Read Hall (SRH [1], [2], [3]) generation rate, which is detrimental to the device performance [4], [5].

To comprehensively model and understand the impact of defects on electronic devices, many physical processes must be taken into account, each of them requiring its own space and time scale to be representative of the phenomenon. These processes include: *(i)* the density functional theory (DFT [6], [7]) to describe the electronic structure, *(ii)* the GW approximation [8], [9] to solve the band gap problem and determine ionization potential and electronic affinities of the defect (electronic trap energy E_t), *(iii)* the MPE theory [10] for evaluating the electrons and holes capture cross sections (σ_n and σ_p), and finally *(iv)* the well-established SRH theory and the detailed balance to calculate the generation rate R_{SRH} induced by each defect as a function of the device properties:

$$R_{SRH} = \frac{v_{th}\sigma_n\sigma_p(n_i^2 - np)}{\sigma_n[n + n_i exp(\frac{E_t - E_F}{k_B T})] + \sigma_p[p + n_i exp(\frac{E_t - E_F}{k_B T})]}$$

Concerning SRH, while most of the parameters are generally known, (Fermi energy E_F, electronic thermal velocity v_{th}, concentration of free carriers p and n, n_i the intrinsic carrier density and E_t trap defect energy), there is a clear lake of knowledge regarding the electronic cross sections σ, which are the theoretical bottleneck of the modeling. The following explains how to calculate them numerically within a full quantum perspective.

II. *Ab initio* calculation of the electronic capture cross sections

The evolution of the system is governed by the Hamiltonian $\hat{\mathscr{H}}$ as $\hat{\mathscr{H}}\Psi_n(\mathbf{r},\mathbf{R}) = E_n^{tot}\Psi_n(\mathbf{r},\mathbf{R})$ where the wave functions $|\Psi_n\rangle$ depends on the electronic $\mathbf{r}$ and ionic $\mathbf{R}$ positions. The generation rate $\Gamma_{i\to f}$ from an initial non excited state $|\Psi_i(\mathbf{r},\mathbf{R})\rangle$ to a final excited state $|\Psi_f(\mathbf{r},\mathbf{R})\rangle$ is governed by the Fermi golden rule:

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar} |\langle \Psi_f(\mathbf{r},\mathbf{R}) | \Delta\mathscr{H} | \Psi_i(\mathbf{r},\mathbf{R}) \rangle|^2 (E_i^{tot} - E_f^{tot})$$

This scalar product is an overlap integral between the modified initial wave function and the final wave function. This scalar product is by definition null when the system is not perturbed ($\Delta\mathscr{H}$=0) and becomes not null for a perturbed system. In the case of the SRH, the electronic excitation is thermally activated, meaning that this perturbation comes from the vibrations modes of the defect, and the fact that the equilibrium positions of the final state differs from the one of the initial state: $\mathbf{R_f} \neq \mathbf{R_i}$. This has induced the name of multiphonon assisted emission (MPE) [10]. The corresponding electronic cross sections C are generally small (C$\approx 10^{-17}$ cm^{-2}) if the needed perturbation (atomic displacement) is large [11], and high otherwise (C$\approx 10^{-13}$ cm^{-2}) as less phonons are needed for this later type of transition. This large versus small cross section is typically determined by

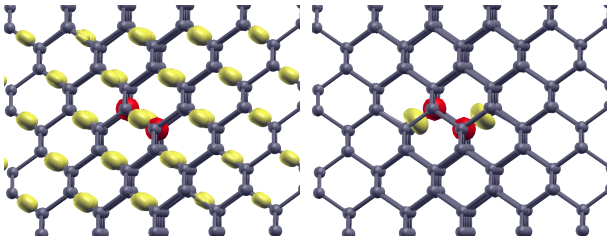
the size (spatial extension) of the defect, an example of calculation is shown in Ref. [12].

To solve the Fermi golden rule, the first step is the Born-Oppenheimer approximation which expresses the fact that the electrons instantaneously follow the ionic positions. The global wave function can be written as a product of an electronic and an ionic one, $\Psi_n(\mathbf{r},\mathbf{R}) = \phi_n(\mathbf{r},\mathbf{R})\chi_n(\mathbf{R})$, which permits to split the initial Hamiltonian into an electronic one at fixed atomic coordinates and a ionic one:

$$\hat{\mathscr{H}}^{elec}_{\mathbf{R}}\phi_n(\mathbf{r},\mathbf{R}) = E_n^{BO}(\mathbf{R})\phi_n(\mathbf{r},\mathbf{R}) \quad (1)$$

$$\hat{\mathscr{H}}^{ion}\chi_s(\mathbf{R}) = E_s\chi_s(\mathbf{R}) \quad (2)$$

where the electronic one is solved within DFT framework, whereas, the ionic one is only solved semi-classically within density functional perturbation theory (DFPT [13]), giving the $3N_{at}$ vibrational modes Ω of the system. MPE is then the transition from a delocalized electronic Bloch wave function of the bulk (Fig. 1a) to an electronic wave function localized near the defect (Fig. 1b). It is followed by the variation of the defect charge state and of the atomic positions.



(a) $|\phi_i|^2$ valence: delocalized (b) $|\phi_f|^2$ defect: localized

Figure 1: Initial and final states for the electronic wave functions (yellow) around a defect site (the divacancy, red balls) in silicon (gray balls).

In the diabatic-static model and first order Taylor-expansion, the Fermi golden rule can be rewritten [14], as:

$$\Gamma_{im\to fn} = \frac{2\pi}{\hbar}\Big|\sum_{k}^{N_{modes}} \underbrace{\langle\phi_i|\frac{\partial H}{\partial \mathbf{Q_k}}|\phi_f\rangle}_{electrons}\underbrace{\langle\chi_{im}|\Delta\mathbf{Q_k}|\chi_{fn}\rangle}_{ions}\Big|^2 \times\left(E_i^{BO}-E_f^{BO}+(\frac{1}{2}+m)\hbar\Omega_i-(\frac{1}{2}+n)\hbar\Omega_f\right) \quad (3)$$

where $\mathbf{Q_k}$ are known as the generalized coordinates associated to the vibrational mode k, and m (n) the vibrational state of the initial (final) modes presented in Fig. 3. The electronic overlap can then be calculated within DFT in the one dimension approximation [14] as shown in Fig.2, whereas the ionic overlap is generally treated analytically [15], [16].

We also note incidentally that using the Fermi golden rule in this case is an approximation because its validity is limited to the transitions between quantum states of the same Hamiltonian, whereas the wave functions used in Eq. 3 have been calculated using Eq. 1 with two different Hamiltonians: the first is expressed at the initial atomic positions and the second at the final ones.

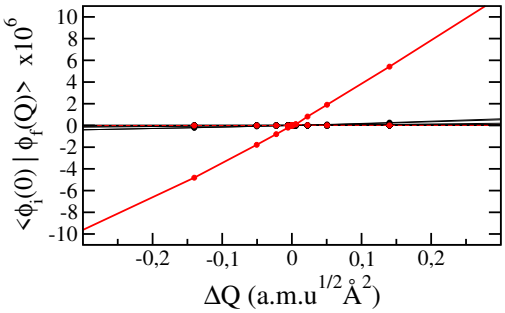


Figure 2: Numerical calculation of the electronic overlap using $\langle\phi_i|\frac{\partial H}{\partial \mathbf{Q}}|\phi_f\rangle = \frac{\partial\langle\phi_i|\phi_f\rangle}{\partial \mathbf{Q}}(E_f^{BO}-E_i^{BO})$. Only the 5 first valence bands and the defects state are shown.

III. Construction of the ionic wave function and its overlap integral

To construct the ionic wave function of a defective supercell and evaluate the ionic part of the scalar product, the first step is to calculate the dynamical matrix in the ground and excited states. As each eigenvector $\mathbf{e_i}$ of the dynamical matrix is a 1D direction into a $3N_{at}$D space, the idea to obtain the corresponding ionic wave functions (needed for ionic overlap calculation) is to expand the total atomic displacement along the independent eigenvectors. The ionic Hamiltonian $\mathscr{H}^{ions}$ can then be split into separated Hamiltonians in each eigenvector dimension:

$$\mathscr{H}^{ions} = \mathscr{H}_1 \otimes \mathscr{H}_2 \otimes \cdots \otimes \mathscr{H}_{3N_{at}} \quad (4)$$

This permits to give for each mode i the analytic solution of the 1D harmonic oscillator known as the Hermit polynomials $H_i(x)$ represented in one dimension Fig. 3. In this 1D solution, the order of the Hermit polynomial of each mode is the Bose-Einstein phonon population of the eigenmode. The subscript m (n) that describes the vibronic state of the initial (final) ionic state Eq. 3 then contains the vibronic state of all the 1D oscillators. This assumption is relevant only for small atomic displacements that typically preserve correct the harmonic approximation.

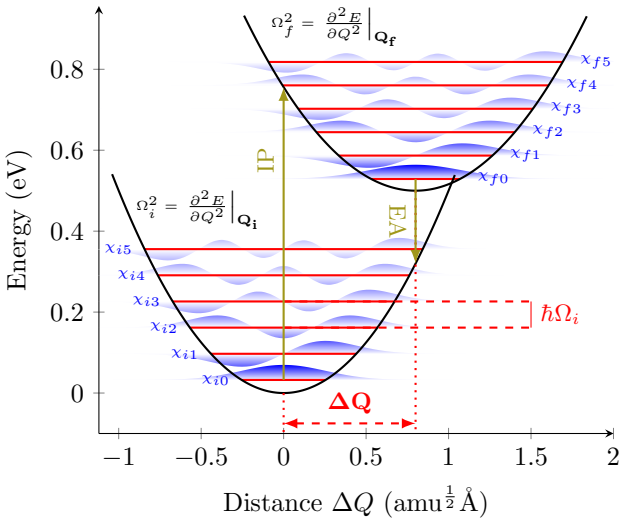


Figure 3: One dimension model of harmonic oscillators in the initial and final states. The blue clouds represents the Hermit polynomials, exact analytic solutions of the 1D oscillator. The general coordinate separating the two equilibrium positions is ΔQ. The order of the polynomial is determined by the temperature as the Bose-Einstein phonon population. The ω^2 are the eigenvalues of the dynamical matrix, second order derivative of the parabola. IP and EA are respectively the ionization potential and electronic affinity as calculated within the GW approximation.

The $3N_{at}$ ionic wave function can then be written as the product of each 1D wave function H_i:

$$\chi(\mathbf{R}) = \prod_{i=1}^{3N_{at}} H_i(x_i) \tag{5}$$

and permits to recover the known snapshots of Figs. 5b and 5c. The x_i are the coordinates of $\mathbf{R}$ into the eigenvector basis: $\mathbf{R} = \sum_i^{3N_{at}} x_i \mathbf{e_i}$ where the equilibrium positions are the origin of the $3N_{at}$D space.

The ionic wave function $\chi(\mathbf{R})$ is a $3N_{at}$ dimensions function that requires a huge amount of memory to be stored. Ideally, it should be stored into a tabular containing $N_s^{3N_{at}}$ terms where N_s represents the number of elements used to sample each dimension. However, to study a defective supercell containing at least 100 atoms, N_s should be at least 250. The storage of this tabular is then clearly out of the present computational possibilities.

To solve this problem, it is possible to make the further assumption that each atom has its own wave function which do not overlap with the ones of the other atoms. Once again, this assumption is still valid in the limit of small temperatures, as the displacement of the atoms is restricted around their equilibrium position.

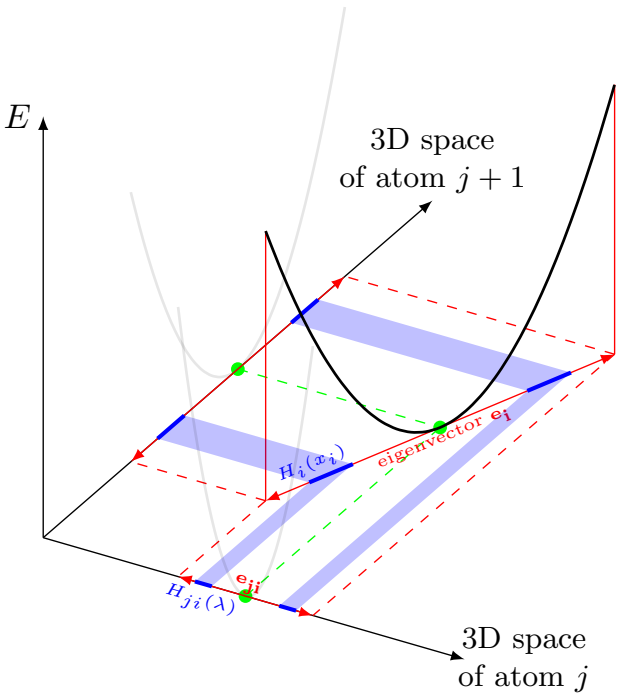


Figure 4: Schematic representation of the 1D solution in the $3N_{at}$D space projected into the 3D space of each atom. The red arrow annotated "eigenvector" represents a 1D vector into the $3N_{at}$D space, the blues lines on it are the highest values of the Hermit polynomial, and the green dot the equilibrium position. These are projected into the 3D space of each atom (the two axis). The black parabola represents the variation of the energy along the displacement induced by the eigenvector, for which the second order derivative (frequencies squared as eigenvalues of the dynamical matrix) gives the degree of Hermit polynomial (phonon population) and its amplitude.

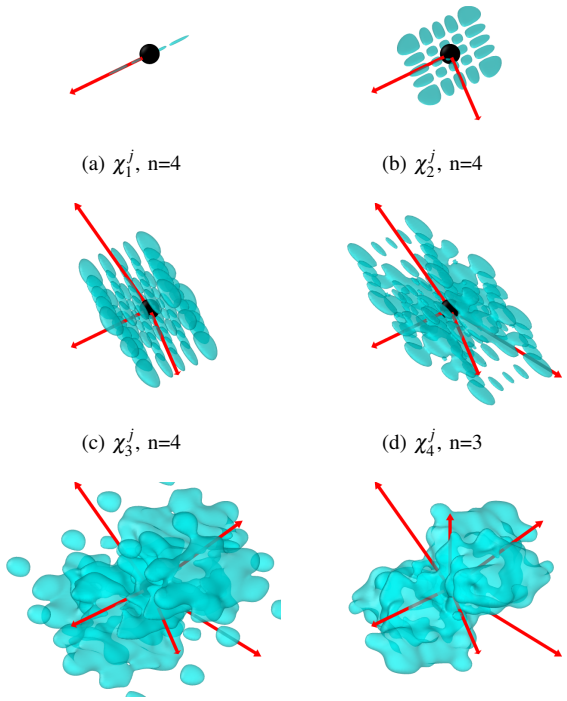
In the 3D space, the 1-atom wave function of the j-th atom $\chi^j(\mathbf{R})$ is then constructed iteratively as:

$$\chi_0^j(\mathbf{R}) = \begin{cases} 1 & \text{if} \quad \mathbf{R} = \mathbf{R}_{at_j} \\ O & \text{if} \quad \mathbf{R} \neq \mathbf{R}_{at_j} \end{cases} \tag{6}$$

$$\chi_{i+1}^j(\mathbf{R}) = \int \chi_i(\mathbf{R} - \lambda \mathbf{e_{ji}}) H_{ji}(\lambda) d\lambda \tag{7}$$

$$\chi^j(\mathbf{R}) = \chi_{3N_{at}}^j(\mathbf{R}) \tag{8}$$

Computationally, the wave function is initialized by sampling the 3D supercell space into $N_x \times N_y \times N_z$ meshes with value 0 everywhere except on the j-th atomic position (Equ. 6). This starting wave function is then extended along the first eigenvector $\mathbf{e_1}$ projected on the 3D subspace of the j-th atom $\mathbf{e_{j1}}$ (red arrow Fig. 5a) and multiply by H_{j1} (Equ. 7). H_{j1} is the Hermit polynomial solution of the first 1D oscillator H_1 (used in equ. 5) but projected on $\mathbf{e_{j1}}$, as represented in Fig. 4. This step is



(a) χ_1^j, n=4 (b) χ_2^j, n=4

(c) χ_3^j, n=4 (d) χ_4^j, n=3

(e) χ_5^j, n=2 (f) χ_6^j, n=1

Figure 5: Iterative construction of a 1-atom ionic wave function including the $3N_{at}$ eigenmodes contributions. For clarity, only 6 modes have been taken into account : the ones for which the displacement of this atom is the highest. Blue cloud: wave function isovalue. Red arrows: displacement of the atom along each eigenmode. n: Bose-Einstein population.

recursively done for each of the $3N_{at}$ eigenvectors. The resulting construction of the 1-atom ionic wave function is presented in Fig. 5.

This method scales in $N_x \times N_y \times N_z \times N_\lambda \times 3N_{at}$. where N_λ is the sampling of the integral in Equ. 7. To reduce the difficulty of this task, the recursion can be limited to the most relevant eigenvectors of larger displacement. This reduce the $3N_{at}$ to roughly a dozen of phonons. In addition, the sampling of the 3D space and the boundaries of the integral can be both restricted around the studied atom, instead of covering all the supercell. To represent χ^{proj}, the $3N_{at}$D wave function projected into a 3D space, a simple sum of each atomic projection χ^j can be done, as their non zero value are restricted around the atomic equilibrium position, *i.e.* they do not overlap:

$$\chi^{proj}(\mathbf{R}) = \sum_{j=1}^{N_{at}} \chi^j(\mathbf{R}) \tag{9}$$

Note that this method completely loses the information concerning the motion of the atoms relatively to each others. Hopefully, this information it is not needed for the evaluation of the ionic overlap integral. Its validity is also restricted to the wave functions calculated at the Brillouin zone center only, which is the case for defects in supercells. Based on this approach, a rigorous calculation of the overlap is numerically possible by summing the contribution of the scalar product of each atomically projected wave function:

$$\langle \chi_{im} | \Delta\mathbf{Q_k} | \chi_{fn} \rangle \sim \sum_{j=1}^{N_{at}} \langle \chi_i^j | \mathbf{Q_{jk}} | \chi_f^j \rangle \tag{10}$$

References

[1] W. Shockley and W. Read, "Statistics of the recombinations of holes and electrons," Phys. Rev., vol. 87, no. 5, pp. 835–842, 1952.

[2] R. N. Hall, "Electron-hole recombination in germanium," Phys. Rev., vol. 87, p. 387, 1952.

[3] C. Sah, R. Noyce, and W. Shockley, "Carrier generation and recombination in p-n junctions and p-n junction characteristics," IRE Proc., vol. 45, pp. 1228–1243, 1957.

[4] D. Garetto, Y. Mamy Randriamihaja, A. Zaka, R. D, A. Schmid, H. Jaouen, and Y. Leblebici, "Analysis of defect capture cross sections using non-radiative multiphonon-assisted trapping model," Sol. St. Elec., vol. 71, pp. 74–79, 2012.

[5] D. Garetto, Y. M. Randriamihaja, D. Rideau, A. Zaka, A. Schmid, Y. Leblebici, and H. Jaouen, "Modeling stressed mos oxides using a multiphonon-assisted quantum approach—part ii: Transient effects," IEEE Trans. Elec. Dev., vol. 59, no. 3, pp. 621–630, 2012.

[6] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," Phys. Rev., vol. 136, p. B864, 1964.

[7] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," Phys. Rev., vol. 140, p. A1133, 1965.

[8] B. Fabien, "Gw approximation of the many-body problem and changes in the particle number," Phys. Rev. Lett., vol. 103, p. 176403, 2009.

[9] P. Rinke, A. Janotti, M. Scheffler, and C. V. de Walle, "Defect formation energies without the band-gap problem: Combining density-functional theory and the gw approach for the silicon self-interstitial," Phys. Rev. Lett., vol. 102, p. 026402, 2009.

[10] K. Huang and A. Rhys, "Theory of light absorption and non-radiative transitions in f-centres," Proc. Roy. Soc. A, vol. 204, pp. 406–423, 1950.

[11] J. Bourgoin and M. Lannoo, Point Defects in Semiconductors II. Springer Berlin Heidelberg, 1983.

[12] A. Jay, A. Hémeryck, F. Cristiano, D. Rideau, P. Julliard, V. Goiffon, A. LeRoch, N. Richard, L. Martin-Samos, and S. De Gironcoli, "Clusters of defects as a possible origin of random telegraph signal in imager devices: a dft based study," Simul. Semicond. Proc. and Devices, pp. 128–132, 2021.

[13] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory," Rev. Mod. Phys., vol. 73, p. 515, 2001.

[14] A. Alkauskas, M. McCluskey, and C. V. de Walle, "Tutorial: Defects in semiconductor: Combining experiment and theory," J. App. Phys., vol. 119, p. 181101, 2016.

[15] D. Goguenheim and M. Lannoo, "Theoretical and experimental aspects of the thermal dependence of electron capture coefficients," J. App. Phys., vol. 68, pp. 1059–1069, 1990.

[16] W. Goes, Y. Wimmer, A.-M. El-Sayed, G. Rzepa, M. Jech, A. Shluger, and T. Grasser, "Identification of oxide defects in semiconductor devices: A systematic approach linking dft to rate equations and experimental evidence," Microelec. Rel., vol. 87, pp. 286–320, 2018.