Ph.D. DISSERTATION

A Deterministic Solution of the Wigner Transport Equation with Infinite Correlation Length

무한한 상관길이를 가지는 위그너 수송 방정식의 결정론적 해

BY

KYOUNG YEON KIM

AUGUST 2022

DEPARTMENT OF ELECTRICAL AND

CMPUTER ENGINEERING

COLLEGE OF ENGINEERING

SEOUL NATIONAL UNIVERSITY

Deterministic Solution of the Wigner Transport Equation with Infinite Correlation Length

무한한 상관길이를 가지는 위그너 수송 방정식의 결정론적 해

지도교수 최 우 영
이 논문을 공학박사 학위논문으로 제출함
2022 년 8 월

서울대학교 대학원
전기정보공학부
김 경 연

김경연의 공학박사 학위논문을 인준함
2022 년 8 월

위 원 장 정덕균 (인)

부위원장 최우영 (인)

위 원 이종호 (인)

위 원 진성훈 (인)

위 원 홍성민 (인)

Abstract

We propose a new discrete formulation of the Wigner transport equation (WTE) with infinite correlation length of potentials. Since the maximum correlation length is not limited to a finite value, there is no uncertainty in the simulation results, and Wigner-Weyl transformation is unitary in our expression. For general and efficient simulation, the WTE is solved self-consistently with the Poisson equation through the finite volume method and the fully coupled Newton-Raphson scheme. By applying the proposed model to resonant tunneling diodes and double gate MOSFET, transient and steady-state simulation results including scattering effects are shown.

Keyword : Wigner equation, Quantum transport, TCAD

Student Number : 2016-20866

Contents

List of Figures

Chapter 1

Introduction

1-1. Various models for device simulation

TCAD simulation is divided into material/process simulation, device simulation, and circuit simulation. Among them, device simulation performs characteristic analysis such as IV curve or electrical data according to device geometry, doping profiles, and various device operation parameters. When the MOSFET gate length is sufficiently long, a continuum model such as the drift-diffusion equation has been mainly used for device characterization. However, as device scaling continues and the physical gate length is shortened to several nm, classical models are no longer valid and an atomistic model is needed [1]-[7]. Therefore, quantum transport models have been used for the analysis of nanoscale devices, and the Non-Equilibrium Green's Function (NEGF) method [8], Wigner transport equation (WTE) [9],[10], and Pauli mater equation

(PME) are representative quantum transport models [11],[12]. However, in the semiconductor industry, the drift-diffusion equation is still widely used for faster and simpler simulation [13]. Quantum transport simulation is difficult to apply practically because it requires a lot of computational power, so these sophisticated models are generally used to calibrate the mobility of the drift-diffusion equation. However, if the computing power gets better and the numerical efficiency of quantum transport models is improved, it will gradually be able to replace the classical models.

NEGF is a method of solving the Schrödinger equation considering the inflow/outflow boundary condition at the source/drain contact. Of these, the NEGF method is the most widely used, but it has difficulties in transient simulation, and there are limitations in considering various scattering mechanisms if local approximation is not used. In NEGF formalism, it is hard to consider the microscopic scattering mechanism because it requires the inversion of matrix of huge rank because the self-energy terms are generally nonlocal function [14]. Electron-phonon scattering can be efficiently calculated through local approximation, but large computational cost is required to include other scattering mechanisms [15]-[17]. Also, although this method is well defined in steady state, it is not suitable for transient simulation which is

very important in device characterization [18]. Recently, there is a study that conducted AC simulation in extremely scaled nanosheet MOSFETs using NEGF formalism considering first-order perturbation [19]. [20]. However, since this is a method based on small-signal analysis, only high-frequency AC simulation is possible, and general transient simulation is still difficult.

Recently, Pratik B. Vyas et al reported a simulation of the dissipative quantum transport through the Pauli master equation (PME) [21]. They show successful simulation results in an ultra-thin body double-gate FET based on the quantum transmitting boundary method (QTBM). This is an attractive model for efficiently handling the scattering mechanism, but it is also limited to a steady state solution and can be applied only when the perturbation is weak and the device length is sufficiently short. They show that dissipative electron transport is also important in nanoscale MOSFETs, so it is important to consider various scattering effects rather than ballistic ones.

As an alternative to the above two methods, we used the WTE for the simulation of quantum transport in this work [22]- [27]. Transient simulation and dissipative transport simulation are possible based on the WTE. Since WTE is a form that includes a differential term with respect to time, both transient simulation and

frequency domain analysis are freely possible. In this thesis, we propose a method to solve the problems that occur in solving WTE numerically, and apply it to one-dimensional and two-dimensional device simulation.

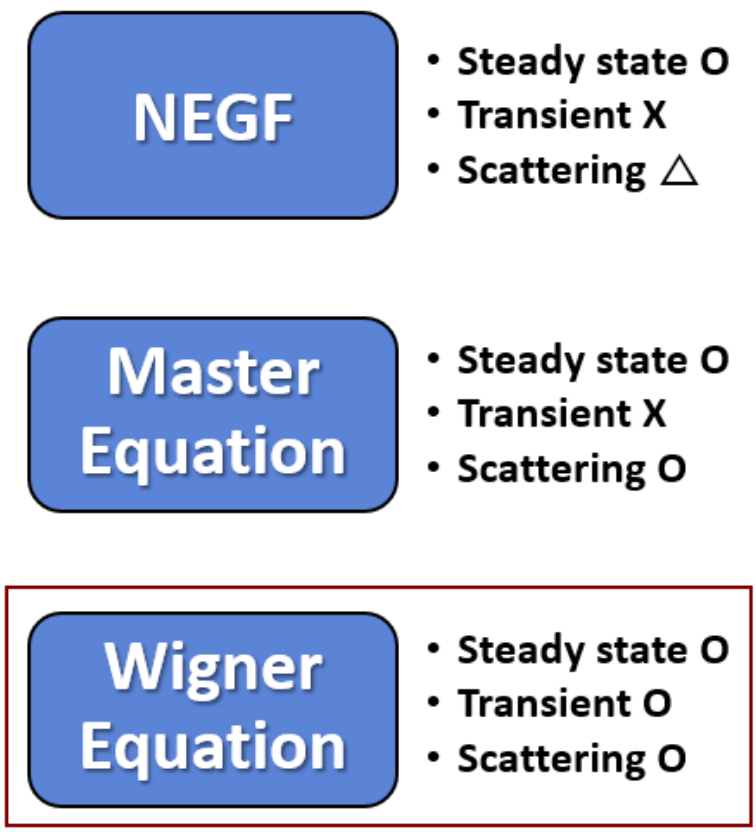


Fig. 1. Three methods for quantum transport simulation. WTE is capable of steady-state simulation and transient simulation, and can conveniently incorporate scattering mechanisms.

1-2. Numerical Problems in Solving WTE

The WTE is obtained through the coordinate transformation and the Fourier transform of the density matrix and quantum Liouville equation:

$$\rho(x,x') = \sum_i w_i < x \mid i >< i \mid x' >, \tag{1}$$

$$\frac{\partial \rho}{\partial t} = \frac{1}{i\hbar}\left[-\frac{\partial \varepsilon}{2\partial k}\left[\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial x'^2} \right]\rho + p[v(x) - v(x')]\rho \right], \tag{2}$$

where i is a complete set of states, w_i is a probability, and v is the potential energy. The Wigner distribution function is obtained through the following coordinate transformation and the Fourie transform:

$$\chi = \frac{1}{2}(x + x'), \quad \zeta = x - x' \tag{3}$$

$$f(\chi,k) = \int_{-\infty}^{\infty} d\zeta e^{-ik\zeta} \rho\left(\chi + \frac{1}{2}\zeta, \chi - \frac{1}{2}\zeta \right). \tag{4}$$

Then the Liouville equation of the Wigner distribution function can be expressed as follow:

$$-\frac{\partial f}{\partial t} = \frac{\partial \varepsilon}{\hbar \partial k}\frac{\partial f}{\partial \chi} + \frac{1}{\hbar}\int_{-\infty}^{\infty} \frac{dk'}{2\pi} \hat{V}(\chi, k - k') f(\chi, k'), \tag{5}$$

where the nonlocal potential term given by

$$\hat{V}(\chi,k)=\frac{1}{i}\int_{-\infty}^{\infty}d\zeta e^{i\zeta k}\left[U\left(\chi+\frac{1}{2}\zeta\right)-U\left(\chi-\frac{1}{2}\zeta\right)\right]. \qquad (6)$$

The WTE can be obtained by simply adding the collisional term to include the scattering mechanism [10]:

$$\left(\frac{\partial f}{\partial t}\right)_C-\frac{\partial f}{\partial t}=\frac{\partial\varepsilon}{\hbar\partial k}\frac{\partial f}{\partial\chi}+\frac{1}{\hbar}\int_{-\infty}^{\infty}\frac{dk'}{2\pi}\hat{V}(\chi,k-k')f(\chi,k'), \qquad (7)$$

At the semi-classical limit, the second term of the RHS of Eq. (7) is reduced to the classical force term, so that the well-known Boltzmann transport equation (BTE) can be obtained [1]-[4]. In the BTE, the calculation can be simplified by using energy-space grid and H-transformation instead of k-space [1], [2]. However, in the WTE, since non-locality in k-space cannot be eliminated in such a way, it requires more memory and longer computation time than the BTE. However, if various scattering mechanisms are included in BTE, nonlocality in energy space occurs. Therefore, the system matrix becomes much denser than in the ballistic case, and there is no significant difference in terms of numerical efficiency compared to WTE.

In general, to solve the WTE numerically, the integral range of the nonlocal potential term (Eq. (6)) is limited to a finite range. But in this way, some of the information in the density operator is lost,

and the Wigner-Weyl transformation is not unitary [10]. For example, if mesh spacing of $d\chi = dx$ and $d\zeta = 2dx$ is used, information on corner triangles of the density operator in x space is lost, and half of the information is lost among other components. Main and Haddad devised a method to use a mesh of $d\chi = 1/2dx$ and $d\zeta = 2x$, but the loss of corner triangles still cannot be recovered. There have been attempts to increase the maximum correlation length to obtain more reliable results. However, even if the maximum correlation length (integral range) is increased, the calculation result has uncertainty unless the integration range is infinite [28], [29]. A. S. Costolanski and C. T. Kelley confirmed that different simulation results were obtained depending on the correlation length, and mentioned that the appropriate correlation length is different depending on the device structure and there is no simple physics-based rule to determine it [29]. For a more accurate simulation, a high order differential scheme is also widely used, but this is a physically distinct problem from the maximum correlation length [30]. As such, the simulation based on the finite correlation length has a problem in that there is uncertainty in the simulation result depending on the correlation length and may not be physically consistent with the density operator. Therefore, we need a method to correctly calculate the nonlocal potential term

without uncertainty.

In this paper, we propose a new formulation with an infinite correlation length by assuming an ideal semi-infinite reservoir in the contact region. Through simple reconstruction of the nonlocal potential term, an equivalent equation with a finite integral range is derived. Since our new formula considers the integral range of nonlocal potential terms up to infinity, it can solve the problem of uncertainty of simulation results according to the finite correlation length. To avoid the statistical fluctuation of solution and for a more general-purpose simulator, we adopt the deterministic method with finite volume method (FVM) rather than the stochastic Monte Carlo method. We use the fully coupled Newton Raphson method to obtain excellent convergence. Through these, quantum transport steady-state and transient simulation with excellent convergence are successfully implemented. By applying our simulator to resonant tunneling diode (RTD), it was confirmed that reliable results are obtained by showing the plateau region and transient oscillation in unstable bias. A more practical structure, the double gate MOSFET, was also simulated through the mode space method, and the device characteristics for various conditions were confirmed. In addition, since the WTE has a similar shape with the BTE except for quantum mechanical term, transient simulation, small signal (noise)

analysis and including of scattering are also readily possible in a similar manner as in BTE [31].

Chapter 2

Simulation methods

2-1. WTE with infinite correlation length

The WTE can be expressed as follow [10]:

$$\left(\frac{\partial f}{\partial t}\right)_C - \frac{\partial f}{\partial t} = \frac{\partial \varepsilon}{\hbar \partial k}\frac{\partial f}{\partial \chi} + \frac{1}{\hbar}\int_{-\infty}^{\infty}\frac{dk'}{2\pi}\hat{V}(\chi, k-k')f(\chi,k'), \qquad (8)$$

where the first term of LHS is scattering integral, t is time, $\hbar$ is Dirac's constant, and ε is energy level. In general, to solve the WTE numerically, the integral range of the nonlocal potential term (Eq. (6)) is limited to a finite range.

$$\hat{V}(\chi,k) = \frac{1}{i}\int_{-L_C}^{L_C} d\zeta e^{i\zeta k}\left[U\left(\chi+\frac{1}{2}\zeta\right) - U\left(\chi-\frac{1}{2}\zeta\right)\right]. \qquad (9)$$

A method such as Eq. (9) causes simulation uncertainty according to L_C. Therefore, to consider the infinite correlation length, we first assume an ideal contact condition in which the reservoir is semi-infinitely long and has a constant potential energy. Such boundary

conditions are commonly used in quantum transport simulation. In a one-dimensional space, the contact is located at χ=0 and χ=L, and the potential energy at the right contact is higher than the left one by U_{ex} as Fig.2. We assume one-dimensional transport and derive the equation, but more complex boundary conditions will be needed to solve the nonlocal potential term in two or three dimensions.

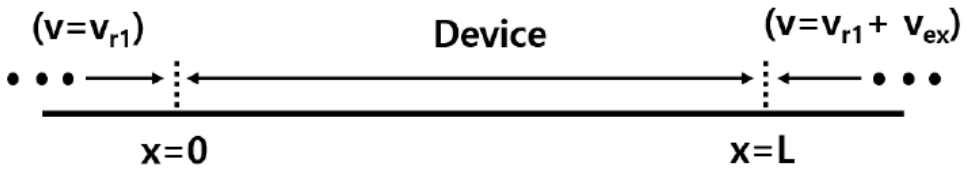


Fig.2. The contact is located at χ=0 and χ=L, and the potential energy at the right contact is higher than the left one by U_{ex}

To reconstruct the equation, the nonlocal potential term is divided into the sum of the two terms as follows:

$$\hat{V}(\chi,k)=\frac{1}{i}\int_{-\infty}^{\infty}d\zeta e^{i\zeta k}\left[U_{cor}(\chi,\zeta)-U_{ex}u(\zeta)+U_{ex}u(-\zeta)\right] + \frac{1}{i}\int_{-\infty}^{\infty}d\zeta e^{i\zeta k}\left[U_{ex}u(\zeta)-U_{ex}u(-\zeta)\right], \tag{10}$$

$$U_{cor}(x,\zeta)=U\left(\chi+\frac{1}{2}\zeta\right)-U\left(\chi-\frac{1}{2}\zeta\right), \tag{11}$$

where u is a unit step function. In this representation, we just add and subtract the product of U_{ex} and the unit step function to the

nonlocal potential term. The reason that the expression is divided into two terms as in Eq. (10) is that each term can be calculated analytically in this form. The integrand function of the first term in Eq. (10) becomes an odd function for ζ, and the second term can be calculated through the Fourier transform relational expression of the unit step function as follow:

$$u(t) \xleftrightarrow{F.T.} \frac{1}{jw} + \pi\delta(w), \tag{12}$$

Thus, Eq. (10) can be rewritten as:

$$\hat{V}(\chi,k) = 2\int_0^{\infty} d\zeta \sin(\zeta k)\left[U_{cor}(\chi,\zeta) - U_{ex}\right] + 2\frac{U_{ex}}{k}. \tag{13}$$

When χ is between 0 and L, the integrand of the first term is always 0 if ζ is greater than 2L because U_{cor} is equal to U_{ex}. Therefore, the integration range can be reduced to [0, 2L]:

$$\hat{V}(\chi,k) = 2\int_0^{2L} d\zeta \sin(\zeta k)\left[U_{cor}(\chi,\zeta) - U_{ex}\right] + 2\frac{U_{ex}}{k} \tag{14}$$

Since the integral range is finitely limited through reformulation of nonlocal potential terms, it is possible to solve WTE with infinite correlation length numerically. Now, if you separate an analytically integrable term, it can be expressed by the following formula.

$$\hat{V}(\chi,k) = 2\int_0^{2L} d\zeta \sin(\zeta k)\left[U_{cor}(\chi,\zeta)\right] + 2\frac{U_{ex}}{k}\cos(2Lk) \tag{15}$$

When the integral range is finitely limited in the nonlocal potential

term, it can be expressed by the following equation.

$$\hat{V}(\chi,k)=\frac{1}{i}\int_{-L_C}^{L_C} d\zeta e^{i\zeta k}\left[U\left(\chi+\frac{1}{2}\zeta\right)-U\left(\chi-\frac{1}{2}\zeta\right)\right]. \tag{16}$$

Here, if the integration range is set to [−2L 2L], it can be seen that this is most similar to Eq.(15), which is an expression with infinite correlation length. However, Eq.(15) is a form of adding the cosine function to Eq.(16), and we can see that this is clearly different from the case where the integration range is limited to 2L. Simulation results according to correlation length will be shown later in the results section.

2−2. Numerical Methods

To solve our new formulation numerically, we use the finite volume method (box integration method). In the steady state, WTE can be expressed as

$$\frac{\partial\varepsilon}{\hbar\partial k}\frac{\partial}{\partial\chi}f(\chi,k)+W_{\chi,k}-C_{\chi,k}=0, \tag{17}$$

where W is quantum evolution term (the second term in the RHS of Eq. (8)) and C is collisional term (the first term in the LHS of Eq. (8)). To apply the upwind scheme, the formula can be divided into two cases according to the direction of the group velocity:

$$\pm v_{\chi,k}\frac{\partial}{\partial \chi} f^{\pm}(\chi,k) + W^{\pm}{}_{\chi,k} - C^{\pm}{}_{\chi,k} = 0, \tag{18}$$

where v is group velocity and the + and − sign represents when the group velocity is positive and negative, respectively. For easy box integration, the equation is transformed as follows using partial differentiation:

$$\begin{aligned} &\pm\frac{\partial}{\partial \chi} v_{\chi,k} f^{\pm}(\chi,k) \mp f^{\pm}(\chi,k)\frac{\partial v_{\chi,k}}{\partial \chi} \\ &+W^{\pm}{}_{\chi,k} - C^{\pm}{}_{\chi,k} = 0. \end{aligned} \tag{19}$$

Through this, the first term can be modified in a form that can be more easily integrated in the control volume.

In χ space with a uniform mesh size, the box integration at node x_I can be obtained by integrating Eq. (19) from $x_{i-0.5}$ to $x_{i+0.5}$:

$$\begin{aligned} &\pm\left[v_{\chi,k} f^{\pm}(\chi,k) \right]_{x_{i-0.5}}^{x_{i+0.5}} \mp \int_{x_{i-0.5}}^{x_{i+0.5}} \partial\chi f^{\pm}(\chi,k)\frac{\partial v_{\chi,k}}{\partial \chi} \\ &+(W^{\pm}{}_{\chi,k} - C^{\pm}{}_{\chi,k})\Delta x = 0. \end{aligned} \tag{20}$$

To calculate the first term of Eq. (20), we need to know the distribution function at $x_{i-0.5}$ and $x_{i+0.5}$. The simplest way to do this is to use the average value of two adjacent nodes. However, in this work we use the Quadratic upstream interpolation for convective kinematics (QUICK) scheme for high numerical accuracy [32]. This method takes the second order derivative into account and has third

order accuracy. The value at the cell face can be calculated as follows through the QUICK scheme:

$$\left[vf^{+}\right]_{x_{i-0.5}} = \frac{3}{4}\left[vf^{+}\right]_{x_{i-1}} + \frac{3}{8}\left[vf^{+}\right]_{x_i} - \frac{1}{8}\left[vf^{+}\right]_{x_{i-2}}, \tag{21a}$$

$$\left[vf^{+}\right]_{x_{i+0.5}} = \frac{3}{4}\left[vf^{+}\right]_{x_i} + \frac{3}{8}\left[vf^{+}\right]_{x_{i+1}} - \frac{1}{8}\left[vf^{+}\right]_{x_{i-1}}, \tag{21b}$$

$$\left[vf^{-}\right]_{x_{i-0.5}} = \frac{3}{4}\left[vf^{-}\right]_{x_i} + \frac{3}{8}\left[vf^{-}\right]_{x_{i-1}} - \frac{1}{8}\left[vf^{-}\right]_{x_{i+1}}, \tag{21c}$$

$$\left[vf^{-}\right]_{x_{i+0.5}} = \frac{3}{4}\left[vf^{-}\right]_{x_{i+1}} + \frac{3}{8}\left[vf^{-}\right]_{x_i} - \frac{1}{8}\left[vf^{-}\right]_{x_{i+2}}. \tag{21d}$$

If the mesh nodes are outside of simulation domain, $vf^{\pm}$ is assumed to have the same value as in the boundary. Dirichlet boundary conditions apply only to the left if the group velocity is greater than 0, and only to the right if it is less than 0 as upwind method. After integration over a perpendicular wave vector, the boundary condition of the Wigner function in one dimension without transverse consideration is as follows:

$$f^{+}(0,k) = \frac{m^{*}k_B T}{\pi\hbar^2}\ln\left[1+\exp\left(-\frac{1}{k_B T}\left(\frac{\hbar^2 k^2}{2m^{*}} - \mu_0\right)\right)\right], \tag{22a}$$

$$f^{-}(L,k) = \frac{m^{*}k_B T}{\pi\hbar^2}\ln\left[1+\exp\left(-\frac{1}{k_B T}\left(\frac{\hbar^2 k^2}{2m^{*}} - \mu_L\right)\right)\right], \tag{22b}$$

where k_B is the Boltzmann constant, T is the temperature, and u_0 and u_L are the Fermi energies at the ends of the device.

The second term of Eq. (17) vanishes if the group velocity is constant. However, if the group velocity changes at any point for reasons such as partial varying effective mass, the second term should be calculated. If the group velocity at x_i abruptly changes from A to B, the second term can be written as

$$\begin{aligned} &\mp\int_{x_{i-0.5}}^{x_{i+0.5}} \partial\chi f^{\pm}(\chi,k)(B-A)\delta(\chi-\chi_i) \\ &= \mp(B-A) f^{\pm}(x_i,k), \end{aligned} \tag{23}$$

where δ is the Dirac delta function which is the derivative of the unit step function. For example, assuming a parabolic band, if the effective mass at x_i changes from m_1 to m_2, Eq. (23) becomes

$$\mp\left(\frac{\hbar k}{m_2}-\frac{\hbar k}{m_1}\right) f^{\pm}(x_i,k). \tag{24}$$

Before describing the quantum evolution term, mesh spacing in k-space should be considered. When Eq. (17) is integrated over k-space, the equation becomes a continuity equation for charge density. In order to satisfy the charge conservation, the integral of the quantum evolution term must also be 0. If a uniform mesh size is used, the integral can be expressed discretely as

$$\int W^{\pm}{}_{\chi,k}\,dk = \sum_i W^{\pm}{}_{\chi,k_i}\Delta k = 0. \tag{25}$$

The above equation holds when the Fourier completeness

relation is satisfied, and the mesh size at that time can be expressed as follow [10]:

$$\Delta k = \frac{\pi}{N_k \Delta x}, \tag{26}$$

where N_k is the number of meshes in k−space. And the k−space is discretized as follow:

$$k_l = \frac{\pi}{\Delta x}\left\{\frac{(l+0.5)}{N_k} - 0.5\right\}, l = 0,1,\ldots,N_{k-1} \tag{27}$$

For calculating the quantum evolution term, we first need to calculate the nonlocal potential term. To calculate the first term of RHS, we assume that the potential changes linearly between adjacent mesh points. In this way, the nonlocal potential is calculated through direct integration rather than discrete integration in order to accurately account for changes in the sine function by position. Since a linear potential is assumed, the equation can be expressed in the form of a product of a sine function and a linear function, so that analytical calculations are possible. Also, even after integration within the mesh, the nonlocal potential consists of sin and cos functions with the same period as before integration, so the Fourier completeness relation is still satisfied. A more precise integration into k−space is not considered. It was difficult to analytically integrate the equation integrated in the real−space

again in the k−space, and the accuracy of the simulation is increased by using a mesh size in the k−space that is small enough. If more accurate integration is possible for k−space, more accurate and efficient simulation will be possible. The detailed calculation method for this will be shown in the appendix.

Eventually, the quantum evolution term is calculated as follow:

$$W^{\pm}{}_{\chi,k} = \frac{\Delta k}{2\pi\hbar}\sum_i \hat{V}(\chi, k-k_i) f^{\pm}(\chi, k_i) \tag{28}$$

To simply include the scattering effect we use the relaxation time approximation [1]:

$$C^{\pm}{}_{\chi,k} \approx -\frac{1}{\tau}\left[f^{\pm}(\chi,k) - \frac{f_{eq}(\chi,k)}{\int dk'\, f_{eq}(\chi,k')}\int dk'\, f^{\pm}(\chi,k') \right] \tag{29}$$

Where τ is the relaxation time and f_{eq} is the local equilibrium distribution function. The equilibrium distribution function uses the solution when the applied bias is zero. In solving this problem, if the integral of the Wigner function is considered as a known value and non−locality between k−space in the Jacobian matrix is not considered, the simulation does not converge well and the current continuity is broken. Therefore, when constructing the system matrix for this, an accurate simulation was carried out considering all the interactions between k−space. When the scattering effect is considered through relaxation time approximation, inaccurate or

unphysical results may be obtained [33], [34], so more complex nonlocal scattering models will be needed for more accurate simulation [34].

Electrostatic potential V_p is obtained through the Poisson equation as

$$\frac{d^2V_p}{dx^2} = -\frac{q}{\varepsilon}\left[N_d(x) - n(x)\right], \tag{30}$$

where N_d is the doping concentration and $n(x)$ is the electron density obtained from the Wigner function:

$$n(x) = \frac{1}{2\pi}\int f(x,k)dk. \tag{31}$$

The potential energy U used in the nonlocal potential term can be calculated as follow:

$$U(x) = -qV_p(x) + U_c(x), \tag{32}$$

where U_c is the band structure function which considers the band offset considering the barriers and wells.

An iterative solver is required because WTE is a nonlinear system that needs to be solved together with the Poisson equation. We simulate with two methods: the Gummel method [35], which is mainly used as a decoupled scheme, and the Newton Raphson method, which is a fully coupled scheme. Through the Newton Raphson method, quadratic convergence is obtained, and since an

accurate response function can be obtained, small signal analysis is also readily possible. In the Newton-Raphson method, most of the computational time is required to update the Jacobian matrix every iteration. Therefore, we use the Newton-Richardson method, which only updates the Jacobian matrix when the convergence is poor. On the other hand, in the case of the Gummel iteration method, since the computational cost for each iteration is small, it is more useful when the system matrix size is too large, such as in 2D or 3D simulations. Also, since the simulation converges well even if the initial value is relatively far from the solution in the Gummel method, it can be used to obtain the approximated initial condition before using the Newton-Raphson method. Most calculations are matrix operations, and their size is usually very large. Therefore, in order to reduce the memory occupancy rate and accelerate the calculation speed, a sparse matrix solver was used rather than directly inversion of the matrix.

In the case of transient simulation, the nonlinear simulation method is the same as that of steady-state simulation, and the backward Euler method is applied for the implicit time integration. Compared to the forward Euler (explicit) method, a much larger time step size can be used and more stable simulation is possible.

The WTE in the transient state is expressed by adding the time

derivative term to Eq. (17):

$$\frac{\partial}{\partial t} f(\chi,k,t) + \frac{\partial \varepsilon}{\hbar \partial k} \frac{\partial}{\partial \chi} f(\chi,k,t) + W_{\chi,k}(t) - C_{\chi,k}(t) = 0, \tag{33}$$

The simplest way to solve this is to use the forward Euler method (explicit method). The upwind scheme is stable only when the Courant–Friedrichs–Lewy condition (CFL) is satisfied:

$$C = \left| \frac{v \Delta t}{\Delta x} \right| \leq C_{\max}. \tag{34}$$

Here, if the explicit method is used, C_{max} becomes 1, so a very small time step is required. For example, if dx is 0.5 nm and v is $5*10^5$ m/s, dt must be smaller than 1 fs, and with such a small time step size, general transient simulation over a wide time range is difficult because of the long calculation time. Therefore, to avoid this problem, the backward Euler method (explicit method) is used in this work. In this method, since C_{max} has a value much larger than 1, a sufficiently large time step size is allowed. The solution at $t=t_1$ is known, and the solution at $t=t_2$ can be calculated through the following equation:

$$\frac{f(\chi,k,t_2) - f(\chi,k,t_1)}{t_2 - t_1} + \frac{\partial \varepsilon}{\hbar \partial k} \frac{\partial}{\partial \chi} f(\chi,k,t_2) + W_{\chi,k}(t_2) - C_{\chi,k}(t_2) = 0. \tag{35}$$

Since the implicit method is used, more complex calculations are required differently from the explicit method. As in the steady state, a fully coupled scheme through the Newton-Raphson method is used to obtain a self-consistent solution. Also, all the other discretization methods are the same as in the steady state. However, empirically, it takes too long to find a solution through iterative calculation for each time step in fully implicit transient simulation. Therefore, we propose and use a semi-implicit method to avoid iterative calculations. The system to be solved is a non-linear system because it is necessary to find a solution that satisfies both the Poisson equation and WTE. However, if the electrostatic potential is handled explicitly, the potential becomes constant value at each time step, and only the implicit WTE needs to be solved. If only WTE is solved, since this is a linear system, the solution in the next time step can be obtained without iteration using the given potential. Even using this semi-implicit method, stable simulation results were obtained, and a sufficiently small time step size was used to reduce the truncation error.

For both the explicit method and the implicit method, the error for exact solution increases as the time step increases. It is also known that the implicit method has a smaller error than the explicit method for the same time step. However, there is a problem that

the explicit method overestimates the transient energy and the implicit method underestimates the transient energy. Therefore, for accurate transient simulation, a sufficiently small time step should be used in both methods. In addition, in the case of simulation for high frequency oscillations, in the case of implicit method, if the time step is large, the transient energy is underestimated to damp the oscillation, and proper characteristics may not appear in the simulation results. However, in our method, the implicit method is used for WTE and the explicit method is used for the Poisson equation, so the shortcomings of both methods can be compensated.

Because explicit methods are mixed, a small time step should be used unlike the fully-implicit method under the influence of the CFL condition. However, it is suitable for efficient transient simulation because it is desirable to use a sufficiently small time step to reduce the truncation error.

2-3. Multi-dimensional Simulation Methods

The previously derived WTE with infinite correlation length and numerical methods assume a one-dimensional simulation. In order to extend it to two or three dimensions, the equation must be reconsidered according to several directions. Also, the Jacobian

matrix size for solving the WTE becomes too large, and calculating the nonlocal potential term in multiple dimensions is also challenging and requires a lot of calculations.

Therefore, in an alternative way, we limited the multidimensional simulation to devices with double gate (DG) structures. And a mode space approach is used to solve the one-dimensional transport equation in the source-drain direction and solve the Schrödinger equation for each cross section in the confinement direction.

We can solve the Schrödinger equation for the confinement direction as follow:

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi \qquad (36)$$

In this study, an effective mass Hamiltonian with a parabolic band is assumed, and Eq. (36) is solved for each valley. The wave function in the gate oxide is assumed to be zero, and the boundary condition at the interface is zero. Since Eq. (36) is an eigen value and eigen vector problem, subband energy (eigen value) and wave vector (eigen vector) are calculated after constructing a system matrix for a given device structure. Eigen value and vector form a pair, and the number is the same as the number of mesh nodes. The subband energy calculated in each valley is arranged in ascending

order, and the lowest 5 subbands are considered.

This subband profile is used as potential energy in WTE to obtain WTE solution in the same way as described above. Since the electron density calculated through WTE is the 1D electron density, the 3D electron density is calculated as follows by multiplying the subband and the wave function corresponding to the position:

$$n_{3D}(x,y,z) = \sum_{\nu} \left| \chi_{\nu}^{z}(x,y) \right|^2 n_{\nu}(z) \tag{37}$$

where χ is the wavefunction of the subband ν. The calculated 3D electron density is used to update the electric potential by solving the Poisson equation.

Chapter 3.

Simulation Results

3-1. Simulation results according to the correlation length

First, we compared simulation results according to correlation length. When the correlation length is finitely limited, the following nonlocal potential term is used.

$$\hat{V}(\chi,k)=\frac{1}{i}\int_{-L_C}^{L_C} d\zeta e^{i\zeta k}\left[U\left(\chi+\frac{1}{2}\zeta\right)-U\left(\chi-\frac{1}{2}\zeta\right)\right]. \qquad (38)$$

For infinite correlation length, we used the new expression we derived earlier as follows.

$$\hat{V}(\chi,k)=2\int_0^{2L} d\zeta \sin(\zeta k)\left[U_{cor}(\chi,\zeta)\right]+2\frac{U_{ex}}{k}\cos(2Lk) \qquad (39)$$

Fig. 3 shows the simulation results according to the correlation length. It was confirmed that all simulation results were different according to the correlation length. Considering that the formula at

$L_c = \infty$ derived in this study is the most accurate solution, it showed the most similar results when $L_c = 2L$, and shows a negative current in low bias when $L_c = L$, and too large current when $L_c = 4L$. It showed nonphysical results with a large current and two peak currents. It can be seen that using a finite integral range sometimes shows nonphysical results, and all of them show different results from $L_c=\infty$. Although only the results up to the integration range of 4L are shown, if a finite integration range is used, the results will not converge to the results with infinite correlation length no matter how much the integration range is increased. To explain the reason, when Lc > 2L, Eq. (38) is converted into the following form:

$$\hat{V}(\chi,k) = 2\int_0^{2L} d\zeta \sin(\zeta k)\left[U_{cor}(\chi,\zeta)\right] + 2\frac{U_{ex}}{k}(\cos(2Lk) - \cos(2L_C k)) \tag{40}$$

Looking at Eq.(40), we can know that no matter how much the correlation length is increased, the entire calculated value oscillates in the form of a cosine function. Also, as the correlation length increases, the period of the last term becomes smaller, so it becomes a function that changes very quickly according to *k*. Therefore, the longer the correlation length, the smaller the mesh size in k−space is required for accurate calculation. For this reason,

assuming the same mesh size, as the correlation length increases, the nonlocal potential term is overestimated or underestimated to obtain a non-physical simulation result. Of course, this problem can be solved if the nonlocal potential term is analytically integrated inside the mesh not only in the real space but also in the k space. However, the formula integrated in the real space has a form that is difficult to integrate analytically up to the k space. And since the calculated value oscillates even if the correlation length is continuously increased as described above, it is preferable to use the formula with infinite correlation length for accurate calculation.

In addition, the simulation result through NEGF was also added as a comparison group under the same conditions, and it was confirmed that reliable resonant tunneling characteristics were shown. Compared to NEGF, it shows peak current at lower voltage and higher valley current. This is a phenomenon that has also appeared in other paper comparing WTE and NEGF in resonant tunneling diodes, and is expected to occur because the formulation and detailed boundary conditions of WTE and NEGF are different [28].

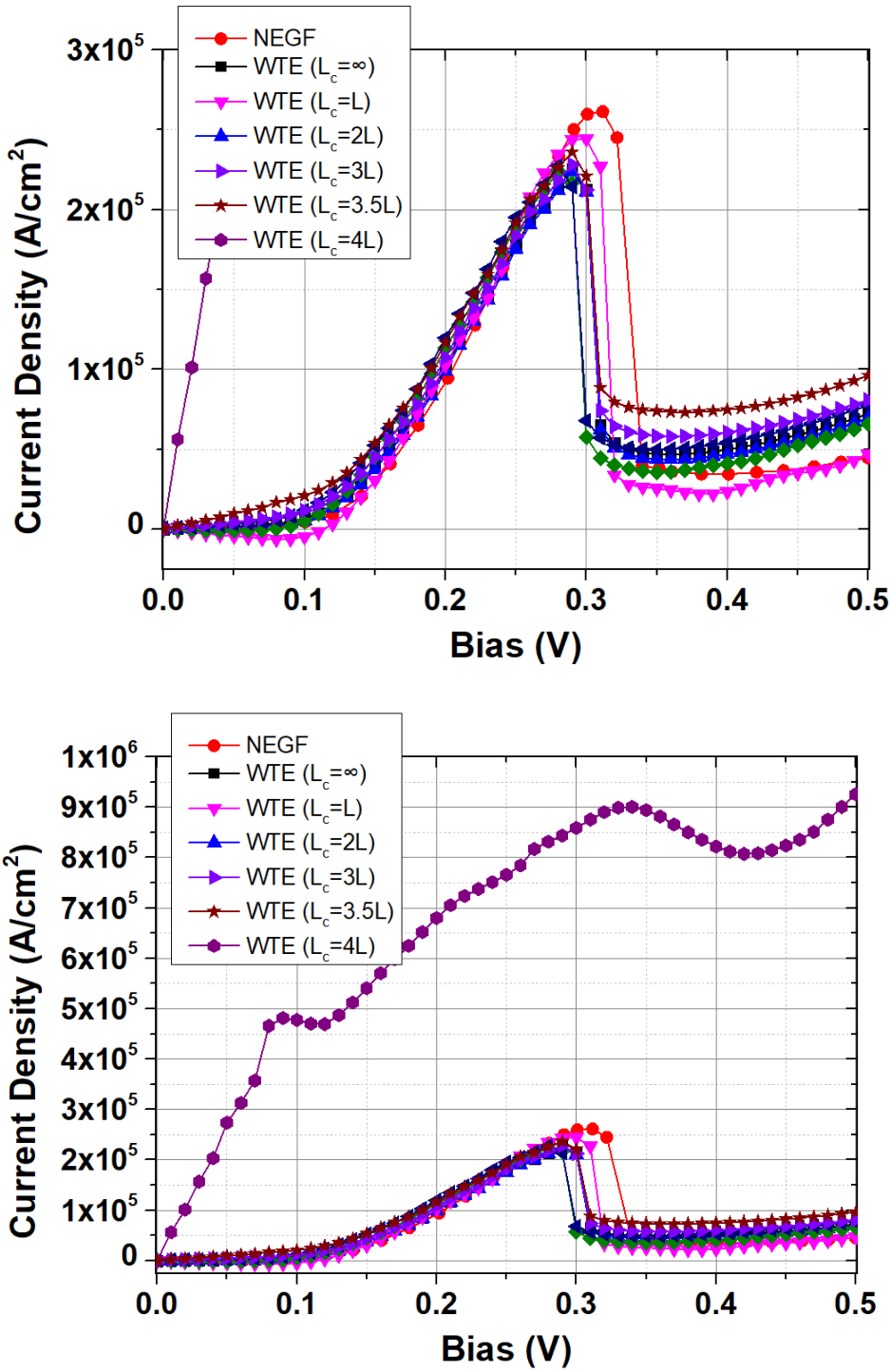


Fig. 3. Current characteristics according to correlation length.

3-2. Simulation for resonant tunneling diode

For the verification of the new model, we simulate a GaAs-AlGaAs-GaAs resonant tunneling diode as an example. Many previous studies have solved WTE by assuming the linear potential drop in the active region of RTD. In addition, even in the previous studies on the solution of deterministic Poisson-Wigner equation, detailed simulation convergence has not been mentioned. In this part, we compare the convergence of the coupled scheme and the decoupled scheme, and show the results for small signal analysis and transient simulation. We only simulated resonant tunneling diode and double gate MOSFET in this work, but we are currently working on applying it to 3D structures such as nanowire through the mode space approach.

As shown in Fig. 4, the emitter and collector have a doping concentration of $2*10^{18}/cm^3$, and the barrier and well are not doped. A 3nm barrier, 4nm well, and 7nm spacer are used. The band offset at the barrier is assumed to be 0.3eV, and the band diagram at the zero bias is shown in Fig. 5. In this work, we use 0.5nm for dx and 150 for N_k. We only consider one-dimensional transport in the z direction, and the boundary condition of the Wigner function is as

follows:

$$f^{+}(0,k)=\frac{m^{*}k_{B}T}{\pi\hbar^{2}}\ln\left[1+\exp\left(-\frac{1}{k_{B}T}\left(\frac{\hbar^{2}k^{2}}{2m^{*}}-\mu_{0}\right)\right)\right], \quad (41)$$

$$f^{-}(L,k)=\frac{m^{*}k_{B}T}{\pi\hbar^{2}}\ln\left[1+\exp\left(-\frac{1}{k_{B}T}\left(\frac{\hbar^{2}k^{2}}{2m^{*}}-\mu_{L}\right)\right)\right], \quad (42)$$

where k_b is the Boltzmann constant, T is the temperature, and u_0 and u_L are the Fermi energies at the ends of the device. We use a parabolic band approximation, and when k>0, the group velocity is greater than 0, so the boundary condition of (41) is applied, and when k<0, the boundary condition of (42) is applied.

Fig. 6 shows the convergence characteristics of our simulation method. A bias step of 0.01V is used, and the solution from the previous bias is used as the initial guess of the solution.

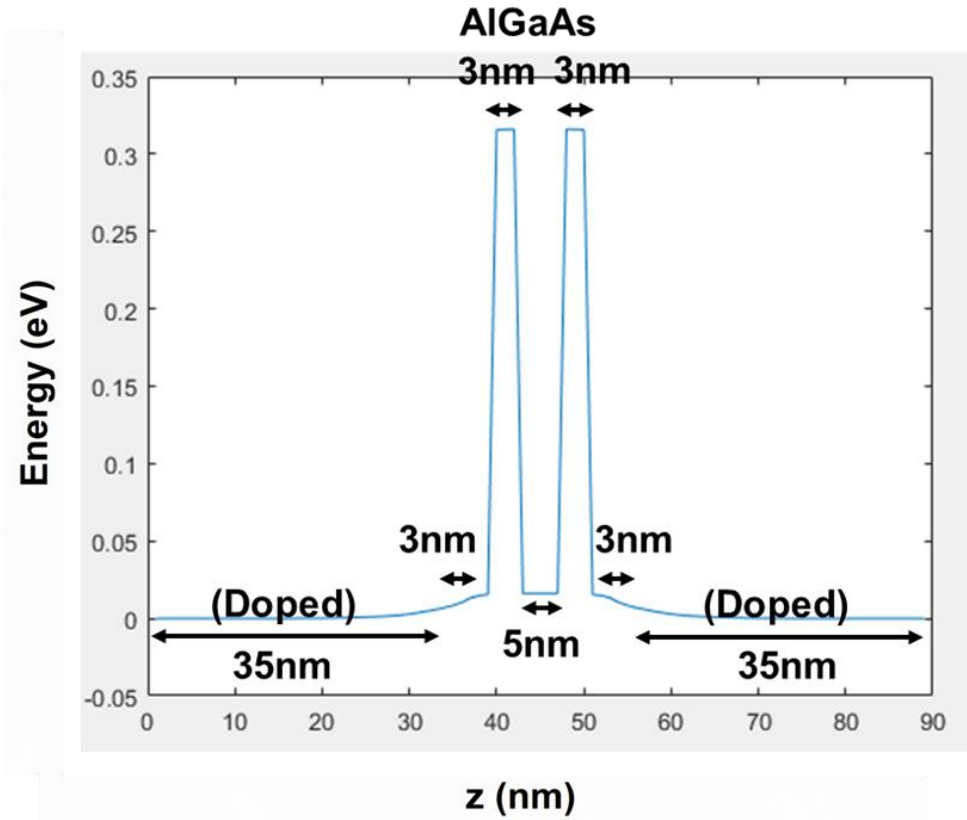


Fig. 4. The emitter and collector have a doping concentration of $2*10^{18}/cm^3$, and the barrier and well are not doped. A 3nm barrier, 4nm well, and 7nm spacer are used. The band offset at the barrier is assumed to be 0.27eV

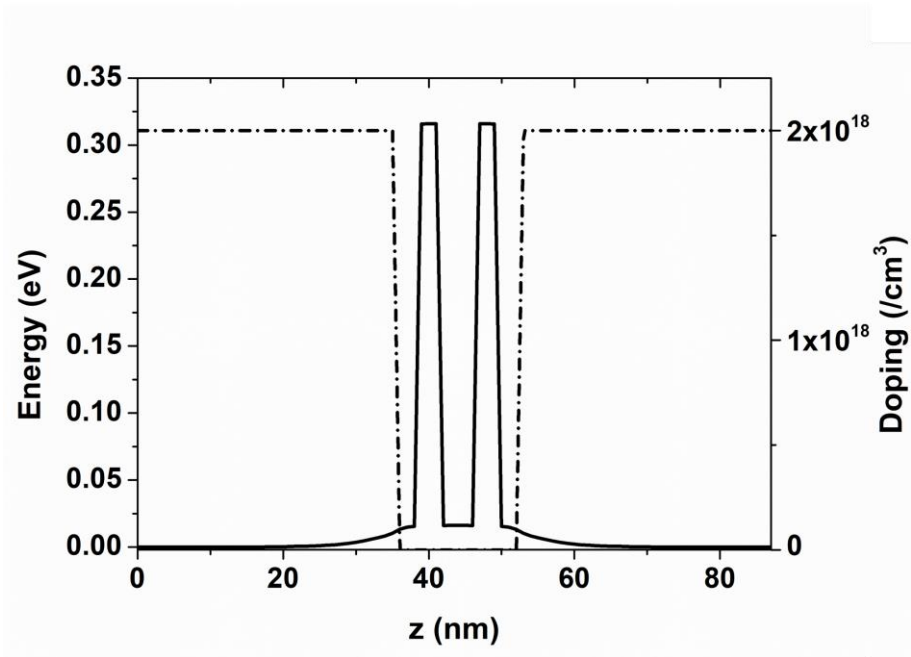


Fig. 5. Band structure at equilibrium and doping density.

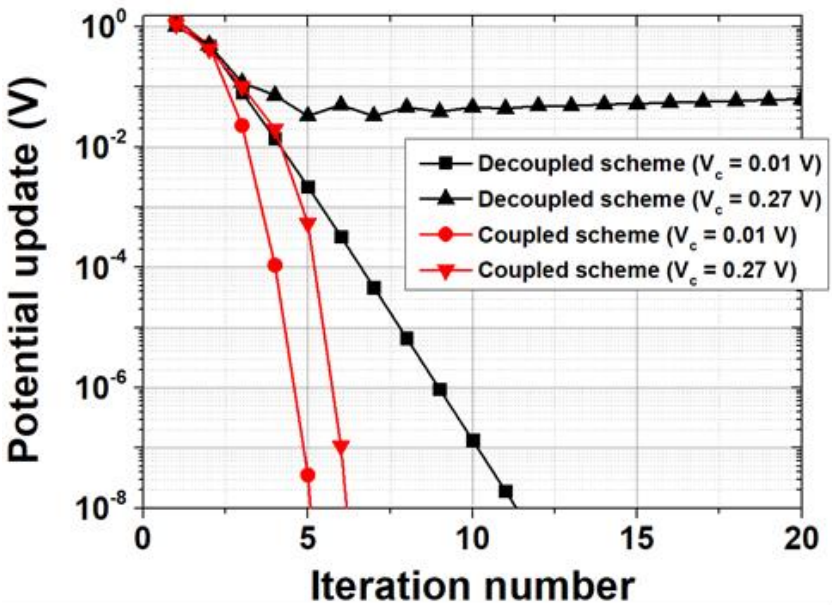


Fig. 6. Convergence of steady-state simulation according to iteration method. Coupled scheme (red line) shows better convergence than decoupled scheme (black line)

Since the exact linear response (Jacobian matrix) for the WTE and Poisson equations is calculated, it shows much more robust convergence when compared to the Gummel iteration method (decoupled scheme). As the bias increases, the convergence speed of the Gummel iteration method is significantly slower. In particular, as shown in Fig. 6, it can be seen that the simulation results oscillate at V_C=0.27 V. This is because the assumption used for the potential derivative term for the charge density becomes increasingly inaccurate as the electrical field increases and the resonance (quantum mechanical effect) occurs. Multiplying by a suitable damping constant or dividing the voltage step finer will result in convergence of simulation results, but this greatly increases the computational time. On the other hand, if the Newton-Raphson method (coupled scheme) is used, accurate results can be obtained with only a few iterations even at a high bias. When using the Newton Raphson method, it takes about twice as much time to solve a linear system equation using a sparse matrix solver. However, we confirm that it is generally more efficient because solution can be found with much fewer iterations. Also, since the linear response extracted in this process can be used for small-signal analysis or noise simulation, this simulation method is highly versatile [2].

Fig. 7 is the result showing the current continuity. As explained in numerical methods, it shows that the Fourier completeness relation is still satisfied even if the nonlocal potential term is integrated in the real space in the mesh for accurate calculation.

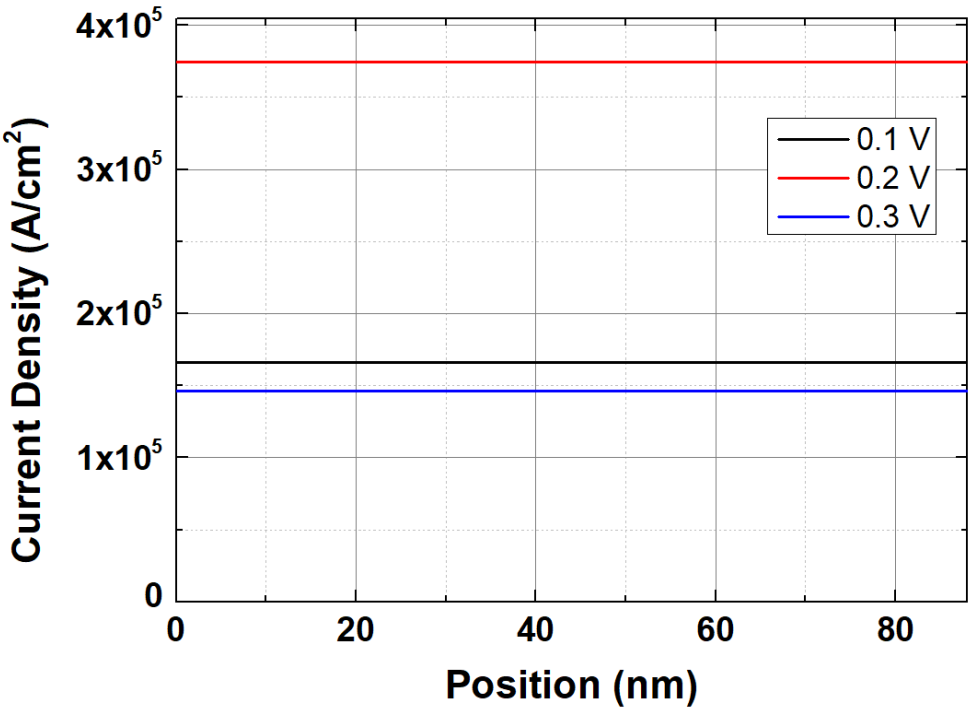


Fig. 7. Current density according to location within the device.

Fig. 8 shows the I-V characteristics for forward bias sweep and backward bias sweep. It can be seen that the current shows different currents in the range of 0.25 V to 0.32 V depending on the direction of the bias sweep and exhibits bistable characteristics. A plateau region in a positive bias sweep can also be found in previous studies [36]-[40]. However, since we performed calculations through the steady-state simulation method, there is a possibility that the wrong solution was found among several possible solutions. Therefore, to verify that the IV curve is calculated correctly, transient simulations are performed for the same biases. Fig. 9 is the simulation result when the bias is changed in step function form by 0.01V. We used 10fs as the time step size. Figure A shows the transient characteristics in the case of a positive bias sweep, and Figure B shows the transient characteristics in the case of a negative bias sweep. In all other biases except 0.27V, if transient simulation is performed for a long time, it converges to a steady-state solution. 0.27 V is a negative resistance region in the plateau, and it is known that intrinsic oscillations occur in this region in the previous studies. Fig. 10 shows the transient characteristics in a wider time domain. The solid line represents the transient current during forward bias sweep, and the dotted line represents the steady-state solution under the same bias. As in the red line, 0.28V,

in all other biases, the transient simulation results converge to a steady-state solution, but in the black line, 0.27V, the transient current oscillates. Also, this oscillation occurs centered on the steady-state solution, so it is shown that the same result as the steady-state solution can be obtained even by time-averaging the oscillating current. When sweeping the bias in the opposite direction, it was also confirmed that all bias converges to a steady state through transient simulation as shown in Fig. 11, and oscillation like the previous situation did not occur.

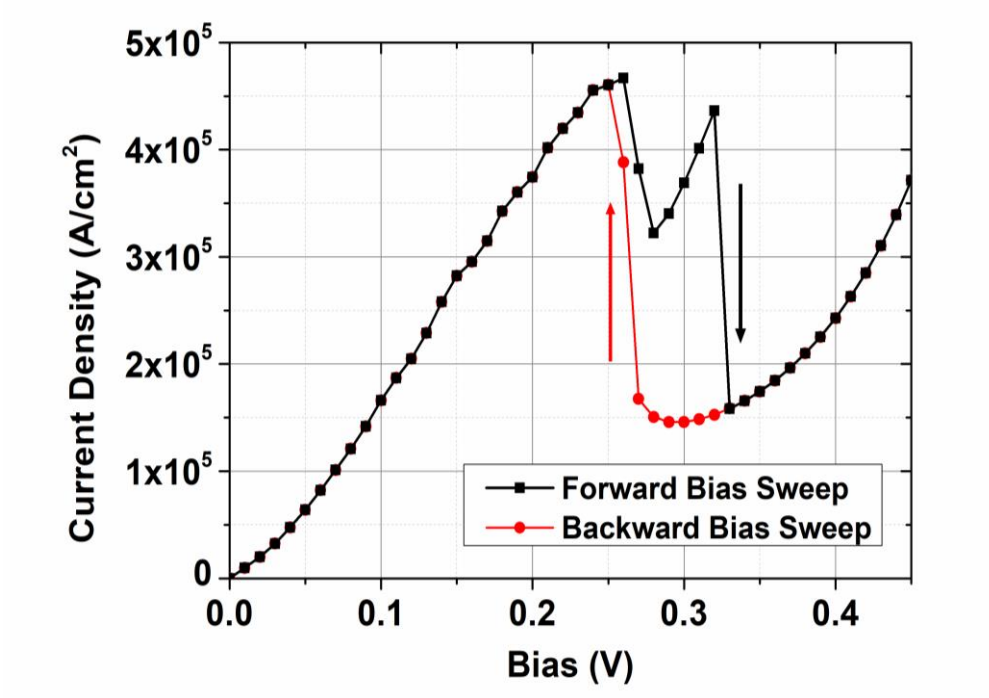


Fig. 8. Current-voltage characteristics obtained through steady-state simulation. 0.01V is used as the bias step, and bistability is shown in the case of forward bias sweep (black line) and backward bias sweep (red line).

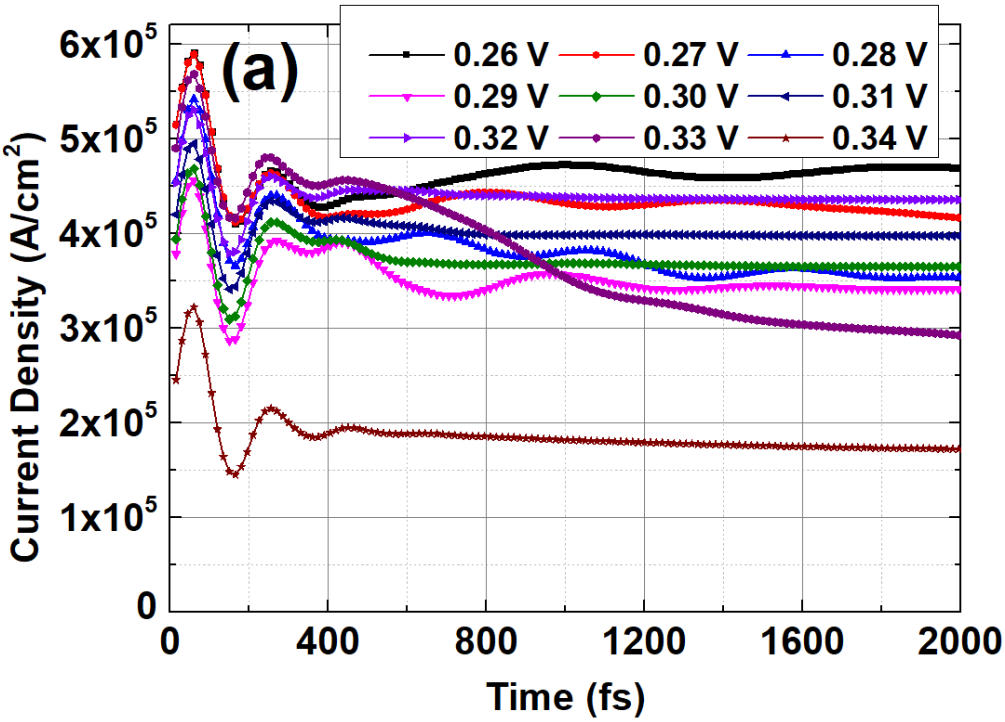


Fig. 9. Transient current characteristics up to 2000 fs.

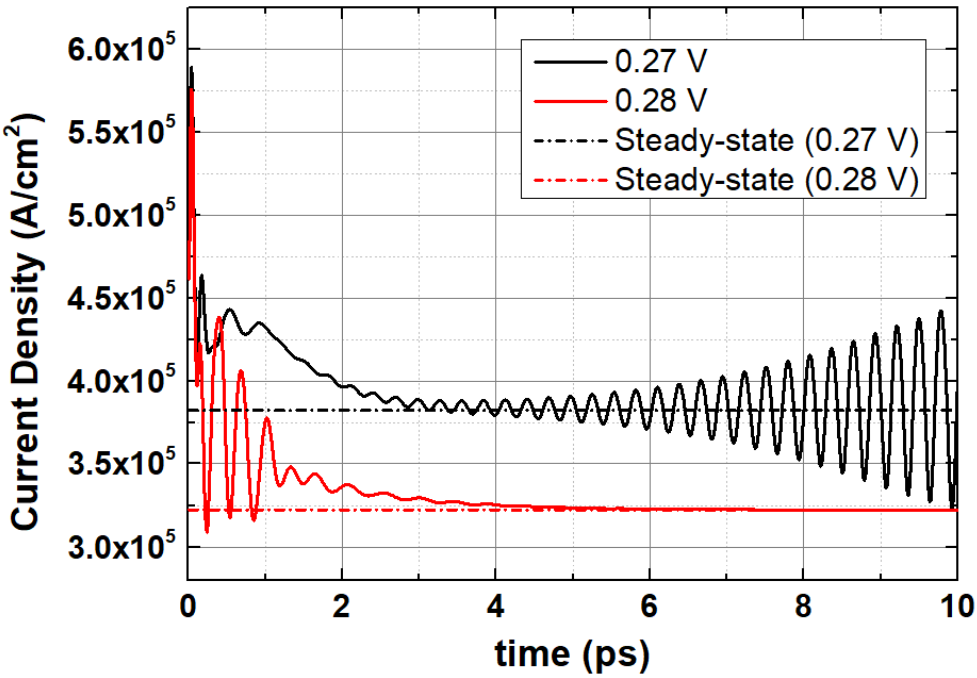


Fig. 10. Transient current characteristics up to 10 ps. At all biases except 0.27 V, the current converges to a steady-state solution. (b) At 0.27 V, the current oscillates even after a long time, and the average value of the oscillation current is almost the same as that of the steady-state solution.

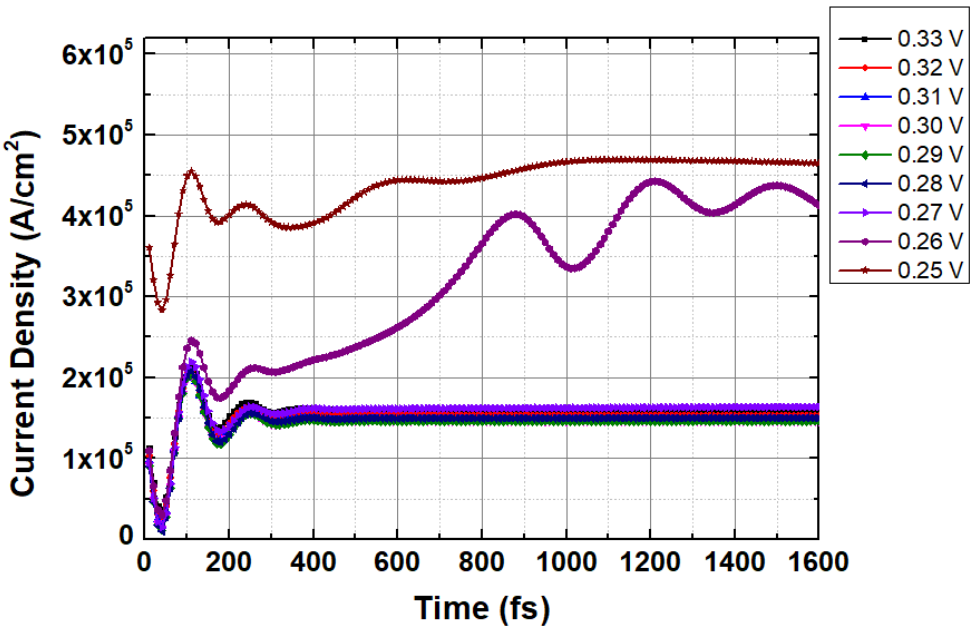


Fig. 11. When sweeping the bias in the opposite direction, it was confirmed that all bias converges to a steady state.

The previous results are simulation results for time step 10fs. In order to check the dependence of simulation results on different time step sizes, we conducted simulations for several time steps. Fig. 12 shows the forward bias sweep for 6 time steps and the transient current at 0.27V. Since we used the semi-implicit method when the time step size is 20 fs, the transient power is underestimated and the oscillation characteristic is not appear. From the time step size of 10fs, oscillation characteristics appeared, and the smaller the time step size, the closer to the real solution, the larger the oscillation amplitude. Through this, it can be confirmed that the time step size does not significantly affect the frequency, but mainly affects the oscillation size. Fig. 13 shows the transient current characteristics in section A (0~2ps) and section B (10~15ps). In section A, we can know that the difference in simulation results according to the time step size is not large. Therefore, if there is no intrinsic oscillation, high-accuracy simulation can be performed even with a relatively large time step size. However, since the characteristic difference according to the time step is clearly seen in section B, it can be seen that the truncation error can be minimized by using a time step size that is sufficiently small compared to the frequency when performing simulation of high-frequency oscillation. And the result when using

a time step size of 1fs showed an oscillation size of about 3.5 A/cm^2, which is much larger than the results of previous papers, and the frequency is about 2.5 THz, showing similar results.

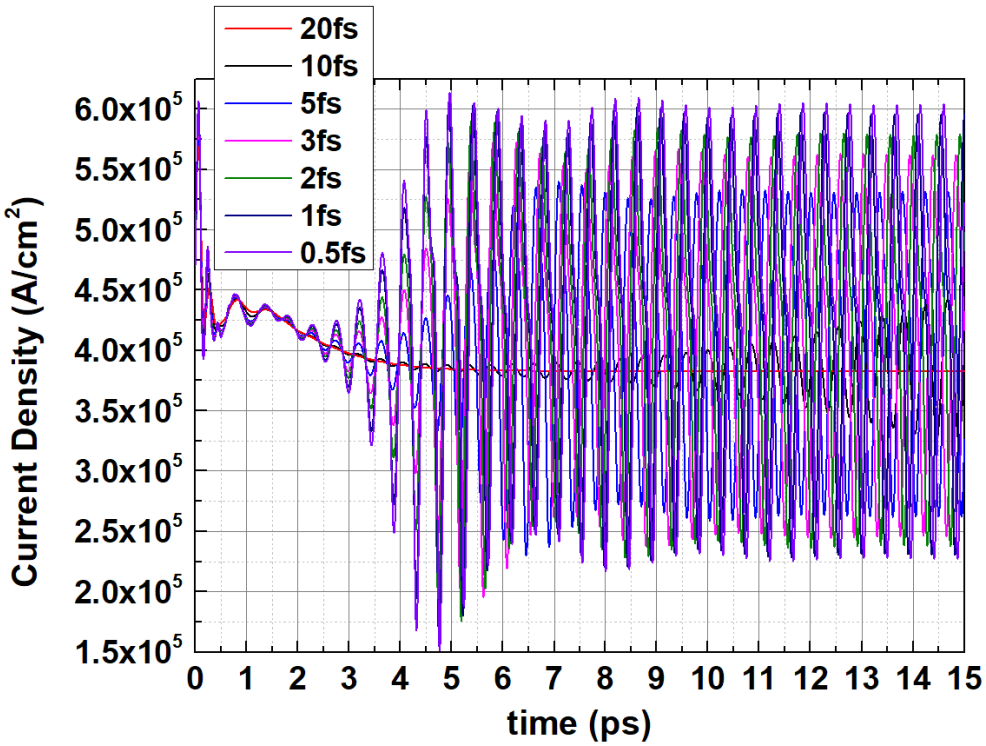


Fig. 12. Transient simulation results according to various time step sizes at forward sweep and bias 0.27V.

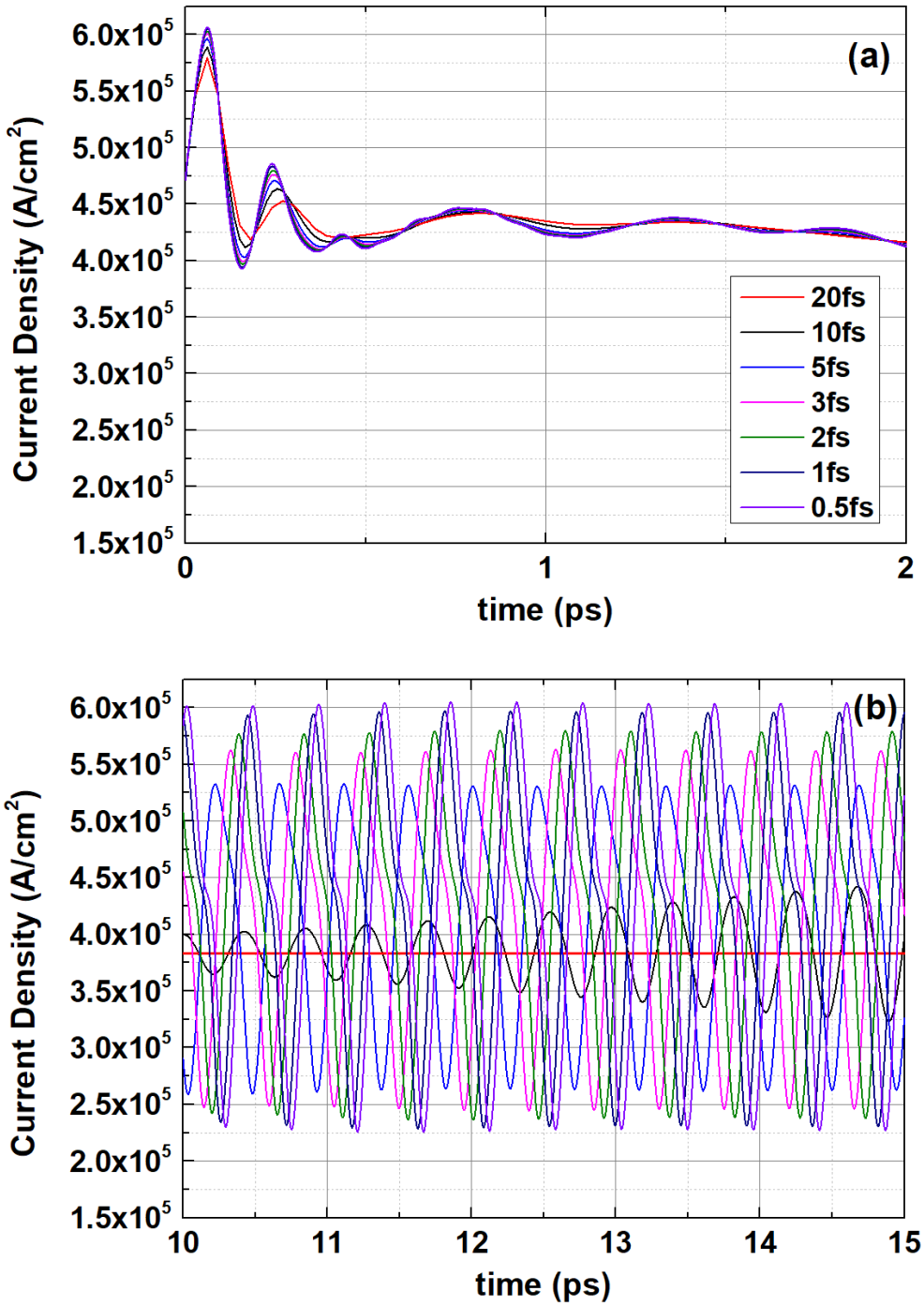


Fig. 13. Transient simulation results for section A (0~2ps) and section B (10~15ps).

Fig. 14 and Fig. 15 show the results of calculating the Wigner function under various bias conditions. It can be confirmed that there is a section showing a fairly large negative value due to the quantum interference effect at 0.24V, which is near the peak current. This is a characteristic of the qausi-probabilty density in which a negative value appears due to the uncertainty principle, When the electron density is calculated by integrating over the k-space, the positive electron density is well defined in all positions.

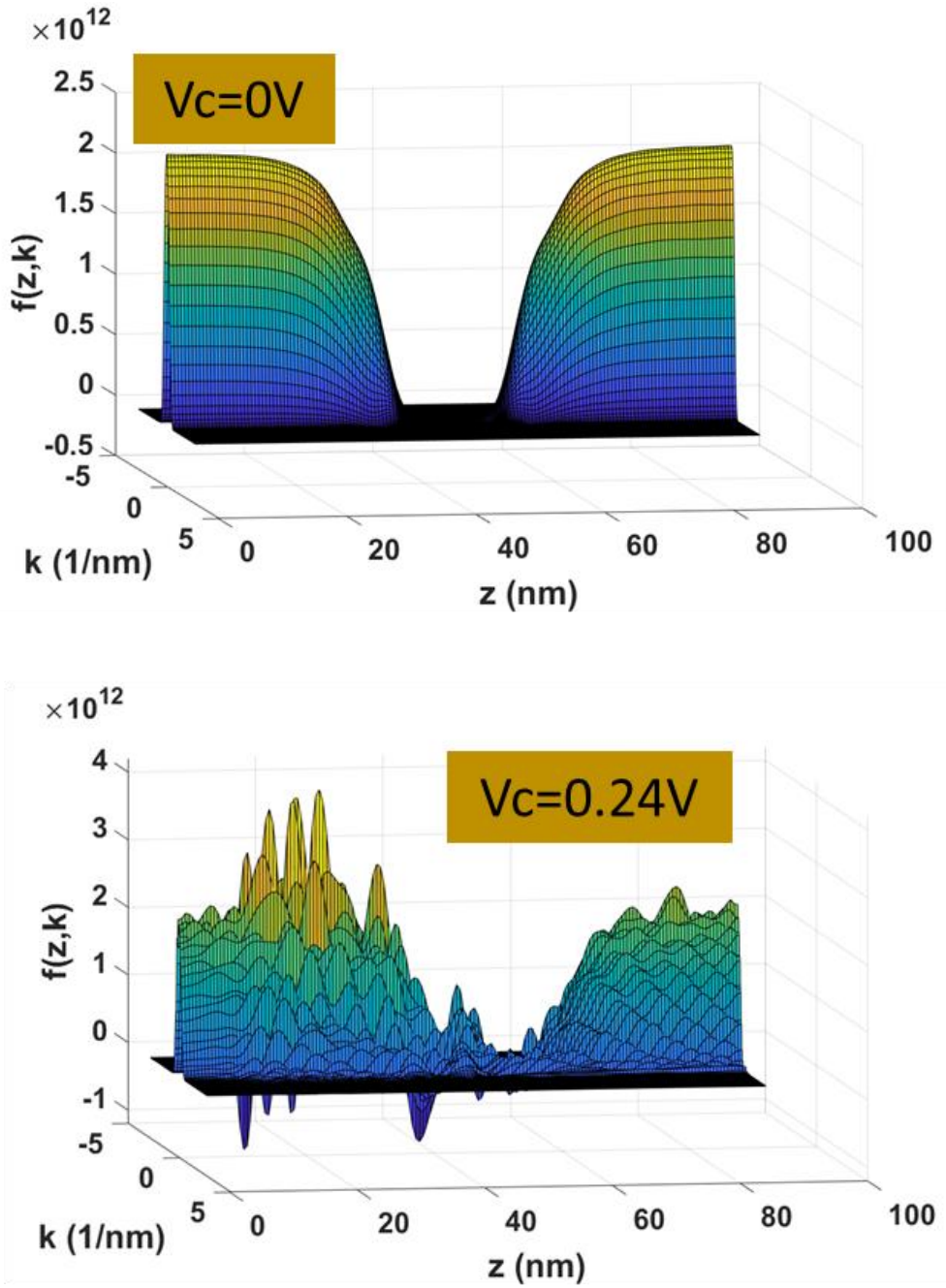


Fig. 14. Wigner function at (a) Vc=0V and (b) Vc=0.24V

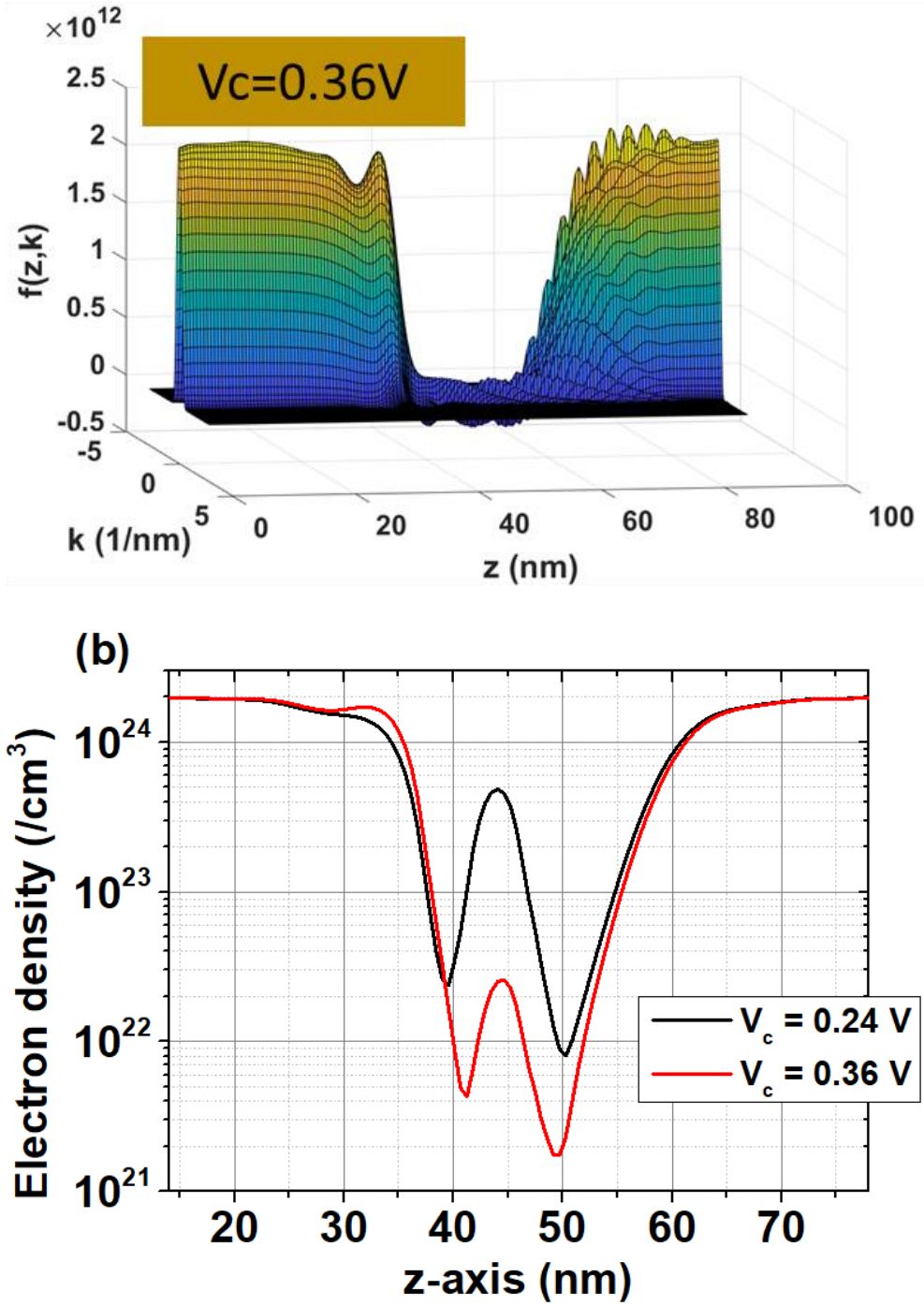


Fig. 15. (a) Wigner function at Vc=0.36. (b) Electron density for Vc=0.24V and Vc=0.36V.

Fig. 16 shows the band diagram at the plateau and the electron density at several biases. Figure A shows the band diagram in steady-state when forward bias sweep is performed. The solid line from 0.27 V to 0.32 V is the plateau region, and it can be seen that the characteristic is clearly different from the normal operation region, which is the dotted line. At the plateau, we can see that band banding occurs in the emitter region in front of the first barrier. Therefore, the emitter region also shows characteristics like another quantum well, and a quantized state exists. And when this state is similar to the resonance energy level between the double barriers, a new current path is formed as shown schematically in Fig. 16. Therefore, as shown in Fig. 17, not only the current in the plateau region but also the electron density in the quantum well does not drop significantly compared to the peak current.

These results show that our proposed new numerical formulation method shows good convergence and is reliable because it shows the same tendency as previous papers for RTD simulation.

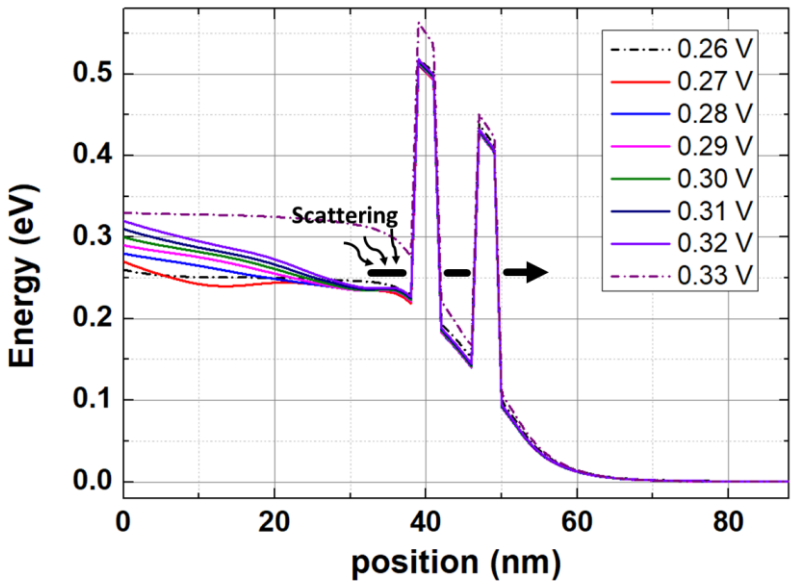


Fig. 16. Conduction band diagram when forward bias sweep is performed.

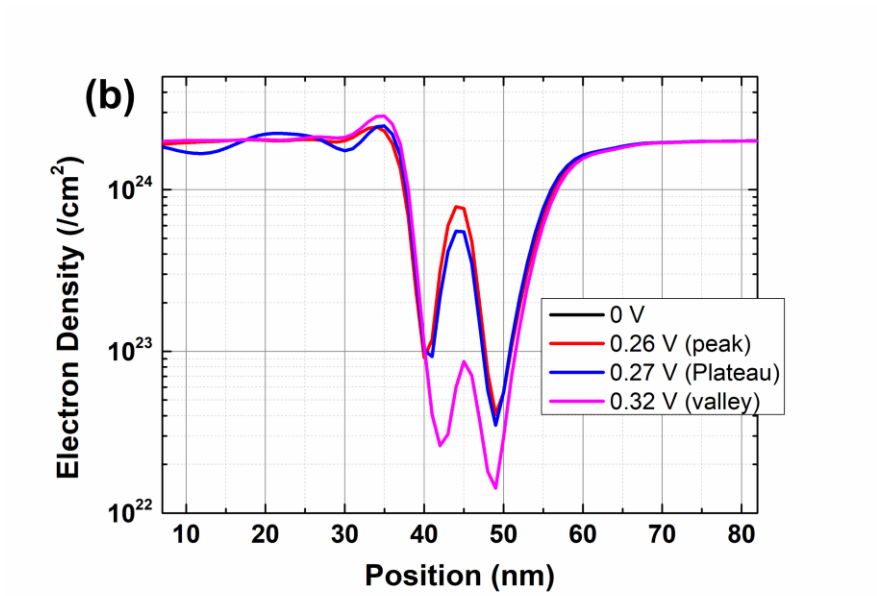


Fig. 17. Electron density in steady-state when forward bias sweep is performed.

3-3. Simulation for double gate MOSFET

Mode space method is used for simulation of double gate MOSFET. As described above, 1D WTE is solved in the source-drain direction and the Schrödinger equation is solved in the confinement direction. The boundary condition of WTE in Contact is as follows:

$$f_{\nu,\eta}^{+}(0,k)=\left[1+\exp\left(\frac{1}{k_BT}\left(\frac{\hbar^2k^2}{2m^*}+E_{\nu,\eta}^{z=0}-\mu_L\right)\right)\right]^{-1}, \quad (41)$$

$$f_{\nu,\eta}^{-}(0,k)=\left[1+\exp\left(\frac{1}{k_BT}\left(\frac{\hbar^2k^2}{2m^*}+E_{\nu,\eta}^{z=L}-\mu_R\right)\right)\right]^{-1}, \quad (42)$$

Where u_L and u_R are the quasi-Fermi levels at the source/drain contacts.

The length of the source and drain region is 9 nm and the doping concentration is $10^{20}/cm^3$. The gate oxide uses SiO_2 and has a thickness of 1 nm. The channel length is 7nm and 10nm, and the body thickness is 3nm and 5nm. To produce an off-current of 1~100A/m, the gate work function is set to 4.22eV.

Unlike simulation in RTD, Schrödinger equation in cross section also needs to be solved, so using Newton-Raphson method makes the system matrix size too large. Therefore, for efficient simulation,

the Gummel method, which consumes less memory, is used. Since the exact Jacobian matrix used in the Newton-Raphson method is required to obtain the correct linear response of the device by any source, in this case, the matrix can be calculated directly. However, since the size of the matrix can be larger than the DRAM memory size of the simulation computer, only the non-zero elements of the matrix are stored and the sparse matrix solver is used.

Fig. 18 shows the IV curve when the gate length is 7 nm and Fig. 19 shows the gate length is 10 nm. Obviously, when the channel length is short, we show that the leakage current increases and the on/off ratio decreases. In addition, when the body thickness is thin, gate controllability is improved, resulting in lower off-current and higher on-current characteristics.

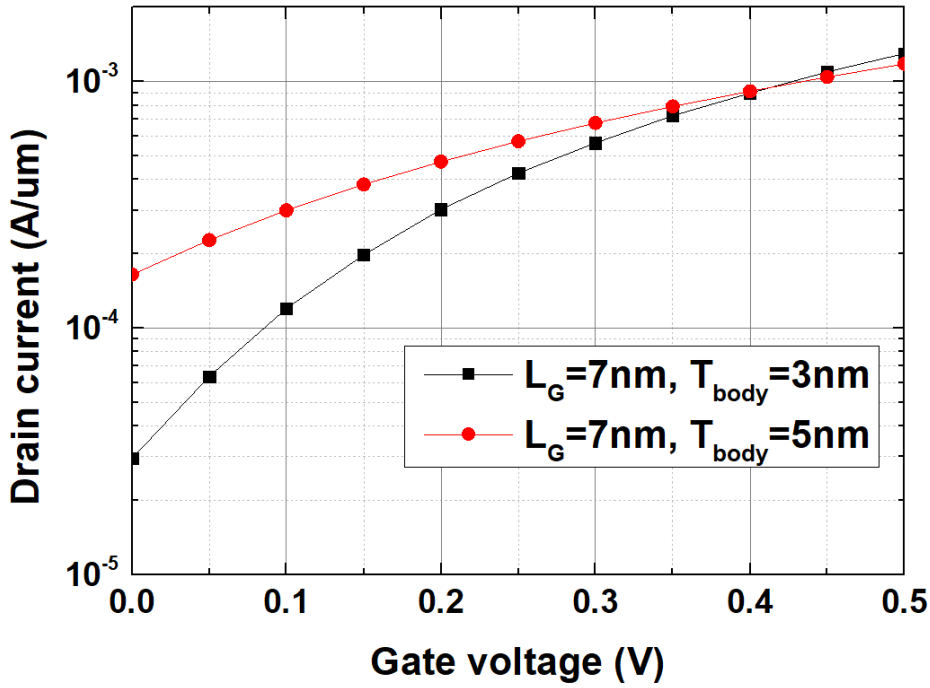


Fig. 18. IV curve when the gate length is70nm. (V_{DS}=0.4V)

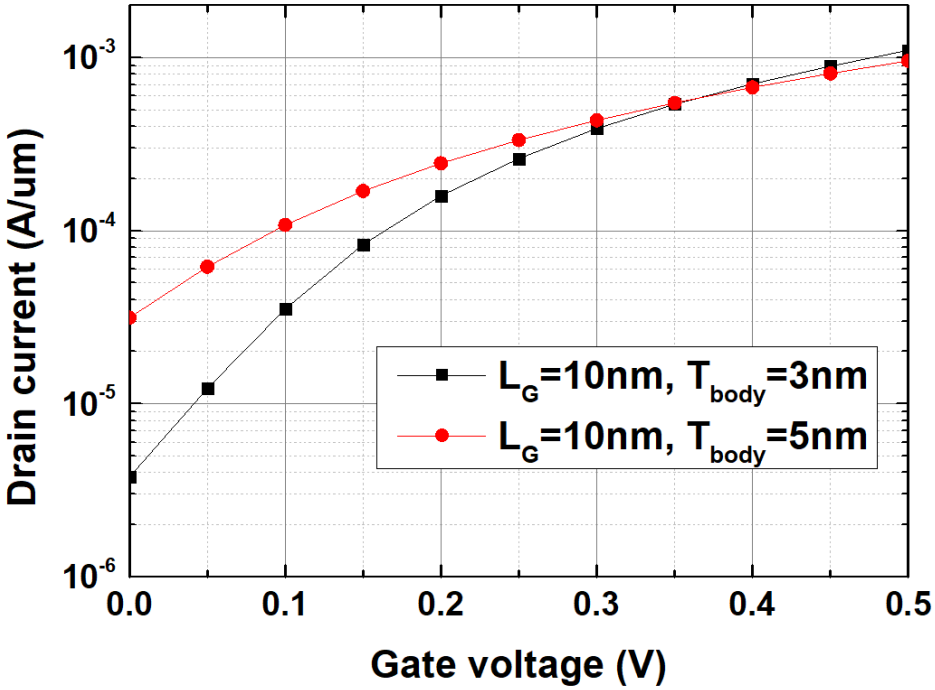


Fig. 19. IV curve when the gate length is 10nm. (V_{DS}=0.4V)

Fig. 20 and Fig. 21 shows the lowest subband pofile in on-state (Vgs=0.5V) and off-state (Vgs=0V) when the gate length is 7nm. Fig. 20 shows a body thickness of 3 nm and Fig. 21 is when the body thickness is 5 nm. Fig. 22 and Fig. 23 shows the lowest subband pofile according to the body thickness when the gate length is 10 nm.

We can know that the thinner the body thickness for the same gate length, the lower the energy barrier in the on-state, so that a higher current flows, and in the off-state, the energy barrier becomes higher and thus a lower leakage current flows. As such, when the body thickness is thin, electrostatically improved gate controllability improves device characteristics.

For the same body thickness, the shorter the channel length, the lower the energy barrier in off-state. Structurally, because it is electrostatically affected by source and drain, the energy barrier is lowered, and at the same time, the channel length is also shortened, so that a higher leakage current including source-drain tunneling flows.

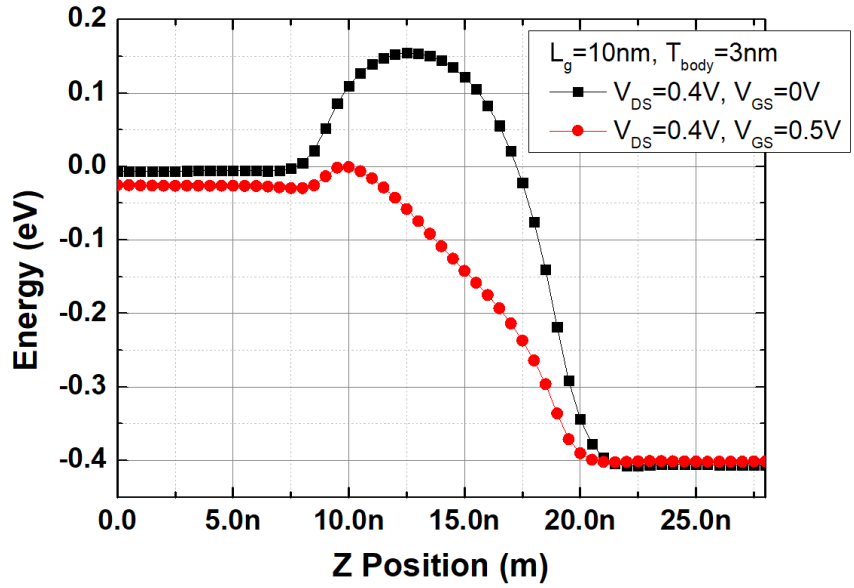


Fig. 20. The lowest subband pofile in on-state (Vgs=0.5V) and off-state (Vgs=0V) when the gate length is 10nm and body thickness 3nm.

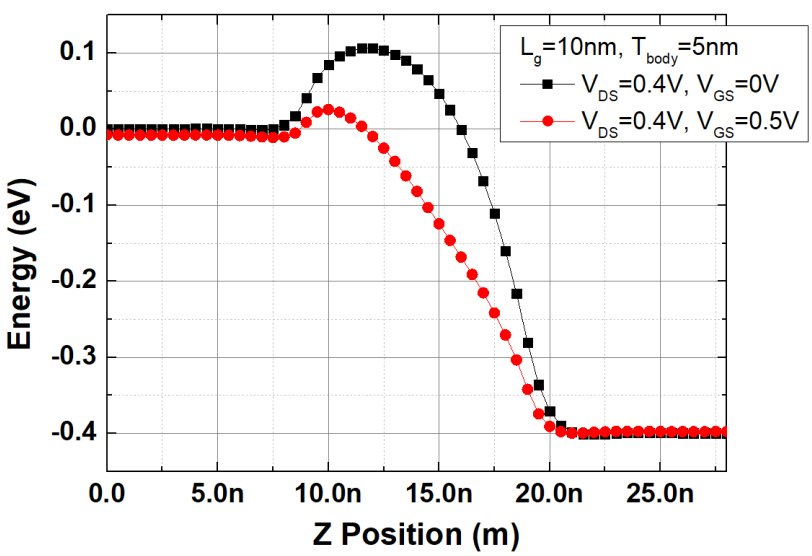


Fig. 21. The lowest subband pofile in on-state (Vgs=0.5V) and off-state (Vgs=0V) when the gate length is 10nm and body thickness 5nm.

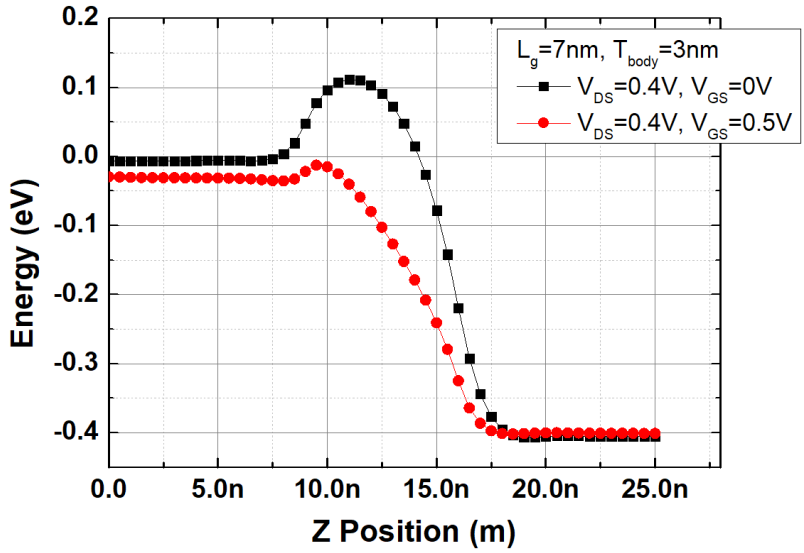


Fig. 22. The lowest subband pofile in on-state (Vgs=0.5V) and off-state (Vgs=0V) when the gate length is 7nm and body thickness 3nm.

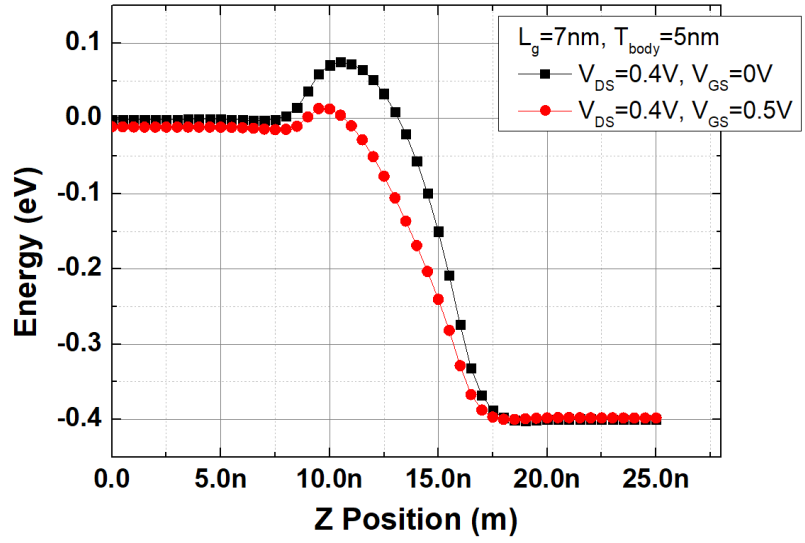


Fig. 23. The lowest subband pofile in on-state (Vgs=0.5V) and off-state (Vgs=0V) when the gate length is 7nm and body thickness 5nm

Fig. 24−27 shows the Wigner function in each condition. Simulation was performed under the same conditions as Fig. 20−23, and this is the result of WTE calculation in the lowest subband. The figures show that in the on−state, the current flows well with the Wigner function showing a high value even near the channel in all cases. In off−state, the Wigner function shows a high value only in the source and drain regions, but when the gate length is 7 nm and the body thickness is 5 nm, the Wigner function increases near the channel and the leakage current flows clearly. The reason why only the results for the lowest subband are shown in the previous results is that, since the largest Wigner function is injected as the boundary condition in the lowest subband, it has the largest electron density and has a major influence on the results. The 3D electron density and electrostatic potential for each condition are shown in the appendix.

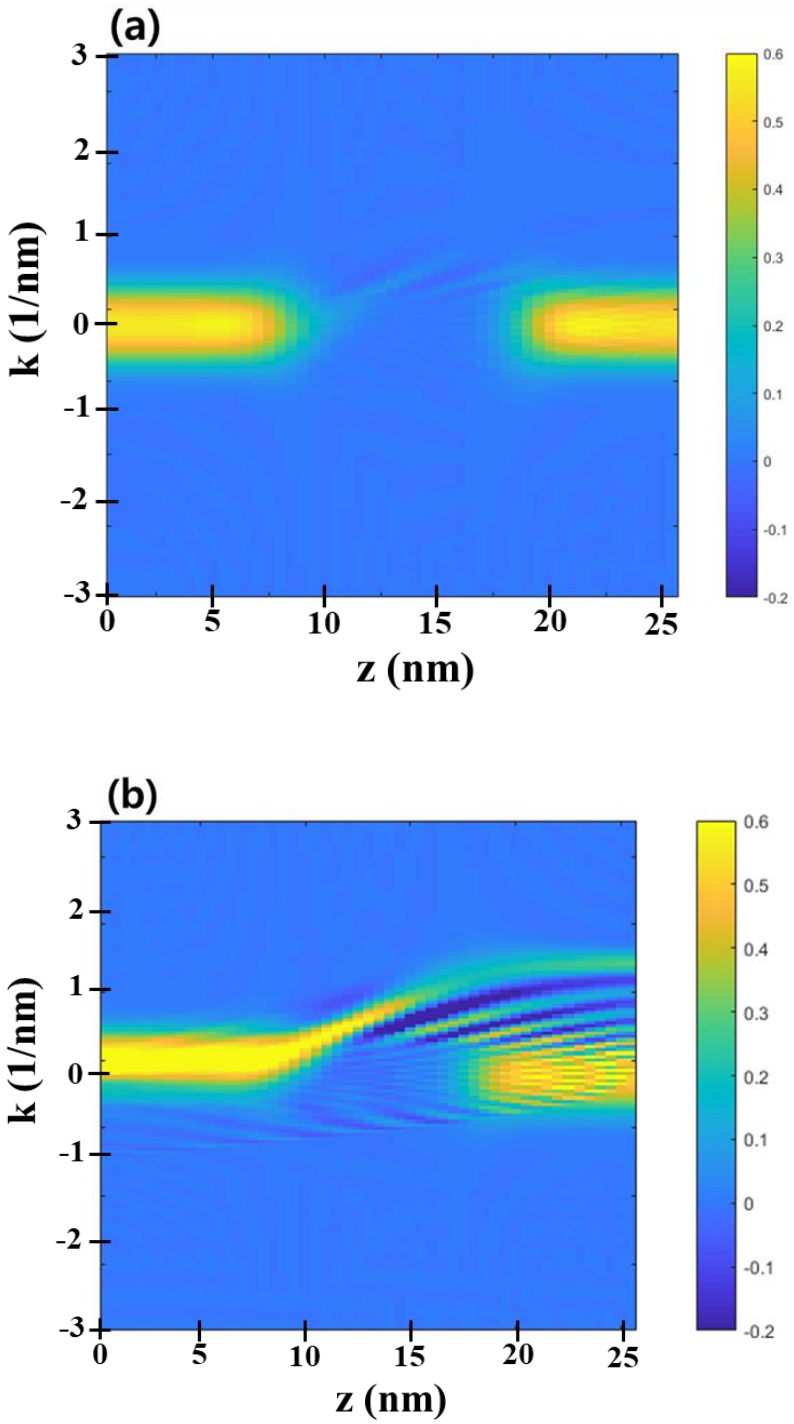


Fig. 24. Wigner function in (a)off-state (Vgs=0V) and (b)on-state (Vgs=0.5V) when the gate length is 7nm and body thickness 3nm.

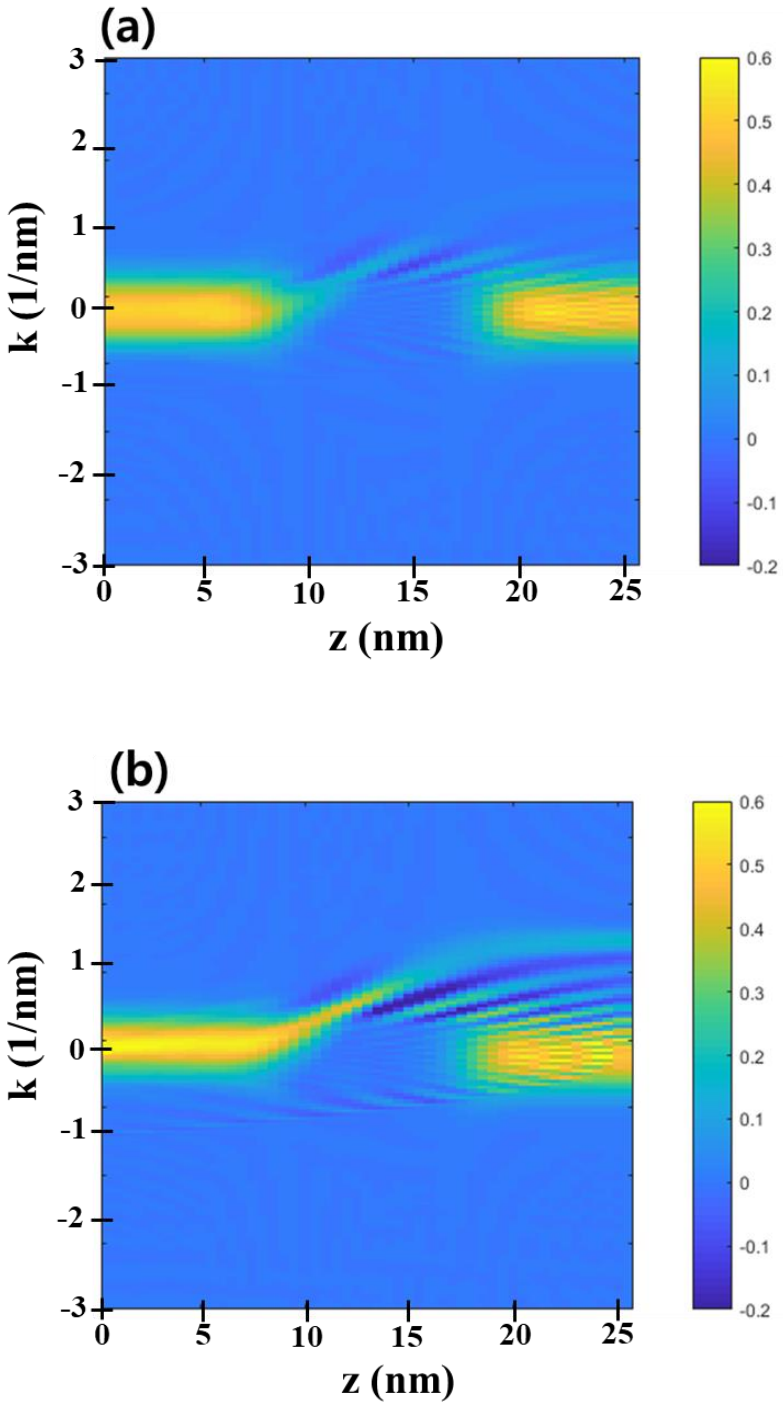


Fig. 25. Wigner function in (a)off-state (Vgs=0V) and (b)on-state (Vgs=0.5V) when the gate length is 7nm and body thickness 5nm.

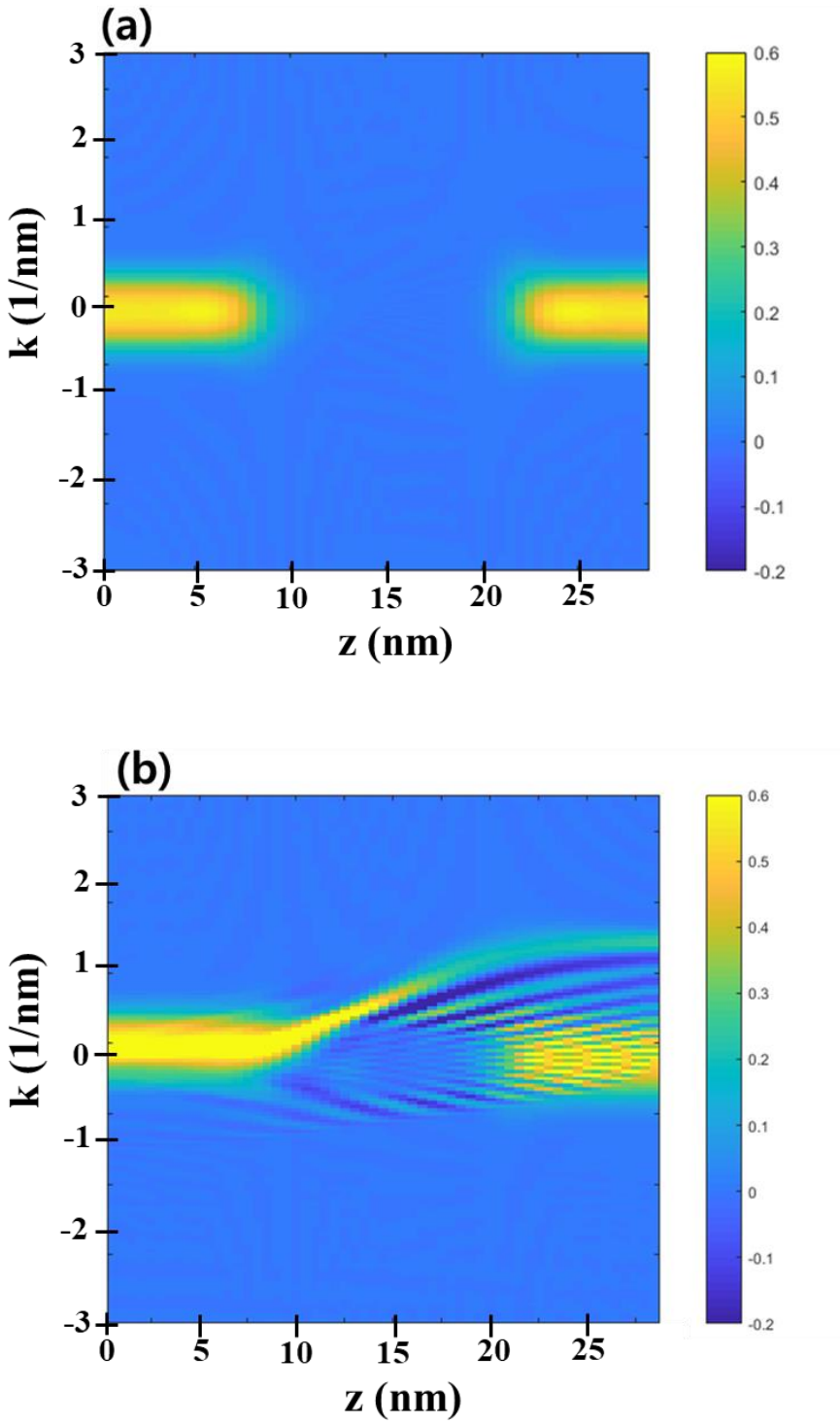


Fig. 26. Wigner function in (a)off-state (Vgs=0V) and (b)on-state (Vgs=0.5V) when the gate length is 10nm and body thickness 3nm.

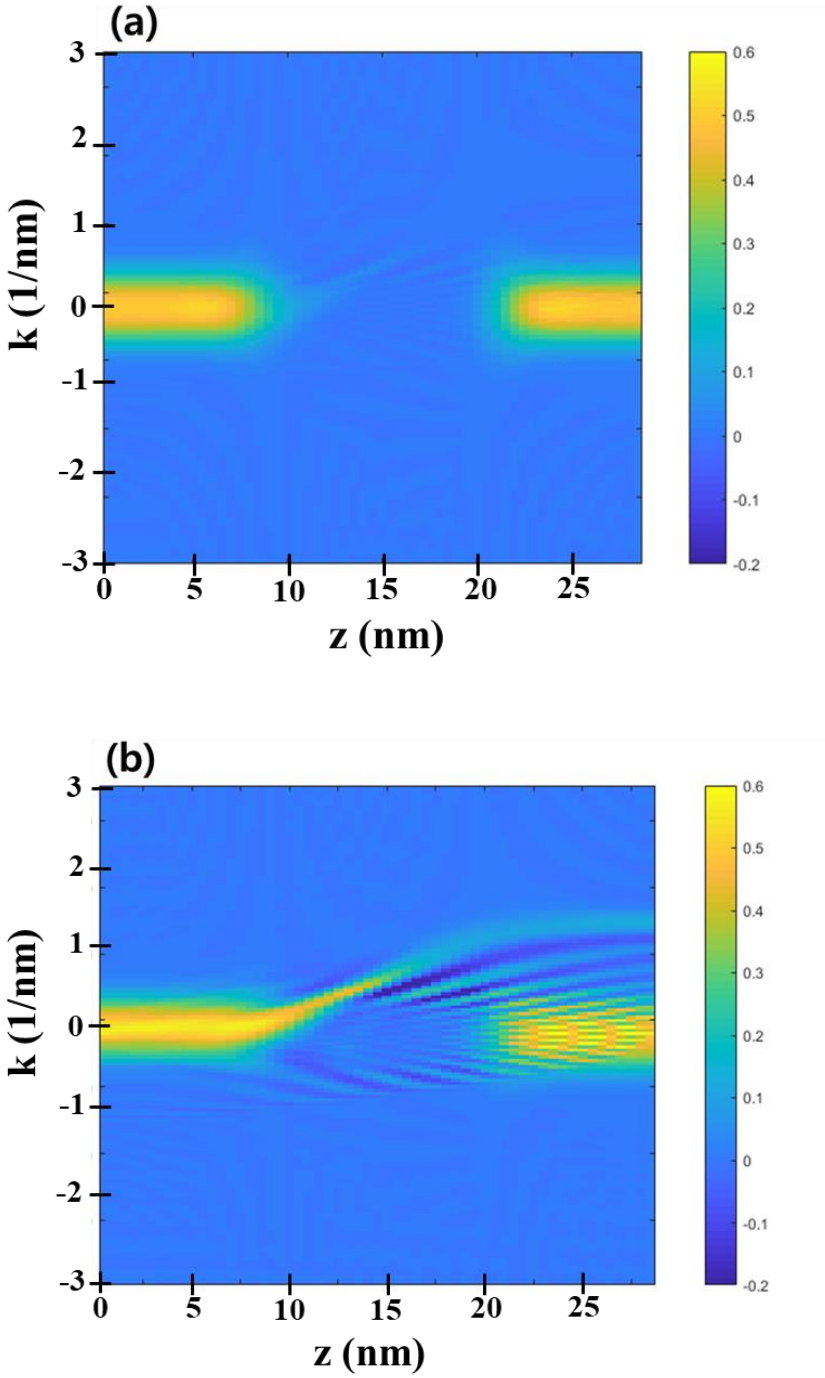


Fig. 27. Wigner function in (a)off-state (Vgs=0V) and (b)on-state (Vgs=0.5V) when the gate length is 10nm and body thickness 5nm.

The gate length of 10 nm, body thickness of 3 nm, and gate oxide thickness of 1 nm, which showed the best characteristics among the previous results, also showed a much higher subthreshold slope (SS) than 60 mV/dec of about 95 mV/dec, even though ballistic transport was assumed. Therefore, in order to predict better performance, a simulation was conducted with a thinner equivalent oxide thickness (EOT). To reduce EOT, there is a method of depositing a thinner silicon oxide or a method of using a high-k gate oxide. When a high-k material is used as the gate oxide, EOT can be expressed as

$$EOT = t_{high-k}\left(\frac{k_{SiO_2}}{k_{high-k}}\right) + t_{SiO_2} \tag{43}$$

We conducted the simulation by reducing the EOT from 1 nm to 0.5 nm. In order to make the simulation simpler without changing the device structure, the oxide thickness was fixed at 1 nm and only the dielectric constant was changed. As shown in Fig. 28, it can be seen that as EOT decreases, the on current increases and the off current decreases. For example, looking at the lowest subband profile, we can know that the gate controllability is improved when EOT is 0.5nm compared to when it is 1nm. As a result, SS is also improved to 95 mV/dec when EOT is 1 nm and to 82 mV/dec when EOT is 0.5 nm.

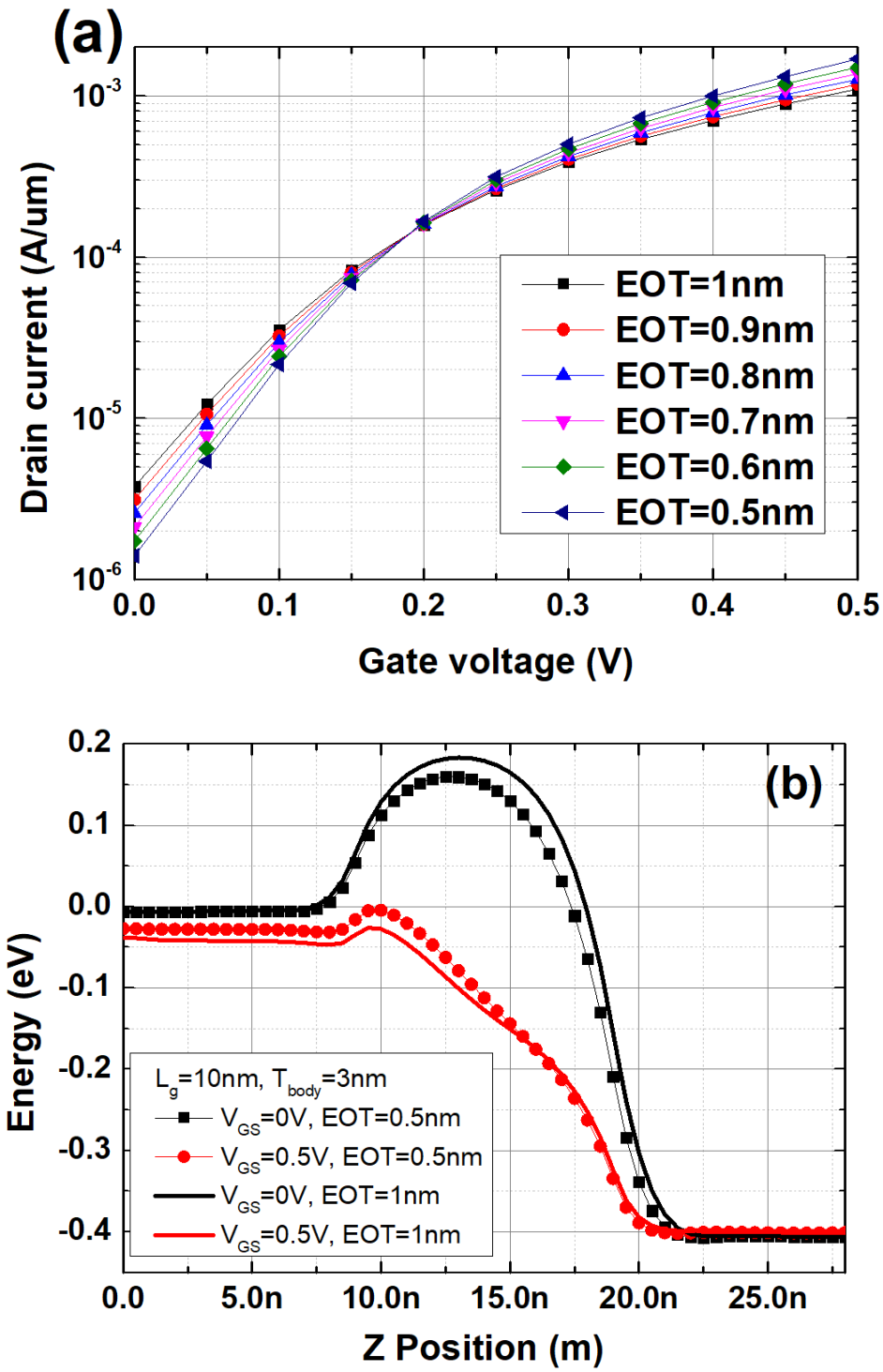


Fig. 28. (a) gate voltage – drain current charateristic and (b) lowest subband diagram according to EOT.

Fig. 29 is a diagram showing all subbands in each valley. As shown in the figure, except for a few subbands, all of them have high energy, so only 5 modes for each valley were considered. As mentioned above, since subbands with high energy hardly contribute to electron density, it can be confirmed that sufficiently accurate simulation results can be obtained by considering only a few subbands.

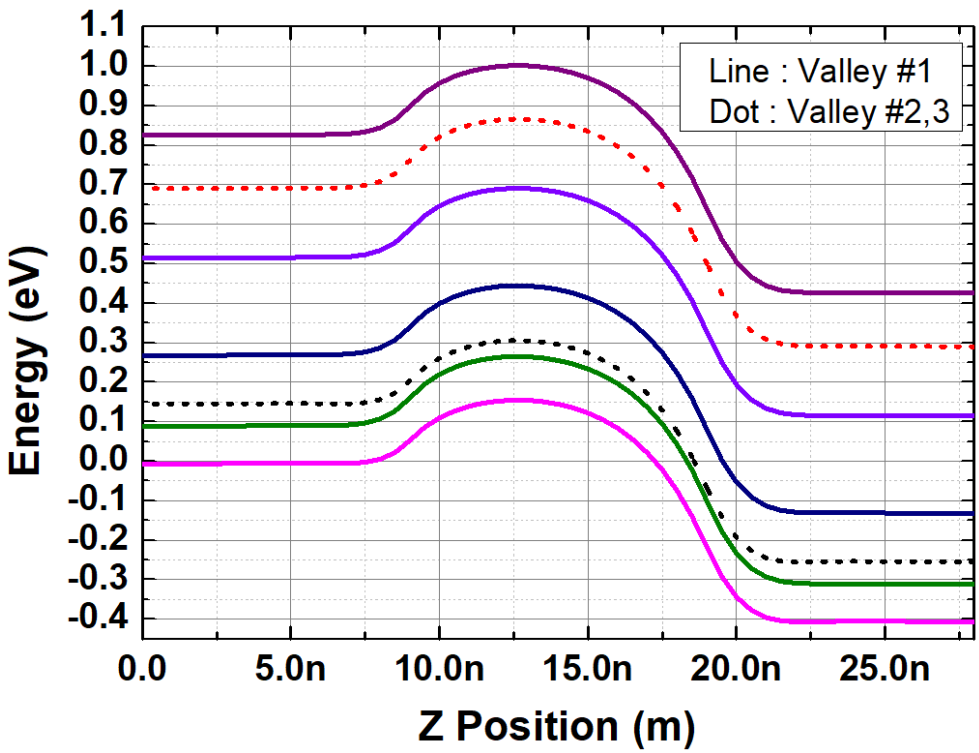


Fig. 29. (a) Subband energy for each valley.

Next, transient simulation was performed on a double gate MOSFET. With Vds of 0.4V, a bias in the form of a step function was applied to Vgs from 0V to 0.5V. As a result, Fig. 30 shows that the drain current converges to a steady state at several ps. However, this is only an attempt at transient simulation in a two-dimensional simulation, and there are several issues to consider. Among all equations, only WTE has a transient form, and the Schrödinger equation in the cross section is a steady-state equation. Therefore, the transient characteristics of WTE can be considered for each time step, but in the case of Schrödinger's equation, only a steady-state solution for a given potential can be derived. Therefore, it is necessary to verify whether the method we used is valid through another method.

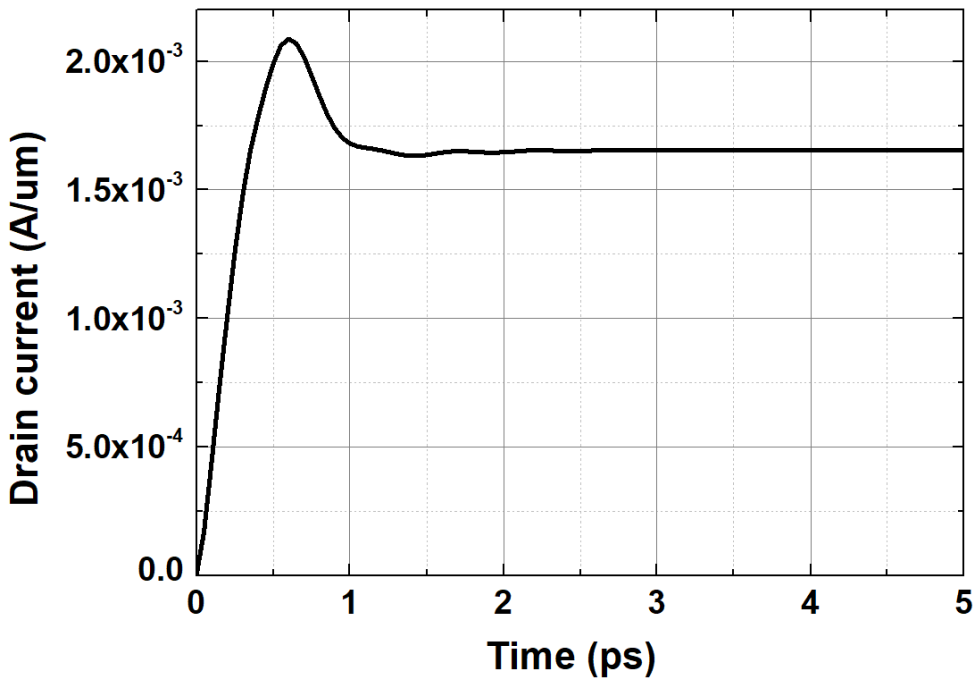


Fig. 30. Transient current characteristic. With Vds of 0.4V, a bias in the form of a step function is applied to Vgs from 0V to 0.5V. Gate length is 10nm, body thickness is 3nm, and T_{OX} is 0.5nm.

The previous simulation results are for ballistic transport. As in the simulation for RTD, we analyzed the device characteristics considering the scattering effect through relaxation time approximation. Relaxation time can be expressed as follows for effective mobility.

$$\tau = m_{eff} \times \mu / e \tag{44}$$

We performed simulation for u=386cm^2/Vs. Fig. 31, it was confirmed that the on-current decreased by about 31% when scattering was considered. Fig. 32 also shows that it has a broadened Wigner function when there is scattering in the on-state. However, since relaxation time approximation is used in this study, the shape of the Wigner function may be different if more sophisticated scattering models are used.

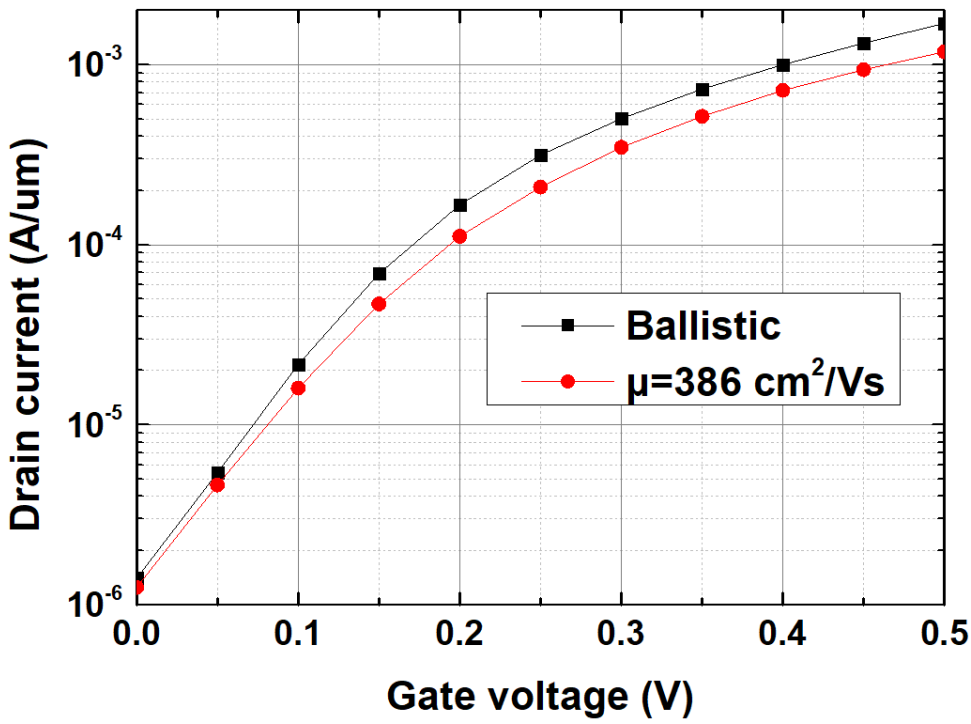


Fig. 31. I-V curve with and without scattering.

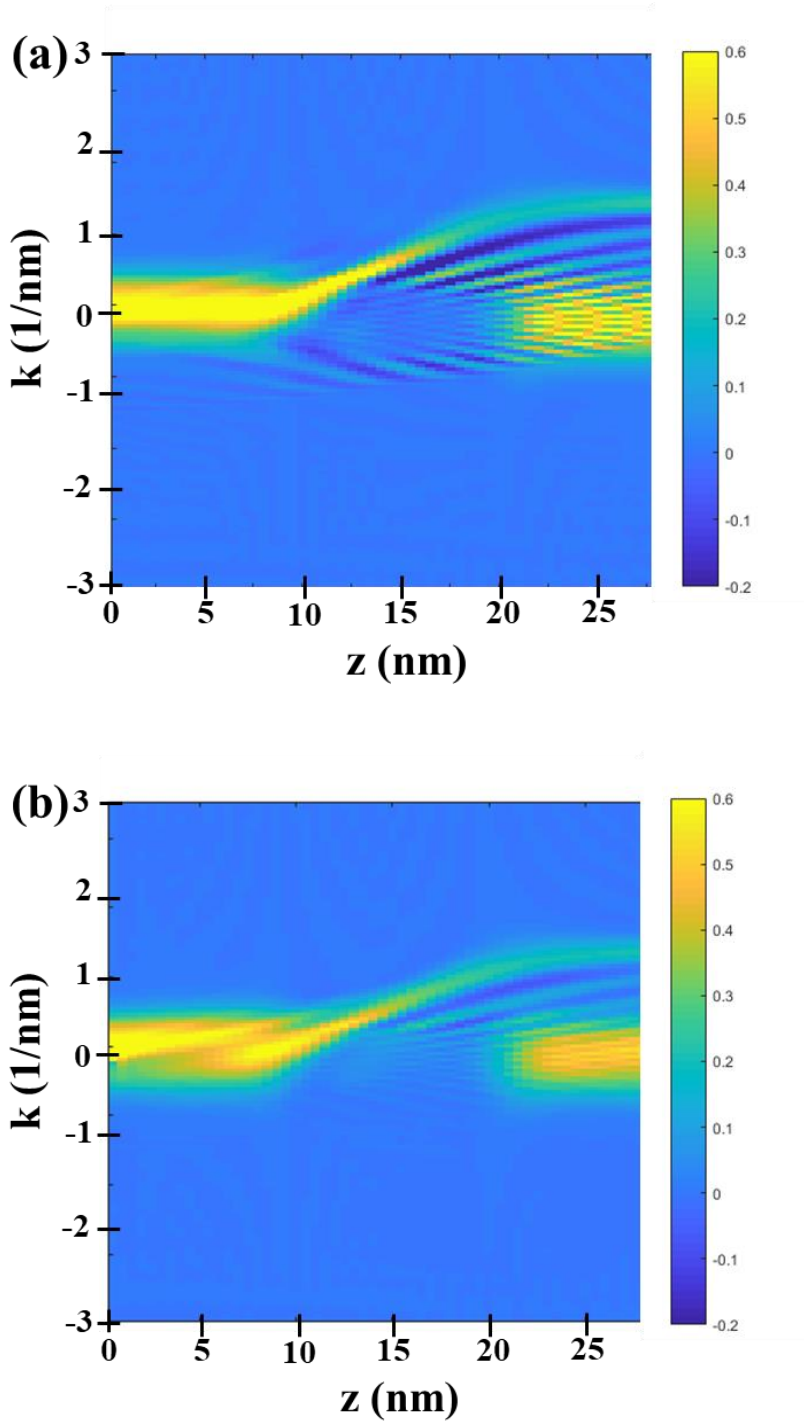


Fig. 32. Wigner function in lowest subband at V_{GS}=0.5V. (a) is when there is no scattering and (b) is when there is scattering.

3. Conclusion

We derive a novel representation of nonlocal potential terms with infinite correlation length with the assumption of ideal contact. Even if the correlation length is not finitely limited, the integral range can be finitely limited through the modification of the equation, and thus it is possible to solve it numerically. Through this, more accurate simulation is possible without uncertainty of the WTE solution due to the finite correlation length. In addition, robust convergence was obtained using fully-coupled scheme in one dimension, and it was confirmed that reliable simulation results were obtained by well showing unique characteristics such as plateaus in RTD. Simulation was applied not only to one-dimensional but also to two-dimensional simulation of double gate structure, device characteristics were confirmed, and quantum transport transient simulation in multi-dimensional simulation was also attempted for the first time. Unlike NEGF and PME, not only steady-state simulation but also transient simulation are possible, and since the Newton-Raphson method is used, the accurate linear response of the equation can be calculated, and thus small signal or

noise analysis will be readily possible.

Appendix

A-1. Numerical integration method of the nonlocal potential terms

Assuming a linear potential drop between meshes as shown in Fig. 8, the potential in the device can be expressed in a linear function form for each section.

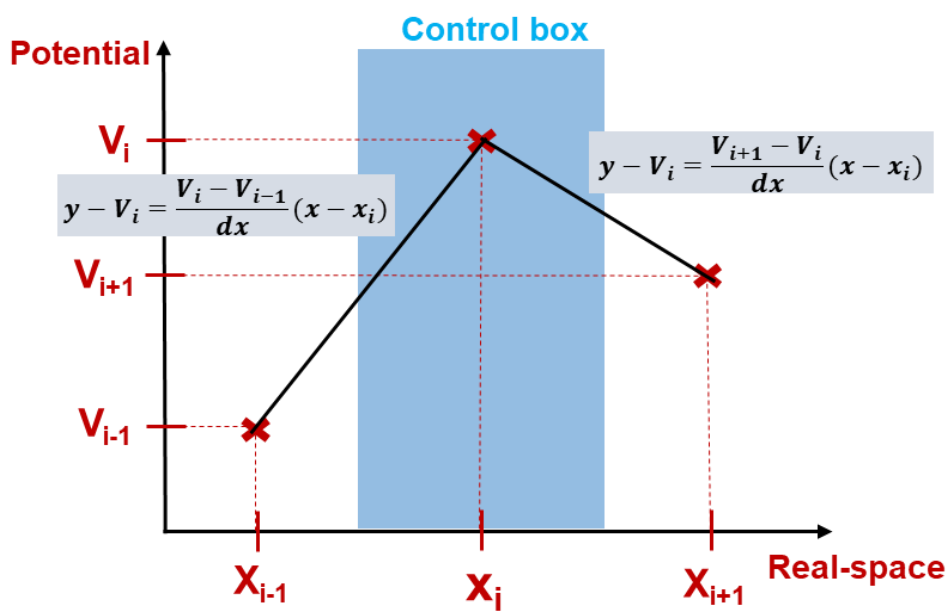


Fig. 33. Assuming a linear potential drop between meshes, the potential can be expressed as a linear function.

Then, we can integrate the nonlocal potential term analytically as follow:

$$\hat{V}(\chi,k) = 2\int_0^{2L} d\zeta \sin(\zeta k)\left[U_{cor}(\chi,\zeta)\right] + 2\frac{U_{ex}}{k}\cos(2Lk)$$

$$= \left[\sum_{A=1}^{N_X} 2\int_{(A-1)dx}^{Adx} d\zeta \sin(\zeta k)\left[U_{cor}(\chi,\zeta)\right]\right] + 2\frac{U_{ex}}{k}\cos(2Lk)$$

$$= \left[\sum_{A=1}^{N_X} 2\int_{(A-1)dx}^{Adx} d\zeta \sin(\zeta k)\left[\begin{array}{l}\frac{V_{i+A}-V_{i+A-1}}{dx}(\zeta - x_{i+A}) + V_{i+A} \\ -\frac{V_{i-A+1}-V_{i-A}}{dx}(\zeta - x_{i-A}) - V_{i-A}\end{array}\right]\right]$$

$$+2\frac{U_{ex}}{k}\cos(2Lk)$$

$$= \left[\sum_{A=1}^{N_X} 2\left|\begin{array}{l}\frac{1}{k}\cos(\zeta k)\left(\begin{array}{l}\frac{V_{i+A}-V_{i+A-1}}{dx}x_{i+A} - \\ \frac{V_{i-A+1}-V_{i-A}}{dx}x_{i-A} + V_{i-A} - V_{i+A}\end{array}\right) \\ +\left(\frac{V_{i+A}-V_{i+A-1}}{dx} - \frac{V_{i-A+1}-V_{i-A}}{dx}\right)\left(\frac{\sin(\zeta k)-\zeta k\cos(\zeta k)}{k^2}\right)\end{array}\right|_{A-1}^{A}\right]$$

$$+2\frac{U_{ex}}{k}\cos(2Lk)$$

(40)

This is difficult to integrate in k−space, so we assume that nonlocal potential terms are uniform in k−space. Therefore, when the finite correlation length is used, the longer the correlation length is, the shorter the period of the sine cos function is, so that the nonlocal potential term changes rapidly in the k−space and a non−physical

result is obtained. Fig. 34 shows the I-V curve according to the number of meshes in k-space when $Lc=4L$. Since the mesh size in k-space decreases as the number of meshes increases, non-physical results due to the rapidly changing nonlocal potential term do not appear.

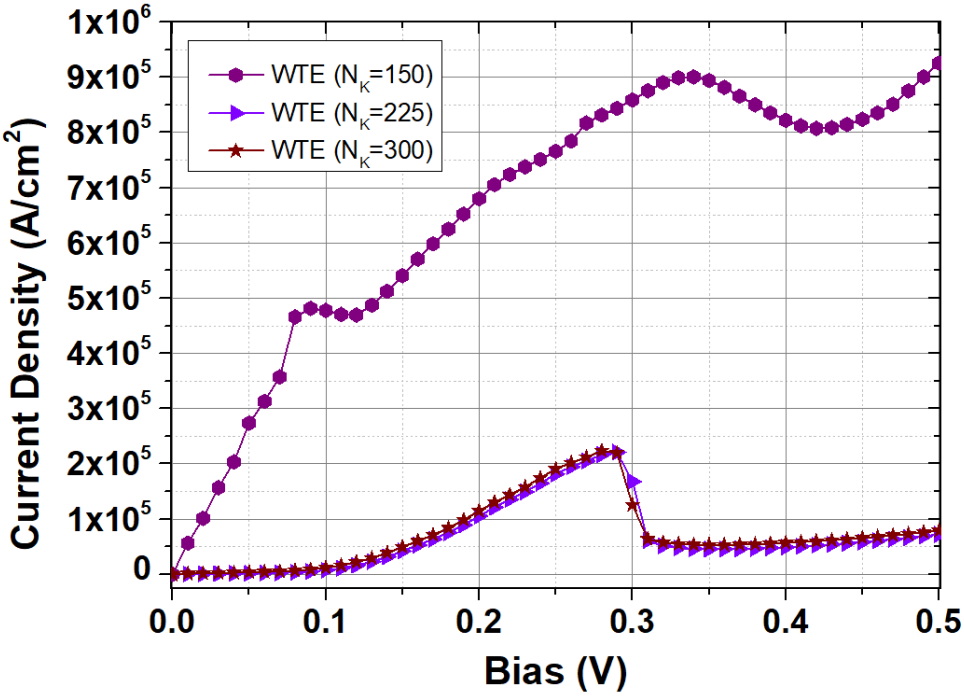


Fig. 34. I-V curve according to the number of meshes in k-space when $Lc=4L$.

A-2. 2D electron density and electric potential results

For better understanding, 2D electron density and potential values are shown when the gate length is 10 nm, the body thickness is 3 nm, and the oxide thickness is 1 nm. It can be confirmed that the electron density and potential energy in the channel are increased in the on-state.

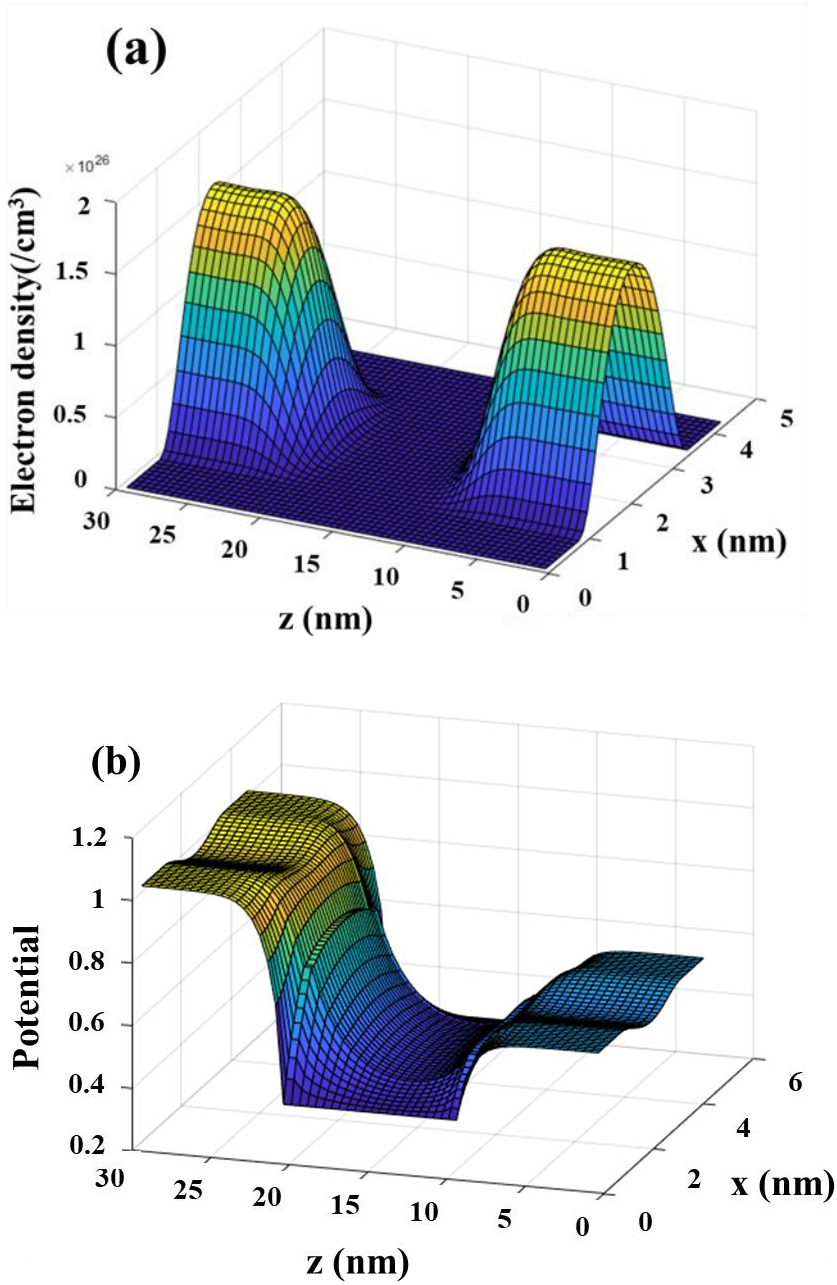


Fig. 35. (a) Electron density and (b) potential when gate length is 10nm and body thickness is 3nm. Drain voltage is 0.4V and gate voltage is 0V (off-state).

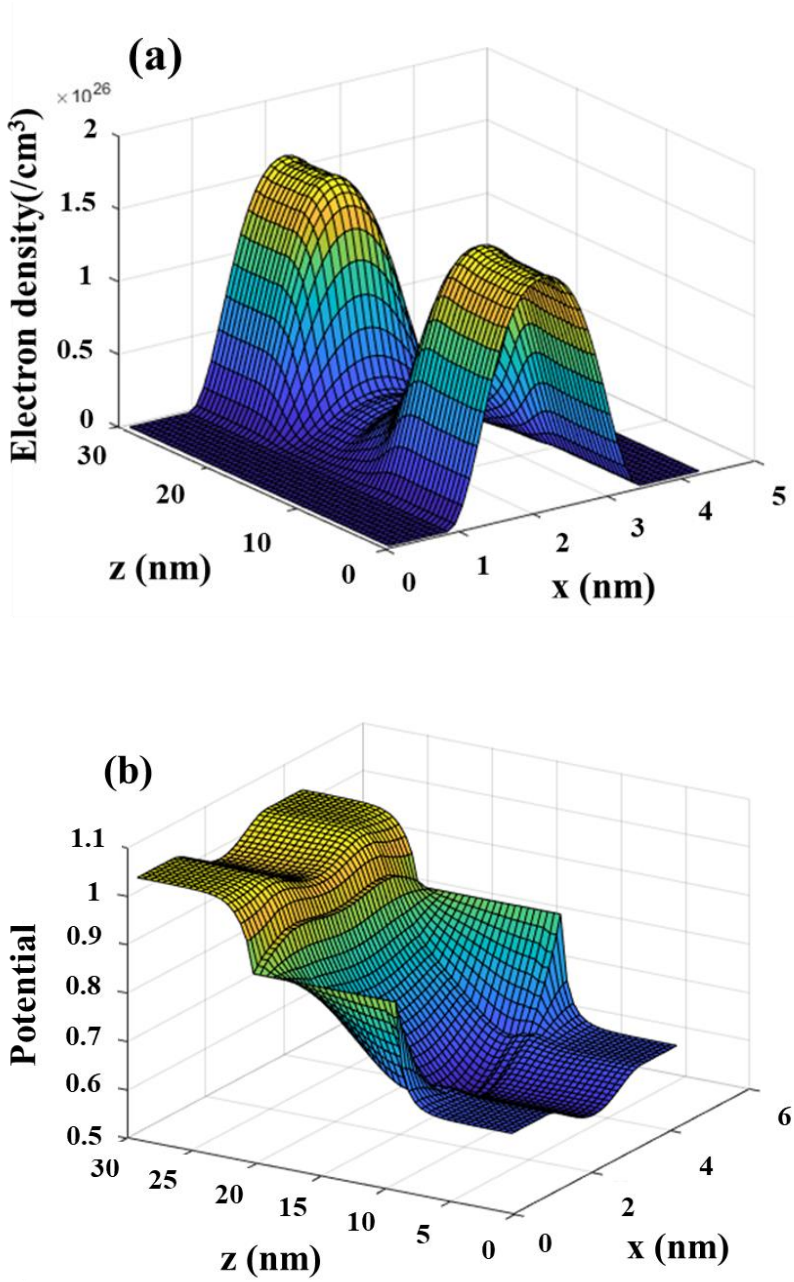


Fig. 36. (a) Electron density and (b) potential when gate length is 10nm and body thickness is 3nm. Drain voltage is 0.4V and gate voltage is 0.5V (on-state).

A-3. Wigner function for each subband

The Wigner function in the on/off state is shown in figures. The figures show that the second or more subbands in each valley show a low Wigner function and thus do not significantly affect the solution. Also, it can be seen that the 3rd valley (t, t, l) shows the highest Wigner function value and has the greatest influence.

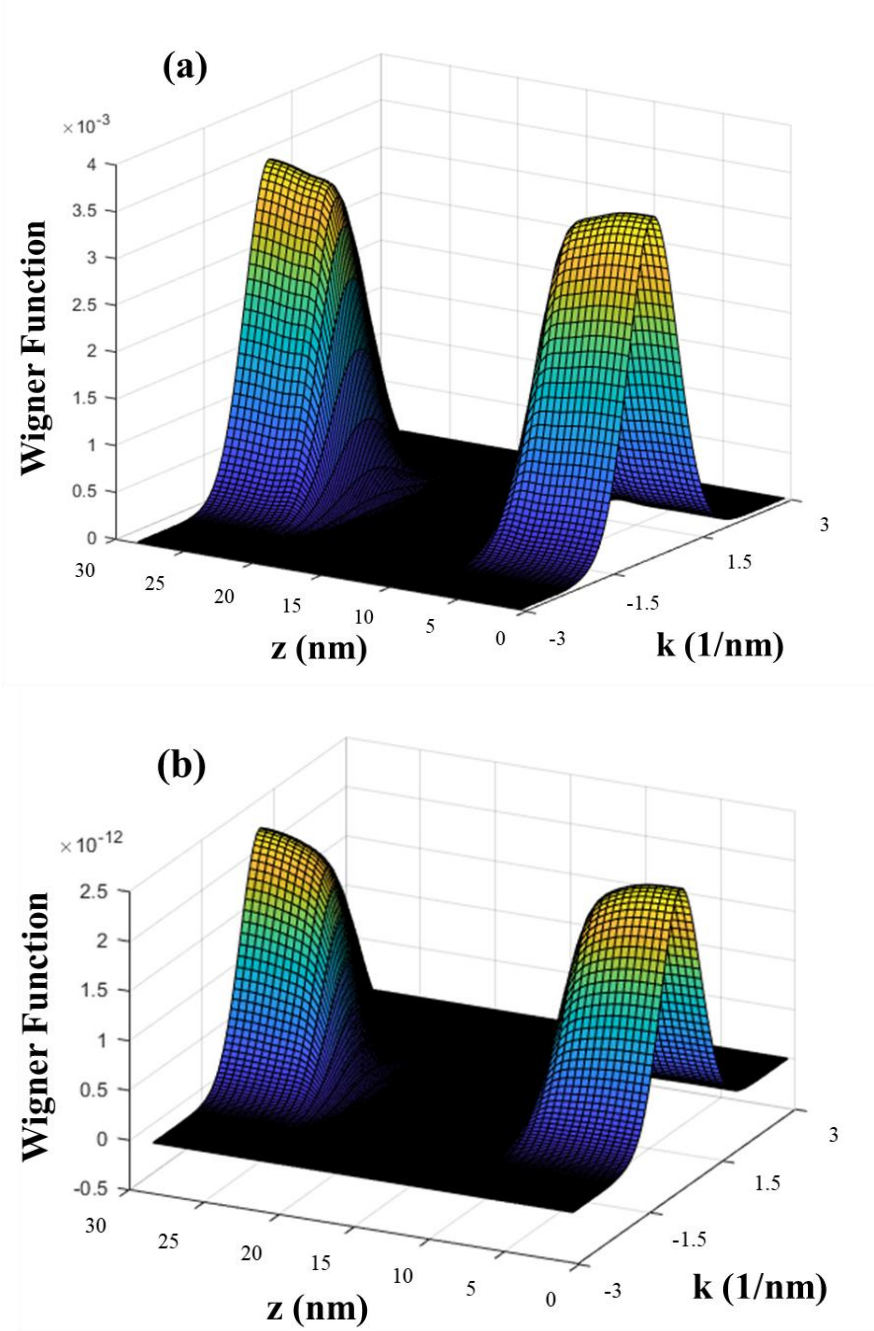
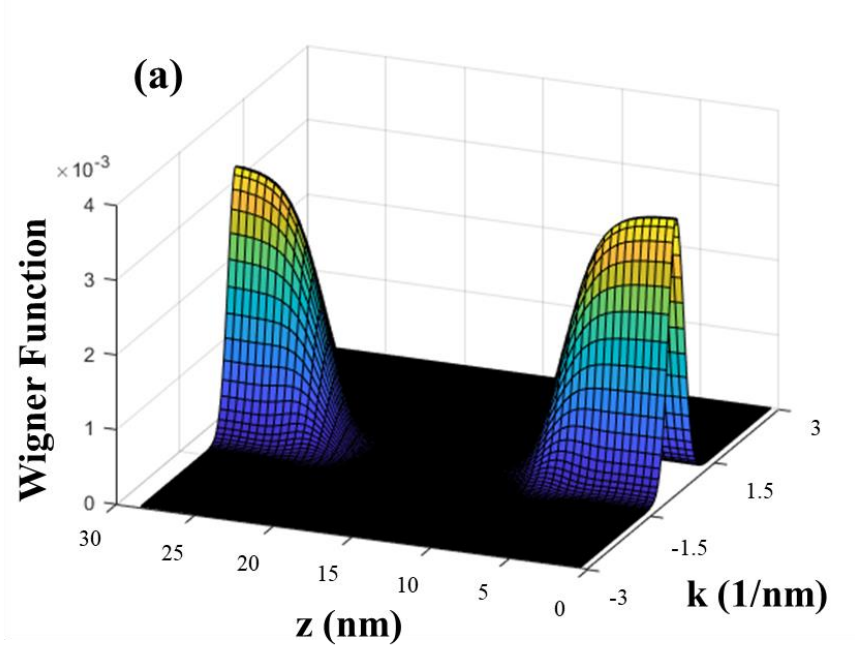


Fig. 37. Wigner function at (a) first and (b) second subband of the 1st valley (l, t, t) when gate length is 10nm and body thickness is 3nm. Drain voltage is 0.4V and gate voltage is 0V (off-state).



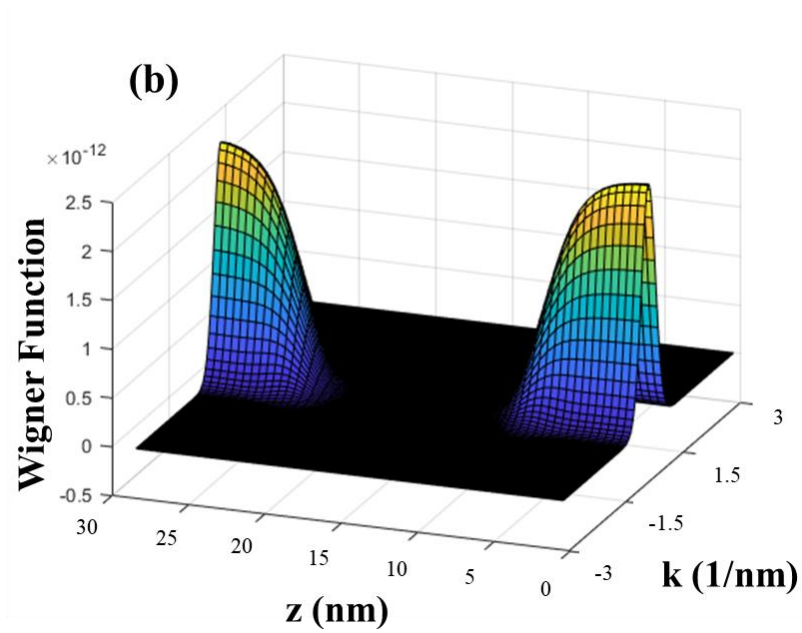


Fig. 38. Wigner function at (a) first and (b) second subband of the 2nd valley (t, l, t) when gate length is 10nm and body thickness is 3nm. Drain voltage is 0.4V and gate voltage is 0V (off-state).

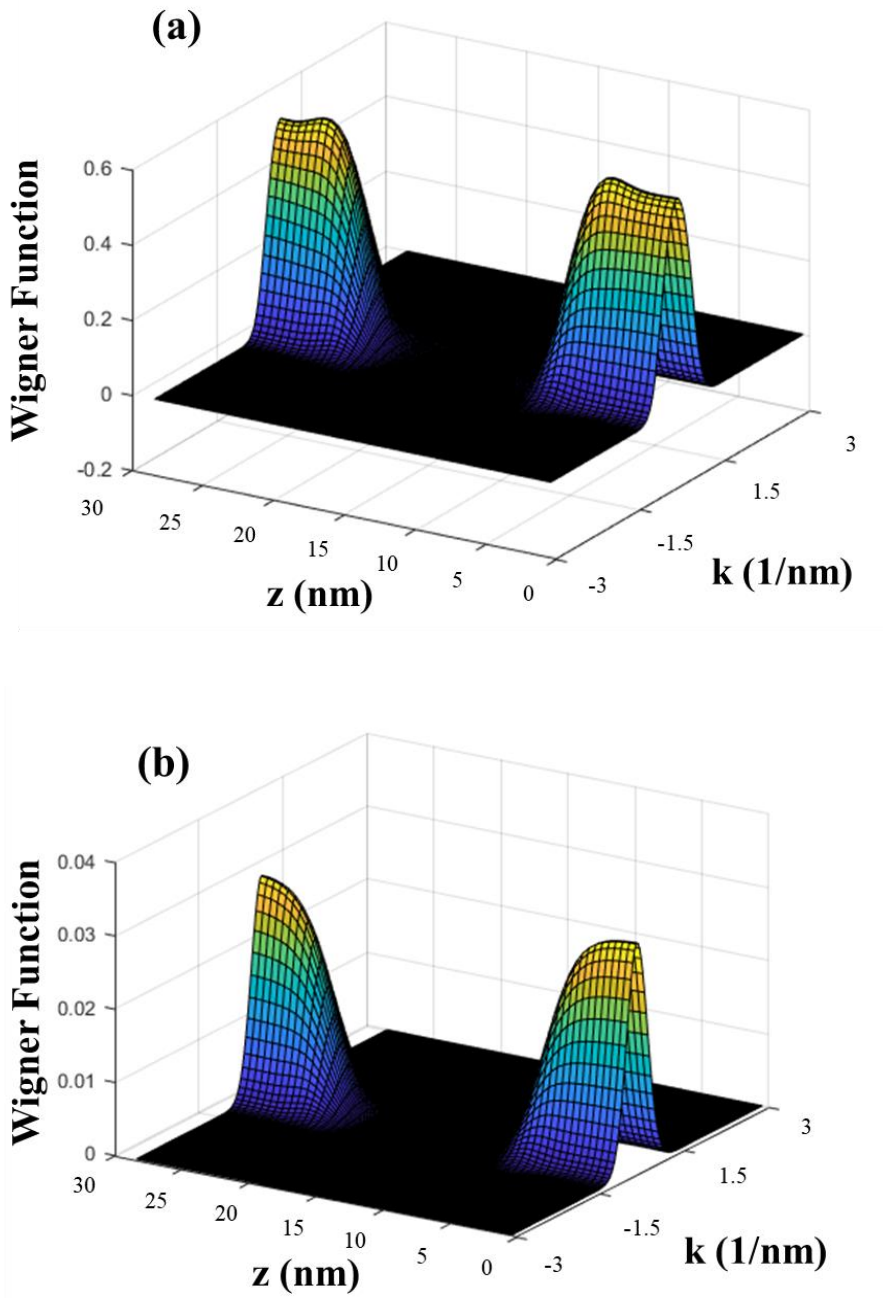


Fig. 39. Wigner function at (a) first and (b) second subband of the 3rd valley (t, t, l) when gate length is 10nm and body thickness is 3nm. Drain voltage is 0.4V and gate voltage is 0V (off-state).

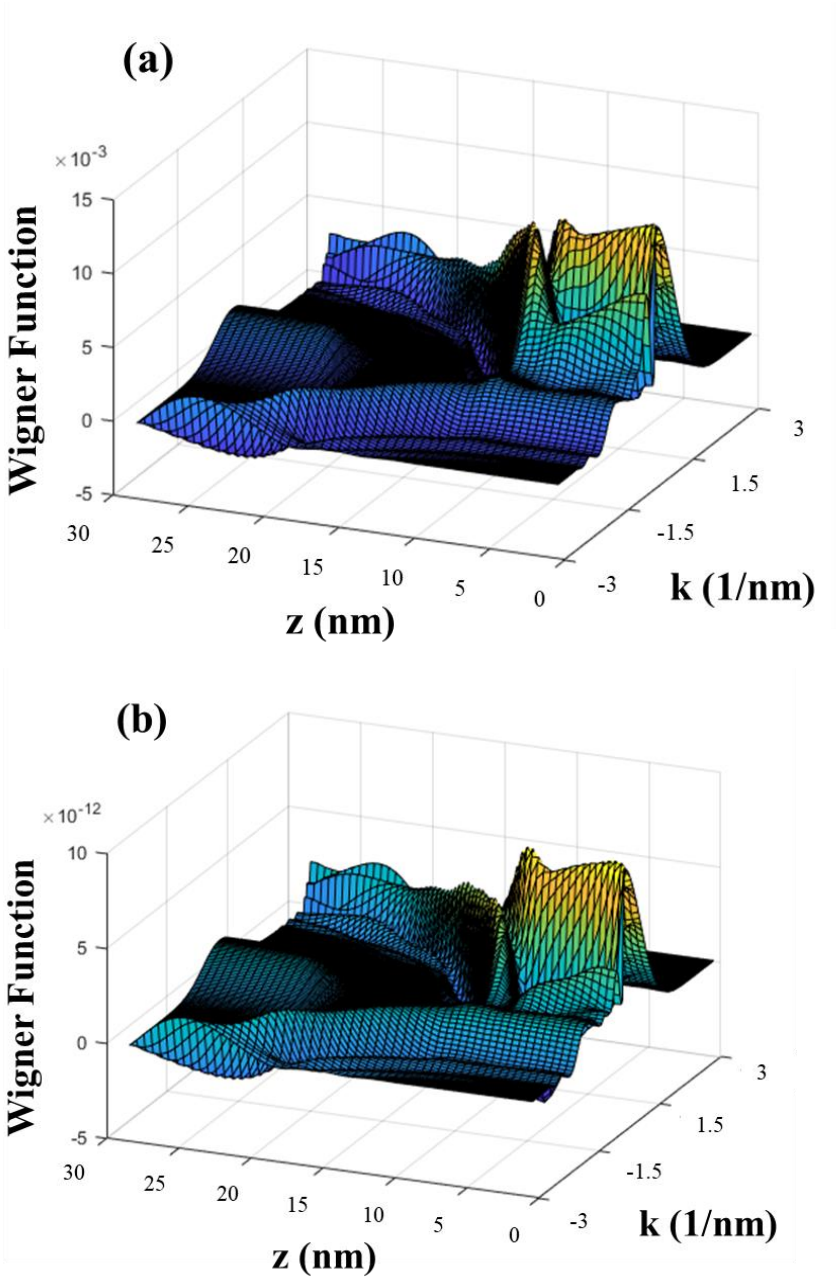


Fig. 40. Wigner function at (a) first and (b) second subband of the 1st valley (l, t, t) when gate length is 10nm and body thickness is 3nm. Drain voltage is 0.4V and gate voltage is 0.5V (on-state).

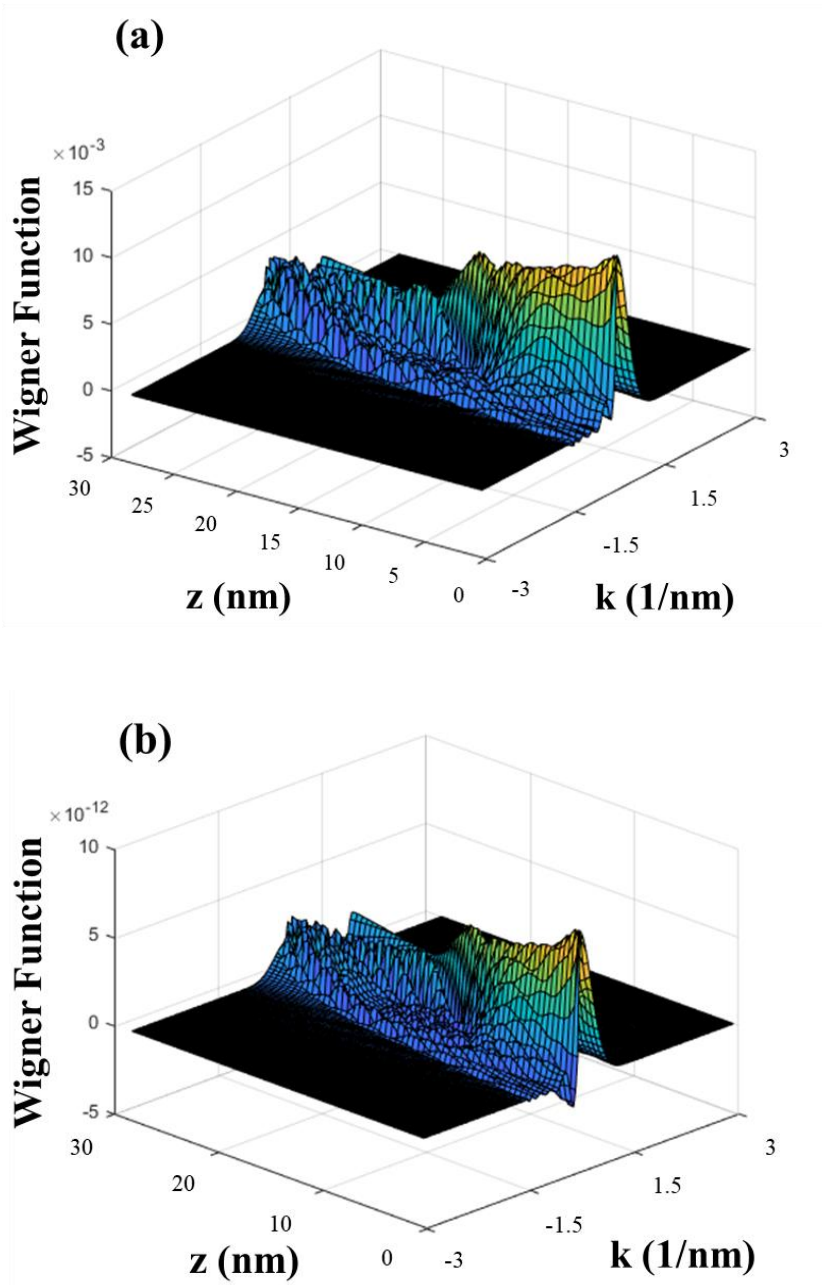


Fig. 41. Wigner function at (a) first and (b) second subband of the 2dn valley (t, l, t) when gate length is 10nm and body thickness is 3nm. Drain voltage is 0.4V and gate voltage is 0.5V (on-state).

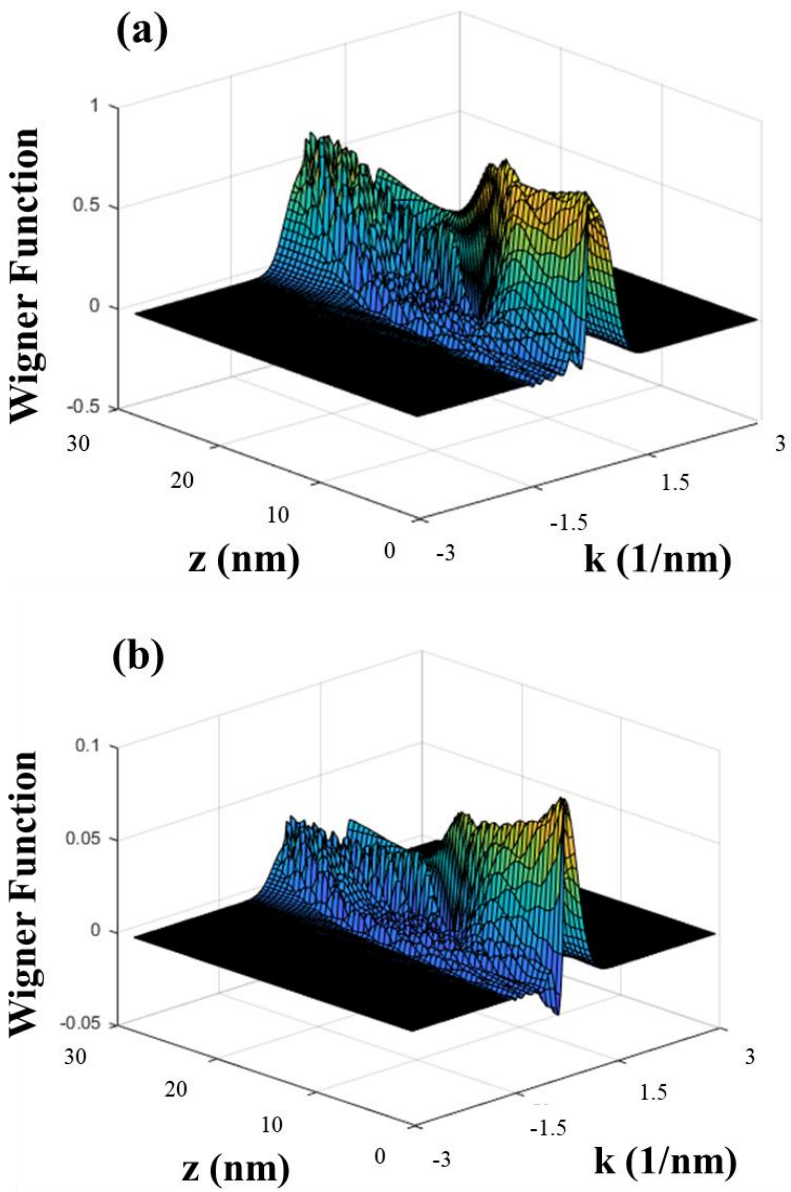


Fig. 42. Wigner function at (a) first and (b) second subband of the 3rd valley (t, t, l) when gate length is 10nm and body thickness is 3nm. Drain voltage is 0.4V and gate voltage is 0.5V (on-state).

References

[1] S. Jin, T. Tang and M. V. Fischetti, "Simulation of Silicon Nanowire Transistors Using Boltzmann Transport Equation Under Relaxation Time Approximation," IEEE Transactions on Electron Devices, vol. 55, no. 3, pp. 727-736, Mar. 2008, DOI: 10.1109/TED.2007.913560.

[2] S. -M. Hong and C. Jungemann, "A fully coupled scheme for a Boltzmann-Poisson equation solver based on a spherical harmonics expansion," Journal of Computational Electronics 8, pp. 225-241, Oct. 2009, DOI: https://doi.org/10.1007/s10825-009-0294-y.

[3] S. Cha and S. -M. Hong, "Theoretical Study of Electron Transport Properties in GaN-Based HEMTs Using a Deterministic Multi-Subband Boltzmann Transport Equation Solver," IEEE Transactions on Electron Devices, vol. 66, no. 9, pp. 3740-3747, Sept. 2019, DOI: 10.1109/TED.2019.2926857.

[4] C. Jungemann, A. T. Pham, B. Meinerzhagen, C. Ringhofer, and M. Bollhofer, "Stable discretization of the Boltzmann equation based on spherical harmonics, box integration, and a maximum entropy dissipation principle," Journal of Applied Physics, Volume 100, Issue 2, July. 2006, DOI: https://doi.org/10.1063/1.2212207.

[5] Jing Wang and M. Lundstrom, "Does source-to-drain tunneling limit the ultimate scaling of MOSFETs?," Digest. International Electron Devices Meeting, 2002, pp. 707-710, DOI: 10.1109/IEDM.2002.1175936.

[6] D. Yadav and D. R. Nair, "Impact of Source to Drain Tunneling on the Ballistic Performance of Si, Ge, GaSb, and GeSn Nanowire

p−MOSFETs," IEEE Journal of the Electron Devices Society, vol. 8, pp. 308−315, 2020, DOI: 10.1109/JEDS.2020.2980633.

[7] K. −H. Kao T. R. Wu, H. −L. Chen, W. −J. Lee, N. −Y. Chen, W. C. −Y. Ma, C. −J. Su, and Y. −J. Lee, "Subthreshold Swing Saturation of Nanoscale MOSFETs Due to Source−to−Drain Tunneling at Cryogenic Temperatures," IEEE Electron Device Letters, vol. 41, no. 9, pp. 1296−1299, Sept. 2020, DOI: 10.1109/LED.2020.3012033.

[8] L. P. Kadanoff and G. Baym, Quantum Statistical Mechanics (Benjamin, New York, 1962), DOI: https://doi.org/10.1201/9780429493218.

[9] E. Wigner, “On the Quantum Correction for Thermodynamic Equilibrium”, Physical Review, 40, 749, June. 1932, DOI: https://doi.org/10.1103/PhysRev.40.749.

[10] W. R. Frensley, “Boundary Conditions for Open Quantum Systems Driven Far from Equilibrium,” Reviews of Modern Physics. 63, 215, July. 1990, DOI: https://doi.org/10.1103/RevModPhys.62.745.

[11] M. V. Fischetti, “Theory of electron transport in small semiconductor devices using the Pauli master equation,” J. Appl. Phys. 83, 270, 1998, DOI: https://doi.org/10.1063/1.367149.

[12] M. V. Fischetti, “Master−equation approach to the study of electronic transport in small semiconductor devices,” Phys. Rev. B 59, 4901, 1999, DOI: https://doi.org/10.1103/PhysRevB.59.4901.

[13] M. A. Stettler et al., "Industrial TCAD: Modeling Atoms to Chips," in IEEE Transactions on Electron Devices, vol. 68, no. 11, pp. 5350−5357, Nov. 2021, DOI: 10.1109/TED.2021.3076976.

[14] P. Mahdi, “Numerical Study of Quantum Transport in Carbon Nanotube−based Transistors,” dissertation, Institute for

Microelectronics, Vienna University of Technology, Vienna, 2007.
[15] S. Jin, Y. J. Park, and H. S. Min, “A three-dimensional simulation of quantum transport in silicon nanowire transistor in the presence of electron-phonon interactions,” Journal of Applied Physics 99, 123719, 2006, DOI: https://doi.org/10.1063/1.2206885.
[16] M. Luisier and G. Klimeck, “Atomistic full-band simulation of silicon nanowire transistors: Effects of electron-phonon scattering,” Physical Review B 80, 155430, 2009, DOI: https://doi.org/10.1103/PhysRevB.80.155430.
[17] T Gunst, T. Markussen, M. L. Palsgaard, K. Stokbro, and M. Brandbyge, “First-principles electron transport with phonon coupling: Large scale at low cost,” Physical Review B 96, 161404, 2017, DOI: https://doi.org/10.1103/PhysRevB.96.161404.
[18] M. R. Hirsbrunner, T. M. Philip, B. Basa, Y. Kim, M. J. Park, and M. J. Gilbert, “A review of modeling interacting transient phenomena with non-equilibrium Green functions,” Reports on Progress in Physics 82, 046001, March. 2019, DOI: https://doi.org/10.1088/1361-6633/aafe5f.
[19] S-M. Hong, P-H. Ahn, “AC NEGF Simulation of Nanosheet MOSFETs,” 2020 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Sep. 2020 pp. 289-292, DOI: 10.23919/SISPAD49475.2020.9241656.
[20] P. -H. Ahn and S. -M. Hong, "Quantum Transport Simulation with the First-Order Perturbation: Intrinsic AC Performance of Extremely Scaled Nanosheet MOSFETs in THz Frequencies," 2021 IEEE International Electron Devices Meeting (IEDM), 2021, pp. 18.4.1-18.4.4, DOI: 10.1109/IEDM19574.2021.9720598.
[21] P. B. Vyas, M. L. Van de Put, and M. V. Fischetti, “Master-Equation Study of Quantum Transport in Realistic Semiconductor

Devices Including Electron-Phonon and Surface-Roughness Scattering," Physical Review Applied 13, 014067, Jan. 2020, DOI: https://doi.org/10.1103/PhysRevApplied.13.014067.

[22] M. Hillery, R. F. O' Connell, M. O. Scully, and E. P. Wigner, "Distribution Functions in Physics: Fundamentals", Physics Reports, vol 106, issue 3, pp 121-167, April. 1984, DOI: https://doi.org/10.1016/0370-1573(84)90160-1.

[23] C. Jacoboni and P. Bordone, "The Wigner-function approach to non-equilibrium electron transport," Reports on Progress in Physics, 67, 1033-1071, 2004, DOI: 10.1088/0034-4885/67/7/R01.

[24] M. Nedjalkov, D. Vasileska, D. K. Ferry, C. Jacoboni, C. Ringhofer, I. Dimov, and V. Palankovski, "Wigner transport models of the electron-phonon kinetics in quantum wires," Physical Review B 74, 035311, 2006, DOI: 10.1103/PhysRevB.74.035311.

[25] F. Rossi, C. Jacoboni, and M. Nedjalkov, "A Monte Carlo solution of the Wigner transport equation," Semiconductor Science and Technology, 9 934, 1994.

[26] D. Querlioz, J Saint-Martin, V.-N. Do, A. Bournel, and P. Dollfus, "A Study of Quantum Transport in End-of-Roadmap DG-MOSFETs Using a Fully Self-Consistent Wigner Monte Carlo Approach," IEEE Transactions on Nanotechnology, Vol 5, No 6, November 2006, DOI: 10.1109/TNANO.2006.883477.

[27] S. Barraud, "Dissipative quantum transport in silicon nanowires based on Wigner transport equation," Journal of Applied Physics, vol 110, issue 9, Nov. 2011, DOI: https://doi.org/10.1063/1.3654143.

[28] H. Jiang, W. Cai, and R. Tsu, "Accuracy of the Frensley inflow boundary condition for Wigner equations in simulating resonant

tunneling diodes," Journal of Computational Physics, vol. 230, issue 5, pp 2031–2044, March 2011, DOI: https://doi.org/10.1016/j.jcp.2010.12.002.

[29] A. S. Costolanski and C. T. Kelley, "Efficient Solution of the Wigner–Poisson Equations for Modeling Resonant Tunneling Diodes," in IEEE Transactions on Nanotechnology, vol. 9, no. 6, pp. 708–715, Nov. 2010, DOI: 10.1109/TNANO.2010.2053214.

[30] Y. Yamada, H. Tsuchiya and M. Ogawa, "Quantum Transport Simulation of Silicon–Nanowire Transistors Based on Direct Solution Approach of the Wigner Transport Equation," in IEEE Transactions on Electron Devices, vol. 56, no. 7, pp. 1396–1401, July 2009, DOI: 10.1109/TED.2009.2021355.

[31] S. Hong and J. Jang, "Transient Simulation of Semiconductor Devices Using a Deterministic Boltzmann Equation Solver," IEEE Journal of the Electron Devices Society, vol. 6, pp. 156–163, 2018, DOI: 10.1109/JEDS.2017.2780837.

[32] B. P. Leonard, "A stable and accurate convective modelling procedure based on quadratic upstream interpolation," Computer Methods in Applied Mechanics and Engineering, vol 19, issue 1, pp 59–98, June 1979, DOI: https://doi.org/10.1016/0045–7825(79)90034–3.

[33] R. Rosati, F. Dolcini, R. C. Iotti, and F. Rossi, "Wigner–function formalism applied to semiconductor quantum devices: Failure of the conventional boundary condition scheme," Physical Review B 88, 035401, 2013, DOI: 10.1103/PhysRevB.88.035401.

[34] R. C. Iotti, F. Dolcini, and F. Rossi, "Wigner–function formalism applied to semiconductor quantum devices: Need for nonlocal scattering models," Physical Review B 96, 115420, 2017, DOI: https://doi.org/10.1103/PhysRevB.96.115420.

[35] H. K. Gummel, “A self-consistent iterative scheme for one-dimensional steady state transistor calculations,” IEEE transactions on Electron Devices, vol 11, issue 10, pp 455-465, Oct 1964, DOI: 10.1109/T-ED.1964.15364.

[36] K. L. Jensen and F. A. Buot, “Numerical simulation of intrinsic bistability and high-frequency current oscillations in resonant tunneling structures,” Physical Review Letters. 66, 1078, Feb. 1991, DOI: https://doi.org/10.1103/PhysRevLett.66.1078.

[37] P. Zhao, H. L. Cui, and D. L. Woolard, “Dynamical instabilities and I-V characteristics in resonant tunneling through double-barrier quantum well systems,” Physical Review B. 63, 075302, Jan. 2001, DOI: https://doi.org/10.1103/PhysRevB.63.075302.

[38] B. A. Biegel, “Wigner Function Simulation of Intrinsic Oscillations, Hysteresis, and Bistability in Resonant Tunneling Structures,” Porc. SPIE 3277, Ultrafast Phenomena in Semiconductors, April. 1998, DOI: https://doi.org/10.1117/12.306152.

[39] B. A. Biegel and J. D. Plummer, “Comparison of self-consistency iteration options for the Wigner function method of quantum device simulation,” Physical Review B. 54, 8070, Sept. 1996, DOI: https://doi.org/10.1103/PhysRevB.54.8070.

[40] B. A. Biegel and J. D. Plummer, "Applied bias slewing in transient Wigner function simulation of resonant tunneling diodes," IEEE Transactions on Electron Devices, vol. 44, no. 5, pp. 733-737, May 1997, DOI: 10.1109/16.568033.

[41] A. S. Costolanski and C. T. Kelley, “Efficient Solution of the Wigner-Poisson Equations for Modeling Resonant Tunneling Diodes,” IEEE transactions on Nanotechnology, vol. 9, no. 6, Nov 2010, DOI: 10.1109/TNANO.2010.2053214.

[42] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom, “nanoMOS 2.5: A Two-Dimensional Simulator for Quantum Transport in Double-Gate MOSFETs,” IEEE Transactions on Electron Devices, vol. 50, no. 9, 2003, DOI: 10.1109/TED.2003.816524.

초 록

본 연구에서는 무한한 상관 길이를 가지는 위그너 수송 방정식의 새로운 수치해석적 풀이법을 제시하였다. 최대 상관 길이가 한정된 값으로 제한되지 않기 때문에, 시뮬레이션 결과에 불확실성이 발생하지 않으며, 제안된 표현법에서는 Wigner-Weyl 변환이 unitary하다. 일반적이고 효율적인 시뮬레이션을 위해, 위그너 수송 방정식을 푸아송 방정식과 유한 체적법과 뉴턴-랩슨 방식을 통해 self-consistent하게 풀었다. 제안된 모델을 resonant tunneling diode와 double gate MOSFET에 적용하여, 산란효과를 고려한 동적 그리고 정적 시뮬레이션 결과를 보여주었다.

주요어 : 위그너 방정식, 양자전송, 티캐드 (TCAD) 시뮬레이션

학번 : 2016-20866