

MULTISCALE MODELING AND COMPUTATION OF NANO-ELECTRONIC
TRANSISTORS AND TRANSMEMBRANE PROTON CHANNELS

By

Duan Chen

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Applied Mathematics

2010

ABSTRACT

MULTISCALE MODELING AND COMPUTATION OF NANO-ELECTRONIC TRANSISTORS AND TRANSMEMBRANE PROTON CHANNELS

By

Duan Chen

The miniaturization of nano-scale electronic transistors, such as metal oxide semiconductor field effect transistors (MOSFETs), has given rise to a pressing demand in the new theoretical understanding and practical tactic for dealing with quantum mechanical effects in integrated circuits. In biology, proton dynamics and transport across membrane proteins are of paramount importance to the normal function of living cells. Similar physical characteristics are behind the two subjects, and model simulations share common mathematical interests/challenges. In this thesis work, multiscale and multiphysical models are proposed to study the mechanisms of nano-transistors and proton transport in transmembrane at the atomic level.

For nano-electronic transistors, we introduce a unified two-scale energy functional to describe the electrons and the continuum electrostatic potential. This framework enables us to put microscopic and macroscopic descriptions on an equal footing at nano-scale. Additionally, this model includes layered structures and random doping effect of nano-transistors.

For transmembrane proton channels, we describe proton dynamics quantum mechanically via a density functional approach while implicitly treat numerous solvent molecules as a dielectric continuum. The densities of all other ions in the solvent are assumed to obey the Boltzmann distribution. The impact of protein molecular structure and its charge polarization on the proton transport is considered in atomic details. We formulate a total free energy functional to include kinetic and potential energies of protons, as well as electrostatic energy of all other ions on an equal footing.

For both nano-transistors and proton channels systems, the variational principle is employed to derive nonlinear governing equations. The Poisson-Kohn-Sham equations are derived for nano-transistors while the generalized Poisson-Boltzmann equation and Kohn-Sham equation are obtained for proton channels. Related numerical challenges in simulations are addressed: the matched interface and boundary (MIB) method, the Dirichlet-to-Neumann mapping (DNM) technique, and the Krylov subspace and preconditioner theory are introduced to improve the computational efficiency of the Poisson-type equation. The quantum transport theory is employed to solve the Kohn-Sham equation. The Gummel iteration and relaxation technique are utilized for overall self-consistent iterations.

Finally, applications are considered and model validations are verified by realistic nano-transistors and transmembrane proteins. Two distinct device configurations, a double-gate MOSFET and a four-gate MOSFET, are considered in our three-dimensional numerical simulations. For these devices, the current fluctuation and voltage threshold lowering effect induced by discrete dopants are explored. For proton transport, a realistic channel protein, the Gramicidin A (GA) is used to demonstrate the performance of the proposed proton channel model and validate the efficiency of the proposed mathematical algorithms. The electrostatic characteristics of the GA channel is analyzed with a wide range of model parameters. Proton channel conductances are studied over a number of applied voltages and reference concentrations. Comparisons with experimental data are utilized to verify our model predictions.

ACKNOWLEDGMENT

Within the limited space, I wish to acknowledge a few amazing people who contribute to my success. Before this, I would like to thank all my comprehensive exam and dissertation thesis defense committee members, Dr. Guowei Wei, Dr. Keith Promislow, Dr. Moxun Tang, Dr. Changyi Wang and Dr. Zhengfang Zhou. It is their mathematical wisdom, help and personal humor that guide me through my Ph.D. career with happiness, make me mature in academics and give my doctoral study a perfect end.

My first thanks go to my advisor, Dr. Guowei Wei, who always has intense faith in the capabilities of his students and encourages them to pursue the next higher goal in research. Prof. Wei brought me into my current research five years ago and it is his interdisciplinary knowledge, great passion and infinite patience that lead me from a young man to a potential researcher. Besides mathematics, Prof. Wei is also an expert in physics, chemistry, biology and engineering. Five years' study with him has accumulated interdisciplinary knowledge for me and given me strong confidence to face many challenges in applied mathematics. He gave us all kinds of computational training to strengthen our basic skills and then lead us wandering in different worlds by bringing one and other brilliant ideas. He is also the greatest mentor from whom I can get help when I had difficulties, challenges or lost directions in my research. Most importantly, as a group leader, he always shares the same faith and dream with us. Not only providing guidance in our research topics, Dr. Wei gives us supports and help in our future research career.

I would like to thank my wife, Shaoyu, the most import person in my life for supporting me all the years. For my career, she is one of the reasons that I always keep myself moving forward. In my life, she is my soulmate with whom I can share my everything. We cheer up for each other's successes and make the happiness doubled, while considerate consolation is given in time when one of us is frustrated. Because

of her unbounded and continuous love for me, our tough PhD life in a foreign country becomes full of fun and happiness. After my graduation, we are expecting a more wonderful life in the future.

I also appreciate my father and mother for their continuous support and encouragement, not only in my career but also in daily life. It is my luck to be born and raised in such a harmonic and happy family. From the first day in the college to my study aboard, they always respect my decisions and encourage me to pursue my interests. Every week we talk to each other through the phone or web camera, they care about every detail of our lives, give good advise and help us as possible as they can. They have credits in my every achievement.

My thanks also go to a lot of lovely friends. Dr. Shan Zhao in University of Alabama, Dr. Yongcheng Zhou in Colorado State University, Dr. Weihua Geng in University of Michigan, Dr. Sining Yu and Dr. Yuhui Sun are great researchers from whom I have learned a lot when I was new in the group. It is their fundamental work that my research is based on. I would like to thank other current and former members in our group: Dr. Siyang Yang, Mr. Zhan Chen, Mrs. Qiong Zheng, Mr. Langhua Hu, Mrs. Jin Kyoung Park, Mrs. Yajie Yu, Mrs. Weijuan Zhou, Mr. Wangheng Liu, Mr. Manfeng Hu, Mrs. Shuailing Wang, Mr. Qi Zhao, Mr. Kelin Xia, Mr. Huibing Zhu, for their valuable help and suggestions in my research and daily life. Our group is like a family, we celebrate every new year's day together and BBQ in Dr. Wei's house each summer. We also had wonderful trips to Maryland and Toronto. I also thank Dr. Maureen Morton, Mr. Nick Boros and Mr. Li Yang for their generous help. At last but not least, I thank my friends Mr. Yanming Li, who is now in University of Michigan, Mr. Linkan Bian in Georgia Tech, Mr. Shangbang Rao in University of North Carolina and Mr. Lening Kang in Statistics, MSU. We got to know each other when I arrived in the U.S., the days we spent together as new comers will be the valuable memory in my life.

TABLE OF CONTENTS

List of Tables		ix
List of Figures		x
1 Introduction		1
1.1 Introduction to Nano-electronic transistors		1
1.1.1 Physical background		1
1.1.2 Review of the current models		2
1.1.3 Existing challenges		5
1.2 Introduction to ion channels and proton transport		9
1.2.1 Biological background		9
1.2.2 Review of current models		11
1.2.3 Existing challenges		13
1.3 Multiscale modeling theory of nano-electronic transistors and trans- membrane proton channels		15
1.4 Implicit solvent theory for structural and electrostatic analyses of bio- molecules		17
1.5 Mathematical issues and numerical challenges in model simulations		19
1.5.1 Highly accurate and efficient solver for linear and nonlinear Poisson equations with singularities		19
1.5.2 Quantum computation for particle transport		21
1.5.3 The self-consistent iteration		23
2 Modeling and computation of nano-electronic transistors		25
2.1 Theory and models		25
2.1.1 Multiscale energy functional on an equal footing		25
2.1.2 Electronic system		29
Double-gate MOSFET		30
Four-gate MOSFET		33
2.1.3 Transport system		35
The Non-equilibrium Green's function (NEGF) formalism		36
Electron density		38
2.1.4 Electrostatic system		41
Individual dopant model		41
Silicon-insulator-interface modeling		42
2.2 Numerical analysis and implementation		44
2.2.1 Decomposition of the Poisson equation		44
2.2.2 Numerical implementation of the self-consistent iterations		49

2.2.3	Analysis of the model well-posedness, convergence and iteration efficiency	53
2.3	Numerical simulation of the nano-electronic transistors	60
2.3.1	Device configurations	60
2.3.2	Simulation results for double-gate MOSFETs	63
2.3.3	Simulation results for four-gate MOSFETs	71
2.4	Conclusion remarks	75
3	Modeling and simulations of transmembrane proton channels	79
3.1	Theory and model	79
3.1.1	General description of the model	79
3.1.2	Free energy components	82
	Electrostatic free energy in the biomolecular region	83
	Electrostatic free energy in the solvent region	84
	Proton free energies and interactions	85
	Kinetic energy.	85
	Electrostatic potential.	86
	Generalized-correlation potential.	87
	External potentials	88
	Proton total energy functional.	89
3.1.3	Total free energy functional of the system	89
3.1.4	Governing equations	91
	Generalized Poisson-Boltzmann equations	91
	Generalized Kohn-Sham equations	92
3.1.5	Proton density operator for the non-hermitian Hamiltonian	93
3.1.6	Proton transport	95
3.2	Computational algorithms	97
3.2.1	Proton density structure and transport	97
	The solution of the generalized Kohn-Sham equation	97
	Boundary treatment of the transport calculation	99
3.2.2	Dirichlet-to-Neumann mapping for the generalize PB equation	104
3.2.3	The self-consistent iteration	107
3.2.4	The work flow of the self-consistent iteration	110
3.2.5	Model parameter selection	111
	The selection of generalized-correlation potential	111
	Choices of the dielectric constants	113
	Effective mass of the proton	115
3.3	Numerical simulations	116
3.3.1	Electrostatic properties of the Gramicidin A channel	118
3.3.2	Conductivity properties of the Gramicidin A channel	121
3.3.3	Model limitations	128
3.4	Conclusion remarks	129

4	Structure and electrostatic analysis for bio-molecules	132
4.1	The Krylov subspace theory (KSP) and preconditioner based MIBPB solver	133
4.1.1	Review of the Poisson-Boltzmann equation	133
4.1.2	KSP based and preconditioner accelerated MIBPB solver	137
4.2	Application to solvation energy calculations	145
4.3	Application to salt effects on protein-protein binding	150
4.4	Conclusion remarks	154
5	Thesis achievements and future work	155
5.1	On the modeling and simulations of nano electronic transistors	156
5.1.1	Contributions	156
5.1.2	Future work	158
5.2	On the modeling and simulations of transmembrane proton channels and biomolecule structures	159
5.2.1	Contributions	159
5.2.2	Future work	160
A	The MIB method	163
A.1	Dimension splitting	164
A.2	Derivative elimination	167
A.3	Derivative evaluation	170
B	Krylov subspace method and preconditioning for the MIB scheme	173
B.1	Linear equation systems and MIB matrices	173
B.2	The Krylov subspace theory and preconditioners	177
C	A short user manual for the MIBPB package	180
C.1	Work flow of the MIBPB package	180
C.2	Work flow for the display of the surface electrostatic potential	184
	Bibliography	188

LIST OF TABLES

2.1	Computational L_∞ error of the model with MIB scheme	59
4.1	Physical units notations	134
4.2	Some physical constants	134
4.3	Iteration numbers and CPU time for the discretization matrices of proteins	142
4.4	Iteration time for different combinations of KSP solvers and preconditioners	143
4.5	Convergence test of MIBPB solver with a set of proteins	143
4.6	Comparison of CPU time for the PETSc and the SLATEC schemes .	145
4.7	Solvation energies and CPU time for proteins	149
4.8	Solvation energies (kcal/mol) and CPU time (second) for large proteins	150
4.9	Binding affinity	153

LIST OF FIGURES

2.1	Illustration of a double-gate MOSFET with its y -direction being infinitely long. (a) Configuration of the double-gate MOSFET; (b) Computational domain. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.	30
2.2	An illustration of a four-gate MOSFET (i.e., silicon nanowire transistor). (a) Configuration; (b) Cross section at $y = 0$ for the computational domain; (c) Cross section at $x = 0$	34
2.3	Work flow of the overall self-consistent iteration.	50
2.4	Computational errors in simulating a double-gate MOSFET. (a) Individual doping in the source and drain with the dopant distribution shown in the lower panel of Fig. 2.5(a); (b) Individual doping in the channel.	53
2.5	Dopant distributions and iteration errors for the double-gate MOSFET. (a) Two distributions of 10 dopants. The distribution in the lower panel may lead to a divergence in the iteration. While the distribution in the upper panel leads to convergence; (b) Relaxation-factor-dependent convergence behaviors for the dopant distribution shown in the left-lower panel.	55
2.6	Errors of simulating the double-gate MOSFET in the line of $y = 0$ and $z = 0$	56
2.7	Contour plot of the electrostatic potential (a) and its absolute errors obtained with three sets of meshes(b, c, d) (10 dopants).	57

2.8	Electrostatic potential energy and difference of potentials for the double-gate MOSFET. (a) Potential landscape obtained with the standard finite difference method; (b) Difference of electrostatic potentials between computed with the standard finite difference method and with the MIB method.	64
2.9	Profiles of electrostatic potential and electron density obtained with the continuous doping approximation. (a) Potential function; (b) Electron density.	64
2.10	Profiles of electrostatic potential energy and electron density obtained with 5 individual dopants ($N = 5$). (a) Potential function; (b) Electron density.	65
2.11	Profiles of electrostatic potential energy and electron density obtained with 10 individual dopants ($N = 10$). (a) Potential function; (b) Electron density.	66
2.12	Profiles of electrostatic potential energy and electron density obtained with 40 individual dopants ($N = 40$). (a) Potential function; (b) Electron density.	66
2.13	Subband energies of the double-gate MOSFET obtained with individual dopant model.	67
2.14	Subband energy profiles under different gate voltage biases. (a) Continuum dopant; (b) Discrete dopants (10).	68
2.15	Subband energy profiles under different source/drain voltage biases. (a) Continuum Dopant; (b) Discrete Dopants (10).	69
2.16	The dependence of I-V profiles on the number of individual dopants. (a) Fluctuation of 5 dopants; (b) Fluctuation of 10 dopants; (c) Fluctuation of 40 dopants; (d) Fluctuation of different number of dopants.	70
2.17	The comparison of I-V profiles under different doping models revealing the individual dopant induced voltage threshold lowering effect. (a) $V_{S/D} = 0.2V$; (b) $V_{S/D} = 0.4V$	71
2.18	Subband energies for the four-gate MOSFET. The dopants break the symmetry and energy degeneracy in the second and third subbands.	72
2.19	Cross-section of potential profiles for the four-gate MOSFET obtained with (a) Continuum doping; (b) Discrete dopants (10).	73

2.20	Current fluctuations of the four-gate MOSFET with (a) 15 dopants; (b) 40 dopants; (c) 80 dopants; (d) Different number of dopants.	74
3.1	(a) Illustration of multiscale model of a proton channel; (b) Computational domains of the multiscale model with Ω_m being the channel molecule and membrane domain and Ω_s being the solvent domain. Here z -direction is regarded as the transport direction.	80
3.2	Work flow of the overall self-consistent iteration.	112
3.3	3D illustration of the Gramicidin A (GA) channel structure and surface electrostatic potential. The negative surface electrostatics as indicated by the intensive red color on the channel upper surface and inside the channel pore implies that the GA selects positive ions. (a) Top view of the GA channel; (b) Side view of the GA channel.	117
3.4	Electrostatic potential and charge density of the GA channel along the z -axis obtained with $\epsilon_m = 2$ and $n_p^0 = 0.1$ molar (Red: $\epsilon_{ch} = 20$; Green: $\epsilon_{ch} = 40$; Blue: $\epsilon_{ch} = 80$). (a) Electrostatic potential profiles in channel; (b) Proton density profiles in the channel.	118
3.5	Electrostatic potential and charge density of the GA channel along the z -axis obtained with $\epsilon_m = 5$ and $n_p^0 = 0.1$ molar (Red: $\epsilon_{ch} = 20$; Green: $\epsilon_{ch} = 40$; Blue: $\epsilon_{ch} = 80$). (a) Electrostatic potential profiles in the channel; (b) Proton density profiles in the channel.	120
3.6	Electrostatic potential and charge density of the GA channel along the z -axis obtained with $\epsilon_m = 10$ and $n_p^0 = 0.1$ molar (Red: $\epsilon_{ch} = 20$; Green: $\epsilon_{ch} = 40$; Blue: $\epsilon_{ch} = 80$). (a) Electrostatic potential profiles in the channel; (b) Proton density profiles in the channel.	121
3.7	Electrostatic potential profiles of the GA channel under different ion reference densities n_p^0 . Red: $n_p^0 = 0.1$ molar; Green: $n_p^0 = 0.2$ molar; Blue: $n_p^0 = 1.0$ molar; Black: $n_p^0 = 2.0$ molar. $\epsilon_m = 5$ and $\epsilon_{ch} = 40$	122
3.8	The total potential of the GA channel which includes electrostatic and generalized-correlation contributions under various voltage biases. Dielectric constants are $\epsilon_m = 5$ and $\epsilon_{ch} = 30$. The pH value of the solution is 2.75. (a) Total potential of monovalent cations; (b) Total potential of monovalent anions.	122
3.9	The first 15 eigenvalues (the $U^j(z)$ in Eq. 3.41) of the effective potentials along the transport direction used in the transport calculations at the voltage bias of 0.2V.	123

3.10	Voltage-current relation of proton translocation of GA at different concentrations. Blue dots: experimental data of Eisenman et al [59]; Red curve: model prediction.	125
3.11	Conductance-concentration relation of proton translocation at a fixed voltage. Voltage bias=0.05V; Blue dots: experimental data of Eisenman et al[59]; Red curve: model prediction.	126
4.1	The implicit protein-solvent system	134
4.2	Condition numbers over 15 proteins (PDB IDs from protein 1 to protein 15: 1ajj, 1vii, 1cbn, 1bbl, 1fca, 1sh1, 1vjw, 1fxd, 1bpi, 1a2s, 1frd, 1svr, 1a63, 2erl, 2pde). (a) Condition numbers for unpreconditioned (unPCed) MIBPB matrices;	139
4.3	Condition numbers over 15 proteins (PDB IDs from protein 1 to protein 15: 1ajj, 1vii, 1cbn, 1bbl, 1fca, 1sh1, 1vjw, 1fxd, 1bpi, 1a2s, 1frd, 1svr, 1a63, 2erl, 2pde). (b) Comparisons of condition numbers under three settings.	140
4.4	Comparison of solvation energies of proteins (From protein 1 to 19: 1ajj, 1bbl, 1vii, 1cbn, 2pde, 1sh1, 1fca, 1fxd, 1vjw, 1bor,1hpt,1bpi, 1mbg, 1r69, 1neq, 1a2s, 1svr, 1a63, 1a7m) calculated by using the MIBPB and the APBS methods.	147
4.5	Comparison of CPU time of preconditioned (PCed) and un-preconditioned (unPCed) MIBPB methods for 19 proteins (from protein 1 to 19: 1ajj, 1bbl, 1vii, 1cbn, 2pde, 1sh1, 1fca, 1fxd, 1vjw, 1bor, 1hpt, 1bpi, 1mbg, 1r69, 1neq, 1a2s,1 svr,1a63, 1a7m).	148
4.6	(a) Binding affinity of 1emv; (b) Binding affinity of 1beb.	153
C.1	Work flow of the MIBPB package	182
C.2	Visualizations of surface electrostatic potentials of protein 1beb. From left to right: (a) Surface electrostatic potential ($I = 0, h = 1.0\text{\AA}$); (b): Surface electrostatic potential with the ionic strength $I = 1 (h = 1.0\text{\AA})$; (c): The difference of surface electrostatic potentials between in an ionic solvent ($I = 1$) and in a water solvent ($I = 0$); (d): The difference of surface electrostatic potentials in water ($I = 0$) between grid meshes $h = 1.0\text{\AA}$ and $h = 0.5\text{\AA}$	185

Chapter 1

Introduction

1.1 Introduction to Nano-electronic transistors

1.1.1 Physical background

The continuous demand in rising the performance of electronic devices has led to the reduced geometric dimension and supply voltage of metal oxide semiconductor field effect transistors (MOSFETs), or complementary metal oxide semiconductors (CMOSs), which are fundamental building blocks of large scale integrated circuits used in almost all electronic equipments. At present, MOSFETs are designed, manufactured and operating on much less than 100nm scale. According to “International Technology Roadmap for Semiconductors (ITRS) (<http://www.itrs.net/>)”, the channel length of CMOSs will be down scaled from the present 45 to about 22 nm in 2016. The down-scaling of the transistor channel length also requires simultaneous down-scaling of the gate oxide, connecting material, doping concentration and operation voltages [55, 40, 19]. The ultimate channel length is expected to be around 16 nm. At such a channel length, most critical design parameters quickly approach the atomic scale and associated physical limits. Many down-scaling associated devices characteristics have been studied by Vasileska, et al [4, 3, 70, 72, 90, 71, 91]

for MOSFET, FinFET and various other silicon on insulator (SOI) devices. The most important feature of a nano-scale transistor is that quantum mechanical effects become significant and will dramatically impact the macroscopic quantities, such as current-voltage characteristics and conductance. In particular, at 22 nm or less, channel tunneling and gate leakage may devastate the classical function of the MOSFET. Electrostatic control and suppression of quantum effects are important issues [116, 129, 54, 43, 137, 83, 6, 51, 120, 49, 10, 150]. Nano transistors with the range of channel lengths being around 20 nm and 10 nm are referred as “ultimate CMOSs” and “functionally enhanced CMOSs”, respectively. Ultimate CMOSs are the smallest CMOSs that still operate with the classical principle while severe quantum effects have to be suppressed by appropriate electrostatic potentials and designs. Functionally enhanced CMOSs are nano-quantum transistors which utilize the fundamental properties of nature that do not have direct analogs in classical physics. Some of these properties are quantum coherence, i.e., a possibility for a quantum system to occupy several states simultaneously, and quantum correlation or entanglement. Presently, the majority of these quantum structures, such as nano-mechanical resonators, quantum dots, quantum wires, single electron transistors, and similar low dimensional structures, exist only as prototypes in research laboratories or just being contemplated [86, 64]. The working principle and physical function of quantum devices are subjects of extensive research. Practical realization of quantum transistors faces a number of challenges in design, test, material selection, lithography, interconnect, process integration, metrology, assembly, packaging, plus device modeling and simulation.

1.1.2 Review of the current models

The main purpose of the device modeling is to predict device characteristics and performance. This amounts to the understanding of transport features, including

current-voltage (I-V) characteristics at the source, drain or gate contacts of the device. For ultimate MOSFETs and other nano-quantum transistors, quantum effects, such as gate leakage and channel tunneling under various voltage settings will be of main concerns of the modeling and simulation [96, 143]. Currently, the non-equilibrium Green's functions (NEGF) formalism is the main workhorse for nano-device transport modeling [153, 74]. The NEGF formalism was originally developed by Schwinger [136], Kadanoff and Baym [89], and has been revived recently for device modeling [45, 96, 44, 143]. This is a general and useful formalism using the Fermi-Dirac statistics for electrons. It allows the description of interactions, including scattering processes of particles (i.e., electrons and phonons) and relaxation due to the surroundings. An equivalent approach is the Dyson integral equation representation. However, computational aspects for differential and integral equations are quite different.

Another practical transport model is the Boltzmann equation, or the Boltzmann-Vlasov equation, which describes the kinetic of a typical particle (e.g., electron, phonon or photon), due to the two-body scattering with another particle and/or external field effect [73, 22]. The inherent Boltzmann distribution can be a good approximation to the Fermi-Dirac distribution at high temperature. Transport properties, such as current density, conductance and tunneling rate, can be computed as expectation values of physical observables with the distribution function, the Wigner distribution [85] or density operator [3]. In fact, the transport equation derived from the quantum Boltzmann equation, known as the Waldmann-Snyder equation [151, 142], can provide quantum correction to the classical drift-diffusion expression. The Waldmann-Snyder equation can be derived from the BBGKY hierarchy with an appropriate scattering closure for the two-body density operator. Other density matrix methods, such as the Master equation [82, 65], describe the time evolution of the probability function. Assuming a continuous-time Markov process, the integrated master equation obeys a Chapman-Kolmogorov equation. However, to describe the

transport of electronic devices, the transition matrix of the master equation has to be evaluated by other reliable means. The semiconductor Bloch equation is derived by using the slowly varying envelope and rotating wave approximation, providing a diagonally-dominated calculation of currents and allowing a simple approximation of scattering. Yet another approach is the Fokker-Planck equation describing the rate change of the probability density function of a particle in terms of drift-diffusion processes [116, 129, 109]. This equation can be used to model the electron transport in the quantum ballistic regime [113]. Additionally, Monte Carlo methods have also been applied to electron transport [22, 4, 90]. The electron scattering effect from the devices interface roughness was studied via an ensemble Monte Carlo device simulation technique [91]. A new scheme was proposed and applied to study the role of the discrete impurities in the device terminal characteristics [70, 71, 72]. By using a corrected Coulomb force, this approach prevents the double-counting of the electron-electron and electron-ion long-range interaction.

To account for the quantum effect, the electronic structure in terms of wavefunctions is required in most transport evaluations. The quantum mechanical theory is indispensable for electron structures at nano scale. To this end, one has to select the level of the description of the quantum system and the level of the approximation to governing equations. The description of electron structures depends on the level of approximations and the size of the quantum system depends on the level of sophistication of the model. Although formally a fully quantum mechanical first principle description is desirable for a given device feature at nano scale, it normally involves a large number of atoms, molecules and electrons. Therefore, the resulting full scale quantum system is intractable and appropriate approximations are required. Currently, practical models describe the dynamics of a few electrons or even a single electron in which the indispensable essential feature of the quantum effect is retained and different levels of approximations can still be derived depending on how the in-

teraction of the single electron with other particles is handled. At the lowest level of approach, a single electron dynamics in a band structure of the solid is governed by the Schrödinger equation, which is coupled back to the Poisson equation as charge sources [104, 146, 15, 87, 99, 8, 62, 133, 13]. The interaction of many bands can also be considered by using a general $\mathbf{k} \cdot \mathbf{p}$ method derivation of many-body Schrödinger equation [162, 64, 106]. Recently, linear combination of bulk band (LCBB) method [83], which relies on the expansion of the confined states in terms of periodic Bloch wave functions, was used for a large number of atoms. Some quantum corrected classical methods, such as quantum drift-diffusion (QDD) models or Schrödinger-Poisson drift-diffusion (SPDD) models, are employed and summarized in a unified framework [48, 51]. The well-posedness of these models and numerical efficiency are analyzed mathematically in the fashion of solution fixed point maps [51].

1.1.3 Existing challenges

Apart from difficulties with device fabrication and testing which typically require nano scale resolution and high precision control, there are numerous modeling and computational problems associated with ultimate and functionally enhanced nano transistors as discussed by the ITRS. The essences of these problems are quantum effects, geometric interface effects and dopant effects.

1) Quantum effects include electron confinement, resonance states, source-drain off state quantum tunneling current, channel barrier tunneling, gate leakage, many body correlations and channel-channel interference at nano scale. These effects are commonly modeled by the coupled Poisson-Schrödinger equations. However, the consistence and validity of these equations have rarely been examined. There is a pressing need for innovative methods, models and algorithms that contribute to the prediction and design of nano-quantum transistors whose channel lengths are in the range of 8-22 nm.

The solution of the many-electron Schrödinger equation, including atomic information, is extremely expensive. Semi-empirical approaches which make use of parameters from experimental data are often used. More rigorous but expensive methods are ab-initio approaches, including the Hartree-Fock method [141] and the density functional theory (DFT) [76, 93, 117]. The size of the system is limited when ab-initio methods are used. To increase the computational capability, pseudopotential methods can be used to remove core electrons and singularities in calculations. The resulting quantum mechanical system is still formidably expensive to solve for nano devices. The DFT is associated with the Kohn-Sham equation and it can be accelerated by using the linear scaling divide and conquer method [156, 101], and the tight binding approximation [79]. In general, there is a pressing need to develop innovative modeling strategies and efficient computational methods for realistic device problems.

To improve the accuracy of the electron structure, it is necessary to consider atomistic models. The core electrons on different atoms are essentially independent of the state of the surrounding atoms. Therefore, only valence electrons participate effectively in interactions between atoms. Thus, the core electron states can be assumed to be fixed and a pseudo-potential may be constructed for each atomic specie which takes into account the effects of the nucleus and core electrons. As such, one only needs to explicitly consider the valence electrons. Furthermore, the tight binding model assumes that the full Hamiltonian of the system can be approximated by the Hamiltonian of an isolated atom centered at each lattice point for a solid-state lattice of atoms. This simplifies the formulation and offers a further saving in computational effort.

2) Geometric interface effects refer to the impact of (layered) material variations within a device and interconnects between devices to the device performance. These effects become crucial to ultimate CMOSs and functionally enhanced CMOSs. For example, dielectric interfaces of metal-oxide, metal-semiconductor, and oxide-

semiconductor will induce non-ballistic transport behavior even if there is no other interaction [91]. However, with few exceptions [103], most present simulation models are based on simplified rectangular geometric shapes, homogeneous dielectric media, and even reduced dimensions. The impact of realistic geometry, including gate dielectric layers and interconnects, has hardly been investigated in the past and calls for new modeling strategies and innovative methods.

Typically, dielectric constants of different components in MOSFETs vary dramatically. For example, the dielectric constant of the silicon dioxide insulator is a few times smaller than that of the bulk silicon substrate. The ratio of dielectric constants in different layers is also important to the device scaling. According to device scaling physics [66], the scale length Λ of the device, for the first order approximation, depends on the insulator thickness (T_I) and the ratio (ϵ_{Si}/ϵ_I) of dielectric constants of the silicon and the insulator in the way

$$\Lambda = W_{dm} + T_I \frac{\epsilon_{Si}}{\epsilon_I},$$

where W_{dm} is the maximum channel depletion depth relating to the channel doping concentration. The above theory predicts that the proper minimum design length lies between Λ and 2Λ . It is clear that smaller value of T_I and larger value of ϵ_I help device scaling. Replacing the silicon dioxide gate dielectric with a high-k material allows increased gate capacitance without the concomitant leakage effects. The proper formation of distinct interfaces is a stringent requirement for ultimate CMOSs and functional electronic devices to suppress leakage currents due to tunneling, as the thickness scales much below 2 nm. Computationally, it is important to be able to simulate the interface roughness and irregularity due to the device fabrication processes [91]. The use of interface description is indispensable for modeling of ultimate and functional CMOSs.

3) Introducing appropriate impurity atoms (known as dopants) into a semiconductor provides electron reservoirs and can increase the electrical conductivity by many orders of magnitude. By doping a semiconductor device, we can engineer its electrical properties, e.g., its conductivity, electrostatic potential and its charge carrying mode. Doping is a key to our understanding of semiconductor devices and a strategy for the design and manufacture of desirable devices [70, 71, 72]. Doping effects are often described by distribution functions in continuum device models without explicit consideration of individual dopant atoms and traps [25]. This continuum approach works very well for electronic devices of large sizes but will lead to severe errors in electron structure and transport for ultimately scaled nano devices. These errors are often seen as statistical fluctuations [92, 63]. When the device size is reduced to 22 nm or less, it becomes indispensable to consider individual dopant atoms and traps. However, with few exceptions [83, 70, 71, 72], this issue has been hardly addressed in the literature. As individual dopants are fundamental to the function of ultimate MOSFETs and nano-quantum transistors, it is imperative to develop innovative models and efficient methods to analyze their impact.

In continuum modeling, dopants have either been described as continuous distributions in p-n regions or been formulated as a change in the dielectric effect, leading to different dielectric values in different p-n regions. These treatments work mostly well for the prediction of device properties. However, when the channel length reduces to about 10 nm, the quantum effect becomes important. Thus, each doping atom may have a dramatic impact to the quantum state of nearby electrons. Atomistic model for dopants becomes indispensable. Wong and Taur [155] provided a classical study of discrete random dopants. Recently, quantum random dopant models are applied to the channel of sub- $0.1\mu\text{m}$ [9] and 25nm [84] MOSFETs for threshold voltage lowering and fluctuations. The impact of random dopant aggregation in source and drain is studied via the NEGF formalism [108]. It is found in these studies that doping is

only macroscopically controllable when the discrete microscopic dopant distribution is also controlled. Macroscopically, identical devices may suffer from strong performance variations because of the microscopic differences. Therefore, it is important to understand individual dopant effect in nano electronic devices.

1.2 Introduction to ion channels and proton transport

1.2.1 Biological background

There are a couple of seemingly conflicting fundamental requirements for a living cell to survive and function properly: On one hand, the cell needs the protection of the plasma membrane, which works as a potential barrier and maintains the intracellular electrolyte composition that may be different from that of the extracellular environment. On the other hand, information communication and material exchange must be established between the intracellular and extracellular environments for all living cells. A wide variety of biological processes, such as signal transduction, nerve impulse and so on, are modulated and sometimes, initiated by the intra/extra-cellular information and material exchanges. These two conflicting tasks are accomplished by ion channels, which are proteins with pores and embedded in lipid bilayers, selectively permitting the permeation of specific ions. Because of these important biological roles, as well as frequently serving as the target for drug designing, ion channels have attracted great research interest in experimental, theoretical and computational explorations. Most research activities are focused on a few ion channel properties [75]:

- (i) The gating of ion channels. Ion channels are not always open or close. Based on the mechanism controlling the open/close status, they are categorized as ligand-gated ion channel (the channel is open only when the specific ligand is bound to the ex-

tracellular receptor domain), voltage-gated ion channel (the channel is open/close by the regulating membrane potential) and other gating channels, such as mechanical, sound, and thermal stimuli. It is worthwhile to point out that the present work does not focus on the ion channel gating mechanism — channels discussed here are all assumed open. (ii) The selectivity of the ion channel. When an ion channel is open, it is not open to all the ion species, only certain ions can penetrate. In this sense, ion channels are also classified by the permeating ions, such as potassium channels, sodium channels, and proton channels, etc. (iii) The efficiency of ion conductance. When an ion channel is open and conducts a specific ion species, the efficiency of ion conductance is of major interest, which is measured by the current-voltage (I-V) curve. Technological advance in the past few decades makes it possible to measure I-V curves through a single channel for a variety of ion channels under physiological conditions. These techniques are considerably empowered by the genetic engineering technology to identify the gating mechanism. (iv) Structural analysis. Many channel protein structures have been discovered by X-ray crystallography, nuclear magnetic resonance (NMR) and cryoelectron microscopy. Channel protein structural information is deposited in the Protein Data Bank (PDB). (v) Theoretical and computational research. Abundant knowledge about ion channels accumulated by experimental means has created an excellent testbed for theoretical modeling and prediction of ion channel transport. Various mathematical/physical models have been proposed for numerical simulations. However, there are still many important theoretical problems in the field [37]. One of the problems concerns the dynamical detail of the ion permeating process. Due to the relative narrowness of the pore size, the ion-water geometry needs to be rearranged in order for ions to successfully cross the channel. Therefore, the orientation and polarity of water molecules, the interaction between partially dehydrated ions and fixed charges on the protein wall must be significantly different from those under the bath condition. Another problem is the precise role

of quantum effects in many proton channels, such as the narrow M2 channels of Influenza A. These problems pose challenges for theoretical/mathematical modelings. Commonly used approaches include molecular dynamics, Brownian dynamics, and the Poisson-Nernst-Planck (PNP) equations. There are a number of excellent reviews [37, 95, 58, 132, 127, 126] for various theoretical models at a variety of levels of descriptions and approximations.

1.2.2 Review of current models

Molecular dynamics (MD) provides one of the most detailed descriptions in modeling biomolecular systems and there are several user-friendly packages available, such as AMBER [118], CHARMM [105], etc. In fact, MD is one of viable models which are able to predict the ion selectivity in ion channel modeling. However, the use of MD in modeling ion permeation is still limited. The most significant barrier for MD applications in ion channels is the difficulties of predicting the channel conductance, which is the primary physical observable. Extremely small time step (around 1 or 2 femto seconds) has to be employed in the numerical integration of the Newton's equation to obtain the necessary accuracy because the fast time scale of molecular bond motions. Whereas, a typical channel current (with the magnitude of the order of pico Ampere) corresponds to average transit time of tens of nanoseconds for a single ion. Therefore, the MD simulation must last around microseconds in order to obtain sufficiently accurate conductance calculations. Due to the total simulation time needed and the necessarily small time step, the MD computation without invoking crude approximations is still not affordable with current computers for accurate conductance prediction. Therefore, the full scale MD simulation of ion channels is not feasible. In practice, it is still very useful for MD simulations to obtain alternative channel configurations, solvent polarizations, diffusion coefficients, etc., in assisting other approaches for the transport estimation [37].

Brownian dynamics (BD) [36] based on the Langevin equation treats ions as explicit particles in the ion channel modeling, while describes the surrounding environments (channel proteins and lipid bilayers) implicitly by a continuum approach. In Brownian dynamics, there are many forces which act on the target ion [37]. First, there is a force from fixed charges in the protein and membrane, as well as the applied external field. Additionally, there is a force from the self-induced charge by an ion on the channel boundary. When ions pass through the channel, there is always a repelling force induced at the channel boundary against the ion motion. Finally, there is a force from mobile ions in bath regions. Among these three components, the force from fixed charges can be obtained by solving the Poisson equation in the absence of mobile ions, whereas forces due to mobile ions can be evaluated by solving the Poisson equation while switching off fixed charge and applied field, and allowing ions to move around all the grid points. Once these Poisson equations are solved numerically, the forces are pre-stored in the grid and ready to be used to determine ion trajectories.

By assuming a mean-field approximation, the Poisson-Nernst-Planck (PNP) model [140] is a continuum electro-diffusion theory which treats not only the protein, lipid layer, bath solution as continuums, but also the ions of interest. The Poisson equation provides the electrostatic potential profile in the whole computational domain based on charge sources from mobile ions in the solution and fixed charges in the channel protein and lipid layer. The gradient of the electrostatic potential gives rise to the driven force, which, together with the gradient of ion density, is used in the Nernst-Planck equation to determine ion density flux. Therefore, the ion density distribution is governed by both the electrostatics induced drifting and the density gradient induced diffusion. The ion conductance is computed from the charge flux. Obviously, both the BD and PNP models have a number of similarities in their initial setups and computational approaches [107, 122, 57].

1.2.3 Existing challenges

Due to its computational efficiency, the PNP model has been widely implemented for various ion channels [77, 80, 88, 161] embedded in different lipid bilayers [80, 24]. Many mathematical analyses, for example, derivation of the NP equation from Boltzmann equation via perturbation theory [140], asymptotic expansions of the I-V relations [1], accelerating algorithms [56] and inverse problems related to the ion selectivity [23], are also popular research topics in the field. However, the validity of the PNP model has been questioned in many aspects, particularly for narrow ion channels [38, 111, 69]. Arguments root from the theoretical defect that ions are treated as continuum instead of particles in the narrow channel. This continuum assumption is only reasonable under bulk concentration condition or a channel pore with a sufficiently large diameter. First of all, it is conceptually difficult to define ion “concentration” when the diameter of a channel pore is comparable to that of an ion. Secondly, when the scale is down to a couple of angstroms, non-electrostatic factors such as Brownian motion, may become important or even dominant. The screening effect is significant when the channel diameter is smaller than the Debye length of the realistic electrolyte. In this situation, ion particles induce dielectric boundary charges, which result in a dielectric self-energy (DSE) barrier. The PNP model neglects these energy barrier factors [24, 37], and usually overestimates biological quantities of interest. The PNP model also ignores the non-electrostatic forces and self-energy, and employs an artificially reduced diffusion coefficient (about a factor of 1/50) to fit experimental data [77]. Several modified PNP models have been proposed, in which the ion self-energy is obtained either by using the Poisson equation [39, 69] or the MD [107] simulation, and is added to the Nernst-Planck equation.

Apart from the ion transport of sodium, potassium and calcium, the long range proton transfer (LRPT) across biomembranes is also of central importance and plays a

major role in many biochemical processes, such as cellular respiration, ATP synthase, photosynthesis and denitrification [94]. The LRPT is usually realized via proton channels or proton nanowires, where water molecules are connected in a chain to conduct protons. Two common examples of proton channels are the Gramicidin A (GA) and the newly discovered M2 proton channel of Influenza A [134]. Theoretical investigation has been extensively carried out and various experimental data about the proton flux are available [122, 123, 121]. However, the main mechanism of the LRPT is not fully understood yet [144], with the belief that protons are totally different from other ions and have larger conductance. In the Grotthuss-type mechanism theory [112, 2], protons achieve the translocation in the channel through a succession of hops along a single chain of hydrogen-bonded water molecules, i.e., an existing hydrogen bonded network, compared to other ions for which the permeation occurs mainly via hydrodynamic diffusion. The actual transfer through the hydrogen bonded network is usually fast and both the rearrangement of the hydrogen bonded network and energy barrier are considered as rate limiting factors. There is an agreement that the aforementioned BD theory and PNP model may be expected to work well for heavy ions but not for protons, which have lighter mass and whose transfer involves the hydrogen bonds making and breaking. These processes need to be studied quantum mechanically. Some investigators have explored proton channels via Feynman path integral simulations and quantum energy levels of protons are computed by the Schrödinger equation [121, 123, 122]. Several theoretical models are proposed in the last decade [135, 144, 21].

1.3 Multiscale modeling theory of nano-electronic transistors and transmembrane proton channels

It is not appropriate to apply any single modeling strategy to interpret such complex nano-transistor and proton channel systems. Fully atomic models such as molecular dynamics keep the number of degrees of freedom for the model but impose unaffordable computational burden to model simulations. While for the pure continuum approximation, many detailed information will be lost. Therefore, multiscale or multiphysical modeling techniques become sharp tools to analyze these systems. Much of the modeling effort is devoted to the proton channel because of the biological complexity.

The ion channel system is divided into the solvent and the molecule subdomains. The former includes the extra/intra cellular bulk aqueous environment, as well as the solvent in the channel pore. Because of the huge number and fast rotating of water molecules, it is computationally expensive to trace their motions individually. Additionally, water molecules are also not the object of main interests. Therefore, the water is implicitly treated as a structureless dielectric continuum. The molecule subdomain consists of channel protein and lipid bilayers. The structure and property of the channel protein is paramount to the proton permeation, so the continuum model is not appreciated. Fortunately, the total number of atoms of the channel protein is not high (usually up to thousands), one can record the atomic details, such as the positions, radii and partial charges of atoms to give explicit representation for the channel protein. On the contrast, the structure of the lipid bilayer is relatively simpler, it can also be approximated by a dielectric continuum.

Besides the multiscale treatment for the model, different physical principles are

employed to study objects of different interest. The most important ion is the proton. Due to the special transport mechanism plus unique properties, the quantum mechanics is used to illustrate the permeation process of protons. For other ion species, they are all treated in classical mechanics and their densities are assumed to obey the Boltzmann distribution. The motion of protons is under intensive electrostatic and other interactions with the mobile ions in the solvent, fixed charges in the channel protein, the water molecules, and other surrounding environments. Ion species are modeled with various mechanisms and corresponding approximations, it is the mutual interactions among ions that recover the whole system as a reasonable approach to the primitive biological process.

The multiscale/domain/physical idea can be also applied to the modeling of nano-transistors. We introduce a two-scale variational framework that, upon energy optimization, generates new self-consistently coupled Poisson-Kohn-Sham equations which allow the easy incorporation of linear scaling tight-binding, pseudopotential, atomic charges and dopants, divide and conquer methods. The proposed framework puts macroscopic description of electrostatic potentials and the microscopic description of electronic structures on an equal footing at nano-scale. MOSFETs are made of several materials, for example, the silicon as semiconductor material and the silicon dioxide as insulator. These materials have different dielectric permittivities. Therefore, interface models and associated elliptic interface techniques are introduced to the nano-electronic device modeling and computation. Because of the nanometer down-scaling of devices and the strong confinement of the channel, electron transport must be illustrated in a quantum mechanical way. Finally, we provide a new mathematical model to account for the random individual dopant effect in semiconductor material. The Dirac delta function is used to represent the dopant position and eliminate the finite-size effect in the previous discrete dopant models.

In work of this thesis, the multiscale modeling method and multi-physical tools

are attempted to study nano-transistors and proton channels. Energy components of both systems are integrated on an equal footing framework, in view of total free energy functional. Governing equations can be derived from the energy functional by the variational principle for both problems. Generally, one elliptic type equation (the Poisson equation for nano-transistors and the generalized Poisson-Boltzmann equation for proton channels) and the generalized Kohn-Sham equation are derived, the former usually provides the electrostatic landscape of the system while the latter describes the motion of quantum particles.

1.4 Implicit solvent theory for structural and electrostatic analyses of bio-molecules

Implicit treatment, or continuum approximation of the solvent is a ubiquitous technique of multiscale modeling for ion channel systems. It is also important for structural and electrostatic analyses of general bio-molecular systems.

Under physiological conditions, almost all important biological processes, for example, signal transduction, DNA specification, transcription, post transcription modification, translation, protein folding and protein ligand binding, occur in water which comprises 65-90% of cellular mass. An elementary prerequisite for the quantitative description and analysis of the above-mentioned processes is the understanding of solvation, which involves energetics of interactions between solute molecules and solvent molecules or ions in aqueous environment. Solute-solvent interactions are typically classified as the polar type and the non-polar type. Although widely used, this classification is arbitrary and has caveats associated with the non-unique descriptions, as well as the intrinsic coupling between these two types of interactions. The polar type of solute-solvent interactions is the main interest of the present work. It originates from electrostatic effects, which play important roles in biophysics, biochemistry,

structural biology, electrochemistry and electrophoresis.

The solvent has a substantial volume and a significant contribution to electrostatics via numerous mobile ions. However, it is the solvated solute molecule that is the focus of the interest in most research. As such, the solute is described in atomic or electronic detail, while atomic details of the solvent and mobile ions are approximated by a mean-force description and probability distribution, respectively. This implicit solvent method can greatly reduce the computational cost of the traditional explicit solvent methods, in which a microscopic description of the solvent is retained. Various implicit solvent models are available to describe polar solvation [128, 138, 47, 29, 12]. The most widely-used methods are currently the generalized Born method [53, 61, 166, 110, 29], polarizable continuum [35, 145, 81] and Poisson-Boltzmann equation (PBE) [11, 97, 138, 47] models. The use of polarizable continuum models is mostly restricted to small molecular systems. Generalized Born methods are very fast but are only heuristic models for estimating polar solvation energies of biomolecular structures. These methods are often used in high-throughput applications such as molecular dynamics simulations [147, 139, 60]. PBE models can be formally derived from Maxwell's equations [16] and offer a somewhat slower, but more accurate way for evaluating polar solvation properties [46, 114, 14]. Additionally, PBE techniques are often used to parameterize and assess the accuracy/performance of generalized Born models [114, 114, 148]. Finally, unlike most generalized Born methods, PB models provide a global solution for the electrostatic potential and field within and around a biomolecule, therefore make them uniquely suited to visualization and other analysis [102, 18] that require global information about electrostatic properties.

1.5 Mathematical issues and numerical challenges in model simulations

Other than sharing similar modeling methodologies, implementations of both nano-transistors and proton channel models encounter common numerical challenges, which have attracted great mathematical interests for the last decades. For example, the multidomain and multiscale treatment of both systems results in interface problems and delta functions in the partial differential equation. Highly accurate and efficient numerical schemes are required to handle these singularities. Additionally, desirable computational theory and algorithms for the scattering of quantum particle in finite region are indispensable. Furthermore, numerical convergence and efficiency of the self-consistent iteration need to be explored for the coupled governing equations that are derived from the total free energy functional. The numerical challenges and associated mathematical treatments in model implementations are outlined as follows:

1.5.1 Highly accurate and efficient solver for linear and non-linear Poisson equations with singularities

One of the governing equations, the Poisson equation for nano-transistors or the generalized Poisson-Boltzmann equation for proton channels, admits interface and delta source singularities. These singularities in the elliptic equation are successfully solved by the evolution of the MIBPB solver, which is a MIB algorithm based Poisson-Boltzmann solver package but can be easily extended to general elliptic equations. The MIBPB-I [163], the MIBPB-II [157] and the MIBPB-III [67] have been developed (<http://www.math.msu.edu/~wei/MIBPB/>). The MIBPB-I is the first PB solver that directly enforces the flux continuity conditions at the dielectric interface in the biomolecular context. However, it cannot maintain its designed order of accuracy in

the presence of MS singularities, such as cusps and self-intersecting surfaces. This problem was addressed in the MIBPB-II by utilizing an advanced MIB technique developed by Yu et al. [158] who offered special treatments for geometric singularities. However, the MIBPB-II loses its accuracy when the mesh size is as large as half of the smallest van der Waals radius, due to the interference of the interface and singular charges. To split the singular charge part of the solution, a Dirichlet to Neumann mapping approach [33] was designed in the MIBPB-III, which is by far the most accurate and reliable PB solver. This new solver remains accurate at the smallest van der Waals radius, i.e., about 1.0 Å grid resolution for proteins. Comparing to traditional PB solvers, the MIBPB-III is a few orders of magnitude more accurate at a given mesh size and about three times faster at a given accuracy [157, 67]. The MIBPB is the first and still the only known second-order convergent PB solver for the singular molecular surfaces of biomolecules, where the second order convergence means that the accuracy of the solution improves four times when the mesh size is halved.

The MIBPB solver serves as a powerful tool that provides the electrostatics analysis for nano-transistor and ion channel studies. Apart from the accuracy, the efficiency of the MIBPB solver is another important issue crucial to many applications. Previous MIBPB solvers are typically slow in comparing with other standard finite difference or finite element methods that do not invoke an interface treatment. This comes from the trade-off of the high accuracy and convergency. Since detailed local interface information (the intersection and normal direction, etc) are included, the matrices which represent the discretization of the elliptic operator loss their nice properties, such as symmetry and positive definiteness, especially for complicated surfaces of biomolecules. The latest version of MIBPB solvers is equipped with Krylove subspace (KSP) theory and preconditioner techniques to increase the solver efficiency. Special efforts have been paid on the strategies for the selection

of most suitable linear system solvers for the resulting MIBPB matrices. Two linear solver libraries, the SLATEC (http://people.sc.fsu.edu/~burkardt/f_src/slatec/slatec.html) and the PETSc (<http://www.mcs.anl.gov/petsc/petsc-as/>) are considered in the exploration of linear solvers.

With the development of the MIBPB, a complete software package for solving the Poisson-Boltzmann equation and its generalization is established for bio-molecular electrostatics analysis. It has been applied to the molecular solvation energy calculation, salt effect based binding energy and other biological applications. The MIBPB package is also modified and adopted in the ion channel calculation to supply the electrostatic background of the system. Furthermore, the treatments on the interface singularities, geometric singularities and source charge singularities in MIBPB have been implanted to the algorithms of solving Poisson equation in nano-devices simulations.

1.5.2 Quantum computation for particle transport

To study the transport of quantum particles, the full Schrödinger equation for N particles is computationally unfeasible because of the extremely high number of degrees of freedom. Instead, we consider the density functional theory and utilize the generalized Kohn-Sham equation from the total energy functional. Although being greatly simplified by the one-particle approximation, fully solving the Kohn-Sham equation still gives computational obstacles. However, based on specific physical properties, the Kohn-Sham operator presents distinguish behaviors under different circumstances. Given reasonable assumptions, the computation complexity might be reduced in further.

In usual cases, one can define three types of boundary conditions for the Kohn-Sham equation. The first one is the infinite boundary condition, which is related to a translation invariance of the potential. Under this circumstance or approximations,

the spectrum of the Kohn-Sham operator is absolute continuous and wavefunctions can be represented by plane waves. The second type of boundary condition is the open boundary conditions, which is for the particle transport through a position dependent potential landscape. The Kohn-Sham operator also has an absolute continuous spectrum while the wavefunction need to be carefully solved. Finally, the confined boundary condition is defined when a particle is confined in a given potential surface and the probability of finding the particle out of the region is zero. Under this circumstance, the energy level of the target particle is quantized, i.e., the Kohn-Sham operator yields a discrete spectrum.

For proton channel computation, the primary motion is the proton transport process, so the transport boundary condition should be imposed along the transport direction, in which the channel pore is aligned. The bath regions are considered as ion reservoirs which are assumed to have infinitely large dimension and perpendicular to the transport direction. Therefore, the Hamiltonian of the proton has an absolutely continuous spectrum in the corresponding plane. In contrast, the channel regions give strong confinement to protons because they are assumed not to penetrate the channel wall. Consequently, a discrete spectrum results in the non-transport direction in the channel region. Similar treatments are also applied in the nano-devices simulations. For MOSFETs, the major direction is also the transport direction in which the transport boundary conditions are applied. For others, in the direction not confined by the insulator materials, the potential may be assumed position invariance. Thus, wavefunctions can be regarded as planewaves and corresponding energy could be integrated out. While for directions along which the material is confined by insulators, confined boundary conditions are applied and discrete sub-bands result. In both proton channel and nano-devices simulations, a series of 1D transport problems eventually need to be solved after the operator splitting. Additionally, corresponding statistical distribution functions indicate that higher energies actually contribute less

to the whole system. Therefore, the computation complexity of the Kohn-Sham equation is greatly reduced since the high energy components can be abandoned within numerical tolerance. Meanwhile, because of the classification of the boundary condition or dimension decomposition of the Kohn-Sham operator, the particle density has different structures in corresponding regions of proton system and various MOSFETs.

1.5.3 The self-consistent iteration

As introduced in earlier sections, two partial differential equations are derived from the total free energy functional via the variational principle. The Poisson-type equations give the electrostatic background for particle transport and the Kohn-Sham equation provides the number density of quantum particles. These two types of equations are strongly coupled. The electrostatics servers part of the potential energy in the Hamiltonian of the Kohn-Sham equation, while the number density of quantum particles forms part of charge sources in the Poisson-type equation. Therefore, these two equations need to be solved in a self-consistent manner till converges are achieved. There are several interesting mathematical issues about the self-consistent system. For example, the well-posedness, i.e., the existence and uniqueness of solutions of the PDE system needs to be carefully justified. There are several preliminary results available in the literature, on the existence of solutions from the fixed point mapping view [51, 120].

For the overall self-consistent iteration, directly inserting the solution of one PDE to another usually does not give good convergence, as verified in many applications[149, 27]. Instead, the Gummel iteration [50] converts the whole self-consistent iteration into an outer and an inner iteration, and is proved in practices to be a more efficient way to solve the system. Essentially, it is pointed out in this thesis that the Gummel iteration is one type of Newton's method. Therefore, various inexact Newton's methods could be employed to enhance the numerical efficiency. Theoretical analy-

sis on the iteration convergence of the Newton's method is also available. Another technique to increase the self-consistent convergence is the relaxation factor scheme [7]. It origins from the fixed point map and converts the process of finding the fixed point (solution) of the iteration to the equivalent process of seeking the steady state of an ordinary differential equation (ODE). As a result, many theories and schemes of solving ODEs can be borrowed to study the self-consistent iteration. To summarize, other than modeling and computation, there are several interesting mathematical issues in the bio-nano system, which are worthwhile to be studied and explored in the future research.

Chapter 2

Modeling and computation of nano-electronic transistors

In this chapter we present the theoretical formulation of our model of material interface and individual dopants for nano-electronic transistors.

2.1 Theory and models

2.1.1 Multiscale energy functional on an equal footing

In the continuum mechanical description, the electric field $\mathbf{E}(\mathbf{r})$ can be expressed as the negative gradient of the electrostatic potential $u(\mathbf{r})$, i.e., $\mathbf{E}(\mathbf{r}) = -\nabla u(\mathbf{r})$. The standard Poisson equation can be derived from Gauss's law describing how electric charge can create and alter electric fields

$$\nabla \cdot \epsilon(\mathbf{r})\mathbf{E}(\mathbf{r}) = -\nabla \cdot \epsilon(\mathbf{r})\nabla u(\mathbf{r}) = n_{\text{total}}(\mathbf{r})q \quad (2.1)$$

where n_{total} is the free charge number density, q is the electron charge and $\epsilon(\mathbf{r})$ is the permittivity. The electrostatic energy functional induced by the given free number

density of charge $n_{\text{total}}(\mathbf{r})$ can be given by

$$E_{\text{Electrostatic}}[u] = \int \left[\frac{\epsilon(\mathbf{r})}{2} |\nabla u(\mathbf{r})|^2 - u(\mathbf{r}) n_{\text{total}}(\mathbf{r}) q \right] d\mathbf{r}. \quad (2.2)$$

where the integration is over \mathbb{R}^3 . The variation of $E_{\text{Electrostatic}}[u]$ with respect to u via the Euler-Lagrange equation recovers the Poisson equation

$$\frac{\delta E_{\text{Electrostatic}}[u]}{\delta u} \Rightarrow -\nabla \cdot \epsilon(\mathbf{r}) \nabla u(\mathbf{r}) - n_{\text{total}}(\mathbf{r}) q = 0. \quad (2.3)$$

In the quantum mechanical description, the electron density is given by

$$n(\mathbf{r}) = \sum_j |\Psi_j(\mathbf{r})|^2 f(E_j - \mu), \quad (2.4)$$

where Ψ_j are the Kohn-Sham orbitals [93, 117], and

$$f(E_j - \mu) = \frac{1}{1 + e^{(E_j - \mu)/k_B T}}, \quad (2.5)$$

is the Fermi-Dirac distribution with μ being the Fermi energy, k_B the Boltzmann constant and T the temperature. The electron energy functional is

$$E_{\text{Electronic}}[n] = \int \left[\sum_j \frac{\hbar^2 f(E_j - \mu)}{2m(\mathbf{r})} |\nabla \Psi_j(\mathbf{r})|^2 + \frac{1}{2} \int \frac{n(\mathbf{r}) n(\mathbf{r}') q^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right. \\ \left. - \sum_j \frac{Z_j n(\mathbf{r}) q^2}{|\mathbf{r} - \mathbf{r}_j|} + E_{\text{XC}}[n(\mathbf{r})] - \sum_j E_j f(E_j - \mu) |\Psi_j(\mathbf{r})|^2 \right] d\mathbf{r}, \quad (2.6)$$

where $m(\mathbf{r})$ is the position dependent electron mass, $\hbar = h/(2\pi)$ with h being the Planck constant, Z_j is the nuclear charge at position \mathbf{r}_j , $E_{\text{XC}}[n(\mathbf{r})]$ is the exchange correlation term and E_j are eigenvalues. Energy minimization with respect to Ψ_j^* ,

which is the complex conjugate of Ψ_j , leads to the Kohn-Sham equation [93]

$$\begin{aligned} \frac{\delta E_{\text{Electronic}}[n]}{\delta \Psi_j^*} &\Rightarrow \left(-\nabla \cdot \frac{\hbar^2}{2m} \nabla + \int \frac{n(\mathbf{r}')q^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - \sum_j \frac{Z_j q^2}{|\mathbf{r} - \mathbf{r}_j|} + U_{\text{XC}}[n(\mathbf{r})] \right) \\ &\times f(E_j - \mu) \Psi_j(\mathbf{r}) - E_j f(E_j - \mu) \Psi_j(\mathbf{r}) = 0 \end{aligned} \quad (2.7)$$

where $U_{\text{XC}}[n(\mathbf{r})] = \frac{\delta E_{\text{XC}}[n(\mathbf{r})]}{\delta n(\mathbf{r})}$ is the exchange correlation potential. It is convenient to cast the Kohn-Sham equation in the form of the Schrödinger equation

$$\left(-\nabla \cdot \frac{\hbar^2}{2m(\mathbf{r})} \nabla + U(\mathbf{r}) \right) \Psi_j(\mathbf{r}) = E_j \Psi_j(\mathbf{r}), \quad (2.8)$$

where the potential energy $U(\mathbf{r})$ includes all the interaction potential energies in Eq. (2.7).

At nano scale, there should be a unified framework to bring the continuum mechanics and quantum mechanics on an equal footing. To establish the relation between the Poisson equation and the Kohn-Sham equation, let set

$$n_{\text{total}}(\mathbf{r})q = n(\mathbf{r})(-q) + n_n(\mathbf{r})q$$

where we recognize that the electron charge is negative and nuclear charge is positive. Here, the nuclear number density is given by $n_n(\mathbf{r}) = \sum_k Z_k \delta(\mathbf{r} - \mathbf{r}_k)$. Then, the solution to the Poisson equation in the free space is

$$u(\mathbf{r}) = - \int \frac{n(\mathbf{r}')q}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \sum_k \frac{qZ_k}{|\mathbf{r} - \mathbf{r}_k|}. \quad (2.9)$$

where Z_k and \mathbf{r}_k indicate the charge magnitude and position of k th nuclear. Therefore, we introduce a multiscale total energy functional as

$$E_{\text{Total}}[u, n] = \int \left[\frac{\epsilon(\mathbf{r})}{2} |\nabla u(\mathbf{r})|^2 - u(\mathbf{r}) n_{\text{total}}(\mathbf{r}) q - \sum_j \frac{\hbar^2 f(E_j - \mu)}{2m(\mathbf{r})} |\nabla \Psi_j(\mathbf{r})|^2 - E_{\text{XC}}[n(\mathbf{r})] + \sum_j E_j f(E_j - \mu) |\Psi_j(\mathbf{r})|^2 \right] d\mathbf{r}. \quad (2.10)$$

To optimize the total energy, we consider the variation of $E_{\text{Total}}[u, n]$ with respect to u

$$\frac{\delta E_{\text{Total}}[u, n]}{\delta u} \Rightarrow -\nabla \cdot \epsilon(\mathbf{r}) \nabla u(\mathbf{r}) - n_{\text{total}}(\mathbf{r}) q = 0. \quad (2.11)$$

This is the standard Poisson equation. Similarly, by variation of $E_{\text{Total}}[u, n(\mathbf{r})]$ with respect to Ψ_j^* , we have

$$\frac{\delta E_{\text{Total}}[u, n]}{\delta \Psi_j^*} \Rightarrow - \left(-\nabla \cdot \frac{\hbar^2}{2m(\mathbf{r})} \nabla + u(\mathbf{r})(-q) + U_{\text{XC}}[n(\mathbf{r})] \right) \Psi_j(\mathbf{r}) + \Psi_j(\mathbf{r}) = 0. \quad (2.12)$$

This is exactly the Kohn-Sham equation, Eq. (2.7), because the electrostatic potential, $u(\mathbf{r})$, includes the Coulomb potential effects of both electrons and nuclei, as shown in Eq. (2.9). Since the electron density in the Poisson equation depends on the solution of the Kohn-Sham equation, which, in turn, depends on the solution of the Poisson equation for the interaction potential, we have a system of self-consistently coupled Poisson-Kohn-Sham equations. The electronic structure obtained from the present theory will be used to evaluate device transport via the NEGF formalism presented in Section 2.1.3.

To our knowledge, this is the first time that the coupled Poisson-Kohn-Sham equations have been derived from the optimization of the total energy functional, Eq. (2.10). Note that the proposed Poisson-Kohn-Sham equations differ from the Poisson-Schrödinger equations commonly used in device modeling [104, 146, 15, 87, 99, 8, 62, 133, 13] in the following aspects: (i) The solution of the Poisson equation

with the nature boundary condition reproduces the correct Coulomb potential for the Kohn-Sham equation. This consistence does not exist in the commonly used Poisson-Schrödinger equations. (ii) The inclusion of the exchange correlation functional allows the construction of various density functional approximations. (iii) While the electron density is computed from the Kohn-Sham equation, the nuclear density is prescribed as point charges. This framework can be used as a starting point for formulating other linear scaling approximations, such as the pseudopotential method, density functional tight binding method [79] and divide and conquer method [156, 101]. (iv) As the mass is a function of position, the effective mass approximation can be generalized to describe different band structures in different regions, including interconnects. (v) Moreover, the present approach also enables the treatment of individual doping atoms, defects and traps as additional charge sources, i.e., replacing $n_n(\mathbf{r})$ with $\sum_{\beta} n_{\beta}(\mathbf{r})$, where $\beta = \text{nuclei, n-dopants, p-dopants and defects}$. The doping density functions are discussed in Section 2.1.4. (vi) Interactions of electrons and other particles, such as phonons, photons and excitations can also be allowed in the total energy functional. Finally, it is emphasized that the main contribution of the proposed total energy functional framework is that it enables the treatment of continuum mechanism and quantum mechanics on an equal footing at nano scale and provides a unified derivation of coupled Poisson-Kohn-Sham equations in a consistent manner.

2.1.2 Electronic system

The difficulty of solving the full-scale Kohn-Sham equation is one of the major obstacles in the modeling and simulation of nano electronic devices. In particular, the consideration of many-body and multiband interactions is very time consuming. Typical computations are often conducted under the single electron approximation, which can provide a reasonable account of the quantum effect in nano electronic device. By inspecting the device geometry and possible symmetry, one can further reduce the

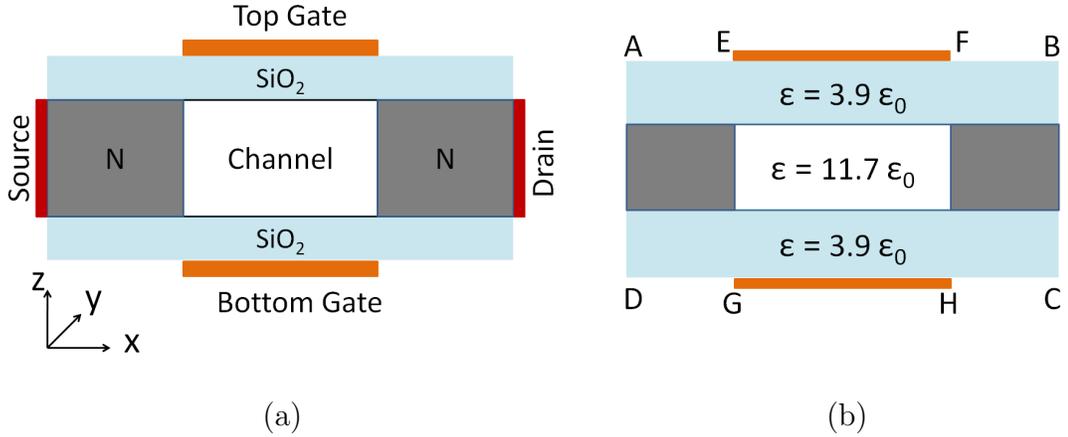


Figure 2.1: Illustration of a double-gate MOSFET with its y -direction being infinitely long. (a) Configuration of the double-gate MOSFET; (b) Computational domain. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

computational dimensions and/or domain by the decomposition of the Kohn-Sham equation and using appropriate transport models [44, 124, 120]. These aspects are discussed with a double-gate MOSFET and a four-gate MOSFET in this subsection. An efficient transport model, the NEGF method, is described in the next subsection.

Double-gate MOSFET

For a double-gate MOSFET [44, 124], the computational dimensions can be split into a confined direction, a transport direction and an infinite direction. Figure 2.1 depicts the geometric configuration and computational domain of the double-gate MOSFET. A detailed description of the involved parameters for the double-gate MOSFET is provided in section 2.3.1. Along the infinite direction, the potential is assumed as translation invariant and the associated Schrödinger operator has an absolutely continuous spectrum. It can be solved with the plane waves. In the confined direction where the gate voltages are applied, the associated Schrödinger operator which in-

cludes the electrostatic potential, is essentially compact, despite that the potential may admit a finite number of point singularities. Therefore, the energy spectrum is discrete in the confined direction. The transport direction accounts for the charge carrier motion in the channel connecting the source and drain contacts. The associated Schrödinger operator also has an absolutely continuous spectrum. The corresponding scattering states can be evaluated with appropriate incoming and outgoing waves, and are subject to the potential consisting of the eigenvalues computed from the confined direction. Without the loss of generality, we denote x , y and z directions for transport, infinite and confined directions, respectively. Under this setting, the wavefunction can be expressed as

$$\Psi_{j,k_x,k_y}(x, y, z) = \psi_{j,k_x}(x, z)\chi_{k_y}(y),$$

where the index j denotes the j th eigenmode of the discrete spectrum and $\chi_{k_y}(y) = e^{ik_y y}$ is the plane wave of given wavenumber k_y . As such, the Kohn-Sham equation of a single electron can be split into two parts

$$\left[-\frac{\hbar^2}{2} \left(\frac{\partial}{\partial x} \frac{1}{m_x} \frac{\partial}{\partial x} + \frac{\partial}{\partial z} \frac{1}{m_z} \frac{\partial}{\partial z} \right) + U \right] \psi_{j,k_x}(x, z) = E_{j,k_x} \psi_{j,k_x}(x, z) \quad (2.13)$$

$$-\frac{\hbar^2}{2} \frac{\partial}{\partial y} \frac{1}{m_y} \frac{\partial}{\partial y} \chi_{k_y}(y) = \frac{\hbar^2 k_y^2}{2m_y} \chi_{k_y}(y), \quad (2.14)$$

where $E_{j,k_x} = E_{j,k_x,k_y} - \frac{\hbar^2 k_y^2}{2m_y}$ with E_{j,k_x,k_y} being the total energy of the system, and $U(x, z)$ is the electrostatic potential energy satisfying the Poisson equation. The solution of Eq. (2.13) requires the information of U .

Note that for the double-gate MOSFET, the y direction is homogeneous, we therefore can solve the Poisson equation in two dimensions. Furthermore, on the $x - z$ plane, due to the confinement along the z direction provided by the insulator layers,

combined with the assumption that the device geometry does not change significantly along transport direction, the system yields discrete states only in the z direction. We split the wavefunction as

$$\psi_{j,k_x}(x, z) = \psi_j \psi_{k_x}, \quad (2.15)$$

where $\psi_j(x_0; z)$ is the discrete eigenstate in the z direction for a given x_0 label and $\psi_{k_x}(x)$ is a scattering state in the transport direction x . The eigenvalue problem with the Dirichlet boundary condition is

$$\left[-\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{m_z} \frac{d}{dz} + U(x_0; z) \right] \psi_j(x_0; z) = \varepsilon_j(x_0) \psi_j(x_0; z), j = 1, 2, \dots \quad (2.16)$$

$$\psi_j(x_0; z) = 0 \quad \text{on} \quad \partial\Omega_D \quad (2.17)$$

where $\varepsilon_j(x_0)$ represents the energy of the j th discrete subband at x_0 . Here $\partial\Omega_D$ is the boundary of the silicon layer at where the wavefunction is forced to be zero because of the confinement. Based on this formalism, one only needs to calculate the quantum transport along the x -direction. We therefore end up with a scattering problem

$$\left[-\frac{\hbar^2}{2} \frac{d}{dx} \frac{1}{m_x} \frac{d}{dx} + \varepsilon_j(x) \right] \psi_{k_x}^j(x) = E_{j,k_x} \psi_{k_x}^j(x), \quad (2.18)$$

Since $\varepsilon_j(x)$ varies along the x axis, it serves as the potential for the scattering problem. The superscript in the scattering wavefunction $\psi_{k_x}^j(x)$ indicates that the potential is associated with the j th discrete subband. Similarly, the subband label j on E_{j,k_x} indicates the scattering potential used in Eq.(2.18). From the physical point of view, the scattering energy is conserved during the scattering process. The transmission and reflection coefficients can be computed based on each given initial energy E_{j,k_x} ,

which is a part of the absolute continuum spectrum.

Equation (2.18) can be solved in many ways, such as time dependent and time independent means. However, it is convenient to use the NEGF strategy [44] as described in Section 2.1.3, which provides not only the solution to Eq. (2.18), but also the desirable transport quantities.

To solve the Poisson equation for $U(x, z)$, one needs to determine the electron density $n(x, z, y)$ according to Eq. (2.4)

$$n(\mathbf{r}) = \sum_{j, k_x, k_y} |\psi_j(x_0; z)|^2 |\chi_{k_y}(y)|^2 |\psi_{k_x}^j(x)|^2 f(E_{j, k_x, k_y} - \mu) \quad (2.19)$$

where $n(\mathbf{r})$ is homogeneous in y direction. In fact, the scattering part $|\psi_{k_x}^j(x)|^2$ does not need to be explicitly evaluated in the NEGF formulation. The further simplification of $n(\mathbf{r})$ is discussed in Section 2.1.3.

Four-gate MOSFET

For a four-gate MOSFET, we denote x , y and z directions for the transport, confined and confined directions, respectively. In the confined $y - z$ directions where the gate voltages are applied, the associated Schrödinger operator is essentially compact, which leads to discrete energy states for the charge carrier. Figure 2.2 depicts the geometric configuration and computational domain of the four-gate MOSFET. A detailed description of the four-gate MOSFET parameters can be found in Section 2.3.1. As in the double-gate case, the associated Schrödinger operator has an absolutely continuous spectrum in the transport direction. The total energy is given by E_{j, k_x} and total wavefunction is Ψ_{j, k_x} . By assuming $\Psi_{j, k_x} = \Psi_j(x : y, z) \Psi_{k_x}^j(x)$, we

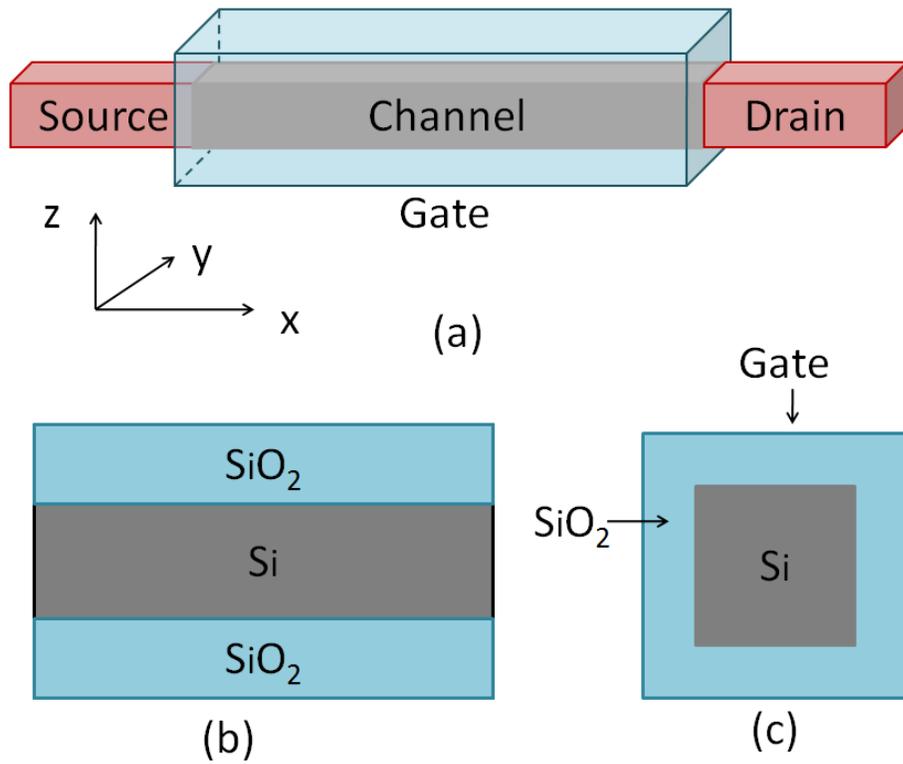


Figure 2.2: An illustration of a four-gate MOSFET (i.e., silicon nanowire transistor). (a) Configuration; (b) Cross section at $y = 0$ for the computational domain; (c) Cross section at $x = 0$.

therefore can split the Kohn-Sham equation as

$$\left[-\frac{\hbar^2}{2} \left(\frac{\partial}{\partial y} \frac{1}{m_y} \frac{\partial}{\partial y} + \frac{\partial}{\partial z} \frac{1}{m_z} \frac{\partial}{\partial z} \right) + U(x_0; y, z) \right] \psi_j = \varepsilon_j(x_0) \psi_j \quad (2.20)$$

$$\psi_j(x_0; y, z) = 0 \quad \text{on} \quad \partial\Omega_D;$$

$$\left[-\frac{\hbar^2}{2} \frac{\partial}{\partial x} \frac{1}{m_x} \frac{\partial}{\partial x} + \varepsilon_j(x) \right] \psi_{k_x}^j(x) = E_{j,k_x} \psi_{k_x}^j(x), \quad (2.21)$$

where $\varepsilon_j(x_0)$ is the j th eigenvalue of the 2D problem at position x_0 , and $\psi_j = \psi_j(x_0; y, z)$ is the corresponding eigenfunction. Here $\psi_{k_x}^j(x)$ is the scattering wavefunction associated with the scattering potential $\varepsilon_j(x)$ and energy E_{j,k_x} . The transport equation (2.21) is the same as that of Eq. (2.18) and can be solved with the NEGF method. For a given set of scattering energies, $E_{k_x}^j = \frac{\hbar^2 k_x^2}{2m_x}$, the Poisson equation (2.11) is solved in 3D and its electronic density source term is

$$n(\mathbf{r}) = \sum_{j,k_x} |\psi_j(x_0; y, z)|^2 |\psi_{k_x}^j(x)|^2 f(E_{j,k_x} - \mu). \quad (2.22)$$

The further evaluation of the density is discussed in Section 2.1.3.

In both double-gate and four-gate MOSFET calculations, it is possible to consider subband interactions by using a variety of combinations of the discrete energy levels. As such, subband states belonged to different levels are used along the x axis. Obviously, this approach can lead to improved transport properties and an increase in computational cost. However, high energy modes, particularly those modes whose energies are significantly higher than the scattering barrier, do not play much role in the transport calculation.

2.1.3 Transport system

This section briefly describes the NEGF formalism in a setting that is consistent with the double-gate MOSFET and four-gate MOSFET studied in the present work.

Additionally, a detailed description of the electron density is also given.

The Non-equilibrium Green's function (NEGF) formalism

Without the loss of generality, we consider the NEGF formalism in a multichannel setting in \mathbb{R}^{3-l} . Here l is the number of non-scattering dimensions, it equals 2 for double-gate and four-gate MOSFETs discussed in the last subsections. We define the whole open system on the domain $\Omega = \Omega_D \cup (\sum_{\alpha} \Omega_{\alpha})$, which consists of the device domain Ω_D and the union of (multiple) contact domains Ω_{α} , such as the source and drain. Let $\Gamma_{\alpha} = \Omega_D \cap \Omega_{\alpha}$ denotes the intersection boundaries of the device domain and contacts. Here, Ω_{α} may extend to infinity but only the Ω_D (or plus small portion of Ω_{α}) is the computational domain of interest. In the framework of the NEGF, the Green's operator (function) on Ω_D is defined as the inverse operator

$$\mathbf{G}^{-1}(E) = E\mathbf{I} - H = E\mathbf{I} - H^0 - \sum_{\alpha} \Sigma^{\alpha} \quad (2.23)$$

where E is the total energy of the scattering system, \mathbf{I} is the identity operator, $H = H^0 + \sum_{\alpha} \Sigma^{\alpha}$ is the full *scattering* Hamiltonian and H^0 is the Hamiltonian of the single charge carrier associated with the scattering potential. In order to reduce the computational cost, the infinite (or large) contact domain Ω_{α} needs to be chopped off and is restricted on the domain of interest Ω_D . Each of the self-energy operators Σ^{α} is solely defined on the corresponding Γ^{α} and reveals the coupling effect of the contacts to the device [44]. In practice, Σ^{α} takes different forms for different numerical discretizations, and more details can be found in Ref. [83]. In the position representation

$$H^0 = -\frac{\hbar^2}{2} \nabla \cdot \left(\frac{1}{m(\mathbf{r})} \nabla \right) + U_{\text{scat}}(\mathbf{r}), \quad \mathbf{r} \in \mathbb{R}^{3-l}, \quad (2.24)$$

where $m(\mathbf{r})$ is the space dependent effective mass of the charge carriers and $U_{\text{scat}}(\mathbf{r})$ is the interaction potential for scattering and $U_{\text{scat}}(\mathbf{r}) = \varepsilon_j(x)$ for double-gate and four-gate MOSFETs discussed in the last subsections.

Once the Green's function/operator is defined, all quantities of interest can be calculated. Among these quantities, the scattering wavefunction is given by

$$\psi_E = \lim_{\varepsilon \rightarrow 0} i\varepsilon \mathbf{G}(E + i\varepsilon) \phi_E, \quad (2.25)$$

where ϕ_E is the incoming wavefunction of energy E . The scattering wavefunction satisfies the Schrödinger equation $H\psi_E = E\psi_E$. Additionally, the non-equilibrium charge carrier density operator is given by

$$\rho = \sum_{\alpha} f(H_{\text{full}} - \mu_{\alpha}) \mathbf{A}^{\alpha}(E). \quad (2.26)$$

In the NEGF theory, each contact is assumed in equilibrium state and α is for contact index, μ_{α} is the contact Fermi level for each contact, and f is the Fermi-Dirac distribution. Here $H_{\text{full}} = H + H_{\text{ns}}$ is the full Hamiltonian of the electron system with H_{ns} being the Hamiltonian of the non-scattering system. The spectral function $\mathbf{A}^{\alpha}(E)$ in Eq. (2.26) is given by

$$\mathbf{A}^{\alpha}(E) = \mathbf{G}(E) \mathbf{\Gamma}^{\alpha}(E) \mathbf{G}^{\dagger}(E), \quad (2.27)$$

where $\mathbf{\Gamma}^{\alpha}(E)$ is the broadening operator, which reflects the dissipative effects on the transport from contact region Ω_{α} , and is defined by

$$\mathbf{\Gamma}^{\alpha}(E) = \mathbf{i}[\Sigma^{\alpha}(E) - (\Sigma^{\alpha}(E))^{\dagger}]. \quad (2.28)$$

Moreover, the number density of charge in the system is given by the position

representation of the density operator (2.26)

$$n(\mathbf{r}) = \langle \mathbf{r} | \rho | \mathbf{r} \rangle, \quad (2.29)$$

where $\langle \cdot |$ and $|\cdot \rangle$ are Dirac notations.

Furthermore, the current in the quantum device from the source (S) to drain (D) is calculated via

$$I = \frac{-q}{h} \text{Tr} T_{\text{DS}}(E) [f(H_{\text{full}} - \mu_{\text{S}}) - f(H_{\text{full}} - \mu_{\text{D}})] \quad (2.30)$$

where μ_{S} and μ_{D} are the Fermi levels of source and drain, respectively. The Tr is the trace and $T_{\text{DS}}(E)$ is the transmission operator.

$$T_{\text{DS}}(E) = \mathbf{\Gamma}^{\text{D}}(E) \mathbf{G}(E) \mathbf{\Gamma}^{\text{S}}(E) \mathbf{G}^{\dagger}(E). \quad (2.31)$$

Finally, the number density contributed from the scattering process can also be computed by using the NEGF formalism. This aspect is elaborated in the next subsection.

Electron density

Electron density is an important quantity in the coupled Poisson- Kohn-Sham theory and is used in the Poisson equation to compute the potential. The general expression of the density is given in Eq. (2.29). Due to the subband decomposition, the evaluation of density becomes slightly subtle. One needs to distinguish the subband degree of freedom, the scattering degree of freedom, and the degree of freedom due to the infinity dimensions. In particular, the energy associated with infinity dimensions should be integrated. Whereas the discrete subband energies serving as the potentials for the scattering system and the discrete subband states are summed over. We illustrate the density evaluation in the double-gate MOSFET and the four-gate MOSFET.

As shown in Eq. (2.29), the electron density is given by the position representation of the density operator $n_{\text{scat}}(\mathbf{r}) = \langle \mathbf{r} | \rho | \mathbf{r} \rangle$. For the double-gate MOSFET, the y direction is infinity and homogeneous. The total energy E_{j,k_x,k_y} is distributed to the plane wave, the eigenstate, and the scattering kinetic energy. Since the eigenstate energy acts as the scattering potential energy, the total energy available to the scattering system in Eq. (2.23) is $E_{j,k_x}(x)$. The energy associated with the plane wave can be integrated. To evaluate ρ , we need to use some notation and identities

$$\langle \mathbf{r} | \psi_E \rangle = \psi_E(\mathbf{r}), \quad \frac{1}{2\pi} \int_{\mathbb{R}} |\psi_E \rangle \langle \psi_E| dE = 1, \quad (2.32)$$

and

$$\sum_k g(k) \rightarrow \frac{2}{(2\pi)^d} \int g(k) (dk)^d, \quad (2.33)$$

for a function $g(k)$. We therefore write the electron density of the double-gate MOSFET as

$$n(\mathbf{r}) = \sum_{j,k_x,k_y} |\psi_j(x_0; z)|^2 |\chi_{k_y}(y)|^2 |\psi_{k_x}^j(x)|^2 f(E_{j,k_x,k_y} - \mu) \quad (2.34)$$

$$= \sum_{j,k_x} \sum_{k_y} |\psi_j(x_0; z)|^2 |\psi_{k_x}^j(x)|^2 f(E_{j,k_x} + \frac{\hbar^2 k_y^2}{2m_y} - \mu) \quad (2.35)$$

$$= \sum_{j,k_x} |\psi_j(x_0; z)|^2 |\psi_{k_x}^j(x)|^2 P^j \quad (2.36)$$

$$= \sum_j |\psi_j(x_0; z)|^2 n_{\text{scat}}^j(x, y) \quad (2.37)$$

where $P^j = \sum_{k_y} f(E_{j,k_x} + \frac{\hbar^2 k_y^2}{2m_y} - \mu)$. Since k_y is continuous, we should change the

summation into an integral

$$P_j = \sum_{k_y} f\left(E_{j,k_x} + \frac{\hbar^2 k_y^2}{2m_y} - \mu\right) = \frac{1}{\pi} \int_{-\infty}^{\infty} f\left(E_{j,k_x} + \frac{\hbar^2 k_y^2}{2m_y} - \mu\right) dk_y \quad (2.38)$$

$$= \frac{1}{\sqrt{\pi}} \left(\frac{2m_y k_B T}{\hbar^2}\right)^{\frac{1}{2}} \mathcal{F}_{-\frac{1}{2}}(\mu - E_{j,k_x}) \quad (2.39)$$

where $\mathcal{F}_{-\frac{1}{2}}(\mu - E_{j,k_x})$ is the Fermi integral of order $-\frac{1}{2}$ given by

$$\mathcal{F}_{-\frac{1}{2}}(y) = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{x^{-\frac{1}{2}}}{1 + e^{x-y/k_B T}} dx. \quad (2.40)$$

from Eq. (2.29), we have

$$n(\mathbf{r}) = \sum_{j,k_x} |\psi_j(x_0; z)|^2 |\psi_{k_x}^j(x)|^2 P_j \quad (2.41)$$

$$= \frac{1}{\sqrt{\pi}} \left(\frac{2m_y k_B T}{\hbar^2}\right)^{\frac{1}{2}} \sum_j |\psi_j(x_0; z)|^2 \quad (2.42)$$

$$\times \frac{1}{2\pi} \int_{\mathbb{R}} \sum_{\alpha} \mathcal{F}_{-\frac{1}{2}}(\mu_{\alpha} - E_{j,k_x}) \mathbf{A}^{\alpha}(E_{j,k_x}) dE_{j,k_x}. \quad (2.43)$$

From this expression one can identify that

$$n_{\text{scat}}^j(x, y) = \frac{1}{\sqrt{\pi}} \left(\frac{2m_y k_B T}{\hbar^2}\right)^{\frac{1}{2}} \frac{1}{2\pi} \int_{\mathbb{R}} \sum_{\alpha} \mathcal{F}_{-\frac{1}{2}}(\mu_{\alpha} - E_{j,k_x}) \mathbf{A}^{\alpha}(E_{j,k_x}) dE_{j,k_x} \quad (2.44)$$

At each given location along the x direction, E_{j,k_x} is given by $E_{j,k_x}(x) = \varepsilon_j(x) + E_{k_x} = \varepsilon_j(x) + \frac{\hbar^2 k_x^2}{2m_x}$. Here, $\varepsilon_j(x)$ are the eigenvalues of the 1D confined system. Based on this expression and Eq.(2.15), the density $n(\mathbf{r})$ given in Eq. (2.41) can be calculated. The Kohn-Sham equation and the Poisson equation are completely coupled and their solutions have to be pursued iteratively.

For the four-gate MOSFET, the density can be evaluated as the follow

$$n(\mathbf{r}) = \sum_{j,k_x} |\psi_j(x_0; y, z)|^2 |\psi_{k_x}^j(x)|^2 f(E_{j,k_x} - \mu) \quad (2.45)$$

$$= \sum_j |\psi_j(x_0; y, z)|^2 n_{\text{scat}}^j(x), \quad (2.46)$$

where $n_{\text{scat}}^j(x)$ is given by

$$n_{\text{scat}}^j(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \sum_{\alpha} f(E_{j,k_x} - \mu_{\alpha}) \mathbf{A}^{\alpha}(E_{j,k_x}) dE_{j,k_x}, \quad (2.47)$$

where $E_{j,k_x}(x) = \varepsilon_j(x) + E_{k_x} = \varepsilon_j(x) + \frac{\hbar^2 k_x^2}{2m_x}$. Here, $\varepsilon_j(x)$ are the eigenvalues of the 2D confined systems, which change over different position x in the transport direction. Note that in Eqs. (2.42) and (2.47), the integrations over energy E_{j,k_x} are carried out at each given position x .

2.1.4 Electrostatic system

Individual dopant model

In earlier individual models, the discrete dopants are approximated by either dilated Gaussian functions [84] or constant charges supported by small regions. All these models are parameter-dependent. In this paper, we propose a point doping model for individual dopants and define the doping density as

$$n_{\beta}(\mathbf{r}) = \sum_j c_{j\beta} \delta(\mathbf{r} - \mathbf{r}_{j\beta}), \quad (2.48)$$

where where $\beta =$ n-dopants and p-dopants, c_{β} are charges of doping atoms and $\delta(\mathbf{r} - \mathbf{r}_{j\beta})$ is the Dirac delta function at position $\mathbf{r}_{j\beta}$. Theoretically, this doping charge source can be added to the total energy functional, Eq. (2.10). This model

provides a better description for microscopic dopants. In fact, it has a connection to the usual Gaussian function model characterized by the influence domain (σ_β) of each dopant

$$n_{\beta, \sigma_\beta}(\mathbf{r}) = \sum_j \frac{c_{j\beta}}{(2\pi\sigma_\beta^2)^{3/2}} e^{-(\mathbf{r}-\mathbf{r}_{j\beta})^2/2\sigma_\beta^2} \xrightarrow{\sigma_\beta \rightarrow 0} n_\beta(\mathbf{r}) = \sum_j c_{j\beta} \delta(\mathbf{r} - \mathbf{r}_{j\beta}). \quad (2.49)$$

The point doping model is recovered if the influence domains are set to zero. Computationally, the delta functions give rise to unbounded electrostatic values locally and is numerically difficult to deal with. This singularity can be alleviated by the Dirichlet to Neumann mapping (DNM) method, which analytically resolves the delta functions and leads to a set of flux jump conditions at the interface. The effective use of the DNM method requires the careful enforcement of additional flux jump conditions. To this end, the MIB framework developed for the Poisson equation will be used. Tens of thousands of atoms can be handled in this approach.

Silicon-insulator-interface modeling

The recognition of material interfaces in MOSFETs implies the acknowledgment of discontinuous material properties or coefficients across the interfaces. This has profound consequences in the well-posedness and numerical convergence of the Poisson equation. For simplicity, the electron is assumed as the majority of charge carriers and then the hole density is neglected. In the present work, we consider the Poisson equation of the form

$$-\nabla \cdot (\epsilon \nabla u) = \rho(\mathbf{r}) \quad \text{in } \Omega \quad (2.50)$$

$$[u] = u \chi_{\Omega_{\text{Si}}} - u \chi_{\Omega_{\text{SiO}_2}} = 0 \quad \text{along } \Gamma_{\text{Si/SiO}_2} \quad (2.51)$$

$$[\epsilon u_n] = \epsilon_{\text{Si}} u_n \chi_{\Omega_{\text{Si}}} - \epsilon_{\text{SiO}_2} u_n \chi_{\Omega_{\text{SiO}_2}} = 0 \quad \text{along } \Gamma_{\text{Si/SiO}_2}, \quad (2.52)$$

where

$$\rho(\mathbf{r}) = q \left[N_{\text{D}} - N_{\text{A}} - n(\mathbf{r}) + \sum_j^{M_{\text{D}}} c_{j_{\text{D}}} \delta(\mathbf{r} - \mathbf{r}_{j_{\text{D}}}) - \sum_j^{M_{\text{A}}} c_{j_{\text{A}}} \delta(\mathbf{r} - \mathbf{r}_{j_{\text{A}}}) \right]$$

and $u_n = (\nabla u) \cdot \mathbf{n}$ with \mathbf{n} being the interface normal direction, letters “A” and “D” denote acceptor and donor, respectively, and M_{A} and M_{D} are numbers of discrete acceptors and donors, respectively. Being a multiscale model, some doping regions in the device may still be modeled as continuum, and thus the continuous doping functions N_{A} and N_{D} are reserved for these specific parts. This treatment is reasonable. For example, when the doping in the channel is small and the system is dominated with electron ballistic transport, the continuum doping treatment is a good approximation. Another case is that when the voltage threshold lowering effect is studied, individual dopants are used for the channel region while continuum doping treatment can be used in the source and drain regions. The computational domain Ω is divided into silicon (Si) region Ω_{Si} and silicon dioxide (SiO_2) insulator layers Ω_{SiO_2} . The interface is defined as $\Gamma_{\text{Si/SiO}_2}$, i.e., $\Omega = \Omega_{\text{Si}} \cup \Omega_{\text{SiO}_2}$, $\Gamma_{\text{Si/SiO}_2} = \Omega_{\text{Si}} \cap \Omega_{\text{SiO}_2}$. It follows that the dielectric constant ϵ is set to as ϵ_{Si} and ϵ_{SiO_2} in corresponding regions. The solution of the Poisson equation $u(\mathbf{r})$ is restricted to the Si and SiO_2 regions, and denoted as $u\chi_{\Omega_{\text{Si}}}$ and $u\chi_{\Omega_{\text{SiO}_2}}$ respectively, where χ_{Ω} is the characteristic function on set Ω . Models (2.51) and (2.52) indicate that the Poisson equation is subject to the jump conditions along the Si/ SiO_2 interface, where the jump conditions reveals the continuities of the potential landscape and its flux.

Although the jump conditions are trivial in physical sense, if no specific numerical algorithm is applied, the discontinuity induced non-smoothness in the solution will be smeared. As such, the numerical scheme is of low accuracy and convergence. A novel numerical scheme, the matched interface and boundary (MIB) method will be illustrated in the next section. The boundary conditions of the Poisson equation

are the following: it takes Dirichlet boundary conditions where the gate voltages are applied and Neumann boundary conditions for the rest of the device.

2.2 Numerical analysis and implementation

Multigate MOSFETs have become an important means to alleviate channel tunneling and gate leaking in ultimate CMOSs. In this section, we consider two multigate MOSFETs, the double-gate MOSFET and the four-gate MOSFET. The proposed interface model and individual doping model are evaluated based on these two types of MOSFETs. Some mathematical algorithms and numerical analysis are presented in the section.

2.2.1 Decomposition of the Poisson equation

In this section, the Dirichlet-to-Neumann mapping (DNM) formalism is introduced to achieve an accurate and efficient account of randomly distributed dopants formulated in Section 2.1.4.

The proposed interface and individual dopant models (2.50) treat the dopants as Dirac delta functions that pose computational difficulties. As an approximation, delta functions can be distributed to the neighboring grid points [157]. However, due to the application of the interface techniques, the interference of the interface treatment and the distributed delta functions leads to the reduction of accuracy to the first order. As such, the combination of an interface technique and the Dirichlet-to-Neumann mapping (DNM) strategy becomes necessary [67]. This combination can substantially improve computational accuracy and the speed of convergence. The essence of the DNM technique is to split the solution into certain parts, such that the singular source terms can be accounted in one part of the solution analytically. Such a treatment will in general produce an additional Neumann condition either on the

interface or on the boundary.

In the present problem of nonlinear iterations, we can further take the advantage of the solution splitting to accelerate the convergence of the iterations. It is also noted that in the Poisson equation, only the electron density $n(\mathbf{r})$ is directly involved in the loop of self-consistent iterations and it is a regular function. Therefore, we first decompose the solution of the Poisson equation into two parts, a slow varying part and a fast varying part: $u = u^{\text{slow}} + u^{\text{fast}}$. The fast varying part is up-dated during the iterations and solves

$$\begin{cases} -\nabla \cdot (\epsilon \nabla u^{\text{fast}}) = -qn(\mathbf{r}) & \text{in } \Omega \\ u^{\text{fast}} = u_{\text{voltage}} & \text{on } \Gamma_{\text{SiO}_2/\text{Gate}} \\ u_n^{\text{fast}} = 0 & \text{on other boundary} \end{cases} \quad (2.53)$$

where $\Gamma_{\text{SiO}_2/\text{Gate}}$ is the interface between the insulator and the metal contact. At the Si/SiO₂ interface $\Gamma_{\text{Si}/\text{SiO}_2}$, Eq. (2.53) is subject to the jump conditions

$$[u^{\text{fast}}] = u^{\text{fast}} \chi_{\Omega_{\text{Si}}} - u^{\text{fast}} \chi_{\Omega_{\text{SiO}_2}} = 0 \quad \text{along } \Gamma_{\text{Si}/\text{SiO}_2} \quad (2.54)$$

$$[\epsilon u_n^{\text{fast}}] = \epsilon_{\text{Si}} u_n^{\text{fast}} \chi_{\Omega_{\text{Si}}} - \epsilon_{\text{SiO}_2} u_n^{\text{fast}} \chi_{\Omega_{\text{SiO}_2}} = 0 \quad \text{along } \Gamma_{\text{Si}/\text{SiO}_2}. \quad (2.55)$$

The slow varying part u^{slow} solves

$$\begin{cases} -\nabla \cdot (\epsilon \nabla u^{\text{slow}}) = \rho(\mathbf{r}) & \text{in } \Omega \\ u^{\text{slow}} = 0 & \text{on } \Gamma_{\text{SiO}_2/\text{Gate}} \\ u_n^{\text{slow}} = 0 & \text{other boundary} \end{cases} \quad (2.56)$$

subject to jump conditions

$$[u^{\text{slow}}] = u^{\text{slow}} \chi_{\Omega_{\text{Si}}} - u^{\text{slow}} \chi_{\Omega_{\text{SiO}_2}} = 0 \quad \text{along } \Gamma_{\text{Si/SiO}_2} \quad (2.57)$$

$$[\epsilon u_n^{\text{slow}}] = \epsilon_{\text{Si}} u_n^{\text{slow}} \chi_{\Omega_{\text{Si}}} - \epsilon_{\text{SiO}_2} u_n^{\text{slow}} \chi_{\Omega_{\text{SiO}_2}} = 0 \quad \text{along } \Gamma_{\text{Si/SiO}_2}. \quad (2.58)$$

Here u^{slow} is solely contributed from the fixed doping terms, either continuous doping or individual dopants. It is noticed that jump conditions in both (2.53) and (2.56) are decoupled. Therefore, solutions u^{slow} and u^{fast} are decoupled too. As such, we only need to update u^{fast} during the nonlinear iterations.

For a given $n(\mathbf{r})$, the system (2.53) with its jump condition is readily to be solved with the MIB method. Whereas, u^{slow} should be further decomposed as

$$\begin{aligned} u_{\text{slow}} &= u^{\text{cont}} + u^{\text{disc}} \\ &= u^{\text{cont}} \chi_{\Omega_{\text{Si}}} + u^{\text{cont}} \chi_{\Omega_{\text{SiO}_2}} + u^{\text{disc}} \chi_{\Omega_{\text{Si}}} + u^{\text{disc}} \chi_{\Omega_{\text{SiO}_2}} \\ &= u^{\text{cont}} \chi_{\Omega_{\text{Si}}} + (u^{\text{cont}} + u^{\text{disc}}) \chi_{\Omega_{\text{SiO}_2}} + u^{\text{disc}} \chi_{\Omega_{\text{Si}}} \end{aligned}$$

where u^{cont} and u^{disc} are the solution associated with continuous doping and individual doping, respectively. Since there is no doping in Ω_{SiO_2} , we set $u^1 = u^{\text{cont}} \chi_{\Omega_{\text{Si}}} + (u^{\text{cont}} + u^{\text{disc}}) \chi_{\Omega_{\text{SiO}_2}}$ and $u^2 = u^{\text{disc}} \chi_{\Omega_{\text{Si}}}$. Here u^1 solves

$$\begin{cases} -\nabla \cdot (\epsilon \nabla u^1) = q [N_{\text{D}} - N_{\text{A}}] & \text{in } \Omega \\ u^1 = 0 & \text{on } \Gamma_{\text{SiO}_2/\text{Gate}} \\ u_n^1 = 0 & \text{on other} \end{cases} \quad (2.59)$$

subject to jump conditions

$$[u^1] = u^1 \chi_{\Omega_{\text{Si}}} - u^1 \chi_{\Omega_{\text{SiO}_2}} = 0 \quad \text{along } \Gamma_{\text{Si/SiO}_2} \quad (2.60)$$

$$[\epsilon u_n^1] = \epsilon_{\text{Si}} u_n^1 \chi_{\Omega_{\text{Si}}} - \epsilon_{\text{SiO}_2} u_n^1 \chi_{\Omega_{\text{SiO}_2}} = -\phi \quad \text{along } \Gamma_{\text{Si/SiO}_2}. \quad (2.61)$$

Similarly, u^2 solves

$$\begin{cases} -\nabla \cdot (\epsilon \nabla u^2) = q \left[\sum_j^{M_{\text{D}}} c_{j_{\text{D}}} \delta(\mathbf{r} - \mathbf{r}_{j_{\text{D}}}) - \sum_j^{M_{\text{A}}} c_{j_{\text{A}}} \delta(\mathbf{r} - \mathbf{r}_{j_{\text{A}}}) \right] & \text{in } \Omega \\ u^2 = 0 & \text{on Gate} \\ u_n^2 = 0 & \text{on other} \end{cases} \quad (2.62)$$

subject to the jump conditions

$$[u^2] = u^2 \chi_{\Omega_{\text{Si}}} - u^2 \chi_{\Omega_{\text{SiO}_2}} = 0 \quad \text{along } \Gamma_{\text{Si/SiO}_2} \quad (2.63)$$

$$[\epsilon u_n^2] = \epsilon_{\text{Si}} u_n^2 \chi_{\Omega_{\text{Si}}} - \epsilon_{\text{SiO}_2} u_n^1 \chi_{\Omega_{\text{SiO}_2}} = \phi \quad \text{along } \Gamma_{\text{Si/SiO}_2}. \quad (2.64)$$

In this manner, we have decomposed u^{slow} into two systems (2.59) and (2.62) with corresponding boundary conditions and jump conditions (2.60-2.61) and (2.63-2.64), respectively. The boundary conditions appear trivial, following the homogeneous Dirichlet and Neumann boundary conditions from u^{slow} . However, it can be seen that the jump conditions, specifically the flux jump conditions in Eqs. (2.61) and (2.64), have been revised. They need to be carefully evaluated.

Since u^2 is zero on Ω_{SiO_2} , we can restrict u^2 on Ω_{Si} . As such, interfaces Si/SiO₂ play the role of boundaries, where the homogeneous Dirichlet boundary conditions are applied, which also results in homogeneous jump conditions in Eqs. (2.60) and (2.63). However, one cannot generally have the homogeneous Dirichlet condition and the Neumann boundary condition simultaneously. The Neumann boundary of u^2 on Si/SiO₂ interfaces creates nonhomogeneous jump conditions in Eqs. (2.61) and

(2.64), which is denoted by ϕ and is computed as the follow.

For $u^2 = u^2 \chi_{\Omega_{\text{Si}}}$, it can be written as u^2 for simplicity and it satisfies:

$$\begin{cases} -\epsilon_{\text{Si}} \Delta u^2 = q \left[\sum_j^{M_{\text{D}}} c_{j_{\text{D}}} \delta(\mathbf{r} - \mathbf{r}_{j_{\text{D}}}) - \sum_j^{M_{\text{A}}} c_{j_{\text{A}}} \delta(\mathbf{r} - \mathbf{r}_{j_{\text{A}}}) \right] & \text{in } \Omega_{\text{Si}} \\ u^2 = 0 & \text{on } \Gamma_{\text{Si/SiO}_2} \\ u_n^2 = 0 & \text{on other.} \end{cases} \quad (2.65)$$

To solve equation (2.65), we set $u^2 = u^* + u^0$, in which

$$u^* = \sum_j^{M_{\text{A}}} \frac{q c_{j_{\text{A}}}}{2\pi\epsilon_{\text{Si}}} \ln(|\mathbf{r} - \mathbf{r}_j|) - \sum_j^{M_{\text{D}}} \frac{q c_{j_{\text{D}}}}{2\pi\epsilon_{\text{Si}}} \ln(|\mathbf{r} - \mathbf{r}_j|) \quad (2.66)$$

for 2D simulation, and

$$u^* = - \sum_j^{M_{\text{A}}} \frac{q c_{j_{\text{A}}}}{4\pi\epsilon_{\text{Si}}} \frac{1}{|\mathbf{r} - \mathbf{r}_j|} + \sum_j^{M_{\text{D}}} \frac{q c_{j_{\text{D}}}}{4\pi\epsilon_{\text{Si}}} \frac{1}{|\mathbf{r} - \mathbf{r}_j|} \quad (2.67)$$

for 3D cases. Finally, u^0 solves the harmonic equation with the corresponding boundary condition

$$\begin{cases} -\Delta u^0 = 0 & \text{in } \Omega_{\text{Si}} \\ u^0 = -u^* & \text{on } \Gamma_{\text{Si/SiO}_2} \\ u_n^0 = -u_n^* & \text{on } \Gamma_{\text{Si/Source}} \cup \Gamma_{\text{Si/Drain}}. \end{cases} \quad (2.68)$$

It follows that the jump ϕ in Eqs. (2.61) and (2.64) reads

$$\phi = \epsilon_{\text{Si}}(u_n^* + u_n^0). \quad (2.69)$$

Note that Eqs. (2.66) and (2.67) are actually the fundamental solutions of the Laplacian operator with the δ function source in an unbounded domain. Harmonic equation (2.68) is used to restrict the fundamental solution in the bounded domain

and match the boundary conditions. This procedure of rendering a Neumann boundary condition $u_n^* + u_n^0$ on the interface from the original Dirichlet boundary condition u^* is called Dirichlet-to-Neumann mapping. Meanwhile, the $\phi = \epsilon_{\text{Si}}(u_n^* + u_n^0)$ is used as the jump at Si/SiO₂ interfaces of the flux of u^2 , if the whole domain $\Omega = \Omega_{\text{Si}} \cup \Omega_{\text{SiO}_2}$ is considered. It is easy to identify that the δ functions for individual dopants are exactly treated without any approximation by this decomposition of the whole problem into sub-systems (2.53), (2.59) and (2.62) with corresponding boundary conditions. Other than the naturally induced homogeneous interface jump conditions, the decomposition of the problem also introduces nonhomogeneous jump conditions. Systems (2.53), (2.59) and (2.62) are typical interface problems and are to be solved by the MIB method.

2.2.2 Numerical implementation of the self-consistent iterations

To achieve efficient convergence, we present an inner-outer iteration procedure in this section. The Gummel iterative scheme was proposed in [50] for solving nonlinear coupled equations in all kinds of device applications. The numerical implementation of the iteration scheme for the nano device simulation is provided below.

- Step 0 (Solution of u^{slow}) : This step is out of the main iteration loop to solve for u^{slow} , which is related to the fixed discrete and continuous doping functions. u^* , the singular part of u^{slow} , is calculated by using Eq. (2.66) for 2D or Eq. (2.67) for 3D. The Harmonic part u^0 is solved on the silicon region by using Eq. (2.68). Therefore, $u^* + u^0$ is the part of u^{slow} from the discrete doping, and their corresponding interface jump ϕ is derived via Eq. (2.69). The continuous doping part, u^1 , is solved from via Eq. (2.59) and the corresponding jump condition (2.60). Finally one obtains $u^{\text{slow}} = u^1 + u^* + u^0$.

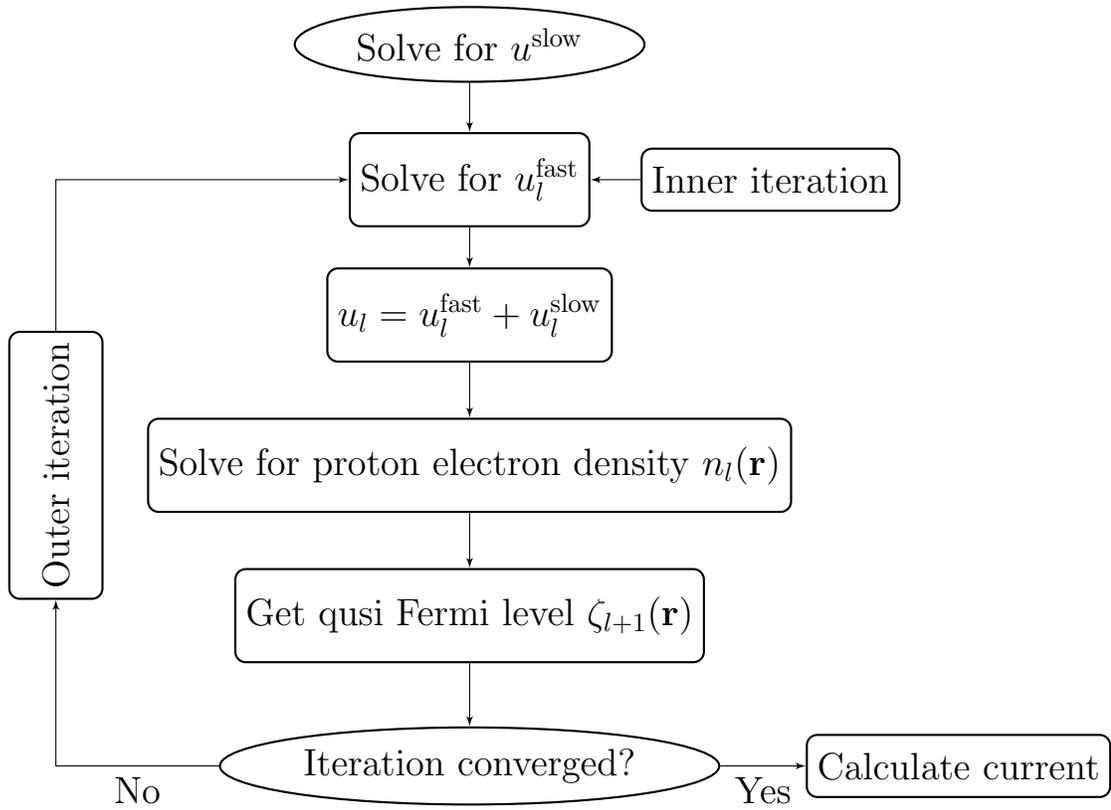


Figure 2.3: Work flow of the overall self-consistent iteration.

- Step 1 (Inner iteration for the Poisson equation): Given the quasi-Fermi level function in l th step, ζ_l , and the calculated u^{slow} , solve the nonlinear Poisson equation for u_l^{fast}

$$\begin{cases} -\nabla \cdot (\epsilon \nabla u_l^{\text{fast}}) = -qn_0 \mathcal{F}_{1/2}(\zeta_l - u^{\text{slow}} - u_l^{\text{fast}}) \\ [u_l^{\text{fast}}] = u_l^{\text{fast}} \chi_{\Omega_{\text{Si}}} - u_l^{\text{fast}} \chi_{\Omega_{\text{SiO}_2}} = 0 \\ [\epsilon \nabla u_l^{\text{fast}} \cdot \mathbf{n}] = \epsilon_{\text{Si}} \nabla u_l^{\text{fast}} \cdot \mathbf{n} \chi_{\Omega_{\text{Si}}} - \epsilon_{\text{SiO}_2} \nabla u_l^{\text{fast}} \cdot \mathbf{n} \chi_{\Omega_{\text{SiO}_2}} = 0 \end{cases} \quad (2.70)$$

where n_0 is the intrinsic density-of-state of an electron system

$$n_0 = \frac{1}{\sqrt{2}} \left(\frac{mk_B T}{\pi \hbar^2} \right)^{3/2} \quad (2.71)$$

where $\mathcal{F}_{1/2}(x)$ is the Fermi-Dirac integral of order 1/2, which takes the form of

$$\mathcal{F}_{1/2}(x) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{y^{1/2} dy}{1 + e^{y - qx/k_B T}}. \quad (2.72)$$

Here, $\zeta(\mathbf{r})$ is the quasi-Fermi potential and will be used as an index function to determine the convergence. The initial value $\zeta_0(\mathbf{r})$ is obtained via a linear interpolation of the source and drain Fermi levels (voltages) over the channel. This initial guess is found to be very effective in our numerical simulations.

- Step 2 (Inner iteration for the Kohn-Sham equation): The outcome of the Step 1 is the electronic potential at l th step, $u_l(\mathbf{r}) = u^{\text{slow}}(\mathbf{r}) + u_l^{\text{fast}}(\mathbf{r})$. Along the transport direction (say x -direction), we slice the electronic potential for each fixed x_0 , and solve the eigenvalue problem either in 1D for the double-gate

MOSFET or in 2D for the four-gate MOSFET

$$\begin{cases} \left(-\frac{\hbar^2}{2m} \nabla_{\mathbf{r}}^2 + U_l(x_0; \mathbf{r}) \right) \psi_j^l(x_0; \mathbf{r}) = \varepsilon_j^l(x_0) \psi_j^l(x_0; \mathbf{r}), \quad j = 1, 2, \dots, \\ \psi_j^l(x_0; \mathbf{r}) = 0 \quad \text{on} \quad \partial\Omega_D^{n-1} \\ \mathbf{r} \in \mathbb{R}^{n-1}, n = 2, 3 \end{cases} \quad (2.73)$$

where $U_l(\mathbf{r}) = u_l(\mathbf{r})(-q)$. Results from this step are set of ψ_j^l and ε_j^l .

- Step 3 (Update density and quasi-Fermi potential): With the available of sub-band energies $\varepsilon_j^l(x)$, we calculate the 3D electron density $n_l(\mathbf{r})$ by using Eq. (2.42) or Eq. (2.46) at the l th step. Once $n_l(\mathbf{r})$ is obtained, one can update the quasi-Fermi level to the $(l + 1)$ th step ζ_{l+1} by inverting the expression $n_l(\mathbf{r}) = n_0 \mathcal{F}_{1/2}(\zeta_{l+1}(\mathbf{r}) - u_l(\mathbf{r}))$

$$\zeta_{l+1}(\mathbf{r}) = u_l(\mathbf{r}) + \mathcal{F}_{1/2}^{-1}(n_l(\mathbf{r})/n_0). \quad (2.74)$$

- Step 4 (Convergence check): The convergence is checked by the criterion of $\|\zeta_{l+1}(\mathbf{r}) - \zeta_l(\mathbf{r})\| < \varepsilon$, where ε is a given small positive number. If the inequality is satisfied, one calculates the current via Eq. (2.30), otherwise go to Step 1.

Figure 2.3 gives the work flow of the present self-consistent iteration scheme.

Remark 1: The reason of a nonlinear Poisson equation applied here is that, during the outer self-consistent iteration loop, the fluctuation of $n(\mathbf{r})$ may undermine the iteration convergence, according to [149]. The use of the Fermi-Dirac integral will average or normalize this kind of fluctuations and thus lead to efficient convergence. This scheme is known as the Gummel iteration [50] when Boltzmann statistics is applied.

Remark 2: There is no analytical form for the Fermi-Dirac integral of order 1/2. Ref. [98] provides a very accurate approximation to the integral by polynomials. It

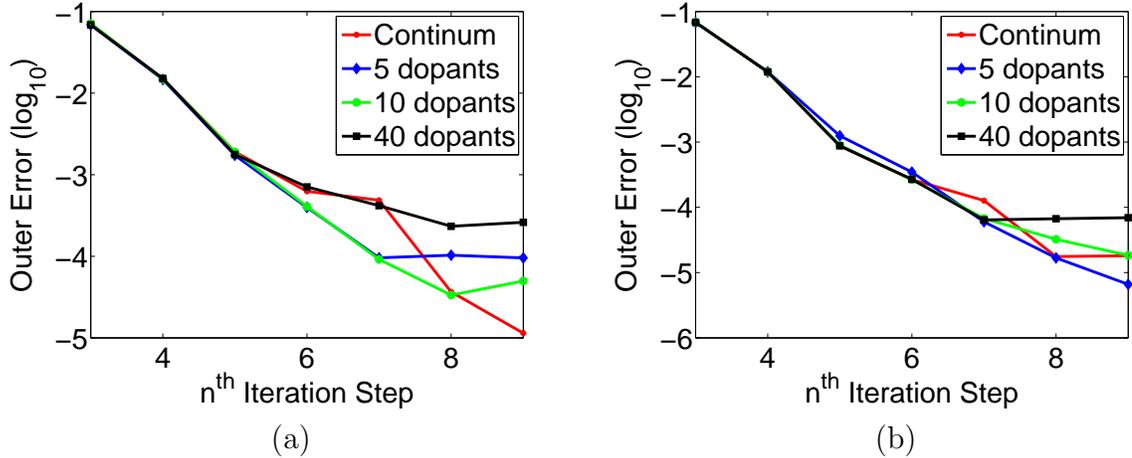


Figure 2.4: Computational errors in simulating a double-gate MOSFET. (a) Individual doping in the source and drain with the dopant distribution shown in the lower panel of Fig. 2.5(a); (b) Individual doping in the channel.

is based on these approximations in this algorithm that all the related calculations are carried out.

Remark 3: Once the nonlinear Poisson equation is employed, it is noted that the initial guess for the whole self-consistent iteration is the quasi-Fermi potential $\zeta(\mathbf{r})$. The choice of the initial guess is based on the physical assumption that $\zeta(\mathbf{r})$ are constants of the voltages at the source/drain and taken as a linear interpolation of two constants along the channel. This initial guess is found to be very effective in our numerical simulations.

2.2.3 Analysis of the model well-posedness, convergence and iteration efficiency

Before numerical simulations are implemented under various physical conditions, some analyses on the robustness and efficiency of numerical algorithms are presented, based on the double-gate MOSFET for simplicity. The detailed device configurations and parameter values are presented in next section.

The well-posedness of the numerical implementation consists of the analysis of the outer iteration loop and inner solution of the nonlinear Poisson equation. We define the K_s, U_s and N_s as the spaces to which the quasi-Fermi functions $\zeta(\mathbf{r})$, electrostatic function $u(\mathbf{r})$, and electron density $n(\mathbf{r})$ belong, respectively. For the whole self-consistent Poisson-NEGF system, it can be interpreted as the application of the fixed point map $\mathcal{T} : K_s \rightarrow K_s$ to the quasi-Fermi potential function

$$\zeta(\mathbf{r}) = \mathcal{T}(\zeta(\mathbf{r})). \quad (2.75)$$

To characterize the details of the map $\mathcal{T} : K_s \rightarrow K_s$, we have the operator $\mathcal{L} : K_s \rightarrow U_s$, which indicates the action of the inner iteration (2.70) from the quasi-Fermi potential to the system potential landscape. Then it is followed by $\mathcal{G} : U_s \rightarrow N_s$, the procedure of using the NEGF scheme and subband decomposition to calculate the electron density from potential. Finally the map $\mathcal{F}_{1/2}^{-1} : N_s \rightarrow K_s$ represents the step that recovers the new quasi-Fermi level from the obtained electron density. The composition of the actions of the above operators yields the definition of the operator \mathcal{T} , representing the outer iteration

$$\mathcal{T} := \mathcal{F}_{1/2}^{-1} \circ \mathcal{G} \circ \mathcal{L}. \quad (2.76)$$

We define a closed convex set

$$K_s = \{\zeta \in L^\infty(\Omega) : C_2 \leq \zeta(\mathbf{r}) \leq C_1, \text{ a.e. in } \Omega\} \quad (2.77)$$

where C_1 and C_2 are lower and upper bounds of ζ in Ω . By analyzing the continuities of these operators and assuming the continuous selection hypothesis, one can prove the existence of the fixed point of the map \mathcal{T} by the Schauder fixed point theorem [51].

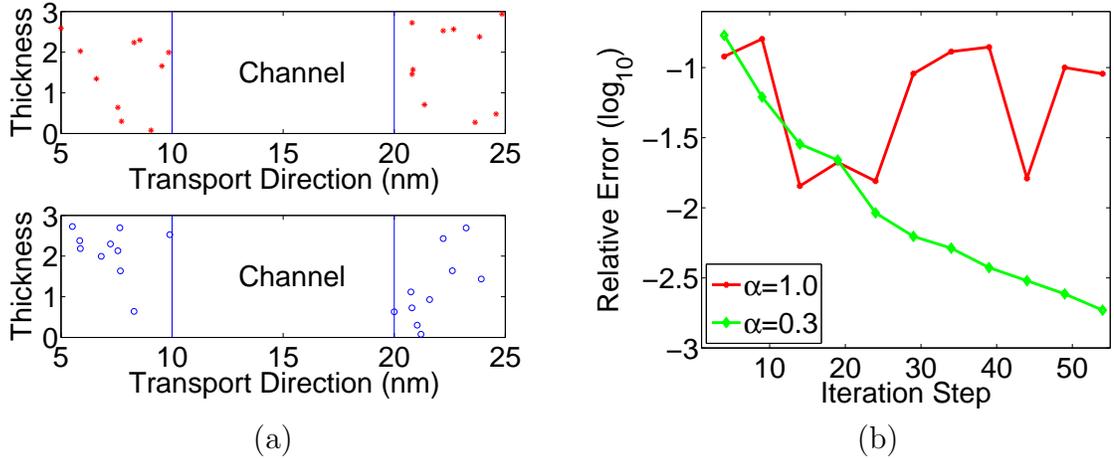


Figure 2.5: Dopant distributions and iteration errors for the double-gate MOSFET. (a) Two distributions of 10 dopants. The distribution in the lower panel may lead to a divergence in the iteration. While the distribution in the upper panel leads to convergence; (b) Relaxation-factor-dependent convergence behaviors for the dopant distribution shown in the left-lower panel.

Another way to check the existence of the solution is the equivalence of the problem with the Poisson-Schrödinger system [120]. The existence of the Poisson-Schrödinger system is proved in [119] and the solution is unique with the constraint of the positivity of carrier concentration.

For the nonlinear Poisson equation, the approximation of the Fermi-Dirac integral has no stronger nonlinearity than the polynomial of order 2 [98]. In this approximation, writing the equation in a variational form, standardly checking the Palais-Smale condition and applying the Soblov embedding theorem, one can easily get the well-posedness analysis.

During the numerical implementation, the Newton-Raphson method guaranteed the convergence of the inner iteration, by picking a reasonable quasi-Fermi level. For the outer iteration, the nonlinearity is too complicated to analyze and no specific analyzing scheme has been designed so far. Many simple potential-density self-consistent loops are employed and proved numerically efficient in many engineering applications.

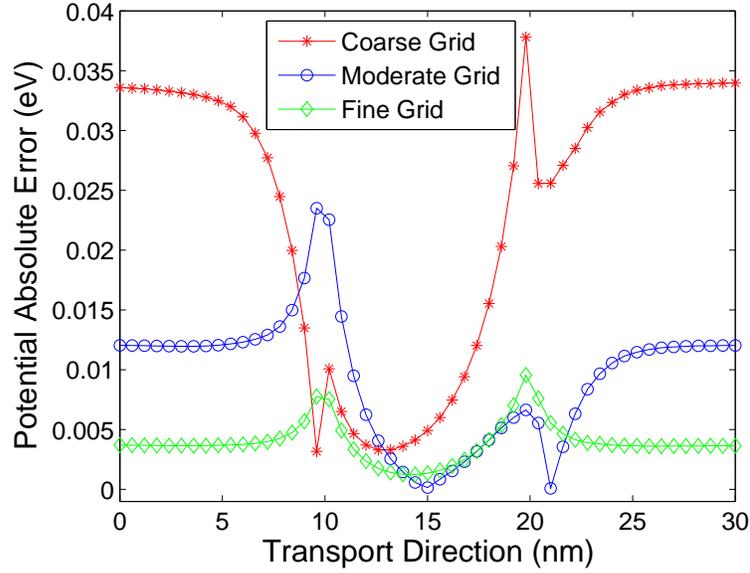


Figure 2.6: Errors of simulating the double-gate MOSFET in the line of $y = 0$ and $z = 0$.

In our model, the convergence of the scheme needs to be re-examined since the introduction of the interface and delta function in the Poisson equation significantly reduce the regularity of the solution.

Figure 2.4 records the outer loop iteration of the system. The left one lists the situation when discrete dopants occur in source and drain ends, and the right one is for individual dopants in the channel. Results for a different number of dopants are compared to the continuous doping model. The horizontal axis is for the iteration loop steps. The first two steps are considered as starting steps and therefore skipped. The vertical axis is the \log_{10} of the absolute convergence error of the potential. Generally, one can conclude that the convergence pattern of the individual dopants model follows that of the continuous model. It is found that the convergence is slight better for the situation that there are dopants in the channel, the reason might be that the gate voltage is applied on the channel part. The δ functions cause high variation of the potential landscape and Dirichlet boundary condition gives stronger suppress to these

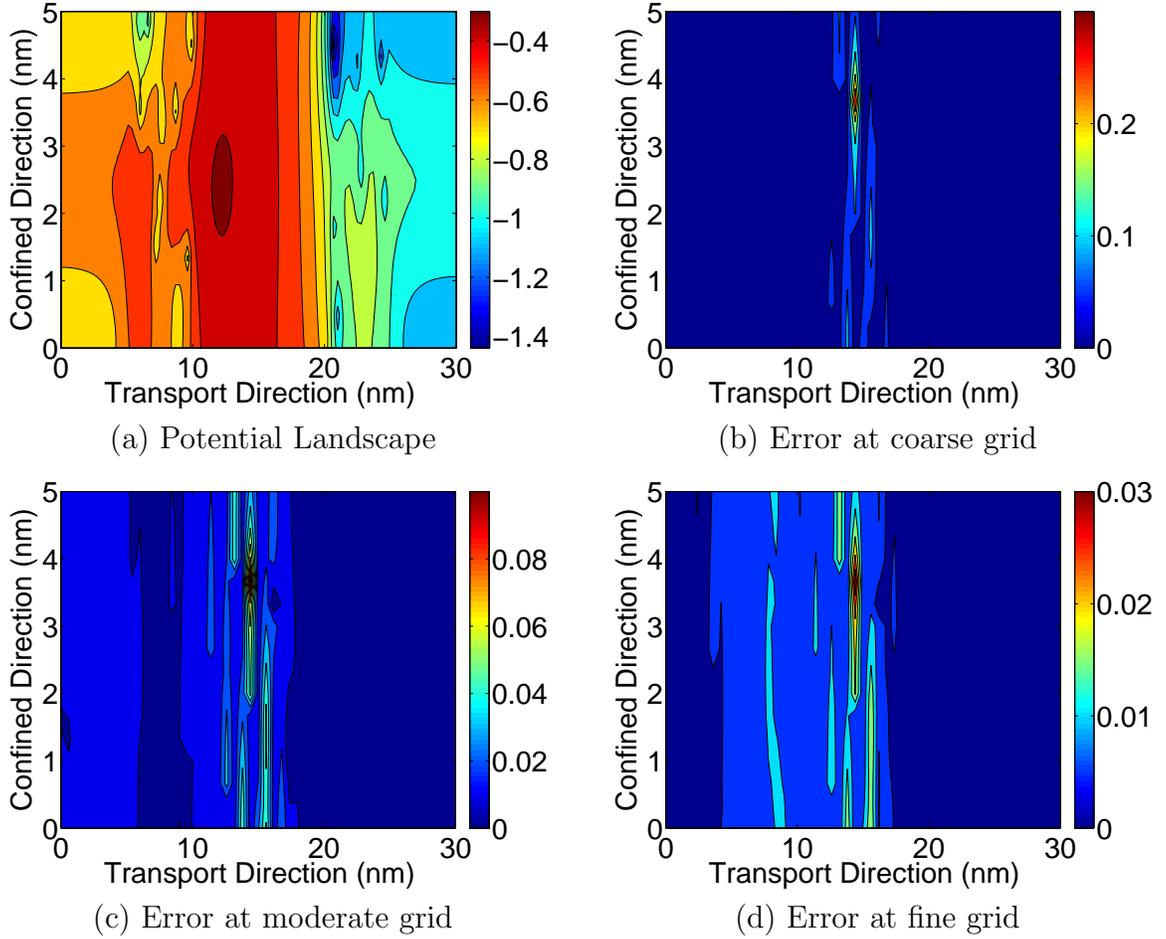


Figure 2.7: Contour plot of the electrostatic potential (a) and its absolute errors obtained with three sets of meshes(b, c, d) (10 dopants).

potential variations than the Neumann boundary condition does.

In our numerical experiments, it is also noticed that for certain number of randomly located dopants, some specific dopant configurations fail the simple self consistent outer iteration loop. The reason of the convergence failure, although not completely clear, might be the relative positions of the discrete dopants. Locally crowded individual dopants may lead to local significant variation of the potential landscape and undermine the convergence efficiency. Although the map (2.75) issues a fixed point, it promises no contraction property. Therefore the usually used outer

iteration

$$\zeta^{n+1} = \mathcal{F}_{1/2}^{-1} \circ \mathcal{G} \circ \mathcal{L}(\zeta^n) \quad (2.78)$$

may not converge. It is difficult to verify the contraction or construct contract mapping based on these operators because of the complex nonlinearity of the NEGF calculation \mathcal{G} . Figure 2.5(a) reveals these situations: The upper panel is one of the position distribution of 10 dopants that one can easily reach the steady state. The convergence behavior of the position configuration in upper panel of Fig. 2.5(a) can be found in Fig. 2.4(a). However, the distribution in the lower panel of Fig. 2.5(a) may lead to convergence failures. It can be seen that in the lower panel, dopants are very crowdedly distributed near the left top corner in the source. To deal with this numerical difficulty, we convert Eq. (2.78) into the steady-state problem of an ordinary differential equation (ODE) [7]

$$\frac{\partial \zeta}{\partial t} = \mathcal{F}_{1/2}^{-1} \circ \mathcal{G} \circ \mathcal{L}(\zeta) - \zeta. \quad (2.79)$$

Therefore many ODE related techniques such as the Runge-Kutta method can be used to improve the convergence properties. One simple treatment is the discretization of Eq. (2.79) as

$$\frac{\zeta^{n+1} - \zeta^n}{\alpha} = \mathcal{F}_{1/2}^{-1} \circ \mathcal{G} \circ \mathcal{L}(\zeta^n) - \zeta^n, \quad (2.80)$$

which leads to a self-consistent iteration with a relaxation factor α

$$\begin{aligned} \zeta^* &= \mathcal{F}_{1/2}^{-1} \circ \mathcal{G} \circ \mathcal{L}(\zeta^n) \\ \zeta^{n+1} &= \alpha \zeta^* + (1 - \alpha) \zeta^n. \end{aligned} \quad (2.81)$$

The traditionally used outer loop iteration actually is the special case of Eq. (2.81) with $\alpha = 1$. By carefully choosing the relax factor α , one can reach the steady

Table 2.1: Computational L_∞ error of the model with MIB scheme

5 dopants	error	order	10 dopants	error	order	40 dopants	error	order
Coarse	0.350	-	Coarse	0.301	-	Coarse	0.218	-
Moderate	0.222	0.7	Moderate	0.123	1.3	Moderate	0.133	0.7
Fine	0.089	1.3	Fine	0.038	1.7	Fine	0.047	1.5

state (fix point) by the self-consistent iteration for arbitrarily distributed individual dopants.

Figure 2.5(b) compares the convergence of the self-consistent iterations with different relaxation factors corresponding to the situation in the upper panel of Fig. 2.5(a). It indicates that $\alpha = 1$ does not work for the convergence loop, while $\alpha = 0.3$ leads to the convergence of the electron current within 0.2% relative error in around 50 steps.

One can easily come to the conclusion that although smaller relaxation factors promise the convergence, they result in more iteration steps. The exact reason of the position-dependent convergence and the choice of the optimal relaxation factor need to be further analyzed in the future.

The numerical solutions of the system are supposed to converge to the real solutions as meshing grids get smaller and smaller. The standard finite difference method is of second order convergence for the inner iteration. However, the strong outer non-linearity may ruin this theoretical rate. Moreover, the discontinuity of the dielectric constant and δ source function singularities in the Poisson equation further reduce the regularity of the solution. The standard FD method will not maintain its second order convergence. It may even diverge. The MIB scheme is designed not only for facilitating the Dirichlet-to-Neumann mapping but also for maintaining the high order convergence in interface problems. For the following convergence analysis, the numerical result obtained with a fine resolution of $h_x = 0.075\text{nm}$, $h_z = 0.025\text{nm}$ is considered as a reference solution, where h_x and h_z are for the grid step along the transport and confined direction, respectively. Three sets of mesh sizes of $(h_x, h_z) = (0.15\text{nm}, 0.05\text{nm})$,

(0.3nm, 0.1nm), and (0.6nm,0.2nm) are denoted as fine, moderate and coarse, respectively.

Figure 2.6 gives the errors under different grids resolutions in the central line of the silicon layer. Numerical results suggest the convergence of the present MIB method.

Furthermore, we examine the 2D errors of the simulation. Figure 2.7 presents the 2D errors for the situation when the dopant number is 10 in both the source and the drain.

Combined with Figs. 2.6 and 2.7, one can conclude that the major convergent errors occur at the junctions of discrete and continuous doping regions. Errors decrease as the grid is refined.

Finally, the L_∞ norm error (E_h) is considered and the convergence order is defined as $\log_2(E_h/E_{h/2})$ for three sets of meshes. The numerical convergence orders are calculated for several cases with different dopants numbers and listed in Table 2.1. The targeted second order of the MIB method is not achieved, partially because the strong nonlinearity of the coupled equation system.

2.3 Numerical simulation of the nano-electronic transistors

2.3.1 Device configurations

In the following sections, the proposed model and numerical implementation are used to examine the effects of random individual dopants and material interface. The impact of individual dopant random fluctuation was first recognized by Hoeneisen and Mead in 1970s and has been studied for many years via semiclassical or quantum mechanical means. Among the quantum models, Martinez et al [108] explored the impact of random dopant aggregation in the source and drain. Jiang et al [84] stud-

ied the gate threshold voltage lowering and fluctuation induced by random dopants in the channel. These studies are based on the smooth function approximation of individual dopants. The results from this treatment depend on empirical parameters and discretization mesh sizes. By using the Dirichlet-to-Neumann mapping, it can be found in the numerical simulation that the Dirac function model proposed in this work treats the individual dopants exactly, is of parameter free and does not require fine grids.

Multi-gate MOSFETs will play a dominant role in quantum devices because their ability to suppress channel tunneling and gate leaking effects. A variety of related investigations are carried out for such device models. We consider multi-gate MOSFETs in this work. Figure 2.1(a) gives an illustration of the 2D double-gate MOSFET and Fig. 2.1(b) is the corresponding computational domain. All relevant components of the device are presented in the graph. The x -direction is taken as the transport direction, the z -direction as the confined direction and all physical profiles are assumed invariant along y -direction. In the gray region, the source and drain are heavily doped while the channel is assumed near ballistic, in which case no doping is imposed. It is noticed from Fig. 2.1 (b) that the dielectric constants are different in the silicon and silicon dioxide layers, which are separated by the interfaces. For the Poisson equation on the computational domain, Dirichlet boundary conditions are taken at the double-gate region (EF and GH). For other boundaries (AE, FB, BC, CH, DG and AD), homogeneous Neumann boundary conditions are employed. The parameters of the device are the following: The total length (AD) of the device is 30 nm, with 10nm for each doping area and channel. The thickness of the silicon layer is 3nm, while the upper and lower silicon dioxide layers are 1nm each.

Figure 2.2(a) is the 3D sketch of the silicon nanowire transistor (SNWT), which is a MOSFET with all-around gates. The x -direction is the transport direction, while the other two are confined directions, where insulator layers exist and gate voltages

are applied. In our simulation, for simplicity, the cross-section of the SNWT is taken as a square, so it is also called a four-gate MOSFET. Figure 2.2(b) gives the cross-section of the SNWT in the y -direction ($y = 0$). It is similar to the structure of the planar double-MOSFET. Figure 2.2(c) presents the cross-section of the SNWT in x -direction ($x = 0$). The total length of the device is 30 nm, with 10nm for each doping area and the channel. The thickness of the silicon layer is 3nm. The thickness of the upper and lower silicon dioxide layers is 1nm each. Treatments that are similar to those for the double-gate MOSFET are taken for the boundary condition. If the slice has gate all around it, the Dirichlet boundary condition is imposed, otherwise homogeneous Neumann boundary condition is imposed.

The parameters of the device are the following: The source and drain voltage bias is $V_{DS} = 0.4V$, the double gate voltage is $V_G = 0.4V$. Three subbands are accounted for electron density calculations. In the double-gate MOSFET simulation, the electron effective mass is taken as $m_x = 0.50m_0$ in the transport direction and as $m_z = 0.20m_0$ in the confined direction. The dielectric constants for silicon layer and silicon dioxide layer are $\epsilon_{Si} = 11.7\epsilon_0$ and $\epsilon_{SiO_2} = 3.9\epsilon_0$, respectively, where $\epsilon_0 = 8.85 \times 10^{-12}Fm^{-1}$. The reference continuous n-doping is taken as $N_D(\mathbf{r}) = 2 \times 10^{20}/cm^3$ in the source and drain. For individual dopants in channel tests, the p-doping is take as $N_A(\mathbf{r}) = -1 \times 10^{20}/cm^3$. Room temperature of $T = 300K$ is assumed.

The strategies of placing individual dopants are as following: The discrete dopants are located for the 5nm long region right before the channel in the source contact and 5nm long region right after the channel in the drain contact. These regions are called discrete regions. The rest areas of the source/drain are self-averaging areas with continuum doping because a sufficient number of doping is necessary [108]. The number of individual dopants M_D and each dopant charge quantity c_j are chosen to

match the self-averaging doping concentration in the sense of an integration

$$\int_{\Omega_D} \sum_k c_k \delta(\mathbf{r} - \mathbf{r}_k) d\mathbf{r} = \int_{\Omega_D} N_D(\mathbf{r}) d\mathbf{r}, \quad (2.82)$$

where $N_D(\mathbf{r})$ is the continuous doping. Individual dopants are randomly and evenly distributed in the discrete region, i.e. the x - and y -coordinates of the dopants are independently generated by a uniform quasi-random number generator.

2.3.2 Simulation results for double-gate MOSFETs

In this section, we investigate the impact of the proposed interface model and the random individual doping model using the double-gate MOSFET. Before proceeding to the numerical results of the discrete dopant model, we show in Fig. 2.8(a) the landscape of the device electrostatic potential calculated from the standard finite difference method without the interface technique. The difference of the electrostatic potentials between obtained from the standard finite difference method and from the MIB method is plotted in Fig. 2.8(b). It is seen that two methods have about 3-4% of differences. The largest difference occurs around the interface region near insulator layers. We believe that the MIB scheme gives a more accurate calculation of potential landscape. Although the advantage of the MIB scheme over the standard FD scheme is not that tremendous in this case because of the relatively simple interface geometry, the MIB technique also has the ability of handling more complex interfaces which may occur in the future MOSFET design. Additionally, the DNM treatment of the discrete dopants can not be realized without the MIB scheme.

We next study the physical profiles of fluctuations under different amount of dopants and positions. The individual dopants in source or drain are anticipated to produce discrete aggregation effects: The on-state current fluctuates due to individual dopants and their relative positions. The fluctuations are not directly caused

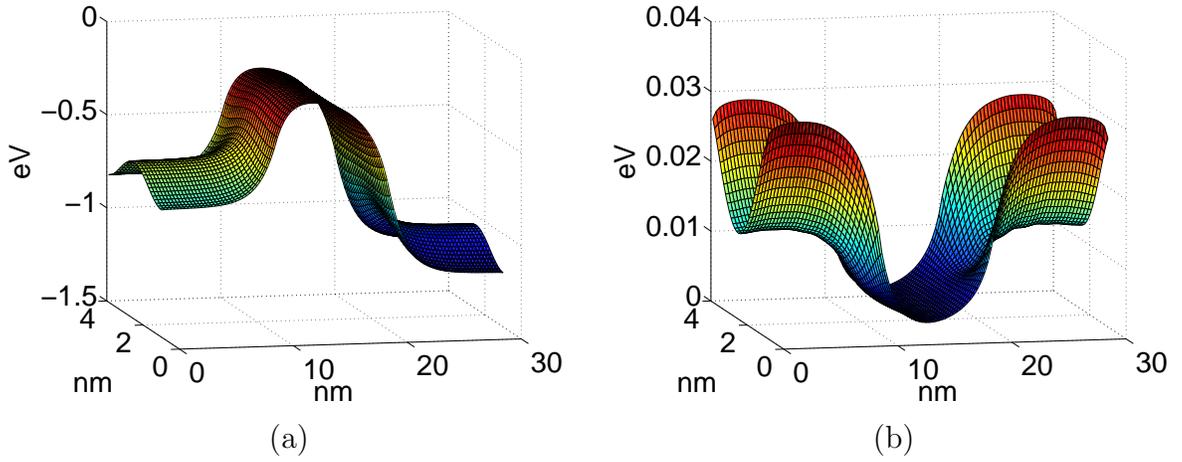


Figure 2.8: Electrostatic potential energy and difference of potentials for the double-gate MOSFET. (a) Potential landscape obtained with the standard finite difference method; (b) Difference of electrostatic potentials between computed with the standard finite difference method and with the MIB method.

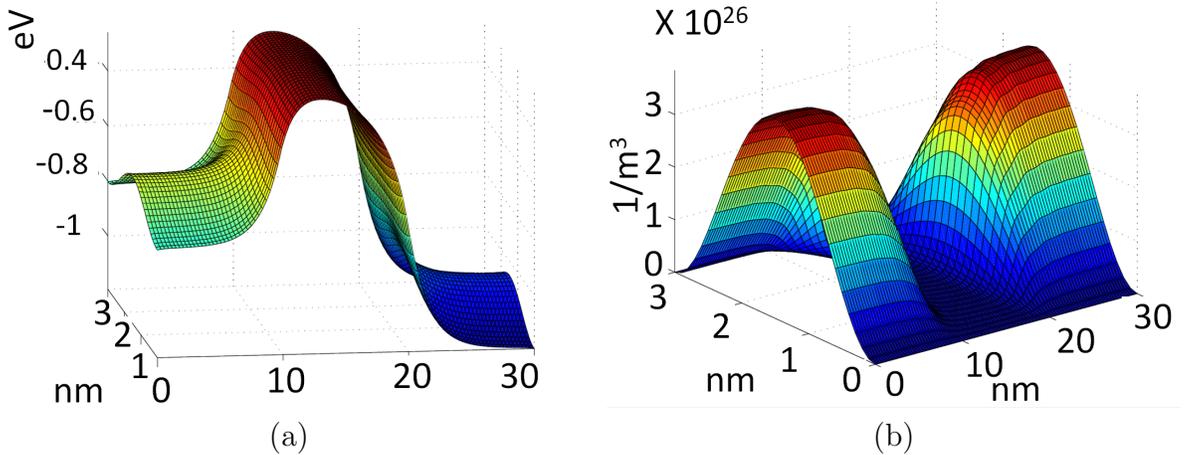


Figure 2.9: Profiles of electrostatic potential and electron density obtained with the continuous doping approximation. (a) Potential function; (b) Electron density.

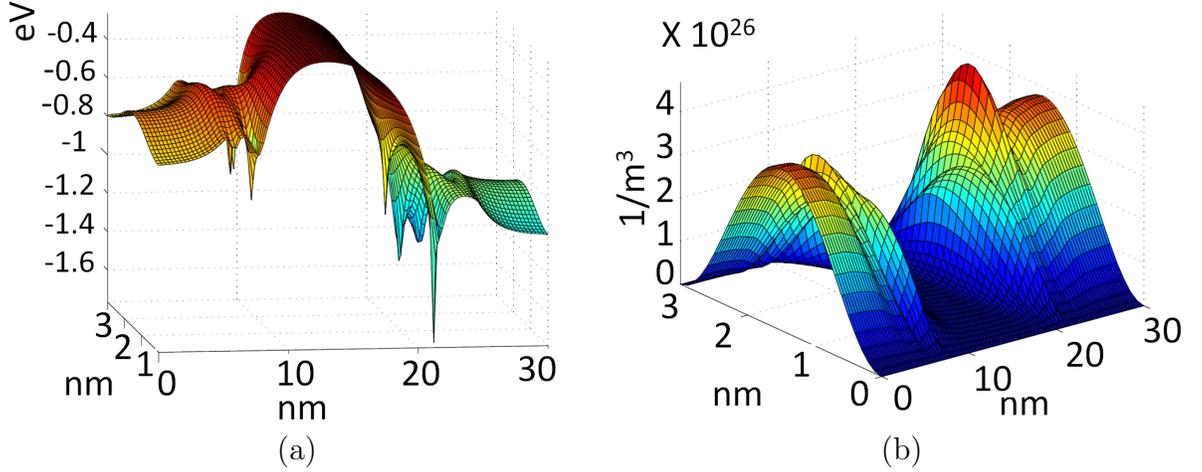


Figure 2.10: Profiles of electrostatic potential energy and electron density obtained with 5 individual dopants ($N = 5$). (a) Potential function; (b) Electron density.

by the spatially resolved individual dopant atoms but mesoscopically by the self-consistent electrostatic potential. Physical parameters are taken as stated in the previous section. To validate the individual dopant model, we employ the constraint of Eq. (2.82). This constraint may appear unphysical. However, it is useful in verifying the validity of the present model. It is expected that simulation profiles of discrete dopants will converge to those of the continuous doping when the number of dopants is large. The reason is that as the number of dopants increases, the amplitude of each dopant becomes smaller and the doping becomes evenly distributed.

For a comparison, Fig. 2.9 illustrates the steady-state potential and electron density obtained using the continuum doping approximation. It is predictable that the presence of the individual dopants will drag down the potential landscape and therefore trap the electron where a discrete dopant presents. This fact is revealed in our simulations: if there were no individual dopants in the region 5-10nm and 20-25nm along the x -direction, the potential landscape and electron density there would be similar to those in the channel (10-20nm in the transport direction). Due to the presence of discrete dopants, the potential landscape is dragged down and becomes

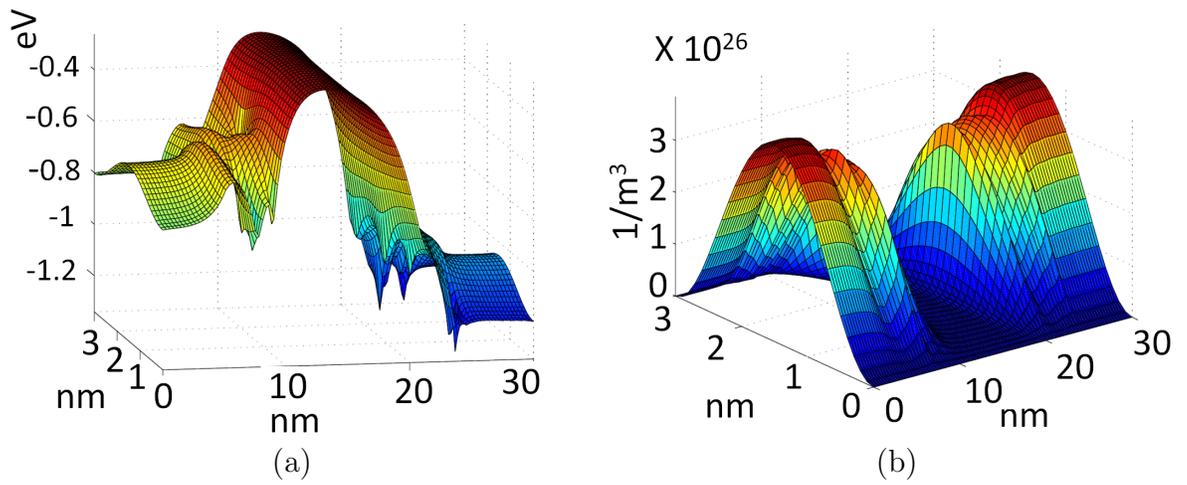


Figure 2.11: Profiles of electrostatic potential energy and electron density obtained with 10 individual dopants ($N = 10$). (a) Potential function; (b) Electron density.

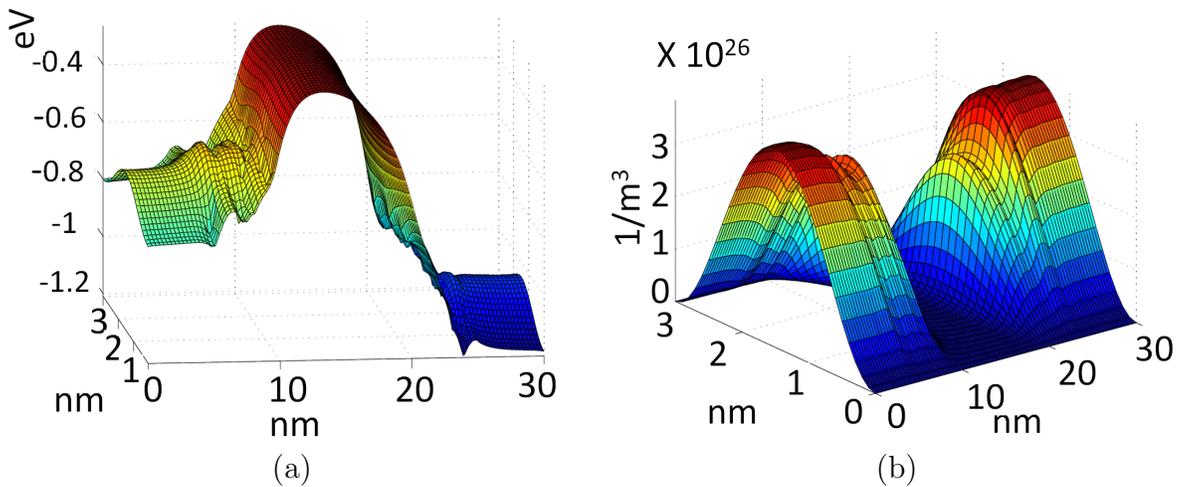


Figure 2.12: Profiles of electrostatic potential energy and electron density obtained with 40 individual dopants ($N = 40$). (a) Potential function; (b) Electron density.

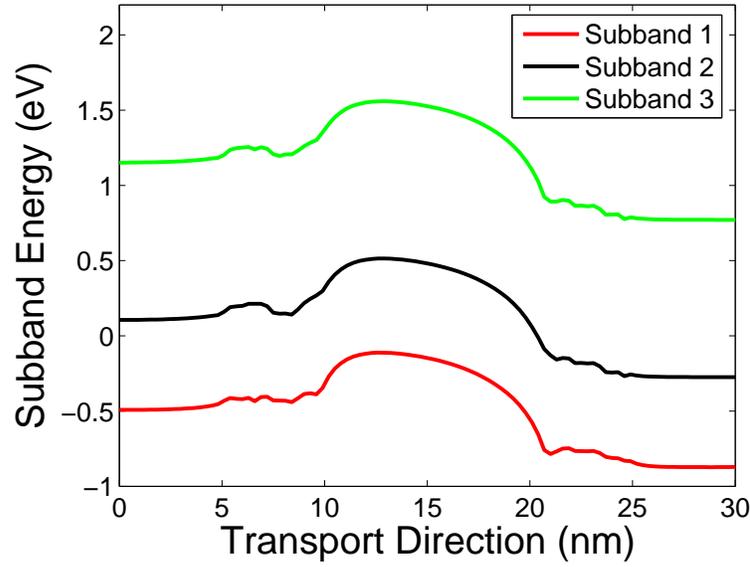


Figure 2.13: Subband energies of the double-gate MOSFET obtained with individual dopant model.

flat, as that presented in the continuous doping region (0-5nm or 25-30nm in the transport direction). As a consequence, the electron probability distribution (density) is increased near the discrete dopant site, comparing to the distribution profile of the continuous regions. Figures 2.10, 2.11 and 2.12 depict the potential landscapes and electron densities obtained with 5, 10 and 40 discrete dopants, respectively. Furthermore, it is observed that the more individual dopants presented in the field, the smoother the potential is, and the closer to the continuum doping model the present discrete doping will be. This is consistent with our expectation.

Detailed physical profiles of the device can be demonstrated by the subband energy. Figure 2.13 provides one of the subband energy profiles of the double-gate MOSFET with discrete dopants in the source and drain. The electron density is calculated based on the first three subband energies. Due to the Fermi-Dirac distribution, the higher the subband is, the less it contributes to the total amount of the electron density. From the subband energies profiles one can also find the flattening out

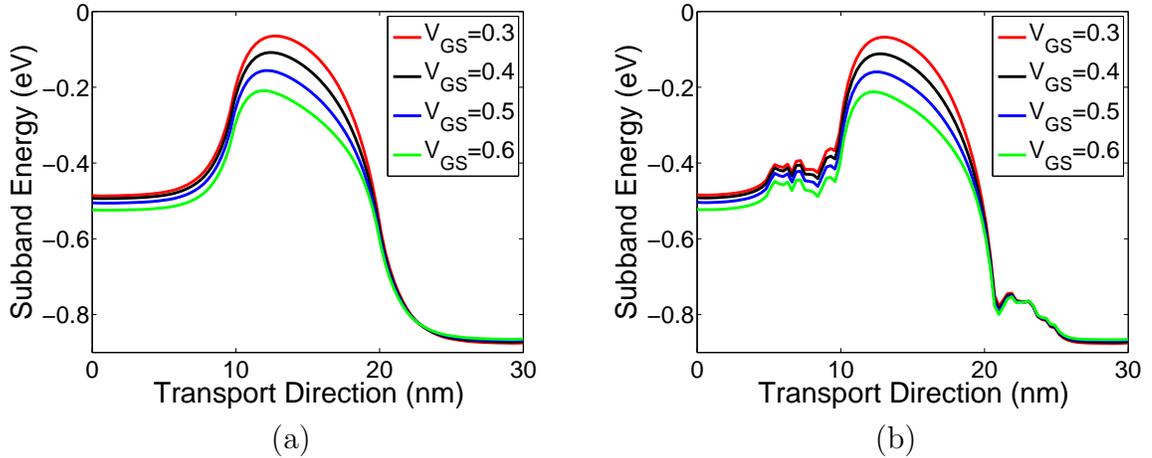


Figure 2.14: Subband energy profiles under different gate voltage biases. (a) Continuum dopant; (b) Discrete dopants (10).

effect of the individual dopants to the potential in the discrete region. Different from the continuous doping, the “flatting” is oscillatory.

Figure 2.14 illustrates the first subband energy profiles of the double-gate MOSFET at different gate voltages. The increase in the gate voltage will push down the potential barrier in the device channel and thus lead to larger charge current. In this study the source/drain bias is fixed under 0.4V.

Figure 2.15 shows the first subband energy of the double-gate MOSFET under different source/drain voltage biases. The differences in the subband profiles at two ends reflect the voltage biases. The width of potential barrier is reduced when the voltage bias is increased. The increase of the voltage bias leads to the increase of charge current in the double-gate MOSFET. The gate voltage is fixed at 0.4V in this study.

Macroscopically, the individual dopants introduce the fluctuation in on-state currents. Figures 2.16(a), (b) and (c) demonstrate the current fluctuations with 5, 10 and 40 dopants, respectively. It can be observed that the fluctuation is about 10% when there are 5 dopants in different positions. The fluctuation decreases gradually

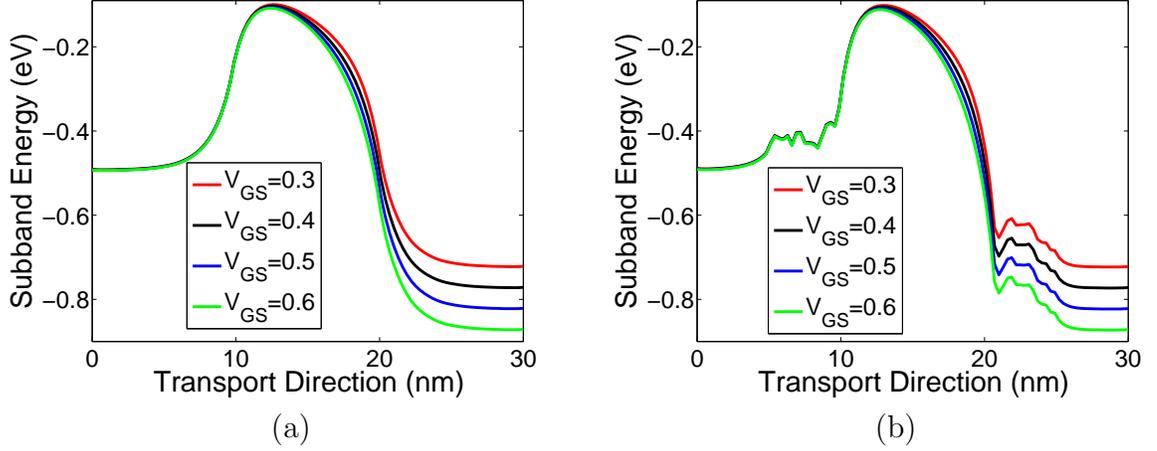


Figure 2.15: Subband energy profiles under different source/drain voltage biases. (a) Continuum Dopant; (b) Discrete Dopants (10).

when there are more dopants. From Fig. 2.16(d), one can draw the conclusion that the fluctuation is small when there is a large number of discrete dopants with small charge amplitude, as expected.

Another interesting aspect of the individual dopants is the lowering effect of the device voltage threshold. The voltage threshold of a MOSFET is defined as the gate voltage applied in order to generate the inversion layer between the insulator layer and the silicon layer, then electron current follows. The presence of the atomistic, or individual dopants has an effect of lowering the voltage threshold. In order to check the ability of our model to detect this effect, we introduce the p-type doping to the equivalent concentration of $-2 \times 10^{19}/\text{cm}^3$ in the channel and compare the I-V curves obtained from the continuous model and individual dopants model.

To obtain the I-V curve presented in Fig. 2.17, we set the source/drain voltage biases as 0.2V (left) and 0.4V (right), and gradually increase the gate voltage from 0 to 0.4V. It is clear from Fig. 2.17 that, to attain the same amount current, the present individual dopants model requires less amount of gate voltage than the continuous model does. State differently, the same gate voltage gives a higher charge current in

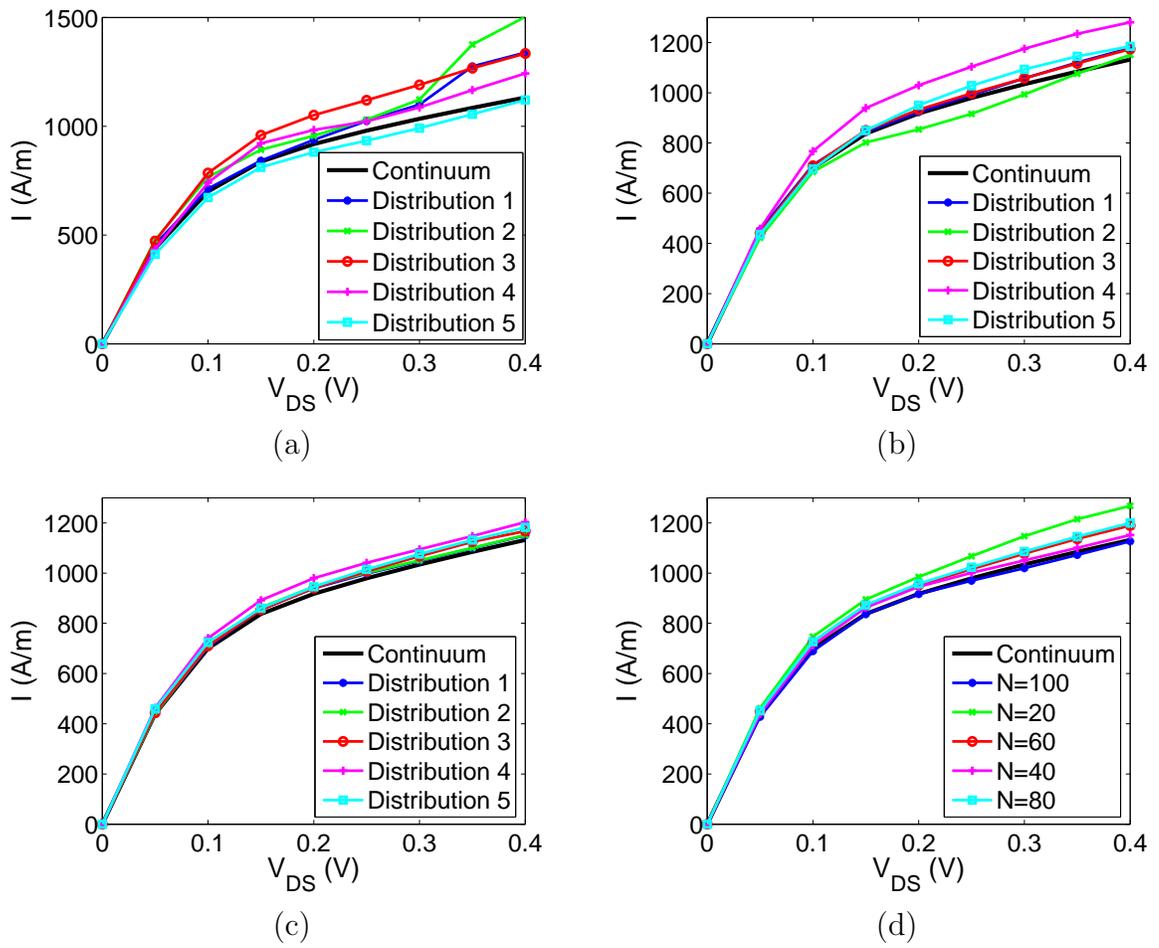


Figure 2.16: The dependence of I-V profiles on the number of individual dopants. (a) Fluctuation of 5 dopants; (b) Fluctuation of 10 dopants; (c) Fluctuation of 40 dopants; (d) Fluctuation of different number of dopants.

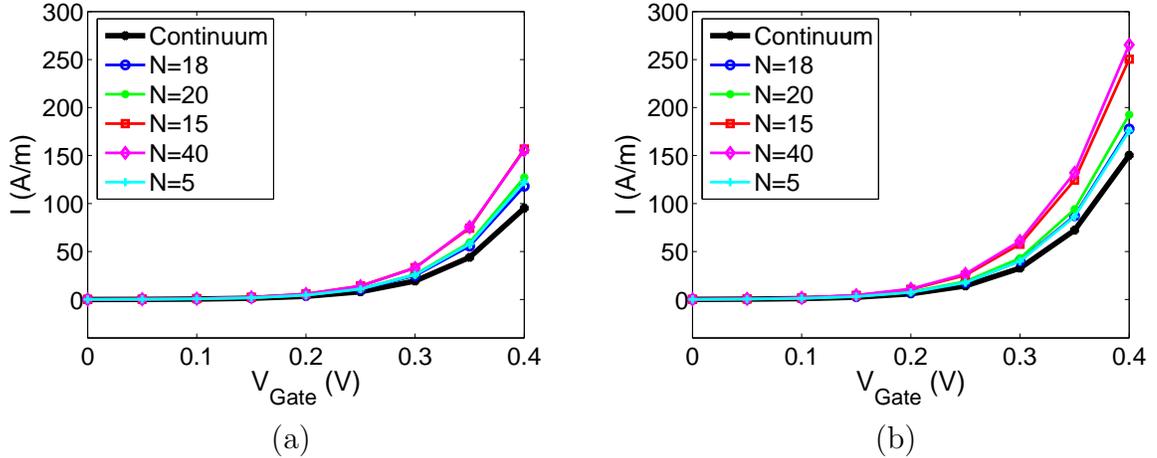


Figure 2.17: The comparison of I-V profiles under different doping models revealing the individual dopant induced voltage threshold lowering effect. (a) $V_{S/D} = 0.2V$; (b) $V_{S/D} = 0.4V$.

the present individual model than that in the continuous model. This fact reveals the individual dopant induced voltage threshold lowering effect.

It is noted that without using the constraint of Eq. (2.82), the impact of individual dopants and impurity can be very significant for small nano devices. The proposed Dirac delta function model can be used for the further investigation of this aspect.

2.3.3 Simulation results for four-gate MOSFETs

In order to achieve good device performance, more effective gate control strategy is required for nano-scale MOSFETs. In this section, the four-gate MOSFET, one kind of silicon nanowire transistors (SNWTs) which resemble multi-gate or gate-all-around devices, is explored. The transport profile of an electron in the SNWT is similar to that of the double-gate MOSFET and is simulated via the NEGF as well. While being different from the planar double-gate MOSFET, the charge and potential profiles in the transverse direction are no longer invariant. 3D simulations of the SNWT with discrete dopants are implemented in this section by using the technique of the

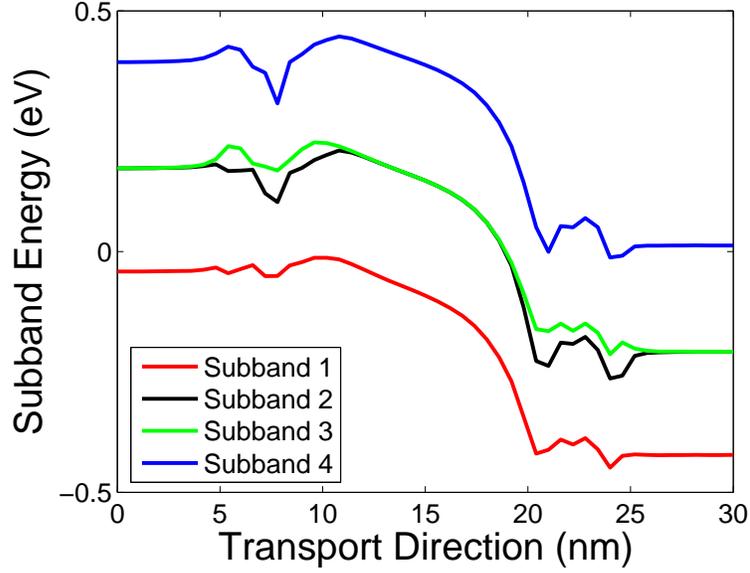


Figure 2.18: Subband energies for the four-gate MOSFET. The dopants break the symmetry and energy degeneracy in the second and third subbands.

subband decomposition. The MIB method and the Dirichlet-to-Neumann mapping in 3D versions described in the previous sections are employed. The configuration of the SNWT is given in Fig. 2.2 and the detailed parameters are almost the same as those in the double-gate MOSFET, except that the y -direction width is 5nm, with 3nm for the silicon layer and 1nm for each SiO_2 layer. The effective mass in the y -direction is taken as $m_z = 0.318m_0$.

The main difference of the present simulation and the simulation of the double-gate MOSFET is that one needs to solve the Schrödinger eigenvalue problem in 2D for each slice of the device. Additionally, the Green's formula for δ - functions takes a different form, as discussed in an earlier section.

Main results from simulating the four-gate MOSFET are quite similar to those of simulating the double-gate MOSFET. For example, the relation between the subband energy and S/D bias (or gate voltage) are generally similar to those of the double-gate MOSFET. Figure 2.18 shows the first 4 subband energies of the SNWT with 15

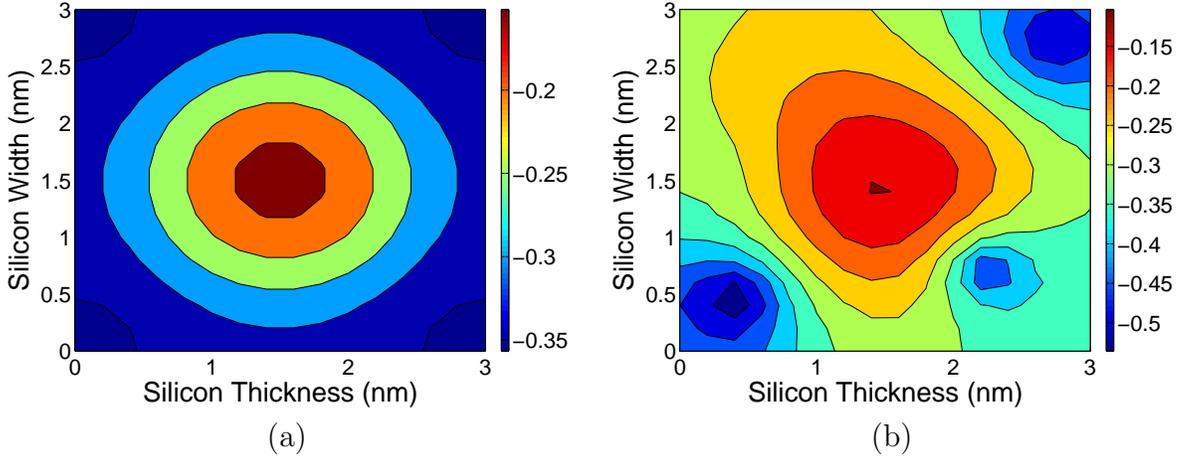


Figure 2.19: Cross-section of potential profiles for the four-gate MOSFET obtained with (a) Continuum doping; (b) Discrete dopants (10).

discrete dopants in source and drain under 0.4V S/D bias and gate voltages. Because a 2D eigenvalue problem is solved, the second and third subbands are almost identical (they are identical for continuous dopants). The fourth and fifth subbands are very similar too. Only the fourth subband is plotted.

Figure 2.19 presents the potential profile in the four-gate MOSFET at certain slice in the source region. It is found from Fig. 2.19(a) that the profile is symmetric due to the uniform doping and the symmetry of the device. While due to the presence of the randomly distributed individual dopants, the potential profile is no longer symmetric.

Finally, as in the last section, the current fluctuation of the four-gate MOSFET with individual dopants is studied and the I-V characteristics is presented in Fig. 2.20. For comparison to previous section, the current is re-scaled to current density according to the device geometry. Figures 2.20(a), (b) and (c) are I-V curves of the SNWT with 15, 40 and 80 dopants in different positions, respectively. While Fig. 2.20(d) shows the fluctuations with different numbers of dopants. For each number of dopants, the current values are averaged out with different positions. It can be concluded that as the number of individual dopants increases, the I-V curve is getting

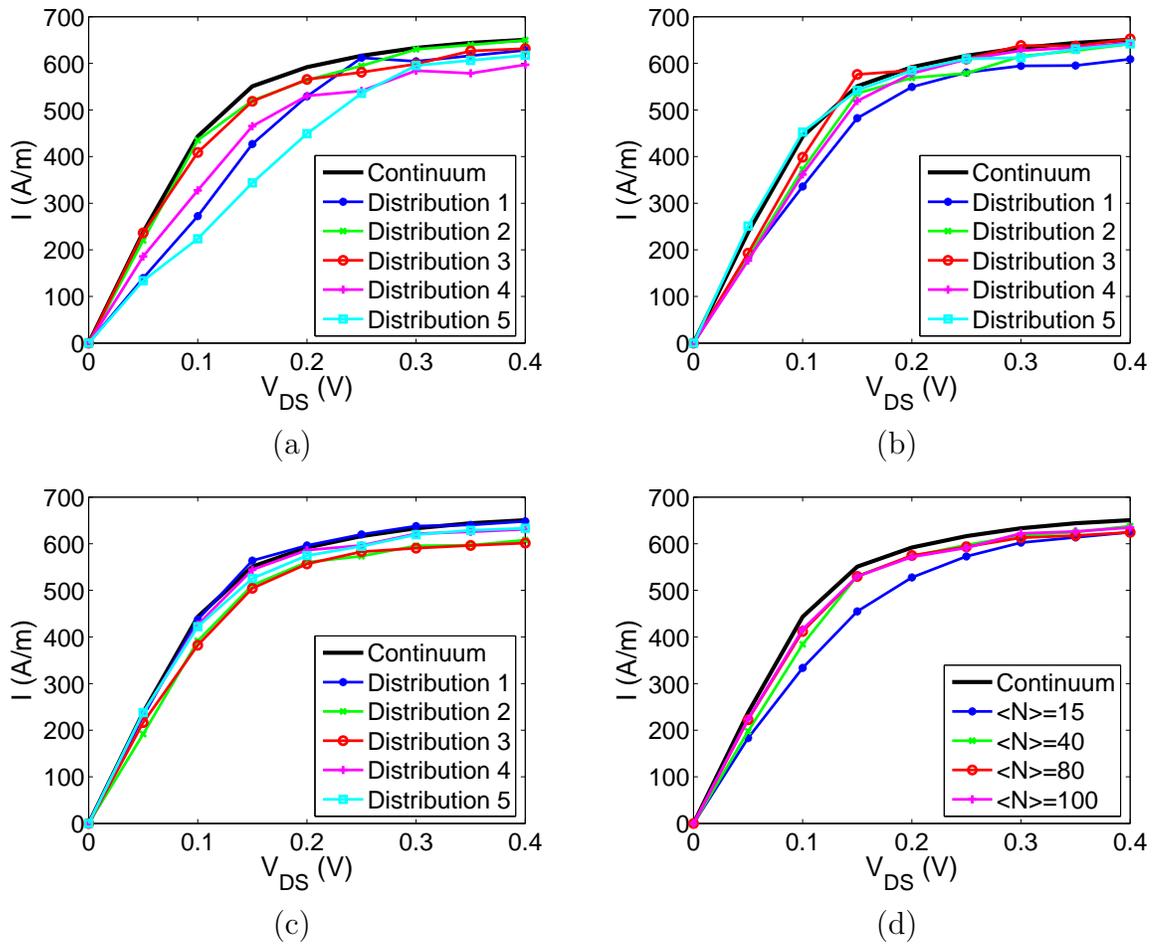


Figure 2.20: Current fluctuations of the four-gate MOSFET with (a) 15 dopants; (b) 40 dopants; (c) 80 dopants; (d) Different number of dopants.

closer to the profile of the continuous doping. This verifies the validity of our modeling for the same reason stated in the last section.

2.4 Conclusion remarks

This work presents mathematical modeling and computational algorithms to nano-scale electronic devices, or ultimate metal oxide semiconductor field effect transistors (MOSFETs). MOSFETs are the elementary building blocks of integrated circuits, such as microcircuits, microchips, or silicon chips, which are used in almost all electronic devices or equipments presently in use and have revolutionized the world of electronics in the past few decades. For years, MOSFETs are designed and functioning according to classical mechanical laws. The continuous miniaturization of nano transistors has led to a new era in design, manufacture, modeling and simulation of nano electronic devices where the quantum effects start to play a very important role. The benefits from the success of these new devices can be tremendous: On one hand, the new generation of ultimate MOSFETs will dramatically advance current semiconductor technology and still operate with the classical principle after suppressing severe quantum effects . On the other hand, functionally enhanced MOSFETs, or nano-quantum transistors, that utilize the fundamental properties of nature which do not have direct analogs in classical physics, will lead to new breakthroughs in device science and technology. However, enormous degrees of freedom of nano electronic devices make their first principle quantum mechanical modeling and simulation essentially intractable. Efficient mathematical modeling, approximation, and computation for this class of problems promise an important topic in applied and computational mathematics. This work addresses a few important issues in the modeling and simulation of nano electronic devices.

A simple and perhaps the most popular model is the effective mass description

of a single electron in band structures governed by the Kohn-Sham equation with an electrostatic potential solved from the Poisson equation. Even at this level of description, current modeling and simulation hardly take into account the significant impact of individual dopants, irregular geometric designs and material interfaces, due to the complexity and challenge of the problem. Additionally, inconsistency in many approximations undermines their performance. The present work addresses the above mentioned problems. First, we introduce a two-scale variational framework that, upon energy optimization, generates new self-consistently coupled Poisson-Kohn-Sham equations, or Poisson-Schrödinger equations. In this framework, the quantum dynamics of microscopic particles is determined by a nano-scale environment, which models the continuum electrostatics governed by the classical Maxwell's equation. As such, classical theory and quantum mechanics are put on an equal footing at nano-scale. Additionally, the magic and art of semiconductors are, by and large, due to the manipulation of layered structures. Such structures have great impact in electrostatic potentials and thus, electronic properties. In this work, we introduce material interfaces in coupled Poisson- Kohn-Sham equations to properly describe the interface effect to the electrostatic potential. Finally, random dopants and impurity have a dramatic impact to the electronic structures and electronic transport. We present a new individual dopant model by utilizing the Dirac delta function. Unlike previous individual dopant models [155, 9, 108, 84] that depend on tunable parameters, the proposed model is free of computational parameters.

This work also introduces two computational algorithms for the simulation of nano electronic devices. An efficient elliptic interface method, the matched interface and boundary (MIB) method [160, 165, 164, 159, 158], is employed for solving the Poisson equation with semiconductor interfaces. Although the MIB method is of arbitrarily high-order accuracy in principle, we only utilize the second order MIB scheme in the present work. However, due to the strong nonlinearity of the coupled

Poisson- Kohn-Sham equations, the numerical order was found to be about 1.5 in the present Gummel-like iteration. The other computational algorithm employed in the present work is the Dirichlet to Neumann mapping (DNM). This approach provides a rigorous treatment of the singular charges sources in the Poisson equation due to the individual dopant model proposed in the present work. The basic idea of the DNM is to separate the singular part of the solution from the regular part, and then solve the singular part analytically. The analytical treatment of the singular part gives rise to a Neumann interface condition or boundary condition for the regular part of the Poisson equation. By splitting the Poisson equation, we have taken the advantage that the amplitudes of the both discrete dopants and continuum doping are fixed, while those of the electronic density is changing during the Gummel-like iteration. The DNM technique works well in the present device simulation.

Two multi-gate MOSFET systems, a double-gate MOSFET and a four-gate MOSFET, are considered in the numerical simulations of the present work. Both problems are modeled in the three-dimensional (3D) setting. The subband decomposition approach is utilized to simplify the computation of the electronic structures. The non-equilibrium Green's functions (NEGF) formalism is employed for the description of electronic transport. The convergence and well-posedness of the present models are analyzed. A relaxation technique is developed to improve the convergence of the iteration scheme which is subject to possible numerical divergence induced by the charge aggregation of randomly distributed individual dopants. In our double-gate MOSFET simulations, the basic characteristics and the quantum effect of the I-V curves are similar to those in the literature. The impact of randomly distributed individual dopants to electronic structure and transport is studied. In particular, individual dopant induced voltage threshold lowering effect is clearly demonstrated. In our four-gate MOSFET simulations, individual dopants effectively break the symmetry of the device. Due to the 2D quantum confinement, the density of quantum states

that are relevant to the electronic transport become larger than that of the double-gate MOSFET. Effects of multigate voltage settings and irregular device geometry are under our investigation.

Chapter 3

Modeling and simulations of transmembrane proton channels

In this chapter we provide the theoretical formulation of our model of quantum dynamics in continuum.

3.1 Theory and model

3.1.1 General description of the model

An ion channel system is complex in terms of biological structure, dynamics and proton transport. Our goal is to model the dynamics and to predict the transport. To this end, we propose a multiscale, multiphysics and multidomain model. The computational domain Ω is divided into two subdomains, i.e., the solvent domain Ω_s consisting of the extracellular/intracellular solvent regions and the ion channel or channel pore region, and the biomolecular subdomain Ω_m including the membrane protein(s) as well as lipid bilayers. Therefore, we have $\Omega = \Omega_s \cup \Omega_m$. A detailed graph of these subdomains is given in Fig. 3.1. The interface Γ between solvent-membrane protein is defined by the molecular surface generated by the MSMS software package

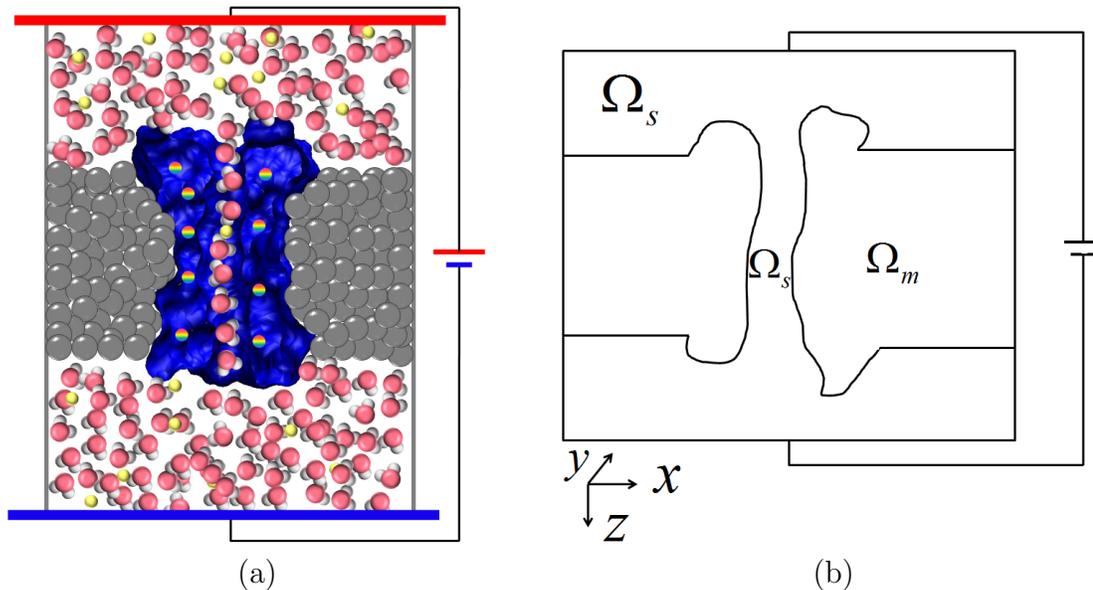


Figure 3.1: (a) Illustration of multiscale model of a proton channel; (b) Computational domains of the multiscale model with Ω_m being the channel molecule and membrane domain and Ω_s being the solvent domain. Here z -direction is regarded as the transport direction.

[131]. It is interesting to note that the physics in each subdomain is very different and there are multiphysics phenomena even in a single subdomain. For the biomolecular subdomain, the membrane protein and lipid bilayer structural data are either generated for molecular dynamics simulations, or downloaded from the Protein Data Bank (PDB) which are collected from X-ray crystallography or nuclear magnetic resonance (NMR) experiments. The force field parameters, such as atomic van der Waals radii and point charges, are obtained from the CHARMM force field [105]. This structural information is utilized in solving the Poisson-type equation for the electrostatic potential. The electrostatic potential distribution near the channel pore is crucial to the channel selectivity, gating, and ion conductance. The interactions between the channel protein and transmission channel ions are accounted in the present model.

In the solvent subdomain, there are three types of materials, ions of interest (i.e., protons), all other ion species and water molecules. In this system, the charge-charge

interactions contribute to the predominate potential energy landscape. Whereas, the strength of other interactions, such as ion-water dipolar interactions, water-water interactions and molecular van der Waals interactions, is much weaker than that of direct charge-charge interactions. This feature provides us a ground to take a multiscale approach to the multiphysics situation in ion channel dynamics. To reduce the number of degrees of freedom, we treat solvent (water) molecules as continuum background or bath. The formation of ion and water clusters and possible ion-water correlations are modeled partially as a dielectric constant effect and partially as a generalized-correlation potential effect. Except for the ions of interest, other ions usually have a small population in the channel pore of a selective channel. Whereas in the bath region, all ions are essentially in a quasi-equilibrium state and their densities are well described by the Boltzmann distribution except for at the solvent-membrane protein interface. Near the solvent-membrane protein interface, the density distribution of ions might be better described by the density functional theory of solution, or integral equations, in which the dispersion interaction between solvent and solute can be better accounted. This effect is modeled as generalized-correlation potential effect in the present work.

The physics in the channel pore region is of central interest and is sharply different from those of other regions. The ions of interest are selected as those which have significant population inside the channel region. There are many evidences which indicate the quantum mechanical behavior of proton transfer in biomolecular systems and proton channels [41, 20]. The first reason is the small mass of a proton which enhances the quantum tunneling effect in the proton transport. Additionally, a narrow channel morphology in many proton channels, such as the Influenza A M2 proton channel [28, 134] leads to severe quantum confinement, which consequently promotes quantum effects. Finally, hydrogen-bonded chain (proton nanowire) of water molecules assisted proton translocation is quantum mechanical in origin [121, 123, 122]. Although

theoretical models were proposed in the last decades [135, 144, 21], the detailed mechanism of proton dynamics and transport is not fully understood. For these reasons, we treat protons quantum mechanically via a scattering formalism which describes how a quantum mechanical proton scatters through electrostatic and generalized-correlation potential fields. The electrostatic potentials include interactions between protons represented by a self-consistent mean field approximation, the interactions between protons and fixed ions from membrane proteins and lipid bilayers, and the interactions between protons and other ion species. The generalized-correlation potential is due to the impacts of the continuum solvent, the van der Waals interaction between the solvent and biomolecules, the effect of ion-water clusters, dispersion effect, and possible break-down of hydrogen-bonded chain in a narrow channel, etc. We utilize a total energy functional framework [154, 27, 31] to incorporate quantum mechanical description and continuum description. Coupled Kohn-Sham equation for the proton dynamics and Poisson-Boltzmann equation for the electrostatic potential are derived from the variational principle. Solutions to these coupled equations give rise to proton structure dynamics, and transport in the ion-channel process, which describes how a quantum mechanical proton scatters through electrostatic and generalized-correlation potential fields.

3.1.2 Free energy components

This subsection describes various free energy components in our multiscale model of quantum dynamics in continuum. In order to give a clear description, Fig. 3.1(a) is reduced to a sketch in Fig. 3.1(b) in $x - z$ cross section, where the z direction represents the proton transport direction: the system is restricted to a rectangular cuboid with appropriate size and partitioned into two different computational domains. The

permittivity $\epsilon(\mathbf{r})$ has different values in two domains

$$\epsilon(\mathbf{r}) = \begin{cases} \epsilon_s(\mathbf{r}) & \forall \mathbf{r} \in \Omega_s \\ \epsilon_m(\mathbf{r}) & \forall \mathbf{r} \in \Omega_m \end{cases} . \quad (3.1)$$

Since both the membrane and channel protein are treated with same dielectric medium, the interface between them is erased and a constant dielectric constant is assumed on Ω_m . On the contrast, the solvent in the bath regions and in the channel pore have different biological characteristics. Therefore the position dependent dielectric constant is imposed on the solvent domain Ω_s . In fact, $\epsilon_s(\mathbf{r})$ in the channel region can differ much from that in the bulk region. The detailed discussion about the dielectric constants is given in Section 3.2.5.

Electrostatic free energy in the biomolecular region

The biomolecular region consists of membrane protein and lipid bilayer. Their structures determine the channel selectivity and gating efficiency. In the present treatment, we assume that structures of membrane protein and lipid bilayer are given and do not change during the ion transport process. This is certainly an approximation and will be easily removed in our future work by a combination of the present formulation with MD simulations [154]. Without structural cooperation, the biomolecules still significantly contribute to ion dynamics and transport by electrostatic interactions. The fixed charges in the channel protein and nearby lipid bilayers determine the fundamental characteristics of the channel and provide the primary environment for ions' permeation. Since the total number of fixed charges is not too large (i.e., up to thousands), with the assumption that the positions of them are essentially fixed, the explicit discrete description is actually affordable. In this sense, they serve as a

source term in the electrostatic potential calculation

$$\rho_f(\mathbf{r}) = \sum_{i=1}^{N_a} Q_i \delta(\mathbf{r} - \mathbf{r}_i) \quad (3.2)$$

where N_a is the total number of fixed charges, Q_i and \mathbf{r}_i are the point charge and position of the i th atom. Therefore, the electrostatic free energy in biomolecular domain is given by

$$G_{\text{Mol}}[\Phi, n] = \int \left[\frac{\epsilon_m(\mathbf{r})}{2} |\nabla\Phi|^2 - \Phi\rho_f \right] d\mathbf{r}, \quad (3.3)$$

where $\Phi(\mathbf{r})$ is the electrostatic potential and is defined on the whole domain $\Omega_s \cup \Omega_m$.

Electrostatic free energy in the solvent region

The ions in the solvent region also contribute to the electrostatic potential. Protons and other ion species are treated in different manners. Let us denote the proton number density in the solvent region as $n(\mathbf{r})$ and the charge density as $\rho_p = qn(\mathbf{r})$, q is the elementary charge or charge carried by a single proton. The charge density serves as a source term in the electrostatic free energy.

In the solvent region, particularly, in the extracellular and intracellular solvent regions, apart from ions of interest, there are many other ions. In the present model, all other ions are treated in a different manner from the ions of interest. Specifically, no detailed description is given to individual ions except for the ions of interest. However, other ions contribute considerably to the electrostatic property of the whole system. To account for their electrostatic effort, we describe other ions by using the Boltzmann distribution. The charge density of other ions is given by

$$\rho' = \sum_j^{N'_c} q_j n'_j(\mathbf{r}) = \sum_j^{N'_c} q_j n_j^0 e^{-q_j(\Phi(\mathbf{r}) - V_{\text{Ext}})/k_B T}, \quad (3.4)$$

where N'_c is the total number of other ion species, n_j^0 and q_j are the bulk constant density and charge of the j th ion species. Here $n'_j = n_j^0 e^{-q_j(\Phi(\mathbf{r}) - V_{\text{Ext}})/k_B T}$ is the number density of j th ion species, it can be noticed that the Boltzmann distribution of the other ion species with respect to the potential has been modified with the generalized chemical potential V_{Ext} , which represents the effects of the chemical potential of j th ion species and the external electric field [125, 127].

The corresponding electrostatic free energy in the solvent region is given by

$$G_{\text{Sol}}[\Phi] = \int \left[\frac{\epsilon_s(\mathbf{r})}{2} |\nabla \Phi(\mathbf{r})|^2 - \Phi(\mathbf{r}) \rho_p(\mathbf{r}) + k_B T \sum_j^{N'_c} n_j^0 \left(e^{-\frac{q_j(\Phi(\mathbf{r}) - V_{\text{Ext}})}{k_B T}} - 1 \right) \right] d\mathbf{r} \quad (3.5)$$

Note that the electrostatic free energy of other ions in Eq. (3.5) is similar in spirit to Sharp and Honig [138], Gilson et al [68], Chen et al [31] and Wei [154].

Proton free energies and interactions

The solvent region might admit a number of ion species, of which a full quantum model can be technically complicated and computationally time consuming. We therefore only treat the ions of interest, i.e., protons, quantum mechanically and assume a continuum description of other ion species. To simplify the problem further, we consider a generalized density functional theory for protons.

Kinetic energy. The proton density operator n_H is given by

$$n_H = e^{-(H - E_{\text{Ext}})/k_B T}. \quad (3.6)$$

where H is the Hamiltonian of the system and E_{Ext} is the external electrical field energy. We define the proton density $n(\mathbf{r})$ as

$$n(\mathbf{r}) = \langle \mathbf{r} | n_H | \mathbf{r} \rangle = \int |\Psi_E(\mathbf{r})|^2 e^{-(E-E_{\text{Ext}})/k_B T} dE, \quad (3.7)$$

where Ψ_E and E are the wavefunction and corresponding energy associated with H . The Boltzmann statistics is adopted in the present work. The kinetic energy is given by $\frac{\mathbf{p}^2}{2m(\mathbf{r})}$ where \mathbf{p} is the momentum and m is proton effective mass. In the coordinate representation, the kinetic energy of protons can be given as

$$\int \int \frac{\hbar^2 e^{-(H-E_{\text{Ext}})/k_B T}}{2m(\mathbf{r})} |\nabla \Psi_E(\mathbf{r})|^2 dE d\mathbf{r}, \quad (3.8)$$

where the Boltzmann factor weights different energy contributions.

Electrostatic potential. Protons have a number of electrostatic interactions. First, protons interact repulsively among themselves

$$U_{\text{Ion-Ion}}(\mathbf{r}) = \frac{1}{2} \int \frac{q^2 n(\mathbf{r}) n(\mathbf{r}')}{\epsilon(\mathbf{r}) |\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \quad (3.9)$$

These interactions lead to a term that is nonlinear in density n and the resulting equations are to be solved iteratively.

Additionally, interactions between protons in the solvent and fixed charges in biomolecules are described as

$$U_{\text{Ion-Fix}}(\mathbf{r}) = \sum_{i=1}^{N_a} \frac{qn(\mathbf{r})Q_i}{\epsilon(\mathbf{r})|\mathbf{r} - \mathbf{r}_i|}. \quad (3.10)$$

This contribution can be handled by the so called Dirichlet to Neumann mapping approach [27].

Finally, interactions between protons and other ion species are of the form

$$U_{\text{Ion-Other}}(\mathbf{r}) = \sum_{j=1}^{N'_c} \int \frac{qq_j n(\mathbf{r}) n'_j(\mathbf{r}')}{\epsilon(\mathbf{r}) |\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \quad (3.11)$$

where the other ionic densities are determined from the continuum Boltzmann distribution in the solvent region with a given profile of electrostatic potential as shown in Eq. (3.4). Therefore, the electrostatic potential energy functional of protons is

$$\int [U_{\text{Ion-Ion}}(\mathbf{r}) + U_{\text{Ion-Fix}}(\mathbf{r}) + U_{\text{Ion-Other}}(\mathbf{r})] d\mathbf{r}.$$

Generalized-correlation potential. The electrostatic potential plays a dominant role in the ion transport process. However, generalized-correlation effects are also important to ion conductance efficiency. Sometimes, generalized-correlation effects can even determine the channel selectivity. Generalized-correlation effects physically originate from van der Waals interactions, dispersion interactions, ion-water dipolar interactions, ion-water cluster formation/dissociation, temperature and entropy effects, etc. For example, one of generalized-correlation effects is an energy barrier to the ion transport due to the change in the solvation environment from the bulk water to a relatively dry channel pore. However, due to the lack of a comprehensive understanding of the ion behavior in channel region, the modeling of generalized-correlations is less quantitative, compared to the electrostatic modeling. In the Brownian dynamics model and the PNP theory, these generalized-correlation effects are encapsulated in the relaxation time and diffusion coefficients, respectively, which are obtained from experimental data and tuned in a reasonable biological range to predict new results. In the present work, we consider a reduced model for generalized-correlation potential energy. We assume that generalized-correlation potential is also a functional of the local ion density $n(\mathbf{r})$ and the density gradient ∇n , i.e., $U_{\text{GC}}[n, \nabla n]$. It includes

two contributions: One is the interaction among the target ions themselves, which represents those short range interactions and possible collisions; the other is the interaction between the ion and the surrounding water molecules, which may include many-body ion-water collisions and dehydration effects. In an analogous structure of energy (3.9), the former should be a quadratic form while the latter is a linear form like Eq. (3.10) of the ion density $n(\mathbf{r})$. Based on these considerations, we assume that the generalized-correlation potential energy functional has the following form

$$\int \frac{\partial U_{\text{GC}}[n, \nabla n]}{\partial n} d\mathbf{r} = \int V_{\text{GC}}[n] d\mathbf{r} = \int (\alpha n(\mathbf{r}) k_B T + V_{\text{Ion-sur}}[n]) d\mathbf{r}, \quad (3.12)$$

where the ∇n dependence has been omitted for simplicity. The first term of Eq. (3.12) is linear in the ion density. Intuitively, if more ions exist in the system, the possibility of the ion-ion generalized-correlation interaction is higher. The energy resulting from the ion-surrounding interaction is simply modeled as energy $V_{\text{Ion-sur}}[n]$, which can be considered as related to the density and relaxation time of ions. The range of $V_{\text{Ion-sur}}[n]$ value is discussed in Section 3.2.5. Here α is a relative weighting parameter for balancing the contribution of two components in the overall $U_{\text{GC}}[n]$.

External potentials Since the extracellular and intracellular surroundings can be infinitely large, it is impossible to include them in a detailed description. In the present work, we make appropriate truncation of the surrounding system. As such, the interaction of channel protons with extracellular and intracellular surroundings are described by external potentials U_{Exter} . In addition to the truncation effect, the external potentials also describe the experimental conditions such as the effect of given extracellular and intracellular bulk concentrations. We denote channel potential

energy functional as

$$\int U_{\text{Exter}}[n]d\mathbf{r} = \int V_{\text{Exter}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} = \int [V_{\text{Extra}}(\mathbf{r})n(\mathbf{r}) + V_{\text{Intra}}(\mathbf{r})n(\mathbf{r})] d\mathbf{r} \quad (3.13)$$

where $V_{\text{Extra}}n(\mathbf{r})$ and $V_{\text{Intra}}n(\mathbf{r})$ are extracellular and intracellular positions, respectively. Because much of extracellular and intracellular surrounding is not explicitly described, V_{Exter} must be non-hermitian. This aspect is discussed in Section 3.1.5.

Proton total energy functional. The total proton potential consists of electrostatic, generalized-correlation and external potentials

$$\begin{aligned} U(\mathbf{r}) &= U_{\text{Elec}}(\mathbf{r}) + U_{\text{GC}}(\mathbf{r}) + U_{\text{Exter}}(\mathbf{r}) \\ &= \frac{1}{2} \int \frac{q^2 n(\mathbf{r})n(\mathbf{r}')}{\epsilon(\mathbf{r})|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \sum_{i=1}^{N_a} \frac{qn(\mathbf{r})Q_i}{\epsilon(\mathbf{r})|\mathbf{r} - \mathbf{r}_i|} + \sum_{j=1}^{N'_e} \int \frac{qq_j n(\mathbf{r})n'_j(\mathbf{r}')}{\epsilon(\mathbf{r})|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \\ &+ U_{\text{GC}}[n(\mathbf{r})] + U_{\text{Exter}}[n(\mathbf{r})]. \end{aligned} \quad (3.14)$$

Thus, the total free energy functional of protons includes kinetic and potential contributions

$$G_{\text{Ion}}[\Phi, n] = \int \left[\int \frac{\hbar^2 e^{-(E-E_{\text{Ext}})/k_B T}}{2m(\mathbf{r})} |\nabla \Psi_E(\mathbf{r})|^2 dE + U(\mathbf{r}) \right] d\mathbf{r}, \quad (3.15)$$

where each kinetic energy term is weighted by the Boltzmann distribution, which is similar to the treatment in our recent work [27].

3.1.3 Total free energy functional of the system

To understand the behavior of protons and their interactions, we consider a total free energy functional that includes all significant kinetic and potential energies. Similar energy framework has been developed in our recent work for biomolecular systems

and nano-electronic devices [154, 27, 31]. The total free energy functional of the present system is given by the combination of the electrostatic energy of the system and the quantum mechanical energy of protons. However, it is important to avoid double counting when one constructs the total energy functional [154]. For the present system, it is interesting to note that had the charge sources $qn(\mathbf{r}') + \sum_{i=1}^{N_a} Q_i \delta(\mathbf{r} - \mathbf{r}') + \sum_{j=1}^{N'_c} q_j n'_j(\mathbf{r}')$ been independent of Φ , we would have

$$qn(\mathbf{r})\Phi(\mathbf{r}) = \frac{1}{2} \int \frac{q^2 n(\mathbf{r})n(\mathbf{r}')}{\epsilon(\mathbf{r})|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \sum_{i=1}^{N_a} \frac{qn(\mathbf{r})Q_i}{\epsilon(\mathbf{r})|\mathbf{r} - \mathbf{r}_i|} + \sum_{j=1}^{N'_c} \int \frac{qq_j n(\mathbf{r})n'_j(\mathbf{r}')}{\epsilon(\mathbf{r})|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (3.16)$$

in a homogeneous dielectric medium. Therefore, the charge source for the electrostatic potential also serves the electrostatic potential energy for protons. With this consideration, we propose the total free energy functional

$$\begin{aligned} G_{\text{Total}}[\Phi, n] &= \int \left\{ \left[\frac{\epsilon(\mathbf{r})}{2} |\nabla\Phi|^2 - \Phi(\rho_p + \rho_f) + k_B T \sum_j^{N'_c} n_j^0 \left(e^{-\frac{q_j(\Phi - V_{\text{Ext}})}{k_B T}} - 1 \right) \right] \right. \\ &\quad - \left[\int \frac{\hbar^2 e^{-(E - E_{\text{Ext}})/k_B T}}{2m(\mathbf{r})} |\nabla\Psi_E(\mathbf{r})|^2 dE + U_{\text{GC}}[n] + U_{\text{Exter}}[n] \right] \\ &\quad \left. - \lambda \left[\int e^{-(E - E_{\text{Ext}})/k_B T} |\Psi_E(\mathbf{r})|^2 dE - N_p \right] \right\} d\mathbf{r}. \quad (3.17) \end{aligned}$$

where the last term in Eq. (3.17) is the Lagrange multiplier for the constraint of proton density. The quantity N_p is the total number of protons in the system, i.e.,

$$\int_{\Omega_s} n(\mathbf{r}) d\mathbf{r} = N_p, \quad (3.18)$$

However, in most experimental set-ups, one does not know N_p . Instead, the bulk concentration or the bulk number density, n_p^0 , is given. When the solvent domain is

sufficiently large compared to the channel pore region, one has two approximations

$$N_p \cong n_p^0 \int_{\Omega_s} e^{-q(\Phi(\mathbf{r})-V_{\text{Ext}})/k_B T} d\mathbf{r} \cong n_p^0 \int_{\Omega_s} d\mathbf{r}, \quad (3.19)$$

where the second approximation is a crude estimation.

The energy functional (3.17) is a truly multiphysical and multiscale framework that contains the continuum approximation for solvent and membrane while explicitly takes into account for the channel protein in a discrete fashion. More importantly, it mixes the classical theory and quantum mechanical descriptions on an equal footing.

Note that Eq. (3.17) is a typical minimization-maximization problem, where the electrostatic free energy is to be minimized while the kinetic energy of protons is to be maximized. Fortunately, this situation does not create a problem as the optimization of the total free energy functional is achieved with two governing equations as described in the next section.

3.1.4 Governing equations

The present system has two unknown functions: the electrostatic potential Φ and the wavefunction Ψ_E . All other functions either are to be explicitly given or depend on Φ and Ψ_E . The governing equations for Φ and Ψ_E are to be derived from the free energy functional by variational principle via the Euler-Lagrange equation. This multiscale variational framework approach was developed in our recently work [154, 27]. It offers successful predictions of the solvation free energies of proteins and small compounds[31, 32].

Generalized Poisson-Boltzmann equations

The total free energy functional given above determines the density distribution and dynamics of protons. The governing equation for electrostatic potential can be derived

by the variation of the functional with respect to the potential Φ

$$\frac{\delta G_{\text{Total}}[\Phi, n]}{\delta \Phi} \implies -\nabla \cdot (\epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r})) = \rho_p(\mathbf{r}) + \rho_f(\mathbf{r}) + \rho'(\mathbf{r}). \quad (3.20)$$

Equation (3.20) is a generalized Poisson-Boltzmann (GPB) equation describing the electrostatic potential generated from three types of charge sources: the ions of interest, other ions species in the solvent described by the continuum approximation and the fixed point charges in biomolecules. This equation is not closed because $n(\mathbf{r})$ needs to be evaluated from another governing equation.

A special case of Eq. (3.20) is also very interesting. Let us assume that all ions in the system are described either by fixed point charges from biomolecules, or by the continuum treatment. Therefore, the system is closed and we arrive at the classical Poisson-Boltzmann equation

$$-\nabla \cdot (\epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r})) = \rho_f(\mathbf{r}) + \rho_s(\mathbf{r}), \quad (3.21)$$

where $\rho_s(\mathbf{r}) = \sum_{j=1}^{N_c} q_j n'_j(\mathbf{r})$, and N_c is for all ions in the continuum solvent.

Generalized Kohn-Sham equations

In the present multiscale model, the density n of protons in Eq. (3.20) is governed by generalized Kohn-Sham equations. This set of equations is obtained by the variation of the total free energy functional with respect to wavefunction Ψ_E^*

$$\frac{\delta G_{\text{Total}}[\Phi, n]}{\delta \Psi_E^*} \implies -\nabla \cdot \frac{\hbar^2}{2m(\mathbf{r})} \nabla \Psi_E(\mathbf{r}) + V(\mathbf{r}) \Psi_E(\mathbf{r}) = \lambda \Psi_E(\mathbf{r}) = E \Psi_E(\mathbf{r}) \quad (3.22)$$

where the multiplier λ is chosen as the eigenvalue E and

$$V(\mathbf{r}) = q\Phi(\mathbf{r}) + V_{\text{GC}}[n] + V_{\text{Exter}}(\mathbf{r})$$

is the effective potential, which includes electrostatic, generalized-correlation and external interactions. The effective potential is discussed in Section 3.1.2.

Equation (3.22) appears to be the conventional Kohn-Sham equation. However, there are some important differences. First, the exchange-correlation potential, which is crucial to electrons, is not presented in Eq. (3.7). The origin of the exchange-correlation potential is from the Fermi-Dirac distribution, spin and many other unknown effects. In the present theory, we use the generalized-correlation potential to represent many unaccounted effects. We assume the Boltzmann statistics for the ions of interest at ambient temperature. Additionally, we define the density as in Eq. (3.7), instead of the conventional choice for electrons: $n_{\text{electron}}(\mathbf{r}) = \sum_j |\Psi_j(\mathbf{r})|^2$. This definition is partially due to the Boltzmann statistics and partially due to the spectrum of the present Kohn-Sham operator, which is bounded from below. Technically, the Hamiltonian of the generalized Kohn-Sham equation (3.22) has not only discrete spectra, but also absolute continuum spectrum. As such, a Boltzmann factor in the density definition is indispensable. Finally, unlike the conventional Kohn-Sham equation, the present generalized Kohn-Sham equation is not a closed one. It is inherently coupled to the generalized Poisson-Boltzmann equation (3.20). This coupled Kohn-Sham and Poisson-Boltzmann system endows us the flexibility to deal with complex multiphysics in a multiscale fashion — the quantum dynamics in continuum.

3.1.5 Proton density operator for the non-hermitian Hamiltonian

As mentioned earlier, the external potential has a non-hermitian component to describe the interaction with truncated extracellular and intracellular surroundings. Let

us explicitly separate the anti-hermitian (or skew hermitian) components

$$V_{\text{Extra}} = V_{\text{Extra}}^h + V_{\text{Extra}}^{ah}, \quad V_{\text{Intra}} = V_{\text{Intra}}^h + V_{\text{Intra}}^{ah}, \quad (3.23)$$

where

$$V_{\alpha}^h = \frac{1}{2}(V_{\alpha} + V_{\alpha}^{\dagger}), \quad V_{\alpha}^{ah} = \frac{1}{2}(V_{\alpha} - V_{\alpha}^{\dagger}), \quad \alpha = \text{Extra, Intra}. \quad (3.24)$$

The non-hermitian parts of the external potentials describe the relaxation effect or spectral line shape broadening due to the interaction with the surroundings. Accordingly, we split the Hamiltonian as

$$H = H^h + V^{ah} = H^h + V_{\text{Extra}}^{ah} + V_{\text{Intra}}^{ah}. \quad (3.25)$$

We first note that the density of protons can be further given by

$$n_H = \int e^{-(E-E_{\text{Ext}})/k_B T} \delta(E - H) dE. \quad (3.26)$$

In this work, we define the spectral operator $\delta(E - H)$ as

$$\delta(E - H) = \frac{i}{2\pi} \lim_{\varepsilon \rightarrow 0} \lim_{\|V^{ah}\| \rightarrow 0} \left[\frac{1}{E - (H - i\varepsilon)} - \frac{1}{E - (H - i\varepsilon)^{\dagger}} \right] \quad (3.27)$$

We therefore approximate the proton density operator by

$$n_H = \frac{i}{2\pi} \int e^{-(E-E_{\text{Ext}})/k_B T} \left[G(E) - G^{\dagger}(E) \right] dE, \quad (3.28)$$

where G is the Green's function (operator)

$$G(E) = (E - H)^{-1}. \quad (3.29)$$

We therefore arrive at a useful expression for the proton density

$$n_H = \frac{i}{\pi} \int e^{-(E-E_{\text{Ext}})/k_B T} \left[\sum_{\alpha} G(E) V_{\alpha}^{ah} G^{\dagger}(E) \right] dE \quad (3.30)$$

$$= \frac{i}{\pi} \sum_{\alpha} \int e^{-(E-E_{\alpha})/k_B T} G(E) V_{\alpha}^{ah} G^{\dagger}(E) dE, \quad (3.31)$$

where $\alpha = \text{Extra, Intra, } E_{\text{Extra}}$ and E_{Intra} are the external electrical field energies at extracellular and intracellular electrodes, respectively. Note that E_{Ext} behaves like an operator such that its value is chosen according to the nearest external interaction. Equation (3.31) provides an appropriate expression for computing the total proton density.

3.1.6 Proton transport

Typically, external electrical field is applied as the difference of electrical potentials, $(E_{\text{Extra}}/q - E_{\text{Intra}}/q)$. The experimental measurements are given as the current and voltage curve, or the so called I-V curve. Therefore, a major goal of our theoretical model is to provide predictions of the current under different external voltages. The current in the standard quantum mechanics is given by

$$I = q \text{Tr} \frac{1}{2} (n_H v^{\dagger} + v n_H) \quad (3.32)$$

$$= q \int \int \frac{\hbar}{2mi} [\Psi_E^*(\mathbf{r}) \nabla \Psi_E(\mathbf{r}) - \Psi_E(\mathbf{r}) \nabla \Psi_E^*(\mathbf{r})] e^{-\frac{(E-E_{\text{Ext}})}{k_B T}} d\mathbf{r} dE \quad (3.33)$$

where Tr is the trace operation and $\frac{1}{2} (n_H v^{\dagger} + v n_H)$ is the symmetrized current operator with v being the velocity vector. Equation (3.33) requires the evaluation of the full scattering wavefunction $\Psi_E(\mathbf{r})$. The spatial derivative can be carried out at a location consistent with the specific feature of the external electrical field E_{Ext} .

An alternative current expression can be given by examining the transition rates

due to the anti-hermitian parts of the external interaction potential. Let us evaluate the transition rate according to the interaction potential V_{Extra}^{ah}

$$I = q \frac{1}{i\hbar} \text{Tr} \frac{1}{2} \left[n_H \left(V_{\text{Extra}}^{ah} \right)^\dagger + V_{\text{Extra}}^{ah} n_H \right] \quad (3.34)$$

$$= \frac{q}{h} \text{Tr} \left\{ \int e^{-(E-E_{\text{Ext}})/k_B T} \sum_{\alpha} G(E) V_{\alpha}^{ah} G^\dagger(E) \left(V_{\text{Extra}}^{ah} \right)^\dagger dE \right. \\ \left. + \int V_{\text{Extra}}^{ah} e^{-(E-E_{\text{Ext}})/k_B T} \sum_{\alpha} G(E) V_{\alpha}^{ah} G^\dagger(E) dE \right\} \quad (3.35)$$

Now we need to make a decision for E_{Ext} because each term involves two interaction potentials. In this work, we systematically choose E_{Ext} according to the nearest external interaction

$$I = \frac{q}{h} \text{Tr} \left\{ \int e^{-(E-E_{\alpha})/k_B T} \sum_{\alpha} G(E) V_{\alpha}^{ah} G^\dagger(E) \left(V_{\text{Extra}}^{ah} \right)^\dagger dE \right. \\ \left. + \int V_{\text{Extra}}^{ah} e^{-(E-E_{\text{Extra}})/k_B T} \sum_{\alpha} G(E) V_{\alpha}^{ah} G^\dagger(E) dE \right\} \quad (3.36)$$

$$= \frac{q}{h} \text{Tr} \int G V_{\text{Intra}}^{ah} G^\dagger V_{\text{Extra}}^{ah} \left[e^{-\frac{E-E_{\text{Extra}}}{k_B T}} - e^{-\frac{E-E_{\text{Intra}}}{k_B T}} \right] dE \quad (3.37)$$

Similarly, we obtain a current expression by using the interaction potential V_{Intra}^{ah}

$$I = q \frac{1}{i\hbar} \text{Tr} \frac{1}{2} \left[n_H \left(V_{\text{Intra}}^{ah} \right)^\dagger + V_{\text{Intra}}^{ah} n_H \right] \\ = \frac{q}{h} \text{Tr} \int G V_{\text{Extra}}^{ah} G^\dagger V_{\text{Intra}}^{ah} \left[e^{-\frac{E-E_{\text{Intra}}}{k_B T}} - e^{-\frac{E-E_{\text{Extra}}}{k_B T}} \right] dE \quad (3.38)$$

Equations (3.37) and/or (3.38) can be used for current evaluations under different external electrical field strengths and concentrations.

3.2 Computational algorithms

The implementation of the theoretical model described in Section 3.1.4 involves a number of computational issues. The present section is devoted to the computational implementation of our quantum dynamics in continuum model.

3.2.1 Proton density structure and transport

Proton density structure concerns the solution of the generalized Kohn-Sham equation whereas the proton transport offers the current-voltage curves, which are to be compared with experimental measurement. This subsection describes the solution strategy of the generalized Kohn-Sham equation and theoretical prediction of experimental data.

The solution of the generalized Kohn-Sham equation

Typically, solving the full-scale Kohn-Sham equation can be a major obstacle in the simulation. Due to the fact that biological characteristics for each subdomain of the ion channel system are quite different and the Kohn-Sham operator will have distinct properties correspondingly. In this subsection, we make use of various decomposition schemes to reduce the computational complexity in solving Eq. (3.22).

Motions of quantum particles in the present system can be generally classified into three categories: scattering along transport directions, confined motion and free motion. The channel pore direction (i.e., the z direction) is designated as the transport direction, in which protons cross the transmembrane protein or scatter back to the solvent. Along the z direction, the Kohn-Sham operator has an absolutely continuous spectrum. In the $x - y$ directions, the Kohn-Sham equation possesses different behaviors. In the extracellular and intracellular regions where the solvent domains are sufficiently large, proton motions are essentially unconfined in the $x - y$ direc-

tions. They undergo intensive electrostatic and generalized-correlation interactions although the system can be regarded as near the equilibrium. The associated Kohn-Sham operator for protons also has an absolutely continuous spectrum. In contrast, in channel pore region, the protons are confined in $x - y$ plane by the channel wall. In the confined plane, the Kohn-Sham operator is essentially compact and has a discrete spectrum. For two different regions, formulations and corresponding treatments of the proton density are different.

The proton density structure in the channel pore is crucial to the proton transport. Whereas, the behavior of protons in the bath is relatively less important. Therefore, as a good approximation, we can truncate the computational domain in the bath regions. Consequently, the Kohn-Sham operator becomes compact for all $x - y$ directions and has discrete eigenvalues. As a good approximation for many ion channels, we split the total wavefunction $\Psi_E(\mathbf{r})$ as

$$\Psi_E(\mathbf{r}) = \psi^j(x, y; z)\psi_k^j(z) \quad (3.39)$$

where $\psi^j(x, y; z)$ is the j -th eigen-mode in the confined directions at a specific location z , and $\psi_k^j(z)$ is the wavefunction along the transport direction, with transport wave number k . Under this circumstance, it is convenient to relabel the total energy E as E_k^j , where j and k are related to the energies for confined and transport directions, respectively. If the mode-mode interaction along the confined direction is neglected, it is easy to verify that ψ^j and ψ_k^j satisfy the following decomposed Kohn-Sham equations,

$$\left[-\frac{\hbar^2}{2} \left(\frac{\partial}{\partial x} \frac{1}{m_x} \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \frac{1}{m_y} \frac{\partial}{\partial y} \right) + V(x, y; z) \right] \psi^j = U^j(z)\psi^j \quad (3.40)$$

$$\psi^j(x, y; z) = 0 \quad \text{on} \quad \partial\Omega_D(z);$$

$$\left[-\frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{m_z} \frac{\partial}{\partial z} + U^j(z) \right] \psi_k^j(z) = E_k^j \psi_k^j(z), \quad j = 1, 2, \dots, \quad (3.41)$$

where $V(x, y; z)$ is the restriction of the potential operator $V(x, y, z)$ at position z , $U^j(z)$ is the j th eigenvalue of the 2D problem at position z , and $\psi^j = \psi^j(x, y; z)$ is the corresponding eigenfunction. Here $\psi_k^j(z)$ is the scattering wavefunction associated with the scattering potential $U^j(z)$. Here $\partial\Omega_D(z)$ is the boundary for the cross section at z . The transport equation (3.41) can be solved as a scattering problem. Finally the proton density (3.7) can be modified as

$$\begin{aligned} n(\mathbf{r}) &= \sum_j \int |\psi^j(x, y; z)|^2 |\psi_k^j(z)|^2 e^{-(E_k^j - E_{\text{Ext}})/k_B T} dE_k^j \\ &\doteq \sum_j |\psi^j(x, y; z)|^2 n_{\text{scat}}^j(z). \end{aligned} \quad (3.42)$$

Equation (3.42) only gives the symbolic proton density structures for an unspecified E_{Ext} . More detailed consideration of E_{Ext} requires the further treatment of the scattering boundary conditions as shown in Sections 3.1.5 and 3.1.6. However, the 2D wavefunction $|\psi^j(x, y; z)|^2$ in Eq. (3.42) can be evaluated from the Kohn-Sham equation (3.40). The solution to this equation is quite standard — it is just the eigenvalue problem of an equation of elliptic type. While to solve the transport problem, as indicated in the theory, one needs to find appropriate expressions of the non-hermitian external operators. The corresponding computational aspects are presented in the next subsection.

Boundary treatment of the transport calculation

Although the quantum confinement Eq. (3.40) only happens in finite channel region, the transport problem Eq. (3.41) is associated with infinitely large surroundings, in principle. Since the same procedure is used to solve Eq. (3.41) for different j , let us

drop the j label

$$\left(-\frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{m_z} \frac{\partial}{\partial z} + U\right) \psi_k(z) = E\psi_k(z), \quad z \in (-\infty, \infty), \quad (3.43)$$

where $-\frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{m_z} \frac{\partial}{\partial z} + U$ is the scattering Hamiltonian and E is the scattering energy. In practical computations, the extracellular and intracellular surroundings have to be truncated. Suppose $[z_{\text{Extra}}, z_{\text{Intra}}]$ is the finite transport interval of interest and the regions $(-\infty, z_{\text{Extra}})$ and $(z_{\text{Intra}}, \infty)$ are assumed as infinitely long extracellular and intracellular environments. We assume that in regions $(-\infty, z_{\text{Extra}})$ and $(z_{\text{Intra}}, \infty)$, the interaction potential U is independent of position due to the spatial average of homogenization type over the large scale. Consequently, Eq. (3.43) admits planewave solutions asymptotically. For instance, if one considers the wavefunctions $\psi_k(z)$ in the extracellular environment, it has the following form

$$\begin{aligned} \psi_k(z) &= e^{ikz} + r_m e^{-ikz} & \text{if } z \in (-\infty, z_{\text{Extra}}) \\ \psi_k(z) &= t_m e^{ikz} & \text{if } z \in (z_{\text{Intra}}, \infty) \end{aligned} \quad (3.44)$$

where r_m and t_m are reflection and transmission coefficients, respectively. Given the specific formulation of the wavefunction in the extracellular bath, Eq. (3.44) can be employed as boundary conditions of Eq. (3.43) to obtain the proton density originated from the extracellular part. Similar boundary conditions for the intracellular part can be derived in the same fashion.

Suppose that the interval $[z_{\text{Extra}}, z_{\text{Intra}}]$ is discretized as $z_{\text{Extra}} = z_1, z_2, \dots, z_N = z_{\text{Intra}}$, where N is the total number of grid points and the grid size is denoted as $\Delta z = (z_2 - z_1)/N$. For simplicity, let $t = \frac{\hbar^2}{2m_z(\Delta z)^2}$, then for interior points z_i , ($i = 2, \dots, N-1$), the discretization of Eq. (3.43) is quite standard by the finite difference method

$$-t\psi_{i-1} + (2t + U_i - E)\psi_i - t\psi_{i+1} = 0 \quad (3.45)$$

where ψ_i represents the numerical solution of $\psi_k(z_i)$ and U_i is for $U(z_i)$. For the discretization at boundary point z_1 , we first define a fictitious function value of $\psi(z)$ on z_0 , the point ahead of z_1 as ψ_0 , then the discretization at z_1 is

$$-t\psi_0 + (2t + U_i - E)\psi_1 - t\psi_2 = 0. \quad (3.46)$$

Now one needs to determine the fictitious value ψ_0 in terms of ψ_i , ($i = 1, 2, \dots, N$).

From the boundary condition (3.44), we have

$$\begin{aligned} \psi_0 &= e^{ik_0 z_0} + r_m e^{-ik_0 z_0} \\ \psi_1 &= e^{ik_1 z_1} + r_m e^{-ik_1 z_1}. \end{aligned} \quad (3.47)$$

In fact, we have $k_0 = k_1$ since the free motion of the wave in the asymptotic regions.

We can denote k_0 and k_1 by k_1 with $\frac{(\hbar k_1)^2}{2m_z} = E - U_1$. By this notation, we have

$$\begin{aligned} \psi_0 - \psi_1 e^{ik_1 \Delta z} &= e^{ik_1 z_0} - e^{ik_1(z_1 + \Delta z)} \\ &= e^{ik_1(z_1 - \Delta z)} - e^{ik_1(z_1 + \Delta z)}. \end{aligned} \quad (3.48)$$

Inserting Eq. (3.48) into Eq. (3.46), one yields

$$-t\psi_1 e^{ik_1 \Delta z} + (2t + U_1 - E)\psi_1 - t\psi_2 = -2ti \sin(k_1 \Delta z) e^{ik_1 z_1}. \quad (3.49)$$

Applying the same strategy for ψ_N and fictitious function value ψ_{N+1} , we have

$$\psi_{N+1} - \psi_N e^{ik_N \Delta z} = t_m e^{ik_N z_{N+1}} - t_m e^{ik_N z_N} e^{ik_N \Delta z} = 0, \quad (3.50)$$

where $\frac{(\hbar k_N)^2}{2m_z} = E - U_N$ and further

$$-t\psi_{N-1} + (2t + U_N - E)\psi_N - t\psi_N e^{ik_N \Delta z} = 0. \quad (3.51)$$

Follow the same boundary treatment for the intracellular environment, the whole system is discretized in vector and matrix forms as the following

$$\mathbf{G}^{-1}\Psi_{\text{Extra}} = (\mathbf{H}^{\text{s}} - E\mathbf{I})\Psi = \mathbf{b}_{\text{Extra}} \quad (3.52)$$

where $\Psi_{\text{Extra}} = (\psi_1, \psi_2, \dots, \psi_N)^T$, \mathbf{I} is the identity matrix of dimension $N \times N$ and

$$\mathbf{H}^{\text{s}} = \begin{pmatrix} 2t + U_1 - te^{ik_1\Delta z} & -t & \dots & \dots & 0 \\ -t & 2t + U_2 & -t & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & -t & 2t + U_N - te^{ik_N\Delta z} \end{pmatrix}_{N \times N}. \quad (3.53)$$

Here $\mathbf{b}_{\text{Extra}}$ is the source term for the incoming waves from the extracellular surroundings

$$\mathbf{b}_{\text{Extra}} = (2ti \sin(k_1\Delta z)e^{ik_1z_1}, 0, \dots, 0)^T. \quad (3.54)$$

The wavefunction Ψ_{Extra} can be written as

$$\Psi_{\text{Extra}} = \mathbf{G}\mathbf{b}_{\text{Extra}}. \quad (3.55)$$

Let $\Psi_{\text{Extra}}^\dagger$ be the complex conjugate of Ψ_{Extra} . We have

$$\Psi_{\text{Extra}} \Psi_{\text{Extra}}^\dagger = \mathbf{G}\mathbf{b}_{\text{Extra}} \mathbf{b}_{\text{Extra}}^\dagger \mathbf{G}^\dagger = \mathbf{G} \begin{pmatrix} [2t \sin(k_1\Delta z)]^2 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} \mathbf{G}^\dagger. \quad (3.56)$$

Similar derivation can be carried out for the wavefunction Ψ_{Intra} related to intracel-

lular surroundings,

$$\Psi_{\text{Intra}} \Psi_{\text{Intra}}^\dagger = \mathbf{G} \mathbf{b}_{\text{Intra}} \mathbf{b}_{\text{Intra}}^\dagger \mathbf{G}^\dagger = \mathbf{G} \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & [2t \sin(k_N \Delta z)]^2 \end{pmatrix} \mathbf{G}^\dagger. \quad (3.57)$$

Therefore, the total density matrix is

$$\mathbf{D} = \frac{1}{2\pi} \int \left[\sum_{\alpha} e^{-(E-E_{\alpha})/k_B T} \mathbf{G} \mathbf{b}_{\alpha} \mathbf{b}_{\alpha}^\dagger \mathbf{G}^\dagger \right] dk, \quad \alpha = \text{Extra, Intra}. \quad (3.58)$$

Use the relation

$$dE = d \frac{(\hbar k)^2}{2m} + 0 = \frac{\hbar^2 k}{m} dk \quad (3.59)$$

to change the above integral into that with respect to energy E , and use the simple limit $\sin(k\Delta z)/(k\Delta z) \rightarrow 1$ as $\Delta z \rightarrow 0$, the above integral can be easily revised as

$$\mathbf{D} = \frac{i}{\pi \Delta z} \int \left[\sum_{\alpha} e^{-(E-E_{\alpha})} \mathbf{G} V_{\alpha}^{ah} \mathbf{G}^\dagger \right] dE, \quad \alpha = \text{Extra, Intra}, \quad (3.60)$$

where

$$V_{\text{Extra}}^{ah} = \begin{pmatrix} -it \sin(k_1 \Delta z) & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} \quad (3.61)$$

and

$$V_{\text{Intra}}^{ah} = \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & -it \sin(k_N \Delta z) \end{pmatrix}. \quad (3.62)$$

It is clear that V_{Extra} and V_{Intra} are the non-hermitian components in the external potential Eq. (3.23) that are introduced to truncate the surroundings. Since V_{α}^{ah} is solely nonzero for one entry in the matrix and this fact is independent of the discretization, it is easy to verify that $\lim_{\Delta z \rightarrow 0} \|V_{\alpha}^{ah}\| = 0$, as required in Eq. (3.27).

Obviously, Eq. (3.60) is actually the discretization form of Eq. (3.31). Finally, the scattering number density is calculated as

$$n_{\text{scat}}(z) = \text{diag}(\mathbf{D}). \quad (3.63)$$

3.2.2 Dirichlet-to-Neumann mapping for the generalize PB equation

Considering Eq. (3.4) and expression (3.20), the generalized Poisson-Boltzmann equation is

$$-\nabla \cdot (\epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r})) = qn(\mathbf{r}) + \sum_{i=1}^{N_a} Q_i \delta(\mathbf{r} - \mathbf{r}_i) + \sum_{j=1}^{N'_c} q_j n_j^0 e^{-\frac{q_j(\Phi - V_{\text{Ext}})}{k_B T}} \quad (3.64)$$

Recall the fact that the electrostatic potential $\Phi(\mathbf{r})$ is defined throughout the domain Ω , which is inhomogeneous with respect to the dielectric constant $\epsilon(\mathbf{r})$. Therefore, we need to physically impose the continuity matching conditions at the interface Γ of two adjunctive subregions. The continuity matching conditions are given as

$$[\Phi]_{\Gamma} = \Phi^+(\mathbf{r}) - \Phi^-(\mathbf{r}) = 0, \quad (3.65)$$

$$[\epsilon \nabla \Phi \cdot \vec{\mathbf{n}}]_{\Gamma} = \epsilon^+ \nabla \Phi^+(\mathbf{r}) \cdot \vec{\mathbf{n}} - \epsilon^- \nabla \Phi^-(\mathbf{r}) \cdot \vec{\mathbf{n}} = 0 \quad (3.66)$$

where superscripts “+” and “-” represent the limiting values of a certain function at two sides of interface Γ , and $\vec{\mathbf{n}}$ is the unit outward normal direction of Γ . Equations (3.65) and (3.66) guarantee the continuities of the potential function and its flux.

Theoretically, Eq. (3.64) admits the boundary condition $\Phi(\infty) = 0$ at the infinity. However, in practical computation, a finite domain is used and appropriate boundary conditions need to be imposed at the domain boundary $\partial\Omega$. In our studies, the channel protein and the associated membrane are embedded in a rectangular cuboid with appropriate sizes. It is very nature to apply the Dirichlet boundary conditions along the electrode portions of the rectangular cuboid boundary, while for the remainder of the boundary, we apply the Neumann boundary condition (i.e., the zero normal electric field conditions).

Physically, the generalized Poisson-Boltzmann equation (3.64) has two types of charge source terms, i.e., the fixed charges given by the delta functions, and the unsteady charges. Therefore, it is wise to treat these source terms separately such that when we keep updating the unsteady source term, we just need to compute the effect of the fixed charge source term once. Mathematically, the solution of Eq. (3.64) has a singular part due to the delta function (i.e., fixed charges) which may cause computational problems. Thus, we should treat the regular part and the singular part of the solution differently [67]

$$\Phi = \bar{\Phi} + \tilde{\Phi} \quad (3.67)$$

where $\bar{\Phi}$ and $\tilde{\Phi}$ denote the singular part and regular part of Φ , respectively. More specifically, $\bar{\Phi}$ should correspond to the singular delta function term and vanish outside the protein and membrane domain Ω_m , while $\tilde{\Phi}$ is defined in the whole domain. By this consideration, we split $\bar{\Phi}(\mathbf{r})$ as

$$\bar{\Phi}(\mathbf{r}) = \Phi^*(\mathbf{r}) + \Phi^0(\mathbf{r}) \quad (3.68)$$

where

$$\Phi^*(\mathbf{r}) = \sum_{i=1}^{N_a} \frac{Q_i}{\epsilon_m |\mathbf{r} - \mathbf{r}_i|} \quad (3.69)$$

represents the Coulomb's potential from the protein fixed charges. Since $\bar{\Phi}(\mathbf{r})$ is required to vanish outside the Ω_m as well as the boundary $\partial\Omega_m$, the $\Phi^*(\mathbf{r})$ should be corrected by $\Phi^0(\mathbf{r})$, which is a harmonic function on Ω_m and

$$\Phi^0(\mathbf{r}) = -\Phi^*(\mathbf{r}), \quad \forall \mathbf{r} \in \partial\Omega_m. \quad (3.70)$$

For the regular part $\tilde{\Phi}$, we can take the advantage of the fact that n_j^0 is zero in Ω_m , and have the following equation and interface jump conditions:

$$-\nabla \cdot \left(\epsilon(\mathbf{r}) \nabla \tilde{\Phi}(\mathbf{r}) \right) - \sum_{j=1}^{N'_c} q_j n_j^0 e^{-\frac{q_j(\tilde{\Phi} - V_{\text{Ext}})}{k_B T}} = qn(\mathbf{r}) \quad (3.71)$$

$$[\tilde{\Phi}]_{\Gamma} = 0 \quad (3.72)$$

$$[\epsilon \nabla \tilde{\Phi} \cdot \bar{\mathbf{n}}]_{\Gamma} = -[\epsilon \bar{\Phi} \cdot \bar{\mathbf{n}}]_{\Gamma} \quad (3.73)$$

Through Eqs. (3.67) to (3.71), the electrostatic potential Φ is decomposed into a singular part and a regular part. It should be noted that it is $\tilde{\Phi}$ that is coupled to the Kohn-Sham equation since $\bar{\Phi}$ is solely nonzero in the protein and membrane region. The effect of the fixed charges is actually first mapped on the $\partial\Omega_m$ in a Dirichlet sense (Eq. (3.70)) and reflected into the solvent region in a Neumann manner (Eq. (3.73)) at the solvent-protein interface Γ . This Dirichlet-to-Neumann mapping (DNM) analytically takes care of the Dirac delta functions and is successfully employed in various applications [67, 27]. The trade-off of this treatment is that one has to solve an elliptic equation (3.71) with non-homogeneous interface jump conditions.

Traditional finite difference or finite element methods fail to come up with high-order accuracy and convergence in solving Eq. (3.71) due to geometric singularities in

the molecular surface [131] and the need to enforce the interface conditions (3.72) and (3.73). The matched interface and boundary (MIB) method has been developed for elliptic equations with complex interfaces, geometric singularity, and singular charges [158, 159, 67, 165, 163, 26]. It offers second-order accuracy and convergence in solving the Poisson-Boltzmann equation with biomolecular context [159, 67, 163, 26]. Therefore, the combination of DNM and MIB provides a robust and efficient solution to the generalized PB equation with second-order accuracy and convergence, even for complex channel protein geometries.

3.2.3 The self-consistent iteration

In this section we analyze the self-consistent iteration between the generalized PB equation and the Kohn-Sham equation. To focus on the essential idea, Eq. (3.71) is symbolically written as

$$L\tilde{\Phi} + F(\tilde{\Phi}) = \rho_p, \quad (3.74)$$

where $\tilde{\Phi}$ and ρ_p represent the electrostatic potential energy and proton charge density, L represents the linear part of the GPB equation while the $F(\tilde{\Phi})$ is the nonlinear part. Simply substituting the quantity ρ_p into Eq. (3.74) does not offer a clue about the iteration convergence analysis and efficiency. The Gummel iteration [50] proposed in semiconductor device applications was verified practically that it works well for a similar self-consistent iteration problem. The idea of the Gummel iteration is described below.

The proton charge density ρ_p and the electrostatics potential $\tilde{\Phi}$ are assumed to have the following intrinsic connection

$$\rho_p(\mathbf{r}) = \mathcal{F}(\tilde{\Phi}(\mathbf{r}), E_{\text{Ext}}), \quad (3.75)$$

where $\mathcal{F}(\tilde{\Phi}, E_{\text{Ext}}) = qn_0 e^{-(q\tilde{\Phi} - E_{\text{Ext}})/k_B T}$ is the Boltzmann function and n_0 is the reference number density of the protons. Equation (3.75) represents the relation between the electrostatic potential and the particle density in the equilibrium state. However, the relation does not hold any more at non-equilibrium. Nevertheless, we can extend E_{Ext} to a function defined over the entire domain $E_{\text{Ext}}(\mathbf{r})$ such that $\rho_p(\mathbf{r}) = \mathcal{F}(\tilde{\Phi}(\mathbf{r}), E_{\text{Ext}}(\mathbf{r}))$. The intermediate values of $E_{\text{Ext}}(\mathbf{r})$ can be easily found once ρ_p and $\tilde{\Phi}(\mathbf{r})$ are available. Based on this argument, Eq. (3.74) is written as a new nonlinear equation

$$L\tilde{\Phi} + F(\tilde{\Phi}) = \mathcal{F}(\tilde{\Phi}, E_{\text{Ext}}). \quad (3.76)$$

We need to linearize Eq. (3.76) appropriately. Note that

$$\mathcal{F}'(\tilde{\Phi}, E_{\text{Ext}}) = -\frac{q}{k_B T} \mathcal{F}(\tilde{\Phi}, E_{\text{Ext}}) = -\frac{q}{k_B T} \rho_p$$

with $\mathcal{F}'(\tilde{\Phi}, E_{\text{Ext}})$ being the Fréchet derivatives of \mathcal{F} with respect to $\tilde{\Phi}$. Similarly, $F'(\tilde{\Phi})$ can be evaluated.

Suppose $\tilde{\Phi}^l$, E_{Ext}^l and ρ_p^l are the values of $\tilde{\Phi}$, E_{Ext} and ρ_p at l th step iteration, then the Newton's method for solving Eq. (3.76) is naturally reduced to the Gummel iteration:

$$\left(L + F'(\tilde{\Phi}^l) + \frac{q}{k_B T} \rho_p^l \right) \Delta \tilde{\Phi}^l = \rho_p^l - L\tilde{\Phi}^l - F(\tilde{\Phi}^l) \quad (3.77)$$

where we update $\tilde{\Phi}^{l+1}$ as $\tilde{\Phi}^{l+1} = \tilde{\Phi}^l + \lambda \Delta \tilde{\Phi}^l$ and $0 < \lambda \leq 1$ is chosen through a line search to guarantee

$$\|L\tilde{\Phi}^{l+1} + F(\tilde{\Phi}^{l+1}) - \rho_p^{l+1}\| < \|L\tilde{\Phi}^l + F(\tilde{\Phi}^l) - \rho_p^l\|. \quad (3.78)$$

Once $\tilde{\Phi}^{l+1}$ and ρ_p^{l+1} are obtained, E_{Ext}^{l+1} can be modified, and whole iteration can continue till the convergence is achieved. It is worthwhile to point out that in order to improve numerical efficiency, Eq. (3.77) can be solved by applying various inexact

Newton's methods. There is plenty of literature about the convergence order discussion so it is necessary for us to generalize the Gummel iteration to the Newton's method.

Another technique to enhance the self-consistent convergence is the relaxation method [27]. If we define the K_s , U_s and N_s as the spaces which the external potential $E_{\text{Ext}}(\mathbf{r})$, electrostatics $\tilde{\Phi}(\mathbf{r})$ and proton charge density $\rho(\mathbf{r})$ belong to, respectively. For the whole iteration of the generalized Poisson-Boltzmann Kohn-Sham system, it can be interpreted as the application of the fixed point map \mathcal{T} on any of the above spaces, say $\mathcal{T} : U_s \rightarrow U_s$ for the electrostatics

$$\tilde{\Phi}(\mathbf{r}) = \mathcal{T}(\tilde{\Phi}(\mathbf{r})). \quad (3.79)$$

To characterize the details of the map \mathcal{T} , we denote the operator $\mathcal{G} : U_s \rightarrow N_s$, which indicates the process of using the Kohn-Sham equation to solve for proton charge density based on the electrostatic potential. Such a process is followed by $\mathcal{F}^{-1} : N_s \rightarrow K_s$, which updates $E_{\text{Ext}}(\mathbf{r})$ by $\rho_p(\mathbf{r})$ and $\tilde{\Phi}(\mathbf{r})$. Finally $\mathcal{L} : K_s \rightarrow U_s$ represents solving the nonlinear GPB equation. The combination of all the above operations yields the definition of the operator \mathcal{T} , which shows the outer iteration

$$\mathcal{T} := \mathcal{L} \circ \mathcal{F}^{-1} \circ \mathcal{G} \quad (3.80)$$

and

$$\tilde{\Phi}^{l+1} = \mathcal{L} \circ \mathcal{F}^{-1} \circ \mathcal{G}(\tilde{\Phi}^l). \quad (3.81)$$

The relaxation scheme converts Eq. (3.81) into the steady-state problem of an ordinary differential equation (ODE)

$$\frac{\partial \tilde{\Phi}}{\partial t} = \mathcal{L} \circ \mathcal{F}^{-1} \circ \mathcal{G}(\tilde{\Phi}) - \tilde{\Phi}. \quad (3.82)$$

Therefore many ODE related techniques such as the Runge-Kutta method can be used to improve the convergence properties. One simple treatment is the discretization of Eq. (3.82) as

$$\frac{\tilde{\Phi}^{l+1} - \tilde{\Phi}^l}{\beta} = \mathcal{L} \circ \mathcal{F}^{-1} \circ \mathcal{G}(\tilde{\Phi}^n) - \tilde{\Phi}^n, \quad (3.83)$$

which leads to a self-consistent iteration with a relaxation factor β [27, 31]

$$\begin{aligned} \tilde{\Phi}^* &= \mathcal{L} \circ \mathcal{F}^{-1} \circ \mathcal{G}(\tilde{\Phi}^n) \\ \tilde{\Phi}^{n+1} &= \beta \tilde{\Phi}^* + (1 - \beta) \tilde{\Phi}^n. \end{aligned} \quad (3.84)$$

The traditionally used outer loop iteration actually is the special case of Eq. (3.84) with $\beta = 1$. By carefully choosing the relax factor β , one can reach the steady state (fix point) by self-consistent iterations.

3.2.4 The work flow of the self-consistent iteration

In previous sections algorithms and related mathematical treatments for solving the GPB equation and the Kohn-Sham equation individually are introduced. Here we assemble all the components together and give a main work flow for the numerical simulation of these coupled equations.

- **Step 0.** Preparation. All the necessary preparations for the whole loop are accomplished in this step, which include:
 - 1. The channel protein of interest is downloaded from the Protein Data Bank. The partial charges, positions, radii of all atoms as well as molecular surfaces are determined by CHARMM force field [105] and related software packages, such as PDB2PQR, see Section 3.3 for detail. The prepared channel structure and surface are then embedded in a proper computational domain.

- 2. Use Eqs. (3.69) and (3.70) to solve for $\bar{\Phi}$, then the quantity in Eq. (3.73) is obtained. Implement the DNM and the MIB schemes to discretize the Laplace operator as matrix L .
- **Step 1.** Solving the generalized PB equations (3.71) and (3.73). Given ρ_p^m (taken an initial guess if $m = 0$), use the inexact Newton’s method, Eq. (3.77) and Eq. (3.78) to obtain $\tilde{\Phi}^m$. Note that the index l in Eq. (3.77) is for the Newton’s method or inner iteration and the index m is for the outer or whole self-consistent iteration loop.
- **Step 2.** Solving the Kohn-Sham equation. The solution of the Kohn-Sham equation consists of two parts, the eigenvalue problem and the scattering problem with the evaluated electrostatic potential energy operator $U = q\tilde{\Phi}^m$.
 - 1. Solving the eigenvalue problem Eq. (3.40).
 - 2. Solving the transport problem Eq. (3.41).
 - 3. Assembling the total charge density ρ_p^{m+1} by Eqs. (3.42) and (3.63).
- **Step 3.** Convergence check. Go to Step 1 to obtain $\tilde{\Phi}^{m+1}$, if $\|\tilde{\Phi}^{m+1} - \tilde{\Phi}^m\| < \varepsilon_1$ and $\|\rho_p^{m+1} - \rho_p^m\| < \varepsilon_2$, where ε_1 and ε_2 are predefined error tolerances, then go to Step 4; otherwise go to Step 2.
- **Step 4.** Current calculation by Eq. (3.37).

Figure 3.2 gives an explicit illustration of the above work flow.

3.2.5 Model parameter selection

The selection of generalized-correlation potential

Generalized-correlation effects are important to ion conductance efficiency. Unfortunately, it is expensive to give a full quantitative description for $U_{GC}[n]$. In current

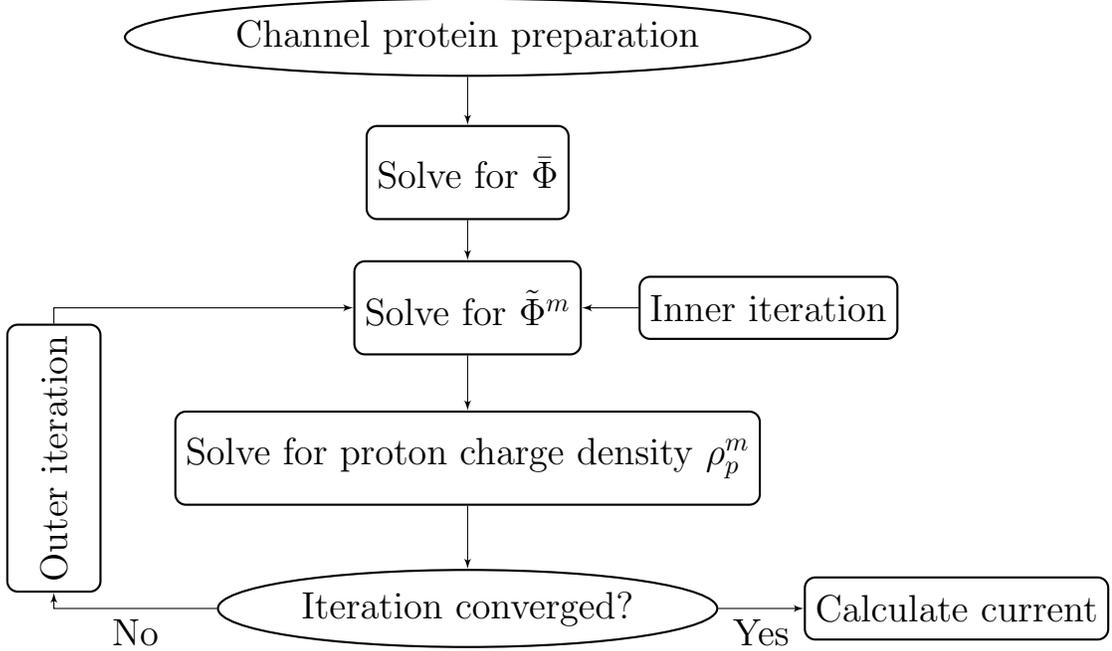


Figure 3.2: Work flow of the overall self-consistent iteration.

existing models, such as PNP based ones, the generalized correlation is integrated as an overall effect and represented implicitly by the phenomenologically reduced diffusion coefficients in the channel pore region. While in BD based models, the effect of generalized correlations is included in the ion friction factor, which is also related to the diffusion coefficient by Einstein's relation [95]. All these treatments indicate that $U_{GC}[n]$ should be related to the diffusion coefficient of ions, which is a physical observable. Based on this discussion, we ignore all detailed components while describe the generalized-correlation interactions as one effective, overall component in the mean field manner. As indicated by Eq. (3.12), the $U_{GC}[n]$ is also a density functional of the $n(\mathbf{r})$, and the first term represents the connection between $U_{GC}[n]$ and given reference ion density. It is quite obvious that α is a tunable parameter. Here we focus on how to choose parameter $V_{\text{Ion-sur}}$.

For a simple start, let this energy be related to the relaxation time τ of an ion by $V_{\text{Ion-sur}} = \hbar/2\tau$, according to the Einstein's relation $D = k_B T \tau / m$, where D and m

are the diffusion coefficient and mass of the particle. Then the energy $V_{\text{Ion-sur}}$ can be given by

$$V_{\text{Ion-sur}} = \frac{\hbar k_B T}{2mD} \quad (3.85)$$

for protons. With a appropriate proton mass and the diffusion coefficient in the bath, one yields $V_{\text{Ion-sur}} \approx 3.4k_B T$. However, the value of diffusion coefficient in the channel is commonly believed reduced, but is inconclusive due to the variation of the channel pore structure diameters and solvation conditions. According to Table 1 of Ref. [42], proton diffusion coefficients reduce to 1/2 to 1/7 of that in the bath condition in various lipid layers. We take the resulting reduction accordingly in the channel region. This argument gives the U_{GC} a range of $6 \sim 20k_B T$.

Choices of the dielectric constants

The Poisson equation describes the electrostatic potential function due to existence of free charges. The left hand side of the Poisson equation can be written as

$$-\nabla^2 \Phi(\mathbf{r}) + \nabla \cdot P(\mathbf{r}) \quad (3.86)$$

$P(\mathbf{r})$ is the polarization field vector which describes the density of permanent or induced electric dipole moments in a dielectric material. For an isotropic medium that has linear response, the polarization field can be defined by

$$P(\mathbf{r}) = \chi E(\mathbf{r}) = -\chi \nabla \Phi(\mathbf{r}) \quad (3.87)$$

where $\chi(\mathbf{r}) = \epsilon(\mathbf{r}) - 1$ is the dielectric susceptibility of the medium. Then Eq. (3.86) can be written as

$$-\nabla \cdot \epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r}). \quad (3.88)$$

Therefore, the permittivity $\epsilon(\mathbf{r})$, which is also called dielectric constant, represents the polarizability of the medium. In biomolecular calculations, $\epsilon(\mathbf{r})$ is generally assumed as piecewise constants in most applications. It is noted that in charge neutral molecules, electric polarization corresponds to the rearrangement of electrons in molecules. In most popular force field packages, some of the polarizations of a charge neutral macromolecule are often treated as partial charges located at the centers of individual atoms. These partial charges give rise to most of the fixed charge source term ρ_f in the generalized Poisson-Boltzmann equation. Due to this treatment of the polarization effect, a relatively small $\epsilon(\mathbf{r})$ value is normally assigned to the biomolecular domain. For example, when calculating the solvation energy of proteins, $\epsilon(\mathbf{r})$ is set to 1 or 2 for the biomolecular domain while 80 for the solvent domain. These values are commonly accepted and vary in only small ranges for different purposes. However, in the application of ion channels, choices of dielectric constants in different regions of interest are worthwhile to be carefully explored.

First, although the ion permeation is a dynamical process, dielectric constants are all assumed time independent due to the fact that the electrolytic solution is a fast relaxing bath, i.e., the relaxation time of the solvent water is extremely short. Secondly, the dielectric constants are approximated as piecewise constants for computational simplicity. In the bulk concentration, a widely used dielectric constant as 80, which is the experimental measurement at room temperature for water. The value of ϵ is usually set to 1 or 2 in the protein domain, which partially accounts for the field-induced atomic polarization of the protein. However, two features about protein structures are neglected in the continuum approximation for ion channels and should be partially compensated by the dielectric constant of the channel protein. One is the re-organization of the protein and water in extremely confined channels and the other is the protein's response to ion's presence in the channel, since the ion permeation takes places at the same time scale. Therefore, in order to encapsulate these features

in a continuum model with a single dielectric coefficient, the value of $\epsilon(\mathbf{r})$ for channel proteins is suggested to be greater than 2.

There are also some issues in assigning the dielectric coefficient for the aqueous region in the ion channel. A general conclusion is that $\epsilon(\mathbf{r})$ in the bulk aqueous region should be much higher than that in the channel region. The main reason is the high confinement of the channel geometry. In ion channel pores which are usually very narrow, water molecules are highly ordered, and their motions are restricted, so are their response to external fields. Therefore, the value of $\epsilon(\mathbf{r})$ should be much smaller than 80, and can be as small as 3 for a dry channel pore. However, these extreme values do not work well in practical computations. In fact, the dielectric coefficient in the channel pore region is still taken as 80 in most existing models despite the above arguments. In the present work, $\epsilon(\mathbf{r})$ values are set to be smaller than 80, but are not too small in order to model the biological environment.

Effective mass of the proton

The choice of effective mass $m(\mathbf{r})$ of the particle in the total Hamiltonian H as in Eq. (3.22) is an important issue to be discussed. The concept of effective mass originates from the solid state physics, which describes the response of the charge carrier to the electric or magnetic fields when quantum mechanics is applied. It is defined by analogy with Newton's second law but in the quantum mechanical framework

$$m = \hbar^2 \left[\frac{d^2 E}{dk^2} \right]^{-1} \quad (3.89)$$

where E and k are the energy and the corresponding wavenumber of the particle, respectively. Generally the effective mass is chosen in the range of 0.001 or 10 times the real mass of the particle and depends on the material and the experimental condition. However, little research has been done, to our knowledge, on the choice of

the effective mass of protons in proton channels or proton experiments. In the present model, we describe protons by quantum mechanics while treat many other particles by classical mechanics and/or continuum description. Therefore, an effective mass approximation is appropriate for our model. We set effective mass $m(r)$ as a model parameter and its value is chosen from 0.01 to 1.0 time of the real proton mass.

3.3 Numerical simulations

In this section, the validity of the proposed model and related performance analysis are presented based on a specific channel protein, the Gramicidin A (GA, PDB code: 1MAG). The GA channel protein is obtained from the soil bacterial species *Bacillus brevis* and is one of the best studied molecular channels, both structurally and functionally. In a bilayer membrane, the GA is dimers and consists of two head-to-head β -helical parts. Each part of the dimer has the sequence of FOR-VAL-GLY-ALA-DLE-ALA-DVA-VAL-DVA-TRP-DLE-TRP-DLE-TRP-DLE-TRP-ETA, and forms a narrow pore of about 4Å in diameter and 25Å in length. It appears to select small monovalent cations, bind bivalent cations, while reject anions. In our approach, the GA structure is downloaded from the PDB, and the pdb file is processed by the PDB2PQR [52], in which the radii and partial charges are adopted from the CHARMM force field values [105]. The molecular surface of the GA is generated via the MSMS package [131] with water probe radius 1.3 Å and density 10. Figure 3.3 gives an illustration of the GA in a 3D display of the structure, surface and electrostatics distribution. From Fig. 3.3(a), one can see that a complete channel pore is formed after the generation of the molecular surface. Although the GA is neutral in general, its surface electrostatics is negatively distributed near the channel mouth as indicated by the red color. Furthermore, as shown in Fig. 3.3(b), the inner part of the channel pore is also intensively negatively charged. This fact indicates the

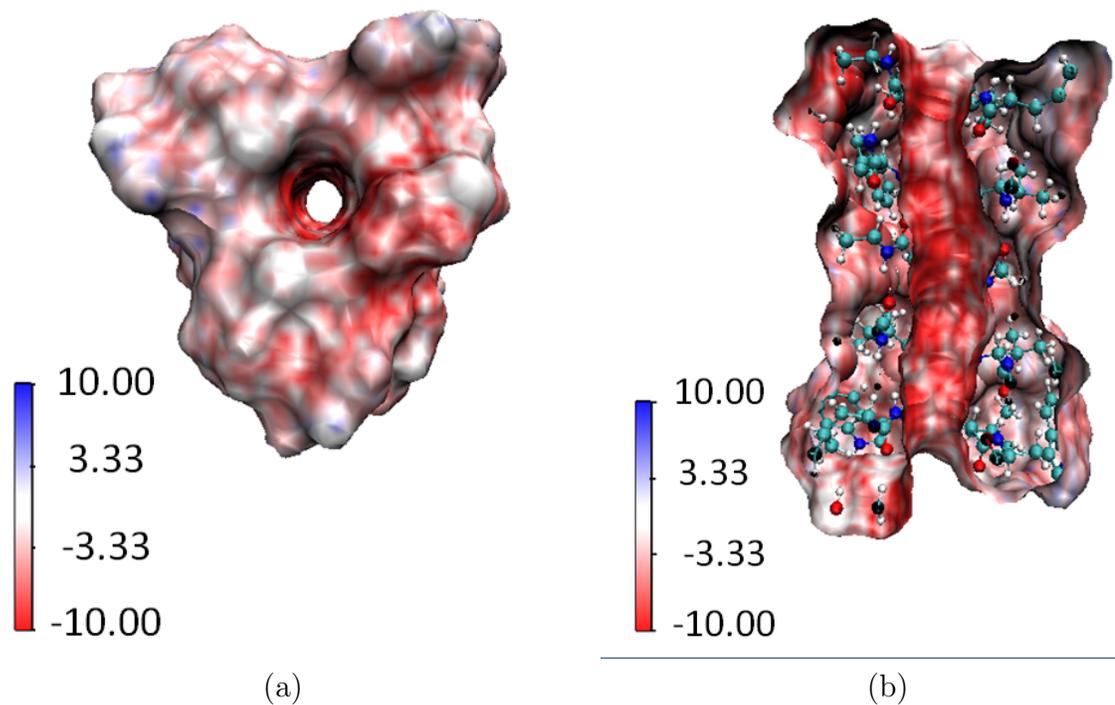
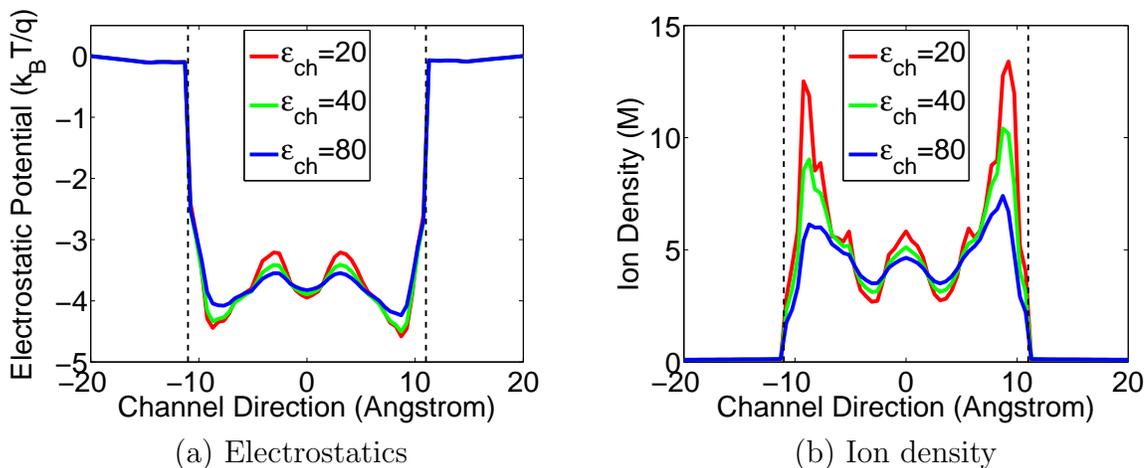


Figure 3.3: 3D illustration of the Gramicidin A (GA) channel structure and surface electrostatic potential. The negative surface electrostatics as indicated by the intensive red color on the channel upper surface and inside the channel pore implies that the GA selects positive ions. (a) Top view of the GA channel; (b) Side view of the GA channel.

selectivity of GA channel to positive ions. Having prepared the GA structure and surface, the channel pore is aligned to z -direction. The simulation grid resolution is taken as 0.5\AA . Under this discretization all the grid points are classified as either in the solvent domain or in the molecular domain. Furthermore, the molecular surface is projected on each layer along the transport direction to determine the beginning and the end of the channel respectively, by the first layer and the last layer on which closed projections can be found. An artificial membrane slab is added along the transport direction between the beginning and end of the channel, see Fig. 3.1(b).



0

Figure 3.4: Electrostatic potential and charge density of the GA channel along the z -axis obtained with $\epsilon_m = 2$ and $n_p^0 = 0.1$ molar (Red: $\epsilon_{ch} = 20$; Green: $\epsilon_{ch} = 40$; Blue: $\epsilon_{ch} = 80$). (a) Electrostatic potential profiles in channel; (b) Proton density profiles in the channel.

3.3.1 Electrostatic properties of the Gramicidin A channel

This subsection presents the electrostatic analysis of the GA channel over a wide range of $\epsilon(\mathbf{r})$ values in the present model. At the atomic level, the motion of an ion when it is passing through the channel is determined by a number of factors, such as electrostatic interactions and generalized-correlation interactions. The electrostatic interactions include the Columbic interactions between ions, and between ions and fixed charges of the channel. The generalized-correlation interactions consist of ion-ion excluded volume effects, the thermal fluctuation of the solvent, van der Waals interactions, and other short range interactions such as the frequent collisions and associations between water molecules and ions. One more factor is the structural cooperation of the channel protein during ion permeation. In the present model, the quantitative description of electrostatic interactions is the major ingredient while the degrees of freedom of generalized-correlation interactions are suppressed to reduce the computational cost.

The electrostatics of the channel system depends on the dielectric constants. In the present work, we carefully test the effect of dielectric constants within an appropriate biological range in order to obtain a reasonable prediction. It is also worth checking the dependence or changing trend of the electrostatics upon these parameters for model training and validity verification. Before the transport problem is simulated, the mathematical algorithms, choices of dielectric constants are carefully examined via the generalized Poisson-Boltzmann equation.

As discussed earlier, $\epsilon_m(\mathbf{r})$ is given as a constant in Ω_m and its value is tested over a range. However, $\epsilon_s(\mathbf{r})$ is strongly position dependent, having different values in the bulk solvent and the channel pore. For simplicity, we take $\epsilon_s(\mathbf{r})$ as piecewise constants, i.e., impose a constant value denoted as ϵ_{bath} in the bulk solvent, whereas another for the channel pore denoted as ϵ_{ch} . There is no controversy upon the choice of $\epsilon_{\text{bath}} = 80$, which is employed in all the following simulations. Figures 3.4-3.6 display the electrostatic potential profiles and (positive) ion density in GA protein with various combinations of ϵ_{ch} and ϵ_m within the range discussed in the earlier section. The reference ion density is taken as 0.1 molar.

All quantities in Figs. 3.4-3.6 are averaged on each cross section along the channel axis. The vertical dash lines in these figures indicate the entrance (left) and exit (right) of the channel. The GA protein is overall neutral in charge, but possesses a negative environment in the channel region and this fact leads to potential well. Near the entrance and the exit of the channel, there are two local potential minima (the valley near the dash line) and a major barrier in the middle of the channel. Accordingly, for the density profile, there are two peaks at the positions where two energy minima present and the density is lower in the middle of the channel. These electrostatic profiles agree with the biological properties of the GA channel.

For each fixed ϵ_m , the magnitude of the electrostatic potentials responds directly to the change of ϵ_{ch} value, as showed in Fig. 3.4(a). When the ϵ_{ch} decreases from 80,

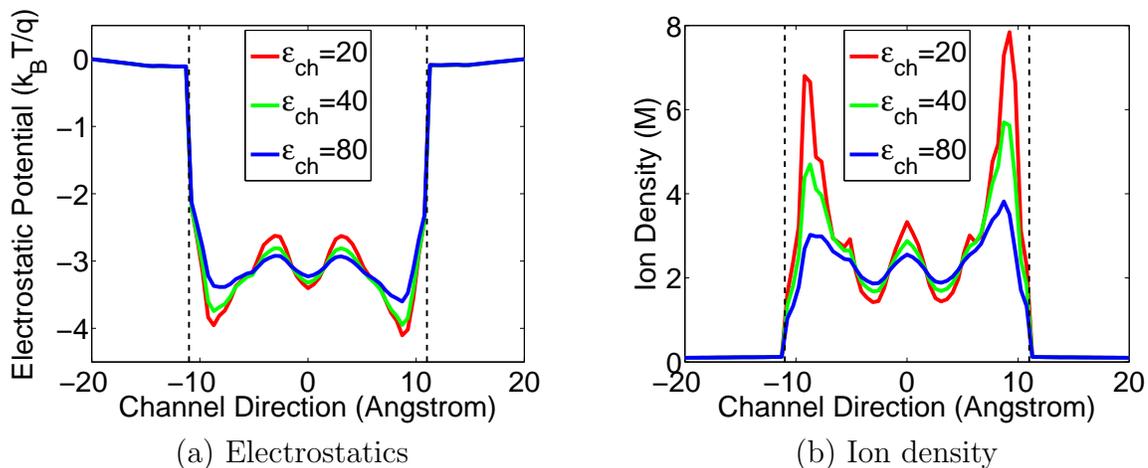


Figure 3.5: Electrostatic potential and charge density of the GA channel along the z -axis obtained with $\epsilon_m = 5$ and $n_p^0 = 0.1$ molar (Red: $\epsilon_{ch} = 20$; Green: $\epsilon_{ch} = 40$; Blue: $\epsilon_{ch} = 80$). (a) Electrostatic potential profiles in the channel; (b) Proton density profiles in the channel.

which is the commonly used value for the solvent, to the lower values suggested by biological observations, the contrast between the energy wells near the entrance/exit and the barrier in the middle becomes sharper. This phenomenon verifies the impact of ϵ_{ch} value and leads us to prefer the lower value in our model. For the ion density profile shown in Fig. 3.4(b), the changes in the peaks with respect to the changes of ϵ_{ch} are very clear. As ϵ_{ch} doubles, the magnitudes of the density at the peaks decrease half accordingly.

The impact of ϵ_m can be examined by fixing ϵ_{ch} , i.e., checking the same color curves throughout Figs. 3.4-3.6. It can be found out that changes in ϵ_m do not affect the potential structure but solely change the magnitudes. When ϵ_m increases, the absolute value of electrostatic potential decreases, and consequently the proton density becomes smaller.

Figure 3.7 depicts the electrostatics profile change with respect to reference proton densities at a certain combination of dielectric constants ($\epsilon_m = 5$ and $\epsilon_{ch} = 40$). It is easy to see that the higher the proton reference concentration, the higher the sources

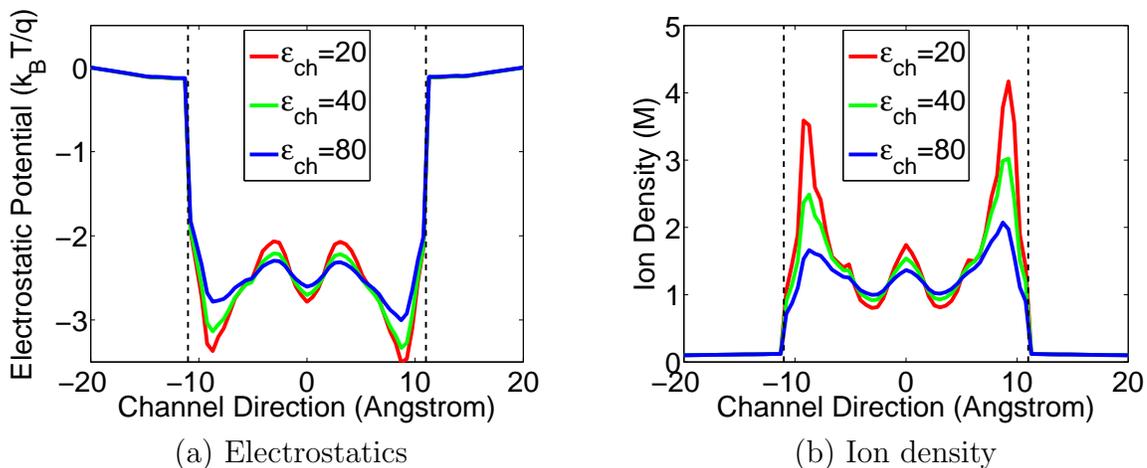


Figure 3.6: Electrostatic potential and charge density of the GA channel along the z -axis obtained with $\epsilon_m = 10$ and $n_p^0 = 0.1$ molar (Red: $\epsilon_{ch} = 20$; Green: $\epsilon_{ch} = 40$; Blue: $\epsilon_{ch} = 80$). (a) Electrostatic potential profiles in the channel; (b) Proton density profiles in the channel.

in the Poisson equation and the results in electrostatic potential profiles are.

3.3.2 Conductivity properties of the Gramicidin A channel

The mechanism of the selectivity of the GA channel can be easily explained in view of the overall potential landscape. Figure 3.8 shows the total effective potential with both the electrostatic and generalized-correlation contributions. Figure 3.8(a) is for the monovalent cations while Fig. 3.8(b) is for monovalent anions. According to the previous discussion, the generalized-correlation potential serves as an energy barrier while the GA protein provides a negatively charged environment for cations in the channel region. Two energy components with opposite signs cancel each other and result in an overall potential landscape that permeates a monovalent cation. However, the overall potential gives rise to a huge barrier for the anions since the positive generalized-correlation potential adds up with the positive electrostatic potential, as Fig. 3.8(b) shows.

Conductance reveals the efficiency of the ion channel transport of some specific

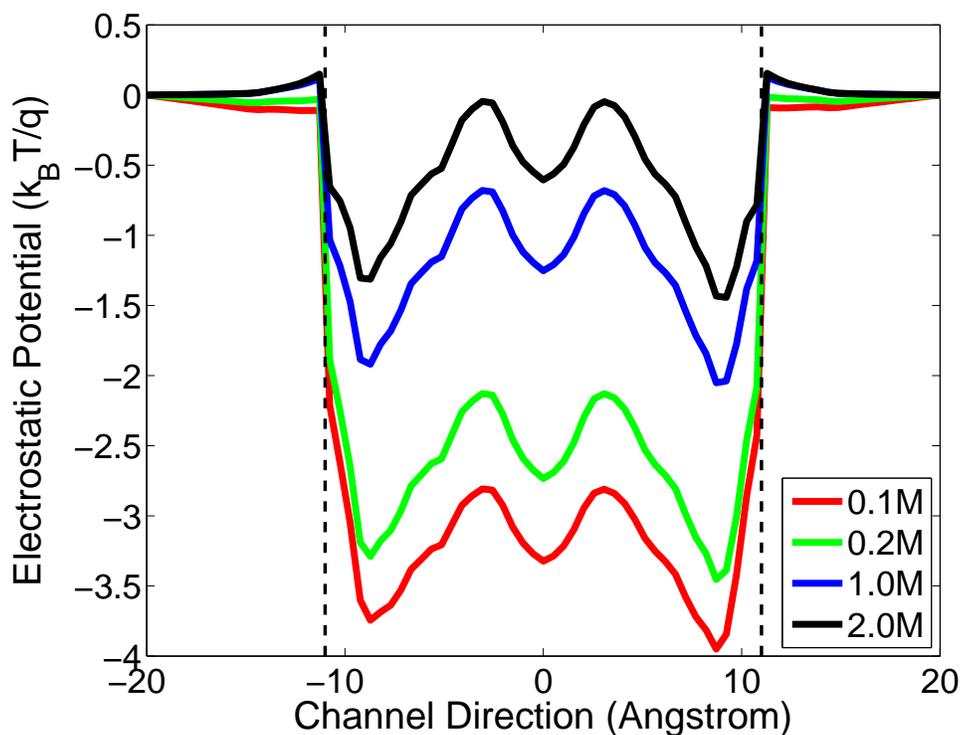


Figure 3.7: Electrostatic potential profiles of the GA channel under different ion reference densities n_p^0 . Red: $n_p^0 = 0.1$ molar; Green: $n_p^0 = 0.2$ molar; Blue: $n_p^0 = 1.0$ molar; Black: $n_p^0 = 2.0$ molar. $\epsilon_m = 5$ and $\epsilon_{ch} = 40$.

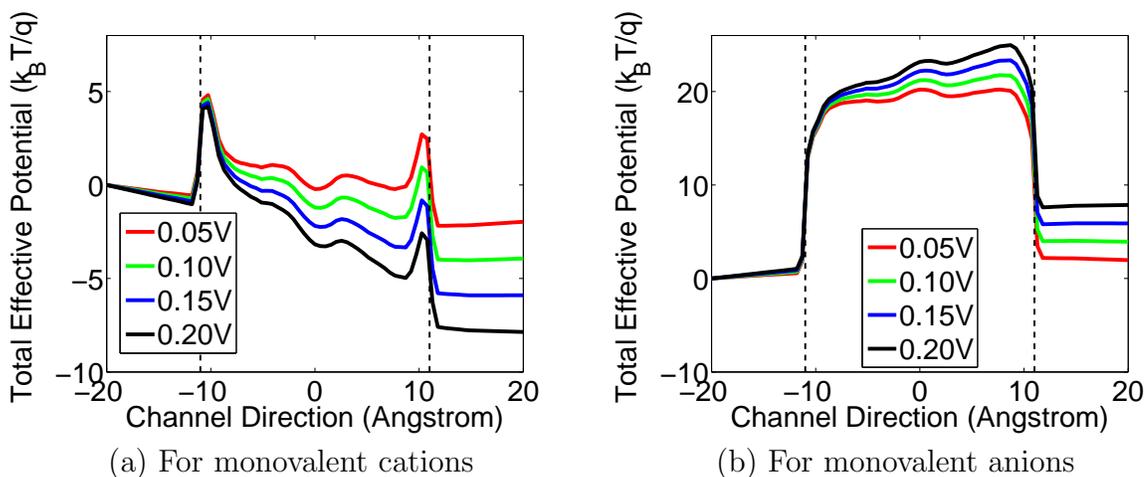


Figure 3.8: The total potential of the GA channel which includes electrostatic and generalized-correlation contributions under various voltage biases. Dielectric constants are $\epsilon_m = 5$ and $\epsilon_{ch} = 30$. The pH value of the solution is 2.75. (a) Total potential of monovalent cations; (b) Total potential of monovalent anions.

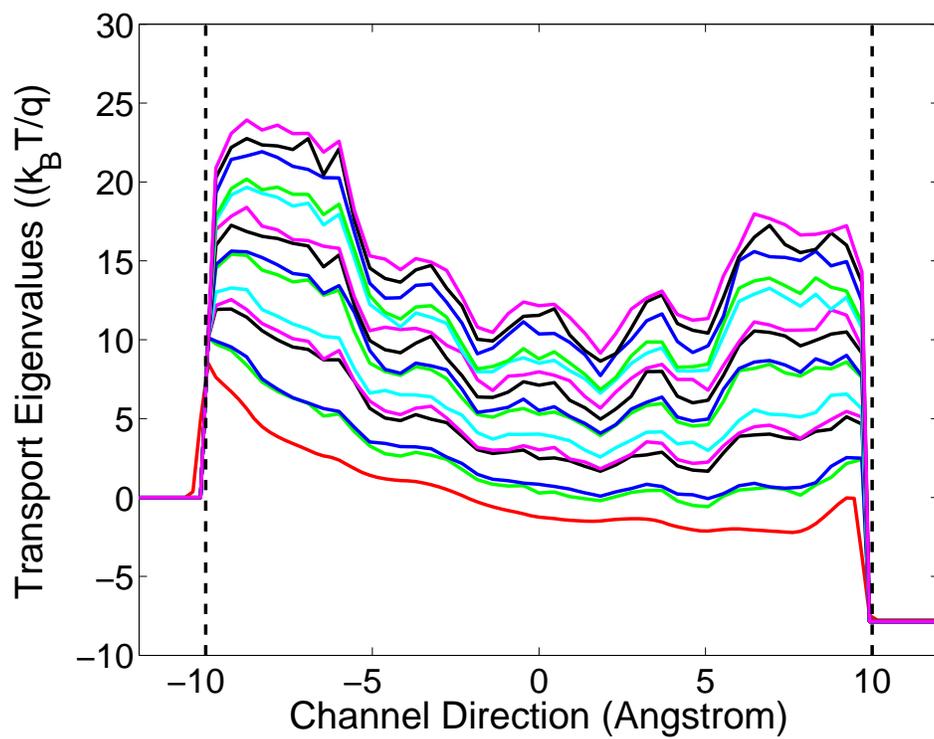


Figure 3.9: The first 15 eigenvalues (the $U^j(z)$ in Eq. 3.41) of the effective potentials along the transport direction used in the transport calculations at the voltage bias of 0.2V.

ions. Due to the fast development of experimental technologies in the past several decades, the single-channel conductance can be measured and becomes one of the prevalent descriptor of the channel function. The simulation of channel conductance mainly focuses on calculating the channel current within the physiological ranges of membrane potentials (i.e., $-0.2\text{V} < V < 0.2\text{V}$) and bath concentrations (up to molar). The channel conducting current is measured at the scale of pico-Ampere (pA) for ion channels. The corresponding characteristics of channel conductance is observed at the scale of pico-Siemens (pS) and is recorded in the voltage-current (I-V) curves and concentration-current (C-C) curves. Based on experimental observations, the I-V curves are expected to be in linear or sub-linear form while the C-C curves are supposed to exhibit saturation behavior, i.e., when the concentration increases, the conductance increases linearly at beginning and then becomes saturated later on.

The conductivity of the proton channel mainly depends on the proton scattering process. Thus we first present the effective potential profile along the transport direction. Figure 3.9 depicts the first 15 effective potential eigenvalues (i.e., $U^j(z)$ in Eq. (3.41)) used in the current calculation under the voltage bias of 0.2V . Similarly, the channel region is presented between two black dash lines. The channel region is essentially confined by the protein surface and a tube-like pore is formed. As displayed in Fig. 3.9, the potential energy profile in the channel pore region has discrete eigenstates, due to the small area confinement at each cross section and the light mass of the proton. For each specific location along the transport direction, the discrete ascending energies correspond to the eigenvalues of the operator in Eq. (3.40). In theory, the total number of the eigenvalues is infinite, but is finite in practical computations, and depends on the discretization of the cross section. In principle, all the eigenvalues should be accounted in computations. However, numerically, due to the Boltzmann distribution, higher energy components contribute little in the total transport quantity. In practise, only a few low lying eigenvalues need to be included

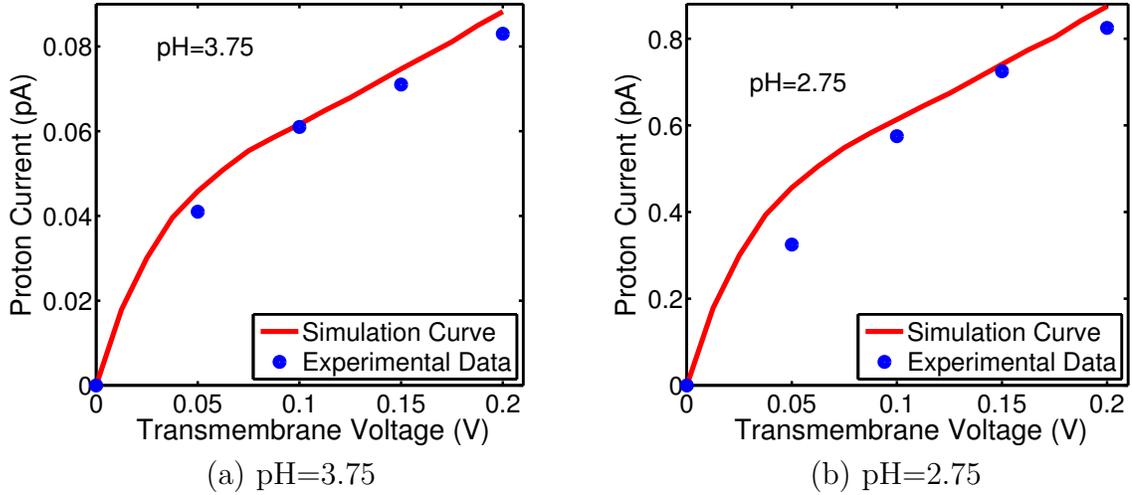


Figure 3.10: Voltage-current relation of proton translocation of GA at different concentrations. Blue dots: experimental data of Eisenman et al [59]; Red curve: model prediction.

in numerical simulations. In our case, the first 15 eigenstates are sufficient to obtain a good degree of convergence in calculating the proton density and current.

Figure 3.10 illustrates the simulation results of the present multiscale model for proton transport, compared with the experimental data from the literature [135] for the GA channel. The blue dots in each figure represent the available experimental observations for certain voltage biases while the red curves are our model predictions calculated with sufficiently many voltage samples. The model parameters are chosen to match the experimental data but all of the choices are taken within the range of physical measurements. The dielectric coefficients are taken as $\epsilon_m = 5$, $\epsilon_{ch} = 30$ and $\epsilon_{bath} = 80$, according to the discussion in previous sections. To determine the generalized correlations, the diffusion coefficients of protons are taken as 3.6×10^{-9} m²/s in the channel, less than a half of the value in the bulk environment, and the relative weighting parameter is set to $\alpha = 0.03$. Taking into account above considerations, we can conclude that experimental data and the present predictions agree quite well and this agreement verifies the validity of our quantum dynamics in

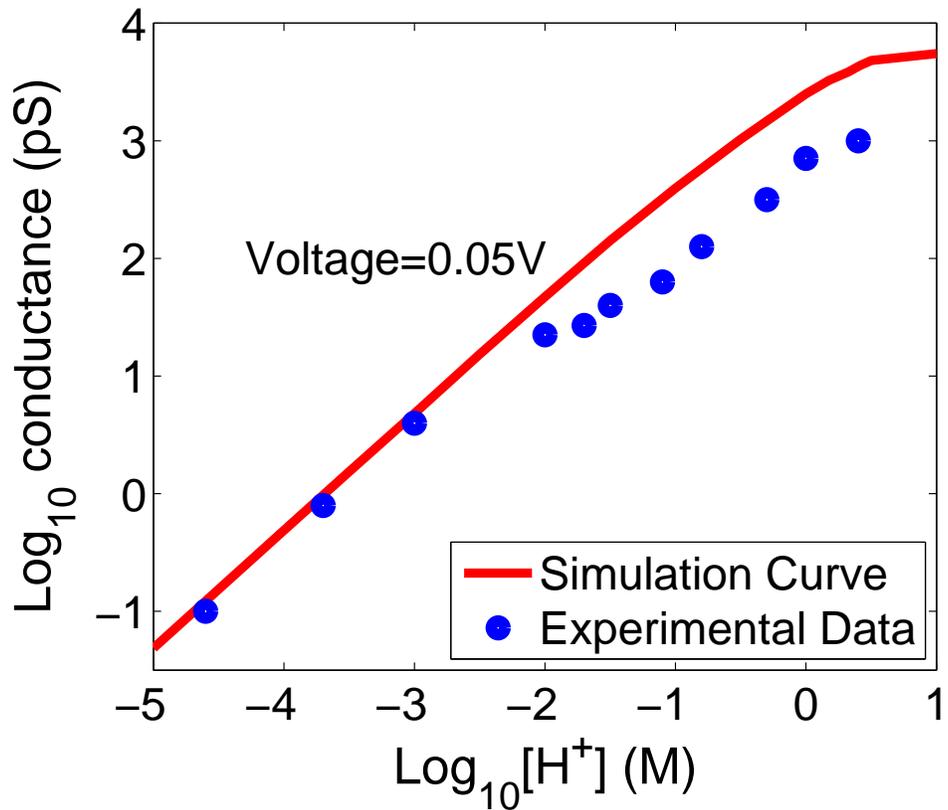


Figure 3.11: Conductance-concentration relation of proton translocation at a fixed voltage. Voltage bias=0.05V; Blue dots: experimental data of Eisenman et al[59]; Red curve: model prediction.

continuum model.

Apart from I-V curves, there are also experimental data available about the conductance-concentration relation (C-C curve) of the proton transport under given voltages. Figure 3.11 displays such a relationship with a comparison between experimental data and model predictions. At a given voltage bias, the conductance of the channel is calculated with various proton concentrations as indicated by the horizontal axis. Using the same set of parameters as those in Fig. 3.10, the computed conductance-concentration relation also agrees fairly well with experimental data. At lower proton concentrations (i.e., pH value being greater than 2), the agreement between our prediction and experimental data is quite good. At relatively higher concentrations, although the numerical simulations slightly overestimate the observed conductance, the conductance saturation against the concentration can still be observed in simulations and it corresponds to the sub-linear characteristics or the flat tail of the C-C curve.

The experimental data used in this work are reported by Eisenman et al[59] and are also employed to verify another proton transport model by Schumaker et al. [135]. There are other experimental data on proton conductance available [5, 42, 34] but under different experimental conditions. First, the experimental data provided by Cukierman et al. [42] offer proton conductions recorded with natural Gramicidin A and with its Dioxolane-Linked dimer in different lipid bi-layers (phosphatidylethanolamine-phosphatidylcholine, or PEPC and glycerymonooleate, or GMO). Their experimental studies were carried out for low (9.8 mM) and high (1578 mM) proton concentrations against the transmembrane voltages. Additionally, in another piece of work [34], the attenuation of proton transfer in Gramicidin water wires by phosphoethanolamine was investigated and a number of I-V curves were provided. It is impossible to fit all the experimental data by a single group of parameters because of the difference in experimental conditions and lipid membrane types.

Nevertheless, it can be observed that our simulation curves under the current set of parameters have shown similar qualitative shapes. Therefore, the present model can fit to these experimental data by slightly adjusting model parameters to reflect the different experimental conditions. Finally, Akeson and Deamer [5] also reported I-V curves of proton conductance for the F_1F_0 ATPases studies. In their results, a severe saturating or sublinear character is found for proton concentration of 10 mM and there were an obvious superlinear pattern for 1.0 M hydrogen chloride (HCl). Our model can not capture these characteristics by just tuning the parameter values. In fact, this set of experimental data was also found difficult for another theoretical model of proton transport [135].

3.3.3 Model limitations

Based on the multiscale approximation, the present model captures the most important factors which have large impacts on the proton permeation. Meanwhile, the quantum treatment of protons provides a potential analysis tool to account for the quantum behavior in proton channel transport and proton translocation in biomolecules. The setup of the present model roots from essential biophysical principles with reasonable approximations, and thus the numerical simulations give considerably good agreements with experimental data under appropriate choices of model parameters. However, this model also has a number of limitations, which are to be studied further in the future. First, in this model, the channel protein is assumed to be rigid, i.e. it does not response to the permeation of ions. This is not true in real situation and the configurational change of the channel protein has been found to have fairly important impact on the ion permeation process. Although the omitted ion-protein interaction has been somehow compensated implicitly by adjusting the dielectric constants, this interaction can not be fully accounted unless more sophisticated models, such as the multiscale molecular dynamics [154], are invoked.

Additionally, the plasma membrane where the channel protein is embedded is simplified. There are various types of membranes, some of them have dipoles and others have charges. In our model, the membrane is just approximated by the uniformly distributed dielectric medium and the charges or dipole effects are neglected. However, there is no essential difficulty to improve this aspect in our model. Point charges from membranes can be added in the present model. Otherwise, a position dependent dielectric constants for the biomolecular region can also represent the charge effects in the membrane. Finally, the other limitation of the present model is the simplified local density approximation of generalized correlations, which reduces the number of the degrees of freedom, although. Compared to the electrostatic potential, the generalized-correlation potential plays a less important role in general. However, it may be of crucial importance for channel selectivity in certain situations. Therefore, an emergent task of our future work is to come up with more quantitative modeling of generalized-correlation interactions meanwhile without significantly increasing the number of degrees of freedom. Local spin density approximation, local density gradient approximation and general linear scaling approaches are under our consideration.

3.4 Conclusion remarks

Proton dynamics and transport across membrane proteins are of paramount importance to the normal function of living cells. Although there are a variety of excellent theoretical models and efficient computational methods for ion channels in general, most commonly used models are much less successful when they are applied to the proton channel transport due to the unique characteristics of protons. It is commonly believed that to a certain extent, proton channels demonstrate quantum mechanical properties such as the translocation as shown in the Grotthuss-type mechanism [112, 2]. However, the exact role of quantum mechanics in the atomic mechanism

of proton channels is still unclear despite of a number of elegant theories in the literature, partly due to the complexity of ion channel systems. The present paper introduces a quantum dynamics in continuum (QDC) model for the prediction and analysis of proton density distribution and conductance in proton channels. Our essential ideas are as follows. First, protons behave quantum mechanically due to their light masses and channel geometric confinement in protein channels. Therefore, a quantum mechanical treatment of protons is necessary. Additionally, since the primary interactions in proton channels are of ion-ion electrostatic type and the van der Waals type of interactions involve less energy, a dielectric continuum treatment of solvent medium may provide a reasonable approximation to the effect of numerous solvent molecules. Most importantly, this treatment dramatically reduces the dimensionality of the problem. As such, our approach is called a QDC model. Moreover, since the atomic detail of the protein structure serves as a physical boundary for proton dynamics and transport, the present model returns molecular surface to separate the continuum solvent domain from the discrete charge domain of the protein. Finally, densities of all other ions and counterions in the solvent are described by the Boltzmann distribution, which is a quasi-equilibrium description as the electrostatic potential varies during the process of protons permeating the membrane.

We propose a multiscale variational paradigm to accommodate the aforementioned aspects in a unified framework. The total free energy functional encompasses the kinetic and potential energies of protons, and the electrostatic energies of ions and fixed charges in the channel system. The first variation is carried out via the Euler-Lagrange equation to derive the governing equations for the system. A generalized Poisson-Boltzmann equation is obtained for the electrostatic potential while a generalized Kohn-Sham equation is resulted for the state of protons in the system. The solution to these two coupled nonlinear equations leads to the desirable electrostatic distribution and proton density profile in the channel system. Expressions for

proton density and proton flux across the membrane are derived from fundamental principles.

The computation of the proposed coupled equations involves a number of mathematical issues, such as the linearization of coupled nonlinear partial differential equations (PDEs) using the Gummel iterations and/or inexact Newton iterations, and the solution of elliptic PDEs with discontinuous coefficients (i.e., piecewise dielectric constants), singular sources (i.e., Dirac delta functions for protein charges), and nonsmooth interfaces (i.e., geometric singularities). In the present work, we utilize the Dirichlet to Neumann mapping method to take care of singular charges, and the matched interface and boundary (MIB) method to accurately handle the discontinuous coefficients and geometric singularities.

The Gramicidin A (GA) channel protein, a popular protein structure, is employed in our numerical studies to demonstrate the performance of the proposed QDC model. We give a detailed discussion about the rationale for model parameter selections. The electrostatic property of the GA channel is analyzed with the proposed model against a large number of model parameters. Proton transport properties, i.e., the current voltage (I-V) curves, are investigated over a large number of combinations of applied voltages and reference bulk concentrations. Our model predictions are compared with experimental data, which validates the present QDC model. Finally, we provide detailed discussion of model limitations and possible future improvements.

Chapter 4

Structure and electrostatic analysis for bio-molecules

In both simulations of nano-electronic transistors and proton channels, the highly accurate and efficient calculation of electrostatics is of paramount importance. The electrostatic model in nano-transistors is relatively straightforward because of the simple system components. However, the biological ion channel system requires careful treatment due to the complicated protein structures and inhomogeneous environments. As reviewed in the introduction, the implicit solvent theory is applied to model and analyze the electrostatics for bio-molecules, and the MIBPB solver serves as a powerful tool to solve the generalized Poisson-Boltzmann equation that provides the electrostatics background for proton channel. This chapter is devoted to the overview of the Poisson-Boltzmann equation, detailed description of the efficient analysis of the MIBPB solver and its various applications.

4.1 The Krylov subspace theory (KSP) and preconditioner based MIBPB solver

In this section, we first give the overview of the Poisson-Boltzmann equation and related numerical challenges, then the MIBPB solver and its efficiency analysis are demonstrated. This fills the lacked computational detail in the previous chapter because the GPB equation is the natural extension of the Poisson-Boltzmann equation.

4.1.1 Review of the Poisson-Boltzmann equation

In the implicit solvent model, the solvent is treated as a continuous medium while the description for solute is kept at the atomic level. The electrostatic potential Φ of a solvent-solute system can be determined by the Poisson-Boltzmann equation(PBE) in a regular domain Ω whose dimension usually has the order from 10\AA^3 to 500\AA^3 for biomolecular applications. Figure 4.1 gives the sketch of the protein-solute system and the computational domain. The protein region and the solvent region are denoted as Ω_1 and Ω_2 , respectively. Naturally the whole computational domain is $\Omega = \Omega_1 \cup \Omega_2$, and the molecular surface is labeled as Γ . For simplicity, the ion-exclusive layer is ignored in the present model. Although mobile ions in the solvent are explicitly indicated in the figure, the whole solvent region is actually modeled by an implicit continuum.

Under these assumptions, if one consider the 1:1 electrolyte for simplicity, the Poisson-Boltzmann equation reads:

$$-\nabla \cdot (\epsilon(\mathbf{r})\nabla\Phi(\mathbf{r})) + \bar{\kappa}^2(\mathbf{r}) \left(\frac{k_B T}{q} \right) \sinh \left(\frac{q\Phi(\mathbf{r})}{k_B T} \right) = 4\pi \sum_{i=1}^{N_a} Q_i \delta(\mathbf{r} - \mathbf{r}_i) \quad (4.1)$$

boundary condition of Eq.(4.1) depends on various applications. The q is the elementary charge, k_B is the Boltzmann constant and T is the temperature, ϵ and $\bar{\kappa}$ are

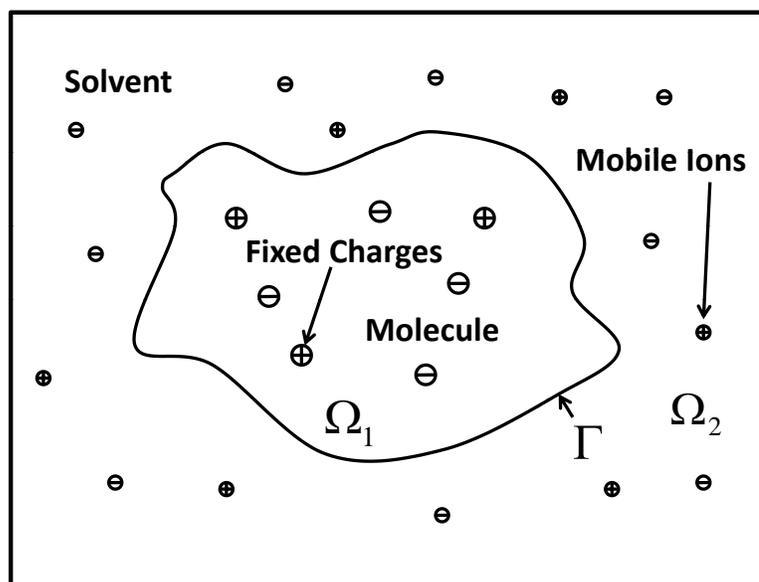


Figure 4.1: The implicit protein-solvent system

Table 4.1: Physical units notations

Abbr.	Meaning	Represents	Equivalent expressions
<i>esu</i>	statcoulomb	charge	fundamental unit
<i>c</i>	coulomb	charge	fundamental unit
<i>cm/m</i>	centermeter/meter	distance	fundamental unit
<i>s</i>	second	time	fundamental unit
K	Kelvin	Temperature	fundamental unit
<i>mol</i>	mole	Quantity	fundamental unit
<i>dyn</i>	dyne	force	esu^2/cm^2
<i>erg</i>	erg	energy	$dyn \cdot cm$
<i>l</i>	liter	volume	$1000cm^3$
molar	mole per liter	concentration	mol/l

Table 4.2: Some physical constants

Label	Name	Values in CGS unit	Values in SI unit
N_a	Avagadro's number	6.022045×10^{23} (no unit)	-
k_B	Boltzmann's constant	$1.380662 \times 10^{-16} erg/K$	$8.617343 \times 10^{-5} eV/K$
q	Fundamental charge	$4.803204 \times 10^{-10} esu$	$1.602176 \times 10^{-19} c$
h	Planck's constant	$6.626068 \times 10^{-27} erg \cdot s$	$4.135667 \times 10^{-15} eV \cdot s$
\AA	Angstrom	$10^{-8} cm$	$10^{-10} m$

dielectric constant and modified Debye-Hückel screening function describing the ion strength, respectively. Here Q_i is the fixed charge in the protein and \mathbf{r}_i denotes the position of the fixed charge, and N_a is the total number of fractional charges.

Eq.(4.1) is presented in the Gauss unit system, in which the units and physical constants are provided in Table 4.1 and 4.2. Usually the electrostatics $\Phi(\mathbf{r})$ is scaled to be dimensionless by $u = q\Phi/k_B T$, consequently, the Eq.(4.1) is revised as

$$-\nabla \cdot (\epsilon(\mathbf{r})\nabla u(\mathbf{r})) + \bar{\kappa}^2(\mathbf{r}) \sinh(u(\mathbf{r})) = C \sum_{i=1}^{N_a} z_i \delta(\mathbf{r} - \mathbf{r}_i) \quad (4.2)$$

The constant $C = 4\pi q^2/k_B T$ results from dimensionless procedure, and $z_i = Q_i/q$ is the charge fraction of the fixed charge.

The hyperbolic term $\sinh(u(\mathbf{r}))$ takes into account the salt effect with the Boltzmann distribution theory in the equilibrium state. Therefore, Eq. (4.2) is a nonlinear partial differential equation (PDE) of elliptic type. Such a nonlinear term can be linearized under the weak potential approximation, i.e, when $u(\mathbf{r}) \ll 1$, $\sinh(u(\mathbf{r})) \sim u(\mathbf{r})$. Thus the linear approximation of Eq. (4.2) is

$$-\nabla \cdot (\epsilon(\mathbf{r})\nabla u(\mathbf{x})) + \bar{\kappa}^2(\mathbf{r})u(\mathbf{r}) = C \sum_{i=1}^{N_a} z_i \delta(\mathbf{r} - \mathbf{r}_i) \quad (4.3)$$

Typically, for biomolecular systems of given ranges of temperature and ionic strength, the PBE is solved with the following coefficient bounds [78]

$$1 \leq \epsilon(\mathbf{r}) \leq 80$$

$$0 \leq \bar{\kappa}^2 \leq 127$$

$$5249 \leq C \leq 10500$$

$$-1 \leq z_i \leq 1.$$

The spatial-dependent coefficients $\epsilon(x)$ and $\bar{\kappa}(\mathbf{x})$ are discontinuous across the molecular surface. It is a challenge to solve such an elliptic equation with high accuracy because the regularity of its solution is reduced due to the interface and geometric singularity. For this class of problems, numerical accuracy and convergence rate are typically low without special interface treatments. Another challenge is the singular source term which contains many Delta functions, which are infinity at their spatial locations. Accurate approximation to the point-supported singular functions is an important topic in computational mathematics. The above two difficulties hinder the accurate numerical solution to the PB equation. To maintain a given accuracy, the grid spacing of the discretization has to be sufficiently small because of the low regularity of the solution. On the other hand, a small grid spacing implies millions of variables even for a middle-size protein. For example, the cube embedding a 2800-atom protein may have a dimension of $50 \times 50 \times 50(\text{\AA}^3)$, which leads to 1×10^6 variables if the resolution is 0.5\AA . This gives rise to a major obstacle for PB applications, especially for the calculation of thermodynamic properties via either the molecular dynamics or pre-equilibrium approaches. The MIB method is introduced and applied to the solution of the PBE. The source term singularity is removed by the DNM. In Ref. [67], the solution $u(\mathbf{r})$ of the PBE is decomposed into a singular part and a regular part. The singular part of the solution comes from singular delta functions, and is obtained analytically as the Green's function. As a consequence, this separation generates an extra Neumann jump condition at the interface for the regular part. Therefore, after the separation, one only needs to solve the remaining homogeneous Poisson-Boltzmann equation subject to corresponding Neumann jump conditions at the interface. This procedure is also called Dirichlet to Neumann mapping. Consequently, truly second-order accurate solution to the PBE with molecular surfaces and singular charges can be obtained with a relatively large grid spacing [67].

4.1.2 KSP based and preconditioner accelerated MIBPB solver

The discretization of the nonlinear PB equation results in a nonlinear equation system of the form

$$F(U_h) = L_h U_h + N(U_h) - f_h = 0, \quad (4.4)$$

where h is the discretization resolution, L_h and f_h represent the matrix and right hand side generated via the MIB and DNМ schemes, $U_h = [u_1, u_2, \dots, u_i, \dots]^T$ is the solution vector. The nonlinear term $N(\cdot)$ is diagonal and $N(U_h) = [N_i(u_i)] = [\bar{\kappa}^2 \sinh(u_i)]$. The inexact-Newton method is perhaps one of the most efficient ways to solve nonlinear system (4.4). If the h is dropped, it reads

$$F'(U_n)V_n = -F(U_n) + R_n, \quad \frac{\|R_n\|}{\|F(U_n)\|} \leq \eta_n \quad (4.5)$$

$$U_{n+1} = U_n + V_n, \quad (4.6)$$

where n is the iteration step, F' is the Jacobian matrix $[\partial F_i(U)/\partial u_j]$ and takes the form $F'(U) = L + N'(U)$. Here N' is the Jacobian matrix of $N(U)$ and is also diagonal $N'_i(U) = N'_i(u_i) = \bar{\kappa}^2 \cosh(u_i)$. It is easy to see that the inexact-Newton method is a two-layer iterative algorithm. The correction term V_n in outer iteration (4.6) is considered as a rough solution of inner iteration (4.5). The scheme converges linearly when η_n , the ratio of the residual R_n between the function value $F(U_n)$, is less than 1, and converges super-linearly as the sequence η_n has the property that $\lim_{n \rightarrow \infty} \eta_n = 0$.

The overall numerical efficiency of solving the nonlinear system strong depends on the efficiency of solving the linear system (4.5), which in further depends on the complexity of the matrix L_h . It is worth pointing out that in standard FDMs, the matrix L_h only depends on the grid resolution L and the dielectric constants. However,

in the MIBPB scheme, the structure of L_h also depends on the molecular surface of a specific protein. Due to this reason, we also call L_h the matrix of a protein for simplicity. The MIB and DNM successfully overcome the equation singularities and promise a high accuracy convergence order by taking into account all the local interface information. However, as a trade-off, the structure of matrix L_h is much more complicated than that from standard FDMs. Specifically, the matrix loses symmetry and may not be positive-definite any more. The lose of these properties will lead to more computational time and memory. Therefore, the selection of appropriate linear solvers becomes subtle when computational efficiency is sought as well.

The review of several basic linear solvers are summarized in Appendix B. However, the matrices from the MIB and the DNM scheme can barely take any advantage from the described methods due to their complicated structures. Therefore, we put our emphasis on choosing other methods and accelerating techniques. In Appendix B, we also include a brief description of Krylov subspace (KSP) techniques. Based on the KSP theory, proper linear solvers and acceleration techniques (preconditioners) are chosen and compared in this section for the numerical efficiency of MIBPB linear systems. Two KSP solvers, the stabilized biconjugate gradient method (BiCG) and the generalized minimal residual method (GMRES), are potentially effective iterative solvers for matrices with general structures. Several preconditioning strategies, the Jacobi preconditioner (JAC), the blocked Jacobi preconditioner (BJAC) and the incomplete LU factorization preconditioner (ILU) are available to incorporate with the two solvers to accelerate the solution of the linear system.

Matrices generated from a set of proteins are employed to test the performance of various KSP solver-preconditioner combinations. For each matrix, the condition number, linear system iteration number and iteration time are used to characterize numerical efficiency. All these measurements of matrices are analyzed numerically by the PETSc (<http://www.mcs.anl.gov/petsc/petsc-as/>). The grid resolution is

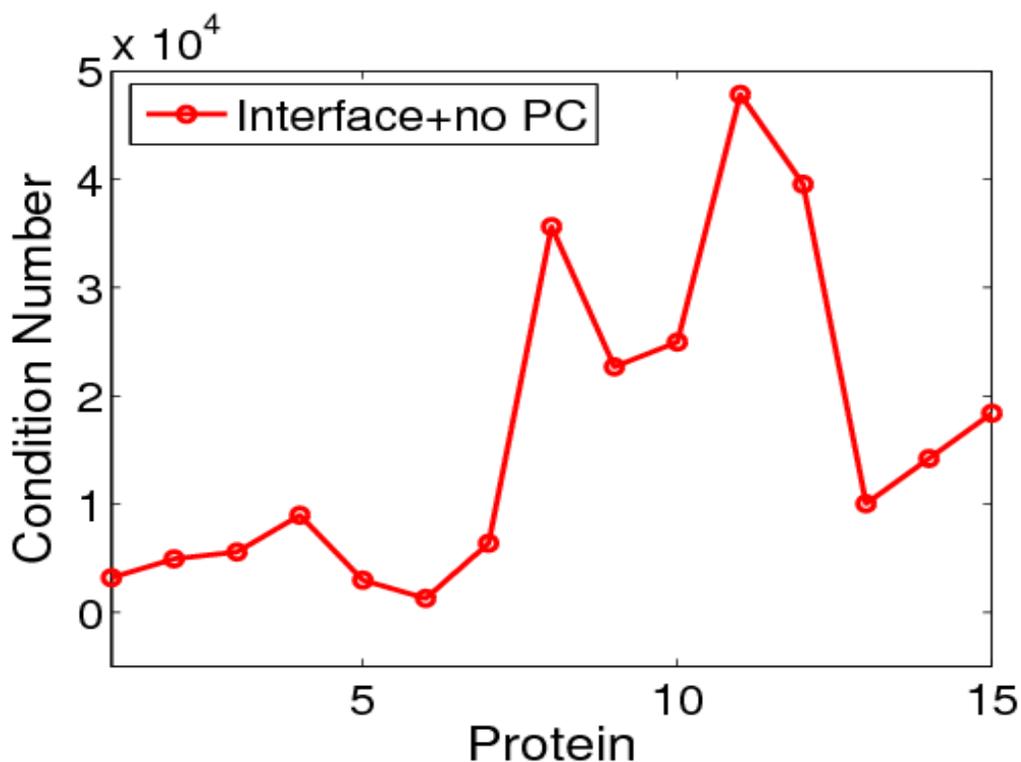


Figure 4.2: Condition numbers over 15 proteins (PDB IDs from protein 1 to protein 15: 1ajj, 1vii, 1cbn, 1bbl, 1fca, 1sh1, 1vjw, 1fxd, 1bpi, 1a2s, 1frd, 1svr, 1a63, 2erl, 2pde). (a) Condition numbers for unpreconditioned (unPCed) MIBPB matrices;

taken as 1.0\AA in the following tests unless otherwise specified. The stopping criterion of all KSP solvers are taken as 1×10^{-6} in order to get more accurate solutions while in practical biological applications the criterion can be relaxed to 1×10^{-3} to save CPU time but satisfactory results are also achieved.

First of all, the matrix condition numbers are examined. The condition number can predict the level of difficulty in solving the linear system before it is really solved. The magnitude of the condition number of a matrix generated via the MIB and DNM scheme depends on the size and structure of a biomolecule. More specifically, under the same grid resolution, a molecule which has a larger number of atoms needs a larger computational domain and a larger matrix size. Meanwhile, a molecule which has a more complex surface geometry leads to more interface conditions and a larger

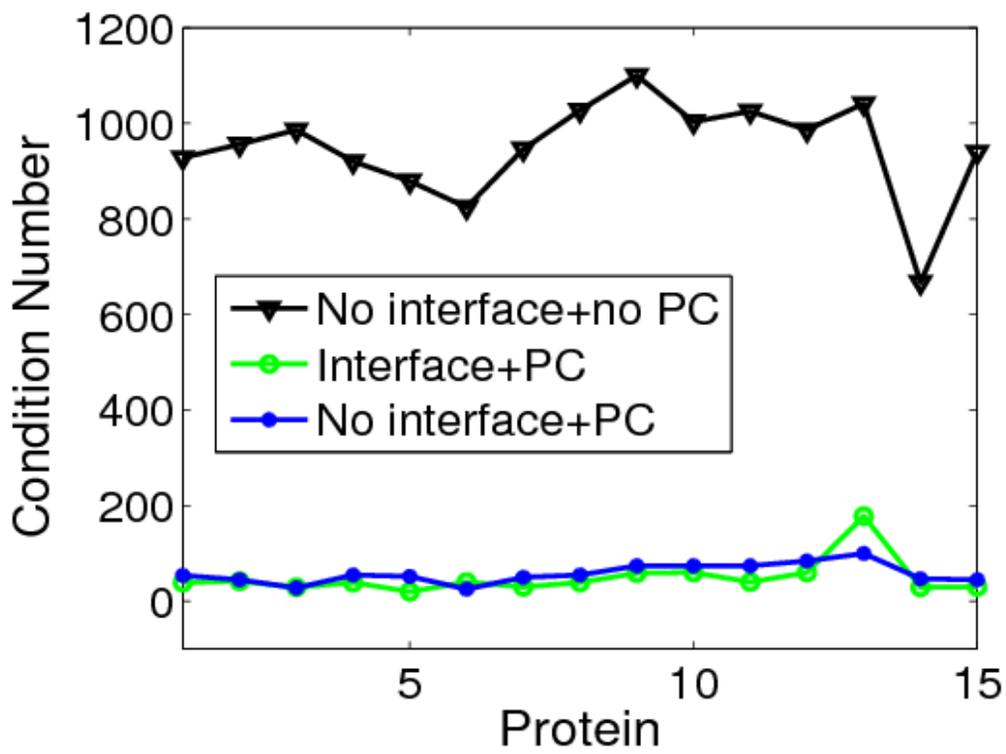


Figure 4.3: Condition numbers over 15 proteins (PDB IDs from protein 1 to protein 15: 1ajj, 1vii, 1cbn, 1bbl, 1fca, 1sh1, 1vjw, 1fxd, 1bpi, 1a2s, 1frd, 1svr, 1a63, 2erl, 2pde). (b) Comparisons of condition numbers under three settings.

matrix size. Both cases contribute to higher condition numbers. Therefore, the size and complexity of a biomolecule usually affect the numerical efficiency of the MIBPB solver.

Figure 4.2 presents condition numbers of matrices corresponding to 15 protein structures and indicates the numerical difficulties of solution without proper acceleration techniques. The horizontal axis lists proteins. Protein data bank (PDB) identification numbers (IDs) are listed in the figure. The numbers of atoms of these proteins range from 500 to 2000. Discretizing the PBE with the MIB scheme, without any preconditioner (PC) applied, the condition numbers are usually in the order of 10^4 , about one order larger than those of the matrices generated from the standard FD discretization, i.e., without the interface treatment. This is expected because the use of the molecular surface as the interface and all included local information around the interface, the MIBPB matrices do not maintain the symmetry and are not positive definite. The MIB matrices generally have larger condition numbers and require more CPU time [163, 157, 67].

By using of preconditioners (PCs), the magnitudes of condition numbers of MIBPB matrices are significantly reduced to less than one hundred, as shown in the circle plot of Fig. 4.3. The triangle plot in Fig. 4.3 gives the condition number magnitudes of the matrices from standard FDMs without PC, revealing the huge differences among different treatments. The circle and dot plots are condition number magnitudes for matrices with PC, from both the MIB scheme and the FDM, respectively. Interestingly, it can be concluded that the condition number magnitudes of two schemes are reduced to almost the same level by using the PCs. We can safely say that the difficulty of solving the linear system generated from the interface based MIBPB scheme is actually comparable to that from standard FD discretization. Under almost the same numerical efficiency, MIB scheme and DNM are able to obtain higher accuracy because all the local geometry information of the molecular surface has been taken

Table 4.3: Iteration numbers and CPU time for the discretization matrices of proteins

Proteins		Preconditioned iteration			Un-Preconditioned iteration		
ID	atoms	number	time	condition #	number	time	condition #
1mbg	903	19	0.3	40	5404	54	118900
1r69	997	18	0.3	40	5400	58	250400
1bor	832	20	0.3	30	2152	23	138850
1vii	596	17	0.2	42	3963	28	4963
1fxd	824	19	0.3	39	7084	80	35637
2erl	573	17	0.2	29	4858	36	14223
1a2s	1272	23	0.6	61	10000	156	24981

into account.

Quantitatively, for a specific KSP solver such as GMRES, the iteration numbers and computing time of linear systems for 7 proteins are listed in Table 4.3. It is well-known that condition numbers can only be mathematically estimated for large matrices, then the listed condition numbers calculated by PETSc solvers may not be exact. Despite this fact, we can still have a sense from the numbers how the PC significantly reduces the difficulties of solving the linear systems.

As stated earlier, two KSP iterative methods, the stabilized BiCG and the GMRES, are associated with three types of preconditioners, JAC, BJAC and ILU. Table 4.4 compares the effect of combinations of these KSP solvers and preconditioners. For different preconditioning strategies, since the ways of counting iteration numbers are different, only the iteration time for each combination is listed in the table. Sample proteins of various sizes are presented in this table, from small size (less than 1000 atoms), middle size (1000-3000 atoms) to large size (around 8000 atoms). It can be concluded that the GMRES performs slightly better than the stabilized BiCG does for small-sized proteins but stabilized BiCG takes the lead in middle- and big-sized proteins. Among the three kinds of preconditioners, the BJAC and the ILU almost have the same effects and are slightly better than the JAC. Therefore, the combination of stabilized BiCG and BJAC is recommended and set as the default option in

Table 4.4: Iteration time for different combinations of KSP solvers and preconditioners

Protein		BiCG			GMRES		
ID	atom	BJAC	ILU	JAC	BJAC	ILU	JAC
1ajj	519	0.24	0.24	0.26	0.23	0.23	0.38
1vjw	828	0.29	0.29	0.35	0.26	0.26	0.44
1a2s	1272	0.56	0.56	0.51	0.57	0.56	0.84
1a7m	2809	1.69	1.67	1.77	1.93	1.91	2.85
1f6w	8243	3.90	3.88	4.48	4.70	4.65	7.19

Table 4.5: Convergence test of MIBPB solver with a set of proteins

Proteins ID	Error		Order	Error		Order
	h=1.0Å	h=0.5Å		h=0.25Å	h=0.25Å	
1ajj	6.52E-2	1.13E-2	2.52	1.76E-3	2.68	2.68
1a23	1.026E1	1.72E-1	2.57	2.74E-2	2.65	2.65
1b4l	1.19E-1	1.25E-2	3.25	2.07E-3	2.59	2.59
1bbl	1.32E-1	1.86E-2	2.82	1.81E-3	3.36	3.36
1bor	9.44E-2	1.31E-2	2.84	1.97E-3	2.73	2.73
1fca	1.20E-1	1.20E-2	3.31	1.78E-3	2.76	2.76
1frd	7.93E-2	1.24E-2	2.67	2.02E-3	2.61	2.61
1fxd	7.66E-2	1.19E-2	2.68	2.00E-3	2.57	2.57
1hpt	8.05E-2	1.37E-2	2.50	1.77E-3	2.90	2.90
1mbg	1.35E-1	1.08E-2	3.64	1.69E-3	2.67	2.67
1neq	8.52E-2	1.27E-2	2.74	1.83E-3	2.79	2.79
1r69	7.95E-2	1.15E-2	2.39	1.92E-3	2.96	2.96
1svr	7.94E-2	1.17E-2	2.21	1.94E-3	3.13	3.13
1uxc	7.55E-2	1.27E-2	2.57	2.02E-3	2.65	2.65
1vjw	7.22E-2	1.20E-2	2.58	2.23E-3	2.43	2.43
2pde	1.12E-1	1.64E-2	2.77	5.46E-3	1.58	1.58

the MIBPB package.

As indicated at the beginning, all the mathematical algorithms and techniques are enforced to maintain the high order convergence of the MIBPB solver. Table 4.5 lists the numerical evidence of the second order convergence through a set of given protein surfaces, atomic coordinates, radii and charges, where protein surfaces are generated by MSMS, and the standard CHARMM force field parameters are used. A special analytical solution was designed and given in [67] for the convergence order check of all proteins. In this table, the numerical error is defined as $\|u_h^{\text{num}} - u^{\text{exact}}\|_{L_\infty}$,

where u_h^{num} is the numerical solution of the PBE at grid resolution h while u^{exact} is the designed exact solution. The numerical experiments are implemented under resolutions $h = 1.0\text{\AA}$, 0.5\AA and 0.25\AA . The numerical error is supposed to be reduced by four times as the grid size is halved and this is clearly demonstrated in the table.

The above mentioned tests are carried out in conjunction with the PETSc software package, whose installation may not be so straightforward. An alternative is to use the SLATEC, which is easier to implement and also includes tens of linear system solvers with different preconditioners. To compare the performance of the PETSc and the SLATEC, we show the computation time of ten methods in the SLATEC for five proteins, whose atom number varies from five hundreds to eight thousands in Table 4.6. All methods are listed in the form: preconditioner/solver. Here GS, DS, BiGS, and OM represent the Gauss-Seidel, the diagonal scaling, the biconjugate gradient squared method, and the orthomin sparse iterative method, respectively. The combination of the ILU/BiCG is used in the PETSc. From the table, one can see that the iteration time of the PETSc is slightly shorter than that of most solvers in the SLATEC for small-sized proteins. The last column of the table lists the averaged CPU time for the PETSc and solvers in SLATEC. The averaged time, which in some sense could reflect the abilities of solvers for proteins in various sizes, is the sum of the CPU time for each corresponding protein and weighted by the atom number. By checking the averaged CPU time one can generally conclude that the ILU/BiCG of the PETSC takes less iteration time than the SLATEC schemes do. Moreover, according to our experience, the PETSc is more stable than the SLATEC for large proteins. However, the SLATEC can be easily incorporated in the MIBPB package.

Table 4.6: Comparison of CPU time for the PETSc and the SLATEC schemes

Protein ID Atoms	1ajj 519	1vjw 828	1a2s 1272	1a7m 2809	2ade 8344	Averaged CPU time
PETSc	0.235	0.272	0.529	1.729	3.777	2.72
GS/GS	0.866	1.222	2.225	9.512	55.016	35.58
ILU/ILU	0.523	0.883	1.344	5.854	32.479	21.07
DS/BiCG	0.331	0.467	1.041	3.140	14.015	9.27
ILU/BiCG	0.262	0.401	0.701	2.038	7.846	5.27
DS/BiGS	0.243	0.313	0.602	2.900	8.879	6.05
ILU/BiGS	0.187	0.393	0.410	1.433	6.575	4.34
DS/OM	0.206	0.420	0.496	3.338	21.993	14.08
ILU/OM	0.179	0.291	0.389	1.25	5.993	3.95
DS/GMRES	0.417	0.559	0.999	3.856	26.262	16.84
ILU/GMRES	0.198	0.279	0.439	1.615	7.685	5.05

4.2 Application to solvation energy calculations

One of the most important applications of the PBE model is solvation energy calculations for the protein-solvent systems. In this section, solvation energies of 28 proteins are calculated and compared with popular PBE solvers to examine the feasibility, usefulness and robustness of the linear solver in the MIBPB package. These proteins have a wide range of numbers of atoms, from around 500 up to 10,000. The corresponding spatial dimensions extend from about $30\text{\AA} \times 30\text{\AA} \times 30\text{\AA}$ to almost $100\text{\AA} \times 100\text{\AA} \times 100\text{\AA}$. Among these calculations, the dielectric constant is set to 1 for proteins and 80 for the solvent. The ion strength $\bar{\kappa}$ is set to zero because no ion is considered for the moment.

The calculation of electrostatic solvation energy ΔG_{elec} is to sum all the fixed charges $\{q_i\}$ of the solute in the solvent, weighted by the reaction field potential $\phi_{\text{rf}}(\mathbf{x})$:

$$\Delta G_{\text{elec}} = \frac{1}{2} \sum_i Q_i \Phi_{\text{rf}}(\mathbf{x}_i) \tag{4.7}$$

where \mathbf{x}_i is the position of each charge. Based on the continuum electrostatics, the reaction field potential is the difference between the electrostatic potential in

the solvent environment $\Phi_s(x)$ and the reference electrostatic potential $\Phi_{\text{ref}}(\mathbf{x})$, i.e., $\Phi_{\text{rf}}(\mathbf{x}) = \Phi_s(\mathbf{x}) - \Phi_{\text{ref}}(\mathbf{x})$. Here $\Phi_{\text{rf}}(\mathbf{x})$ can be computed by solving the PBE twice with corresponding settings. Specifically, $\Phi_{\text{ref}}(\mathbf{x})$ is calculated by setting a uniform dielectric constant in the whole computational domain, while $\Phi_s(\mathbf{x})$ is calculated by setting the dielectric constants for solute and solvent differently. Therefore, $\Phi_{\text{ref}}(\mathbf{x})$ can be obtained by the standard linear PB equation with the Dirichlet boundary condition via the standard finite difference or FFT methods but $\Phi_s(\mathbf{x})$ is solved by using the MIBPB algorithm.

The performance of the MIBPB method for calculating solvation energies has already been examined in our previous work [67]. It has been shown that the MIBPB solver has high accuracy and good convergence order because of the use of interface treatments but has relatively low numerical efficiency due to the absence of appropriate matrix acceleration techniques. The MIBPB matrix requires longer CPU time to solve. The Krylov theory and associated preconditioners discussed in the present work make the MIBPB solver more efficient. Here the new results are presented for various proteins.

Figure 4.4 gives the comparison of the calculated solvation energies from the MIBPB and the APBS packages. It is seen that the solvation energies calculated from the MIBPB agree very well with those from the APBS. The mesh sizes of $h = 1\text{\AA}$ is used in the MIBPB and $h = 0.25\text{\AA}$ is used in the APBS methods, respectively. The reader is referred to Ref. [67] for a more detailed comparison among the MIBPB, the APBS and the PBEQ methods.

Once the preconditioning techniques are applied, the required CPU time is significantly reduced. Figure 4.5 illustrates the differences of the CPU time needed to calculate solvation energies for nineteen moderately large proteins at three different grid resolutions. The solid lines are the CPU time for preconditioned (PCed) systems and dashed lines are for unpreconditioned (unPCed) systems. It can be concluded

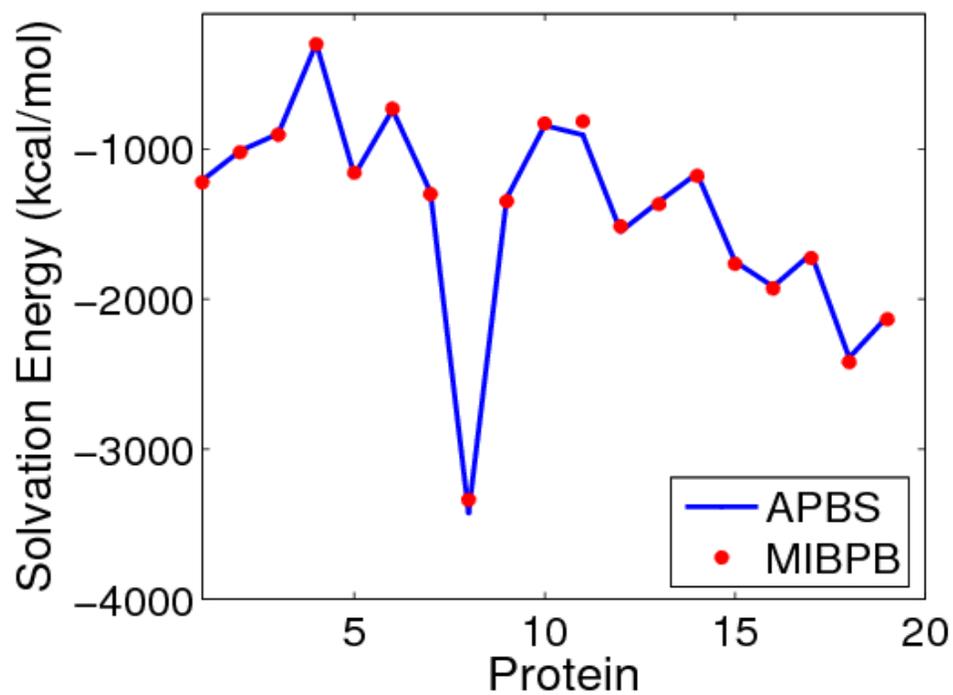


Figure 4.4: Comparison of solvation energies of proteins (From protein 1 to 19: 1ajj, 1bbl, 1vii, 1cbn, 2pde, 1sh1, 1fca, 1fxd, 1vjw, 1bor,1hpt,1bpi, 1mbg, 1r69, 1neq, 1a2s, 1svr, 1a63, 1a7m) calculated by using the MIBPB and the APBS methods.

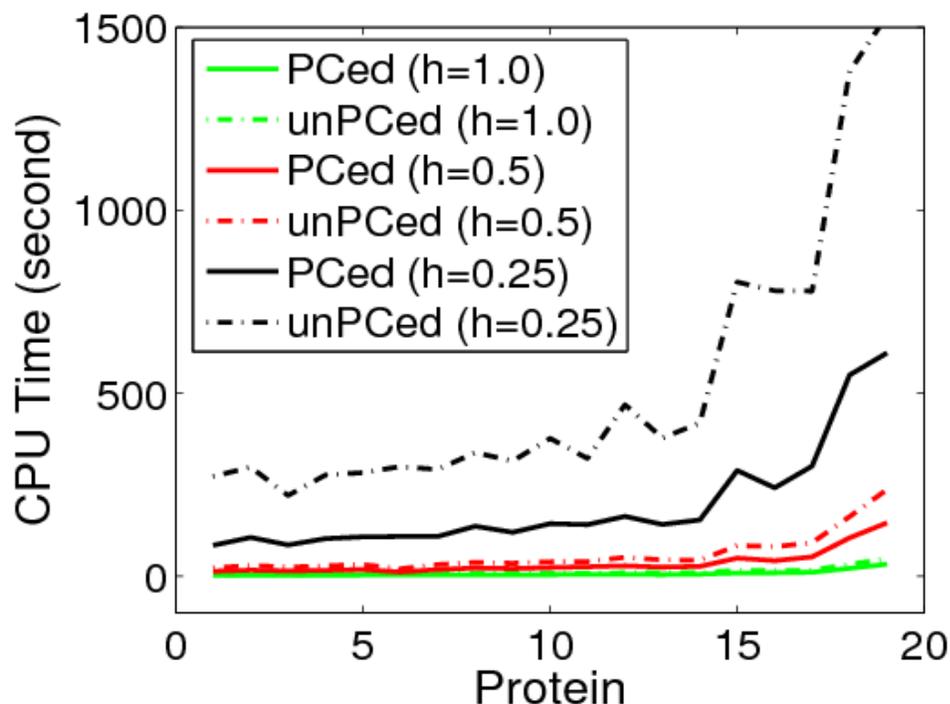


Figure 4.5: Comparison of CPU time of preconditioned (PCed) and un-preconditioned (unPCed) MIBPB methods for 19 proteins (from protein 1 to 19: 1ajj, 1bbl, 1vii, 1cbn, 2pde, 1sh1, 1fca, 1fxd, 1vjw, 1bor, 1hpt, 1bpi, 1mbg, 1r69, 1neq, 1a2s, 1svr, 1a63, 1a7m).

that at each grid resolution, preconditioners can save more than half of the overall CPU time.

Table 4.7 lists the results for the tested proteins at different grid resolutions and compares the values with the original MIBPB-III scheme in terms of solvation energies and CPU time. For each protein case from different grid resolutions, the CPU time increases in nonuniform pattern from less than 10 seconds for $h = 1.0\text{\AA}$, several tens of seconds for $h = 0.5\text{\AA}$, to a few hundreds of seconds for $h = 0.25\text{\AA}$. Note that there is an eight times increase in the number of unknowns when the mesh size is halved. The increase in the CPU time is roughly linear with the increase in the number of unknowns.

It is found that, at resolution of 0.25\AA , the results from the MIBPB+KSP and

Table 4.7: Solvation energies and CPU time for proteins

h(Å)	Solvation energy (kcal/mol)				CPU time (second)			
	MIBPB+KSP			MIBPB	MIBPB+KSP			MIBPB
	1.0	0.5	0.25	0.25	1.0	0.5	0.25	0.25
1ajj	-1141.9	-1136.3	-1136.6	-1137.2	2.3	15	85	273
2pde	-826.2	-819.9	-817.2	-820.9	3.4	17	108	283
1vii	-914.6	-901.5	-902.8	-901.2	2.8	15	86	221
1cbn	-311.1	-303.6	-303.7	-303.8	2.9	16	102	277
1bor	-858.8	-854.3	-857.9	-853.7	4.5	24	143	377
1bbl	-998.3	-986.9	-988.7	-986.8	2.8	16	106	298
1fca	-1215.5	-1199.9	-1200.0	-1200.1	3.5	19	109	292
1uxc	-1157.6	-1138.1	-1139.1	-1138.7	4.2	21	127	347
1sh1	-755.7	-728.0	-751.4	-753.3	3.3	12	109	300
1mbg	-1368.4	-1349.8	-1352.4	-1346.1	4.8	25	142	378
1ptq	-893.5	-871.8	-872.2	-873.1	3.9	22	133	376
1vjw	-1250.9	-1236.9	-1236.9	-1237.9	4.1	26	120	315
1fxd	-3309.7	-3299.7	-3301.6	-3300.0	4.2	31	138	338
1r69	-1111.0	-1086.5	-1087.9	-1089.5	5.5	32	154	419
1hpt	-827.3	-810.9	-812.7	-814.3	4.9	26	141	322
1bpi	-1320.8	-1298.9	-1301.3	-1301.9	5.4	50	164	469
1a2s	-1928.8	-1913.1	-1913.6	-1913.5	9.6	47	242	780
1frd	2879.8	-2851.4	-2856.3	-2851.9	10.8	51	284	707
1svr	-1741.6	-1709.8	-1710.7	-1711.2	11.1	57	301	779
1neq	-1765.6	-1729.1	-1732.7	-1730.1	9.1	50	289	804
1a63	-2420.8	-2371.2	-2370.2	-2373.5	22	113	550	1376
2erl	-964.2	-948.2	-949.3	-948.8	2.3	15	101	276

from the original MIBPB-III have less than 0.1% disagreement. This is due to the use of different convergence norms in the KSP solvers and the regular solver. The solvation energy calculations show a correct convergence tendency. The values from resolutions of 0.25\AA and 0.5\AA are pretty close, while calculations at $h = 0.25\text{\AA}$ cost much more CPU time. Therefore, we can conclude that grid resolution between 0.5\AA and 1.0\AA is sufficient for the calculation and can guarantee the accuracy.

Table 4.8 shows the robustness and efficiency of the MIBPB package for calculating solvation energies of large proteins which exceed 7000 atoms. For time efficiency, all the calculations are carried out at the grid resolution of $h = 1.0\text{\AA}$. Note that the reliability of these solvation free energies has been cross-checked with other popular

Table 4.8: Solvation energies (kcal/mol) and CPU time (second) for large proteins

Protein		MIBPB ($h = 1.0\text{\AA}$)	
ID	atoms	Solvation energy	CPU time (s)
1cbg	7838	-5659.9	181
1c4k	11439	-9901.9	398
1e24	7776	-9506.4	231
1f6w	8243	-5611.2	225
1po7	7796	-5471.2	206

PB solvers. The reported CPU time can be used as a reference.

4.3 Application to salt effects on protein-protein binding

In this section, the ability of the MIBPB package to solve the nonlinear PBE is tested by solvent salt effect on protein-protein binding. The nonlinear PBE has had considerable success in describing biomolecular electrostatics with salt effects on the binding of ligands, peptides and proteins to nucleic acids, membranes and proteins. The binding free energies reflect the non specific salt dependence of the formation of macromolecular complex and the measurement is the binding affinity. Some experimental data are available and the binding affinity is calculated as the ratio between salt dependent binding energy $\Delta\Delta G_{el}(I)$ at a specific salt strength I and the natural logarithm of I . In the present work, we have implemented the nonlinear version of the PBE solver in the MIBPB package. Simulation results are obtained by varying the ionic strengths.

The binding energy (ΔG_{el}) has several components while the one related to electrostatics is the difference of the electrostatic free energies of the complex and each

of its free molecules [17]

$$\Delta G_{\text{el}}(I) = G_{\text{el}}^{\text{AB}}(I) - G_{\text{el}}^{\text{A}}(I) - G_{\text{el}}^{\text{B}}(I), \quad (4.8)$$

where $G_{\text{el}}^{\text{AB}}(I)$, $G_{\text{el}}^{\text{A}}(I)$ and $G_{\text{el}}^{\text{B}}(I)$ represent the electrostatic free energies of the complex AB, component A and component B, respectively, at a given ionic strength I .

The electrostatic free energy can be further split into three components

$$G_{\text{el}}(I) = G_{\text{coul}} + G_{\text{rxn}} + G_{\text{salt}}(I), \quad (4.9)$$

where G_{coul} is the Coulomb energy calculated in a homogeneous medium, G_{rxn} is the corrected reaction field energy and $G_{\text{salt}}(I)$ is the electrostatic energy contributed by mobile ions. Among the three terms in Eq. (4.9), only $G_{\text{salt}}(I)$ is salt dependent. Thus, the salt dependence of the binding free energy $\Delta\Delta G_{\text{el}}(I)$ is electrostatic component of the binding energy in Eq. (4.8) calculated at some salt strength I minus the one calculated at the zero salt concentration [17]

$$\begin{aligned} \Delta\Delta G_{\text{el}}(I) &= \Delta G_{\text{el}}(I) - \Delta G_{\text{el}}(I = 0) \\ &= \left\{ G_{\text{el}}^{\text{AB}}(I) - G_{\text{el}}^{\text{AB}}(I = 0) \right\} \\ &\quad - \left\{ G_{\text{el}}^{\text{A}}(I) - G_{\text{el}}^{\text{A}}(I = 0) \right\} \\ &\quad - \left\{ G_{\text{el}}^{\text{B}}(I) - G_{\text{el}}^{\text{B}}(I = 0) \right\}, \end{aligned} \quad (4.10)$$

where various energy terms are calculated at different ionic strengths by using the MIBPB package.

To verify our nonlinear solver, one hetero-dimeric and one homo-dimeric complexes are studied in this work. The experiments on these two protein complexes can be found in [130, 152] and biological features(1emv and 1beb) are listed in Table 4.9.

The first four columns describe the properties of proteins and the last two columns are the slopes (binding affinity) of the lines in Fig. 4.6. It can be seen in a quantitative view that the slopes obtained from experiments and simulations are very close to each other. The calculations were performed assuming that all Arg, Asp, Glu and Lys residues are ionized in both free and bound states. The results are obtained with dielectric constants of 2 and 80 for the solute and the continuum solvent, respectively. The parameter $\bar{\kappa}^2$ is determined by:

$$\bar{\kappa}^2 = \left(\frac{8\pi^2 N_a q^2}{1000 k_B T} \right) I \quad (4.11)$$

where q , $k_B T$ are the same as those defined in Eq. (4.2), N_a is the Avogadro's number. After a simple derivation, $\bar{\kappa}^2$ is given by

$$\bar{\kappa}^2 = 8.486902807 \text{\AA}^{-2} I \quad (4.12)$$

for $T = 298K$. Here the ion strength I is in the unit of molar.

Figure 4.6 depicts the experimental and calculated salt dependence of the binding free energies $\Delta\Delta G_{el}(I)$ for the two complexes. The $\Delta\Delta G_{el}(I)$ are plotted against the logarithm of the salt strength I . The salt dependence is assumed as in a linear pattern therefore the least square fitting line is applied to calculate the binding affinity, which is the slope of the line. It should be explained that experimental data dots for $\Delta\Delta G_{el}(I)$ are read from the graphs in Ref. [17], while the fitting line slope is explicitly given based on the experimental data with error bars. The diamond points and solid line are experimental data and the corresponding fitting line, respectively. The circle points and dashed line are numerical stimulations.

In the homo-dimeric complex, the experimentally observed binding free energies decrease with the increase of ionic strength, while for the hetero-dimeric complex, the

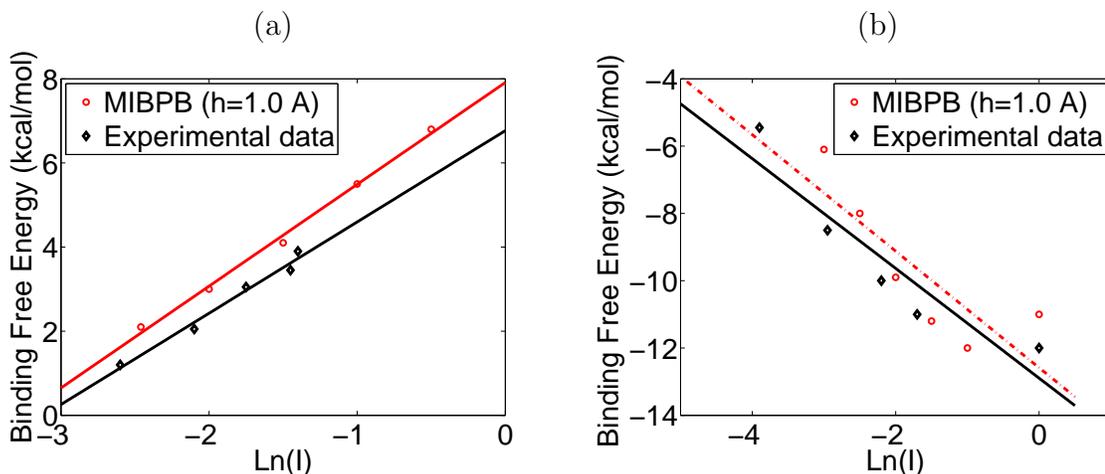


Figure 4.6: (a) Binding affinity of 1emv; (b) Binding affinity of 1beb.

Table 4.9: Binding affinity

Complex	PDB ID	Complex charge	Surface area (\AA^2)	Charge of the free monomers	Exp. data	MIBPB $h = 1.0\text{\AA}$
E9Dnase-Im9 (10) (B-A)	1emv	-3	1465	B=+5; A=-8	2.17	2.42
Lactoglobulin dimer (57)(A-B)	1beb	+26	1167	A=B=+13	-1.62	-1.90

experimental measurement had detected an increase. Our results obtained from the MIBPB package reproduced these observations. The calculated magnitudes of the slopes of the salt dependence are in quite good agreement with experimental results within the range of errors, as the fitting lines are almost parallel. The discrepancies between the experimental data and simulation results are also expected: the PBE, no matter in linear or full nonlinear form, only gives the ideal situation of the solute-solvent system but many details, such as the “protein conformation change, pK_a shifts upon complexation or possible ionization states”, are lacking [17]. Despite these approximation, the application of PBE for static protein structures is suggested in general by these good agreements with experimental data.

4.4 Conclusion remarks

The MIBPB has built in advanced interface techniques that are able to deal with discontinuous solvent-solute interfaces and geometric singularities of molecular surfaces. Therefore, it provides reliable electrostatic potentials at the mesh of about 1\AA , whereas traditional methods have to resort to about 0.25\AA mesh resolution to achieve a similar level of reliability. In the present work, we further equip the MIBPB solver with advanced Krylov subspace techniques to accelerate the speed of the convergence of solving linear equation systems originated from the MIBPB discretization. The performances of various solver-preconditioner combinations have been carefully examined through mathematical analysis and numerical experiments. Dramatic reductions in condition numbers are found when appropriate preconditioners are utilized. Upon the use of appropriate combinations of preconditioners and solvers, significant reductions in the CPU time are found in solving the PB equation for large proteins. Both the PETSc and the SLATEC are employed in the present MIBPB package to speed up the convergence rate of the iterations of the linear systems. The PETSc package is found to be more reliable and efficient. In the present work, the structure preparation of proteins is conducted via the PDB2PQR software package, while the MSMS software package is utilized for the molecular surface generation. Additionally, the nonlinear MIBPB solver has been developed in the present work. This is achieved via the standard inexact Newton method, assisted by the Krylov subspace acceleration techniques. The present nonlinear MIBPB solver has been tested and applied to the salt dependence analysis of protein-protein binding interactions. Our results of binding affinities are compared with experimental data.

Chapter 5

Thesis achievements and future work

The major contribution of this thesis is the mathematical modeling and highly efficient computations for nano-electronic transistors and transmembrane proton channels. The former subject includes electronic semi-conductor devices that have continuous demand nowadays in rising the performance and reducing the dimension; while the latter belongs to the study of ion channels in molecule-based biology. The two topics share similar physical characteristics, attract same mathematical interests and encounter common simulation challenges. In both fields, motions of quantum particles (electrons, protons) are studied under intensive electrostatic and general-correlation potential interactions at nano-scale. With reasonable assumptions, energy components are assembled on an equal-footing energy functional from which governing equations are derived. For the two governing equations, the generalized Poisson-Boltzmann equation and the generalized Kohn-Sham equations, there are a number of singularities and numerical challenges that attracted numerous mathematical interests in recent decades. Corresponding mathematical algorithms are adopted to overcome these difficulties in simulations.

Most of the materials of this thesis are from the following publications:

- **Duan Chen** and Guo-wei Wei, “Quantum dynamics in continuum for proton channel transport.”, *submitted to Journal of computational Physics*
- **Duan Chen** and Guo-Wei Wei, “Modeling and simulation of electronic structure, material interface and random doping in nano-electronic devices” *Journal of Computational Physics*, 229, Vol. 229, pp. 4431-4460, 2010.
- **Duan Chen**, Zhan Chen, Changjun Chen, Weihua Geng and Guo-Wei Wei, “MIBPB: A software package for electrostatic analysis”, *Journal of Computational Chemistry*, in press , 2010
- **Duan Chen**, G.W. Wei, X. Cong and Ge Wang, “Computational methods for optical molecular imaging”, *Commun. Numer. Methods Engng*, vol. 25, pp.1137-1161, 2008

5.1 On the modeling and simulations of nano electronic transistors

5.1.1 Contributions

This thesis contributes to the studies on nano-MOSFETs with the following aspects: First of all, it is the first time that a two scale energy functional that includes continuum mechanism and quantum mechanics on an equal footing has been proposed, and the coupled Poisson-Kohn-Sham equations are derived from the optimization of the total energy functional. The macroscopic electrostatics energy functional is defined on the inhomogeneous silicon and silicon dioxide system, while the motion of major charge carriers (electrons in this work) is microscopically described through

interactions between electrons and continuum/discrete dopants as well as the self-interactions.

Secondly, two practical issues are addressed in the current work. One is the material interface problem, the ratio of dielectric constants of the silicon and the insulator has a great impact on the device performance. The recognition of material interfaces in MOSFETs implies the acknowledgment of discontinuous material properties or coefficients across the interfaces. The interface effect study is naturally contained in our Poisson equation with discontinuous coefficients. Another topic is the individual dopant effect. Electrical properties of the nano-transistor, i.e., the conductivity, electrostatic potential and its charge carrying mode can be manipulated easily by doping strategies. In continuum modeling, dopants have always been described as continuous distributions. These treatments work mostly well for the prediction of device properties. However, when the channel length reduces to nano-scale, the quantum effect becomes important. Thus, each doping atom may have a dramatic impact on the quantum state of nearby electrons. Therefore, the atomistic model for dopants becomes indispensable. This work presents a new individual dopant model by utilizing the Dirac delta function.

Corresponding to the aforementioned issues, this work introduces two computational algorithms for the simulation of nano-electronic devices. An efficient elliptic interface method, the matched interface and boundary (MIB) method, is employed for solving the Poisson equation with semiconductor interfaces. The other computational algorithm employed in the present work is the Dirichlet-to-Neumann mapping (DNM). This approach provides a rigorous treatment of the singular charge sources in the Poisson equation due to the individual dopant model proposed in the present work. Necessary numerical analysis are presented. The second order MIB scheme is utilized in the present work; however, due to the strong nonlinearity of the coupled Poisson-Kohn-Sham equations, the numerical order was found to be about 1.5 in

the present Gummel-like iteration. Self-consistent iteration difficulties under certain circumstances are overcome by a relaxation scheme.

Finally, two multi-gate MOSFET systems, a double-gate MOSFET and a four-gate MOSFET, are considered in the numerical simulations of the present work. Both problems are modeled in the three-dimensional (3D) setting. In our double-gate MOSFET simulations, the basic characteristics and the quantum effect of the I-V curves are similar to those in the literature. The impact of randomly distributed individual dopants on electronic structure and transport is studied. In particular, individual dopants induced voltage threshold lowering effect is clearly demonstrated. In our four-gate MOSFET simulations, individual dopants effectively break the symmetry of the device. Due to the 2D quantum confinement, the density of quantum states that are relevant to the electronic transport become larger than that of the double-gate MOSFET.

5.1.2 Future work

In our current study, nano-transistors are studied with relatively simplified assumptions. For our future work, the modeling and simulation can be improved as follows;

1. Currently the electron transport is assumed in the ballistic limit, i.e., the transport of electrons is assumed to have no interaction with other objects. Real devices actually operate below this limit. The interaction of electrons with their surroundings is called scattering. Important scattering mechanisms include the surface roughness scattering, electron-electron scattering, impurity-electron scattering and phonon-electron scattering. A sophisticated model needs to be set up to take into account these effects.

2. The existing theory is for general nano-transistor modeling while our numerical implementations, or specifically, the transport simulations are only suitable for simple MOSFET geometries. In fact, there are a wide variety of geometric configurations

for MOSFETs for different performance requirements. Therefore, the Kohn-Sham equation decomposition technique must be dropped. Alternatively, an efficient full 3D algorithm for transport needs to be developed.

5.2 On the modeling and simulations of transmembrane proton channels and biomolecule structures

5.2.1 Contributions

Although there are a variety of excellent theoretical models and efficient computational methods for ion channels in general, most commonly used models are much less successful when they are applied to proton channel transport due to the special features of proton channels. Our present work on proton channel contributes a quantum dynamics in continuum (QDC) model for the prediction and analysis of proton density distribution and conductance in proton channels.

First we consider a necessary quantum mechanical treatment of protons due to their light mass and channel geometric confinement in protein channels. Additionally, primary ion-ion electrostatic interactions in proton channels are given top priority, and then a dielectric continuum treatment of solvent medium works as a reasonable approximation to the effect of numerous solvent molecules. Furthermore, densities of all other ions and counter-ions in the solvent are described by the Boltzmann distribution, which is a quasi-equilibrium description as the electrostatic potential varies during the process of a proton permeating through the membrane.

We proposed a multiscale variational paradigm to accommodate the aforementioned aspects in a unified framework. The total free energy functional encompasses

the kinetic and potential energies of protons, and the electrostatic energies of ions and fixed charges in the channel system. The first variation is carried out via the Euler-Lagrange equation to derive the governing equations for the system. A generalized Poisson-Boltzmann equation is obtained for the electrostatic potential while a generalized Kohn-Sham equation results for the orbitals of protons in the system. The solution to these two coupled nonlinear equations leads to the desired electrostatic distribution and proton density profile in the channel. Expressions for proton density and proton current across the membrane are derived from fundamental principles.

A complete, accurate and efficient Poisson-Boltzmann solver MIBPB, is developed for electrostatics analysis in molecular biology. This solver serves as a powerful tool in proton transport calculation and other solvation applications. The previous MIBPB-I, the MIBPB-II and the MIBPB-III are equipped with matched interface and boundary (MIB) scheme and Dirichlet-to-Neumann mapping (DNM) method. They overcome numerical challenges of discontinuous coefficients (i.e., piecewise dielectric constants), singular sources (i.e., Dirac delta functions for protein charges), and non-smooth interfaces (i.e., geometric singularities). In this thesis, the Krylov subspace theory and preconditioner methods are utilized to enhance the solver efficiency, the inexact Newton's method based on these two techniques is included to extend the PB solver to a nonlinear one. Detailed solver processes that contain protein structure preparation, PB calculation and solution visualization are provided.

5.2.2 Future work

Our model is based on several assumptions which may not hold in realistic biological processes, I intend to continue working on the modeling and simulation of ion channel problems with further details, which may include the following aspects:

1. In the current model set up, the plasma membrane is modeled as a uniform dielectric continuum. While in realistic conditions, there are various types of mem-

branes, some of which have dipoles and others of which have charges. There will be no essential difficulty to improve our model in this aspect for the future work. Point charges from membranes can be added in the present model. Otherwise, a position dependent dielectric constant for the biomolecular region can also represent the charge effects in the membrane;

2. The configurational cooperation of the channel protein needs to be included in the model in a more explicit way. At this stage, the channel structure is assumed to be rigid and the missed ion-protein interaction is somehow compensated by the dielectric constant. In our future work, the multiscale molecular dynamics may be adopted to remove this “frozen” structure assumption.

3. A major limitation of the present model originates from the oversimplified treatment of the general-correlation potential. Comparing to the electrostatic potential, the general-correlation potential is solely related to the relaxation time of particles in the channel. It is more phenomenological and lacks detailed information of ion-water interactions. Therefore, a major task in our future work is to collect the literature regarding the chemical/physical details about ion dehydration and hydrogen-bond breaking/making processes, to use reasonable mathematical formulation, and to eventually establish more quantitative modeling of the general-correlation potential.

APPENDIX

Appendix A

The MIB method

In this section, the MIB method is discussed in general \mathbb{R}^d dimensional space. Without loss of generality, assume that the computational domain is a compact rectangle

$$[a_1, b_1] \times [a_2, b_2] \times \dots \times [a_d, b_d]$$

and uniformly discretized with grid spacing h for simplicity. Then for $j \in \{1, 2, \dots, d\}$, $N_j = \frac{(b_j - a_j)}{h}$ is the number of grid points along the x_j direction. Since the MIB is a dimension splitting method for approximating the differential operator

$$\sum_{j=1}^d \frac{\partial}{\partial x_j} \left(\beta \frac{\partial}{\partial x_j} \right),$$

only the discretization of a specific direction, say x_{j^*} , is illustrated for simplicity. Discretizations of other directions can be easily derived in a similar way.

Let $e_1 = (1, 0, 0, \dots, 0)$ and $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ be respectively the 1st and j^{th} standard coordinate vectors of \mathbb{R}^d space and the d -tuple $E = (c_1, c_2, \dots, c_d)$ be the index of a specific grid point. We denote a non-boundary grid point as X_E (i.e., $c_j \in \{2, 3, \dots, N_j - 1\}$), and its function value as $u_E = u(X_E)$. If X_E is near the interface, its discretization scheme depends whether it is a regular point or an

irregular point according to the designed convergence order [164]. At a regular point where the discretization scheme does not involve any grid point across the interface, the approximation for the differential operator is the standard central finite difference scheme

$$\frac{\partial}{\partial x_{j^*}} \left(\beta \frac{\partial u}{\partial x_{j^*}} \right) = \frac{\beta_{E-\frac{1}{2}e_{j^*}}^+}{h^2} u_{E-e_{j^*}} - \frac{\beta_{E-\frac{1}{2}e_{j^*}}^+ + \beta_{E+\frac{1}{2}e_{j^*}}^+}{h^2} u_E \quad (\text{A.1})$$

$$+ \frac{\beta_{E+\frac{1}{2}e_{j^*}}^+}{h^2} u_{E+e_{j^*}} + O(h^2). \quad (\text{A.2})$$

The MIB scheme is applied at an irregular point where the discretization scheme involves at least one grid point across the interface and it will be described in three major steps in the following three subsections.

A.1 Dimension splitting

The MIB method takes the dimension splitting approach which reduces a multidimensional interface problem into 1D ones. Consequently, the MIB method simplifies the local topological relation near an interface, which is crucial for 3D problems with complex interface geometries. The MIB method provides special schemes around each intersecting point of the given interface and prescribed meshlines. Therefore, at each intersecting point which is not a grid point, there is one meshline, say x_{j^*} , that is a primary direction. In a second order MIB scheme, if the interface Γ passes between two grid points X_E and $X_{E+e_{j^*}}$, then X_E and $X_{E+e_{j^*}}$ are a pair of irregular points. For simplicity, we denote the consecutive grid points $X_{E+ke_{j^*}}$ ($k \in \{\dots, -1, 0, 1, \dots\}$) on the x_{j^*} axis as $i+k$ and function values $u(X_{E+ke_{j^*}})$ on these grid points as u_{i+k} .

Fictitious values f_1 and f_2 are constructed as smoothly extended function values of u^+ and u^- on both sides of the interface along the x_{j^*} direction. The fictitious

value f_1 is the extension of the solution in the left domain to the grid point $i + 1$, and f_2 is the extension of the solution in the right domain to the grid point i , while the vertical line indicates the interface location. Unlike the original function value u which might be discontinuous, both u^+ (left) and u^- (right) are well-behaved across the interface. With the assistance of fictitious values, u^\pm are matched by the jump conditions at the interface by uniform interpolation.

Assume $X_0 \in \mathbb{R}^d$ is the position where the interface intersects the mesh line. We discretize the interface jump conditions as the follows,

$$[u] = u^+ - u^- = \phi(X_0) \quad (\text{A.3})$$

where

$$\begin{aligned} u^+ &= w_{-1}^{+,0} u_{i-1} + w_0^{+,0} u_i + w_1^{+,0} f_1 + O(h^3) \\ u^- &= w_{-1}^{-,0} f_2 + w_0^{-,0} u_{i+1} + w_1^{-,0} u_{i+2} + O(h^3) \end{aligned} \quad (\text{A.4})$$

and

$$\begin{aligned} [\beta u_n] &= \beta^+ \nabla u^+ \cdot \vec{n} - \beta^- \nabla u^- \cdot \vec{n} \\ &= \begin{pmatrix} \beta^+ u_{x_1}^+ - \beta^- u_{x_1}^- \\ \beta^+ u_{x_2}^+ - \beta^- u_{x_2}^- \\ \dots \\ \beta^+ u_{x_{j^*}}^+ - \beta^- u_{x_{j^*}}^- \\ \dots \\ \beta^+ u_{x_d}^+ - \beta^- u_{x_d}^- \end{pmatrix} \cdot \vec{n} = \varphi(X_0) \end{aligned} \quad (\text{A.5})$$

Here $w_k^{\pm,l}$ ($l \in \{0, 1\}, k \in \{-1, 0, 1\}$) are interpolation weights which can be easily calculated by interpolations. The superscripts \pm present the two subdomains, and

0, 1 are for the interpolation and the first order derivative respectively, and the set of subscripts $-1, 0, 1$ is the index of grid points.

In jump condition (A.5), the first derivatives for all directions and from the two domains are all involved due to the $\nabla u^\pm = (u_{x_1}^\pm, u_{x_2}^\pm, \dots, u_{x_d}^\pm)^T$. Among them, $u_{x_{j^*}}^\pm$ can be easily obtained by interpolation with function values and fictitious values

$$\begin{aligned} u_{x_{j^*}}^+ &= w_{-1}^{+,1} u_{i-1} + w_0^{+,1} u_i + w_1^{+,1} f_1 + O(h^2) \\ u_{x_{j^*}}^- &= w_{-1}^{-,1} f_2 + w_0^{-,1} u_{i+1} + w_1^{-,1} u_{i+2} + O(h^2). \end{aligned} \quad (\text{A.6})$$

The evaluation of $u_{x_j}^\pm$ by the interpolation formulation for all $j = 1, 2, \dots, d$ and $j \neq j^*$ is presented in later sections A.3.

Symbolically, we assume $u_{x_j}^\pm$ are solved and the fictitious values f_1 and f_2 can totally determined by equations (A.3)(A.5) Once fictitious values f_1 and f_2 are determined, modified discretizations for $\frac{\partial}{\partial x_{j^*}} \left(\beta \frac{\partial u^\pm}{\partial x_{j^*}} \right)$ at grid points u_i and u_{i+1} are given by

$$\frac{\partial}{\partial x_{j^*}} \left(\beta^+ \frac{\partial}{\partial x_{j^*}} u^+ \right) = \frac{\beta_{i-\frac{1}{2}}^+}{h^2} u_{i-1} - \frac{\beta_{i-\frac{1}{2}}^+ + \beta_{i+\frac{1}{2}}^+}{h^2} u_i + \frac{\beta_{i+\frac{1}{2}}^+}{h^2} f_1 + O(h) \quad (\text{A.7})$$

and

$$\frac{\partial}{\partial x_{j^*}} \left(\beta^- \frac{\partial}{\partial x_{j^*}} u^- \right) = \frac{\beta_{i-\frac{1}{2}}^-}{h^2} f_2 - \frac{\beta_{i-\frac{1}{2}}^- + \beta_{i+\frac{1}{2}}^-}{h^2} u_{i+1} + \frac{\beta_{i+\frac{1}{2}}^-}{h^2} u_{i+2} + O(h). \quad (\text{A.8})$$

Methods for the determination of f_1 and f_2 are described in the next two subsections.

Remarks: (i) Grid points $i - 1$ and i are required to be in the same subdomain and so are $i + 1$ and $i + 2$. The assumption can always be satisfied by refining the grid mesh for a smooth interface. For sharp-edged interfaces, the grid refinement might not work. A special MIB scheme has been developed to deal with interface

singularities in [159].

(ii) From the discussion above, the 2nd order MIB scheme can be generalized to higher order ones (4th and 6th order) by extending the solution to more fictitious values near the interface. In the MIB method, this is done by repeatedly using lowest order jump conditions instead of creating higher order derivatives.

(iii) In Ref. [164], the fictitious values f_1 and f_2 have been shown to be of $O(h^3)$, which guarantees the first order local truncation error and the global second order convergence of the MIB scheme.

A.2 Derivative elimination

In general, at every intersecting point of the interface and mesh lines, there are two original interface conditions and $d - 1$ additional first order interface conditions for an elliptic interface in \mathbb{R}^d . These $d + 1$ interface conditions involve $2d$ first order derivatives. In the MIB method, we simultaneously determine as fewer fictitious values around an intersecting point as possible so that we have the maximal flexibility in avoiding the determination of many first order derivatives, which are often difficult to evaluate due to geometric constraints. Nevertheless, we have to determine at least two fictitious values so that both of the original two interface conditions can be implemented at least indirectly. Consequently, we determine only two fictitious values around an intersecting point and thus, use $d - 1$ interface conditions to eliminate $d - 1$ first order derivatives. The remaining $d + 1$ derivatives are to be approximated using appropriate grid function values near the intersecting point.

Specifically, jump conditions (A.3) and (A.5) are employed to determine two fictitious values f_1 and f_2 . However the first order derivatives along all directions on each subdomain are involved in Eq. (A.5). Two derivatives along the primary direction x_{j^*} can be approximated by Eq. (A.6). For $2d - 2$ derivatives along other directions,

i.e., $u_{x_j}^\pm$ ($j = 1, 2, \dots, d, j \neq j^*$), $d-1$ interface conditions, which are obtained by differentiating Eq. (A.3) along tangential directions, are used together with Eq. (A.5) to eliminate $d-1$ derivatives. There are flexibilities in the selection of $d-1$ derivatives from $2d-2$ derivatives. A general principle is to avoid evaluating those derivatives that are difficult to compute due to local geometric constraint and to optimize the resulting matrix of linear algebraic equations. In most cases, we normally evaluate one of $u_{x_j}^\pm$, either $u_{x_j}^+$ or $u_{x_j}^-$.

A closed interface Γ in \mathbb{R}^d can be considered as a \mathbb{R}^{d-1} manifold embedded in a \mathbb{R}^d space. Consider a map Φ from \mathbb{R}^{d-1} to \mathbb{R}^d :

$$\Phi : (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{d-1}) \rightarrow (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{d-1}, I(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{d-1})), \quad (\text{A.9})$$

where I is a hyper-surface function. The tangent space of Φ is given as

$$T_s\Phi = \text{span}(\Phi_{\tilde{x}_1}, \Phi_{\tilde{x}_2}, \dots, \Phi_{\tilde{x}_{d-1}}) =: \text{span}(\vec{t}_1, \vec{t}_2, \dots, \vec{t}_{d-1}), \quad (\text{A.10})$$

where \tilde{x}_k ($\forall k \in \{1, 2, \dots, d-1\}$) are manifold parameters, and $\Phi_{\tilde{x}_k}$ are derivatives of Φ with respect to \tilde{x}_k . The tangent vectors \vec{t}_k are not necessarily orthonormal to each other. However, in 2D and 3D practices, these tangent vectors can be orthonormal with appropriate parametrization of the manifold. From Eq. (A.3), the gradient jump at the interface can be decomposed into $d-1$ tangential jump conditions

$$[\nabla u] \cdot \vec{t}_j = \lim_{h \rightarrow 0} \frac{\phi(x + \vec{t}_j) - \phi(x)}{h}, \quad j = 1, 2, \dots, d-1. \quad (\text{A.11})$$

With the jump condition (A.5), a $d \times 2d$ system results in

$$(P^+, -P^-) \begin{pmatrix} \nabla u^+ \\ \nabla u^- \end{pmatrix} = \begin{pmatrix} \nabla\phi \cdot \vec{t}_1 \\ \nabla\phi \cdot \vec{t}_2 \\ \dots \\ \nabla\phi \cdot \vec{t}_{d-1} \\ \varphi \end{pmatrix}, \quad (\text{A.12})$$

where

$$P^+ = \begin{pmatrix} \vec{t}_1^T \\ \vec{t}_2^T \\ \dots \\ \vec{t}_{d-1}^T \\ \beta^+ \vec{n}^T \end{pmatrix}, \quad P^- = \begin{pmatrix} \vec{t}_1^T \\ \vec{t}_2^T \\ \dots \\ \vec{t}_{d-1}^T \\ \beta^- \vec{n}^T \end{pmatrix}. \quad (\text{A.13})$$

In Eq. (A.12), only one of each remaining $u_{x_j}^\pm$ is to be approximated. Therefore, $d-1$ variables need to be eliminated and it is always possible to eliminate $d-1$ variables from d independent equations. Since the tangent vectors and the normal vector are linearly independent, there is a well defined equation after the elimination. Without loss of generality, assume $u_{x_j}^+$ is to be eliminated from the system and the equation left from the elimination is

$$c_1 u_{x_{j^*}}^+ + c_2 u_{x_{j^*}}^- + \sum_{j=3, j \neq j^*}^d c_j u_{x_j}^- = \sum_{j=1}^{d-1} \tilde{c}_j \nabla\phi \cdot \vec{t}_j + \tilde{c}_d \varphi, \quad (\text{A.14})$$

where c_j and \tilde{c}_j are coefficients after the elimination. Finally, Equations. (A.3) and (A.14) are used to determine the two fictitious values f_1 and f_2 .

A.3 Derivative evaluation

After the elimination of $d - 1$ derivatives, we have to evaluate $d - 1$ derivatives. This is a difficult task for complex interfaces and is a challenging task for interfaces with geometric singularities. The evaluations are pursued along x_j ($j = 1, 2, \dots, d, j \neq j^*$) direction but in general not along any x_j axis. In 2D, the evaluation must compute in the x_{j^*} - x_j plane. In higher dimensions, the evaluation can be pursued with more flexibilities. In general, we evaluate a derivative in a specific x_j - x_i plane so that the resulting matrix is relatively diagonal and symmetric with respect to the given geometric constraint. Let us assume that x_{j^*} - x_j be such a choice and denote the consecutive grid points $X_{E+me_{j^*}+ne_j}$ along x_{j^*} axis and x_j axis as $(i+m, j+n)$, and the function values $u(X_{E+me_{j^*}+ne_j})$ as $u_{i+m,j+n}$, where $m, n \in \{\dots, -2, -1, 0, 1, 2, \dots\}$.

When the irregular points are off the interface, three auxiliary function values $u_{o,E}$, $u_{o,1}$ and $u_{o,2}$ along the auxiliary line $x_{j^*} = x_o$ are needed to approximate $u_{x_j}^-$, where x_o is the j^* component of X_o . Hence, $u_{x_j}^-$ is approximated by the first order derivative scheme

$$u_{x_j}^- = p_1 u_{o,E} + p_2 u_{o,1} + p_3 u_{o,2}, \quad (\text{A.15})$$

where p_1, p_2 and p_3 are derivative weights.

Unfortunately, $u_{o,E}$, $u_{o,1}$ and $u_{o,2}$ are not available since they do not locate on any grid point. Therefore, they are to be obtained from function values on other nearby grid points on the corresponding mesh lines. The auxiliary value $u_{o,E}$ locates between the irregular points (i, j) and $(i+1, j)$, which can be obtained with fictitious values

$$u_{o,E} = (w_{-1}^{-,0}, w_0^{-,0}, w_1^{-,0}) \cdot (f_2, u_{i+1,j}, u_{i+2,j})^T. \quad (\text{A.16})$$

However, the auxiliary values $u_{o,1}$ and $u_{o,2}$ have to be interpolated by other nearby

values

$$\begin{aligned}
u_{o,1} &= (\tilde{w}_{-1}^{-,0}, \tilde{w}_0^{-,0}, \tilde{w}_1^{-,0}) \cdot (u_{i+1,j+1}, u_{i+2,j+1}, u_{i+3,j+1})^T \\
u_{o,2} &= (\bar{w}_{-1}^{-,0}, \bar{w}_0^{-,0}, \bar{w}_1^{-,0}) \cdot (u_{i,j+2}, u_{i+1,j+2}, u_{i+2,j+2})^T.
\end{aligned} \tag{A.17}$$

Note that all the grid points used in interpolating the auxiliary points have to be in the same subdomain. By the assumption that after eliminating the $u_{x_j}^+$ and approximating $u_{x_j}^-$, all the grid points used to interpolate $u_{o,1}$ and $u_{o,2}$ should be in the Ω^- domain (there is no restriction for $u_{o,E}$ since fictitious values are involved). Therefore, the grid points $(i, j+2)$, $(i+1, j+2)$ and $(i+2, j+2)$ are the closest grid points to interpolate $u_{o,2}$. While for $u_{o,1}$, the grid point $(i, j+1)$ is in the Ω^+ domain, and the closest grid points for $u_{o,1}$ are $(i+1, j+1)$, $(i+2, j+1)$ and $(i+3, j+1)$. Therefore, the final expression for $u_{x_j}^-$ is:

$$\begin{aligned}
u_{x_j}^- &= [p_1, p_2, p_3] \begin{bmatrix} w_{-1}^{-,0} & w_0^{-,0} & w_1^{-,0} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{w}_{-1}^{-,0} & \tilde{w}_0^{-,0} & \tilde{w}_1^{-,0} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \bar{w}_{-1}^{-,0} & \bar{w}_0^{-,0} & \bar{w}_1^{-,0} \end{bmatrix} \\
&\times [f_2, u_{i+1,j}, u_{i+2,j}, u_{i+1,j+1}, u_{i+2,j+1}, u_{i+3,j+1}, u_{i,j+2}, u_{i+1,j+2}, u_{i+2,j+2}].
\end{aligned}$$

From the above equation we can see that $u_{x_j}^-$ is actually a linear combination of some unknown function values on grid points around the irregular point $u_{i,j}$. By repeating the same procedure to other coupled x_{j^*} - x_j planes, all similarly structured $u_{x_j}^-$ can be determined. Together with jump condition (A.5), two unknown fictitious values f_1 and f_2 can be determined.

If the target irregular point is right on the interface, the calculation of fictitious values is the similar. At the end of this paper, applications of the MIB scheme are given. The 2D MIB scheme is useful to interface problems in plastic membrane,

electromagnetic wave propagation, etc. When $d = 2$, the following normal vector and tangential vector are usually implemented as

$$\begin{aligned}
\vec{n} &= (\cos \theta, \sin \theta) \\
\vec{t} &= (-\sin \theta, \cos \theta) \\
P^\pm &= \begin{pmatrix} -\sin \theta & \cos \theta \\ \beta^\pm \cos \theta & \beta^\pm \sin \theta \end{pmatrix}, \tag{A.18}
\end{aligned}$$

where θ is the angle between the positive direction of the x -axis and the normal vector of the interface. It is easy to show that \vec{n} and \vec{t} are orthonormal. For the MIB scheme used in the implicit solvent model for electrostatics analysis in molecular biology or in biomedical image computation, we consider a 3D MIB scheme. The normal vector and tangential vectors can be chosen as:

$$\begin{aligned}
\vec{n} &= (\sin \psi \cos \theta, \sin \psi \sin \theta, \cos \psi) \\
\vec{t}_1 &= (-\cos \psi \cos \theta, -\cos \psi \sin \theta, \sin \psi) \\
\vec{t}_2 &= (-\sin \theta, \cos \theta, 0) \\
P^\pm &= \begin{pmatrix} -\cos \psi \cos \theta & -\cos \psi \sin \theta & \sin \psi \\ -\sin \theta & \cos \theta & 0 \\ \beta^\pm \sin \psi \cos \theta & \beta^\pm \sin \psi \sin \theta & \beta^\pm \cos \psi \end{pmatrix}, \tag{A.19}
\end{aligned}$$

where θ and ψ are the azimuth and zenith angles with respect to the normal direction. The explicit expressions of the fictitious values are referred to Refs. [165, 159].

Appendix B

Krylov subspace method and preconditioning for the MIB scheme

B.1 Linear equation systems and MIB matrices

A system of linear algebraic equations is formed after discretizing the elliptic equation

$$L_h u_h = f_h \tag{B.1}$$

where L_h is a real non-singular n by n matrix under grid spacing h , u_h is the numerical solution vector and f_h is the source term vector. The matrix L_h is viewed as a linear operator mapping \mathbb{R}^n into \mathbb{R}^n , the space \mathbb{R}^n being a linear space equipped with an inner-product (\cdot, \cdot) inducing a norm $\|\cdot\|$ defined as follows

$$(u, v) = \sum_{i=1}^n u_i v_i, \quad \|u\| = (u, u)^{1/2}, \quad \forall u, v \in \mathbb{R}^n.$$

where u_i represents the i -th component of the vector u .

Generally, systems of linear algebraic equations are commonly solved by using direct methods and iterative methods. Direct methods, such as Gaussian elimination, and LU decomposition work for general matrices with arbitrary structure but require large computer memory. Therefore, they are not computationally efficient and hence unsuitable for solving the 3D PB model of biomolecules, even for small proteins.

Some of the iterative methods such as Richardson, Jacobi, Gauss-Seidel and SOR iterations, also work well for general structured matrices but they are barely employed due to the reduced robustness for large protein system. The classic linear iteration methods for solving Eq.(B.1) can be viewed as the following form

$$u_h^{j+1} = u_h^j - BL_h u_h^j + Bf_h, \quad (\text{B.2})$$

where B is matrix approximating L_h^{-1} in some sense. Different construction of matrix B results in a different iterative method. The necessary and sufficient condition for the convergence of algorithm (B.2) is that the spectrum ρ of the error propagation operator must be smaller than 1, i.e., $E = I_h - BL_h$ and $\rho(E) < 1$ [115], where I_h is the identity operator associated with the grid resolution h . The smaller value of $\rho(E)$ indicates the better convergence of the method. The spectra of this family of iteration methods can be expressed as $\rho(E) = 1 - O(h^2)$, which implies that as grid spacing gets smaller, these methods converge more and more slowly. This property severely restricts the wide applications of these methods for large linear systems.

The conjugate gradient (CG) method is a very efficient iterative method if the matrix is symmetric and positive definite. Actually it is the main workhorse of most popular PBE solvers since the matrices from the standard FDMs or FEMs satisfy these good properties. The multigrid (MG) method is an accelerating technique and can be applied in combination with any of commonly used solvers. Using a hierarchy of discretizations, MG shifts the computation between coarser and finer grids by

extrapolation and restriction, and thus accelerates the convergence. It is almost the fastest accelerating technique known so far and applied in many popular PB solvers, such as the APBS.

Unfortunately, the matrix L_h from the MIB can barely take advantages from the described methods due to its complicated structure. For the discretization of the Laplace operator in the PBE by standard FDMs, each grid point except the boundary ones takes the following form:

$$\begin{aligned}
-\nabla \cdot (\epsilon \nabla u)|_{\mathbf{x}=x_i, y_j, z_k} &= c_0 u_{i,j,k} + c_1 u_{i-1,j,k} + c_2 u_{i+1,j,k} \\
&+ c_3 u_{i,j-1,k} + c_4 u_{i,j+1,k} \\
&+ c_5 u_{i,j,k-1} + c_6 u_{i,j,k+1}
\end{aligned} \tag{B.3}$$

where i, j, k represent the discretization indices along the x, y, z directions, respectively. The coefficients $c_m, m = 0, 1, \dots, 6$ only depend ϵ and grid spacing h . The symmetric structure of Eq. (B.3) and the facts $\sum_{m=0}^6 c_m = 0$ and $c_1 = c_2 = c_3 = c_4 = c_5 = c_6$ make the whole matrix symmetric and positive definite.

However, since the MIB scheme takes into account the interface treatment and at all the irregular grid points near the interface, discretizations are modified. For the simplest case, assume that only one fictitious point is needed and without the loss of generality, the modification is in the form:

$$\begin{aligned}
-\nabla \cdot (\epsilon \nabla u)|_{\mathbf{x}=x_i, y_j, z_k} &= c_0 u_{i,j,k} + c_1 f^* + c_2 u_{i+1,j,k} \\
&+ c_3 u_{i,j-1,k} + c_4 u_{i,j+1,k} \\
&+ c_5 u_{i,j,k-1} + c_6 u_{i,j,k+1}
\end{aligned} \tag{B.4}$$

Note that the fictitious value f^* is used in Eq. (B.4) for the smooth extension of the function. The fictitious value f^* can further be expanded as the linear combination

of the unknown function values.

$$f^* = \sum_{m=1}^M \tilde{c}_m u_{i'_m, j'_m, k'_m} \quad (\text{B.5})$$

where $u_{i'_m, j'_m, k'_m}$ is the nearby function values around $u_{i, j, k}$, $\tilde{c}_m, m = 1, 2, \dots, M$ are the corresponding coefficients. Usually $M = 10$ in second order MIB scheme for a smooth interface but could be bigger for interface with singularities. The choice of $u_{i'_m, j'_m, k'_m}$ and calculation of \tilde{c}_m totally depend on the local information of the interface. The introduction of the fictitious values gives high accuracy for the interface problems but also ruins the good properties such as symmetry and positive-definiteness of the overall matrix.

To solve the matrices generated from the MIB scheme, the direct methods and regular iterative methods will be ruled out from the beginning due to the poor convergence for huge systems. The CG method also does not work because the unpredictably general matrix structures. Meanwhile, the direct application of the multigrid method, which is an important accelerating technique, also has a potential problem due to the shift of irregular point locations during grid refinement cycles. Ref [30] showed the poor behavior of the algebraic multigrid method (AMD) and proposed a new multigrid scheme based on the local interface problem but the interpolation operator at the interface will cost much extra work.

Therefore, we put more emphasis on looking for suitable solvers and accelerating techniques in the Krylov subspace theory. Stabilized biconjugate gradient method (BiCG) and generalized minimal residual method (GMRES) are two examples in Krylov subspace methods, which deal with the general nonsingular matrix that does not have to be symmetric and positive definite. In the following section, the Krylov subspace (KS) methods, idea of preconditioners are briefly introduced.

B.2 The Krylov subspace theory and preconditioners

Suppose u^0 is an initial guess for the solution u in system

$$L_h u_h = f_h \quad (\text{B.6})$$

and defines the initial residual $r^0 = f - Lu^0$. For notation simplicity, the subscript h is dropped here. As shown in Ref. [100], the Krylov subspace can be derived from the following projection method. The m^{th} iteration $u^m, m = 1, 2, \dots$ is of the form

$$u^m \in r^0 + \mathcal{S}_m, \quad (\text{B.7})$$

where \mathcal{S}_m is some m -dimensional space, called the search space. Strictly speaking, Eq. (B.7) is an abused notation, it means that u^m can be decomposed as the residual r^0 and an element in space \mathcal{S}_m . Because of m degrees of freedom, a total of m constraints is required to make u^m unique. This is achieved by choosing an m -dimensional space \mathcal{C}_m , called the constraint space, and by requiring that the m^{th} residual is orthogonal to that space, i.e.,

$$r^m = f - Lu^m \in r^0 + L\mathcal{S}_m, \quad r^m \perp \mathcal{C}_m. \quad (\text{B.8})$$

Here the orthogonality is in the sense of the inner product in the Euclidean space.

If the space \mathcal{S}_m is defined as the Krylov subspace $\mathcal{K}_m(L, r^0)$, i.e.,

$$\mathcal{S}_m = \mathcal{K}_m(L, r^0) \equiv \text{span}\{r^0, Lr^0, \dots, L^{m-1}r^0\}, \quad m = 1, 2, \dots, \quad (\text{B.9})$$

then the projection method is the so-called Krylov subspace method. More specifi-

cally, if $\mathcal{C}_m = \mathcal{S}_m$, it is the Galerkin method, which includes the CG method and its generalizations, and if $\mathcal{C}_m = L\mathcal{S}_m$, it yields the GMRES. These are the basic idea of Krylov subspace methods.

For the convergence analysis, note that conditions (B.7) and (B.8) imply that the error $u - u^m$ and the residual r^m can be written in the polynomial form

$$u - u^m = p_m(L)(u - u^0), \quad r^m = p_m(L)r^0, \quad (\text{B.10})$$

where p_m is a polynomial of degree at most m and with value one at the origin. Ref. [100] gives the error bound for Krylov subspace methods

$$\frac{\|u - u^m\|}{\|u - u^0\|} \leq \min_{p \in \pi_m} \max_k |p(\lambda_k)|, \quad (\text{B.11})$$

where π_m denotes the set of polynomials of degree at most m and with value one at the origin, λ_k are the eigenvalues of the matrix L . It can be concluded from Eq. (B.11) that the convergence behavior of the Krylov subspace methods is completely determined by their spectra. However, as indicated in Ref. [100], it is always difficult to really evaluate the upper bound. Alternatively, it states that the condition number of the matrix is a criteria which although, only partially reveals the practice convergence behavior but is easier to calculate. For matrix L , the condition number is defined as the ratio of the extreme eigenvalues or spectra

$$\varsigma = \frac{\lambda_{\max}}{\lambda_{\min}}. \quad (\text{B.12})$$

Since the rate of the convergence of Krylov projection methods for a particular linear system is strongly dependent on its spectrum, preconditioner is typically used to alter the spectrum and hence accelerate the convergence rate of iterative techniques.

Preconditioner can be applied to system (B.6) by

$$\left(M_L^{-1}LM_R^{-1}\right)(M_Ru) = M_L^{-1}f, \quad (\text{B.13})$$

where M_L and M_R denote the left and right precondition matrices. Usually if $M_R = I$, the left preconditioned results and the residual is given by

$$r_L = M_L^{-1}f - M_L^{-1}Lu. \quad (\text{B.14})$$

Properly preconditioned matrix $M^{-1}L$ may significantly reduce the condition number of L , hence the rate of convergence is accelerated. The commonly used precondition strategies are Jacobi preconditioner, block preconditioner and incomplete LU factorization. However, preconditioning a large sparse system is an empirical exercise. Different preconditioners work better for different kinds of problems. In practice, the combinations of different Krylov subspace solvers and preconditioners are investigated, and the rate of convergence is analyzed via the spectra of preconditioned and un-preconditioned matrices.

Appendix C

A short user manual for the MIBPB package

C.1 Work flow of the MIBPB package

The MIBPB solver package incorporates with two packages to accomplish the electrostatic potential calculation. First, molecular structures are prepared via Python software package PDB2PQR (<http://pdb2pqr.sourceforge.net/>): it accomplishes many common tasks of preparing structures for continuum electrostatic calculations, such as adding a limited number of missing heavy atoms to biomolecular structures, determining side-chain pK_a s, placing missing hydrogens, etc.

Users can either submit the protein PDB index to the online server (<http://pdb2pqr.sourceforge.net/>) or download the executable file to prepare the molecular structure.

Once the molecular structure is prepared, the computational domain Ω will be automatically generated based on the coordinates of the protein atoms: first a smallest cuboid that contains the protein will be calculated and then each length of the cuboid is symmetrically extend at two ends by 5 to 10Å, depending on the protein size. This

strategy usually employed in many FDMs is verified to be reasonable in practices and also the extension of the cuboid can be customized easily. The larger size of Ω is of course closer to real biological situation. However, the solution of the PBE is not sensitive to this change while the computational cost will be increased.

Additionally, the geometry of the molecular surface used in the MIB scheme is generated by the MSMS (http://www.scripps.edu/~sanner/html/msms_home.html). Given the information of the coordinates and radius of each atom in the molecule, surfaces are generated at given water probe radius in a triangulation form. The intersection of each triangle with the meshing lines and the normal direction extracted from the surface information are key ingredients of the MIBPB scheme. For the MSMS parameter, the water molecule probe radius is recommended as 1.4 and the vertex density is 10. These parameters are enough to generate the molecular surface with good quality, various 3D Cartesian grid resolutions in current use can obtain necessary surface information under this setting.

There are two options for choosing KS solvers and preconditioners in solving MIBPB matrices. One is to use the SLATEC, which has been incorporated in our MIBPB package. The other way is to use the PETSc. According to our tests, the PETSc is generally more stable and reliable than the SLATEC, particularly for large proteins. It needs to be pre-installed by the user if one chooses the PETSc matrix acceleration option.

The current MIBPB package offers half stand-alone solvers in which users have to prepare the molecular structures and generate the surfaces on their own with desired parameters. The package also has one-step solvers which have integrated all the steps with default parameter settings. Either the half stand-alone or one-step solver is further classified into linear solver and nonlinear solver. Therefore, there are in total four executable MIBPB files in the package. Additionally, two other auxiliary small Perl scripts, the `pqr2xyzr.pl` and `dat2dx.pl` are included in the

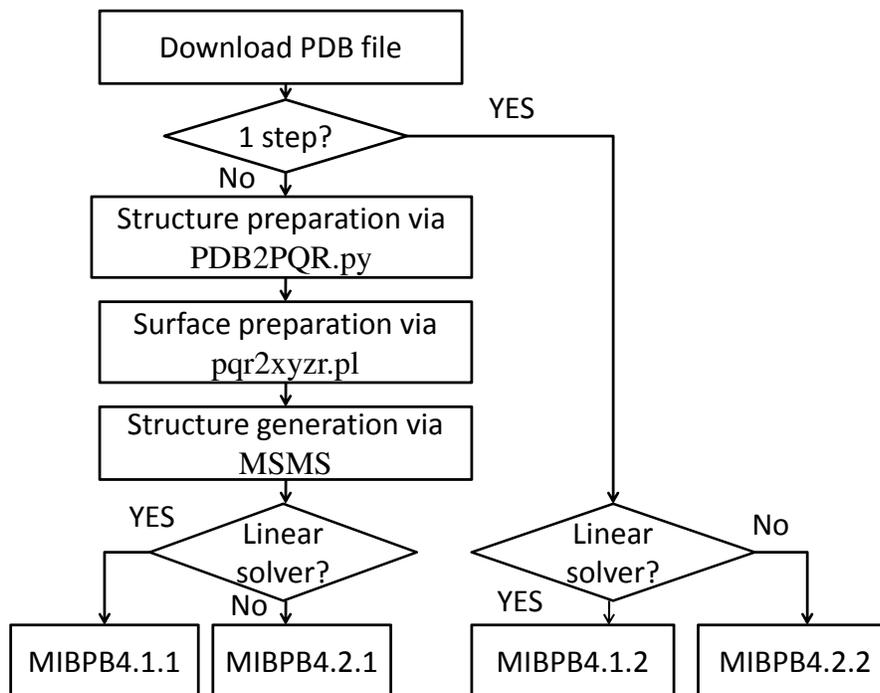


Figure C.1: Work flow of the MIBPB package

MIBPB package to accomplish the molecular surface preparation and ultimate data visualization. Figure (C.1) is the work flow of the MIBPB package usage. Users are referred to <http://www.math.msu.edu/~wei/MIBPB/> for detailed instructions.

For a clearer demonstration, we use one specific protein example to illustrate the procedure. Protein with ID 1ajj is assumed to have been downloaded from the Protein Data Bank and saved as file 1ajj.pdb.

1. Prepare the protein structure

- Input file: 1ajj.pdb
- Command line: `python pdb2pqr.py -ff=CHARMM 1ajj.pdb 1ajj_apbs.pqr`
- Output file: 1ajj_apbs.pqr.

- Remark: For full usage of `pdb2pqr.py`, users are referred to the corresponding link.

2. Molecular surface preparation

- Input file: `1ajj_apbs.pqr`
- Command line: `pqr2xyzr 1ajj`
- Output files: `1ajj.xyzr`, `1ajj.pqr`
- Remark: `1ajj.xyzr` file stores the coordinates and radii of the atoms in the protein, `1ajj.pqr` stores the coordinates and partial charges. They are necessary files for the MSMS to generate molecular surfaces.

3. Molecular surface generation

- Input files: `1ajj.xyzr`, `1ajj.pqr`
- Command line: `msms -if 1ajj.xyzr -prob 1.4 -de 10 -of 1ajj`
- Output files: `1ajj.vert`, `1ajj.face`. Now the molecular surface is generated in the triangulation form. The vertices and normal direction of each triangle are saved in these files.
- Remark: water probe radius and triangulation density are set as default values 1.4 and 10, respectively. They are adjustable parameters.

4. MIBPB implementation

- Linear solver: `mibpb4.1.1 1ajj eps1=1 eps2=80 h=0.5`
- Nonlinear solver: `mibpb4.2.1 1ajj eps1=1 eps2=80 kappa=1.0 h=0.5`
- Output file: `1ajj_pbe.dat`
- Remark: Above command lines give the standard format. Parameters are adjustable.

C.2 Work flow for the display of the surface electrostatic potential

After the electrostatic potential file is obtained by running the MIBPB solver, we can display it on the molecular surface by using the VMD (<http://www.ks.uiuc.edu/Research/vmd/>), a molecular visualization program. We are able to visualize the potential distribution on the surface by implementing a file transformation via the Perl script `dat2dx.pl`. Moreover, by taking the difference of surface electrostatic potentials under different grid resolutions h , we are also able to check the convergence of the solutions and therefore suggest a proper grid resolution for balancing high numerical accuracy and efficiency. The procedure is shown as the following.

1. Visualization file preparation.

- MIBPB package generates output file `[pdbname]_pbe.dat`, in which the electrostatic potentials on grid points of the protein-solvent system are stored. Before displaying the electrostatic potential on the molecular surface, one needs to use `dat2dx.pl` script to transform the data file to the `[pdbname].dx` file.

- For example, for protein `1ajj`, one gets `1ajj_pbe.dat` file from the MIBPB package. Then use the command: `dat2dx.pl 1ajj [dcel] [xleft] [xright] [yleft] [yright] [zleft] [zright]`

where `[dcel]` is the mesh size (we assume a uniform mesh). Here `[xleft]`, `[xright]`, `[yleft]`, `[yright]`, `[zleft]` and `[zright]` prescribe the span of computational domain in x, y, z direction, respectively. Here `xleft`, `xright`, `yleft`, `yright`, `zleft`, and `zright` should be the same as those used in calculating the potential.

2. Molecular surface drawing

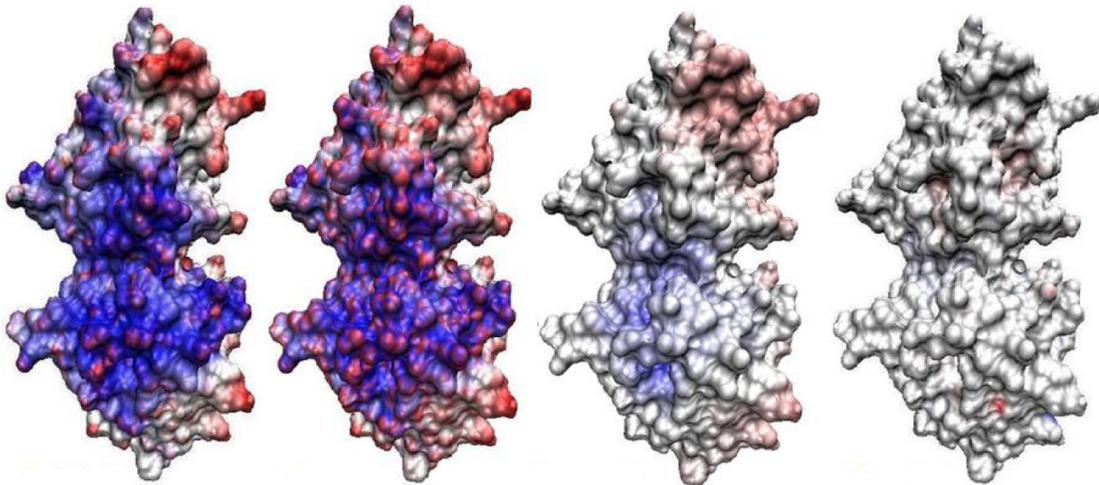


Figure C.2: Visualizations of surface electrostatic potentials of protein 1beb. From left to right: (a) Surface electrostatic potential ($I = 0, h = 1.0\text{\AA}$); (b): Surface electrostatic potential with the ionic strength $I = 1$ ($h = 1.0\text{\AA}$); (c): The difference of surface electrostatic potentials between in an ionic solvent ($I = 1$) and in a water solvent ($I = 0$); (d): The difference of surface electrostatic potentials in water ($I = 0$) between grid meshes $h = 1.0\text{\AA}$ and $h = 0.5\text{\AA}$.

- Load the PDB data file into the VMD
- Set drawing parameters in the Graphical Representation window: choose the “Volume” option for coloring method and the “Surf” option for drawing method.

3. Surface electrostatic potential drawing

- Load the [pdbname].dx format potential file into the VMD. In the Molecular File Browser window, load [pdbname].dx file for the same protein as that in molecular surface (instead of for new molecular).
- Set drawing parameters in the same Graphical Representation window as that in the second step. Choose the “Volume” option for coloring method and the “Surf” option for drawing method. Adjust the Color Scale Data Range to see different color effects.

Figure C.2 illustrates the visualization of electrostatic potential calculated from the MIBPB package, using protein 1beb as an example. The potentials calculated via both the linear MIBPB solver and the nonlinear MIBPB solver are plotted on the molecular surface via the VMD through the above procedure. Figure C.2(a) displays the potential distribution on the surface of protein 1beb when the solvent is water. In this case the linear MIBPB solver is implemented because $\bar{\kappa}$ is set as zero. While Fig. C.2(b) presents the potential distribution when $\bar{\kappa}^2 = 8.48$, in which case the nonlinear MIBPB solver is employed. These two calculations are carried out when grid resolution h is taken as 1.0\AA . Figure C.2(c) gives the difference of electrostatics in (a) and (b), from which the salt effect on electrostatic distribution may be observed. Figure C.2(d) reveals the potential difference in solvent when the calculations are under resolutions $h = 1.0\text{\AA}$ and 0.5\AA , i.e. the error $|\phi_h - \phi_{h/2}|$. It can be found that the error is almost zero around the molecular surface, this fact indicates that at $h = 1.0\text{\AA}$, the result is accurate enough so that reducing grid resolution to 0.5\AA does not give too much improvement. Mathematically speaking, the result is almost convergent between mesh size 1.0\AA and 0.5\AA , which is the recommended grid resolution range in the MIBPB package.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] N. Abaid, R. S. Eisenberg, and W. S. Liu. Asymptotic expansions of I-V relations via a Poisson-Nernst-Planck system. *SIAM Journal on Applied Dynamical Systems*, 7(4):1507–1526, 2008.
- [2] N. Agmon. The grotthuss mechanism. *Chem. Phys. Lett*, 244:456–462, 1995.
- [3] S. S. Ahmed, C. Ringhofer, and D. Vasileska. Parameter-free effective potential method for use in particle-based device simulations. *IEEE Tran. Nanotech.*, 4:456–471, 2008.
- [4] S. S. Ahmend and D. Vasileska. Modelling of narrow-width SOI devices. *Semicond. Sci. Technol*, 19:S131–133, 2004.
- [5] M. Akeson and D. Deamer. Proton conductance by the gramicidin water wire. Model for proton conductance in the F₀F₁ATPases? *Biophys J.*, 60:101–109, 1991.
- [6] M. G. Ancona and G. J. Iafrate. Quantum correction to the equation of state of an electron gas in a semiconductor. *Physical Review B*, 39:9536–9540, 1989.
- [7] C. R. Anderson. Efficient solution of the Schrödinger-Poisson equations in layered semiconductor devices. *J. Comput. Phys.*, 228:4745–4756, 2009.
- [8] P. Andrei and I. Mayergoyz. Analysis of fluctuations in semiconductor devices through self-consistent Poisson-Schrödinger computations. *J. Applied Physics*, 96:2071–2079, 2004.
- [9] A. Asenov. Random dopant induced threshold voltage lowering and fluctuations in Sub-0.1 μm MOSFET: A 3-D “atomistic” simulation study. *IEEE Transactions on Electron Devices*, 45:2505–2513, 1998.
- [10] A. E. Ayyadi and A. Jüngel. Semiconductor simulations using a coupled quantum drift-diffusion Schrödinger-Poisson model. *SIAM J. Appl. Math*, 66:544–572, 2005.

- [11] N. A. Baker. Poisson-Boltzmann methods for biomolecular electrostatics. *Methods in Enzymology*, 383:94–118, 2004.
- [12] N. A. Baker, D. Bashford, and D. A. Case. Implicit solvent electrostatics in biomolecular simulation. In B. Leimkuhler, C. Chipot, R. Elber, A. Laaksonen, A. Mark, T. Schlick, C. Schutte, and R. Skeel, editors, *New Algorithms for Macromolecular Simulation*. Springer, 2006.
- [13] S. Barraud. Quantization effects on the phonon-limited electron mobility in ultrathin SOI, sSOI and GeOI devices. *Semiconductor Science and Technology*, 22:413–417, 2007.
- [14] D. Bashford and D. A. Case. Generalized Born models of macromolecular solvation effects. *Annual Review of Physical Chemistry*, 51:129–152, 2000.
- [15] S. Bednarek, B. Szafran, and J. Adamowski. Solution of the Poisson-Schrödinger problem for a single-electron transistor. *Zeitschrift für Angewandte Mathematik und Physik*, 61:4461–4464, 2000.
- [16] D. Beglov and B. Roux. Solvation of complex molecules in a polar liquid: an integral equation theory. *Journal of Chemical Physics*, 104(21):8678–8689, 1996.
- [17] C. Bertonati, B. Honig, and E. Alexov. Poisson-Boltzmann calculations of non-specific salt effects on protein-protein binding free energy. *Biophysical Journal*, 92:1891–1899, 2007.
- [18] N. Blomberg, R. R. Gabdouliline, M. Nilges, and R. C. Wade. Classification of protein sequences by homology modeling and quantitative analysis of electrostatic similarity. *Proteins*, 37(3):379–387, 1999.
- [19] H. Borli, S. Kolberg, T. A. Fjeldly, and B. Iniguez. Precise modeling framework for short-channel double-gate and gate-all-around mosfets. *IEEE T Electron Devices*, 55:2678–2686, 2008.
- [20] J. Bothma, J. Gilmore, and R. McKenzie. The role of quantum effects in proton transfer reactions in enzymes: quantum tunneling in a noisy environment? *New Journal of Physics*, 12(055002), 2010.
- [21] S. Braun-Sand, A. Burykin, Z. T. Chu, and A. Warshel. Realistic simulations of proton transport along the gramicidin channel: Demonstrating the importance of solvation effects. *J. Phys. Chem. B*, 109:583–592, 2005.
- [22] F. M. Bufler, A. Schenk, and W. Fichtner. Efficient Monte Carlo device modeling. *IEEE T Electron Devices*, 47:1891–1897, 2000.
- [23] M. Burger, R. S. Eisenberg, and E. Heinz. Inverse problems related to ion channel selectivity. *SIAM J Applied Math*, 67(4):960–989, 2002.

- [24] A. E. Cardenas, R. D. Coalson, and M. G. Kurnikova. Three-dimensional Poisson-Nernst-Planck theory studies: Influence of membrane electrostatics on Gramicidin A channel conductance. *Biophysical Journal*, 79:80–93, July 2000.
- [25] D. M. Caughey and R. E. Thomas. Carrier mobilities in silicon empirically related to doping and field. *IEEE Proc.*, 55:2192–2193, 1967.
- [26] D. Chen, Z. Chen, C. Chen, W. H. Geng, and G. W. Wei. MIBPB: A software package for electrostatic analysis. *J. Comput. Chem.*, in press, 2010.
- [27] D. Chen and G. W. Wei. Modeling and simulation of electronic structure, material interface and random doping in nano-electronic devices. *J. Comput. Phys.*, 229:4431–4460, 2010.
- [28] H. Chen, Y. Wu, and G. Voth. Proton transport behavior through the influenza A M2 channel: Insights from molecular simulation. *Biophys J.*, 93:3470–3479, 2007.
- [29] J. Chen and C. L. Brooks, III. Implicit modeling of nonpolar solvation for simulating protein folding and conformational transitions. *Physical Chemistry Chemical Physics*, 10:471–81, 2008.
- [30] T. Chen and J. Strain. Piecewise-polynomial discretization and Krylov-accelerated multigrid for elliptic interface problems. *J. Comput. Phys.*, 16:7503–7542, 2008.
- [31] Z. Chen, N. A. Baker, and G. W. Wei. Differential geometry based solvation models I: Eulerian formulation. *J. Comput. Phys.*, 2010.
- [32] Z. Chen, N. A. Baker, and G. W. Wei. Differential geometry based solvation models II: Lagrangian formulation. *J. Math. Biol.*, 2010.
- [33] I.-L. Chern, J.-G. Liu, and W.-C. Weng. Accurate evaluation of electrostatics for macromolecules in solution. *Methods and Applications of Analysis*, 10(2):309–28, 2003.
- [34] A. Chernyshev and S. Cukierman. Proton transfer in gramicidin water wires in phospholipid bilayers: Attenuation by phosphoethanolamine. *Biophys J.*, 91:580–587, 1997.
- [35] M. Chiba, D. G. Fedorov, and K. Kitaura. Polarizable continuum model with the fragment molecular orbital-based time-dependent density functional theory. *Journal of Computational Chemistry*, 29:2667–2676, 2008.
- [36] S.-H. Chung, T. Allen, and S. Kuyucak. Conducting-state properties of the KcsA potassium channel from molecular and Brownian dynamics simulations. *Biophysical Journal*, 82:628–645, 2002.

- [37] S.-H. Chung and S. Kuyucak. Recent advances in ion channel research. *Biochimica et Biophysica Acta*, 1565:267–286, 2002.
- [38] B. Corry, S. Kuyucak, and S.-H. Chung. Test of Poisson-Nernst-Planck theory in ion channels. *The Rockefeller University Press*, 114(4):597–599, October 1999.
- [39] B. Corry, S. Kuyucak, and S.-H. Chung. Dielectric self-energy in Poisson-Boltzmann and Poisson-Nernst-Planck models of ion channels. *Biophysical Journal*, 84:3594–3606, June 2003.
- [40] Y. Cui, Z. Zhong, D. Wang, W. U. Wang, and C. M. Lieber. High performance silicon nanowire field effect transistors. *Nano Letters*, 3:149–152, 2003.
- [41] R. Cukier. Theory and simulation of proton-coupled electron transfer, hydrogen-atom transfer, and proton translocation in proteins. *Biochimica Et Biophysical Acta-Bioenergetics*, 1655:37–44, 2004.
- [42] S. Cukierman, E. P. Quigley, and D. S. Crumrine. Proton conduction in Gramicidin A and in its dioxolane-linked dimer in different lipid bilayers. *Biophys J.*, 73:2489–2502, 1997.
- [43] G. Curatola, G. Doornbos, J. Loo, Y. V. Ponomarev, and G. Iannaccone. Detailed modeling of sub-100-nm MOSFETs based on Schrödinger DD per sub-band and experiments and evaluation of the performance gap to ballistic transport. *IEEE Trans. Electron Devices*, 52:1851–1858, 2005.
- [44] S. Datta. Nanoscale device modeling: the Green’s function method. *Superlattices and Microstructures*, 28:253–278, 2000.
- [45] S. Datta. *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, 1995.
- [46] L. David, R. Luo, and M. K. Gilson. Comparison of generalized Born and Poisson models: Energetics and dynamics of HIV protease. *Journal of Computational Chemistry*, 21(4):295–309, 2000.
- [47] M. E. Davis and J. A. McCammon. Electrostatics in biomolecular structure and dynamics. *Chemical Reviews*, 94:509–21, 1990.
- [48] C. de Falco, E. Gatti, A. L. Lcacaíta, and R. Sacco. Quantum-corrected drift-diffusion models for transport in semiconductor devices. *J. Comput. Phys.*, 204:533–561, 2005.
- [49] C. de Falco, J. W. Jerome, and R. Sacco. A coupled Schrödinger drift-diffusion model for quantum semiconductor devices simulations. *J. Comput. Phys.*, 181:222–259, 1964.

- [50] C. de Falco, J. W. Jerome, and R. Sacco. A self-consistent iterative scheme for the one-dimensional steady-state transistor calculations. *IEEE Trans. Ele. Dev.*, 11:455–465, 1964.
- [51] C. de Falco, J. W. Jerome, and R. Sacco. Quantum-corrected drift-diffusion models: Solution fixed point map and finite element approximation. *J. Comput. Phys.*, 204:533–561, 2009.
- [52] T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, and N. A. Baker. PDB2PQR: An automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Research*, 32:W665–W667, 2004.
- [53] B. N. Dominy and C. L. Brooks, III. Development of a generalized Born model parameterization for proteins and nucleic acids. *Journal of Physical Chemistry B*, 103(18):3765–3773, 1999.
- [54] B. S. Doyle, S. Datta, M. Doczy, S. Harelend, B. J. Kavalieros, T. Linton, A. Murthy, R. Rios, and R. Chau. High performance fully-depleted tri-gate CMOS transistors. *IEEE Trans. Electron Devices*, 24:263–265, 2003.
- [55] X. F. Duan, Y. Huang, Y. Ciu, J. F. Wang, and C. M. Lieber. Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronics devices. *Nature*, 409:66–69, 2001.
- [56] W. Dyrka, A. T. Augousti, and M. Kotulska. Ion flux through membrane channels: An enhanced algorithm for the Poisson-Nernst-Planck model. *J. Comput Chem*, 29:1876–1888, 2008.
- [57] S. Edwards, B. Corry, S. Kuyucak, and S.-H. Chung. Continuum electrostatics fails to describe ion permeation in the gramicidin channel. *Biophysical Journal*, 83:1348–1360, September 2002.
- [58] B. Eisenberg. Ion channels as devices. *Journal of Computational Electronics*, 2:245–249, 2003.
- [59] G. Eisenman, B. Enos, J. Hagglund, and J. Sandbloom. Gramicidin as an example of a single-filing ion channel. *Ann. N.Y. Acad. Sci.*, 339:8–20, 1980.
- [60] M. Feig and I. Brooks, C. L. Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Curr Opin Struct Biol.*, 14:217 – 224, 2004.
- [61] M. Feig, W. Im, and I. Brooks, C. L. Implicit solvation based on generalized Born theory in different dielectric environments. *Journal of Chemical Physics*, 120(2):903–911, 2004.

- [62] M. Ferrier, R. Clerc, G. Ghibaudo, F. Boeuf, and T. Skotnicki. Analytical model for quantization on strained and unstrained bulk nMOSFET and its impact on quasi-ballistic current. *Solid- State Electronics*, 50:69–77, 2006.
- [63] G. Fiori, S. Di Pascoli, and G. Lannaccone. Three-dimensional simulations of quantum confinement and random dopants effects in nanoscale nmosfets. *J. Comput. Theor. nanoscience*, 5:1115–1119, 2008.
- [64] G. Fiori and G. Lannaccone. Three-dimensional simulation of one-dimensional transport in silicon nanowire transistors. *IEEE T Nanotechnology*, 6:524–529, 2007.
- [65] M. V. Fischetti. Master-equation approach to the study of electronic transport in small semiconductor devices. *Phys. Rev. B*, 59:4901–4917, 1999.
- [66] D. J. Frank, Y. Taur, and H. S. P. Wong. Generalized scale length for two-dimensional effects in MOSFETs. *IEEE Electron Dev. Lett.*, 10:385–387, 1998.
- [67] W. Geng, S. Yu, and G. W. Wei. Treatment of charge singularities in implicit solvent models. *Journal of Chemical Physics*, 127:114106, 2007.
- [68] M. K. Gilson, M. E. Davis, B. A. Luty, and J. A. McCammon. Computation of electrostatic forces on solvated molecules using the Poisson-Boltzmann equation. *Journal of Physical Chemistry*, 97(14):3591–3600, 1993.
- [69] P. Graf, M. G. Kurnikova, R. D. Coalson, and A. Nitzan. Comparison of dynamic lattice Monte Carlo simulations and the dielectric self-energy Poisson-Nernst-Planck continuum theory for model ion channels. *J. Phys. Chem. B*, 108:2006–2015, 2004.
- [70] W. J. Gross, D. Vasileska, and D. K. Ferry. A novel approach for introducing the electron-electron and electron-impurity interactions in partical-based simulations. *IEEE Electron Device Letters*, 20(9):463–465, 1999.
- [71] W. J. Gross, D. Vasileska, and D. K. Ferry. Ultrasmall MOSFETs: The Importance of the Full Coulomb Interaction on Device Characteristics. *IEEE Transactions on Electron Devices*, 47(10):1831–1837, 2000.
- [72] W. J. Gross, D. Vasileska, and D. K. Ferry. Three-dimensional simulations of ultrasmall metal-oxide-semiconductor field-effect transistors: The role of the discrete impurities on the device terminal characteristics. *Journal of Applied Physics*, 91(6):3737–3740, 2002.
- [73] Z. Y. Han, N. Goldsman, and C. K. Lin. Incorporation of quantum corrections to semiclassical two-dimensional device modeling with the Wigner-Boltzmann equation. *Solid-State Electronics*, 49:145–154, 2005.

- [74] U. Harbola and S. Mukamel. Superoperator nonequilibrium Green's function theory of many-body systems; applications to charge transfer and transport in open junctions. *Physics Report - Review Section of Physics Letters*, 465:191–222, 2008.
- [75] B. Hille. *Ionic Channels of Excitable Membranes*. Sunderland, MA: Sinauer Associates, 1992.
- [76] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–871, 1964.
- [77] U. Hollerbach, D. P. Chen, and R. S. Eisenberg. Two- and three-dimensional Poisson-Nernst-Planck simulations of current flow through Gramicidin A. *Journal of Scientific Computing*, 16(4):373–409, Decemeber 2002.
- [78] M. J. Holst. *Multilevel Methods for the Poisson-Boltzmann Equation*. University of Illinois at Urbana-Champaign, Numerical Computing Group, Urbana-Champaign, 1993.
- [79] H. Hu, Z. Y. Lu, M. Elstner, J. Hermans, and W. T. Yang. Simulating water with the self-consistent-charge density functional tight binding method: From molecular clusters to the liquid state. *Journal of Physical Chemistry A*, 111:5685–5691, 2007.
- [80] H. Hwang, G. C. Schatz, and M. A. Ratner. Ion current calculations based on three dimensional Poisson-Nernst-Planck theory for a cyclic peptide nanotube. *Journal of Physical Chemistry B*, 110:6999–7008, 2006.
- [81] R. Improta, V. Barone, G. Scalmani, and M. J. Frisch. A state-specific polarizable continuum model time dependent density functional theory method for excited state calculations in solution. *Journal of Chemical Physics*, 125(054103), 2006.
- [82] H. Ishikuro and T. Hiramoto. Hopping transport in multiple-dot silicon single electron MOSFET. *Solid-State Electronics*, 42:1425–1428, 1998.
- [83] H. Y. Jiang, S. H. Shao, W. Cai, and P. W. Zhang. Boundary treatments in non-equilibrium Green's function (NEGF) methods for quantum transport in nano-MOSFETs. *J Comput. Physics*, 227:6553–6573, 2008.
- [84] X. Jiang, H. X. Deng, J. W. Luo, S. S. Li, and L. W. Wang. A fully three-dimensional atomistic quantum mechanical study on random dopant-induced effects in 25-nm MOSFETs. *IEEE T Electron Devices*, 55:1720–1726, 2008.
- [85] Y. H. Jiang and W. Cai. Effect of boundary treatments on quantum transport current in the Green's function and Wigner distribution methods for a nano-scale DG-MOSFET. *J. Comput. Phys.*, Submitted, 2009.

- [86] 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. *J. Appl. Phys.*, 99:123719–10, 2006.
- [87] B. Jogai. Influence of surface states on the two-dimensional electron gas in algalan/gan heterojunction field-effect transistors. *J. Applied Physics*, 93:1631–1635, 2003.
- [88] Y. W. Jung, B. Z. Lu, and M. Mascagni. A computational study of ion conductance in the KcsA K⁺ channel using a Nernst-Planck model with explicit resident ions. *J. Chem. Phys.*, 131(215101), 2009.
- [89] L. Kadanoff and G. Baym. *Quantum Statistical Mechanics*. 1962.
- [90] H. R. Khan, D. Vasileska, and S. S. Ahmed. Modeling of FinFet: 3D MC Simulation Using FMM and Unintentional Doping Effects on Device Operation. *Journal of Computational Electronics*, 3:337–340, 2004.
- [91] T. Khan, D. Vasileska, and T. J. Thornton. Effect of interface roughness on silicon-on-insulator-metal-semiconductor field-effect transistor mobility and the device low-power high-frequency operation. *J. Vac. Sci. Technol. B*, 23(4):1782–1784, 2005.
- [92] B. Kim, W. Kwon, C. Baek, S. Jin, Y. Song, and D. M. Kim. Three-Dimensional Simulation of Dopant-Fluctuation-Induced Threshold Voltage Dispersion in Nonplanar MOS Structures Targeting Flash EEPROM Transistors. *IEEE T Electron Devices*, 55:1456–1463, 2008.
- [93] W. Kohn and L. J. Sham. Self consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133A1138, 1965.
- [94] L. Krishtalik. The mechanism of the proton transfer: An outline. *Biochimica Et Biophysical Acta-Bioenergetics*, 1458:6–27, 2000.
- [95] S. Kuyucak, O. S. Andersen, and S.-H. Chung. Models of permeation in ion channels. *Rep. Prog. Phys.*, 64:1427–1472, 2001.
- [96] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic. Single and multiband modeling of quantum electron transport through layered semiconductor devices. *Journal of Applied Physics*, 81:7845–7869, 1997.
- [97] G. Lamm. The Poisson-Boltzmann equation. In K. B. Lipkowitz, R. Larter, and T. R. Cundari, editors, *Reviews in Computational Chemistry*, pages 147–366. John Wiley and Sons, Inc., Hoboken, N.J., 2003.
- [98] F. G. Lether. Analytical expansion and numerical approximation of the fermi-dirac integrals of order $-1/2$ and $1/2$. *J. Sci. Comput.*, 4:479–497, 2000.

- [99] G. Li and N. R. Aluru. Hybrid techniques for electrostatic analysis of nanoelectromechanical systems. *J. Applied Physics*, 96:2221–2231, 2004.
- [100] J. Liesen and P. Tichy. Convergence analysis of Krylov subspace methods. *CGAMM-Mitteilungen*, 27:153–173, 2004.
- [101] H. Y. Liu, M. Elstner, E. Kaxiras, T. Frauenheim, J. Hermans, and W. T. Yang. Quantum mechanics simulation of protein dynamics on long timescale. *Proteins-Structure Function and Genetics*, 44:484–489, 2001.
- [102] D. R. Livesay, P. Jambeck, A. Rojnuckarin, and S. Subramaniam. Conservation of electrostatic properties within enzyme families and superfamilies. *Biochemistry*, 42(12):3464–3473, 2003.
- [103] X. Loussier, D. Munteanu, J. L. Autran, and O. Tintori. Impact of Geometrical and Electrical Parameters on Speed Performance Characteristics in Ultimate Double-Gate Metal-Oxide-Semiconductor Field-Effect Transistors. *Japanese J Applied Physics*, 47:3390–3395, 2008.
- [104] J. H. Luscombe, A. M. Bouchard, and M. Luban. Electron confinement in quantum nanostructure - Self-consistent Poisson-Schrödinger theory. *Physical Review B*, 46:10262–10268, 1992.
- [105] J. MacKerell, A. D., D. Bashford, M. Bellot, J. Dunbrack, R. L., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. Reiher, W. E., B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, 102(18):3586–3616, 1998.
- [106] K. Majumdera and N. Bhat. Band structure effects in ultra-thin-body double-gate field effect transistor: A fullband analysis. *J Applied Physics*, 103:114503, 2008.
- [107] A. B. Mamonov, R. D. Coalson, A. Nitzan, and M. G. Kurnikova. The role of the dielectric barrier in narrow biological channels: A novel composite approach to modeling single-channel currents. *Biophysical Journal*, 84:3646–3661, June 2003.
- [108] A. Martinez, J. R. Barker, A. Svizhenko, M. P. Anantram, and A. Asenov. The impact of random dopant aggregation in source and drain on the performance of ballistic DG Nano-MOSFETs: A NEGF study. *IEEE Transactions on Nanotechnology*, 6:438–445, 2007.
- [109] W. R. McKinnon. Fokker-Planck approach to extending the one-flux method of carrier transport in semiconductors to variable energies. *J. Applied Phys.*, 94:4986–4994, 2003.

- [110] J. Mongan, C. Simmerling, J. A. McCammon, D. A. Case, and A. Onufriev. Generalized Born model with a simple, robust molecular volume correction. *Journal of Chemical Theory and Computation*, 3(1):159–69, 2007.
- [111] B. Moy, B. Corry, S. Kuyucak, and S.-H. Chung. Tests of continuum theories as models of ion channels. I. Poisson-Boltzmann theory versus Brownian dynamics. *Biophys. J.*, 78:2349–2363, 2000.
- [112] J. Nagle and H. Morowitz. Molecular mechanisms for proton transport in membranes. *Proc. Natl. Acad. Sci. U.S.A.*, 1458(72):298–302, 1978.
- [113] K. Natori. Ballistic metal-oxide-semiconductor field effect transistor. *J. Appl. Phys.*, pages 4879–4890, 1994.
- [114] A. Onufriev, D. Bashford, and D. A. Case. Modification of the generalized Born model suitable for macromolecules. *Journal of Physical Chemistry B*, 104(15):3712–3720, 2000.
- [115] J. M. Ortega. In *Numerical Analysis: A second course*. Academic Press, NY, 1999.
- [116] J. T. Park and J. P. Colinge. Multiple-gate SOI MOSFETs: Device design guidelines. *IEEE Trans. Electron Devices*, 49:2222–2229, 2002.
- [117] R. Parr and W. Yang. *Density Functional Theory of Atoms and Molecules*. 1989.
- [118] D. Perlman, D. Case, J. Caldwell, W. Ross, T. Cheatham, S. Debolt, D. Ferguson, G. Seibel, and P. Kollman. Amber, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.*, 91:1–41, 1995.
- [119] L. Pisani and G. Siciliano. Note on a Schrödinger-Poisson system in a bounded domain. *App. Math. Lett.*, 21:521–528, 2008.
- [120] E. Polizzi and N. Ben Abdallah. Subband decomposition approach for the simulation of quantum electron transport in nanostructures. *J. Comput. Phys.*, 202:150–180, 2005.
- [121] R. Pomes and B. Roux. Theoretical study of H⁺ translocation along a model proton wire. *J. Phys. Chem*, 100:2519–2527, 1996.
- [122] R. Pomes and B. Roux. Molecular mechanism of H⁺ conduction in the single-file water chain of the gramicidin channel. *Biophysical Journal*, 82:2304–2316, 2002.

- [123] R. Pomes and B. Roux. Structure and Dynamics of a Proton Wire: A Theoretical Study of H^+ Translocation along the Single-File Water Chain in the Gramicidin A Channel. *Biophysical Journal*, 71:19–39, 2002.
- [124] Z. Ren, R. Venugopal, S. Goasquen, S. Datta, and M. S. Lundstrom. nanoMOS 2.5: A two-dimensional simulator for quantum transport in double-gate MOS-FETs. *IEEE Trans. Elec. Dev.*, 50:1914 – 1925, 2003.
- [125] B. Roux. Influence of the membrane potential on the free energy of an intrinsic protein. *Biophysical Journal*, 73:2980–2989, December 1997.
- [126] B. Roux. Computational studies of the gramicidin channel. *Acc. Chem. Res.*, 35:366–375, 2002.
- [127] B. Roux, T. Allen, S. Berneche, and W. Im. Theoretical and computational models of biological ionchannels. *Quarterly Reviews of Biophysics*, 7(1):1–103, 2004.
- [128] B. Roux and T. Simonson. Implicit solvent models. *Biophysical Chemistry*, 78(1-2):1–20, 1999.
- [129] T. Saito, T. Saraya, T. Inukai, H. Majimi, T. Nangumo, and T. Hiramoto. Suppression of short channel effect in triangular parallel wire channel mosfets. *IEICE Trans. Electron*, E85C:1073 – 1080, 2002.
- [130] K. Sakurai, M. Oobatake, and Y. Goto. Salt-dependent monomer-dimer equilibrium of bovine β -lactoglobulin at pH 3. *Protein Sci*, 10:2325–2335, 2001.
- [131] M. F. Sanner, A. J. Olson, and J. C. Spohner. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers*, 38:305–320, 1996.
- [132] M. Saraniti, S. Aboud, and R. Eisenberg. The simulation of ionic charge transport in biological ion channels: an introduction to numerical methods. *Reviews in Computational Chemistry*, 22:229–294, 2006.
- [133] S. Scaldaferrri, G. Curatola, and G. Iannaccone. Direct solution of the Boltzmann transport equation and Poisson-Schrödinger equation for nanoscale MOS-FETs. *IEEE T. Electron Devices*, 54:2901–2909, 2007.
- [134] J. R. Schnell and J. J. Chou. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature*, 451:591–596, January 2008.
- [135] M. F. Schumaker, R. Pomes, and B. Roux. A combined molecular dynamics and diffusion model of single proton conduction through gramicidin. *Biophysical Journal*, 79:2840–2857, December 2000.
- [136] J. Schwinger. Brownian motion of a quantum oscillator. *J. Math. Phys.*, 2:407–432, 1961.

- [137] S. H. Shao, W. Cai, and H. Z. Tang. Accurate calculation of green's function of the schrodinger equation in a block layered potential. *J Comput. Physics*, 219:733–748, 2006.
- [138] K. A. Sharp and B. Honig. Electrostatic interactions in macromolecules - theory and applications. *Annual Review of Biophysics and Biophysical Chemistry*, 19:301–332, 1990.
- [139] T. Simonson. Macromolecular electrostatics: continuum models and their growing pains. *Current Opinion in Structural Biology*, 11(2):243–252, 2001.
- [140] A. Singer, D. Gillespie, J. Norbury, and R. Eisenberg. Singular perturbation analysis of the steady state Poisson-Nernst-Planck system: Applications to ion channels. *European Journal of Applied Mathematics*, 19:541–560, 2008.
- [141] J. C. Slater. A simplification of the Hartree-Fock method. *Phys. Rev.*, 81:385–390, 1951.
- [142] R. F. Snider. Quantum-mechanical modified boltzmann equation for degenerate internal states. *J. Chem. Phys.*, 32:1051–1060, 1960.
- [143] A. Svizhenko, M. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal. Two-dimensional quantum mechanical modeling of nanotransistors. *J Applied Phys.*, 91:2343–2354, 2002.
- [144] M. S. Till, T. Essigke, T. Becker, and G. M. Ullmann. Simulating the proton transfer in Gramicidin A by a sequential dynamical Monte Carlo method. *J. Phys. Chem*, 112:13401–13410, 2008.
- [145] J. Tomasi, B. Mennucci, and R. Cammi. Quantum mechanical continuum solvation models. *Chem. Rev.*, 105:2999–3093, 2005.
- [146] A. Trellakis, A. T. Galick, A. Pacelli, and U. Ravaioli. Iteration scheme for the solution of the two-dimensional Schrödinger-Poisson equations in quantum structures. *J. Appl. Phys.*, 81:7880–7884, 1997.
- [147] V. Tsui and D. A. Case. Molecular dynamics simulations of nucleic acids with a generalized Born solvation model. *Journal of the American Chemical Society*, 122(11):2489–2498, 2000.
- [148] V. Tsui and D. A. Case. Calculations of the absolute free energies of binding between RNA and metal ions using molecular dynamics simulations and continuum electrostatics. *Journal of Physical Chemistry B*, 105(45):11314–11325, 2001.
- [149] F. Venturi, R. Smith, E. C. Sangiorgi, M. R. Pinto, and B. Ricco. A general purpose device simulator coupling Poisson and Monte Carlo transport with applications to deep submicron MOSFETs. *IEEE Trans. Computer-Aided Design*, 8:360, 1989.

- [150] R. Venugopal, Z. Ren, S. Datta, and M. S. Lundstrom. Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches. *J. App. Phys.*, 92:3730–3739, 2002.
- [151] L. Waldmann. Die Boltzmann-Gleichung für Gase mit rotierenden Molekülen. *Z. Naturforsch. Teil A*, 12:660–662, 1957.
- [152] R. Wallis, G. R. Moore, R. James, and C. Kleanthous. Protein-Protein Interactions in Colicin E9 DNase-Immunity Protein Complexes. 1. Diffusion-Controlled Association and Femtomolar Binding for the Cognate Complex. *Biochemistry*, 34:13743–13750, 1995.
- [153] J. Wang, E. Polizzi, and M. Lundstrom. A three-dimensional quantum simulation of silicon nanowire transistors with the effective-mass approximation. *J. Appl. Phys.*, 96:2192–203, 2004.
- [154] G. W. Wei. Differential geometry based multiscale models. *Bulletin of Mathematical Biology*, 72:1562 – 1622, 2010.
- [155] H. S. P. Wong, Y. Tuar, and D. J. Frank. Microelectronics reliability. *SIAM J. Appl. Math.*, 38:1447–1456, 1998.
- [156] W. T. Yang and T. S. Lee. A density matrix divide and conquer approach for electronic structure calculations of large molecules. *J. Chem. Phys.*, 103:5674–5678, 1995.
- [157] S. Yu, W. Geng, and G. W. Wei. Treatment of geometric singularities in implicit solvent models. *Journal of Chemical Physics*, 126:244108, 2007.
- [158] S. Yu and G. W. Wei. Three-dimensional matched interface and boundary (MIB) method for treating geometric singularities. *Journal of Computational Physics*, 227:602–632, 2007.
- [159] S. Yu, Y. Zhou, and G. W. Wei. Matched interface and boundary (MIB) method for elliptic problems with sharp-edged interfaces. *Journal of Computational Physics*, 224(2):729–756, 2007.
- [160] S. Zhao and G. W. Wei. High-order FDTD methods via derivative matching for Maxwell’s equations with material interfaces. *Journal of Computational Physics*, 200(1):60–103, 2004.
- [161] Q. Zheng, D. Chen, and G. W. Wei. Second-order Poisson-Nernst-Planck solver for ion channel transport. *Journal of Comput. Phys.*, 2010.
- [162] Y. Zheng, C. Rivas, R. Lake, K. T. Alam, B. Boykin, and G. Klimeck. Electronic properties of silicon nanowires. *IEEE Trans. Electron Devices*, 52:1097–1103, 2005.

- [163] Y. C. Zhou, M. Feig, and G. W. Wei. Highly accurate biomolecular electrostatics in continuum dielectric environments. *Journal of Computational Chemistry*, 29:87–97, 2008.
- [164] Y. C. Zhou and G. W. Wei. On the fictitious-domain and interpolation formulations of the matched interface and boundary (MIB) method. *Journal of Computational Physics*, 219(1):228–246, 2006.
- [165] Y. C. Zhou, S. Zhao, M. Feig, and G. W. Wei. High order matched interface and boundary method for elliptic equations with discontinuous coefficients and singular sources. *Journal of Computational Physics*, 213(1):1–30, 2006.
- [166] J. Zhu, E. Alexov, and B. Honig. Comparative study of generalized Born models: Born radii and peptide folding. *Journal of Physical Chemistry B*, 109(7):3008–22, 2005.