A Compact Model Based on Bardeen's Transfer Hamiltonian Formalism for Silicon Single Electron Transistors

FABIAN J. KLÜPFEL [ID]

Fraunhofer Institute for Integrated Systems and Device Technology IISB, 91058 Erlangen, Germany

e-mail: fabian.kluepfel@iisb.fraunhofer.de

ABSTRACT Presented is a physics-based compact model for a silicon-nanopillar single-electron transistor (SET). Tunneling currents are calculated using a master equation approach with rates obtained via the transfer Hamiltonian formalism. The quantum confinement of electrons on the quantum dot is taken into consideration by a suitable approximation as required for a nanometer-sized device. Device geometry and material properties enter the model directly as model parameters. Thus, this model enables the investigation of circuits and application scenarios for specific SET technologies in dependence on geometry and material variations. The model was implemented in HSPICE and used to simulate an inverter and a ring oscillator to evaluate the performance of the model. Specific device characteristics for a SET with a semiconducting quantum dot like the gate voltage threshold for the onset of current oscillations are reproduced. Therefore, simulations with the presented model will allow the testing of the SET circuits with more realistic assumptions concerning the device behavior compared to the much more abstract SET compact models available up to now.

INDEX TERMS Circuit simulation, compact model, single electron transistors, SPICE.

I. INTRODUCTION

The miniaturization of conventional silicon CMOS electronics is approaching physical limits that make it increasingly difficult to follow the so-called More-Moore pathway. This has increased the need to explore different device types that offer additional advantages compared to standard CMOS technology in terms of performance, power consumption, or area requirements (More-Than-Moore) [1], [2]. Among others, the single-electron transistor (SET) is considered as a candidate for future electronics due to expected benefits for low-power applications and for its low-area requirements.

SETs make use of a quantum dot (QD) coupled by tunneling contacts to source and drain electrodes. More than one QD can be utilized as well, but this case will not be addressed in this work. The most prominent SET feature is the Coulomb blockade which is caused by the electrostatic repulsion of electrons on the QD towards charge carriers on the electrodes. Application scenarios investigated by means

of circuit simulation include all-SET logic circuits, hybrid SET-FET logic circuits, as well as single or few-electron memories [3].

Due to the fundamentally different mode of operation of a SET compared to a classical transistor like a MOSFET when making use of the Coulomb blockade effect, the applicability of SETs in electronic circuits must be thoroughly tested by simulation. This requires the availability of suitable compact models that allow fast evaluation of the current-voltage (IV) behavior and usage in a circuit simulator for the coupling with other circuit elements. For this purpose, various SET compact models have been proposed [4]–[8]. Basis for most SET compact models is the orthodox theory (reviewed e.g. in [9]). A common assumption of all these models is the use of a metallic quantum dot without band gap that results in a perfectly periodic IV characteristic with respect to the gate voltage. Quantum confinement on the QD is typically neglected.

Experimental works on SETs with metallic QDs demonstrate very regular oscillations for gate voltages around $V_{GS} = 0$ V [10]–[13]. This case is well covered by the


The associate editor coordinating the review of this manuscript and approving it for publication was Sourabh Khandelwal.

orthodox theory and existing compact models. However, SETs employing silicon QDs generally exhibit a threshold behavior. Typically, gate voltages of few volts are necessary to switch the device on. Current oscillations are observed only for gate voltages above this threshold [14]–[17]. Some works use a second gate with a positive voltage in order to demonstrate current oscillations around zero gate voltage. Still, this does not change the device behavior fundamentally but only shifts the threshold to negative gate voltages [18], [19]. For measurements on SETs with silicon source and drain contacts, the threshold behavior could originate from depletion of these contacts. However, the threshold for current oscillations is also observed for metallic contacts [15] and for a small gate contact that covers only the QD area [19]. Thus, the QD bandgap causes a threshold behavior and a theoretical model for description of a SET with semiconducting QD needs to implement the off-region at low gate voltages, regardless of the source/drain material.

Conduction electrons on a QD have discrete energy levels due to confinement. The energy spacing of these levels on a QD with size in the nanometer range is in a similar order than the Coulomb blockade energy. The influence of energy states above the ground state was observed for silicon SETs measured at low temperatures as fine structure in the IV characteristics outside the coulomb diamonds [20], [21]. A compact model completely ignoring the influence of discrete energy states pretends a more regular behavior of the device as to be expected experimentally. Furthermore, quantum confinement and hence the groundstate energy is strongly dependent on the QD size and must be respected when studying the influence of the QD size on device and circuit properties.

The calculation of tunneling currents in orthodox theory relies on an empirical tunneling resistance for each tunneling junction, while electrostatics is described by a set of capacitance parameters. These parameters must be obtained from experiment or from numerical simulations. In order to perform device optimization or variability studies on SET circuits, such empirical model parameters are not well suited, as their dependence on device geometry or on material properties is unknown (see e.g. [22]). Systematic variation of SET geometry is experimentally rarely feasible due to limitations of processing technologies. Numerical simulations of SET characteristics based on three-dimensional device geometries and using the transfer Hamiltonian method proposed by Bardeen [23] for current calculation have been reported beforehand [24]–[27]. Such simulations are based on the actual device geometry and material parameters and can be used for calibration of compact model parameters. However, if multiple device variants are to be studied this is a time-consuming process. Thus, a SET compact model that features material properties and device geometry directly as model parameters could greatly improve the time needed for variability studies and test simulations for SET circuits.

This work uses a simplified version of the transfer Hamiltonian formalism in order to derive an analytic SET model that is made available for circuit simulation via
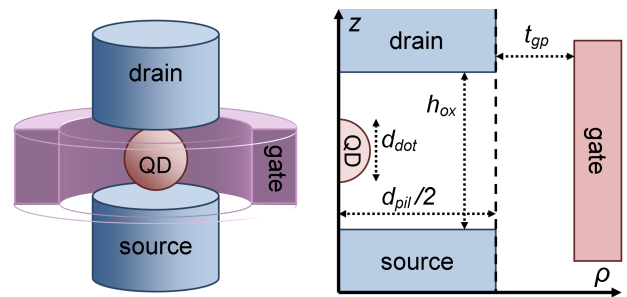
FIGURE 1. Schematic drawing of the SET structure. Left: 3D image, the front half of the all-around gate is hidden for better view of the QD. Right: Cross section, the *z*-axis corresponds to the symmetry axis.

implementation in HSPICE [28]. The SET geometry is based on a silicon nanopillar as proposed by the European Union's H2020 project IONS4SET [29]. Geometrical dimensions and doping levels are available as model parameters. Quantum confinement and the influence of energy states above the groundstate are considered via suitable approximations. The model and its ability to reproduce the correct trends for geometric variations are demonstrated via numerical simulations with the framework described in [27].

II. PROBLEM FORMULATION
For this work, we choose a SET geometry based on a silicon nanopillar with diameter d_{pil}. Embedded in the pillar is an silicon dioxide layer with thickness h_{ox} that contains the spherical QD with diameter d_{dot}. The gate contact completely surrounds the pillar as displayed in Fig. 1. The distance between pillar and all-around gate is given by t_{gp}. Because of the symmetry of the structure we will mostly use the cylindrical coordinates z (direction of SET current flow) and ρ (radial distance from the pillar axis) as shown in Fig. 1 (right side). For development of this model the QD is assumed to be centered between source and drain. A more generalized formulation which allows the asymmetric positioning of the QD will be the topic of future work. The distance between the QD and either electrode is given by

$$t_{ed} = \frac{h_{ox} - d_{dot}}{2}. \tag{1}$$

It should be noted that confinement effects in the pillar-shaped source and drain electrodes are neglected in this work. If not mentioned otherwise, we use the nominal parameter values defined in Tab. 1. The table also specifies the parameter range used for variability studies. QD diameters d_{dot} of around 3 nm are chosen in order to permit room-temperature operation of the SET which requires a QD size smaller than 10 nm [30]. Furthermore, the fabrication of self-aligned silicon QDs in a silicon dioxide layer has been demonstrated experimentally with ion-beam mixing and subsequent annealing for QD sizes of about 3 nm [31]. For obtaining measurably large currents, the distance t_{ed} should not exceed 2 nm while dot formation by ion-beam mixing requires a minimum distance of at least 0.5 nm. This provides a range of interest for the oxide thickness h_{ox}. The pillar
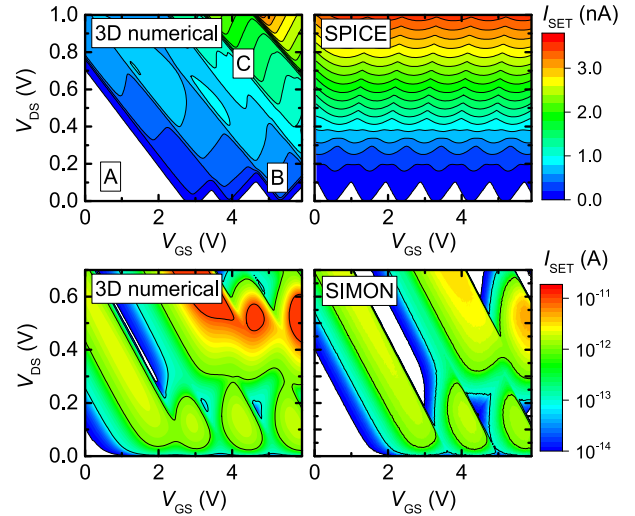
TABLE 1. SET geometry and material parameters.

Symbol	Quantity	Nominal	Variability range
d_{pil}	pillar diameter	10 nm	8 nm - 16 nm
h_{ox}	pillar oxide thickness	5.6 nm	4.8 nm - 6.4 nm
d_{dot}	QD diameter	3.2 nm	2.0 nm - 4.4 nm
t_{gp}	gate-pillar oxide thickness	3 nm	1 nm - 10 nm
h_{g}	gate contact height	10 nm	9 nm - 22 nm
Q_{bg}	QD background charge	0	0 - 4 e
g_{dot}	QD eigenstate degeneracy	12	–
m_{dot}	QD effective electron mass	0.328	–
N_{G}	gate net doping density	10^{20} cm^{-3}	–
N_{S}	source net doping density	10^{20} cm^{-3}	–
N_{D}	drain net doping density	10^{20} cm^{-3}	–
g_{elec}	electrode conduction band degeneracy	6	–
m_{elec}	electrode effective electron mass	0.328	–
m_{ox}	oxide effective electron mass	0.5	–
E_{B}	oxide energy barrier height	3.15 eV	–
ϵ_{barr}	oxide dielectric constant	3.9	–
T	temperature	300 K	100 K - 300 K

FIGURE 2. Simulated current-voltage plots for SETs with silicon QD. Upper left: Numerical simulation of pillar SET with metallic electrodes at 50 K, according to [27]. Marked are operating regions: Off-region (A), Coulomb-Blockade regime (B), Tunneling through higher energy eigenstates (C). Upper right: SPICE simulation of SET at 50 K using Inokawa's SET model [6], calibrated to numerical simulation. Lower left: Numerical simulation of pillar SET with silicon electrodes at 300 K, according to [27]. Lower right: Simulation with SIMON at 300 K, calibrated to numerical simulation.

diameter d_{pil} should be as small as possible in order to allow good electrostatic control of the electrostatic potential of a QD centered within the pillar. Experimentally, fabrication of silicon nanopillars with diameters down to 15 nm has been demonstrated [32]. The specific values given in Tab. 1 are also influenced by the grid used for numerical simulations in this work.

Numerical simulations of IV characteristics for silicon-nanopillar SETs have been reported before [27], based on the commercial 3D Poisson-Schroedinger solver nextnano++ [33] with a self-written implementation of the transfer Hamiltonian formalism. Simulations with this framework for the nominal SET device are presented in Fig. 2. Different operation regimes can be distinguished and are marked in the figure: Off-region (A), Coulomb-Blockade regime (B), and a regime at higher drain voltages where tunneling occurs through higher energy eigenstates in addition to the ground-state (C). A HSPICE simulation shown in Fig. 2 using Inokawa's SET model [6], with capacitances and tunneling resistances calibrated to numerical simulation, exhibits good agreement in the Coulomb-Blockade regime while the other operation regimes are not well described or not featured at all. Simulations with the Monte-Carlo simulator SIMON [34] shown in Fig. 2 exhibit very similar features compared to numerical simulations when using both semiconducting electrodes and QD and calibrating capacitances, tunneling resistances, and energy levels to the numerical simulations. However, SIMON does not offer the same flexibility for circuit simulations compared to SPICE, especially no MOSFET models are available. Aim of this work is the development of a SPICE model that agrees reasonable well with numerical simulations in all operation regimes and features the model parameters summarized in Tab. 1 instead of more abstract capacitance or resistance values.

Every SET consists of source and drain electrodes with tunneling contacts towards the central quantum dot. One or more gate contacts are located in non-tunneling distance to

the QD in order to influence the QD potential V_{dot}. Due to charge conservation of the QD charge Q_{dot}, the potential V_{dot} in the presence of a single gate is given by (see e.g. [1])

$$V_{\mathrm{dot}} = \frac{C_{\mathrm{G,dot}}}{C_{\mathrm{dot}}} V_{\mathrm{G}} + \frac{C_{\mathrm{D,dot}}}{C_{\mathrm{dot}}} V_{\mathrm{D}} + \frac{C_{\mathrm{S,dot}}}{C_{\mathrm{dot}}} V_{\mathrm{S}} + \frac{Q_{\mathrm{dot}}}{C_{\mathrm{dot}}} \quad (2)$$

Here, $C_{\mathrm{dot}} = C_{\mathrm{G,dot}} + C_{\mathrm{D,dot}} + C_{\mathrm{S,dot}}$ is the total QD capacitance and $C_{\mathrm{G,dot}}$, $C_{\mathrm{D,dot}}$, $C_{\mathrm{S,dot}}$ are the capacitances between QD and the respective contacts.

In order to calculate current-voltage characteristics of a SET, the first step is to obtain values for the capacitances in (2) in order to solve the electrostatics of the system (chapter III). Subsequently, the quantum-mechanical treatment of the QD within the isolating matrix is necessary in order to obtain the energy eigenstates and wave functions for charge carriers on the QD (chapter IV). Finally, this information can be used in order to calculate tunneling rates between electrodes and QD (chapter V).

III. ELECTROSTATICS

The QD is the central part of the SET structure. The capacitances of the QD towards the electrodes determine the QD potential as expressed by (2) which in turn controls the current flow in the SET. The QD capacitances are determined by the three-dimensional distribution of the electric field in the structure, however the pillar-SET structure is too complicated to derive a rigorous analytic solution of the field. Thus, we split the problem in two parts. First, we formulate an approximation of the total capacitance C_{dot}. Then, we determine the gate capacitance. Under the assumption of a QD centered between source and drain, we can calculate $C_{\mathrm{S,dot}} = C_{\mathrm{D,dot}} = (C_{\mathrm{dot}} - C_{\mathrm{G,dot}})/2$.
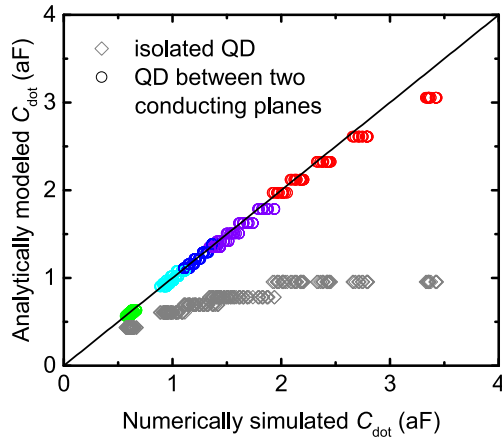
FIGURE 3. Numerical simulations with Synopsys Sentaurus [35] in comparison to the analytical model presented in this work. The model for the isolated QD corresponds to (3) and the model for the QD between two electrodes is given by (5). The colors for the latter model represent different d_{dot} values, from 2 nm (left) to 4.4 nm (right).

A. TOTAL QUANTUM DOT CAPACITANCE

In order to obtain absolute capacitance values, an expression for the total QD capacitance is required. An initial guess might be taken from the capacitance of a QD in an infinite matrix of silicon dioxide. For this case, the solution can be derived simply from Coulomb's law and is given by

$$C_{\mathrm{dot}} = 4\pi \epsilon_r \epsilon_0 \frac{d_{\mathrm{dot}}}{2}. \tag{3}$$

However, comparison to numerical simulations shows that this solution underestimates the real QD capacitance, which is modified by the electrodes in vicinity of the dot. The disagreement compared to numerical simulations can be seen in Fig 3.

The electric potential of a point charge in vicinity of a conducting plane can be easily described by the sum of the potentials of the actual point charge and an image charge of opposite sign. The combined potential satisfies the boundary condition of a constant potential on the surface of the plane. The field of a point charge in between two conducting planes cannot be described in an exact manner by using point-like image charges. However, we use the potential of a point charge and two image charges as an approximation, which will serve the purpose of an analytical approximation reasonably well. The magnitude of the image charges is chosen in order to get a potential of zero at the electrode interfaces close to the dot. The electrostatic potential around the dot is then described by

$$\Phi(z, \rho) = \frac{Q_{\mathrm{dot}}}{4\pi \epsilon_r \epsilon_0} \left(\frac{1}{\sqrt{\rho^2 + z^2}} - \frac{3/4}{\sqrt{\rho^2 + (h_{\mathrm{ox}} - z)^2}} - \frac{3/4}{\sqrt{\rho^2 + (h_{\mathrm{ox}} + z)^2}} \right). \tag{4}$$

The coordinate origin is located at the QD center. For the purpose of capacitance evaluation, the QD potential is defined as the potential in a distance $d_{\mathrm{dot}}/2$ from the dot center
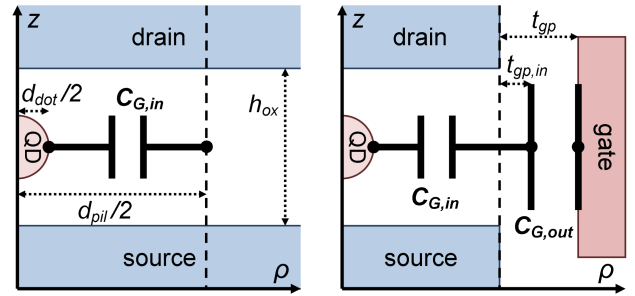
FIGURE 4. Schematic drawings for the derivation of the gate-to-QD capacitance. Left: Capacitance between QD and pillar edge (dashed line). Black circles mark the positions where the electrostatic potential is evaluated. Right: Combination of inner and outer capacitances.

towards the gate electrode, i.e. $V_{\mathrm{dot}} = \Phi(z = 0, \rho = d_{\mathrm{dot}}/2)$. Then, the total QD capacitance is given by

$$C_{\mathrm{dot}} = \frac{\partial Q_{\mathrm{dot}}}{\partial V_{\mathrm{dot}}} = \frac{4\pi \epsilon_r \epsilon_0 \, d_{\mathrm{dot}}}{2 - \dfrac{3}{\sqrt{1 + (2h_{\mathrm{ox}}/d_{\mathrm{dot}})^2}}}$$

$$\approx \frac{4\pi \epsilon_r \epsilon_0 \, d_{\mathrm{dot}}}{2 - \dfrac{3 d_{\mathrm{dot}}}{2 h_{\mathrm{ox}}}}. \tag{5}$$

This result is very similar to (3) but includes a geometry dependent parameter in the denominator that increases the total capacitance. The expression follows the values obtained with numerical simulation much more closely than the capacitance of an isolated dot (see Fig. 3). The approximation in the last step of (5) is based on the fact that the QD must always be smaller than the thickness of the containing oxide layer.

B. GATE-TO-QD CAPACITANCE

The gate-to-QD capacitance $C_{\mathrm{G,dot}}$ describes the electrostatic coupling between the gate contact and the QD. The analytic approximation of the gate-to-QD capacitance is separated into a contribution within the nanopillar, described by $C_{\mathrm{G,in}}$, and a contribution outside the nanopillar, denoted $C_{\mathrm{G,out}}$. For a good description of $C_{\mathrm{G,in}}$, we need to understand how a change of the electrostatic potential at the pillar edge will influence the potential at the QD located between the source and drain contacts. As a starting point, the electrostatic potential between two infinite source/drain contacts at zero voltage in cylindrical coordinates is considered (see Fig. 4 left). The potential must fulfill the Poisson equation $\Delta \Phi = 0$ with the boundary condition $\Phi(z = \pm h_{\mathrm{ox}}/2, \rho) = 0$. The ansatz $\Phi(z, \rho) = R(\rho) \cos(\pi z/h_{\mathrm{ox}})$ is sufficient to satisfy the boundary condition. With the additional condition of a finite value for the potential at the QD center, the radial solution is given by the modified Bessel functions I_n of the first kind. Here, we consider only the first order function I_0, which can be approximated for large radii by an asymptotic expression [36]:

$$\Phi(z, \rho) = I_0 \left(\frac{\pi}{h_{\mathrm{ox}}} \rho \right) \cos \left(\frac{\pi}{h_{\mathrm{ox}}} z \right)$$

$$\approx \sqrt{\frac{h_{\mathrm{ox}}}{2\pi^2 \rho}} \exp \left(\frac{\pi}{h_{\mathrm{ox}}} \rho \right) \cos \left(\frac{\pi}{h_{\mathrm{ox}}} z \right) \tag{6}$$

This expression makes it possible to derive a relative capacitance between the uncharged quantum dot at potential V_{dot} and a gate contact at the pillar edge with potential V_{pil} when taking (2) into account:

$$\frac{C_{G,in}}{C_{dot}} = \frac{V_{dot}}{V_{pil}} = \frac{\Phi(0, d_{dot}/2)}{\Phi(0, d_{pil}/2)}$$

$$= \sqrt{\frac{d_{pil}}{d_{dot}}} \exp\left(\frac{\pi}{2h_{ox}}(d_{dot} - d_{pil})\right). \quad (7)$$

For a gate located at some distance t_{gp} to the pillar, an additional contribution to the capacitance must be considered (see Fig. 4 right). The most simple expression would be a that of a cylindrical capacitor with a height in the order of the oxide layers's height. This is given by

$$C_{G,out} = 2\pi\epsilon_r\epsilon_0 \frac{h_{ox}}{\ln(1 + 2t_{gp}/d_{pil})}. \quad (8)$$

Ideally one would simply connect $C_{G,in}$ and $C_{G,out}$ as stated above in series to obtain the total gate-to-dot capacitance. However, there are two further issues to be considered. First, the electrostatic influence of the source/drain contacts will extend somewhat over the pillar diameter. An empirical way to deal with this fact is to define a boundary between the inner and outer capacitance contributions which is not at the pillar boundary but an additional distance $t_{gp,in}$ further towards the gate (see Fig. 4 right). This means that d_{pil} in $C_{G,in}$ must be replaced by $d_{pil} + 2t_{gp,in}$ and t_{gp} in $C_{G,out}$ must be replaced by $t_{gp} - t_{gp,in}$. The simplest way to implement the idea of a shifted boundary between the two capacitance components is to use a fixed value $t_{gp,in} = l_{gp}$. However, for thin gate oxides it must be ensured that $t_{gp,in} \leq t_{gp}$. This is respected by the following empirical formula:

$$t_{gp,in} = \left(\frac{1}{t_{gp}^{3/2}} + \frac{1}{l_{gp}^{3/2}}\right)^{-2/3}. \quad (9)$$

Here, l_{gp} is a fitting parameter. Good agreement to numerical simulations is obtained for $l_{gp} = 2.5$ nm which will be used in the following.

Second, comparison with numerical simulation shows that the effective height of the outer capacitor is not simply given by h_{ox} as assumed beforehand, but is influenced by pillar diameter and QD size as well. This coupling is well captured by modifying the effective height of the capacitor by an empirical factor $C_{G,in}/C_{dot}$. The considerations in this section lead to the following capacitance contributions which connected in series give the total gate-to-QD capacitance:

$$C_{G,in} = C_{dot}\sqrt{\frac{d_{pil} + 2\, t_{gp,in}}{d_{dot}}}$$

$$\times \exp\left(\frac{\pi}{2h_{ox}}(d_{dot} - d_{pil} - 2t_{gp,in})\right)$$

$$C_{G,out} = 2\pi\epsilon_r\epsilon_0 \frac{h_{ox}C_{G,in}/C_{dot}}{\ln(1 + 2(t_{gp} - t_{gp,in})/d_{pil})}$$

$$C_{G,dot} = (1/C_{G,in} + 1/C_{G,out})^{-1} \quad (10)$$
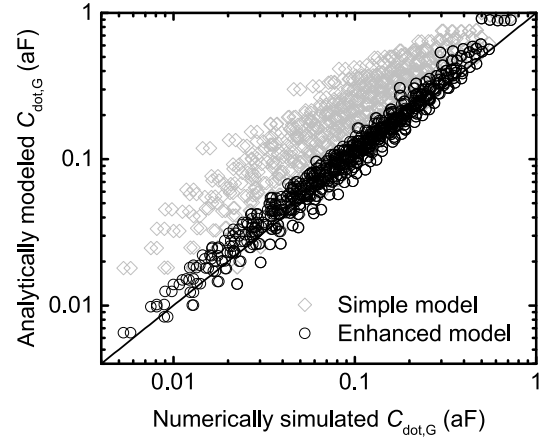
FIGURE 5. Numerical simulations with Synopsys Sentaurus [35] in comparison to the analytical model presented in this work. The simple model corresponds to (7) and (8) while the enhanced model is given by (10).

For validation, analytically calculated capacitance values were compared to numerical simulations of 800 different geometry variants for the parameter range presented in Tab. 1. Fig. 5 demonstrates that the model given by (10) describes the gate capacitances of all variants reasonably well, given the wide spread of values over more than two orders of magnitude.

C. QUANTUM DOT POTENTIAL
The QD capacitances defined above allow to calculate the QD potential according to (2). When using the source potential as reference for all potentials in the system we obtain

$$V_{dot,S} = \frac{C_{D,dot}}{C_{dot}}V_{DS} + \frac{C_{G,dot}}{C_{dot}}V_{GS} + \frac{Q_{bg} - eN}{C_{dot}}. \quad (11)$$

Here, the QD charge Q_{dot} was split into a background charge Q_{bg} and the charge of N additional electrons located on the QD. These electrons correspond to the free charge carriers contributing to the tunneling current and will be addressed in more detail in chapter V. The energetic position of the QD conduction band minimum in reference to the source electrode is then given by

$$E_{dot} = -eV_{dot,S}. \quad (12)$$

D. PARASITIC CAPACITANCES
In addition to the capacitances between QD and contacts treated in the previous sections there are parasitic capacitances C_{DS}, C_{GS}, C_{GD} between the three contacts. These parasitics have no influence on the static SET characteristic, but will strongly affect the dynamic behavior. The drain-source capacitance is basically a classical plate capacitor with circular metallic contacts. In order to separate the influence of the QD from the parasitic capacitance we consider a capacitor with ring-like electrodes with outer diameter d_{pil} and inner diameter d_{dot}:

$$C_{DS} = \pi\epsilon_r\epsilon_0 \frac{d_{pil}^2 - d_{dot}^2}{4h_{ox}}. \quad (13)$$
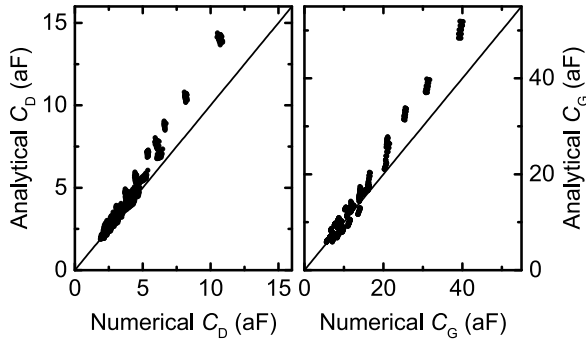
FIGURE 6. Analytical model for total contact capacitances in comparison to numerical simulation with Synopsys Sentaurus [35]. Left: Total drain capacitance. Right: Total gate capacitance.

The gate-source as well as the gate-drain parasitic capacitances are approximated by a cylindrical capacitor similar to (8):

$$C_{GS} = C_{GD} = \pi \epsilon_{barr} \epsilon_0 \frac{h_g/2 + t_{gp}}{\ln(1 + 2t_{gp}/d_{pil})}. \tag{14}$$

The numerator would be $h_g - h_{ox}$ for ideal cylindrical capacitances, but was adjusted empirically to better fit numerical simulations. The parasitic capacitances are difficult to evaluate separately from the QD capacitances, but are contained in the total contact capacitances. The total drain capacitance obtained when using grounded source and floating gate contact is given by

$$C_D = C_{DS} + \frac{C_{D,dot}}{2} + \frac{C_{GD}}{2} \tag{15}$$

under the conditions $C_{D,dot} = C_{S,dot}$ and $C_{GD} = C_{GS}$ given by the symmetry of the system. The total gate capacitance obtained when both source and drain are grounded is given by

$$C_G = 2\,C_{GD} + \left(\frac{1}{C_{G,dot}} + \frac{2}{C_{D,dot}}\right)^{-1}. \tag{16}$$

Numerical simulations on 800 geometry variants of the pillar-SET in the parameter ranges described in Tab. 1 were performed for evaluation of the analytical formulas. The comparison is displayed in Fig. 6 and demonstrates a good agreement. Only for large parasitic capacitances the analytical description overestimates the actual values up to about 25%.

IV. QUANTUM DOT EIGENSTATES
A. ENERGY LEVELS
The quantum dot embedded in an oxide matrix corresponds to the classical quantum mechanical problem of a 3D quantum well with finite walls, as sketched in Fig 7. The barrier height E_B is given by the difference between the conduction band minima of the QD and the barrier material. However, even in the 1D case of a finite quantum well no analytic solution exists for the energies E_n of the bound states. For a 3D quantum well with infinitely high walls, analytic solutions exist only for the case of a box-shaped well and not for the
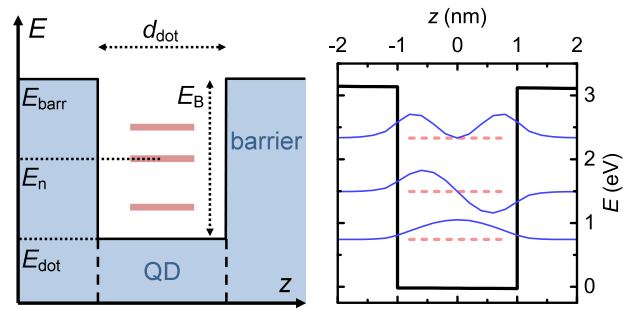
FIGURE 7. Energy landscape around the QD in direction of current flow (z-direction). Left: Schematic drawing with the conduction band edge marked in black and the bandgaps indicated blue, QD energy levels are sketched in red. Right: Numerical simulation with nextnano++ for $d_{dot} = 2$ nm, the first three energy eigenstates are shown in red and the corresponding wave functions in blue (arbitrary units, shifted according to the respective energy levels).
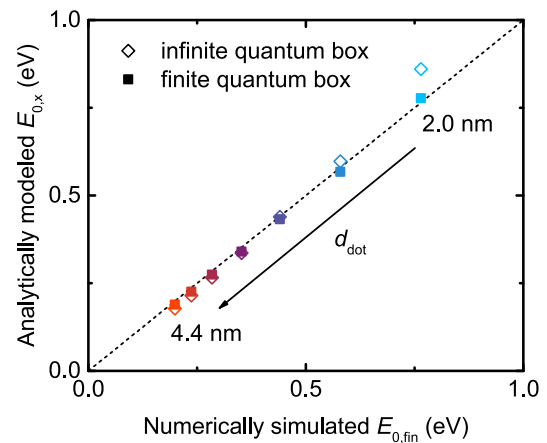
FIGURE 8. Numerical simulation of the QD groundstate with nextnano++ in comparison to the analytical model presented in this work. The model for the infinite quantum box corresponds to (17) and the model for finite quantum box is given by (18).

spherical case. Thus, in order to obtain analytical expressions for energy eigenstates and wavefunctions we approximate the spherical quantum dot by a cube with side length d_{dot} and assume an infinitely high potential barrier around it. The allowed energies are then given by [37]

$$E_{n,inf} = E_{0,1d}\,(n_x^2 + n_y^2 + n_z^2)$$

$$\text{with } E_{0,1d} = \frac{\pi^2 \hbar^2}{2m_{dot}d_{dot}^2}. \tag{17}$$

m_{dot} is the effective mass of an electron on the QD. $E_{0,1d}$ denotes the ground-state energy for the one-dimensional case. The quantum numbers n_x, n_x, and n_x are defined for integer values larger than zero. The ground-state energy for the cubic QD, $E_{0,inf}$, is obtained for $n_x = n_y = n_z = 1$. Comparison of the ground-state energy obtained from numerical simulation for different QD diameters as presented in Fig. 8 shows that the analytical solution for the cubic QD with infinite energy barrier underestimates the actual energies slightly for larger dots, and overestimates the energies significantly for smaller dots.

For a cubic QD with a finite energy barrier, no analytical expressions for energy levels and wave functions exist but approximations are available in literature. For the eigenstate energies we use the so-called parabolic approximation reported in [38] that leads to

$$E_{n,\mathrm{fin}} = E_{n,\mathrm{inf}} \frac{E_{0,\mathrm{1d}}}{4\,E_{\mathrm{B}}} \left(\sqrt{\frac{4\,E_{\mathrm{B}}}{E_{0,\mathrm{1d}}} + 1} - 1 \right)^2 . \tag{18}$$

Strictly, this approximation is only valid for the ground state. However, for simplicity and because the Coulomb blockade regime depends only on the ground state we use this expression for all energy states. $E_{n,\mathrm{fin}}$ denotes the energetic positions of the states above the QD conduction band minimum at energy E_{dot}. Thus, the QD energy levels in reference to the source electrode are given by

$$E_n = E_{n,\mathrm{fin}} - eV_{\mathrm{dot,S}}. \tag{19}$$

In the following, E_0 will denote the ground-state energy with reference to the source electrode according to (19), while $E_{0,\mathrm{fin}}$ denotes the ground-state energy with reference to the QD energy according to (18).

Due to the low tunneling currents through an oxide barrier it can be expected that electrons on the QD thermalize long before a tunneling event occurs. Thus, for a given number of N electrons on the QD, the occupation statistics will be according to the Fermi-Dirac distribution. In silicon, each energy state is 12-fold degenerate (including spin degeneracy) when assuming that the degeneracy is not lifted due to the electric field or other influences. In Fig. 7 on the right it can be seen that the energy spacing between the first eigenstates is significantly larger than $k_{\mathrm{B}}T$ at room temperature (≈ 26 meV). Thus, for a low number of electrons on the QD only the groundstate will be occupied and all higher energy states have occupation probabilities very close to zero. This simplifies the model substantially, because it is not necessary to calculate the QD Fermi level for the given number of electrons N on the dot. For this model, N is considered to be lower than the degeneracy g_{dot} of the ground state. Then, the Fermi-Dirac distribution is approximated by

$$f_{\mathrm{dot}} = \frac{N}{g_{\mathrm{dot}}} \text{ for the ground state}$$
$$f_{\mathrm{dot}} = 0 \text{ for higher energy states.} \tag{20}$$

Hence, it must be noted that this model can only describe the first twelve oscillation periods above the threshold voltage.

B. WAVE FUNCTIONS

For the box-like quantum well with infinite walls, the wave functions consist of separate contributions for each special dimension, i.e. $\Psi_{\mathrm{dot}} = \Psi_x \Psi_y \Psi_z$. Within an energy barrier with finite height the wave function is expected to decrease exponentially. Perpendicular to the tunneling current the exponential tails can be neglected and we use the wave functions for an infinite potential well as approximation. These solutions are given by

$$\Psi_x = \sqrt{\frac{2}{d_{\mathrm{dot}}}} \sin\left(\frac{\pi n_x x}{d_{\mathrm{dot}}}\right) \text{ for } x \in [0; d_{\mathrm{dot}}]$$
$$\Psi_y = \sqrt{\frac{2}{d_{\mathrm{dot}}}} \sin\left(\frac{\pi n_y y}{d_{\mathrm{dot}}}\right) \text{ for } y \in [0; d_{\mathrm{dot}}]$$
$$\Psi_x = \Psi_y = 0 \text{ otherwise.} \tag{21}$$

The origin of the coordinate system has been placed at a corner of the QD for simplicity. For the calculation of the tunneling currents, only integrals over the perpendicular wave functions will be required. For the x-direction the following results are obtained (similar for y-direction):

$$\int_{-\infty}^{+\infty} \Psi_x \mathrm{d}x = \frac{2\sqrt{2d_{\mathrm{dot}}}}{\pi n_x} \text{ for odd } n_x$$
$$\int_{-\infty}^{+\infty} \Psi_x \mathrm{d}x = 0 \text{ for even } n_x \tag{22}$$

In direction of current flow the extension of the wave function into the barrier is decisive, as it enters Bardeen's formula for calculation of the transfer matrix element [23]. Thus, the solution for the infinite well is not sufficient in z-direction, because it drops to zero at the QD interface. We adapt the solution derived by Sée et al. [25] based on the WKB approximation of the wave function within a quantum well barrier of finite height:

$$\Psi_z = \sqrt{\frac{2}{d_{\mathrm{dot}}}} \sqrt{\frac{k_{\mathrm{dot}}^2}{k_{\mathrm{dot}}^2 + \frac{m_{\mathrm{dot}}^2}{m_{\mathrm{ox}}^2} k_{\mathrm{ox}}^2}} \exp(-k_{\mathrm{ox}}z) \text{ for } z > 0$$
$$k_{\mathrm{dot}} = \sqrt{\frac{2m_{\mathrm{dot}}}{\hbar^2}(E_n - E_{\mathrm{dot}})}$$
$$k_{\mathrm{ox}} = \sqrt{\frac{2m_{\mathrm{ox}}}{\hbar^2}(E_{\mathrm{barr}} - E_n)} \tag{23}$$

Here, we assume that the electric field along the z-direction is low enough to treat the barrier as rectangular, in contrast to the cited work of SÃľe et al. that presents the more general form of a trapezoidal barrier. The QD interface is located at $z = 0$ and positive values of z extend into the barrier. For calculation of the tunneling currents in the Coulomb blockade regime, only the lower energy levels are relevant. This justifies the approximation that the tunneling barrier $E_{\mathrm{barr}} - E_{\mathrm{n}}$ is large compared to the QD eigenstate energies $E_{\mathrm{n}} - E_{\mathrm{dot}}$. The QD wave function within the oxide barrier in direction of current flow is then simplified to

$$\Psi_z \approx \sqrt{\frac{2\,m_{\mathrm{ox}}(E_n - E_{\mathrm{dot}})}{d_{\mathrm{dot}}m_{\mathrm{dot}}(E_{\mathrm{barr}} - E_n)}} \exp(-k_{\mathrm{ox}}z). \tag{24}$$

V. TUNNELING CURRENT CALCULATION

The tunneling-current calculation will follow the original idea of Bardeen [23] to determine wave functions for both sides

of the tunneling junction independently and to use them to calculate a matrix element that characterizes the coupling between the two sides of the barrier. This approach has been adapted in the past to the SET problem [24]–[27] and will be described here only as far as necessary in order to follow the procedure to the simplified equations implemented in the compact model.

A. SOURCE/DRAIN ELECTRODES

The electrodes are considered to consist of a metal or a highly doped semiconductor with a partly filled parabolic conduction band. The energetic position of the conduction band minimum is given by E_{elec} and the fermi energy by E_{f}. The effective electron mass in the electrode is given by m_{elec}. Within the framework of the transfer Hamiltonian formalism, the tunneling of electrons between electrode and the QD depends on the extension of the electron wave functions within the barrier. We apply a similar description for the electrode wave functions in the barrier as already described in (24) for the QD in direction of current flow. The major difference is that the prefactor contains the electrode volume V_{elec} which accounts for the nominally infinite extension of the electrodes in x- and y-directions (see the 3D case in [25]). Also, we consider the electrode to fill the volume $z > t_{\text{ed}}$. This changes the exponential factor accordingly:

$$\Psi_{\text{elec}}(E) \approx \sqrt{\frac{2\,m_{\text{ox}}(E - E_{\text{elec}})}{m_{\text{elec}}V_{\text{elec}}(E_{\text{barr}} - E)}} \\ \times \exp(k_{\text{ox}}(z - t_{\text{ed}})). \quad (25)$$

The density of states for the electrode conduction band is obtained from the description of a volume material with g_{elec} conduction band minima in the Brillouin zone and an additional spin degeneracity of 2 [37]:

$$\rho_{\text{elec}}(E) = g_{\text{elec}}\frac{V_{\text{elec}}}{2\pi^2}\left(\frac{2m_{\text{elec}}}{\hbar^2}\right)^{3/2}\sqrt{E}. \quad (26)$$

We can use this definition to remove the explicit dependence of the wave function on the electrode volume. Furthermore, we note that $E_{\text{B}} \approx E_{\text{barr}} - E$ for a sufficiently high energy barrier and low tunneling energy. Then,

$$\Psi_{\text{elec}}(E) \approx \sqrt{\frac{g_{\text{elec}}m_{\text{ox}}}{\pi^2\rho_{\text{elec}}m_{\text{elec}}E_{\text{B}}}} \\ \times \left(\frac{2\,m_{\text{elec}}(E - E_{\text{elec}})}{\hbar^2}\right)^{3/4}\exp(k_{\text{ox}}(z - t_{\text{ed}}))$$

$$\text{with } k_{\text{ox}} \approx \sqrt{\frac{2\,m_{\text{ox}}E_{\text{B}}}{\hbar^2}}. \quad (27)$$

The electrode Fermi energy E_{f} with respect to the conduction-band edge is determined by the doping level of the electrode. For sufficiently high doping, the electric field within the electrode is negligible. Then, the electron density in the electrode is constant and given by [37]

$$n = N_{\text{C}}F_{1/2}\left(\frac{E_{\text{f}}}{k_{\text{B}}T}\right)$$

$$\text{with } N_{\text{C}} = 2\left(\frac{m_{\text{elec}}k_{\text{B}}T}{2\pi\hbar^2}\right)^{3/2} \quad (28)$$

where $F_{1/2}(x)$ is the Fermi-Dirac integral and N_{C} is the conduction-band-edge density of states. For very high doping, i.e. a degenerate semiconductor, the doping density N_{D} can be assumed to be equal to the number of charge carriers n. There are several approximations available for calculation of the Fermi-Dirac integral or its reverse function. Here, an approximation presented by Nilsson [39] is used because of its comparably simple analytical form:

$$\frac{E_{\text{f}}}{k_{\text{B}}T} = \frac{\ln(N_{\text{D}}/N_{\text{C}})}{1 - N_{\text{D}}/N_{\text{C}}} + \left(\frac{3\sqrt{\pi}N_{\text{D}}}{4N_{\text{C}}}\right)^{2/3} \\ + \frac{8\sqrt{\pi}N_{\text{D}}/N_{\text{C}}}{3(4 + \sqrt{\pi}N_{\text{D}}/N_{\text{C}})^2} \quad (29)$$

The energy distribution of the electrode charge carriers is then given straightforwardly by the Fermi-Dirac distribution:

$$f_{\text{elec}}(E) = \frac{1}{1 + \exp\left(\frac{E - Ef}{k_{\text{B}}T}\right)}. \quad (30)$$

B. TUNNELING RATES

We apply Bardeen's equation [23] in the following form to combine wave functions of the QD with the wave function of a single electrode:

$$M(E) = \frac{\hbar^2}{2\,m_{\text{ox}}}\iint_S(\Psi_{\text{dot}}^*\nabla\Psi_{\text{elec}} - \nabla\Psi_{\text{dot}}^*\Psi_{\text{elec}})\overrightarrow{dS}. \quad (31)$$

The integral is over a surface within the barrier with constant potential. In our derivation of the wave function we assumed the barrier to be at constant potential which means that we can place the surface arbitrarily. We choose S to be a plane perpendicular to the z-direction, halfway between electrode and quantum dot. The wave function derivatives in (31) are then simply derivatives in z. The complex conjugation of the dot wave function (denoted Ψ_{dot}^*) can be ignored as all wave function expressions derived earlier are real functions. Thus, we can write the matrix element as

$$M(E) = \frac{\hbar^2}{2\,m_{\text{ox}}}\iint_S(\Psi_{\text{dot}}k_{\text{ox}}\Psi_{\text{elec}} \\ - (-k_{\text{ox}})\Psi_{\text{dot}}\Psi_{\text{elec}})\overrightarrow{dS} \\ = \frac{\hbar^2 k_{\text{ox}}}{m_{\text{ox}}}\iint_S\Psi_{\text{x}}\Psi_{\text{y}}\Psi_{\text{z}}\Psi_{\text{elec}}\overrightarrow{dS} \\ = \frac{\hbar^2 k_{\text{ox}}}{m_{\text{ox}}}(\Psi_{\text{z}}\Psi_{\text{elec}})|_{z=\frac{t_{\text{ed}}}{2}}\int\Psi_{\text{x}}dx\int\Psi_{\text{y}}dy. \quad (32)$$

Using equations (22) (24) (27) we can determine the squared matrix element for tunneling between an electrode and a QD eigenstate with energy E_n to be

$$|M(E_n)|^2 = \frac{512\,g_{\text{elec}}\sqrt{2m_{\text{elec}}}m_{\text{ox}}d_{\text{dot}}}{\pi^6\hbar\rho_{\text{elec}}m_{\text{dot}}E_{\text{B}}}\exp(-2\,k_{\text{ox}}t_{\text{ed}}) \\ \times \frac{(E_n - E_{\text{elec}})^{3/2}(E_n - E_{\text{dot}})}{n_x^2 n_y^2}. \quad (33)$$

Note, that according to (22) the matrix element is only non-zero for both odd n_x and odd n_y. Only the last term of (33) depends on the eigenstate properties and voltages, the first terms depend merely on material properties and the SET geometry. The unusual numerical prefactor of 512 results from multiplying the prefactors of the wave function components when evaluating (32).

According to [25] the tunneling rates between the QD charged with N electrons and an electrode are given by

$$\Gamma_{\text{elec}\to\text{dot}}^N = \sum_n (1 - f_{\text{dot}}(E_n)) f_{\text{elec}}(E_n) \Gamma_n^N$$

$$\Gamma_{\text{dot}\to\text{elec}}^N = \sum_n f_{\text{dot}}(E_n)(1 - f_{\text{elec}}(E_n)) \Gamma_n^N$$

$$\Gamma_n^N = \frac{2\pi}{\hbar} |M(E_n)|^2 \rho_{\text{elec}} g_{\text{dot}}. \tag{34}$$

All quantities in these equations have been defined previously and allow an analytic evaluation of the tunneling rates. Note, that the QD electron charge state N is not explicitly stated in (34) but enters the rate calculation via the definitions of E_{dot}, E_n, and f_{dot}.

C. TUNNELING CURRENTS

In the previous subsection the tunneling rates were calculated in dependence on the QD charge state. In principle, there can be an arbitrary number of electrons on the quantum dot which fluctuates over time during current flow. The probability to find N electrons on the QD at any given time und specific bias conditions is denoted P_N. Under steady-state conditions, the QD occupation probabilities and the tunneling rates between the QD and the source (src) and drain (drn) electrodes must satisfy the following equation system: [25]

$$P_N(\Gamma_{\text{src}\to\text{dot}}^N + \Gamma_{\text{dot}\to\text{src}}^N + \Gamma_{\text{drn}\to\text{dot}}^N + \Gamma_{\text{dot}\to\text{drn}}^N)$$
$$= P_{N-1}(\Gamma_{\text{src}\to\text{dot}}^{N-1} + \Gamma_{\text{drn}\to\text{dot}}^{N-1})$$
$$+ P_{N+1}(\Gamma_{\text{dot}\to\text{src}}^{N+1} + \Gamma_{\text{dot}\to\text{drn}}^{N+1})$$
$$\sum_N P_N = 1 \tag{35}$$

After solving this equation system, the tunneling current from drain to source can be obtained via

$$I_{\text{SET}} = -e \sum_N P_N(\Gamma_{\text{src}\to\text{dot}}^N - \Gamma_{\text{dot}\to\text{src}}^N). \tag{36}$$

An exact solution to (35) and (36) would have to treat all charge states $N \geq 0$. However, if the groundstate energy of the QD is below the conduction band energies of both electrodes, electrons occupying this state are trappend and cannot tunnel to the electrodes. This means that a charge state N_0 with $E_0 \approx -eV_{\text{DS}}$ represents the lowest charge state that can contribute to the tunneling current for given bias conditions. Using (19) and (11) this condition leads to

$$N_0(V_{\text{GS}}, V_{\text{DS}}) = \text{nint}\left(\frac{Q_{\text{bg}}}{e} + \frac{C_{\text{G,dot}}}{e} V_{\text{GS}} \right.$$
$$\left. - \frac{C_{\text{dot}} - C_{\text{D,dot}}}{e} V_{\text{DS}} - \frac{C_{\text{dot}}}{e^2} E_{0,\text{fin}} \right). \tag{37}$$
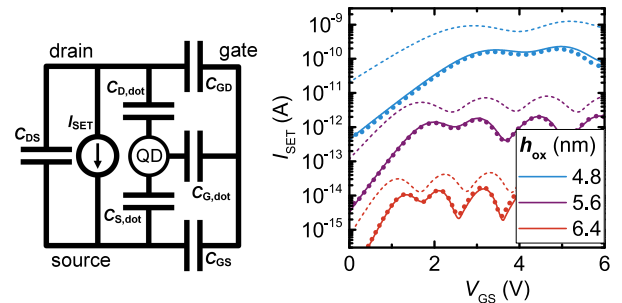
FIGURE 9. Left: Equivalent circuit of the compact model presented in this work. Right: Numerical simulation (dots) in comparison to HSPICE simulations for varied tunneling oxide thickness. Dashed lines denote the basic compact model, solid lines the improved and calibrated model including corrections for the QD capacitance and the dot shape as described in Appendices B and C.

The nint() function rounds to the nearest integer. In case (37) evaluates to a negative number, N must be set to zero. Due to the bandgap of the QD, holes can only appear on the QD for very low gate voltages. This case is not included in the present work.

Equation (37) defines a lower limit for N in dependence on the bias voltages. An approximate upper limit can be obtained by calculating for which N_{max} the dot energy is similar to the source electrode energy. It can be assumed that a higher N would lift the dot energy too high above the source electrode energy in order to contribute to the current. The ansatz $E_0 \approx 0$ leads to

$$N_{\text{max}}(V_{\text{GS}}, V_{\text{DS}}) = N_0 + \text{nint}\left(\frac{C_{\text{dot}} - C_{\text{D,dot}}}{e} V_{\text{DS}} \right). \tag{38}$$

In order to implement the compact model in SPICE syntax and to ensure a low evaluation time for circuit simulations it is mandatory to limit the number of charge states used in the current calculation. For that purpose, a fixed number of four charge states is used in the compact model. The analytical solution to (35) and (36) for the case of four charge states is somewhat lengthy and given in Appendix A. The validity range of this solution can be estimated by solving (38) for V_{DS} and setting $N_{\text{max}} - N_0 \leq 4$. The valid drain voltage range is then given by

$$V_{\text{DS}} \leq \frac{4e}{C_{\text{dot}} - C_{\text{D,dot}}}. \tag{39}$$

In principle, simulations with the model will work for higher V_{DS} as well, but discontinuities will occur as the drain voltage increases above the threshold given by (39).

VI. SPICE MODEL
A. DESCRIPTION OF THE BASIC MODEL
The compact model is based on six capacitances and one voltage-controlled current source (G-element in SPICE syntax). The equivalent circuit is displayed in Fig. 9. The QD capacitances are defined by (5), (9), and (10), the parasitic capacitances by (13) and (14). The current is given by (40) in Appendix A with the tunneling rates defined by (34). Here, we consider the first five QD energy states that provide a
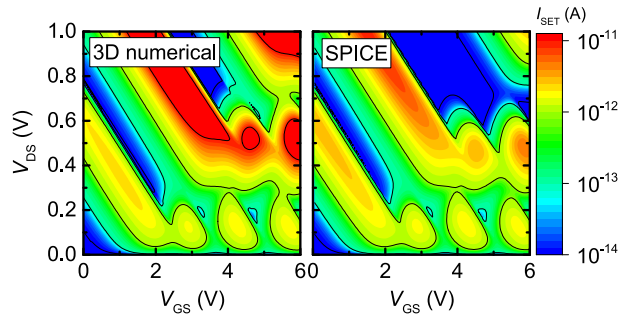
FIGURE 10. Simulated current-voltage plots for SETs with silicon QD. Left: Numerical simulation of pillar SET with silicon electrodes at 300 K, according to [27]. Right: Simulation with HSPICE at 300 K using the calibrated model described in this work.

TABLE 2. Model parameters.

Symbol	Quantity	Nominal	Fitted Value
A	scaling factor for current	1.0	0.55
c_{en}	scaling factor for energy levels	1.0	1.25
c_0	scaling factor total capacitance ($N=0$)	1.0	1.20
c_1	scaling factor total capacitance ($N>0$)	1.0	0.83
l_{gp}	transition length between inner and outer gate capacitance	–	2.5 nm

non-zero contribution to the tunneling rates. In the calculation of tunneling rates enter the calculation of the matrix elements via (33), of the minimum QD charge state via (37), of the QD potential and energy levels via (11), (18), and (19), and the electrode fermi distribution via (29) and (30). This basic form of the model contains only a single empirical parameter, l_{gp}, that was used to divide inner and outer contributions to the gate capacitance. The resulting transfer characteristics for varied tunneling oxide thickness is shown in Fig. 9 on the right side (dashed lines) in comparison to numerical simulations (dots). It can be seen that there is good qualitative agreement between the two simulation approaches. Off-region, current oscillations, and the shift of the characteristics with changing h_{ox} are well reproduced by the SPICE simulations. However, the currents are somewhat too high, and the subthreshold slope as well as the oscillation period are not perfectly matching the numerical simulations. Furthermore, the first oscillation period obtained from numerical simulations is always shorter than the following ones. This has been attributed to different shapes of the electrostatic potential for empty and charged QDs [27], but is not reproduced by the basic compact model.

B. EXTENDED AND CALIBRATED MODEL

In order to allow a proper calibration of the compact model to numerical simulations, few additional corrections and additional empirical parameters were introduced:

The total QD capacitance is considered to be $c_0 C_{dot}$ for the first oscillation period after the gate voltage threshold and $c_1 C_{dot}$ for the following oscillation periods, where c_0 and c_1 are constants and C_{dot} the total QD capacitance as defined beforehand. The resulting modifications to the model equations are summarized in Appendix B and allow adjustment of the total capacitance as well as shortening of the first oscillation period as observed in numerical simulations.

When varying the QD diameter d_{dot} or the pillar oxide thickness h_{ox}, the tunneling current changes exponentially due to the change of the tunneling distance. With the basic model the exponential variation of the current is approximately reproduced (as shown in Fig. 9), but exhibits some deviation that cannot be corrected by simply scaling the

current. This deviation can be traced to the approximation of the spherical QD in numerical simulation by a cubic QD in the compact model. The cubic QD has a full face in tunneling distance t_{ed} of the electrode, while only a single point of the spherical QD has a tunneling distance t_{ed}. The remaining surface curves away from the electrode, which creates an additional sensitivity of the tunneling current on tunneling distance and dot diameter. A correction factor for the current is derived in Appendix C that catches this effect and leads to a better scaling behavior in comparison to numerical simulation when varying d_{dot} and h_{ox}. A proportionality factor A is introduced for the current in order to allow adjustment of the current magnitude.

An additional scaling factor c_{en} was introduced for the QD energy levels. This can be justified as well by the use of a cubic QD in the model, which has a larger volume and hence somewhat lower energy levels compared to the spherical QD.

The solid lines in Fig. 9 show the improvement of the extended model compared to the basic one. The agreement of the compact model to numerical simulations is also illustrated by the IV characteristics in Fig. 10. All operation regimes are well captured with particular focus on the Coulomb blockade regime. For higher drain voltages, the model does at least not pretend periodic behavior where no oscillations are expected physically.

The five empirical model parameters are listed in Tab. 2. The values used within this work are all close to the nominal values, which is a confirmation that the physics of the problem is well reproduced by the compact model. The good agreement of the calibrated model compared to numerical simulations is demonstrated in Fig. 11 with variations in d_{dot}, d_{pil}, Q_{bg}, and T. For all simulations, the same set of model parameters was used with the exception of the varied parameter.

For this work, the empirical model parameters were extracted by hand, focusing foremost on a good fit for the nominal device as defined in Tab. 1 and then on proper behavior under variations as shown in Fig. 11. A better fit between SPICE model and 3D simulations could certainly be obtained using optimization algorithms like the Rosenbrock method or genetic algorithms. However, when using such methods the calibration strategy must certainly depend on the application, especially concerning the question whether calibration should be performed with equal priority for a large number of device variants or whether a nominal device is fitted with higher priority. Hence, the development of suitable
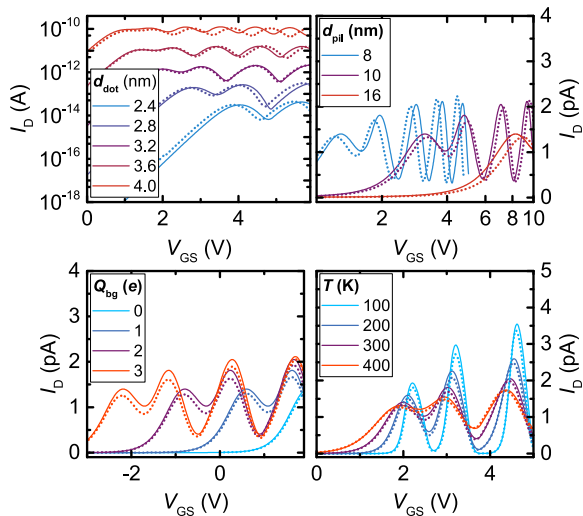
FIGURE 11. Simulated current-voltage plots for SETs with silicon QD, showing numerical simulations according to [27] (dots) in comparison to HSPICE simulations using the calibrated model described in this work (solid). Shown are variations of QD diameter d_{dot}, pillar diameter d_{pil}, positive background charge Q_{bg}, and temperature T.
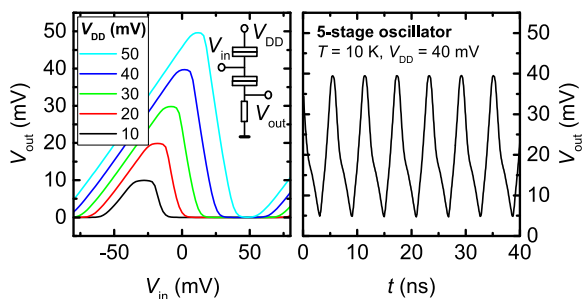
FIGURE 12. Test circuits simulated with HSPICE. Left: Inverter characteristic, the inset shows the circuit consisting of a SET and a resistor. Right: Ringoscillator using five inverters as shown on the left side. Simulation conditions for both examples were: $Q_{bg} = 4e$, $T = 10$ K.

calibration strategies is considered out of the scope of this work which focuses on model development.

C. CIRCUIT SIMULATIONS

The usability of the model in SPICE simulations has been tested with various circuits containing SETs and other elements like FETs or resistors. Although simulation times are certainly somewhat longer compared to simpler SET models, the full analytical description presented in this work shows reasonable convergence and allows to investigate the typical application scenarios presented in literature. It must be noted that not all circuits demonstrated with more ideal models will work with the silicon-based SETs addressed here. This is no shortcoming of the model but reveals the difficulties one must expect when attempting to build such circuits with real devices. An example is the inverter displayed in Fig. 12 on the left side. Around $V_{GS} = 0$ V the SET operates in the off-region when no background charge is considered. Thus, the Coulomb blockade effect cannot be used to realize the low-voltage inverter. Only when a positive background charge Q_{bg} of a few e is assumed to exist, the device

characteristics presented in the figure can be realized. For the simulations shown in Fig. 12, a positive background charge of $Q_{bg} = 4e$ was used. Whether such a device could be fabricated reproducibly remains open and is beyond the scope of this work. However, the inverter simulation demonstrates that logic gates can be characterized by the model. A ring oscillator consisting of five such inverters was used to demonstrate the dynamic capabilities of the model. Here, the parasitic gate capacitances are the dominant cause of delay in the circuit.

VII. CONCLUSION

A compact model for the physical description of a SET with silicon QD is derived from basic principles and tested against numerical simulations. Qualitative agreement between SPICE simulation and numerical simulation with a 3D quantum-mechanics based model is reached even without calibration. With a small number of five fitting parameters we obtain very good quantitative agreement not only for the nominal device but also for variations in geometry, background charge, and temperature. Simulated characteristics of two test circuits are presented to confirm the usability of the model in a SPICE simulation environment.

There are several effects not included in the model that might affect the device characteristics. Examples are confinement effects in the source and drain contacts, scattering by interface traps at the silicon/oxide boundaries, or depletion in low-doped source/drain contacts. However, only experiments on SETs with systematic variations in geometry and doping can give a solid answer on the question what contributions would be really needed for describing real devices. So far, the technological challenges in fabricating working SETs result in low reproducibility and usually allow to realize singular devices only. Thus, no systematic experimental studies on SET characteristics in dependence on geometry or doping exist. The model presented in this work can fill this knowledge gap to a certain extent, as it allows to study effects of geometry and doping variations in a silicon based SET on proposed application scenarios. The model equations might also serve as basis to develop compact models for other SET geometries than the silicon-nanopillar devices presented here.

APPENDIX A
SOLUTION OF THE MASTER EQUATION FOR FOUR CHARGE STATES

For the case of four charge states, denoted N_0, $N_1 = N_0 + 1$, $N_2 = N_0 + 2$, $N_3 = N_0 + 3$ in the following, the solution to (35) and (36) is given by

$$
\begin{aligned}
I_{\text{SET}} = e[&(\Gamma^{N_2}_{\text{dot}\to\text{src}} + \Gamma^{N_2}_{\text{dot}\to\text{drn}})(\Gamma^{N_3}_{\text{dot}\to\text{src}} + \Gamma^{N_3}_{\text{dot}\to\text{drn}}) \\
&\times (\Gamma^{N_0}_{\text{src}\to\text{dot}}\Gamma^{N_1}_{\text{dot}\to\text{drn}} - \Gamma^{N_0}_{\text{drn}\to\text{dot}}\Gamma^{N_1}_{\text{dot}\to\text{src}}) \\
&+ (\Gamma^{N_0}_{\text{src}\to\text{dot}} + \Gamma^{N_0}_{\text{drn}\to\text{dot}})(\Gamma^{N_3}_{\text{dot}\to\text{src}} + \Gamma^{N_3}_{\text{dot}\to\text{drn}}) \\
&\times (\Gamma^{N_1}_{\text{src}\to\text{dot}}\Gamma^{N_2}_{\text{dot}\to\text{drn}} - \Gamma^{N_1}_{\text{drn}\to\text{dot}}\Gamma^{N_2}_{\text{dot}\to\text{src}}) \\
&+ (\Gamma^{N_0}_{\text{src}\to\text{dot}} + \Gamma^{N_0}_{\text{drn}\to\text{dot}})(\Gamma^{N_1}_{\text{src}\to\text{dot}} + \Gamma^{N_1}_{\text{drn}\to\text{dot}}) \\
&\times (\Gamma^{N_2}_{\text{src}\to\text{dot}}\Gamma^{N_3}_{\text{dot}\to\text{drn}} - \Gamma^{N_2}_{\text{drn}\to\text{dot}}\Gamma^{N_3}_{\text{dot}\to\text{src}})]
\end{aligned}
$$

$$/[(\Gamma_{\text{src}\to\text{dot}}^{N_0} + \Gamma_{\text{drn}\to\text{dot}}^{N_0})((\Gamma_{\text{src}\to\text{dot}}^{N_1} + \Gamma_{\text{drn}\to\text{dot}}^{N_1})$$
$$\times (\Gamma_{\text{src}\to\text{dot}}^{N_2} + \Gamma_{\text{drn}\to\text{dot}}^{N_2} + \Gamma_{\text{dot}\to\text{src}}^{N_3} + \Gamma_{\text{dot}\to\text{drn}}^{N_3})$$
$$+ (\Gamma_{\text{dot}\to\text{src}}^{N_2} + \Gamma_{\text{dot}\to\text{drn}}^{N_2})(\Gamma_{\text{dot}\to\text{src}}^{N_3} + \Gamma_{\text{dot}\to\text{drn}}^{N_3}))$$
$$+ (\Gamma_{\text{dot}\to\text{src}}^{N_1} + \Gamma_{\text{dot}\to\text{drn}}^{N_1})(\Gamma_{\text{dot}\to\text{src}}^{N_2} + \Gamma_{\text{dot}\to\text{drn}}^{N_2})$$
$$\times (\Gamma_{\text{dot}\to\text{src}}^{N_3} + \Gamma_{\text{dot}\to\text{drn}}^{N_3})] \tag{40}$$

APPENDIX B
SHORTENING OF FIRST OSCILLATION PERIOD

The electrostatic field within an uncharged QD is approximately constant in absence of an external electric field. The charge density in a charged QD has maximum in the QD center and drops close to zero towards the QD boundary due to the shape of the respective wave functions in the quantum mechanical description. This leads to a curved potential within the QD for the case of the charged QD. The changing shape of the potential in the transition from the uncharged state to the charged state with a single electron leads to a larger shift of the groundstate energy compared to charging with additional electrons. This effect has been illustrated in [27]. For the compact model, the influence on the IV characteristic can be included by using a value $C_{\text{dot},0} = c_0 C_{\text{dot}}$ for the total capacitance of the uncharged QD in comparison for the value $C_{\text{dot},1} = c_1 C_{\text{dot}}$ for the charged QD. This leads to a modified version of (11) for the calculation of the QD potential:

$$V_{\text{dot,S}} = \frac{C_{\text{D,dot}}}{C_{\text{dot},1}} V_{\text{DS}} + \frac{C_{\text{G,dot}}}{C_{\text{dot},1}} V_{\text{GS}} + \frac{Q_{\text{bg}}}{C_{\text{dot},1}}$$
$$- e\left[(N > 0)\frac{1}{C_{\text{dot},0}} + (N > 1)\frac{N-1}{C_{\text{dot},1}}\right] \tag{41}$$

Furthermore, the minimum number of electrons on the QD for given operation conditions, previously defined in (37), is given by

$$N_0(V_{\text{GS}}, V_{\text{DS}}) = \text{nint}\left(\frac{Q_{\text{bg}}}{e} + \frac{C_{\text{G,dot}}}{e} V_{\text{GS}} - \frac{C_{\text{dot},1} - C_{\text{D,dot}}}{e} V_{\text{DS}}\right.$$
$$\left. - \frac{C_{\text{dot},1}}{e^2} E_{0,\text{fin}} + 1 - \frac{C_{\text{dot},1}}{C_{\text{dot},0}}\right). \tag{42}$$

APPENDIX C
CORRECTION FACTOR FOR QD ROUNDING

Using the wave functions for a cubic QD for calculation of currents leads to some difference in the absolute current levels compared to 3D numerical simulations. Especially when varying oxide thickness h_{ox} and quantum dot diameter d_{dot} the trends in the absolute current levels are not correctly reproduced. This can be attributed mostly to the spherical QD shape in numerical simulations. For the cubic QD wave function as described above, we integrate the wave function over a plane S in the oxide barrier in order to obtain the matrix elements. The distance of the integration plane from the QD was $t_{\text{ed}}/2$. Following (33), the explicit dependence of $|M|^2$ on geometrical properties obtained via the QD wave function

are given by

$$|M|_{\text{cube}}^2 \propto \left|\iint_S \Psi_x \Psi_y \Psi_z \vec{dS}\right|^2$$
$$\propto \frac{128\, d_{\text{dot}}}{\pi^4} \exp(-k_{\text{ox}} t_{\text{ed}}). \tag{43}$$

The prefactor consists of the contributions of the lateral wave functions Ψ_x and Ψ_y as well as the normalization constant of Ψ_z. The expression is linearly dependent on d_{dot}, which will be compared to the case of a spherical dot in the following.

For a spherical quantum dot with finite potential boundaries, the wave function in the boundary drops with $\exp(-k_{\text{ox}} r)$ [37]. The normalization constant is not exactly known, but should be proportional to $1/d_{\text{dot}}^{3/2}$ similar to the cubic QD. Then, the matrix element square is proportional to:

$$|M|_{\text{sph}}^2 \propto \left|\iint_S \Psi_{\text{sph}} \vec{dS}\right|^2$$
$$\propto \frac{1}{d_{\text{dot}}^3}\left|\iint_S \exp(-k_{\text{ox}} r)\vec{dS}\right|^2$$
$$= \frac{1}{d_{\text{dot}}^3}\left|2\pi \int_0^\infty \rho \exp(-k_{\text{ox}}\right.$$
$$\left. \times \left(\sqrt{\rho^2 + \left(\frac{t_{\text{ed}}+d_{\text{dot}}}{2}\right)^2} - \frac{d_{\text{dot}}}{2}\right))d\rho\right|^2$$
$$\approx \frac{4\pi^2}{d_{\text{dot}}^3}\left|\int_0^\infty \rho \exp\left(-k_{\text{ox}}\left(\frac{t_{\text{ed}}}{2} + \frac{\rho^2}{t_{\text{ed}}+d_{\text{dot}}}\right)\right)d\rho\right|^2$$
$$\approx \frac{2\pi^2\, h_{\text{ox}}^2}{d_{\text{dot}}^3\, k_{\text{ox}}^2}\exp(-k_{\text{ox}} t_{\text{ed}}). \tag{44}$$

Taylor expansion in ρ was used in order to approximate the square root term within the exponential function. Furthermore, definition (1) for t_{ed} was used in the prefactor. Comparison of $|M|^2$ for cubic and spherical QD reveals a correction factor for the tunneling currents:

$$\frac{|M|_{\text{sph}}^2}{|M|_{\text{cube}}^2} \approx \frac{\pi^6}{64}\frac{h_{\text{ox}}^2}{d_{\text{dot}}^4\, k_{\text{ox}}^2}. \tag{45}$$

The corrected equation for the current is then given by

$$I_{\text{SET}}^{\text{corr}} = A\frac{\pi^6}{64}\frac{h_{\text{ox}}^2}{d_{\text{dot}}^4\, k_{\text{ox}}^2} I_{\text{SET}}. \tag{46}$$

Indeed, this equation reproduces variations of d_{dot} and h_{ox} much better when compared to numerical simulations with a spherical dot. An empirical constant A is introduced to enable calibration of the absolute current. The value of A should be in the order of one.

REFERENCES

[1] W. Haensch, E. J. Nowak, R. H. Dennard, P. M. Solomon, A. Bryant, O. H. Dokumaci, A. Kumar, X. Wang, J. B. Johnson, and M. V. Fischetti, "Silicon CMOS devices beyond scaling," *IBM J. Res. Develop.*, vol. 50, nos. 4–5, pp. 339–361, Jul. 2006. doi: 10.1147/rd.504.0339.

[2] K. Bernstein, R. K. Cavin, W. Porod, A. Seabaugh, and J. Welser, "Device and architecture outlook for beyond CMOS switches," *Proc. IEEE*, vol. 98, no. 12, pp. 2169–2184, Dec. 2010. doi: 10.1109/jproc.2010.2066530.

[3] A. M. Ionescu, M. J. Declercq, S. Mahapatra, K. Banerjee, and J. Gautier, "Few electron devices: Towards hybrid CMOS-SET integrated circuits," in *Proc. 39th Design Autom. Conf.*, Jun. 2002, pp. 88–93. doi: 10.1109/DAC.2002.1012600.

[4] S.-H. Lee, D. H. Kim, K. R. Kim, J. D. Lee, B.-G. Park, Y.-J. Gu, G.-Y. Yang, and J.-T. Kong, "A practical SPICE model based on the physics and characteristics of realistic single-electron transistors," *IEEE Trans. Nanotechnol.*, vol. 1, no. 4, pp. 226–232, Dec. 2002. doi: 10.1109/TNANO.2002.807394.

[5] K. Uchida, J. Koga, R. Ohba, and A. Toriumi, "Programmable single-electron transistor logic for low-power intelligent Si LSI," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2002, pp. 162–453. doi: 10.1109/ISSCC.2002.992194.

[6] H. Inokawa and Y. Takahashi, "A compact analytical model for asymmetric single-electron tunneling transistors," *IEEE Trans. Electron Devices*, vol. 50, no. 2, pp. 455–461, Feb. 2003. doi: 10.1109/TED.2002.808554.

[7] E.-S. A. M. Hasaneen, M. A. A. Wahab, and M. G. Ahmed, "Exact analytical model of single electron transistor for practical IC design," *Microelectron. Rel.*, vol. 51, no. 4, pp. 733–745, Apr. 2011. doi: 10.1016/j.microrel.2010.10.016.

[8] A. Jain, A. Ghosh, N. B. Singh, and S. K. Sarkar, "A new SPICE macro model of single electron transistor for efficient simulation of single-electronics circuits," *Analog Integr. Circuits Signal Process.*, vol. 82, no. 3, pp. 653–662, Jan. 2015. doi: 10.1007/s10470-015-0491-5.

[9] C. Wasshuber, *Computational Single-Electronics* (Computational Microelectronics). Wien, Austria: Springer, 2001.

[10] T. Sato, H. Ahmed, D. Brown, and B. F. G. Johnson, "Single electron transistor using a molecularly linked gold colloidal particle chain," *J. Appl. Phys.*, vol. 82, pp. 696–701, Jul. 1997. doi: 10.1063/1.365600.

[11] K. I. Bolotin, F. Kuemmeth, A. N. Pasupathy, and D. C. Ralph, "Metal-nanoparticle single-electron transistors fabricated using electromigration," *Appl. Phys. Lett.*, vol. 84, pp. 3154–3156, Apr. 2004. doi: 10.1063/1.1695203.

[12] K. Maeda, N. Okabayashi, S. Kano, S. Takeshita, D. Tanaka, M. Sakamoto, T. Teranishi, and Y. Majima, "Logic operations of chemically assembled single-electron transistor," *ACS Nano*, vol. 6, pp. 2798–2803, Mar. 2012. doi: 10.1021/nn3003086.

[13] G. Karbasian, A. O. Orlov, A. S. Mukasyan, and G. L. Snider, "Single-electron transistors featuring silicon nitride tunnel barriers prepared by atomic layer deposition," in *Proc. Joint Int. EUROSOI Workshop Int. Conf. Ultimate Integr. Silicon (EUROSOI-ULIS)*, Jan. 2016, pp. 32–35. doi: 10.1109/ULIS.2016.7440045.

[14] Y. Takahashi, H. Namatsu, K. Kurihara, K. Iwadate, M. Nagase, and K. Murase, "Size dependence of the characteristics of Si single-electron transistors on SIMOX substrates," *IEEE Trans. Electron Devices*, vol. 43, no. 8, pp. 1213–1217, Aug. 1996. doi: 10.1109/16.506771.

[15] B. H. Choi, S. W. Hwang, I. G. Kim, H. C. Shin, Y. Kim, and E. K. Kim, "Fabrication and room-temperature characterization of a silicon self-assembled quantum-dot transistor," *Appl. Phys. Lett.*, vol. 73, pp. 3129–3131, Nov. 1998. doi: 10.1063/1.122695.

[16] A. Dutta, S. P. Lee, Y. Hayafune, S. Hatatani, and S. Oda, "Single-electron tunneling devices based on silicon quantum dots fabricated by plasma process," *Jpn. J. Appl. Phys.*, vol. 39, pp. 264–267, Jan. 2000. doi: 10.1143/JJAP.39.264.

[17] Y. Sun, Rusli, and N. Singh, "Room-temperature operation of silicon single-electron transistor fabricated using optical lithography," *IEEE Trans. Nanotechnol.*, vol. 10, no. 1, pp. 96–98, Jan. 2011. doi: 10.1109/TNANO.2010.2086475.

[18] B.-G. Park, D. H. Kim, K. R. Kim, K.-W. Song, and J. D. Lee, "Single-electron transistors fabricated with sidewall spacer patterning," *Superlattices Microstruct.*, vol. 34, pp. 231–239, Apr. 2003. doi: 10.1016/j.spmi.2004.03.013.

[19] P. J. Koppinen, M. D. Stewart, and N. M. Zimmerman, "Fabrication and electrical characterization of fully CMOS-compatible Si single-electron devices," *IEEE Trans. Electron Devices*, vol. 60, no. 1, pp. 78–83, Jan. 2013. doi: 10.1109/TED.2012.2227322.

[20] H. Ishikuro and T. Hiramoto, "Quantum mechanical effects in the silicon quantum dot in a single-electron transistor," *Appl. Phys. Lett.*, vol. 71, pp. 3691–3693, Dec. 1997. doi: 10.1063/1.120483.

[21] S. Lee, Y. Lee, E. B. Song, and T. Hiramoto, "Observation of single electron transport via multiple quantum states of a silicon quantum dot at room temperature," *Nano Lett.*, vol. 14, pp. 71–77, Dec. 2013. doi: 10.1021/nl403204k.

[22] E. Amat, J. Bausells, and F. Perez-Murano, "Exploring the influence of variability on single-electron transistors into SET-based circuits," *IEEE Trans. Electron. Devices*, vol. 64, no. 12, pp. 5172–5180, Dec. 2017. doi: 10.1109/ted.2017.2765003.

[23] J. Bardeen, "Tunnelling from a many-particle point of view," *Phys. Rev. Lett.*, vol. 6, pp. 57–59, Jan. 1961. doi: 10.1103/PhysRevLett.6.57.

[24] A. Scholze, A. Schenk, and W. Fichtner, "Single-electron device simulation," *IEEE Trans. Electron. Devices*, vol. 47, no. 10, pp. 1811–1818, Oct. 2000. doi: 10.1109/16.870553.

[25] J. Sée, P. Dollfus, S. Galdin, and P. Hesto, "From wave-functions to current-voltage characteristics: Overview of a Coulomb blockade device simulator using fundamental physical parameters," *J. Comput. Electron.*, vol. 5, pp. 35–48, Mar. 2006. doi: 10.1007/s10825-006-7917-3.

[26] S. Illera, J. D. Prades, A. Cirera, and A. Cornet, "Transport in quantum dot stacks using the transfer Hamiltonian method in self-consistent field regime," *EPL (Europhys. Lett.)*, vol. 98, Apr. 2012, Art. no. 17003. doi: 10.1209/0295-5075/98/17003.

[27] F. J. Klüpfel and P. Pichler, "3D simulation of silicon-based single-electron transistors," in *Proc. Int. Conf. Simulation Semicond. Processes Devices (SISPAD)*, Sep. 2017, pp. 77–80. doi: 10.23919/SISPAD.2017.8085268.

[28] *HSPICE Version O-2018.09*, Synopsys, Mountain View, CA, USA.

[29] *European Union Horizon 2020 Grant Agreement No 688072, IONS4SET, 2016–2020*, Eur. Commission, Brussels, Belgium, 2016.

[30] K. K. Likharev, "Single-electron devices and their applications," *Proc. IEEE*, vol. 87, no. 4, pp. 606–632, Apr. 1999. doi: 10.1109/5.752518.

[31] X. Xu, T. Prüfer, D. Wolf, H.-J. Engelmann, L. Bischoff, R. Hübner, K.-H. Heinig, W. Möller, S. Facsko, J. von Borany, and G. Hlawacek, "Site-controlled formation of single Si nanocrystals in a buried SiO_2 matrix using ion beam mixing," *Beilstein J. Nanotechnol.*, vol. 9, pp. 2883–2892, Nov. 2018. doi: 10.3762/bjnano.9.267.

[32] A. Gharbi, P. Pimenta-Barros, O. Saouaf, G. Reynaud, L. Pain, R. Tiron, C. Navarro, C. Nicolet, I. Cayrefourcq, M. Perego, F. Pérez-Murano, E. Amat, and M. Fernández-Regúlez, "Pillars fabrication by DSA lithography: Material and process options," *Proc. SPIE*, vol. 10586, Mar. 2018, Art. no. 105860Q. doi: 10.1117/12.2297414.

[33] *nextnano++ Version 2016_04_29_RedHat_67*, Nextnano GmbH, Munich, Germany, 2006.

[34] C. Wasshuber, "SIMON version 2.0," Cambridge, MA, USA, 1997.

[35] *Sentaurus TCAD Version O-2018.06*, Synopsys, Mountain View, CA, USA, 2018.

[36] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions* (National Bureau of Standards Applied Mathematics), vol. 55. Washington, DC, USA: U.S. Government Printing Office, 1972.

[37] M. Grundmann, *The Physics of Semiconductors*. Berlin, Germany: Springer, 2006.

[38] O. F. de Alcantara Bonfim and D. J. Griffiths, "Exact and approximate energy spectrum for the finite square well and related potentials," *Amer. J. Phys.*, vol. 74, pp. 43–48, Jan. 2006. doi: 10.1119/1.2140771.

[39] N. G. Nilsson, "An accurate approximation of the generalized Einstein relation for degenerate semiconductors," *Phys. Status Solidi A*, vol. 19, pp. K75–K78, Sep. 1973. doi: 10.1002/pssa.2210190159.

FABIAN J. KLÜPFEL received the Diploma degree in physics and the Ph.D. (Dr. rer. nat.) degree from the University of Leipzig, Germany, in 2010 and 2015, respectively.

From 2010 to 2015, he worked on oxide–semiconductor devices at the Semiconductor Physics Group, University of Leipzig. In 2016, he joined Fraunhofer IISB, Erlangen, Germany, where he focuses on TCAD simulation of nanoelectronic devices, including conventional CMOS transistors as well as nanowire and single-electron devices. His further activities include compact model calibration and development for unconventional semiconductor devices.