# *Modeling of trap assisted tunneling through thin dielectric layers*

**Giuseppe Iannaccone**
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni,
Università di Pisa

G. Iannaccone, *Modeling of trap assisted tunneling through thin dielectric layers*, Material Science and Technology, **18**, pp.736-738 (2002).

# Modelling of trap assisted tunnelling through thin dielectric layers

G. Iannaccone

**A detailed model is proposed of the leakage current due to tunnelling through defect traps in thin insulating layers. Such leakage currents are a major concern for the reliability of semiconductor devices based on the insulating properties of dielectric layers, such as metal–oxide–semiconductor transistors and memories, and of nanoscale devices whose operation is based on controlled tunnelling through thin dielectrics. The proposed model allows the addressing of both dc and noise properties of trap assisted currents. Results from a numerical implementation of the model are also presented.**

MST/5217

*The authors are in the Dipartimento di Ingegneria dell'Informazione, Università degli studi di Pisa, Via Diotisalvi 2, I–56122 Pisa, Italy (g.iannaccone@iet.unipi.it). Based on a presentation at the 3rd Euroconference on Nanoscience for Nanotechnology, held in Oxford, UK on 16–19 September 2000.*

## Introduction

The operation of several nanoscale electron devices is based on controlled tunnelling through thin insulating and quasi-insulating barriers: among them are single electron transistors, resonant tunnelling devices, and quantum cellular automata.[1] In addition, the operation of several conventional electron devices, such as metal–oxide–semiconductor (MOS) transistors and memories, relies on the insulating properties of silicon oxide layers whose thickness is in the range 2–10 nm. For this reason, the degradation of silicon oxide has always been a major concern for researchers in the field of semiconductor devices.[2]

The presence of electron traps in the insulator can lead to an intolerable increase of the leakage current, since trap assisted tunnelling can take place, i.e. electrons can tunnel through the dielectric in a two step process, first from one electrode to the trap, and then from the trap to the other electrode.

In particular, the observed non-repeatability of the $I-V$ characteristics of a device structure consisting of a dielectric layer stacked between two electrodes is often due to the presence of traps induced in the dielectric by a large applied electric field. Such traps degrade the insulating properties of the dielectric and lead to an increase of the leakage current by a few orders of magnitude.

This phenomenon was first observed in the context of MOS capacitors in the eighties,[3,4] and was widely studied thereafter, for its relevance to the reliability of non-volatile semiconductor memories and transistors.[5–7]

It is therefore very important to fully understand transport in the trap assisted tunnelling regime. In this paper, the author presents a model for dc properties and noise through an insulating layer in which traps are present with a known distribution in volume and energy. Noise, in particular, can provide deep insights into the transport mechanisms of such devices, since, as is well known, it is extremely sensitive to the presence of defects.

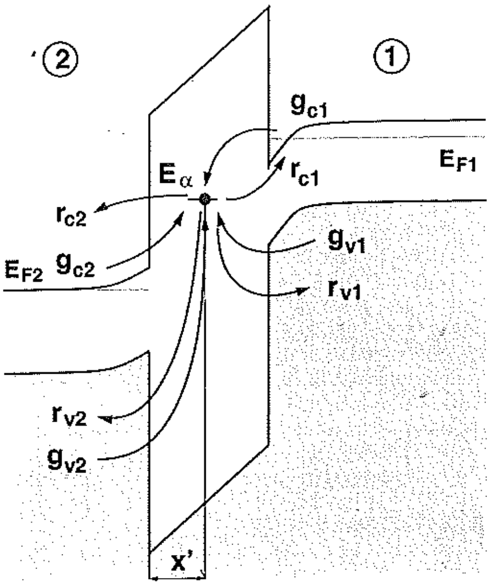
The paper is organised thus: in the next section the proposed model is described and in the subsequent section some results obtained from its numerical implementation are examined and discussed. Although MOS structures are often referred to, since they are the most relevant from an industrial point of view, the theory can be applied to any conductor–insulator–conductor structure.

## Model

For the sake of generality, let us consider the semiconductor–insulator–semiconductor structure whose conduction and valence bands are sketched in Fig. 1. In the case of metal contacts the situation is simpler, since only one band per electrode can be considered. In addition, let us consider a trap in the oxide, consisting of a localised electron state at position $x'$ in the oxide ($0<x'<d$) and at energy $E_\alpha$.

We will assume that the traps behave as acceptors (i.e., neutral when unoccupied) and, because of the Coulomb blockade, can practically be occupied by only one electron. Only a single energy level is considered with two possible states (spin up and down). This greatly simplifies the model, and has no significant effect from a quantitative point of view, since when one electron occupies a trap, the energy levels of other electron states in the same trap are so raised by the Coulomb energy that a second electron has only a vanishing probability of entering the trap.

The notation used in the case of generation–recombination processes[8,9] is followed: the 'generation' rate from this point of view is the transition rate from an electrode to the unoccupied trap, and the 'recombination' rate the transition rate from the occupied trap to one electrode. As can be seen in Fig. 1, four different generation rates are considered on



**1 Band profile of a semiconductor–insulator–semiconductor structure, with a trap of energy $E_\alpha$ at depth $x'$ and sketch of the eight transition rates considered**

DOI 10.1179/026708302225008839

the basis of the location of the initial state: generation rate from the conduction band of electrode 1 ($g_{c1}$), from the valence band of electrode 1 ($g_{v1}$), from the conduction band of electrode 2 ($g_{c2}$) from the valence band of electrode 2 ($g_{v2}$). The four recombination rates are defined analogously, on the basis of the location of the final state (the same subscript notation is used). Let us call $|\alpha\rangle$ the electron state in the trap, and let us consider a state $|\beta\rangle$ in the conduction band of region 1. According to the Fermi 'golden rule' the transition rate from $|\beta\rangle$ to $|\alpha\rangle$ would be

$$\nu_{\beta\to\alpha}=\frac{2\pi}{\hbar}\,|M(\alpha,\beta)|^2 h_\Gamma(E_\alpha-E_\beta) \quad . \; . \; . \; . \; . \; . \; . \; . \quad (1)$$

where $\hbar$ is the reduced Planck's constant, $M(\alpha,\beta)$ is the transition matrix element between state $|\alpha\rangle$ and $|\beta\rangle$, and $E_\alpha$ and $E_\beta$ are the energies of states $|\alpha\rangle$ and $|\beta\rangle$, respectively. the function $h_\Gamma$ is a Lorentian curve of halfwidth $\Gamma$

$$h_\Gamma(E_\alpha-E_\beta)=\frac{\Gamma/\pi}{(E_\alpha-E_\beta)^2+\Gamma^2} \quad . \; . \; . \; . \; . \; . \; . \; . \quad (2)$$

and represents the simplest way to account for inelastic transitions. As can be noticed, $h_\Gamma$ tends to a delta function as $\Gamma$ approaches 0, i.e. when only elastic transitions are considered. The larger $\Gamma$, the larger degree of inelastic transitions is allowed.

The transition rate can also be related to the probability current density $J(\beta,x')$ of state $|\beta\rangle$ on the plane $x'$ where the trap is located through the so called capture cross section $\sigma_{\alpha,\beta}$

$$\nu_{\beta\to\alpha}=\sigma_{\alpha,\beta}J(\beta,x')=\sigma_{\alpha,\beta}T_1(E_l)\nu_1(E_l) \quad . \; . \; . \; . \; . \; . \; . \quad (3)$$

where $E_l$ is the energy in the $x$ direction of state $|\beta\rangle$, $T_1(E_l)$ is the transmission probability of the one-dimensional barrier from $x'$ to $d$, and $\nu_1$ is the so called attempt frequency of the state of longitudinal energy $E_l$. The trap cross-section can depend of course on the trap state and on the initial state in a non-trivial way. However, given our lack of knowledge on the nature of traps, the simplest assumption is made that agrees with equation (1): $\sigma_{\alpha,\beta}=kh_\Gamma(E_\alpha-E_\beta)$, where $k$ is a constant. The cross-section has therefore a Lorentian behaviour as a function of the difference between the initial and final energy: elastic transitions are favoured (zero difference), and the higher the difference between the initial and final energies, the less likely the transition.

The state $|\beta\rangle$ is defined by its longitudinal energy $E_l$, its energy in the transverse plane $E_T$ ($E_\beta=E_l+E_T$), and its spin. The generation rate $g_{cl}$ is obtained by integrating equation (3) over all occupied states in the conduction band of electrode 1

$$g_{cl}=2\int_{E_{cl}}^{\infty}\mathrm{d}E_l\int_0^{\infty}\mathrm{d}E_T k h_\Gamma(E_l+E_T-E_\alpha)T_1(E_l)\nu_1(E_l)$$
$$\times f_1(E_l+E_T)\rho_1(E_l)\rho_T \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (4)$$

The factor 2 takes into account spin conservation, $\rho_l$ and $\rho_T$ are the densities of states in the longitudinal direction and in the transversal plane respectively, and $f_1(E_l+E_T)$ and $E_{c1}$ are the Fermi–Dirac occupation factor and the conduction band edge in the first electrode respectively.

The recombination rate $r_{c1}$ has an expression very similar to equation (4), with the difference that the integral has to be performed over unoccupied states in the conduction band of electrode 1 with the same spin of the trapped electron

$$r_{1c}=\int_{E_{cl}}^{\infty}\mathrm{d}E_l\int_0^{\infty}\mathrm{d}E_T k h_\Gamma(E_l+E_T-E_\alpha)T_1(E_l)\nu_1(E_l)$$
$$\times[1-f_1(E_l+E_T)]\rho_1(E_l)\rho_T \quad . \; . \; . \; . \; . \; . \; . \; . \quad (5)$$

At this point, the expressions of the other transition rates can be derived straightforwardly, and they will not be written in detail. Transition rates can be grouped as follows

$$\left.\begin{array}{ll} g_1=g_{c1}+g_{v1} & r_1=r_{c1}+r_{v1} \\ g_2=g_{c2}+g_{v2} & r_2=r_{c2}+r_{v2} \end{array}\right\} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad (6)$$

The occupation factor $f'$ of the trap in the steady state regime can be readily obtained by imposing the detailed balance of generation and recombination

$$(1-f')(g_1+g_2)=f'(r_1+r_2)\to f'=\frac{g_1+g_2}{g_1+g_2+r_1+r_2} \qquad (7)$$

The average current $I'$ through the trap can therefore be written as

$$I'=qg_1(1-f')-qr_1f'=qr_2f' \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad (8)$$

while the noise spectral density of the noise current at zero frequency is readily obtained using a procedure very close to that used for obtaining generation–recombination noise and noise in resonant tunnelling structures[10]

$$S'=2qI'\left(1-\frac{g_1r_2}{(g_1+r_1+g_2+r_2)^2}\right)=2q\gamma'I' \quad . \; . \; . \quad (9)$$

The shot noise suppression factor $\gamma'$, or Fano factor, is defined as $\gamma'=S'/2qI'$: as can be seen from equation (9), its value is between 0·5 and 1.[9]

Let us assume that traps are distributed with a density $\eta$ per unit volume per unit energy. The total trap assisted current density $J_{TAT}$ and the associated noise spectral density $S_{TAT}$ can be obtained by integrating $I'$ and $S'$ over $E_\alpha$ in the insulator gap, and $x'$, in the longitudinal direction from 0 to $d$, i.e.

$$\left.\begin{array}{l} J_{TAT}=\iint I'\eta(E_\alpha,x')\,\mathrm{d}E_\alpha\,\mathrm{d}x' \\ S_{TAT}=\iint S'\eta(E_\alpha,x')\,\mathrm{d}E_\alpha\,\mathrm{d}x' \end{array}\right\} \quad . \; . \; . \; . \; . \; . \; . \quad (10)$$

$J_{TAT}$ is proportional to the product of the capture cross-section and the trap density, while the Fano factor

$$\gamma_{TAT}\equiv\frac{S_{TAT}}{2qJ_{TAT}}=\frac{\iint\gamma'I'\,\mathrm{d}x'\,\mathrm{d}E_\alpha}{\iint I'\,\mathrm{d}x'\,\mathrm{d}E_\alpha} \quad . \; . \; . \; . \; . \; . \; . \quad (11)$$

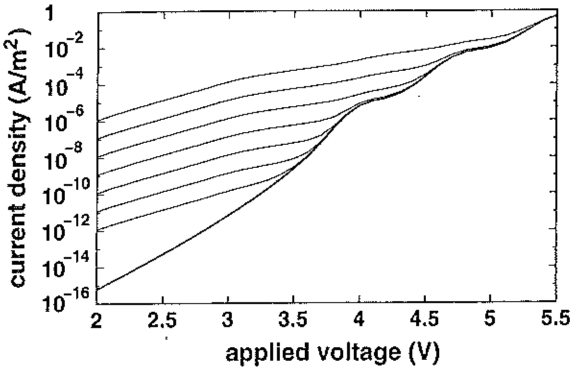
is again between 0·5 and 1, and is independent of any constant factors in equations (4) and (10).

## Numerical results and discussion

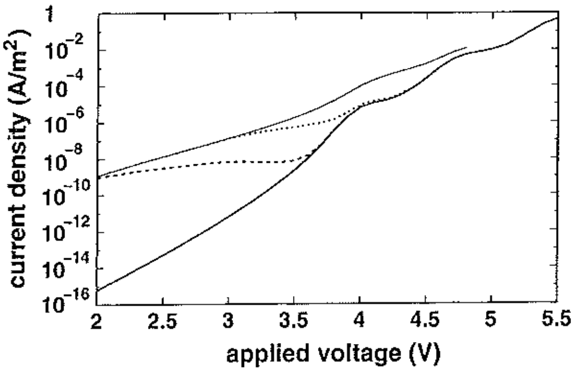
A numerical simulation has been performed of the model for the trap assisted tunnelling just described. First, the non-linear Poisson equation was solved in order to obtain the electron density and the band profiles; then the tunnelling current density and the trap assisted current density were computed for an arbitrary value of $k\eta$. Metal–oxide–semiconductor capacitors realised on (100) oriented $n$ silicon substrates with $5\times10^{15}$ cm$^{-3}$ phosphor doping were considered. The oxide thickness is 6 nm and the top gate is made of polysilicon with $10^{20}$ cm$^{-3}$ donor doping.

Figure 2 shows $I$–$V$ characteristics for a fresh oxide (with no traps), and for increasing values of $k\eta$ in steps of one decade. A trap density per unit volume per unit energy was considered which is uniform in the oxide volume and in an energy range of 1–3 eV below the oxide conduction band. We also consider $\Gamma=10$ meV. Results exhibit good qualitative agreement with the experimental measurement (*see*, for example, Figs. 1–5 in Ref. 4, Fig. 1 in Ref. 5, Fig. 4 in Ref. 7, and Fig. 1 in Ref. 11). The lack of knowledge of the nature and the energy distribution of the traps does not allow quantitative reproduction of the experimental $I$–$V$ curves.

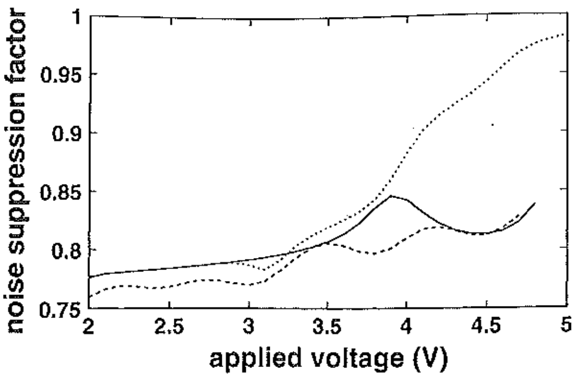
However, the details of the trap density have a strong influence on the dc properties and of noise in trap assisted tunnelling. For example, Fig. 3 is a plot of the $I$–$V$ characteristics obtained with three different energy distributions



**2 Computed current density as function of applied voltage for fresh oxide with thickness of 6 nm (thick line) and for increasing values of $k\eta$ in steps of one decade (thin lines): a uniform trap density was considered with a constant value between 1 and 3 eV below the oxide conduction band, and a value of $\Gamma$ of 10 meV**



**3 Computed current density as function of applied voltage for fresh oxide with thickness of 6 nm (thick solid line) and for a uniform trap density with a constant value between $E^*$ and 3 eV below the oxide conduction band, where $E^*=0$ (thin solid line), 1 (dotted line), and 2 eV (dashed line)**



**4 Computed noise suppression factor as a function of the applied voltage for a uniform trap density with a constant value between $E^*$ and 3 eV below the oxide conduction band, where $E^*=0$ (thin solid line), 1 (dotted line), and 2 eV (dashed line)**

of $\eta$. The thick solid curve is the current of the fresh oxide, while the other curves are obtained assuming a uniform $\eta$ between $E^*$ and 3 eV below the oxide conduction band, where $E^*=0$ eV for the thin solid line, $E^*=1$ eV for the dotted line, $E^*=2$ eV for the dashed line ($E^*$ is the energy gap between the oxide conduction band and the highest energy traps considered) ($E^*$ is the energy gap between the oxide conduction band and the highest energy traps considered). It can clearly be seen that higher energy traps participate in transport at larger gate voltage. At low voltages only the low energy traps contribute.

For the same cases, in Fig. 4 the shot noise suppression factor $\gamma$ is plotted. As expected, it is in the range 0·5–1, pretty close to 0·8, and in good agreement with the few experimental results available.[11] The large differences among the noise suppression factor in the three cases considered need to be studied systematically, by considering the contribution of each energy bin. Maximum suppression of shot noise ($\gamma=0{\cdot}5$) and maximum current through the trap is obtained when $g_1$ and $r_2$ are equal while other transistion rates are negligible. Away from this condition, $\gamma$ quickly approaches 1 and the current decreases. For this reason, a slight variation of trap density in energy can have a significant effect on $\gamma$.

In conclusion, the simple model allows the dc and noise properties of trap assisted tunneling current to be computed and good agreement with experiment to be obtained, considering the poor information available on the detailed distribution of traps. Proper refinement of the model requires accurate information on capture cross-section of traps, and on trap density.

## Acknowledgements

This work has been supported by the Ministry for the University and Scientific and Technological Research of Italy, and by the Italian National Research Council (CNR) through the project MADESS II.

## References

1. L. MOLENKAMP, D. J. PAUL, and R. COMPANO (eds): 'Nanoelectronics technology roadmap'; 1998, European Union.
2. C. T. SAH: *Solid-State Electron.*, 1990, **33**, 147–167.
3. J. MASERJIAN and N. ZAMANI: *J. Appl. Phys.*, 1982, **53**, 559.
4. P. OLIVO, T. N. NGUYEN, and B. RICCÒ: *IEEE Trans. Electron. Devices*, 1988, **35**, 2259–2267.
5. R. MOAZZAMI and C. HU: Tech. Digest 1992 'Electronic device' Meeting, San Francisco, CA, USA, December 1992, IEEE, 139–142.
6. G. J. HEMINK, K. SHIMIZU, S. ARITOME, and R. SHIHOTA: Proc. 1996 International Reliability Physics Symp., Dallas, TX, USA, April 1996, IEEE, 117–120.
7. S. TAKAGI, N. YASUDA, and A. TORIUMI: *IEEE Trans. Electron Devices*, 1999, **46**, 335–341; 348–354.
8. S. SZE: 'Physics of semiconductor devices', 2nd edn; 1981, New York, Wiley. 1981.
9. G. IANNACCONE, M. MACUCCI, and B. PELLEGRINI: *Phys. Rev. B*, 1997, **55**, 4539–4550.
10. A. VAN DER ZIEL: 'Noise in solid state devices and circuits'; 1986, New York, Wiley.
11. G. IANNACCONE, F. CRUPI, B. NERI, and S. LOMBARDO: *Appl. Phys. Lett.*, 2000, **77**, 2876–2878.