Fast Perturbative Treatment for Efficient Nano-Scale Device Simulation Based on Bridge-Function Pseudo-Spectral Method

Yuta Saito, Hiroki Fujikawa, Satofumi Souma, and Matsuto Ogawa
Department of Electrical and Electronic Engineering, Kobe University
1-1 Rokkodai, Nada, Kobe, 657-8501, JAPAN
Email: 111t225t@stu.kobe-u.ac.jp / ogawa@eedept.kobe-u.ac.jp

***Abstract*—A fast perturbation treatment implemented into the bridge-function pseudospectral method (BPSM) is discussed for 3D nano-scaled device simulation. It is shown that the present method, that is, a fast uncoupled mode-space (FUMS) BPSM approach gives enough accurate results compared with the conventional uncoupled BPSM for a Si nano-wire FET structure. In addition, further detailed analysis on the accuracy, speed, and scalability shows the present approach is 40 times faster than the conventional BPSM on a same platform.**

I. Introduction

Numerical solutions of the 3D Schrödinger equation, Poisson's equation, and non-equilibrium Green's function (NEGF) are often required for nano-scaled device design and understanding device physics. They are usually performed numerically with a finite difference method (FDM) or a finite element method (FEM). However, as the dimension of the devises becomes lower, e.g. tri-gate MOSFETs, gate-all-around FETs, and quantum dot devices, which are comprised of quantum wire structures or quantum dots, FDM or FEM becomes inappropriate in terms of computational costs and time. This is because the FDM/FEM, in general, requires many mesh points to achieve high accuracy. To solve this problem, we have developed an effective and powerful method, bridge-function pseudo-spectral method (BPSM) [1], which namely achieves much higher accuracy with much less mesh points than the conventional FDM/FEM.

In other literature, in order to reduce the size of the matrix and the computational time in the conventional FDM, Ren *et al*. have used an uncoupled mode space (UMS) approach [2], [3]. This approach is a kind of a separation of variables method for the solution of partial differential equations, where the Schrödinger equation is solved in the plane perpendicular to the transport direction, then the subband energies and eigenfunctions are calculated. The wave function is expanded in the subbands and the resulting transport equation is simply a one-dimensional equation.

Although the UMS approach is also adopted for the BPSM, it requires to solve the 2D Schrödinger equation-1D NEGF calculation repeatedly on each cross-section of the device under consideration. Because of that, it is found that the computational costs and time rapidly increase as the number of longitudinal mesh points increase. In the present study, we have adopted a perturbative treatment, namely, a fast uncoupled mode space (FUMS) approach into our previous BPSM. In the FUMS approach, the subband profile is approximated up to the first order by considering the change from the average of potential profile as the perturbation. Here, we have only to solve the non-perturbative Schrödinger equation once in the coupled 2D Schrödinger equation-1D NEGF calculation, even when the number of longitudinal mesh points increases.

II. Theory

We have adopted the FUMS approach into our previous BPSM. The BPSM is a method to solve a differential equation such as Schrödinger equation, non-equilibrium Green's function, or Poisson's equation, which enables us to calculate much more accurately with much less computational time or cost [1]. In the BPSM, the region under consideration $[0, L]$ is divided into sub-regions ($i = 0, 1, \cdots, I$) and each sub-region is spanned by a basis set (Fig. 1) [4]. The BPSM is composed of both the orthonormal basis functions to expand physical quantities and the Gauss-Lobatto (GL) quadrature [5]. The basis functions (Fig.2) in each region are comprised of a set of Lagrange polynomials, one of which is called a bridge-function, a connection of two Lagrange polynomials. Mathematically the 1D orthonormal basis functions are expressed as

$$\chi_{i,m}(x) = \begin{cases} \dfrac{f_{i,n}(x) + f_{i+1,1}(x)}{\sqrt{w_n^i + w_1^{i+1}}} & (m = 1) \\ (\text{Bridge function}) & \\ \dfrac{f_{i,m}(x)}{\sqrt{w_m^i}} & (m \neq 1), \end{cases} \tag{1}$$

where $f_{i,m}(x)$ is the $(n-1)$th order Lagrange interpolation polynomial and w_m^i denotes the weights of the GL quadrature [5], respectively. The continuity of electrical flux density and probability current density at each interface or boundary between neighboring regions are naturally guaranteed by the bridge-function (solid lines in Fig.2), which enables us to take the boundary conditions into account in much easier manner than the conventional PSMs. In addition, the FUMS is easily implemented in the BPSM formalism.

 ISBN 978-0-615-71756-2

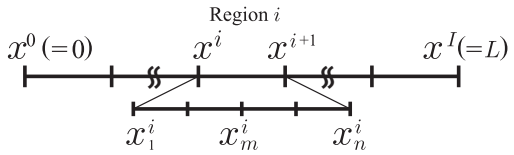

Fig. 1. The device area under consideration $[0, L]$ is divided into regions or elements in the present BPSM. I is the total numbers of the nodal points in x direction, n the numbers of local basis functions and x_m^i's are the Gauss-Lobatto points, respectively.

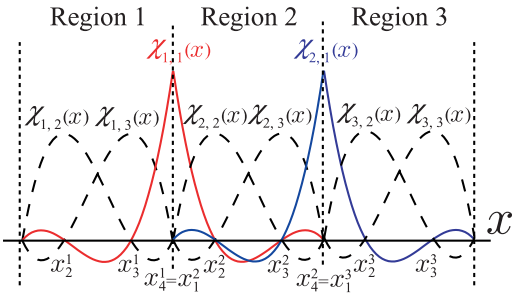

Fig. 2. 1D basis functions $\chi_{i,m}(x)$ ($I = 3$ and $n = 3$), where i indicates the region, and m the basis. The solid lines are the *bridge functions*, which ensure the continuity of any physical quantities between the adjacent regions.

Although the combination of the BPSM and the UMS approach has been adopted for nano-scaled device simulation, it still requires large computational time. This is because it needs to solve the 2D Schödinger equation-1D NEGF calculation repeatedly on each cross-section of the device under consideration to obtain the eigenenergy or subband energy and wave function of each cross-section.

In contrast, in the FUMS, it is assumed that the shape of the cross-section of the device does not change drastically along the longitudinal direction, so that the subband profile can be treated within a first-order perturbation problem by considering the cross-sectional potential change as a perturbation from the average of confinement potential. That is, we solve the non-perturbative 2D Schrödinger equation in the average cross-sectional potential and then obtain the eigenenergy or subband energy $E_k^{\rm sub}$ and wave function $\xi_k(y,z;x)$ of each cross-section along the channel as the solution of a perturbation problem. The FUMS approach, therefore, requires to solve the non-perturbative Schrödinger equation only once in the coupled 2D Schrödinger equation-1D NEGF calculation, even when the number of longitudinal mesh points increase. The non-perturbative Schrödinger equation is expressed as

$$\left[-\frac{\hbar^2}{2}\frac{\partial}{\partial y}\left(\frac{1}{m_y^*(y,z)}\frac{\partial}{\partial y}\right)-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\left(\frac{1}{m_z^*(y,z)}\frac{\partial}{\partial z}\right) +\overline{U_{\rm T}}(y,z)\right]\overline{\xi_k}(y,z)=\overline{E_k^{\rm sub}}\;\overline{\xi_k}(y,z). \quad (2)$$

where variables with bars $(-)$ are non-perturbative physical quantites, namely, $\overline{\xi_k}(y,z)$ is the wave function in the cross-section, $\overline{E_k^{\rm sub}}$ is the eigenenergy of each subband, respectively. $\overline{U_{\rm T}}(y,z)$ is the cross-sectional potential averaged along the longitudinal direction, which is defined as

$$\overline{U_{\rm T}}(y,z)=\frac{1}{L_x}\int_0^{L_x}U(x,y,z)dx, \quad (3)$$

where L_x is the longitudinal length of the device. Since $U(x,y,z)$ changes gradually along the longitudinal direction, the change can be expressed $U_{\rm L}(x,y,z)$ and the potential profile is approximated as

$$U(x,y,z)=\overline{U_{\rm T}}(y,z)+U_{\rm L}(x,y,z). \quad (4)$$

The first-order correction $U_{\rm L}(x,y,z)$ can be viewed as a perturbation added to the non-perturbative problem. The non-perturbative Hamiltonian H_0 and the first-order correction H' are expressed as follows,

$$H = H_0 + H', \quad (5)$$

$$H_0 = -\frac{\hbar^2}{2}\frac{\partial}{\partial y}\left(\frac{1}{m_y^*(y,z)}\frac{\partial}{\partial y}\right)-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\left(\frac{1}{m_z^*(y,z)}\frac{\partial}{\partial z}\right) +\overline{U_{\rm T}}(y,z), \quad (6)$$

$$H' = U_{\rm L}(x,y,z). \quad (7)$$

Then, the zeroth-order equation and the first-order equation are expressed as

$$H_0\overline{\xi_k}(y,z)=\overline{E_k^{\rm sub}}\;\overline{\xi_k}(y,z), \quad (8)$$

$$(H_0+H')\xi_k(y,z;x)=E_k^{\rm sub}\xi_k(y,z;x). \quad (9)$$

Taking advantage of the first order perturbation theory, we obtain both the subband energies and the wave function along the channel as

$$E_k^{\rm sub}(x)=\overline{E_k^{\rm sub}}+\iint\{U(x,y,z)-\overline{U_{\rm T}}(y,z)\}\left|\overline{\xi_k}(y,z)\right|^2dydz, \quad (10)$$

$$\xi_k(y,z;x)=\overline{\xi_k}(y,z)+\sum_{l\neq k}\frac{1}{\overline{E_k^{\rm sub}}-\overline{E_l^{\rm sub}}}\times \left\{\iint\overline{\xi_l}(y,z)\{U(x,y,z)-\overline{U_{\rm T}}(y,z)\}\overline{\xi_k}(y,z)dydz\right\}. \quad (11)$$

In the present study, both the BPSM and the FUMS are simultaneously adopted for the simulation. When the BPSM is applied to the FUMS, the non-perturbative wave function is expanded as

$$\overline{\xi_k}(y,z)=\sum_{i_y'',m_y'',i_z'',m_z''}c^k_{i_y,m_y,i_z,m_z}\chi_{i_y,m_y}(y)\chi_{i_z,m_z}(z) \equiv\sum_{i,m}c^k_{i,m}\chi^{\rm 2D}_{i,m}, \quad (12)$$

where $c^k_{i,m}=c^k_{i_y,m_y,i_z,m_z}\equiv\overline{\xi_k}(y^{i_y}_{m_y},z^{i_z}_{m_z})$ is a function of the weights, $i=(i_y,i_z)$, $m=(m_y,m_z)$, and $\chi^{\rm 2D}_{i,m}=\chi_{i_y,m_y}(y)\chi_{i_z,m_z}(z)$ is the basis functions, respectively. The weak form of the non-perturbative Schrödinger equation can be represented [1] as

$$\sum_{i,m}\left[A^{i,i'}_{m,m'}+\overline{U_{\rm T}}(\boldsymbol{r}^i_m)\delta_{m,m'}\delta_{i,i'}\right]c^k_{i,m}=\overline{E_k^{\rm sub}}c^k_{i,m}, \quad (13)$$

where $\delta_{i,i'}$ and $\delta_{m,m'}$ imply Kronecker's delta, and $\boldsymbol{r}_m^i = (y_{m_y}^{i_y}, z_{m_z}^{i_z})$, respectively. The matrix element $A_{m,m'}^{i,i'}$ in Eq. (13) is written as

$$A_{m,m'}^{i,i'} = (\delta_{i,i'} + \delta_{i,i'\pm 1^*}) \times \iint \left(\frac{\partial \chi_{i',m'}}{\partial y} \frac{1}{m_y^*} \frac{\partial \chi_{i,m}}{\partial y} + \frac{\partial \chi_{i',m'}}{\partial z} \frac{1}{m_z^*} \frac{\partial \chi_{i,m}}{\partial z} \right) dydz. \quad (14)$$

The integration of Eq. (14) is evaluated by the GL quadrature which ensures the higher accuracy than conventional methods. The first term of the integration in Eq. (14) is evaluated as

$$\iint \frac{\partial \chi_{i',m'}}{\partial y} \frac{1}{m_y^*} \frac{\partial \chi_{i,m}}{\partial y} dydz =$$

$$\delta_{i_z,i'_z} \delta_{m_z,m'_z} \sum_{i''_y, m''_y} w_{m''_y}^{i''_y} \frac{1}{m_y^*} \frac{\partial \chi_{i'_y,m'_y}(y_{m''_y}^{i''_y})}{\partial y} \frac{\partial \chi_{i'_y,m'_y}(y_{m''_y}^{i''_y})}{\partial y}. \quad (15)$$

Similarly, $E_k^{\rm sub}(x)$ and $\xi_k(y,z;x)$ are finally expressed within the framework of the BPSM as

$$E_k^{\rm sub}(x) = \overline{E_k^{\rm sub}} + \sum_{i,m} c_{i,m}^k c_{i,m}^k \iint \{U(x,y,z) - \overline{U_{\rm T}}(y,z)\} \chi_{i,m}^2 dydz, \quad (16)$$

$$\xi_k(y,z;x) = \overline{\xi_k}(y,z) + \sum_{l \neq k} \frac{1}{\overline{E_k^{\rm sub}} - \overline{E_l^{\rm sub}}} \times \sum_{i,m} c_{i,m}^l c_{i,m}^k \iint \{U(x,y,z) - \overline{U_{\rm T}}(y,z)\} \chi_{i,m}^2 dydz, \quad (17)$$

respectively. The integrals in Eq. (16) and Eq. (17) are evaluated by the GL quadrature as

$$\iint \{U(x,y,z) - \overline{U_{\rm T}}(y,z)\} \chi_{i,m}^2 dydz = \sum_{i''_y, m''_y, i''_z, m''_z} w_{m''_y}^{i''_y} w_{m''_z}^{i''_z} \{U(x, y_{m''_y}^{i''_y}, z_{m''_z}^{i''_z}) - \overline{U_{\rm T}}(y_{m''_y}^{i''_y}, z_{m''_z}^{i''_z})\} \times \chi_{i_y,m_y}^2(y_{m''_y}^{i''_y}) \chi_{i_z,m_z}^2(z_{m''_z}^{i''_z}). \quad (18)$$

Since the FUMS assumes a gradual potential along the longitudinal direction, it may fail for the abrupt change of the potential. In that case, we may include the mode-coupling approaches [6], [7] into our BPSM to take the abrupt change into account.

III. Result and Discussion

In order to show the superiority of the present method in terms of computational costs and accuracy over conventional UMS approach, we have applied the BPSM-FUMS approach to the analysis of I-V characteristics of a gate-all-around Si nano-wire MOSFET (SiNW FET) shown in Fig. 3 [8], where Schrödinger-Poisson equations and NEGF are solved simultaneously and self-consistently. The orientation dependent effective masses are taken into account in the simulation in the conduction band of Si. Our previous BPSM-UMS approach has been also applied to check and compare the accuracy and the computational costs. Since both approaches utilize the

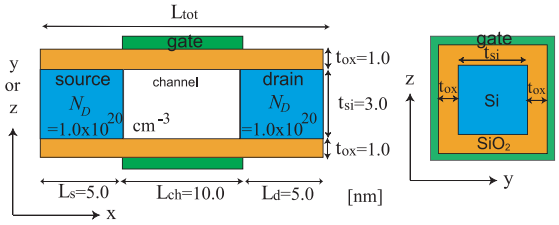

Fig. 3. Model structure of a SiNW FET [8] under investigation. The channel orientation is along $\langle 100 \rangle$.

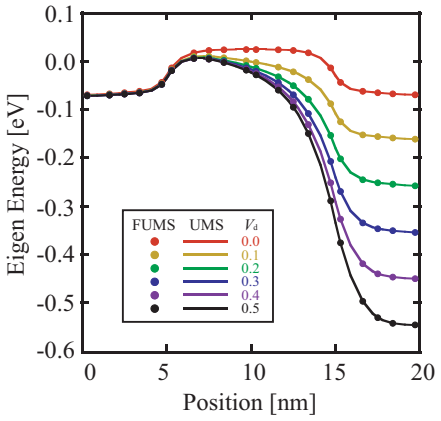

Fig. 4. First subband profile along the channel at $V_{\rm g} = 0.4\ V$ with changing $V_{\rm d}$ from 0.0 to 0.5 V, where the solid lines are obtained by the USM, while dots are by the FUMS.

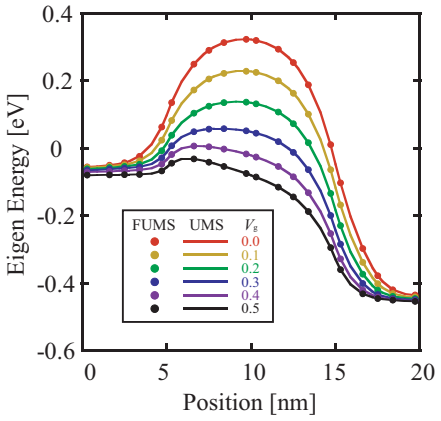

Fig. 5. First subband profile along the channel at $V_{\rm d} = 0.4\ V$ with changing $V_{\rm g}$ from 0.0 to 0.5 V, where the solid lines are obtained by the USM, while dots are by the FUMS.

BPSM in the backbone, we call hereafter the former FUMS and the latter UMS unless any confusion arises.

Figures 4 and 5 show the comparison of the first subband profiles along the channel calculated by the UMS (solid lines) to those calculated by the FUMS (dots) with changing $V_{\rm d}$ (Fig. 4) or changing $V_{\rm g}$ (Fig. 5), respectively. The longitudinal nodal number $N_x = 41$ and the cross-sectional nodal number $N_{yz} = 676$ are fixed, which were checked to give enough accurate results in the UMS. As seen in these figures, the results by the FUMS reproduce those by the UMS even when the bias voltage changes. This feature is maintained as long as the device structure is similar to Fig. 3, that is, the confinement potential varies gradually along the channel, even when the device size is changed. Therefore, the present FUMS enable us to obtain as good an approximation of the subband profiles as produced by the UMS.

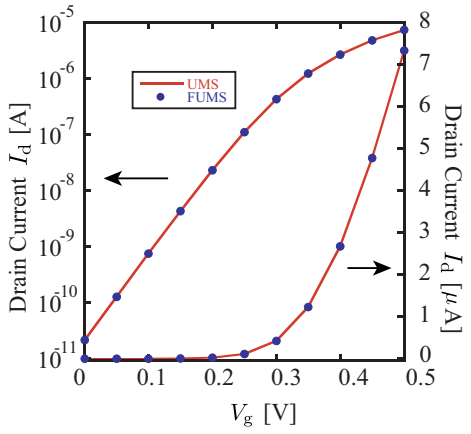

Fig. 6. Comparison of the $I_{\rm d}$-$V_{\rm g}$ characteristics between UMS and FUMS with longitudinal nodal number N_x is 41 and the cross-sectional nodal number N_{yz} is 676, where the solid line is obtained by the UMS, while dots are by the FUMS.

Figure 6 compares the $I_{\rm d}$-$V_{\rm g}$ characteristics by the UMS with those of the FUMS at the drain bias $V = 0.5$ V in the linear scale and the log scale. The nodal numbers are same ($N_x = 41$ and $N_{yz} = 676$) as the case of Figs. 4 and 5. It is shown that in almost all bias region, the results of the FUMS reproduce those by the UMS within the error of 0.03%. More importantly, in the subthreshold region, since the drain current is sensitive to the accuracy of the subband profile, the results show that in the 2D Schrödinger equation-1D NEGF-3D Poisson equation self-consistent calculation, the FUMS can reproduce quite comparable results in accuracy with those by the conventional UMS.

Finally, Fig. 7 shows the comparison of the total computational time to solve Schrödinger equation in the whole self-consistent calculation, which is one of the most important results of the present study. As seen in this figure, the computational time by the UMS is increasing linearly as the longitudinal nodal number N_x increases, which is the numbers of the cross-sections. In contrast, the computational time by

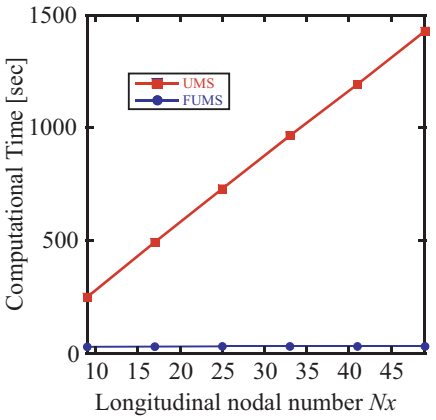

Fig. 7. Comparison of the total computational time to solve Schrödinger equation in the whole self-consistent calculation, with changing the longitudinal nodal number N_x, the cross-sectional nodal number $N_{yz} = 676$ is fixed, where the red line is obtained by the UMS, while the blue line is by the FUMS.

the FUMS is much less than that by the UMS, in addition, almost never changes. To be more precise, the comparison of the slopes in Fig. 7 shows that the FUMS is 570 times faster than the conventional UMS. That is because the FUMS needs to solve the non-perturbative Schrödinger equation only once in the coupled 2D Schrödinger equation-1D NEGF calculation, therefore the computational time does not increase even when N_x increases. It is important to emphasize that at the mesh size of $N_x = 41$ and $N_{yz} = 676$, which were checked to give enough accurate results in the UMS simulation, the present FUMS can calculate 40 times faster than the UMS.

IV. Conclusion

In this work we offer a new fast simulation framework that facilitates the simulation of 3D nano-scaled devices. The speed of the simulation has been drastically reduced by taking advantage of the present UMS-BPSM where both the Schrödinger-Poisson and NEGF equations are treated on equal-footing. This method can be applicable to any 3D device simulation as long as the potential profile is gradual along the transport direction. The drawbacks are for the case of abrupt change of the potential, which may be overcome by the spatial combination of the present method and the conventional techniques.

References

[1] Y. Saito, T. Nakamori, S. Souma, and M. Ogawa, Proc. SISPAD 2011, **11-4**, 311 (2011).
[2] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M.S. Lundstrom, IEEE Trans. Electron Devices, **50** 1914 (2003).
[3] E. Polizzi and N. Ben Abdallah, J. Computational Phys.**202**, 150 (2005).
[4] T.N. Rescigno, and C.W. McCurdy, Phys. Rev. A **62**, 032706 (2000).
[5] A. Gil, J. Segura, and N.M. Temme, Numerical Mehods for Special Functions, SIAM (2007).
[6] M. Luisier, A. Schenk, and W. Fichtner, J. Appl. Phys. **100**, 043713 (2006).
[7] A. Afzalian, N.D. Akhavan, C.W. Lee, R. Yan, I. Ferain, P. Razavi, J.P. Colinge, J. Comput. Electron. **8**, 287 (2009).
[8] S. Bangsaruntip, A. Majumdar, and J.W. Sleight, 2010 Symposium on VLSI Technology Digest, 21 (2010).