Received 3 April 2019; revised 24 May 2019 and 10 June 2019; accepted 24 June 2019. Date of publication 27 June 2019; date of current version 12 July 2019. The review of this paper was arranged by Editor C. Bulucea.

Digital Object Identifier 10.1109/JEDS.2019.2925402

Efficient Atomistic Simulation of Heterostructure Field-Effect Transistors

YONGSOO AHN AND MINCHEOL SHIN

School of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea

CORRESPONDING AUTHOR: M. SHIN (e-mail: mshin@kaist.edu)

This work was supported in part by the Basic Science Research Program through the National Research Foundation (NRF) of Korea funded by the Ministry of Science and ICT under Grant 2017R1A2B2005679, and in part by the BK21 Plus in the Korea Advanced Institute of Science and Technology.

ABSTRACT In this paper, atomistic-level quantum mechanical simulations are performed for nanoscale field-effect transistors (FETs) with lateral or vertical heterojunction, within the non-equilibrium Green's function formalism. For efficient simulation of such heterostructure FETs, a novel approach is developed where the Green's functions are calculated by complementarily using the two algorithms of the recursive Green's function and the *R*-matrix. The *R*-matrix algorithm is extended to seamlessly combine the two methods on the open system and an algorithm for the electron correlation function based on the extended *R*-matrix algorithm is also developed. The proposed method significantly reduces simulation time, making rigorous atomistic simulations of heterojunction FETs possible. As an application, device simulations are performed for the germanane/InSe vertical tunneling FET (VTFET) modeled through the first-principles density functional theory. Our simulation results reveal that the germanane/InSe VTFET is a promising candidate for future low power applications.

INDEX TERMS Heterostructure field-effect transistors, quantum transport, non-equilibrium Green's function, recursive Green's function, *R*-matrix.

I. INTRODUCTION

Devices with heterojunctions are being widely considered for the future field-effect transistors (FETs) such as tunneling FETs (TFETs) [1]–[4] and Schottky barrier FETs [5], [6]. By employing the heterojunctions, the band gap and band offsets can be altered, thereby the device performances can be engineered. In order to properly describe such band properties and to naturally capture the interface effects in heterojunctions, it is essential to model the heterostructure devices in the atomistic level. Although the empirical tight-binding method, one of the most popular atomistic-level approaches, has been widely adopted for nanodevice modeling, it is not adequate for heterojunction [7]. Recently, the first-principles density functional theory (DFT) has been increasingly applied to modeling heterostructure FETs, since the devices with any atomic structures can be rigorously simulated by self-consistently combining DFT with non-equilibrium Green's function (NEGF) that is the most prevailing quantum transport approach.

As a practical approach, recent works [8]–[10] have employed the simulation framework where the DFT and NEGF parts are performed sequentially. Namely, after modeling a device by the DFT method, the Hamiltonians with the basis set of linear combination of atomic orbitals (LCAO) are extracted and imported into a NEGF-Poisson solver. This framework has served a powerful simulation platform for the homogeneous devices and can also be applicable to devices with heterojunctions. For the latter, however, the large size of junction Hamiltonians can become an obstacle to efficient NEGF simulations. Especially, if a junction is modeled in DFT using the supercell approximation [11], to capture the right physics, the supercell should be large enough to prevent the coupling between images [12], [13], resulting in the increase of the size of junction Hamiltonian.

In the recursive Green's function (RGF) method [14]–[16], which is the widely-used NEGF algorithm, the computational burden to handle large-sized junction Hamiltonians can be reduced by regularly slicing the junction region into smaller

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

slabs. It is in principle possible, but it is in practice not efficient when it comes to not-well structured atom arrangement. This kind of problem was tackled in [17], [18] by exploiting parallel computations. Other methods such as hierarchical Shur complement [19], [20] and fast inverse using nested dissection methods [21] are more efficient than RGF, but the usage of these methods is limited to the systems where the matrix decomposition processes (e.g., LDL^{T}, LU) are applicable.

Mil'nikov *et al.* have developed the R-matrix propagation method for the NEGF simulations of both continuum [22] and atomistic devices [23]. The R-matrix method is not restricted to the shape of device geometry or Hamiltonian because the method not only takes a group of atoms arbitrarily from the device region but also uses only basic matrix operations. The R-matrix algorithm is therefore more suitable for handling junction Hamiltonians than the aforementioned methods.

The aim of this paper is to present a methodology to reduce the computational burden for NEGF simulations of FETs with heterojunctions modeled in the atomistic level. Our main idea is to use the R-matrix and RGF methods complementarily; i.e., the R-matrix method is employed for the junctions usually having large-sized Hamiltonians, and for those of the rest of the device (homogeneous unit cells), the RGF method is used. For this purpose, we extend the R-matrix algorithm so that the algorithm directly applies to the open systems. The elapsed time for calculating retarded Green's functions is reported to assess the proposed approach, and an application to an example device of germanane/InSe vertical TFET (VTFET) is demonstrated.

II. METHODOLOGY

A. CONCEPT OF THE PROPOSED METHOD

Consider a heterostructure device that consists of a 'device region' and semi-infinite 'leads,' as shown in Fig. 1(a). The device region can be divided into homogeneous unit cells and junction region. The homogeneous unit cells have usually small-sized Hamiltonian, whereas the junction unit cell has relatively large-sized Hamiltonian. In the widely-used RGF method, the unit cell is usually chosen as the slab, and the numerical cost is cubically proportional to the size (N_s) of Hamiltonian for the slab. Thus, the large-sized junction Hamiltonian gives substantial burden to the NEGF simulation employing the RGF. To alleviate this burden, the R-matrix algorithm is applied to the junction region in the proposed method (termed as RGF+R-matrix method) as illustrated in Fig. 1(b). For the junction region, the operation count of the R-matrix algorithm scales as $N_a^3 n_a N_c^2$ [23], where N_a is the size of Hamiltonian for atom cluster a, n_a is the number of cluster a in the junction, and N_c is the size of Hamiltonian for the boundary atoms of atom cluster C. With a limited interaction range, N_a and N_c are typically much smaller than N_s for the junction region. So, the numerical burden for the junction region can be reduced, thereby the numerical efficiency of whole NEGF simulation can be significantly improved.

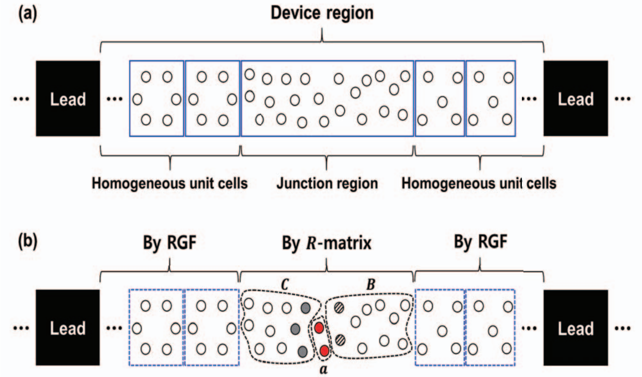

FIGURE 1. (a) A schematic of heterostructure device. The device region of the heterostructure device is divided into homogeneous unit cells and junction region. The circles indicate atoms, and the boxes in the device region denote the unit cells, which are usually chosen as the slabs in the RGF method. (b) A concept of the proposed method where RGF and R-matrix algorithms are applied to homogeneous unit cells and junction unit cell, respectively. The boxes are the slabs for RGF, and the shaded circles indicate the atoms that contribute to the calculation of the R-matrix algorithm at an iteration step.

Because the RGF method includes the open system while the R-matrix method was originally designed for the closed system [23], to combine the RGF and R-matrix methods seamlessly, the R-matrix method should be extended so that the algorithm applies to open system directly. In Sections II-B–II-D, we discuss that in detail, and the extended R-matrix method is then combined with the RGF method in Section II-E.

B. EXTENSION OF R-MATRIX METHOD TO OPEN SYSTEM

Let us start by writing down a linear equation describing an open system as

$$(\varepsilon \mathbf{S}_{DD} - \mathbf{H}_{DD} - \Sigma_{DD})\Psi = \Phi, \tag{1}$$

where $\mathbf{H}_{DD}$ and $\mathbf{S}_{DD}$ are the Hamiltonian and the overlap matrices for the device region, respectively, Σ_{DD} is the self-energy matrix which accounts for semi-infinite leads and scattering effect, if any (throughout this work, we neglect the scattering effects, but the method proposed in this work is applicable even if the scattering effects are included), Ψ is the wave function for the device region, and Φ is the source term that describes the carrier injection. Since we consider the open system, the energy ε in Eq. (1) is usually given as a complex number having a small imaginary part (typically $< 10^{-9}$). Let C be a cluster of atoms arbitrarily taken from the device region and B be the rest of the device region. Then, $\mathbf{H}_{DD}$, $\mathbf{S}_{DD}$, and Σ_{DD} can be represented as

$$\mathbf{H}_{DD} = \begin{pmatrix} \mathbf{H}_{CC} & \mathbf{H}_{CB} \\ \mathbf{H}_{BC} & \mathbf{H}_{BB} \end{pmatrix},$$

$$\mathbf{S}_{DD} = \begin{pmatrix} \mathbf{S}_{CC} & \mathbf{S}_{CB} \\ \mathbf{S}_{BC} & \mathbf{S}_{BB} \end{pmatrix},$$

$$\Sigma_{DD} = \begin{pmatrix} \Sigma_{CC} & \Sigma_{CB} \\ \Sigma_{BC} & \Sigma_{BB} \end{pmatrix}. \tag{2}$$

The R-matrix method is similar to the contact block reduction (CBR) method [24], [25] in that device partitioning is performed, but the partitioning in the two methods is different. In the CBR method, the device region is divided into the boundary regions coupled with the leads and the rest of the device region while in the R-matrix method, the device region is divided into two regions C and B arbitrarily.

Now Eq. (1) can be written as

$$\begin{pmatrix} \mathbf{A}_{CC} & \mathbf{A}_{CB} \\ \mathbf{A}_{BC} & \mathbf{A}_{BB} \end{pmatrix} \begin{pmatrix} \Psi_C \\ \Psi_B \end{pmatrix} = \begin{pmatrix} \Phi_C \\ \Phi_B \end{pmatrix}, \tag{3}$$

with

$$\begin{aligned} \mathbf{A}_{CC} &= \varepsilon \mathbf{S}_{CC} - \mathbf{H}_{CC} - \Sigma_{CC}, \\ \mathbf{A}_{CB} &= \varepsilon \mathbf{S}_{CB} - \mathbf{H}_{CB} - \Sigma_{CB}, \\ \mathbf{A}_{BC} &= \varepsilon \mathbf{S}_{BC} - \mathbf{H}_{BC} - \Sigma_{BC}, \\ \mathbf{A}_{BB} &= \varepsilon \mathbf{S}_{BB} - \mathbf{H}_{BB} - \Sigma_{BB}. \end{aligned} \tag{4}$$

C. RETARDED GREEN'S FUNCTION BY EXTENDED R-MATRIX METHOD

C.1. FORWARD PROPAGATION

Now we define the R matrix $\mathbf{R}_{cc}$ as

$$\mathbf{R}_{cc} = P_c \mathbf{A}_{CC}^{-1} P_c, \tag{5}$$

where c is the boundary atoms of C and P_c is the projection operator from C onto c. The definition of $\mathbf{R}_{cc}$ in this work is different from that in [23] in that Eq. (5) includes the self-energies.

We now pick a group of atoms $a \in B$ and add a into C. Note that a can be selected arbitrarily in principle. The wave functions at c and a can be obtained as

$$\begin{aligned} \Psi_c &= -\mathbf{R}_{cc}\mathbf{A}_{ca}\Psi_a - \mathbf{R}_{cc}\mathbf{A}_{c\tilde{B}}\Psi_{\tilde{B}} + \mathbf{R}_{cc}\Phi_c, \\ \Psi_a &= -\mathbf{R}_{aa}\mathbf{A}_{ac}\Psi_c - \mathbf{R}_{aa}\mathbf{A}_{a\tilde{B}}\Psi_{\tilde{B}} + \mathbf{R}_{aa}\Phi_a, \end{aligned} \tag{6}$$

where the set $\tilde{B} = B - a$ and $\mathbf{R}_{aa} = \mathbf{A}_{aa}^{-1} = (\varepsilon \mathbf{S}_{aa} - \mathbf{H}_{aa} - \Sigma_{aa})^{-1}$. Performing proper matrix operations for Eq. (6) as in [23], we can obtain

$$\begin{pmatrix} \Psi_c \\ \Psi_a \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{R}}_{cc} & \tilde{\mathbf{R}}_{ca} \\ \tilde{\mathbf{R}}_{ac} & \tilde{\mathbf{R}}_{aa} \end{pmatrix} \begin{pmatrix} -\mathbf{A}_{c\tilde{B}}\Psi_{\tilde{B}} + \Phi_c \\ -\mathbf{A}_{a\tilde{B}}\Psi_{\tilde{B}} + \Phi_a \end{pmatrix}, \tag{7}$$

where

$$\begin{aligned} \tilde{\mathbf{R}}_{cc} &= \mathbf{R}_{cc} + \mathbf{R}_{cc}(-\mathbf{A}_{ca})\tilde{\mathbf{R}}_{aa}(-\mathbf{A}_{ac})\mathbf{R}_{cc}, \\ \tilde{\mathbf{R}}_{ca} &= \mathbf{R}_{cc}(-\mathbf{A}_{ca})\tilde{\mathbf{R}}_{aa}, \\ \tilde{\mathbf{R}}_{ac} &= \tilde{\mathbf{R}}_{aa}(-\mathbf{A}_{ac})\mathbf{R}_{cc}, \\ \tilde{\mathbf{R}}_{aa} &= (\mathbf{I}_{aa} - \mathbf{R}_{aa}\mathbf{A}_{ac}\mathbf{R}_{cc}\mathbf{A}_{ca})^{-1}\mathbf{R}_{aa}. \end{aligned} \tag{8}$$

By taking only boundary atoms from the set $c \cup a$ (see Fig. 2(a)), we can construct new atom cluster $\tilde{c}$ and its R matrix ($\mathbf{R}_{\tilde{c}\tilde{c}}$), that is,

$$\mathbf{R}_{\tilde{c}\tilde{c}} = P_{\tilde{c}} \tilde{\mathbf{R}}_{(c\cup a)(c\cup a)} P_{\tilde{c}}, \tag{9}$$

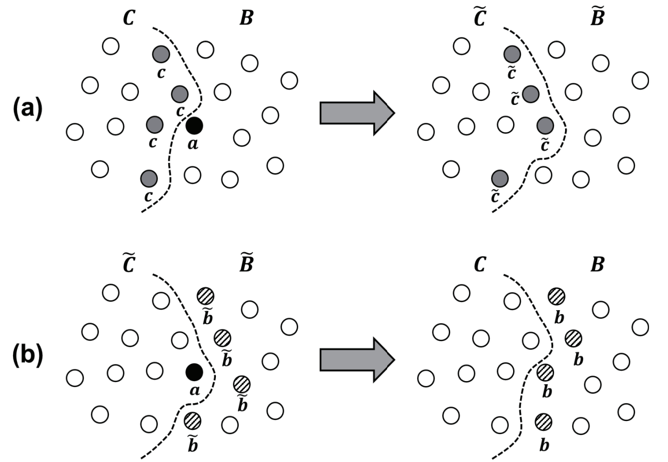

FIGURE 2. A schematic diagram of R-matrix propagation. The gray and patterned circles indicate the boundary atoms of C (or $\tilde{C}$) and B (or $\tilde{B}$), respectively. (a) In the forward propagation, a (black circle) is chosen from B and is added into C. The boundary atoms of the set $c \cup a$ constitute an atom cluster $\tilde{c}$. (b) In the backward propagation, a is subtracted from $\tilde{C}$ and is added into $\tilde{B}$, reconstructing C and B. Here, one atom is excluded from $a \cup \tilde{b}$ after the boundary projection from $a \cup \tilde{b}$ onto b.

where $P_{\tilde{c}}$ is the projection operator from $c \cup a$ to $\tilde{c}$, and

$$\tilde{\mathbf{R}}_{(c\cup a)(c\cup a)} = \begin{pmatrix} \tilde{\mathbf{R}}_{cc} & \tilde{\mathbf{R}}_{ca} \\ \tilde{\mathbf{R}}_{ac} & \tilde{\mathbf{R}}_{aa} \end{pmatrix}. \tag{10}$$

$\tilde{c}$ and $\mathbf{R}_{\tilde{c}\tilde{c}}$ of an iteration step are used as c and $\mathbf{R}_{cc}$ at the next iteration, respectively, and iterations continue. $\mathbf{R}_{\tilde{c}\tilde{c}}$ obtained at the last iteration is equal to the diagonal block of retarded Green's function, corresponding to the set a taken at the last iteration step.

Eqs. (7) and (8) are consistent with Eqs. (17) and (18) in [23], respectively. This implies that, with the same recipes used in the closed system, R matrix can be propagated for the open system.

C.2. BACKWARD PROPAGATION

For the device region divided into $\tilde{C} = C \cup a$ and $\tilde{B} = B - a$ (see the left side of Fig. 2(b)), the retarded Green's function $\mathbf{G}$ can be given from

$$\begin{pmatrix} \mathbf{A}_{\tilde{C}\tilde{C}} & \mathbf{A}_{\tilde{C}\tilde{B}} \\ \mathbf{A}_{\tilde{B}\tilde{C}} & \mathbf{A}_{\tilde{B}\tilde{B}} \end{pmatrix} \begin{pmatrix} \mathbf{G}_{\tilde{C}\tilde{C}} & \mathbf{G}_{\tilde{C}\tilde{B}} \\ \mathbf{G}_{\tilde{B}\tilde{C}} & \mathbf{G}_{\tilde{B}\tilde{B}} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_{\tilde{C}\tilde{C}} & \mathbf{0}_{\tilde{C}\tilde{B}} \\ \mathbf{0}_{\tilde{B}\tilde{C}} & \mathbf{I}_{\tilde{B}\tilde{B}} \end{pmatrix}. \tag{11}$$

$\mathbf{G}$ satisfies the Dyson's equations

$$\mathbf{G} = \mathbf{G}^0 + \mathbf{G}^0\mathbf{U}\mathbf{G} \tag{12}$$

$$\quad = \mathbf{G}^0 + \mathbf{G}\mathbf{U}\mathbf{G}^0, \tag{13}$$

where

$$\mathbf{G}^0 = \begin{pmatrix} \mathbf{G}^0_{\tilde{C}\tilde{C}} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^0_{\tilde{B}\tilde{B}} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{-1}_{\tilde{C}\tilde{C}} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{-1}_{\tilde{B}\tilde{B}} \end{pmatrix}, \tag{14}$$

$$\mathbf{U} = \begin{pmatrix} \mathbf{0} & -\mathbf{A}_{\tilde{C}\tilde{B}} \\ -\mathbf{A}_{\tilde{B}\tilde{C}} & \mathbf{0} \end{pmatrix}. \tag{15}$$

The boundary terms of $\mathbf{G}_{\tilde{C}\tilde{C}}$ in Eq. (12) can be written as

$$\begin{pmatrix}\mathbf{G}_{cc} & \mathbf{G}_{ca}\\ \mathbf{G}_{ac} & \mathbf{G}_{aa}\end{pmatrix} = \begin{pmatrix}\tilde{\mathbf{R}}_{cc} & \tilde{\mathbf{R}}_{ca}\\ \tilde{\mathbf{R}}_{ac} & \tilde{\mathbf{R}}_{aa}\end{pmatrix} + \begin{pmatrix}\tilde{\mathbf{R}}_{cc} & \tilde{\mathbf{R}}_{ca}\\ \tilde{\mathbf{R}}_{ac} & \tilde{\mathbf{R}}_{aa}\end{pmatrix}\begin{pmatrix}-\mathbf{A}_{c\tilde{b}}\\ -\mathbf{A}_{a\tilde{b}}\end{pmatrix} \times \mathbf{G}_{\tilde{b}\tilde{b}}\begin{pmatrix}-\mathbf{A}_{\tilde{b}c} & -\mathbf{A}_{\tilde{b}a}\end{pmatrix}\begin{pmatrix}\tilde{\mathbf{R}}_{cc} & \tilde{\mathbf{R}}_{ca}\\ \tilde{\mathbf{R}}_{ac} & \tilde{\mathbf{R}}_{aa}\end{pmatrix}, \quad (16)$$

where $\tilde{b}$ is the boundary atom set of $\tilde{B}$. So the retarded Green's function corresponding to a is

$$\mathbf{G}_{aa} = \tilde{\mathbf{R}}_{aa} + \Theta_{a\tilde{b}}\mathbf{G}_{\tilde{b}\tilde{b}}\Theta_{\tilde{b}a}, \quad (17)$$

where

$$\Theta_{a\tilde{b}} = \tilde{\mathbf{R}}_{ac}(-\mathbf{A}_{c\tilde{b}}) + \tilde{\mathbf{R}}_{aa}(-\mathbf{A}_{a\tilde{b}}),$$
$$\Theta_{\tilde{b}a} = (-\mathbf{A}_{\tilde{b}c})\tilde{\mathbf{R}}_{ca} + (-\mathbf{A}_{\tilde{b}a})\tilde{\mathbf{R}}_{aa}. \quad (18)$$

Like in [23], $\Theta_{a\tilde{b}}$ and $\Theta_{\tilde{b}a}$ must be stored in the forward propagation to calculate $\mathbf{G}_{aa}$. To compute the off-diagonal parts of $\mathbf{G}$, we move on to $\mathbf{G}_{\tilde{B}\tilde{C}}$. Performing the boundary projection for Eq. (13), we can obtain $\mathbf{G}_{a\tilde{b}}$ and $\mathbf{G}_{\tilde{b}a}$ from $\mathbf{G}_{\tilde{C}\tilde{B}}$ and $\mathbf{G}_{\tilde{B}\tilde{C}}$, respectively, and they are

$$\mathbf{G}_{a\tilde{b}} = \Theta_{a\tilde{b}}\mathbf{G}_{\tilde{b}\tilde{b}},$$
$$\mathbf{G}_{\tilde{b}a} = \mathbf{G}_{\tilde{b}\tilde{b}}\Theta_{\tilde{b}a}. \quad (19)$$

The initial $\mathbf{G}_{\tilde{b}\tilde{b}}$ for the backward propagation is set to be the final $\mathbf{R}_{\tilde{c}\tilde{c}}$ in the forward propagation. To calculate the matrix that plays the same role as $\mathbf{G}_{\tilde{b}\tilde{b}}$ in the next backward iteration, we build the retarded Green's function for the set $a \cup \tilde{b}$ as

$$\mathbf{G}_{(a\cup\tilde{b})(a\cup\tilde{b})} = \begin{pmatrix}\mathbf{G}_{aa} & \mathbf{G}_{a\tilde{b}}\\ \mathbf{G}_{\tilde{b}a} & \mathbf{G}_{\tilde{b}\tilde{b}}\end{pmatrix}. \quad (20)$$

In Eq. (20), by carrying out the boundary projection from the set $a \cup \tilde{b}$ to b (see Fig. 2(b)), where b is the boundary atom set of B, $\mathbf{G}_{bb}$ is obtained as

$$\mathbf{G}_{bb} = P_b\mathbf{G}_{(a\cup\tilde{b})(a\cup\tilde{b})}P_b, \quad (21)$$

where P_b is the projection operator from the set $a \cup \tilde{b}$ to b. $\mathbf{G}_{bb}$ is used as $\mathbf{G}_{\tilde{b}\tilde{b}}$ of the next backward iteration, and the backward iterations continue.

D. CARRIER AND CURRENT DENSITIES BY EXTENDED R-MATRIX METHOD

The electron density by the NEGF method can be calculated using the electron correlation function $\mathbf{G}^n$. At an atom α, the electron density n_α is calculated through

$$n_\alpha = \frac{g_s}{2\pi v_\alpha}\int d\varepsilon\ \mathrm{Re}\left[\mathrm{Tr}\left\{\left(\mathbf{G}^n(\varepsilon)\mathbf{S}_{DD}\right)_{\alpha\alpha}\right\}\right], \quad (22)$$

where g_s is the spin degeneracy, v_α is the volume of α, and the subscript $\alpha\alpha$ indicates the matrix block corresponding to atom α.

To derive the recipes for $\mathbf{G}^n$, let us start with the definition of $\mathbf{G}^n$ of the device region divided into C and B. That is,

$$\begin{pmatrix}\mathbf{A}_{CC} & \mathbf{A}_{CB}\\ \mathbf{A}_{BC} & \mathbf{A}_{BB}\end{pmatrix}\begin{pmatrix}\mathbf{G}^n_{CC} & \mathbf{G}^n_{CB}\\ \mathbf{G}^n_{BC} & \mathbf{G}^n_{BB}\end{pmatrix} = \begin{pmatrix}\Sigma^{\mathrm{in}}_{CC} & \Sigma^{\mathrm{in}}_{CB}\\ \Sigma^{\mathrm{in}}_{BC} & \Sigma^{\mathrm{in}}_{BB}\end{pmatrix}\begin{pmatrix}\mathbf{G}^\dagger_{CC} & \mathbf{G}^\dagger_{CB}\\ \mathbf{G}^\dagger_{BC} & \mathbf{G}^\dagger_{BB}\end{pmatrix}, \quad (23)$$

where the superscript '†' is the Hermitian conjugate of matrix and the in-scattering function Σ^{in} is defined as

$$\Sigma^{\mathrm{in}} = \sum_{j=1}^{N_{\mathrm{lead}}}\Gamma_j f_j, \quad (24)$$

where N_{lead} is the number of leads in the system and Γ_j and f_j are the broadening matrix and Fermi-Dirac distribution function of lead j, respectively. With the self-energy Σ_j, Γ_j is given as

$$\Gamma_j = i\left(\Sigma_j - \Sigma_j^\dagger\right). \quad (25)$$

$\mathbf{G}^n$ satisfies the Dyson's equations

$$\mathbf{G}^n = \mathbf{G}^{n0} + \mathbf{G}^{n0}\mathbf{U}^\dagger\mathbf{G}^\dagger + \mathbf{G}^0\mathbf{U}\mathbf{G}^n \quad (26)$$
$$= \mathbf{G}^{n0} + \mathbf{G}\mathbf{U}\mathbf{G}^{n0} + \mathbf{G}^n\mathbf{U}^\dagger\left(\mathbf{G}^0\right)^\dagger, \quad (27)$$

where

$$\mathbf{G}^{n0} = \mathbf{G}^0\Sigma^{\mathrm{in}}\left(\mathbf{G}^0\right)^\dagger. \quad (28)$$

As in the calculation of $\mathbf{G}$, the necessary parts of $\mathbf{G}^n$ are calculated by the forward and backward recursive algorithms. In the forward propagation, the boundary parts of $\mathbf{G}^{n0}$ are calculated, and on- and off-diagonal blocks of $\mathbf{G}^n$ are obtained by the backward recursion.

We define $\mathbf{R}^n$ as the boundary projection of $\mathbf{G}^{n0}$. Then, from Eq. (28), the R matrix $\mathbf{R}^n_{cc}$ for the boundary atoms of C can be defined as

$$\mathbf{R}^n_{cc} = \mathbf{R}_{cc}\Sigma^{\mathrm{in}}_{cc}\mathbf{R}^\dagger_{cc}. \quad (29)$$

Choosing a from B and adding it into C, we have

$$\begin{pmatrix}\tilde{\mathbf{R}}^n_{cc} & \tilde{\mathbf{R}}^n_{ca}\\ \tilde{\mathbf{R}}^n_{ac} & \tilde{\mathbf{R}}^n_{aa}\end{pmatrix} = \begin{pmatrix}\tilde{\mathbf{R}}_{cc} & \tilde{\mathbf{R}}_{ca}\\ \tilde{\mathbf{R}}_{ac} & \tilde{\mathbf{R}}_{aa}\end{pmatrix}\begin{pmatrix}\Sigma^{\mathrm{in}}_{cc} & \Sigma^{\mathrm{in}}_{ca}\\ \Sigma^{\mathrm{in}}_{ac} & \Sigma^{\mathrm{in}}_{aa}\end{pmatrix}\begin{pmatrix}\tilde{\mathbf{R}}^\dagger_{cc} & \tilde{\mathbf{R}}^\dagger_{ca}\\ \tilde{\mathbf{R}}^\dagger_{ac} & \tilde{\mathbf{R}}^\dagger_{aa}\end{pmatrix}. \quad (30)$$

Eq. (30) is associated with $\mathbf{R}^n_{cc}$ via Eq. (8). By performing the boundary projection from $c \cup a$ onto $\tilde{c}$ for Eq. (30), we can compute the matrix $\mathbf{R}^n_{\tilde{c}\tilde{c}}$, which is the new $\mathbf{R}^n_{cc}$ for the next forward calculation.

Using Eqs. (26) and (27), the necessary parts of $\mathbf{G}^n$ can be calculated in a similar manner to that for $\mathbf{G}$, and they are

$$\mathbf{G}^n_{aa} = \tilde{\mathbf{R}}^n_{aa} + \left(\Theta^n_{a\tilde{b}} + \Omega_{a\tilde{b}}\right)\mathbf{G}^\dagger_{\tilde{b}a} + \Theta_{a\tilde{b}}\mathbf{G}^n_{\tilde{b}a},$$
$$\mathbf{G}^n_{a\tilde{b}} = \left(\Theta^n_{a\tilde{b}} + \Omega_{a\tilde{b}}\right)\mathbf{G}^\dagger_{\tilde{b}\tilde{b}} + \Theta_{a\tilde{b}}\mathbf{G}^n_{\tilde{b}\tilde{b}},$$
$$\mathbf{G}^n_{\tilde{b}a} = \left(\mathbf{G}^n_{a\tilde{b}}\right)^\dagger, \quad (31)$$

where

$$\Theta^n_{a\tilde{b}} = \tilde{\mathbf{R}}^n_{ac}\left(-\mathbf{A}^\dagger_{c\tilde{b}}\right) + \tilde{\mathbf{R}}^n_{aa}\left(-\mathbf{A}^\dagger_{a\tilde{b}}\right),$$
$$\Omega_{a\tilde{b}} = \tilde{\mathbf{R}}_{ac}\Sigma^{\text{in}}_{c\tilde{b}} + \tilde{\mathbf{R}}_{aa}\Sigma^{\text{in}}_{a\tilde{b}}. \quad (32)$$

The initial $\mathbf{G}^n_{\tilde{b}\tilde{b}}$ is equal to the final $\mathbf{R}^n_{cc}$ of the forward propagation, and the matrix $\mathbf{G}^n_{bb}$ ($\mathbf{G}^n_{\tilde{b}\tilde{b}}$ for the next iteration) can be obtained via

$$\mathbf{G}^n_{bb} = P_b\begin{pmatrix}\mathbf{G}^n_{aa} & \mathbf{G}^n_{a\tilde{b}} \\ \mathbf{G}^n_{\tilde{b}a} & \mathbf{G}^n_{\tilde{b}\tilde{b}}\end{pmatrix}P_b. \quad (33)$$

The hole density is obtained from the hole correlation function $\mathbf{G}^p$, using the relation

$$\mathbf{G}^p = i\left(\mathbf{G} - \mathbf{G}^\dagger\right) - \mathbf{G}^n, \quad (34)$$

and the net drain current (J) is calculated by

$$J = \frac{g_s q_0}{h}\int d\varepsilon\ \text{Tr}\Big\{i\left(\mathbf{G}_{\beta\beta}(\varepsilon) - \mathbf{G}^\dagger_{\beta\beta}(\varepsilon)\right)\Sigma^{\text{in}}_L(\varepsilon) - \left(\mathbf{G}^n_{\beta\beta}(\varepsilon)\Gamma_L(\varepsilon)\right)\Big\}, \quad (35)$$

where q_0 is the elementary charge, h is the Planck constant, the subscript L indicates the left lead of the system, and the subscript β is the set of all the atoms coupled with the left lead. Eq. (35) is consistent with [16, eq. (154)].

E. RGF+R-MATRIX METHOD

A heterostructure FET can be divided into two parts: junction regions and homogeneous region consisting of identical unit cells. The junction regions play a critical role in determining the characteristics of the device while the homogeneous unit cells occupy most of the device spatially. The R-matrix method is numerically preferable for treating large-sized junction cells because an atom cluster can be constructed by adding a few number of atoms and only boundary atoms of the cluster contribute to calculating the related R matrices. However, for homogeneous unit cells usually represented by small-sized Hamiltonians, taking partial atoms for cluster construction is inefficient because it may increase the number of matrix operations. So, for the homogeneous unit cells, the RGF method, which is a well-established method for layered structures, is preferable.

In order to sketch how the RGF and R-matrix methods are combined, let us consider a homogeneous device that has structurally identical N_q unit cells, as shown in Fig. 3(a). We assume that all the atoms in each unit cell are coupled with the ones in the nearest unit cells, and we consider the unit cells as the slabs for RGF. Because a unit cell can be considered as an atom group, we can choose all the atoms in the slab q for the atom cluster a. Then, c and $\tilde{b}$ are the slabs $p = q - 1$ and $q + 1$, respectively (see Fig. 3(a)), and they are not coupled each other. So the boundary atoms of $c \cup a$ are equal to a, satisfying $\mathbf{R}_{\tilde{c}\tilde{c}} = \tilde{\mathbf{R}}_{aa}$. Now using $\tilde{\mathbf{R}}_{aa}$ in Eq. (8), we can write $\mathbf{R}_{\tilde{c}\tilde{c}}$ as

$$\mathbf{R}_{\tilde{c}\tilde{c}} = \tilde{\mathbf{R}}_{aa} = (\mathbf{A}_{aa} - \mathbf{A}_{ac}\mathbf{R}_{cc}\mathbf{A}_{ca})^{-1}, \quad (36)$$

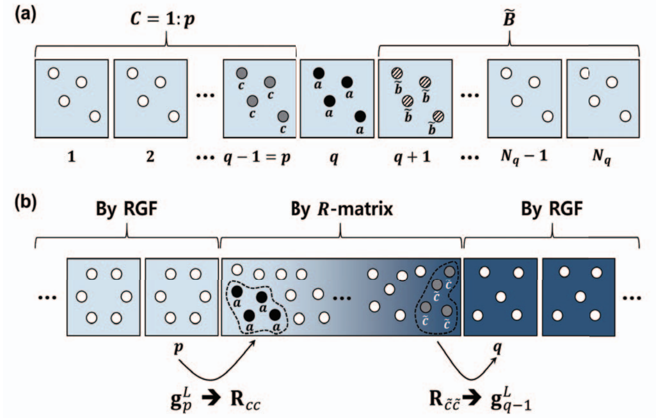

FIGURE 3. (a) Homogeneous device constructed by repeating structurally identical N_q unit cells. Each unit cell enclosed by a box is considered as a slab. (b) A schematic of the RGF+R-matrix method showing the block of cells where RGF and R-matrix algorithms are applied: the former applied to homogeneous unit cells and the latter to junction cell.

or using the corresponding slab indices,

$$\mathbf{R}_q = \tilde{\mathbf{R}}_q = \left(\mathbf{A}_{q,q} - \mathbf{A}_{q,p}\mathbf{R}_p\mathbf{A}_{p,q}\right)^{-1}. \quad (37)$$

Here $\mathbf{R}_p$ has the meaning of the boundary parts of the consecutive slabs $1{:}p$, that is

$$\mathbf{R}_p = P_p\mathbf{A}^{-1}_{1:p,1:p}P_p. \quad (38)$$

One can see that Eq. (38) is the same as the definition of the left connected Green's function ($\mathbf{g}^L_p$) in the RGF [14]–[16]. Now we can write Eq. (37) as

$$\mathbf{g}^L_q = \left(\mathbf{A}_{q,q} - \mathbf{A}_{q,q-1}\mathbf{g}^L_{q-1}\mathbf{A}_{q-1,q}\right)^{-1}. \quad (39)$$

For a device region with open boundary conditions, the R matrices play the same role as the left-connected Green's functions in the RGF algorithm. By utilizing the relation between $\mathbf{R}_{\tilde{c}\tilde{c}}$ and $\mathbf{g}^L_q$, we can combine the R-matrix method and the RGF method, as can be seen in Fig. 3(b). The RGF and R-matrix methods are applied to the homogeneous slabs and the junction region, respectively. For slab p at the left of the junction, $\mathbf{g}^L_p$ is calculated via Eq. (39) and used as $\mathbf{R}_{cc}$ in Eq. (36) for the junction. For the final atom cluster of the junction, $\mathbf{R}_{\tilde{c}\tilde{c}}$ is calculated using Eq. (36) and it is treated as $\mathbf{g}^L_{q-1}$ to calculate $\mathbf{g}^L_q$ of slab q on the right of the junction. The backward recursion can be carried out in a similar manner. When the process changes from RGF to R-matrix, $\mathbf{G}_{q,q}$ is considered as $\mathbf{G}_{bb}$, and for the reverse step, $\mathbf{G}_{\tilde{b}\tilde{b}}$ is used for the calculation of $\mathbf{G}_{q,q}$.

III. RESULTS AND DISCUSSION

A. DFT SIMULATIONS AND HAMILTONIAN EXTRACTION

In this work, the SIESTA [26] code, a LCAO-based DFT tool, was utilized to model heterostructures in the atomistic level. The generalized gradient approximation of Perdew-Burke-Ernzerhof [27] was adopted for the exchange-correlation functional, and double-ζ plus polarization pseudoatomatic orbitals were employed for the basis set. To

JOURNAL OF THE
ELECTRON DEVICES SOCIETY

TABLE 1. Details of the simulated heterostructures where two junctions J1 and J2 are included. a_0 is the lattice constant of the relaxed primitive unit cells.

Structure	Type	a_0 [Å]			Hamiltonian size				
		A	B	C	A	B	C	J1	J2
MoS_2 1T/2H/1T	A/B/A	3.16	3.19	-	164	164	-	1312	1312
Silicene/Silicane/Silicene	A/B/A	3.86	3.88	-	104	144	-	992	992
$MoTe_2/SnS_2$ VTFET	A/C/B	3.57	3.72	3.65	164	156	320	1792	1768
Germanane/InSe VTFET	A/C/B	4.08	4.09	4.09	144	272	416	2116	2490

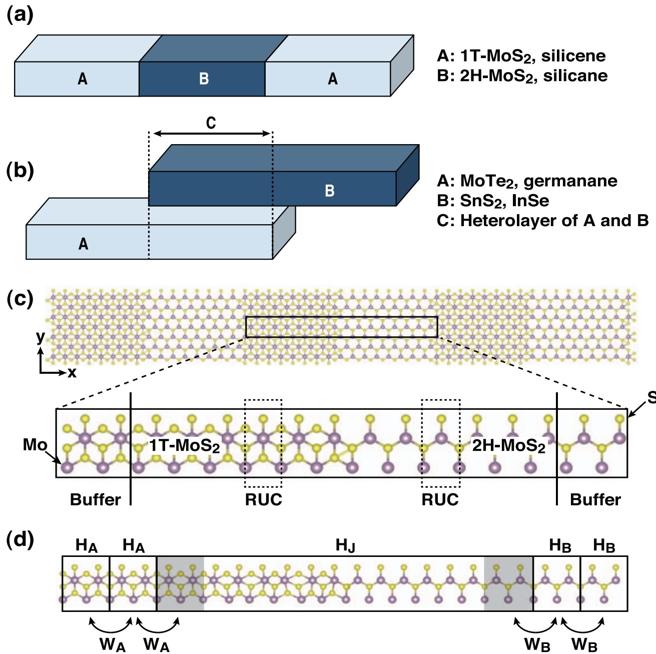

FIGURE 4. A schematic of the two-dimensional heterostructures considered: (a) A/B/A lateral heterojunction and (b) vertical TFET structure, which can be considered as A/C/B lateral heterojunction. (c) A schematic of the MoS_2 1T/2H junction unit cell that is used in the DFT relaxation, which is repeated along the x- and y-directions. (d) The MoS_2 1T/2H heterostructure for Hamiltonian extraction. The shaded regions are the extra RUCs inserted so that the coupling Hamiltonians between the junction and the component materials are the same as $\mathbf{W}_{A/B}$.

demonstrate the efficiency of the RGF+R-matrix algorithm, we consider four heterostructures of two-dimensional (2D) materials: MoS_2 1T/2H/1T and silicene/silicane/silicene as the lateral heterojunctions (Fig. 4(a)), and $MoTe_2/SnS_2$ and germanane/InSe for the 2D VTFET structures (Fig. 4(b)).

For each material, the atomic positions in the primitive unit cell were relaxed first, until the maximum force became less than 0.05 eV/Å with $25 \times 25 \times 1$ Monkhorst-Pack k-grids. Using the structurally optimized primitive cells, we then made the heterojunctions. As an example, let us consider MoS_2 1T/2H heterojunction illustrated in Fig. 4(c). First, for each material, the relaxed primitive cells were taken as the rectangular unit cells (RUCs), and with the RUCs, the junction supercell was constructed. To suppress the coupling between the periodic images, two buffer cells were attached to the left and right so that the junction supercell has total sixteen RUCs (see Fig. 4(c)), and the geometry optimization was conducted with the junction supercell. The periodic boundary conditions were imposed in the x- and y-directions while 10 nm vacuum layer were included in the z-direction. Other considered heterostructures were modeled in a similar way, treating the structure of Fig. 4(b) as a lateral junction of A/C/B [28].

For NEGF simulations, we extracted the necessary Hamiltonian blocks from the supercell consisting of the relaxed junctions and additional RUCs. For example, consider MoS_2 1T/2H heterostructure in Fig. 4(d). $\mathbf{H}_{A/B}$ and $\mathbf{W}_{A/B}$ are the Hamiltonians of the RUC A/B and between the nearest neighbor RUCs, respectively. $\mathbf{H}_J$ is the on-site Hamiltonian of the junction. An extra RUC (shaded in Fig. 4(d)) was inserted between the junction and the left (right) region so that the coupling Hamiltonian between them was the same as $\mathbf{W}_A$ ($\mathbf{W_B}$), as in [29]. The overlap matrices were obtained in the same manner. The relaxed lattice parameters of the component materials and the extracted Hamiltonian sizes are summarized on Table 1.

B. COMPUTATIONAL EFFICIENCY

To assess the numerical efficiency of the RGF+R-matrix method, we have measured the elapsed CPU time per energy point for calculating the retarded Green's functions. Three algorithms, RGF-only, R-matrix-only, and RGF+R-matrix are compared. The computation was conducted using 20 cores, and the elapsed CPU time values for 20 cores were averaged. In case of the R-matrix-only and RGF+R-matrix methods, the number of picked atoms (size of the atom cluster a) were varied. As shown in Fig. 5, for all the considered structures, R-matrix-only and RGF+R-matrix exhibit the better performance than RGF-only. Applying the R-matrix algorithm to junction parts, the speed-up of from 1.7 to 4.3 times (41% to 77% reduction of elapsed CPU time) has been achieved, compared with the results of RGF-only. The RGF+R-matrix further improves the efficiency. This is because the RGF method is more efficient than the R-matrix method for handling the homogeneous unit cells of small-sized Hamiltonians as mentioned in the first paragraph of Section II-E.

To verify whether the retarded Green's functions are computed accurately by the RGF+R-matrix method, we have calculated the transmission function using 1000 energy points. As can be seen in Fig. 6, for all the simulated heterostructures, transmission functions of RGF+R-matrix

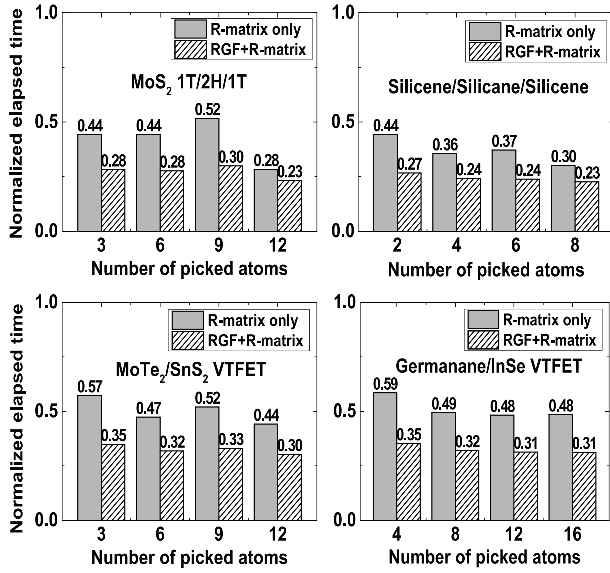

FIGURE 5. The elapsed CPU time per energy point for calculating the retarded Green's functions of the simulated heterostructures. The elapsed time by RGF-only is set to the reference value of 1.0.

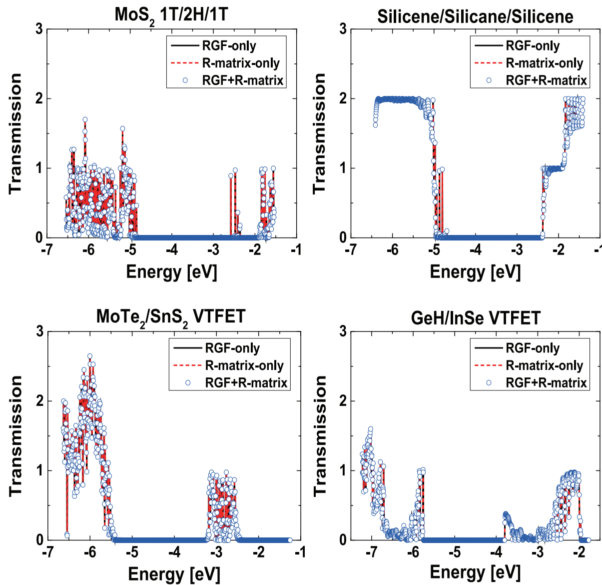

FIGURE 6. The transmission functions of the simulated heterostructures.

excellently match those of RGF-only. For the used 1000 energy points, the maximum difference of the calculated transmission was about 10^{-5}. Note that the three methods of RGF, R-matrix, and RGF+R-matrix are the exact methods, thus mathematically, the final computation results by the three methods should be the same. This is the reason that we have chosen the widely-used RGF algorithm as the reference to verify the proposed RGF+R-matrix algorithm. However, the three methods are also recursive algorithms, so the round-off errors may accumulate during recursive iterations, leading to the different results numerically.

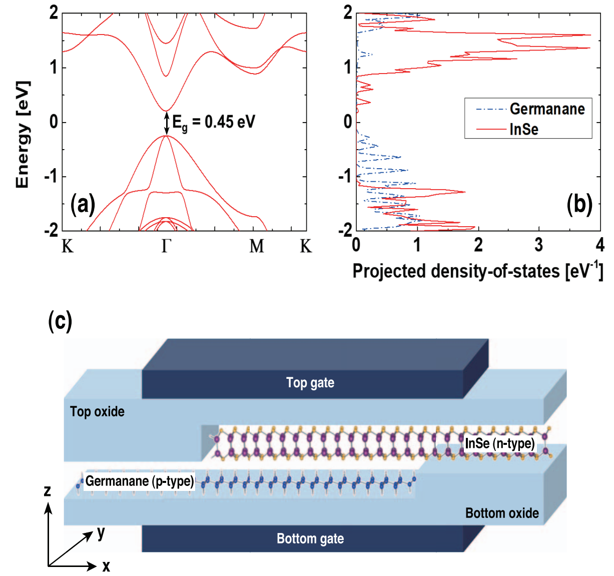

FIGURE 7. (a) Band structure and (b) projected density-of-states of germanane/InSe heterolayer primitive unit cell. Fermi-level is 0 eV. (c) Schematic of the simulated germanane/InSe VTFET.

C. SELF-CONSISTENT DEVICE SIMULATION

As an application, we have performed a device simulation of germanane/InSe 2D VTFET using the RGF+R-matrix method. As can be seen in Figs. 7(a) and (b), germanane/InSe heterolayer has the direct band gap of 0.45 eV and the atomic orbitals of InSe and germanane contribute mainly to the conduction and valence bands, respectively, which implies that the band-to-band tunneling may occur between germanane and InSe. The schematic of the simulated device is illustrated in Fig. 7(c). The length of the overlap region (germanane/InSe heterolayer) is 25 nm and the gate length is 85 nm; the length of gate overlap for each source/drain is 30 nm. The source and drain lengths are 40 nm. Equivalent oxide thickness is 0.5 nm, and the additional thickness of the oxide at the top (bottom) of the source (drain) region is 0.54 nm (0.39 nm), which is equal to the atomic thickness of InSe (germanane) monolayer. The device is periodic in the y-direction, which is accounted for 20 transverse modes in this work. Source (germanane) and drain (InSe) regions are doped to be p-type and n-type, respectively, with the density of 10^{20} cm^{-3}. A low drain voltage V_{DD} of 0.4 V was applied. The threshold voltage (V_{th}) was defined such that OFF-state current was 10^{-6} μA/μm, and we denote V_{GT} as $V_{TG} - V_{th}$, where V_{TG} is the top gate voltage. The bottom gate voltage (V_{BG}) was set to -0.8 V to ensure the TFET operation within the V_{GT} window of 0 V and 0.4 V. The size of the extracted Hamiltonian and overlap matrices for the homogeneous unit cells was reduced by a mode space transformation [9] for more efficient implementation. Ballistic transport was assumed. The NEGF-Poisson self-consistent calculations were carried out until convergence. The charge density was calculated by the algorithm described in Section II-D and the current values (I_D) were computed

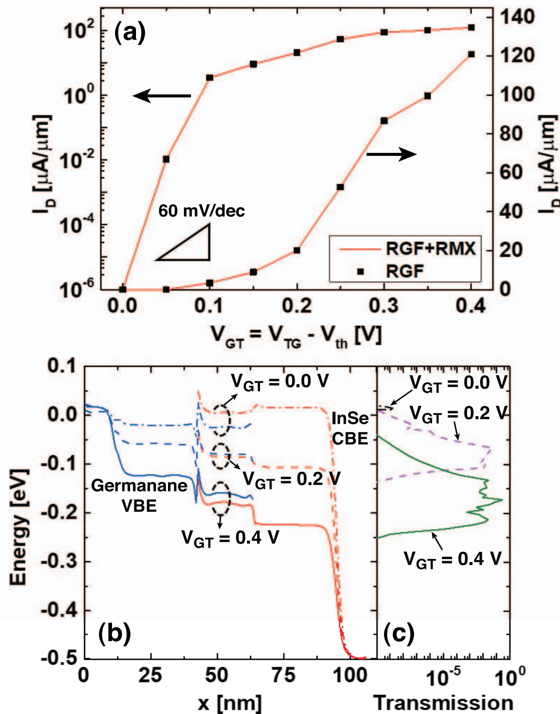

FIGURE 8. (a) The transfer characteristics of germanane/InSe VTFET at $V_{DD} = 0.4$ V and $V_{BG} = -0.8$ V. (b) Band edge profile and (c) transmission (log scale) of the simulated VTFET at various top gate voltage values.

by both the methods of RGF-only and RGF+R-matrix. All the simulations were conducted assuming the temperature of 300 K.

It can be seen in Fig. 8(a) that the current values computed by RGF-only and RGF+R-matrix agree excellently. The simulated transistor has the ON-state current of 121 μA/μm, and the minimum subthreshold swing of the device is 12 mV/dec, offering low power applications. One can see the operation mechanism of the simulated VTFET in Figs. 8(b) and (c). At $V_{GT} = 0$ V, the conduction band edge (CBE) of InSe is higher than the valence band edge (VBE) of germanane, so ultra-low tunneling current flows in the device. As the V_{GT} increases, the CBE of InSe moves down in energy faster than the VBE of germanane, until at on-state, CBE of InSe becomes located below VBE of germanane. Then, the electrons in the valence band of germanane move easily to the conduction band of InSe, giving rise to a large current (see Fig. 8(c)). Note that, in the heterostructure TFETs, the resonances by the quasi-bound states in junction region can contribute to the total current when scattering effects are included [30]. So, the current values in Fig. 8(a) may be underestimated, since ballistic transport was assumed in this work.

IV. CONCLUSION

We have presented a method for computationally efficient NEGF simulations of the heterostructure FETs including large-sized junctions modeled in the atomistic level. The main idea is to employ the RGF and R-matrix algorithms for the small-sized homogeneous unit cells and the large-sized junctions, respectively. For this purpose, the R-matrix method has been extended so that the algorithm applies to open systems directly. Compared with the usual RGF-only method, the Green's functions are calculated more efficiently using the proposed RGF+R-matrix method. As an application, the device simulations on germanane/InSe 2D VTFET have been demonstrated successfully. In this work, although we have considered the heterostructures modeled by DFT, the RGF+R-matrix method can be applied to the heterostructures modeled using other approaches such as tight-binding method and $k{\cdot}p$ method. In addition to the 2D heterostructure FETs that were selected for example devices in this work, the proposed method can be useful for the devices having heterojunctions of three-dimensional materials and also for the case where irregularities such as surface roughness, atom vacancies, and dopants are present.

REFERENCES

[1] U. E. Avci, D. H. Morris, and I. A. Young, "Tunnel field-effect transistors: Prospects and challenges," *IEEE J. Electron Devices Soc.*, vol. 3, no. 3, pp. 88–95, May 2015.
[2] D. Sarkar *et al.*, "A subthermionic tunnel field-effect transistor with an atomically thin channel," vol. 526, no. 7571, pp. 91–95, Oct. 2015.
[3] E. Lind, E. Memišević, A. W. Dey, and L.-E. Wernersson, "III–V heterostructure nanowire tunnel FETs," *IEEE J. Electron Devices Soc.*, vol. 3, no. 3, pp. 96–102, May 2015.
[4] S.-C. Lu, M. Mohamed, and W. Zhu, "Novel vertical hetero- and homo-junction tunnel field-effect transistors based on multi-layer 2D crystals," *2D Mater.*, vol. 3, no. 1, pp. 1–6, Feb. 2016.
[5] J. Kang, W. Liu, and K. Banerjee, "High-performance MoS_2 transistors with low-resistance molybdenum contacts," *Appl. Phys. Lett.*, vol. 104, no. 9, pp. 1–5, Mar. 2014.
[6] H. Tian *et al.*, "Novel field-effect schottky barrier transistors based on graphene-MoS_2 heterojunctions," *Sci. Rep.*, vol. 4, no. 5951, pp. 1–9, Aug. 2014.
[7] S. Brück, M. Calderara, M. H. Bani-Hashemian, J. VandeVondele, and M. Luisier, "Efficient algorithms for large-scale quantum transport calculations," *J. Chem. Phys.*, vol. 147, no. 7, pp. 1–13, Aug. 2017.
[8] Á. Szabó, R. Rhyner, and M. Luisier, "Ab initio simulation of single- and few-layer MoS_2 transistors: Effect of electron-phonon scattering," *Phys. Rev. B, Condens. Matter*, vol. 92, no. 3, pp. 1–10, Jul. 2015.
[9] 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, pp. 1–10, Apr. 2016.
[10] M. Shin, H.-E. Jung, and S. Jung, "First-principles based quantum transport simulations of nanoscale field effect transistors," in *Proc. Int. Electron Device Meeting*, Dec. 2017, p. 352.
[11] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, "Iterative minimization techniques for ab initio total-energy calculations: Molecular dynamics and conjugate gradients," *Rev. Mod. Phys.*, vol. 64, no. 4, pp. 1045–1097, Oct. 1992.
[12] A. Blom and K. Stokbro, "Atomistic modeling of semiconductor interfaces," *J. Comput. Electron.*, vol. 12, no. 4, pp. 623–637, Dec. 2013.
[13] L. Lin, Y. Guo, R. Gillen, and J. Robertson, "Chemical trends of defects at HfO_2:GaAs and Al_2O_3:GaAs/InAs/InP/GaSb interfaces," *J. Appl. Phys.*, vol. 113, no. 13, pp. 1–9, Apr. 2013.
[14] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *J. Appl. Phys.*, vol. 81, no. 12, pp. 7845–7869, Jun. 1997.
[15] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal, "Two-dimensional quantum mechanical modeling of nanotransistors," *J. Appl. Phys.*, vol. 91, no. 4, pp. 2343–2354, Feb. 2002.

[16] 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.

[17] S. Cauley, M. Luisier, V. Balakrishnan, G. Klimeck, and C.-K. Koh, "Distributed non-equilibrium Green's function algorithms for the simulation of nanoelectronic devices with scattering," *J. Appl. Phys.*, vol. 110, no. 4, pp. 1–14, Aug. 2011.

[18] S. Cauley, V. Balakrishnan, G. Klimeck, and C.-K. Koh, "A two-dimensional domain decomposition technique for the simulation of quantum-scale devices," *J. Comput. Phys.*, vol. 231, no. 4, pp. 1293–1313, Feb. 2012.

[19] L. Lin, J. Lu, L. Ying, R. Car, and Weinan E, "Fast algorithm for extracting the diagonal of the inverse matrix with application to the electronic structure analysis of metallic systems," *Commun. Math. Sci.*, vol. 7, no. 3, pp. 755–777, 2009.

[20] U. Hetmaniuk, Y. Zhao, and M. P. Anantram, "A nested dissection approach to modeling transport in nanodevices: Algorithms and applications," *Int. J. Numer. Methods Eng.*, vol. 95, no. 7, pp. 587–607, Aug. 2013.

[21] S. Li, S. Ahmed, G. Klimeck, and E. Darve, "Computing entries of the inverse of a sparse matrix using the FIND algorithm," *J. Comput. Phys.*, vol. 227, no. 22, pp. 9408–9427, Nov. 2008.

[22] G. Mil'nikov, N. Mori, Y. Kamakura, and T. Ezaki, "*R*-matrix theory of quantum transport and recursive propagation method for device simulations," *J. Appl. Phys.*, vol. 104, no. 4, pp. 1–14, Aug. 2008.

[23] G. Mil'nikov, N. Mori, and Y. Kamakura, "*R*-matrix method for quantum transport simulations in discrete systems," *Phys. Rev. B, Condens. Matter*, vol. 79, no. 23, pp. 1–5, Jun. 2009.

[24] D. Mamaluy, A. Mannargudi, D. Vasileska, M. Sabathil, and P. Vogl, "Contact block reduction method and its application to a 10 nm MOSFET device," *Semicond. Sci. Technol.*, vol. 19, no. 4, pp. S118–S121, Apr. 2004.

[25] D. Mamaluy, D. Vasileska, M. Sabathil, T. Zibold, and P. Vogl, "Contact block reduction method for ballistic transport and carrier densities of open nanostructures," *Phys. Rev. B, Condens. Matter*, vol. 71, no. 24, pp. 1–14, Jun. 2005.

[26] J. M. Soler *et al.*, "The SIESTA method for ab initio order-*N* materials simulation," *J. Phys. Condens. Matter*, vol. 14, no. 11, pp. 2745–2779, Mar. 2002.

[27] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 77, no. 18, pp. 3865–3868, Oct. 1996.

[28] J. Zhang, W. Xie, J. Zhao, and S. Zhang, "Band alignment of two-dimensional lateral heterostructures," *2D Mater.*, vol. 4, no. 1, pp. 1–8, Mar. 2017.

[29] Á. Szabó, "Dissipative quantum transport simulations in two-dimensional semiconductor devices from first principles," Ph.D. dissertation, Eng. Phys., ETH Zürich, Zürich, Switzerland, 2016.

[30] T. A. Ameen, H. Ilatikhameneh, J. Z. Huang, M. Povolotskyi, R. Rahman, and G. Klimeck, "Combination of equilibrium and nonequilibrium carrier statistics into an atomistic quantum transport model for tunneling heterojunctions," *IEEE Trans. Electron Devices*, vol. 64, no. 6, pp. 2512–2018, Jun. 2017.