Research Article

Efficient Ab-Initio Electron Transport Calculations for Heterostructures by the Nonequilibrium Green's Function Method

Hirokazu Takaki,¹ Nobuhiko Kobayashi,¹ and Kenji Hirose²

¹ Institute of Applied Physics, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

² Smart Energy Research Laboratories, NEC Corporation, 34 Miyukigaoka, Tsukuba, Ibaraki 305-8501, Japan

Correspondence should be addressed to Nobuhiko Kobayashi; nkoba@bk.tsukuba.ac.jp

Received 14 May 2014; Accepted 29 September 2014; Published 27 October 2014

Academic Editor: Miguel A. Correa-Duarte

Copyright © 2014 Hirokazu Takaki et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present an efficient computation technique for ab-initio electron transport calculations based on density functional theory and the nonequilibrium Green's function formalism for application to heterostructures with two-dimensional (2D) interfaces. The computational load for constructing the Green's functions, which depends not only on the energy but also on the 2D Bloch wave vector along the interfaces and is thus catastrophically heavy, is circumvented by parallel computational techniques with the message passing interface, which divides the calculations of the Green's functions with respect to energy and wave vectors. To demonstrate the computational efficiency of the present code, we perform ab-initio electron transport calculations of Al(100)-Si(100)-Al(100) heterostructures, one of the most typical metal-semiconductor-metal systems, and show their transmission spectra, density of states (DOSs), and dependence on the thickness of the Si layers.

1. Introduction

Recent technological advances have made it possible to fabricate nanometer-scale devices. Since device operations are strongly affected by the interfaces of different materials, understanding of the electronic structures and transport properties from atomistic viewpoints has become indispensable for such nanodevices. Ab-initio electron transport calculations based on density functional theory (DFT) [1, 2] combined with the nonequilibrium Green's function (NEGF) formalism [3, 4] have been widely recognized to be highly advantageous for analyzing the electron transport properties of nanostructures.

Metal-semiconductor-metal heterostructures are one of the most important systems for nanodevices such as field effect transistors (FETs) and capacitors. This system has been studied intensively to clarify the electronic states of interfaces, which determine the performance of nanometer-scale electronic devices. For two-dimensional (2D) interfaces, since the electron transport properties depend not only

on the incident electron energy across the interfaces but also on the 2D momentum parallel to the interfaces, the computational load for calculating the Green's functions at each incident energy using a 2D Bloch wave vector has become severely demanding.

When the electron-phonon interactions are small and can be ignored, the couplings to other states with a different energy and wave vector are neglected and then the Green's functions are computed separately for different energies and wave vectors. This enables us to perform a parallel computation with message passing interface (MPI) and reduce the computational burden significantly. To demonstrate the efficiency of this technique, here, we perform ab-initio electron transport calculations of metal-semiconductor-metal heterostructures with 2D interfaces using our program code based on DFT and the NEGF formalism [5, 6] combined with MPI parallelization with respect to the incident energy and 2D Bloch wave vectors. Specifically, we present electronic states and transmission spectra for Al(100)-Si(100)-Al(100)

heterojunctions in which a small number of Si layers are sandwiched between Al electrodes and discuss the dependence of transport properties on the Si layer thickness.

2. Calculation Methods

2.1. Basis Functions and Matrix Elements with Two-Dimensional Periodicity. In this paper, we use atomically localized basis functions $\zeta_{i\mathbf{k}_\parallel}(\mathbf{r})$ for the expansion of wavefunctions [7]. Here, the numerical pseudoatomic orbitals $|\zeta_{i\mathbf{k}_\parallel}\rangle$ are the eigenfunctions of the pseudopotentials under the confinement potential [8] in the 2D Bloch representation:

$$\langle \mathbf{r} | \zeta_{i\mathbf{k}_\parallel} \rangle = \zeta_{i\mathbf{k}_\parallel}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_n e^{-i\mathbf{k}_\parallel \cdot \mathbf{T}_n} \zeta_{in}(\mathbf{r}), \quad (1)$$

where the index i is used as an abbreviation of atom i_a , multiplicity n_i , and azimuthal and magnetic quantum numbers l_i and m_i , respectively; that is, $i = \{i_a, n_i, l_i, m_i\}$. $\mathbf{k}_\parallel$ and n represent the 2D Bloch wave vector and the index of the unit cells along the direction parallel to the electrode interfaces, respectively. Using these basis functions, the overlap matrix and the Kohn-Sham Hamiltonian matrix are written as

$$S_{ij\mathbf{k}_\parallel} = \sum_n e^{-i\mathbf{k}_\parallel \cdot \mathbf{T}_n} \int d\mathbf{r} \zeta_i(\mathbf{r}) \zeta_{jn}(\mathbf{r}), \quad (2)$$

$$H_{ij\mathbf{k}_\parallel}^{\text{KS}} = \sum_n e^{-i\mathbf{k}_\parallel \cdot \mathbf{T}_n} \iint d\mathbf{r} d\mathbf{r}' \zeta_i(\mathbf{r}) H^{\text{KS}}(\mathbf{r}, \mathbf{r}') \zeta_{jn}(\mathbf{r}').$$

The overlap matrix and the Kohn-Sham Hamiltonian matrix in (2) are calculated by the standard procedure for DFT calculations using the numerical pseudoatomic orbital basis set [9].

2.2. Charge Density Construction with Massive Parallelization. In the DFT-NEGF calculations, the density matrix is obtained by integration of the Green's function with respect to energy and the wave vectors [3, 4]. For a system with 2D periodicity, the density matrix is calculated as follows:

$$\rho_{ijn} = \int \frac{d\mathbf{k}_\parallel}{S_{\text{BZ}}} e^{i\mathbf{k}_\parallel \cdot \mathbf{T}_n} \times \left[\int_C dz G_{ij\mathbf{k}_\parallel}^r(z) + \int_{\mu_{\min}=\min(\mu^L, \mu^R)}^{\mu_{\max}=\max(\mu^L, \mu^R)} d\epsilon G_{ij\mathbf{k}_\parallel}^<(\epsilon) \right]. \quad (3)$$

Here, μ^L and μ^R are the chemical potentials of the left and right electrodes, and G^r and $G^<$ are the retarded and lesser Green's functions, respectively, where S_{BZ} is the area of the Brillouin zone in the 2D reciprocal space. The two Green's functions are calculated from

$$G_{\mathbf{k}_\parallel}^r(z) = \left[zS_{\mathbf{k}_\parallel} - H_{\mathbf{k}_\parallel} - \Sigma_{\mathbf{k}_\parallel}^L(z) - \Sigma_{\mathbf{k}_\parallel}^R(z) \right]^{-1}, \quad (4)$$

$$G_{\mathbf{k}_\parallel}^<(\epsilon) = G_{\mathbf{k}_\parallel}^r(\epsilon) \Sigma_{\mathbf{k}_\parallel}^<(\epsilon) G_{\mathbf{k}_\parallel}^a(\epsilon),$$

FIGURE 1: Integration contour for the Green's functions used to calculate the density matrix. The complex contour C is discretized by the Gauss-Legendre scheme, where the total number of grid points of contour C is typically about 40. The contour along the real axis is discretized by a fine mesh.

where G^a is the advanced Green's function and Σ^L and Σ^R are the self-energy matrices of the left and right electrodes, respectively, which are obtained using the surface Green's function technique:

$$\Sigma_{\mathbf{k}_\parallel}^<(\epsilon) = i \left\{ f^{\text{FD}}(\epsilon - \mu^L) \Gamma_{\mathbf{k}_\parallel}^L(\epsilon) + f^{\text{FD}}(\epsilon - \mu^R) \Gamma_{\mathbf{k}_\parallel}^R(\epsilon) \right\}, \quad (5)$$

$$\Gamma_{\mathbf{k}_\parallel}^{L/R}(\epsilon) = -2 \text{Im} \left[\Sigma_{\mathbf{k}_\parallel}^{L/R}(\epsilon) \right].$$

The first term on the right hand side of (3) is the contribution from equilibrium states and the second term is the contribution from nonequilibrium states. They are integrated along the complex energy z contour and the real energy ϵ axis, respectively, as shown in Figure 1.

The most time consuming part of the computation based on the DFT-NEGF formalism is the calculation of the Green's functions G^r and $G^<$ to obtain the density matrix by integration of the Green's functions with respect to energy and the wave vector. In a typical system, the total number of discretized points for the energy and the wave vector in practical calculations are about 10^2 – 2×10^2 and 10^2 – 10^3 , respectively, and thus the total number of discretized points is of the order 10^4 – 10^5 . This heavy computational load is imposed at each iteration procedure for self-consistent field (SCF) in the DFT-NEGF formalism, causing difficulty in the SCF transport calculation.

This heavy computational load can be avoided by using parallel computing. When the electron-phonon interactions are small and can be ignored, the Green's functions depend only on a single set of the energy and the wave vectors as shown in (4), which means that the calculations of the Green's function are performed independently for various sets of energies and wave vectors. The calculation of (3) is separated by parallelization with MPI as shown in Figure 2(a), and the MPI processes are distributed among multicore processors. The integration of the Green's functions is also performed in parallel.

FIGURE 2: (a) Flowchart of DFT-NEGF calculations with parallelization by MPI. The calculation process surrounded by the dashed line corresponds to the calculations of the Green's functions, which are the central and most time consuming parts of the present formalism. The Green's functions are calculated independently for various sets of complex/real energies and 2D Bloch wave vectors, whose processes are assigned to multicore processor by MPI. During the calculations of the Green's functions, the density matrix is partially calculated by integration with respect to the energy and the wave vectors in each core. After the calculations of the density matrix in each core are completed, the Hamiltonian matrix is constructed by summing the elements with communication among MPI processes, and then all the matrix elements of the density matrix are obtained. (b) Elapsed time and speed-up ratio of total elapsed time resulting from the parallelization by the MPI in the ab-initio electron transport calculations for an Al(100)-Si(100)-Al(100) heterostructure. The speed-up ratio is calculated as T_1/T_p , where T_1 and T_p are the elapsed times for serial and parallel computing, respectively. The number of discretized grid points in the calculations are $N_{\mathbf{k}_{\parallel}} = 121$, $N_z = 50$, and $N_{\epsilon} = 100$. For reference, the ideal speed-up ratio is also represented by the green dotted line. The calculation speed increased linearly while lagging slightly behind the ideal speed-up ratio.

The partially integrated density matrix ρ_{ijn}^p in each core is calculated using

$$\rho_{ijn}^p = \sum_{l=1}^{N_{\mathbf{k}_{\parallel}}^p} w_l^{\mathbf{k}_{\parallel}} \left\{ \sum_{m=1}^{N_z^p} \text{Im} \left[w_m^z G_{ij\mathbf{k}_{\parallel}}^r(z_m) \right] + \sum_{n=1}^{N_{\epsilon}^p} w_n^{\epsilon} G_{ij\mathbf{k}_{\parallel}}^<(\epsilon_n) \right\}, \quad (6)$$

where $N_{\mathbf{k}_{\parallel}}^p$ and $w_l^{\mathbf{k}_{\parallel}/z/\epsilon}$ are the total number of discretized points and the weight factor for the integration with respect to $\mathbf{k}_{\parallel}$, z , and ϵ , respectively. After the partial integration, the density matrix ρ_{ijn}^p in each processor is distributed among all the processors and is summed to obtain the total density matrix ρ_{ijn} . Only values of ρ_n^p are communicated in the parallel computing, which significantly reduces the MPI communication latency. Since the total number of discretized points of energy and wave vectors is more than 10^3 , the computational efficiency is improved as the total number of processors increases. The number of distributed arrays in the communication between the processors required for the calculation of the density matrix becomes larger as a

trade-off in such large-scale parallel computing, and thus the maximum efficiency should be determined.

The speed-up ratio of the present numerical code is shown in Figure 2(b) as a function of the total number of cores for ab-initio SCF electron transport calculations of an Al(100)-Si(100)-Al(100) heterostructure. Here, the speed-up ratio is defined as T_1/T_p , where T_1 and T_p are the total elapsed times for serial and parallel computing, respectively. We see that the speed-up ratio exhibits almost linear behavior with increasing number of cores, showing the efficiency of MPI parallel computing using the present code. This enables us to perform the transport calculations of large-scale systems with 2D interfaces.

After the SCF loop converges, the transmission function $T(\epsilon)$ for the electron transport and the density of states (DOS) $D(\epsilon)$ are calculated from

$$\mathcal{T}(\epsilon) = \int \frac{d\mathbf{k}_{\parallel}}{S_{\text{BZ}}} \text{Tr} \left[\Gamma_{\mathbf{k}_{\parallel}}^L(\epsilon) G_{\mathbf{k}_{\parallel}}^r(\epsilon) \Gamma_{\mathbf{k}_{\parallel}}^R(\epsilon) G_{\mathbf{k}_{\parallel}}^a(\epsilon) \right], \quad (7)$$

$$D(\epsilon) = -\frac{1}{\pi} \text{Tr} \{ \text{Im} [G^r(\epsilon) S] \}. \quad (8)$$

FIGURE 3: Schematic diagram of Al(100)-Si(100)-Al(100) heterostructure. The system is periodic along the x and y directions parallel to the interfaces and the transport direction is taken along the z direction, which is oriented in the (100) direction normal to the interfaces. The square region surrounded by dashed lines represents 2D unit cell. Here, a_0 and d are the lattice constant of a Si monolayer and the total thickness of the Si layers, respectively.

The electric current $I(V)$ flowing through the system at an applied bias voltage V is obtained by integrating the transmission function with respect to the energy within the different chemical potentials of the left and right electrodes as

$$I(V) = \frac{2e}{h} \int_{\mu_{\min}}^{\mu_{\max}} d\varepsilon \mathcal{T}(\varepsilon) \times [f^{\text{FD}}(\varepsilon - \mu_{\max}) - f^{\text{FD}}(\varepsilon - \mu_{\min})]. \quad (9)$$

Note that, in the same way as for the parallel computations of the Green's functions in the SCF calculations, the computations of the Green's functions along the real energy axis for the transmission function $\mathcal{T}(\varepsilon)$ are also performed by parallelization with respect to the energy and the wave vectors.

3. Results

As an application of the present method, we calculate the electron transport properties of Al(100)-Si(100)-Al(100) heterostructures. The system is composed of a small number of Si layers sandwiched between Al electrodes with (100) direction normal to the interfaces in the transport direction. A schematic diagram of the system is shown in Figure 3. For this system, we consider the dependence of the transport properties and the electronic states on the thickness of the Si layers d between the interfaces.

First, note that the system becomes metallic when the Si layers are thin owing to the hybridization of electronic states of metallic electrodes. On the other hand, the system becomes semiconducting with a finite energy gap as the Si layers become thick. It is interesting to consider the Si layer thickness needed for the system to show semiconducting behavior.

Here, we analyze the transport properties of Al(100)-Si(100)-Al(100) heterostructures with various Si layer thicknesses for the bulk lattice constant of $a_0 = 5.431$ [Å]. Single- ζ polarized basis sets [8] with the GGA-PBE functional [10] for the exchange-correlation energy and nonlocal pseudopotentials in the fully separable Kleinman-Bylander form [11] are used in the calculations.

Figure 4 shows the transmission probability (Figure 4(a)) through the two Al electrodes and the DOS (Figure 4(b)) for various thicknesses of Si layers. We see that the system becomes metallic owing to metallic electrode contacts in the case of the minimum Si layer thickness of $d = a_0$. We also note that the direct interaction between the electrodes exists when their distance is less than twice the cutoff radius of the basis functions for Al atoms. In the case of more than three Si layers, finite energy gaps appear around the Fermi level, which originate from the semiconducting properties of the Si layers. Since the lattice constants of Al and Si along the x and y directions are comparable, that is, $a^{\text{Al}}/a^{\text{Si}} = 0.997$ in the present heterostructure, the electron density and effective potential near the center of the Si layers rapidly recover the bulk values as the Si layers become thick. Thus, the electron transport properties at the atomic scale are governed by the thickness of the layers.

4. Conclusions

We have developed an efficient numerical calculation code for ab-initio electron transport with 2D interfaces based on DFT and the NEGF formalism in atomically localized basis sets. The parallel computation techniques with the MPI are employed for efficient computation of the density matrix via the Green's functions, which depend on the energy and 2D wave vectors parallel to the interfaces. Since each matrix element of the Green's functions depends only on a single energy and 2D wave vector point, it is possible to divide the calculations with respect to the energy and the wave vectors, which are carried out separately in different processors.

As a demonstration of the efficiency of the present code, numerical calculations of the electron transport properties of Al(100)-Si(100)-Al(100) heterostructures are presented, where the total number of points of energy and wave vectors is about 10^5 . The computational efficiency improves as the total number of processors increases, with the communication latency between the processors becoming negligible. As an application of the present technique, electron transport properties such as transmission spectra and the DOS were investigated for the above system with different thicknesses of Si layers.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

FIGURE 4: Transmission probability (a) and DOS (b) for various thicknesses of Si layers around the Fermi level. The system shows metallic behavior with finite transmission for the entire energy regime for the minimum Si layer thickness. For a thickness of more than three-layers, finite energy gaps are formed around the Fermi level, which originate from the semiconducting properties of the Si layers.

Acknowledgments

This work was supported by a Grant-in-Aid for JSPS Fellows 231588 and Scientific Research 23360018 from the Ministry of Education, Culture, Sports, Science and Technology, Japan. Numerical calculations are partly performed using the super-computers system B at ISSP, University of Tokyo.

References

- [1] P. Holienberg and W. Kolm, "Inhomogeneous electron gas," *Physical Review*, vol. 136, no. 3, pp. B864–B871, 1964.
- [2] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical Review*, vol. 140, no. 4A, pp. 1133–1138, 1965.
- [3] J. Taylor, H. Guo, and J. Wang, "Ab initio modeling of quantum transport properties of molecular electronic devices," *Physical Review B*, vol. 63, no. 24, Article ID 245407, 13 pages, 2001.
- [4] M. Brandbyge, J.-L. Mozos, P. Ordejon, J. Taylor, and K. Stokbro, "Density-functional method for nonequilibrium electron transport," *Physical Review B*, vol. 65, Article ID 165401, 2002.
- [5] H. Haug and A.-P. Jaulio, *Quantum Kinetics in Transport and Optics of Semiconductors*, Springer, Berlin, Germany, 2nd edition, 2007.
- [6] S. Datta, *Electronic Transport in Mesoscopic Systems*, Cambridge University Press, Cambridge, UK, 1995.
- [7] R. A. Evarestov, *Quantum Chemistry of Solids The LG AO First Principle Treatment of Crystals*, Springer, Berlin, Germany, 2007.
- [8] T. Ozaki and H. Kino, "Numerical atomic basis orbitals from H to Kr," *Physical Review B*, vol. 69, no. 19, Article ID 195113, 2004.
- [9] J. M. Soler, E. Artacho, J. D. Gale et al., "The SIESTA method for ab initio order-N materials simulation," *Journal of Physics: Condensed Matter*, vol. 14, no. 11, p. 2745, 2002.
- [10] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Physical Review Letters*, vol. 77, Article ID 3865, p. 1396, 1996.

- [11] L. Kleiimai and D. M. Bylander, "Efficacious form for model pseudopotentials," *Physical Review Letters*, vol. 48, no. 20, pp. 1425–1428, 1982.