First Principles Based NEGF Simulations of Si Nanowire FETs

Mincheol Shin, Woo Jin Jeong, Jaehyun Lee and Junbeom Seo
School of Electrical Engineering
Korea Advanced Institute of Science and Technology
Daejeon 34141, Rep. of Korea
mshin@kaist.ac.kr

Abstract—In this work, we have performed first-principles density functional theory (DFT) based, self-consistent device simulations for Si nanowire field effect transistors (NW FETs) of cross-section up to 3.24 nm. Through mode space transformation, non-orthogonal DFT Hamiltonian and overlap matrices are reduced in size from a few thousands to around one hundred. Ultra-efficient quantum-mechanical transport calculations in the non-equilibrium Green's function (NEGF) formalism in a non-orthogonal basis are therefore made possible.

Keywords—density functional theory; non-equilibrium Green's function; nanowire field effect transistors.

I. Introduction

As the feature size of the field effect transistors (FETs) approaches the nanometer scale, it has become important to simulate the devices at the atomistic level [1]. While the empirical tight-binding (TB) method is well suited for the purpose [2], it has some limitations such as the transferability of TB parameters in hetero-structures and the difficulty in dealing with surface passivation in nanowire (NW) structure. The use of first-principles density functional theory (DFT) Hamiltonians should give an ultimate solution [3-4], but the great hurdle is that the size of the non-orthogonal Hamiltonian matrices becomes prohibitively large for transport simulations of realistically sized devices.

In this work, we have effectively reduced the size of the DFT Hamiltonians, thereby greatly alleviating the computational burden. By extending the recently introduced equivalent model (EM) method [5] to the case of non-orthogonal DFT Hamiltonians, much smaller a set of basis than the original full set of basis is devised to construct an effective Hamiltonian matrix [6]. Simulations of Si NW p-type FETs containing tens of thousands of atoms or more are therefore made possible.

II. Approach

The schematic structure of the Si NW FET that we simulate in this work is shown in Fig. 1. Simulations are done in the following three steps of (1) relaxation of the NW unit cell by the DFT method and extraction of the Hamiltonian and overlap matrices, (2) reduced basis transformation of the imported Hamiltonian matrices, and (3) the non-orthogonal NEGF transport calculations in a self-consistent manner with the Poisson's equation [6]. See Fig. 2.

In the first step, the atomic and electronic structure of Si NWs is obtained by using the openMX DFT package [3] with a s2p2d1 atomic orbitals basis set (13 orbitals per atom). The local density approximation is adopted for the exchange-correlation energy functional. The dangling bonds of Si atoms are passivated by hydrogen atoms. Using a $6 \times 1 \times 1$ Monkhorst-Pack grid, the structures are fully relaxed until the maximum force becomes less than 4×10^{-3} Hartree/Bohr. Then the DFT Hamiltonian and overlap matrices are extracted for the relaxed structures.

For the second step, we have developed a mode space method for non-orthogonal DFT Hamiltonians by extending the EM developed for orthogonal tight-binding Hamiltonians [5,6]. The idea of the method is: (i) a small number of the Bloch states within an energy window of interest are selected to form the basis transformation matrix, (ii) the Hamiltonian matrices are reduced in size via the basis transformation, and (iii) the spurious states (unphysical branches in the band structure diagram) resulted from using a small number of basis are removed via a variational method in an iterative way.

In the third step, with the reduced-sized, mode-space Hamiltonian matrices, the non-orthogonal NEGF approach is utilized for transport calculations. The retarded, electron and hole Green's functions are computed by the recursive Green's function method [7]. The NEGF-Poisson self-consistent steps are implemented for the device simulations. The Hydrogen passivation atoms are included in the charge density
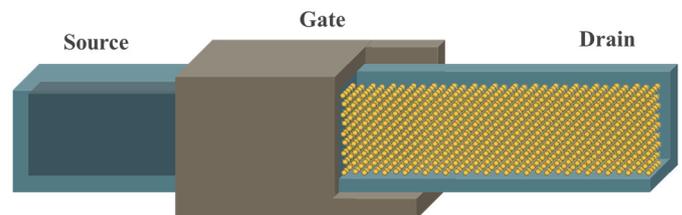

Fig. 1. Schematic structure of Si nanowire field effect transistors. In this work, the channel region is 10 nm long and the source/drain regions are 15 nm long. SiO$_2$ of 1 nm thickness is assumed for the gate oxide. The channel region is surrounded by all-around gates.
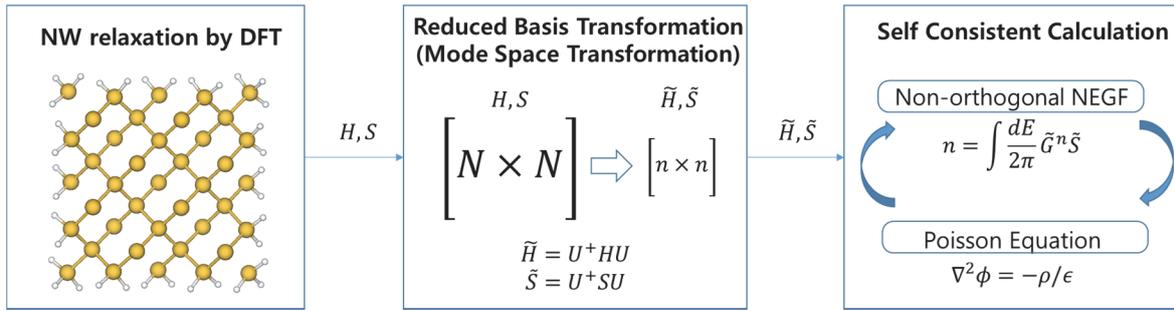
Fig. 2. Simulation steps for DFT-based NEGF transport calculations.

calculations and their electrostatic potential is assumed to be the same as that of Si atom passivated.

III. RESULTS

In Fig. 3, the effect of the basis reduction for Si NW of 3.24 nm×3.24 nm is demonstrated. There are total 413 atoms (313 Si and 100 H atoms) in a unit cell and the size of the imported DFT Hamiltonian matrices is 5369×5369. The energy window of interest is between 1.0 eV above and 0.4 eV below the valence band edge. The number of the initial Bloch states selected is 73 which are marked by open circles in the figure. Fig. 3 (b) shows the E-k diagram at the initial stage, where many unphysical branches are shown to appear. After 11 iteration steps, all the unphysical branches in the energy window are cleared and the final E-k diagram becomes very close to that by the full-sized original DFT Hamiltonians. See Fig. 3 (c). The number of modes at the final stage is 67, which is the size of the mode space Hamiltonians used for the NEGF transport calculations.
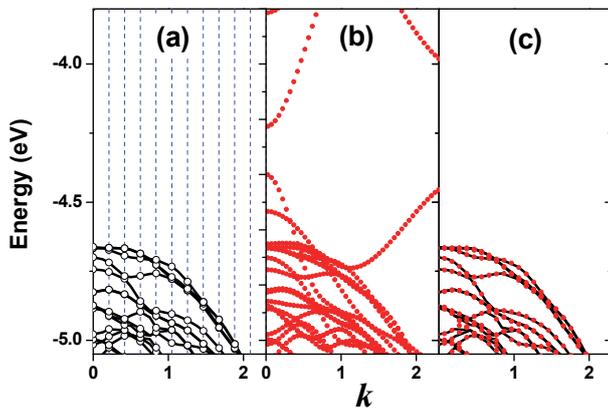
Through the successful size reduction of the imported Hamiltonian and overlap matrices, Si NW FETs are simulated very efficiently. Ballistic transport is assumed in the square Si channel of 3.24 nm width. Fig. 4 shows the transfer characteristics of the p-type device, as the energy window below the valence band edge is gradually expanded. As the window is increased, more number of the modes participate in the transport, resulting in an increase in the size of the mode space Hamiltonian as shown in Fig. 5. What is remarkable in Fig. 4 is that the energy window of 0.4 eV is sufficient because further increase of the window little affects the results, which means that only 1.2 % of the original Hamiltonian in size is needed, for all practical purposes, for transport calculations. A tremendous saving in computational resources and computing time has therefore been achieved.

The same device is simulated using the sp3s*d5 empirical TB Hamiltonian and compared in Fig. 4, whose current-voltage characteristics turn out to be quite close to the DFT-based result. That is not surprising because the overall band structures by the two Hamiltonians are not very different when Si atoms are passivated by H atoms.
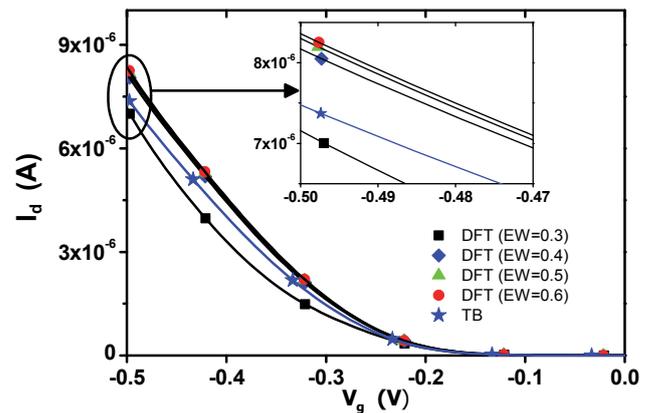

Fig. 3. The Hamiltonian size reduction scheme is demonstrated for Si NW of 3.24nm×3.24nm. The energy window of interest is between 1 eV above and 0.4 eV below the valence band edge. In (a), the selected Bloch states at the initial stage are marked by open circles. Dashed lines indicate the k's that are used for the selection of the Bloch states. In (b) and (c), the energy bands calculated by the reduced-size Hamiltonians (red dots) at the initial and final stages are shown, respectively. The energy bands calculated by the original Hamiltonians are displayed by black solid lines.

Fig. 4. The transfer characteristics of the p-type Si NW FET of 3.24 nm width. $V_{DD} = 0.5\,V$. For DFT-based calculations, the energy window of interest is taken 1.0 eV above and 'EW' below the valence band edge, where EW = 0.3, 0.4, 0.5, and 0.6 eV. In the magnified view, note that EW of 0.4 eV yields the sufficiently close result. The result by using the $sp^3s^*d^5$ empirical TB method is also shown for comparison.

Width (nm)	Energy Window (eV)	Size of Reduced Hamiltonian	Size of Original Hamiltonian	Size-Reduction Ratio (%)
3.24	0.2	28	5369	0.5
	0.3	51		1.0
	0.4	67		1.2
	0.5	89		1.6
	0.6	108		2.0

Fig. 5. The summary of the effect of the Hamiltonian size reduction for Si NW of 3.24 nm width, as the energy window below the valence band edge is gradually expanded. Above the valence band edge, the energy range of 1.0 eV is taken to ensure that there are no unphysical states in the band gap region.

Tremendous speed up in simulation time was achieved. For average 10 self-consistent steps per one bias point, it took less than 1h in wall time to obtain all the data points displayed in Fig. 4 (for the energy window of 0.4 eV). Compared to using the full-size imported DFT Hamiltonians, the extent of savings in CPU time is enormous, estimated to be at least many orders of magnitude faster.

IV. CONCLUSION

We have demonstrated that density functional theory based device simulations of Si nanowire p-type FETs having tens of thousands of atoms or more are feasible within less than 1 hour of CPU time using a small cluster computers. That was possible through successful reduced basis transformation on non-orthogonal DFT Hamiltonians. The methodology introduced in this work can be a useful platform for atomistic device simulations.

REFERENCES

[1] See http://www.itrs2.net/reports.html for "The International Technology Roadmap for Semiconductors."

[2] J. E. Fonseca, T. Kubis, M. Povolotskyi, B. Novakovic, A. Ajoy, G. Hegde, H. Ilatikhameneh, Z. Jiang, P. Sengupta, Y. Tan, and G. Klimeck, "Efficient and realistic device modeling from atomic detail to the nanoscale," J. Comput. Electron., vol. 12, no. 4, pp. 592–600, Oct. 2013.

[3] T. Ozaki, "Variationally optimized atomic orbitals for large-scale electronic structures," Phys. Rev. B, vol. 67, no. 15, p. 155108, Apr. 2003.

[4] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," Phys. Rev., vol. 140, no. 4A, pp. A1133–A1138, Nov. 1965.

[5] G. Mil'nikov, N. Mori, and Y. Kamakura, "Equivalent transport models in atomistic quantum wires," Phys. Rev. B, vol. 85, no. 3, p. 035317, Jan. 2012.

[6] M. Shin, W. J. Jeong, and J. Lee, "Density functional theory based simulations of silicon nanowire field effect transistors," J. Appl. Phys., vol. 119, no. 15, p. 154505, Apr. 2016.

[7] M. P. Anantram, M. S. Lundstrom, and D. E. Nikonov, "Modeling of Nanoscale Devices," Proc. IEEE, vol. 96, no. 9, pp. 1511–1550, Sep. 2008.