ABSTRACT

Title of dissertation: SINGLE ELECTRON TRANSISTOR IN PURE SILICON

Binhui Hu, Doctor of Philosophy, 2009

Dissertation directed by: Professor Chia-Hung Yang, Department of Electrical and Computer Engineering

As promising candidates for spin qubits, semiconductor quantum dots (QDs) have attracted tremendous research efforts. Currently most advanced progress is from GaAs QDs. Compared to GaAs, lateral QDs in ^{28}silicon are expected to have a spin coherence time orders of magnitude longer, because ^{28}Si has zero nuclear spin, and there is no hyperfine interaction between electron spins and nuclear spins.

We have developed enhancement-mode metal-oxide-semiconductor (MOS) single electron transistors (SETs) using pure silicon wafers with a bi-layer gated configuration. In an MOS-SET, the top gate is used to induce a two-dimensional electron gas (2DEG), just as in an MOS field effect transistor. The side gates deplete the 2DEG into a QD and two point-contact channels; one connects the QD to the source reservoir, and the other connects the QD to the drain reservoir.

We have systematically investigated the MOS-SETs at 4.2 K, and separately in a dilution refrigerator with a base temperature of 10 mK. The data show that there is an intrinsic QD in each point-contact channel due to the local potential fluctuations in these SETs. However, after scaling down the SETs, we have found that the

intrinsic QDs can be removed and the electrostatically defined dots dominate the device behavior, but these devices currently only work in the many-electron regime. In order to realize single electron confinement, it is necessary to continue scaling down the device and improving the interface quality.

To explore the spin dynamics in silicon, we have investigated a single intrinsic QD by applying a magnetic field perpendicular to the sample surface. The magnetic field dependence of the ground-state and excited-state energy levels of the QD mostly can be explained by the Zeeman effect, with no obvious orbital effect up to 9 T. The two-electron singlet-triplet (ST) transition is first time directly observed in a silicon QD by excitation spectroscopy. In this ST transition, electron-electron Coulomb interaction plays a significant role. The observed amplitude spectrum suggests the spin blockade effect. When the two-electron system forms a singlet state in the dot at low fields, and the injection current from the lead becomes spin-down polarized, the tunneling conductance is reduced by a factor of 8. At higher magnetic fields, due to the ST transition, the spin blockade effect is lifted and the conductance is fully recovered.

SINGLE ELECTRON TRANSISTOR IN PURE SILICON

by

Binhui Hu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:
Professor Chia-Hung Yang, Chair/Advisor
Dr. Bruce E. Kane
Professor James R. Anderson
Professor Ping-Tong Ho
Professor John Melngailis

Dedication

To my parents

for their love and support.

Acknowledgments

Everything has a beginning, has an end. After six and a half years, now it is an end.

When I look back the road that I have gone through, the first one I want to thank is my advisor, Chia-Huang. His passion, knowledge and wisdom are my great resources. Whenever I have a question, he is right there to help. Without his guidance, it would be difficult to go though all these years, and to reach this point. It has been my pleasure to work with him and to learn from him.

I would like to thank Ming Jey for her help on device bonding and characterization. Her skills have been my goal. I also want to thank my previous labmates, especially Greg for his patience to teach me the device processing.

My special thanks go to Dr. Bruce Kane for his financial support and research advice in the last two years. I do not think that it is possible to have this thesis without his continuous support. His advice always reminds me the possibility beyond my knowledge, and the discussion with him is fruitful.

In addition, I want to thank Luyan. He has been a reliable friend, from discussing on research topics, to having a dinner together. Now it is difficult to do these things any more. I wish he had a wonderful and successful career at Yale. I would like to thank all of my friends. We have so much fun together. Without them, my life would lose color.

Finally, I want to thank my parents and my elder brother. Without their support from their hearts, I don't know whether I can finish this journey or not.

I owe a lot to my parents for not being with them these years, and they never complain anything, but provide me with all of their love.

Table of Contents

List of Figures vii

List of Symbols and Abbreviations xi

1 Introduction 1
1.1 Motivation . . . 1
1.2 Silicon quantum dots for quantum computing . . . 3
1.3 Outline of this thesis . . . 7

2 Theories of single electron transistors and double quantum dot systems 9
2.1 Single electron transistors (SETs) . . . 9
2.1.1 Operating principle of an SET . . . 9
2.1.2 Amplitude and lineshape of Coulomb oscillations . . . 17
2.1.3 Quantum interference in a closed dot and in an open dot . . . 20
2.1.4 Excited states and excitation spectrum . . . 25
2.1.5 Single-particle states in a two-dimensional elliptic harmonic oscillator . . . 29
2.2 Double quantum dot systems . . . 30
2.2.1 Stability diagram of two serial quantum dots . . . 30
2.2.2 Tunnel coupling between two dots . . . 38

3 Device structure and fabrication process 40
3.1 Device structure . . . 40
3.2 Fabrication process . . . 43
3.2.1 Ohmic contacts by ion implantation . . . 43
3.2.2 Trench isolation . . . 47
3.2.3 High density PECVD oxide . . . 49
3.2.4 All e-beam defined side gates . . . 50

4 Silicon metal-oxide-semiconductor field-effect transistors 52
4.1 Introduction . . . 52
4.2 Metal-oxide-semiconductor(MOS) capacitors . . . 53
4.2.1 MOS capacitors at room temperature . . . 53
4.2.2 MOS capacitors at low temperatures . . . 56
4.3 Oxide charges and interface trapped charges . . . 60
4.3.1 Capacitance method (C-V method) . . . 61
4.3.2 Experimental MOS C-V curves . . . 66
4.3.3 Conductance method . . . 72
4.4 2D electron concentration and electron mobility . . . 76
4.5 Summary . . . 80

5 Silicon SET transport characteristics 82
5.1 Introduction . . . 82
5.2 Measurement setup . . . 83
5.3 Transport characteristics at 4.2 K . . . 85
5.3.1 Functions of side gates . . . 85
5.3.2 Stability charts of silicon SETs . . . 86
5.3.3 Double quantum dot like behavior . . . 90
5.4 Transport characteristics in a dilution refrigerator . . . 90
5.4.1 Functions of side gates . . . 92
5.4.2 Multiple quantum dot behavior . . . 92
5.4.3 Stability chart of an intrinsic quantum dot in silicon . . . 96
5.5 Towards an electrostatically well defined quantum dot in silicon . . . 99
5.5.1 Removal of intrinsic quantum dots . . . 100
5.5.2 Functions of side gates . . . 103
5.5.3 The electrostatically defined quantum dot . . . 104
5.6 Origin of intrinsic quantum dots . . . 107
5.7 Coupled double quantum dots in silicon . . . 110
5.8 Summary . . . 114

6 Excited states and magnetic field spectra of a single quantum dot 117
6.1 Introduction . . . 117
6.2 Singlet-triplet transition in a silicon SET . . . 118
6.3 Electron spin blockade in a silicon SET . . . 127
6.4 Further discussion . . . 134
6.4.1 Orbital effect . . . 134
6.4.2 Spin polarization in silicon . . . 140
6.5 Quantum interference and false excited states . . . 144
6.6 Summary . . . 152

7 Conclusion and future work 154
7.1 Conclusion . . . 154
7.2 Future work . . . 157

A Measurement system and techniques 160
A.1 Introduction . . . 160
A.2 Dilution refrigerator . . . 160
A.3 DC measurement system . . . 162
A.4 AC measurement system . . . 162
A.4.1 Basic setup of the AC measurement system . . . 162
A.4.2 Filtering . . . 166
A.4.3 Transimpedance amplifier (pre-amp) . . . 171

B AC measurement system 179

C Recipe for silicon MOS-SETs 181

List of Figures

1.1 The quantum dot quantum computer architecture 2

1.2 Schematic of an MOS-SET 6

2.1 Schematic of a single electron transistor 10

2.2 Coulomb blockade oscillations and diamond chart 16

2.3 Lineshape of Coulomb oscillations 19

2.4 Potential profile of a rectangular double barrier 21

2.5 Transmission coefficient of a rectangular double barrier 24

2.6 Electrons transport through excited states 26

2.7 Schematic of an excitation spectrum. 27

2.8 Schematic of a serial double quantum dot system 31

2.9 Stability diagrams of a double quantum dot system 35

2.10 Coulomb peak spacings in a stability diagram 37

2.11 Stability Diagram with tunnel coupling 39

3.1 Schematic of an MOS-SET 41

3.2 Electric potential profile of an MOS-FET 42

3.3 Concentration of phosphorous dopants after ion implantation 44

3.4 Diffusion of phosphorous dopants after annealing 46

3.5 Trench isolation 47

3.6 Side gate fabrication process 51

4.1 Energy band diagram of a p-type MOS capacitor 53

4.2 Charges and their locations in thermally oxidized silicon 60

4.3 C-V curves of a p-type device wafer 67

4.4 C-V curves of an n-type device wafer 69
4.5 C-V curves of an n-type device wafer with HDPECVD oxide 70
4.6 Equivalent circuits for conductance method 72
4.7 Conductance curves on a p-type device wafer 75
4.8 Basic setup for hall measurements 76
4.9 2D electron concentration v.s. gate voltage 78
4.10 Shubnikov-de Haas oscillations of the 2DEG 78
4.11 Electron mobility v.s. gate voltage 80

5.1 Schematic of the AC measurement system 84
5.2 Functions of side gates at 4.2 K 85
5.3 Stability chart of an SET at 4.2 K 87
5.4 Stability chart of an intrinsic QD at 4.2 K 88
5.5 Double QD like behavior at 4.2 K 89
5.6 Functions of side gates in a DR 91
5.7 Double QD like behavior in a DR 93
5.8 Single intrinsic QD behavior in a DR 94
5.9 Equivalent circuit and schematic potential profile of an MOS-SET . . 97
5.10 Stability chart of an intrinsic QD in a DR 98
5.11 Schematic of the device scale-down 100
5.12 Schematic of the intrinsic QD removal 101
5.13 Removal of the intrinsic QD 102
5.14 Functions of side gates of the scaled-down device 103
5.15 Behavior of an electrostatically defined QD 105
5.16 Stability chart of an electrostatically defined QD 106

5.17 Stability chart of another intrinsic QD 108
5.18 Intrinsic QD due to trapped charges 109
5.19 Schematic of coupled double dots in silicon 111
5.20 Stability chart of a double dot system in Si 112
5.21 Stability chart of a double dot system with different tunnel coupling . 113

6.1 Equivalent circuit and potential profile of an intrinsic QD 119
6.2 Coulomb oscillations and Stability chart of an intrinsic QD 120
6.3 Motion of the excited state in B 121
6.4 Excitation spectrum of the intrinsic QD 122
6.5 Evolution of Coulomb blockade peaks with B 128
6.6 Peak spacing as a function of B 129
6.7 Schematic of spin blockade effect 132
6.8 Peak spacings as a function of $B_{\perp}$ 135
6.9 Stability chart in $B_{||}$ 137
6.10 Peak spacings as a function of $B_{||}$ 138
6.11 Spin and valley splitting in Si MOS 140
6.12 Magnetoconductance v.s. V_g 142
6.13 Magnetoresistance v.s. B 143
6.14 Fine features in stability chart 145
6.15 Fine features at different magnetic fields 146
6.16 Excitation spectrum near the first peak 147
6.17 1D potential profile of an SET 149
6.18 Transmission coefficient of a parabolic potential well 150
6.19 Schematic of quantum interference caused conductance fluctuations . 151

7.1 Center of the point-contact channel as a function of V_C,V_D 156

7.2 A small Si QD from an Australian group 158

A.1 Schematic of the measurement system 161

A.2 Sample wiring near the mixing chamber 163

A.3 Detailed wiring of the DC measurement system 164

A.4 Schematic of the AC measurement system 165

A.5 Frequency response of the AC+DC adder box 167

A.6 Frequency response of an RF powder filter 168

A.7 Frequency response of a homemade low-pass filter 170

A.8 Frequency response of the transimpedance amplifier 172

A.9 Noise spectrum of the transimpedance amplifier 173

A.10 Output noise due to the input capacitance 174

A.11 Coulomb oscillation peak measured by AC lock-in technique 178

B.1 Detailed wiring of the AC measurement system part 1/2 179

B.2 Detailed wiring of the AC measurement system part 2/2 180

List of Symbols and Abbreviations

C	Capacitance
C_{ij}	Direct Coulomb interaction energy
C_{ox}	Oxide capacitance per unit area
D_{it}	Interface trap density
e	Electronic charge (1.6×10^{-19} C), written as q sometimes
E_c	Charging energy (e^2/C), or conduction-band edge
E_d	Ionized-donor energy level
E_f	Fermi energy level
E_i	Intrinsic Fermi level
E_v	Valence-band edge
g	g-factor in the Zeeman effect, or the degeneracy
h	Planck's constant (6.63×10^{-34} J·s)
$\hbar$	Reduced Planck constant ($h/(2\pi)$)
k_B	Boltzmann's constant (1.38×10^{-23} J/K)
K_{ij}	Exchange interaction energy
L_D	Debye length
N_a	Acceptor impurity density
N_d	Donor impurity density
n_i	Intrinsic carrier density
T	Absolute temperature
T_1	Longitudinal relaxation time
T_2	Transverse relaxation time
U	Free energy
V_{fb}	Flatband voltage
V_t	Threshold voltage
α	Ratio of the gate capacitance to the total capacitance (C_g/C_Σ)
ε_0	Vacuum permittivity (8.85×10^{-14} F/cm)
ε_{si}	Silicon permittivity ($11.9\varepsilon_0$)
ε_{ox}	Oxide permittivity ($3.9\varepsilon_0$)
τ	Lifetime
ψ	Wave function, or potential in Chapter 4
ω	Angular frequency ($2\pi f$)
ω_0	Angular frequency of an harmonic oscillator
ω_c	Cyclotron frequency
μ	Electrochemical potential, or mobility

1D	1 dimensional
2D	2 dimensional
2DEG	2 dimensional electron gas
AC	Alternating current
ACE	Acetone

BCB	Benzocyclobutene
BLP-1.9+	a low pass filter from Mini-circuit
BOE	Buffered oxide etch
CHA	CHA industries, one manufacturer of e-beam evaporators
CI model	Constant interaction model
COM	Common
C-V	Capacitance v.s. voltage measurement
DC	Direct current
DI	Deionized water
DQD	Double quantum dot
DR	Dilution refrigerator
DUT	Device-under-test
ES	Excited state
ESR	Electron spin resonance
GPIB	General purpose interface bus
GS	Ground state
HDPECVD	High-density plasma-enhanced chemical-vapor-deposition
HF	High frequency
HMDS	Hexamethyldisilazane
IPA	Iso-propyl alcohol
LF	Low frequency
MIBK	Methyl isobutyl ketone
MMA	Methyl methacrylate
MNK 126	Dilution refrigerator from leiden cryogenic B.V.
MOS	Metal-oxide-semiconductor
MOSFET	Metal-oxide-semiconductor field-effect-transistor
MOSSET	Metal-oxide-semiconductor single-electron-transistor
NIL	Nanoimprint lithography
NIST	National Institute of Standards and Technology
NMR	Nuclear magnetic resonance
OPA602	Operation amplifier from TI
PAR	One manufacturer of lock-in amplifiers
PECVD	plasma-enhanced chemical-vapor-deposition
PMMA	Poly methyl methacrylate
QD	Quantum dot
QED	Quantum electrodynamic
QW	Quantum well
Rashba SO	Rashba spin-orbital effect
RF	Radio-frequency
RIE	Reactive-ion-etching
sccm	Standard Cubic Centimeters per Minute
SET	Single electron transistor
SMU	Source/monitor unit

SOI	Silicon on insulator
SRIM	Stopping and Range of Ions in Matter, a software from IBM.
ST	Singlet-triplet transition
SUPRIM IV	Semiconductor process simulator from Stanford University
TC	Time constant
UCF	Universal conductance fluctuations
UV	Ultra-violet

Chapter 1

Introduction

1.1 Motivation

Exploiting quantum properties of physical systems as a new source of computational power, has been pursued since 1980s.[1][2][3] However, It was in 1994 that Peter Shor discovered a remarkable algorithm to factor large integers fast using a quantum computer.[4] This sparks a tremendous interest in quantum computers, since they can potentially break many cryptosystems in use today. Shor's result is a powerful evidence that quantum computers are indeed more powerful than classical computers. Another example is the Grover database search algorithm, which speeds up the data search process on quantum computers.[5]

There are a number of physical implementations, including nuclear magnetic resonance (NMR) systems, trapped ions, quantum dots (QDs), super conducting circuits, cavity quantum electrodynamic (QED) systems, and other systems.[6] People are already able to demonstrate basic quantum computing operations using NMR systems, trapped ions, but they are difficult to scale up. Silicon single electron transistors belong to the quantum dot approach. As a solid state implementation, it has the potential for upscaling, an important merit for practical applications. It is also compatible with existing microelectronics processing.

The quantum dot quantum computer architecture was first proposed by Daniel

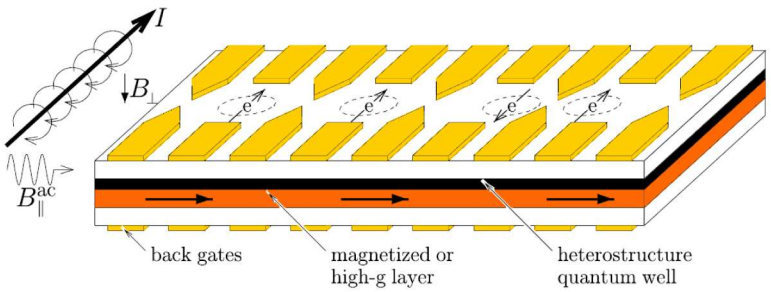

Figure 1.1: Schematic diagram of the quantum dot quantum computer architecture. Single electron spins are manipulated by local electron spin resonance (ESR), and exchange interaction between two nearby quantum dots is controlled by electrostatic gates. (After V. N. Golovach and D. Loss, Ref.[8])

Loss and David P. DiVincenzo in 1997.[7] Their implementation uses the spin of individual excess electrons confined in QDs as qubits. Their idea was further developed by Golovach and Loss.[8] The schematic diagram of such a system is shown in Fig. 1.1. Information is encoded in two-level electron spin states (qubits). Local electron spin resonance (ESR) is used to rotate single electron spins. Exchange interaction between two QDs is controlled by electrostatic gates. By properly turning the exchange interaction on or off, the square root of swap ($\sqrt{SWAP}$) operation on two qubits can be realized. All possible single qubit operations and the square root of swap operation on two qubits can form a universal operation set for quantum computation.[7]

1.2 Silicon quantum dots for quantum computing

Since the scalable quantum dot quantum computer architecture was proposed,[7] tremendous research efforts have been put into the implementation of basic quantum circuits. Various material systems have been explored, including GaAs/AlGaAs systems,[9][10] Si/SiGe systems,[11] silicon on insulator (SOI) systems,[12] and metal-oxide-semiconductor (MOS) systems,[13][14] etc.

Currently the most advanced progress is from GaAs QD approach. Marcus Group of Harvard University has demonstrated the coherent exchange of two electron spins in a double dot system, controlled by fast electrical switching of the tunnel coupling between these two QDs.[9] Thus they have realized the $\sqrt{SWAP}$ operation on two qubits. Kouwenhoven group of Delft University has realized the rotation of a single electron spin in a QD using ESR.[10] Hence a universal operation set has been accomplished.[7] But the prominent problem of GaAs is the short electron spin coherence times T_1 and T_2, which have been measured roughly to be 1 ms and 1 μs respectively. This is due to the hyperfine interaction, because an electron spin couples to randomly oriented nuclear spins of surrounding atoms in the crystal lattice. Since all isotopes of Ga and As have nuclear spin, it can not be avoided. On the contrary, a single electron confined in a silicon QD is expected to have a spin coherence time orders of magnitudes longer than that in GaAs, because ^{28}Si has zero nuclear spin, and there is no hyperfine interaction between electron spins and nuclear spins. This is the major advantage of silicon QDs, as shown in Table 1.1.

Other advantages include the stable and high quality thermal oxide. Silicon

Table 1.1: Electron spin coherence times in different material systems. (After Chia-Hung Yang.)

	Si:P	^{28}Si:P	Si QD	SiGe 2DEG	GaAs QD
T_1	Hours	100 ms	?	2.3 μs	$\sim$1 ms
T_2	-	60 ms	$\sim$1 ms?[1]	3 μs	$\sim 1\mu$s
Temperature	$\sim$2 K	$\sim$7 K	?	$\sim$5 K	$\sim$ 30 mK
Environment	Natural Si	^{28}Si	?	Si/SiGe QW	GaAs/AlGaAs
Technique	ESR [2]	ESR[3]	?	ESR [4]	Transport [5, 6]
Cause of decoherence	Photon	Photon	?	Rashba SO	Hyperfine

[1] Because of the quality of Si QDs, T_1 and T_2 have not been measured. However, coherence times can be upwards of three orders of magnitude longer than in GaAs.[15]

[2] G. Feher and E.A. Gere, Phys. Rev. 114, 1245 (1959).

[3] A. M.Tyryshkin, S. A. Lyon, A. V. Astashkin, and A. M. Raitsimring, Phys. Rev. B 68, 193207 (2003).

[4] A. M. Tyryshkin, S. A. Lyon, W. Jantsch, and F. Schaffler, Phys. Rev. Lett. 94, 126802 (2005).

[5] Kouvenhoven group, Ref. [16].

[6] Marcus group, Ref. [9].

is also backed by the immense resource from the semiconductor industry.

Compared to GaAs, there are also disadvantages of silicon, including the larger effective mass and the lower electron mobility. The effective mass m^* is $0.19m_e$ in silicon, while m^* is $0.067m_e$ in GaAs, which means smaller confinement energy. The lower electron mobility means a shorter coherent length. We may be able to overcome these problems by making smaller devices, using high purity wafers, and optimizing fabrication processes.

Since the first observation of Coulomb blockade oscillations in silicon,[17] there has been steady progress in the development of silicon-based single electron transistors using silicon-on-insulator (SOI) structures,[12] Si/SiGe quantum well structures[11] and metal-oxide-semiconductor (MOS) structures,[13][14] though it remains a considerable challenge due to material properties.[13] In the SOI approach, the thin Si layer is sandwiched by oxide on the top and at the bottom, so the size of the device can be small. But the defects at the silicon/buried oxide interface cause strong localization, and the last few electrons are difficult to deplete.[18] In the Si/SiGe quantum well approach, the sample contains a two-dimensional electron gas (2DEG) from donors in the system prior to nanofabrication. Surface Schottky gates are used to define QDs and to deplete electrons in QDs from many down to 1. Although this depletion mode approach has been applied in GaAs single electron transistors (SETs), fabrication of silicon-based SETs suffers from problems due to material deficiency. For example, gate leakage current due to dislocations in Si/SiGe quantum wells frequently disrupts the single electron transport. However they are making progress.[11] In the MOS approach, some people use the tradi-

tional doped MOS field-effect transistor (FET) structure.[19] Although the process is readily available, dopant induce disorder reduces the electron mobility, and tunnel barriers are also not tunable.

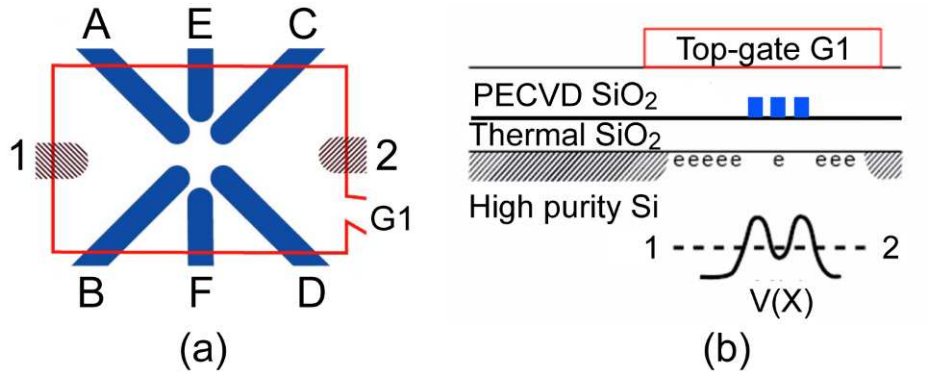

Figure 1.2: (a) Schematic top view and (b) cross sectional view of an MOS-SET.(Not drawn to scale) In (a), the shaded regions, the solid lines, and the rectangular area, depict the heavily phosphorus-doped Ohmic source (1) and drain (2) leads, the 6 side gates (A-F), and the top gate (G1), respectively. The top gate induces a 2DEG at the Si/thermal oxide interface, and the side gates deplete the 2DEG into a quantum dot and two point contact channels. The schematic potential profile along the 1-2 direction is shown as the inset (lower right) in (b).

Our device is based on an enhancement-mode MOS structure with a bilayer-gated configuration. As shown in Fig. 3.1, the top gate can induce a 2DEG at the Si/thermal oxide interface, so the quality of the 2DEG potentially can be high. It also has electrostatically defined tunnel barriers, which can be tuned by side gates. We also use high purity silicon wafers to minimize the concentration of impurities and to improve the electron mobility. Compared to the SOI and Si/Ge quantum

well approaches, the difficulties are how to make the device small and to keep the electron mobility high after fabrication processes.

Because of the flexibility of the bilayer-gated configuration, it has also been actively pursued by other research groups, most notably one group at the University of New South Wales (Australia),[14] and another group at Sandia National Laboratories.[20] The Australian group demonstrated a small device with a 40 nm quantum dot, which works in the few electron regime, although the device quality still needs to be improved.[14] The Sandia group are working on optimizing the fabrication processes. Until now their devices only work in the many electron regime.[20]

1.3 Outline of this thesis

The purpose of this thesis is to investigate QDs in silicon MOS-SETs as potential quantum computing devices. The current understanding of this system is presented. I also discuss the difficulties and the possible solutions for further improvement.

In Chapter 2, I introduce the basic theories of single electron transistors and double quantum dot systems. The characterization techniques are also explained. It serves as a background to understand the data in Chapter 5 and Chapter 6.

In Chapter 3, the device structure and the operating principle of an MOS-SET are explained. The detailed fabrication processes are also discussed.

Our SET device is based on a silicon MOS structure, and the Si/SiO_2 interface

is critical to the device characteristics. In Chapter 4, we examine the quality of the Si/SiO_2 interface in terms of the interface-trap density and the electron mobility. The 2D electron concentration is also measured.

In Chapter 5, Silicon MOS-SETs are systematically investigated at 4.2 K, and separately in a dilution refrigerator. The data show that there is an intrinsic QD in a point-contact channel in addition to the electrostatically defined QD. It is demonstrated that the intrinsic QD can be removed, when the gap between two neighboring side gates of the point-contact channel is reduced. Since electrostatically well defined lateral QDs are a must in order to realize spin qubit devices in silicon, the future possible improvement is discussed.

For the purpose of investigating single spin in silicon, we focus on a single intrinsic QD in Chapter 6. The magnetic field dependence of the ground-state and excited-state energy levels is measured. The two-electron singlet-triplet (ST) transition is first time directly observed in a silicon QD by excitation spectroscopy. The observed amplitude spectrum suggests the spin blockade effect. When the two-electron system forms a singlet state at low magnetic fields, and the injection current from the lead becomes spin-down polarized, the tunneling conductance is reduced by a factor of 8. At higher fields, due to the ST transition, the spin blockade effect is lifted and the conductance is fully recovered.

In Chapter 7, I summarize the major results in the thesis, and discuss the possible improvements in the future.

Chapter 2

Theories of single electron transistors and double quantum dot systems

2.1 Single electron transistors (SETs)

2.1.1 Operating principle of an SET

Like a metal-oxide-semiconductor field–effect transistor (MOSFET), a single electron transistor (SET) has a source, drain and gate [Fig. 2.1(a)]. It also has an island, which is capacitively coupled to the source, drain and gate. There is a special requirement. The capacitor between the island and the gate should have no leakage, but the capacitors between the island and the source, the island and the drain should allow electrons to tunnel through. However, the resistance should be much larger than the resistance quantum h/e^2 ($\sim 25.8\ \mathrm{k}\Omega$),[22] so that the electrons are well localized on the island.

The basic operating principle of an SET is as follows. Let's ignore the gate first. When we apply a bias between the source and the drain, electrons move from the source to the island. This charges the island up and requires a finite energy. Because electrons are discrete particles, there is a situation like this [Fig. 2.1(c)], if we move one more electron from the source to the island, the potential energy of electrons on the island will be higher than the source. This cannot happen because

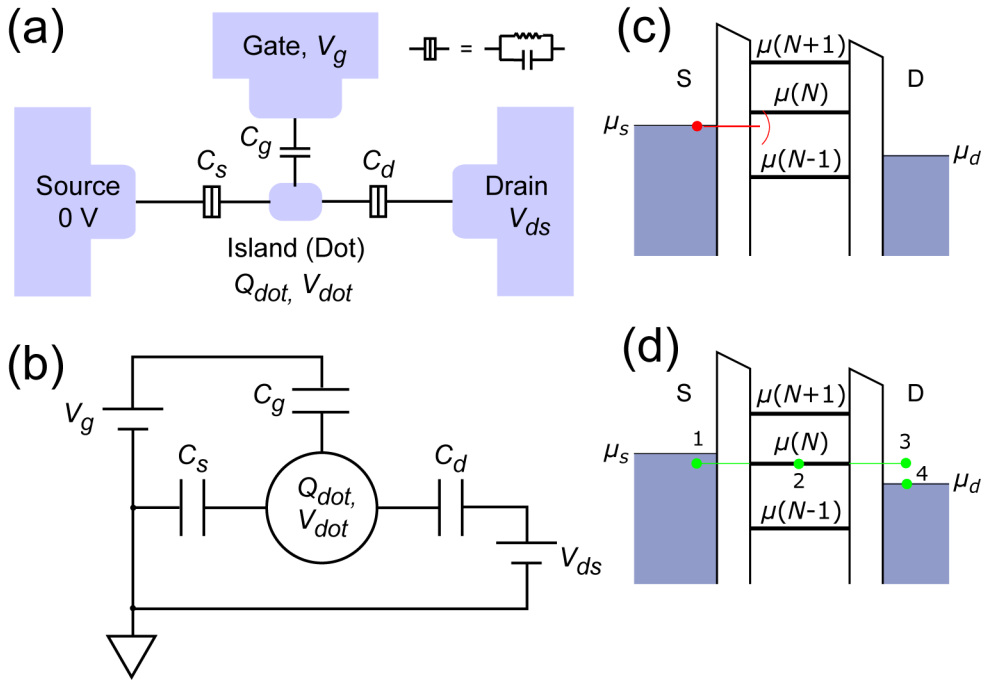

Figure 2.1: (a) Schematic of a single electron transistor (SET). It has two tunnel junctions C_s and C_d, connected to a small area, known as the island. The electrical potential of the island can be tuned by the gate, capacitively coupled to the island. A tunnel junction is characterized by a tunnel resistor and a capacitor, as shown in the inset. (b) In the constant interaction model (CI model), an SET can be modeled as a capacitor network. The capacitor network and voltage sources form an isolated closed system. (c) The SET is in the blocking state. (d) The SET is in the transmitting state. μ_s and μ_d are the electrochemical potentials (Fermi levels) of the source and drain, respectively. $\mu(N)$ is an electrochemical potential level, defined in Eq. 2.7.(After Ref. [21].)

of the energy conservation law. If the next highest potential energy level is lower than the drain potential energy, electrons cannot move to the drain to create a vacancy. So the device is off. Now we consider the gate. If a positive voltage is applied on the gate, it will effectively lowers the potential energy of electrons on the island. If there is a state between the source and drain bias window, the device is on [Fig. 2.1(d)].

When the island of the SET is a small region of semiconductor, and electrons on the island are strongly confined in all three spatial directions, the energy spectrum of the island becomes quantized. This small island is called a quantum dot (QD). It is also called an artificial atom.[22]

To approximately describe an SET, people often refer to the constant interaction model (CI model).[23] There are two basic assumptions in this model. First, the coulomb interactions of an electron on the island with others, in or outside the island, are characterized by a constant capacitance C. Second, the discrete single-particle energy spectrum, calculated for non-interacting electrons, is unaffected by the interactions.[22]

The SET can be modeled as a capacitor network [Fig. 2.1(b)]. If the capacitor network is combined with voltage sources, an isolated closed system is formed. When $V_{ds} = 0$, we can define a free energy as the total energy of the isolated closed system, including the capacitors and the voltages sources, and defined up to an additive constant. Here the free energy is assigned to be 0 when the electric potential of the quantum dot V_{dot} is 0. So the free energy is equal to the change of the total electrostatic energy E_Σ stored in the capacitors minus the work done by voltage

sources when V_{dot} changes from 0 to V, where V is the electric potential of the dot at the final state. This free energy determines the transport of the SET. In the linear regime of conductance, i.e. the difference of the electrochemical potentials of the source and drain leads $\mu_d - \mu_s = -eV_{ds} \approx 0$, and the electric potential of the source $V_s = 0$, so the electric potential of the drain $V_d \approx 0$. The charge on the QD is

$$\begin{aligned} Q_{dot} &= C_s V_{dot} + C_d V_{dot} + C_g(V_{dot} - V_g) \\ &= CV_{dot} - C_g V_g \end{aligned} \tag{2.1}$$

$$CV_{dot} = Q_{dot} + C_g V_g \tag{2.2}$$

where C_s, C_d, and C_g are the capacitances between the dot and the source, drain and gate respectively, the total capacitance $C = C_s + C_d + C_g$, and V_g is the electric potential of the gate. The free energy is equal to the work done by an external force when charges move from the source to the quantum dot, so that V_{dot} changes from 0 to V. We also notice that the work done by voltage source V_g is equal to the charge moving from the voltage source to the upper plate of capacitor C_g times V_g, and this work does not contribute to the process when the charge moves from the

ground to the QD. So the free energy

$$\begin{aligned} U(Q) &= \int_{V_{dot}=0}^{V} V_{dot} dQ_{dot} \\ &= \int_{V_{dot}=0}^{V} V_{dot} d(CV_{dot} - C_g V_g) \\ &= C \int_{V_{dot}=0}^{V} V_{dot} dV_{dot} \\ &= \frac{1}{2} CV^2 \qquad (2.3) \\ &= \frac{(Q + C_g V_g)^2}{2C} \qquad (2.4) \end{aligned}$$

from Eq. 2.1 and Eq. 2.2, where Q and V are the net charge Q_{dot} and the electric potential V_{dot} of the dot at the final state. We can define

$$Q_{ext} = C_g V_g, \qquad (2.5)$$

which is usually called the "external charge".[24] The physical meaning of this definition and Eq. 2.4 becomes straightforward. $-Q_{ext}$ is the polarization charge on the dot, which is induced by the gate voltage, when $V_{dot} = 0$. The work done by the external force moves additional charge $Q - (-Q_{ext}) = Q + Q_{ext}$ from the source to the quantum dot, so the work done by an external force is just Eq. 2.4.

On the QD, the net charge Q is equal to the charge of the free electrons $(-eN)$ plus the background charge $(Q_0 = eN_0)$: $Q = -eN + eN_0 = -e(N - N_0)$, where N is an integer, N_0 can be any real number. If we also consider the quantized energy levels in the dot, the free energy becomes

$$U(N) = \frac{[e(N - N_0) - C_g V_g]^2}{2C} + \sum_N E_n, \qquad (2.6)$$

where N is the number of free electrons on the dot, eN_0 is the background charge,

and $\sum_N E_n$ is a sum over the occupied energy states of the QD. The electrochemical potential of the dot is given by

$$\begin{aligned}\mu(N) &= U(N) - U(N-1) \\ &= (N - N_0 - 1/2)E_c + E_N - e(C_g/C)V_g, \end{aligned} \tag{2.7}$$

the energy needed to add the Nth electron to the dot, where the charging energy $E_c = e^2/C$. So the electrochemical potential $\mu(N)$ includes the charging energy term and the single particle energy E_N. We also notice that $\mu(N)$ can be measured by the gate voltage with a conversion factor eC_g/C. The energy levels of the dot in Fig. 2.1(c) and (d), are actually the electrochemical potential levels (..., $\mu(N-1)$, $\mu(N)$, $\mu(N+1)$, ...). The addition energy E_A is defined as

$$\begin{aligned} E_A(N) = \Delta\mu(N) &= \mu(N+1) - \mu(N) \\ &= E_c + E_{N+1} - E_N \\ &= E_c + \Delta E \end{aligned} \tag{2.8}$$

Electrons can flow through the QD, when there is an electrochemical potential level between the electrochemical potentials of the source and drain leads, i.e. $\mu_s \geq \mu(N) \geq \mu_d$, as shown in Fig. 2.1(d). In the linear region with $\mu_s = 0$ and $\mu_d - \mu_s = -eV_{ds} \approx 0$, when we sweep the gate voltage, the successive peaks happen at $\mu(N; V_g) = 0$, so $e(C_g/C)V_g^N = (N - N_0 - 1/2)E_c + E_N$, as shown in Fig. 2.2(a). Thus the addition energy is related to the peak spacing in the gate voltage by $E_A(N) = E_c + E_{N+1} - E_N = e(C_g/C)(V_g^{N+1} - V_g^N)$, where V_g^N and V_g^{N+1} are the gate voltages of the Nth and $(N+1)$th Coulomb blockade peaks. We also define

the conversion factor

$$\alpha = C_g/C, \tag{2.9}$$

so that the change of the gate voltage can be converted into the change of the electrochemical potential of the dot by $\Delta E_{dot} = e\alpha\Delta V_g$.

If we sweep the DC bias between the drain and the source of an SET while stepping the gate voltage, the differential conductance ($\partial I_{ds}/\partial V_{ds}$) produces a 2D color graph, called a stability chart or "diamond chart" [Fig. 2.2(b)]. The stability chart can be used to characterize the SET tunnel-barrier capacitances as well as the size of the quantum dot. The half height (e/C), the width (e/C_g), and the slopes ($-C_g/C_d$) and $C_g/(C - C_d)$ of the diamond edges can uniquely determine the SET structure. By finding these capacitances, we can model the dot as a disc of radius r and solve for the radius using

$$C = 8\varepsilon r, \tag{2.10}$$

where ε is the semiconductor dielectric constant. Please note that using such a method, the size is usually overestimated, because the quantum dot is modeled by an unscreened disc, whereas in reality the Coulomb interaction of electrons in the dot is partially screened by the nearby leads. For the same size, the capacitance of a screened disc is larger than that of an unscreened one.[22] In Fig. 2.2(b), the diamonds labeled by single numbers represent the Coulomb blockade regions, where electrons can not flow through the SET, and the number of electrons in the dot is stable.

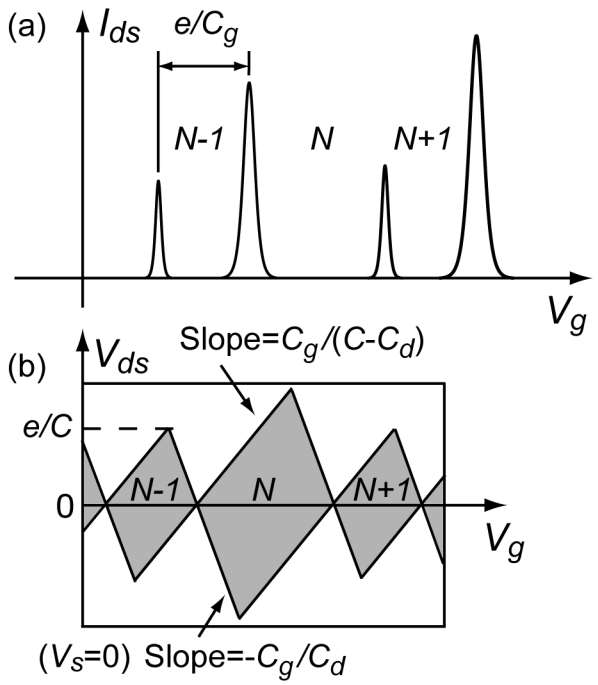

Figure 2.2: (a) Coulomb blockade oscillations of an SET. The distance between two nearest peaks is $\Delta V_g = e/C_g$. (b) Stability chart of an SET, where the source-drain differential conductance G_{ds} is measured against V_{ds} and V_g, and shows as a 2D color graph. The rectangle indicates the range of V_{ds} and V_g in the measurement. G_{ds} is zero in the gray area, and not zero otherwise. The Coulomb blockade regions (gray areas) are diamond shaped. The minimum half height is $\Delta V_{ds} = e/C$. The slopes of the edges are $(-C_g/C_d)$ and $C_g/(C - C_d)$, respectively.

2.1.2 Amplitude and lineshape of Coulomb oscillations

There are two kinds of Coulomb blockade oscillations:[23]

1. $\Delta E \ll k_B T \ll e^2/C$, the classical or metallic Coulomb blockade regime, where many levels are excited by thermal fluctuations.

2. $k_B T \ll \Delta E, e^2/C$, the quantum Coulomb blockade regime, where only one or a few levels participate in transport.

Where ΔE is the energy level spacing due to size quantization in the dot, $k_B T$ is the thermal energy, and e^2/C is the charging energy.

The classical Coulomb blockade regime can be described by the "orthodox" Coulomb blockade theory. The lineshape of a conductance peak is given by [23]

$$\frac{G}{G_\infty} = \frac{\delta/k_B T}{2\sinh(\delta/k_B T)} \approx \frac{1}{2}\cosh^{-2}(\frac{\delta}{2.5k_B T}) \text{ for } h\Gamma, \Delta E \ll k_B T \ll e^2/C, \quad (2.11)$$

where $\delta = e\alpha(V_g - V_{g0})$, V_{g0} is the gate voltage at resonance, k_B is the Boltzmann constant, T is the electron temperature, $\alpha = C_g/C_\Sigma$ is the ratio of the gate capacitance to the total capacitance, and $h\Gamma$ is the lifetime broadening of the energy levels in the dot. $1/G_\infty = 1/G_l + 1/G_r$, is the Ohmic sum of the two barrier conductances, which is independent of the temperature and the size of the dot, and is characterized completely by the two barriers. So the maximum conductance at peak $G_{max} = G_\infty/2$ in this regime, which is independent of temperature.[23]

In the quantum Coulomb blockade regime, electrons tunnel through the dot by a single level. The single peak conductance is given by [25]

$$G/G_\infty = \frac{\Delta E}{4k_B T}\cosh^{-2}(\delta/2k_B T) \text{ for } h\Gamma \ll k_B T \ll \Delta E, e^2/C, \quad (2.12)$$

where ΔE is assumed independent of E and T. The lineshapes in the classical regime and in the quantum regime are the same, except for the different "effective electron temperatures". The maximum conductance at peak $G_{max} = G_\infty \cdot (\Delta E/k_B T)$, increases linearly with decreasing temperature in the quantum regime, while it is constant in the classical regime. This can be used to distinguish a quantum peak from a classical peak. Due to the quantum phase coherence, the quantum conductance G_{max} can exceed the Ohmic value G_∞.

In our silicon SETs, the 2D electron concentration is about $n_{2d} = 1 \times 10^{12}$ cm^{-2}, and the corresponding Fermi wavelength is about 35 nm, which is at the same order as the device dimension ($\sim$ 100 nm). In the dilution refrigerator, the electron temperature is about 400 mK, and $k_B T$ is about 0.035 meV, whereas ΔE is about $0.2 \sim 1.6$ meV, and e^2/C is about $2 \sim 6$ meV. So our silicon SET is in the quantum Coulomb blockade regime.

In the above discussion, the tunnel resistance R_t is assumed to be much larger than the resistance quantum h/e^2 ($\sim$ 25.8 kΩ), so that the lifetime broadening $h\Gamma$ is much less than the charging energy e^2/C, because $(h\Gamma)\tau \sim h$, and the lifetime $\tau = R_t C$ due to tunneling, $h\Gamma \sim h/\tau = h/(R_t C) \ll h/(C \times h/e^2) = e^2/C$. This implies that the tunnel coupling between the dot and the leads is small, and the charge is well defined in the dot. In addition, $h\Gamma$ should be less than the thermal energy $k_B T$. Otherwise higher-order tunneling processes can not be neglected any more.[25] When the tunnel resistance R_t is much larger than the resistance quantum h/e^2, the dot is called a closed dot, and the transport behavior is dominated by the Coulomb blockade effect.

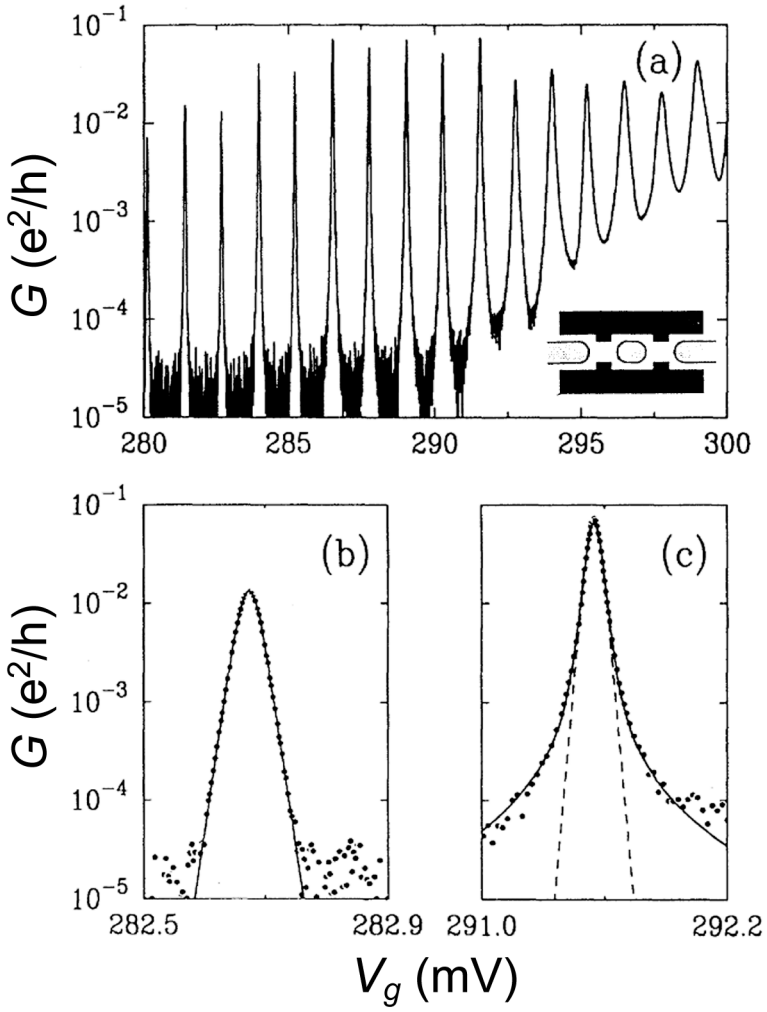

Figure 2.3: (a) Coulomb blockade oscillations measured at $B = 2.53$ T. (b) A low V_g conductance peak from (a) shown fit to a thermally broadened resonance (solid line) when the life time broadening $h\Gamma \ll k_B T$. (c) A conductance peak at higher V_g shown fit to a thermally broadened Lorentzian (solid line). The dash line is the best fit using the same line shape as in (b). (After E.B. Foxman et al., Ref. [26].)

When the barrier conductance increases, the tunnel coupling between the dot and the lead increases. The charge quantization in the dot is gradually lost, and the Coulomb blockade is lifted. When the tunnel conductance G_t is larger than $2e^2/h$, the dot is called a open dot, and the transport behavior is dominated by the quantum interference effect.[27] In Fig. 2.3(a), with an increasing gate voltage, the coupling increases. There are two distinct regimes. When $V_g < 290$ mV, the SET is in the Coulomb blockade regime (a closed dot), where the conductance minima are 0. The lineshape in this regime is determined by Eq. 2.12, as shown in Fig. 2.3(b). When $V_g > 290$ mV, there is no Coulomb blockade any more, and the conductance minima do not go to zero. Please note that the electron temperature is about 60 mK in the measurement, which is much less than the charging energy. The device gradually turns into an open dot. In the transition, the electron-electron Coulomb interaction (the charging energy) decreases, due to the strong coupling between the dot and the leads.[26] The lineshape is approximately Lorentzian [Fig. 2.3(c)]. In an open dot regime, the quantum interference effect plays an important role.

2.1.3 Quantum interference in a closed dot and in an open dot

In order to illustrate the characteristics of quantum interference, I consider an electron tunneling through two identical rectangular potential barriers as shown in Fig. 2.4. A more general asymmetrical rectangular double barrier case can be found

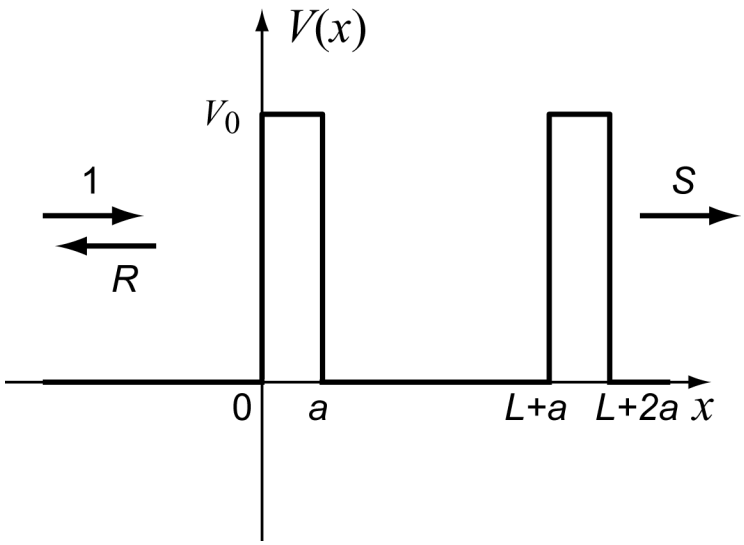

Figure 2.4: The potential profile of a rectangular double barrier, with barrier height V_0.

in Ref. [28]. We have the wave function in the different regions

$$\text{for } x < 0,\ \psi = e^{ikx} + Re^{-ikx}, \tag{2.13}$$

$$\text{for } a < x < L + a,\ \psi = Ae^{ikx} + Be^{-ikx}, \tag{2.14}$$

$$\text{for } x > L + 2a,\ \psi = Se^{ikx}. \tag{2.15}$$

where $k = \sqrt{2m^*E}/\hbar$, m^* is the effective mass of the electron. and E is the energy of the electron.

In the case of a single barrier, the wave function is [29]

$$\text{for } x < 0,\ \phi = e^{ikx} + R_0e^{-ikx}, \tag{2.16}$$

$$\text{for } x > a,\ \phi = S_0e^{ikx}. \tag{2.17}$$

and

$$R_0 = \frac{(k^2 + \beta^2)\sinh\beta a}{(k^2 - \beta^2)\sinh\beta a + i2k\beta\cosh\beta a} = |R_0|e^{i\varphi_0}, \tag{2.18}$$

where $\beta = \sqrt{2m^*(V_0 - E)}/\hbar$, V_0 is the height of the potential barrier, $|R_0|$ and φ_0 are the modulus and the phase angle of R_0, respectively.

The total transmission coefficient

$$\begin{aligned} T = |S|^2 &= \frac{|S_0|^4}{|1 - |R_0|^2 e^{i2(kL+\varphi_0)}|^2} \\ &= \frac{T_0^2}{|1 - (1 - T_0)e^{i2(kL+\varphi_0)}|^2} \end{aligned} \tag{2.19}$$

where $T_0 = |S_0|^2$, $|R_0|^2 + |S_0|^2 = 1$.

$T = 1$, when

$$2(kL + \varphi_0) = 2n\pi, \tag{2.20}$$

where n is an integer. So

$$E_n = \frac{\hbar^2(n\pi - \varphi_0)^2}{2m^*L^2} \text{ at } T = 1. \tag{2.21}$$

Eq. 2.20 is the quantization condition.

In an SET, the transmission coefficient of each barrier has $T_0 \ll 1$, and the wave function in the well can be approximately treated as a bound state. The quantization condition to determine the eigen energy levels inside a rectangular potential well is [30]

$$k \tan \frac{kL}{2} = \beta, \text{ or } k \cot \frac{kL}{2} = -\beta, \tag{2.22}$$

which is exactly the same as Eq. 2.20, when $T_0 \to 0$, i.e. $\beta a \gg 1$, as shown below.

At $\beta a \gg 1$, Eq. 2.18 becomes

$$R_0 = \frac{k^2 + \beta^2}{(k + i\beta)^2} = e^{i\varphi_0}. \tag{2.23}$$

From Eq. 2.20, we get

$$e^{i(kL+\varphi_0)} = (-1)^n$$
$$e^{ikL/2} = i^n \frac{k+i\beta}{\sqrt{k^2+\beta^2}} \tag{2.24}$$

Eq. 2.24 is just the same as Eq. 2.22.

More generally, for an arbitrary potential well in an SET with barriers $T_0 \ll 1$, the total transmission coefficient reaches the maxima when the energy of the electron is equal to the eigen energy levels inside the potential well, which can be understood in the sense of the law of energy conservation.

For an open dot with $E > V_0$, the results are easily obtained by replacing β by $i\beta$ with $\beta = \sqrt{2m^*(E-V_0)}/\hbar$ in Eq. 2.18. Eq. 2.20 holds even for an open dot. The only difference is the minimum total transmission coefficient T_{min}. For a closed dot with $T_0 \ll 1$, $T_{min} \approx T_0^2/4 \sim 0$; for an open dot,

$$T_{min} \approx \frac{T_0^2}{(2-T_0)^2} \quad \text{with} \quad T_0 = \left[1 + \frac{V_0^2}{4E(E-V_0)}\right]^{-1}, \tag{2.25}$$

which is not zero.

Fig. 2.5 shows an example of the total transmission coefficient T v.s. electron energy, where the barrier width $a = 10$ nm, the potential well width $L = 20$ nm, the height of the potential barrier is 50 meV, and $m^* = 0.19m_0$ as in silicon. There are two distinct regimes. One is the closed dot regime ($E < V_0$), in which electrons can be well confined in the potential well. The other is the open dot regime ($E > V_0$), in which the potential well cannot confine electrons. In the closed dot regime ($E < V_0$), the total transmission coefficient T is almost 0, except when the resonant tunneling

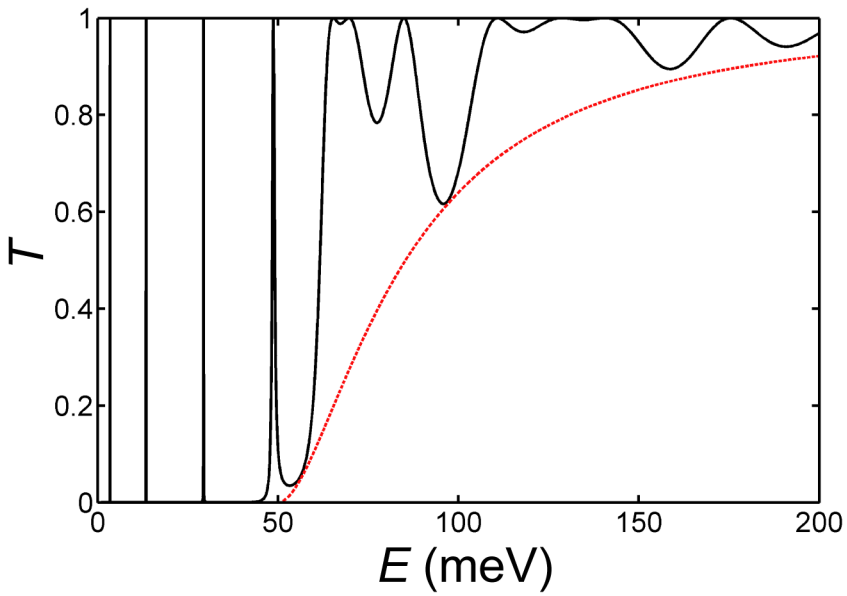

Figure 2.5: The total transmission coefficient T v.s. electron energy E of a a rectangular double barrier. The barrier width is 10 nm, the potential well width is 20 nm, the height of the barriers is $V_0 = 50$ meV, and $m^* = 0.19m_0$ as in silicon. The dash line shows that the minimum transmission coefficient is not zero when $E > V_0$.

happens. In the open dot regime ($E > V_0$), the total transmission coefficient T is always not 0, and there are resonant tunneling peaks due to quantum interference.

In a general asymmetrical rectangular double barriers, the maximum total transmission coefficient [31]

$$T = \frac{4T_l T_r}{(T_l + T_r)^2}, \tag{2.26}$$

where T_l, T_r are the transmission coefficients of the left and the right potential barriers, respectively. When resonant tunneling through the double barriers occurs, the total transmission coefficient reaches its maximum. In the case of $T_l = T_r$, $T = 1$, which means that the conductance is $2e^2/h$ for one 1D channel, according to Landauer formula.[32] When $E < \min(V_l, V_r)$, the device is in the closed dot regime; when $E > \min(V_l, V_r)$, the device is in the open dot regime, where V_l, V_r are the heights of the left and the right potential barriers, respectively.

When we compare Fig. 2.3(a) with Fig. 2.5, they share some similarities — both have a blockade regime and a non-blockade regime. In the case of Coulomb blockade [Fig. 2.3(a)], the peak spacing also includes the energy difference due to the electron-electron Coulomb interaction (the charging energy) in addition to the size quantization [Fig. 2.5]. Another difference is that Fig. 2.3(a) is measured at finite temperature, and Fig. 2.5 is calculated at 0 K.

2.1.4 Excited states and excitation spectrum

The previous discussion mainly focuses on the linear-response regime, i.e. the difference of the electrochemical potentials of the source and drain leads $\mu_d -$

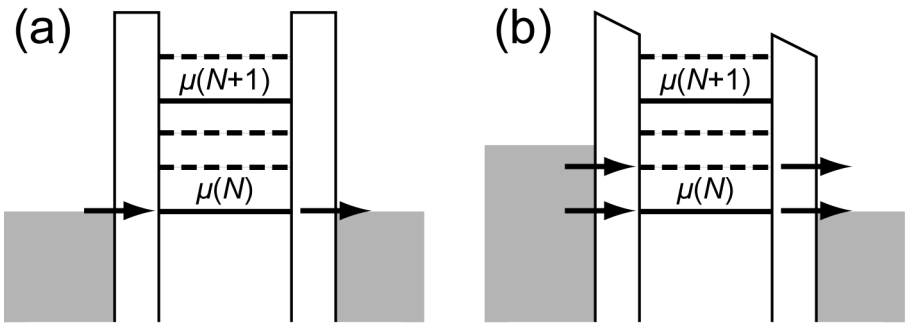

Figure 2.6: Schematic diagrams of the electrochemical potential levels of a quantum dot (a) in the small-bias regime and (b) in the large-bias regime. The solid lines are the ground states, and the dash lines are the excited states. When V_{ds} is large enough, electrons can also tunnel through the dot by excited states, as in (b).

$\mu_s = -eV_{ds} \approx 0$, and only the ground state energy levels participate in the transport[Fig. 2.6(a)]. When V_{ds} increases, the bias window between the source reservoir and the drain reservoir increases. As shown in Fig. 2.6(b), when V_{ds} is so large that higher-lying excited state energy levels can also contribute to the transport. These extra tunneling channels are usually detected as an increase in current. How exactly the current changes depends on the tunnel coupling of these energy levels involved.[33]

In a stability chart where the source-drain differential conductance is measured against V_{ds} and V_g, these excited states are shown as parallel lines along the Coulomb blockade diamond edges. From these signals, the energies of the excited states of the quantum dot can be mapped out. Fig. 2.7(a) shows the free energies of different electronic configurations and possible transitions between them. GS(N)

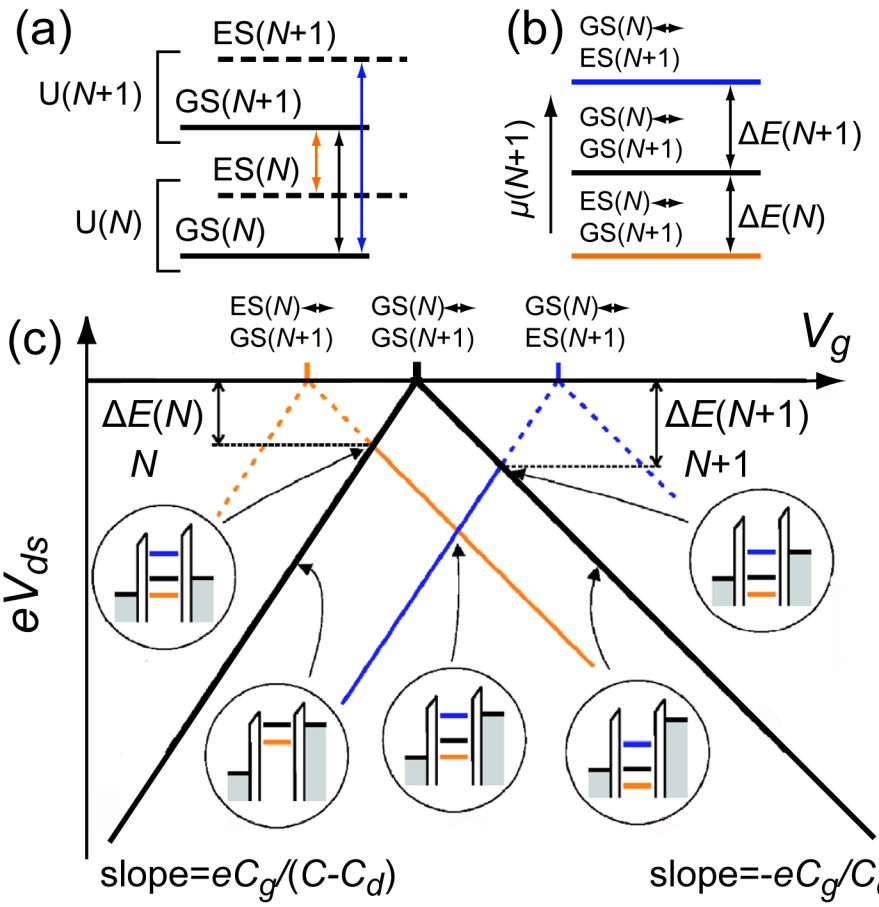

Figure 2.7: Schematic of an excitation spectrum. (a) Free energies for N electrons $U(N)$ and for $N+1$ electrons $U(N+1)$ with possible transitions. (b) The electrochemical potentials for the transitions depicted in (a). (c) Schematic of a stability chart with excited state signals (solid lines). Various alignment of energy levels with μ_s and μ_d at different V_{ds} and V_g is also explained. (After R. Hanson et al., Ref. [33].)

and ES(N) are the ground state and the excited state free energies of the system when the dot has N electrons in it. GS($N+1$) and ES($N+1$) are the ground state and the excited state free energies of the system when the dot has $(N+1)$ electrons in it. The electrochemical potentials associated with these transitions are calculated using Eq. 2.7 at a fixed V_g, and they form an electrochemical potential ladder[Fig. 2.7(b)]. The ladder can be mapped on the gate voltage axis [Fig. 2.7(c)], when the corresponding electrochemical potential is aligned with $\mu_s(=0)$ and μ_d at $V_{ds} = 0$ by tuning the gate voltage V_g. However, at small V_{ds}, electrons tunnel through the dot only at the gate voltage indicated by GS(N) $\leftrightarrow$GS(N+1), i.e. through the ground state energy level. At larger V_{ds}, when both the ground state energy level and the excited state energy levels are in the transport window, the excited states contribute to the transport and show as solid lines in Fig. 2.7(c). Various alignment of energy levels with μ_s and μ_d at different V_{ds} and V_g is also explained with schematic diagrams. From these relationships, we can get the energies of the excited states in the dot. Please note that the electrochemical potential of the transition ES(N) $\leftrightarrow$GS($N+1$) is lower than that of the transition between the two ground states GS(N) $\leftrightarrow$GS($N+1$), because when the dot has N electrons in it, and the system is in the excited state, it will need less energy to add one more electron in the dot.

2.1.5 Single-particle states in a two-dimensional elliptic harmonic oscillator

In our MOS-SETs, electrons are strongly confined along the z direction, and form a 2D electron gas (2DEG) on the $x-y$ plane. Then the 2DEG is depleted into a quantum dot by side gates. So the confining potential can be approximately modeled as a 2D anisotropic parabolic potential $V(x,y) = m^*(\omega_x^2 x^2 + \omega_y^2 y^2)/2$ in general, where $\omega_x \neq \omega_y$, and m^* is the electron effective mass.

The eigen energies of single-particle states in this potential and a magnetic field B perpendicular to the 2DEG can be solved analytically.[34, 35] It is

$$E_{n_1,n_2} = (n_1 + 1/2)\hbar\omega_1 + (n_2 + 1/2)\hbar\omega_2, \tag{2.27}$$

where n_1 and n_2 are quantum numbers, and the frequencies are given by

$$2\omega_{1,2}^2 = (\omega_x^2 + \omega_y^2 + \omega_c^2) \pm [(\omega_x^2 + \omega_y^2 + \omega_c^2)^2 - 4\omega_x^2\omega_y^2]^{1/2}, \tag{2.28}$$

where $\omega_c = qB/m^*$ is the cyclotron frequency.

The right side of Eq. 2.28 becomes

$$\begin{aligned}
&(\omega_x^2 + \omega_y^2 + \omega_c^2) \pm (\omega_x^2 + \omega_y^2 + \omega_c^2 - 2\omega_x\omega_y)^{1/2}(\omega_x^2 + \omega_y^2 + \omega_c^2 + 2\omega_x\omega_y)^{1/2} \\
=&1/2[(\omega_x - \omega_y)^2 + (\omega_x + \omega_y)^2 + 2\omega_c^2] \pm [(\omega_x - \omega_y)^2 + \omega_c^2]^{1/2}[(\omega_x + \omega_y)^2 + \omega_c^2]^{1/2} \\
=&1/2\{[(\omega_x - \omega_y)^2 + \omega_c^2]^{1/2} \pm [(\omega_x + \omega_y)^2 + \omega_c^2]^{1/2}\}^2
\end{aligned} \tag{2.29}$$

From Eq. 2.28 and Eq. 2.29,

$$\omega_{1,2} = 1/2\{[(\omega_x + \omega_y)^2 + \omega_c^2]^{1/2} \pm [(\omega_x - \omega_y)^2 + \omega_c^2]^{1/2}\}, \tag{2.30}$$

When $\omega_x = \omega_y = \omega_0$, from Eq. 2.27 and Eq. 2.30, we can get the well-known Fock-Darwin states

$$E_{n_1,n_2} = 1/2(n_1 - n_2)\hbar\omega_c + (n_1 + n_2 + 1)\hbar(\omega_0^2 + 1/4\omega_c^2)^{1/2}, \tag{2.31}$$

which play an important role in the GaAs QDs.[22]

At low magnetic field B with $\hbar\omega_c \ll \hbar\omega_0$,

$$\frac{\partial E_{n_1,n_2}}{\partial B} \approx (n_1 - n_2)\frac{q\hbar}{2m^*}. \tag{2.32}$$

However, for a strong anisotropic parabolic potential and low magnetic field B with $\hbar\omega_c \ll \hbar(\omega_x + \omega_y),\ \hbar|\omega_x - \omega_y|$, from Eq. 2.30,

$$\frac{\partial E_{n_1,n_2}}{\partial B} \propto B \sim 0 \text{ or } \frac{\partial E_{n_1,n_2}}{\partial(\hbar\omega_c)} \propto \frac{\omega_c}{|\omega_x - \omega_y|} \sim 0. \tag{2.33}$$

So for a strong anisotropic parabolic potential, the orbital effect $\hbar\omega_c$ becomes not important.

2.2 Double quantum dot systems

2.2.1 Stability diagram of two serial quantum dots

A serial double quantum dot (DQD) system can be modeled as a network of tunnel resistors and capacitors. In the case of weak tunnel coupling ($R >> 25.8$ kΩ), the system can also be modeled as a classical capacitor network [Fig. 2.8(b)]. Here we do not consider the discrete quantum states in the QDs. Cross capacitances and stray capacitances are also neglected for simplicity, although they can be easily incorporated in the model. All parameters are defined in Fig. 2.8(b), and consistent with Ref. [36].

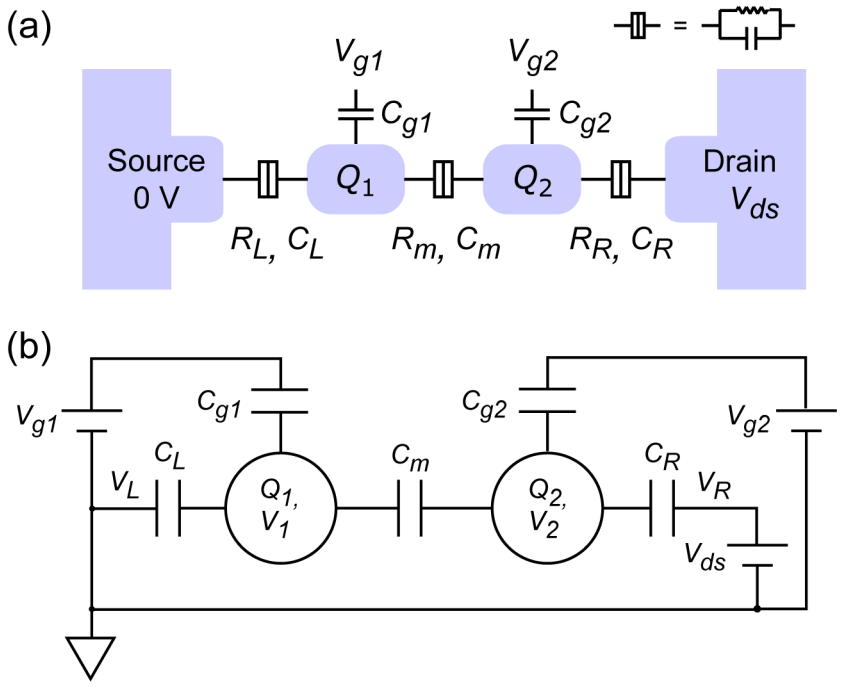

Figure 2.8: (a) Schematic of two coupled quantum dots in series. A tunnel junction is characterized by a tunnel resistor and a capacitor, as shown in the inset. (b) The serial double quantum dot (DQD) system is modeled as a capacitor network. When it combines with voltage sources, an isolated closed system is formed.

Only when the capacitor network is combined with voltage sources, an isolated closed system is formed. When $V_{ds} = 0$, the free energy is defined as the total electrostatic energy E_Σ stored in the capacitors minus the work done by voltage sources, and defined up to an additive constant: one may arbitrarily choose a condition where the free energy is zero. Here we define the free energy of the double dot system to be 0 when the electrostatic potentials of the quantum dots are 0.

The discussion will follow Ref. [36], but Ref. [36] doesn't distinguish the free energy from the electrostatic energy, and is potentially misleading.

$$\begin{aligned} Q_1 &= C_L(V_1 - V_L) + C_{g1}(V_1 - V_{g1}) + C_m(V_1 - V_2), \\ Q_2 &= C_R(V_2 - V_R) + C_{g2}(V_2 - V_{g2}) + C_m(V_2 - V_1). \end{aligned} \tag{2.34}$$

So

$$\begin{pmatrix} Q_1 + C_L V_L + C_{g1} V_{g1} \\ Q_2 + C_R V_R + C_{g2} V_{g2} \end{pmatrix} = \begin{pmatrix} C_1 & -C_m \\ -C_m & C_2 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}, \tag{2.35}$$

where $C_1 = C_L + C_{g1} + C_m$ and $C_2 = C_R + C_{g2} + C_m$. Here again $Q_{1ext} = C_L V_L + C_{g1} V_{g1}$ and $Q_{2ext} = C_R V_R + C_{g2} V_{g2}$ are the "external charge".

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \frac{1}{C_1 C_2 - C_m^2} \begin{pmatrix} C_2 & C_m \\ C_m & C_1 \end{pmatrix} \begin{pmatrix} Q_1 + C_L V_L + C_{g1} V_{g1} \\ Q_2 + C_R V_R + C_{g2} V_{g2} \end{pmatrix}. \tag{2.36}$$

More generally, Eq. 2.35 and Eq. 2.36 can be write as

$$\hat{Q} = C_{cc} \vec{V_c}, \tag{2.37}$$

$$\vec{V_c} = C_{cc}^{-1} \hat{Q}, \tag{2.38}$$

where $\hat{Q} = \begin{pmatrix} Q_1 + C_L V_L + C_{g1} V_{g1} \\ Q_2 + C_R V_R + C_{g2} V_{g2} \end{pmatrix}$, $\vec{V_c} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$, and $C_{cc} = \begin{pmatrix} C_1 & -C_m \\ -C_m & C_2 \end{pmatrix}$.

When $V_L = V_R = 0$, the free energy is equal to the work done by an external force when charges move from the ground to dot 1 and dot 2, so that $\vec{V_c}$ changes from 0 to $(V_1\ V_2)^T$, where T is the transpose of a vector. We also notice that the work done by voltage source V_{g1} is equal to the charges moving from the voltage source to the upper plate of capacitor C_{g1} times V_{g1}, and this work does not contribute to the process when the charges move from the ground to dot 1. This is also true for voltage source V_{g2}. It can be proved that the free energy

$$\begin{aligned} U(Q_1, Q_2) &= \int_{\vec{V_c}=0}^{\vec{V}} \vec{V_c} \cdot d\vec{Q} \\ &= \int_{\vec{V_c}=0}^{\vec{V}} \vec{V_c} \cdot d\hat{Q} \\ &= \int_{\vec{V_c}=0}^{\vec{V}} \vec{V_c} \cdot C_{cc} d\vec{V_c} \\ &= \frac{1}{2}\vec{V}^T \cdot C_{cc}\vec{V} \qquad (2.39) \\ &= \frac{1}{2}\hat{Q}^T \cdot C_{cc}^{-1}\hat{Q} \qquad (2.40) \end{aligned}$$

where $\vec{V} = \begin{pmatrix} V_1 \\ \\ V_2 \end{pmatrix}$ and $\vec{Q} = \begin{pmatrix} Q_1 \\ \\ Q_2 \end{pmatrix}$ are the electric potentials and the net charges of the two dots at the final state.

From Eq. 2.35 and Eq. 2.40, in the case of $V_L = V_R = 0$ and $Q_{1(2)} = -N_{1(2)}e$,

$$U(N_1, N_2) = \frac{1}{2}N_1^2 E_{C1} + \frac{1}{2}N_2^2 E_{C2} + N_1 N_2 E_{Cm} + f(V_{g1}, V_{g2}), \qquad (2.41)$$

and

$$\begin{aligned} f(V_{g1}, V_{g2}) = &-\frac{1}{e}[C_{g1}V_{g1}(N_1 E_{C1} + N_2 E_{Cm}) + C_{g2}V_{g2}(N_1 E_{Cm} + N_2 E_{C2})] \\ &+ U(0,0), \end{aligned}$$

where e is the electron charge, E_{C1} and E_{C2} are the charging energies of dot 1 and dot 2 respectively, and E_{Cm} is the electrostatic coupling energy, with

$$E_{C1} = \frac{e^2}{C_1}(1 - \frac{C_m^2}{C_1 C_2})^{-1},\ E_{C2} = \frac{e^2}{C_2}(1 - \frac{C_m^2}{C_1 C_2})^{-1},\ E_{Cm} = \frac{e^2 C_m}{C_1 C_2}(1 - \frac{C_m^2}{C_1 C_2})^{-1}.$$

Here $C_{1(2)}$ is the sum of all capacitances attached to dot 1(2) as before.

The electrochemical potential $\mu_1(N_1, N_2)$ of dot 1 is the free energy difference when the N_1th electron is added to dot 1, while dot 2 keeps N_2 electrons, so

$$\begin{aligned}\mu_1(N_1, N_2) &\equiv U(N_1, N_2) - U(N_1 - 1, N_2) \\ &= (N_1 - \frac{1}{2})E_{C1} + N_2 E_{Cm} - \frac{1}{e}(C_{g1}V_{g1}E_{C1} + C_{g2}V_{g2}E_{Cm}),\end{aligned} \tag{2.42}$$

Similarly, the electrochemical potential $\mu_2(N_1, N_2)$ of dot 2 is

$$\begin{aligned}\mu_2(N_1, N_2) &\equiv U(N_1, N_2) - U(N_1, N_2 - 1) \\ &= (N_2 - \frac{1}{2})E_{C2} + N_1 E_{Cm} - \frac{1}{e}(C_{g1}V_{g1}E_{Cm} + C_{g2}V_{g2}E_{C2}).\end{aligned} \tag{2.43}$$

The additional energy of dot 1 is the change in $\mu_1(N_1, N_2)$ when N_1 changes by 1, $\mu_1(N_1 + 1, N_2) - \mu_1(N_1, N_2) = E_{C1}$. So the additional energy is equal to the charging energy in this classical regime. The additional energy of dot 2 is also $\mu_2(N_1, N_2 + 1) - \mu_2(N_1, N_2) = E_{C2}$. And $\mu_1(N_1, N_2 + 1) - \mu_1(N_1, N_2) = \mu_2(N_1 + 1, N_2) - \mu_2(N_1, N_2) = E_{Cm}$.

Stability diagrams of a serial double-dot system are shown in Fig. 2.9. In each domain, the charge configuration is stable. These stability diagrams represent three different cases, according to the inter-dot capacitance C_m. When $C_m = 0$, Eq. 2.41 becomes

$$U(N_1, N_2) = \frac{(-N_1 e + C_{g1}V_{g1})^2}{2C_1} + \frac{(-N_2 e + C_{g2}V_{g2})^2}{2C_2}. \tag{2.44}$$

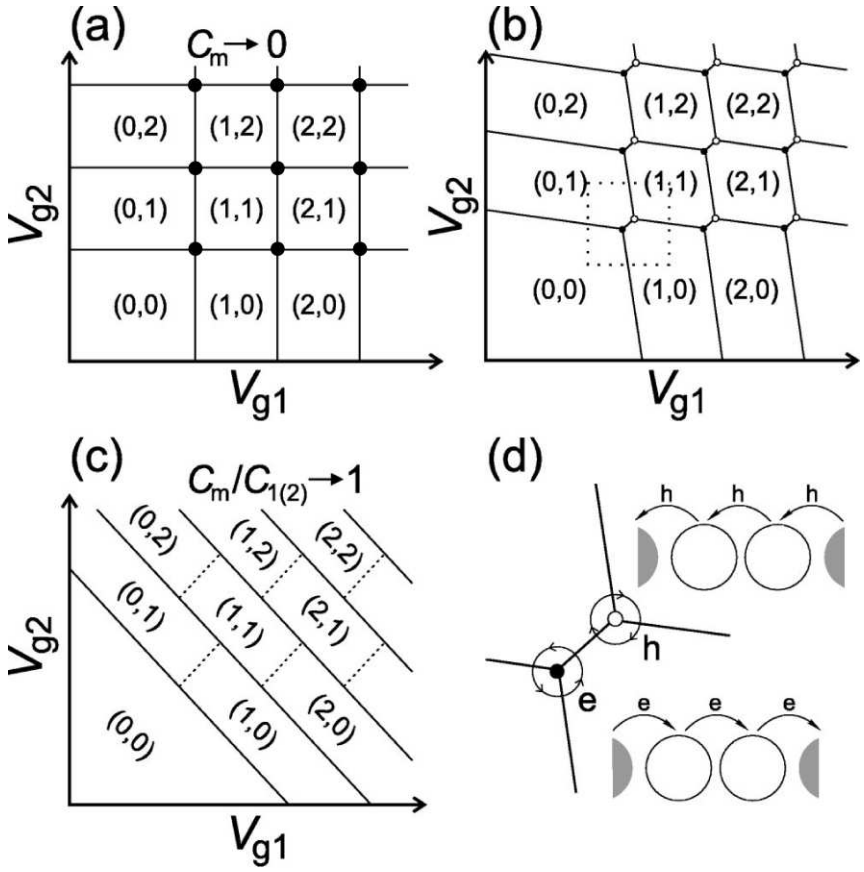

Figure 2.9: Schematic stability diagrams of a double quantum dot system with (a) small, (b) intermediate, (c) large inter-dot capacitive coupling. In each domain, the charge configuration is stable, and denoted by (N_1, N_2), where N_1 and N_2 refer to the numbers of electrons in dot 1 and dot 2, respectively. The two kinds of triple points corresponding with the electron transfer process ($\bullet$) and the hole transfer process ($\circ$) are illustrated in (d). (After W. G. van der Wiel, etc, Ref. [36].)

This means that the double-dot system includes two independent dots in series, and high conductance peaks (the dots) are located at the intersections of horizontal and vertical lines, as shown in Fig. 2.9(a). When C_m becomes the dominant capacitance in C_1 and C_2, and hence $C_m/C_{1(2)} \to 1$,

$$C_{cc}^{-1} = \frac{1}{C_1C_2 - C_m^2} \begin{pmatrix} C_2 & C_m \\ C_m & C_1 \end{pmatrix} \approx \frac{C_m}{C_1C_2 - C_m^2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{2.45}$$

so Eq. 2.40 becomes

$$\begin{aligned} U(Q_1, Q_2) &= \frac{1}{2} \frac{C_m}{C_1C_2 - C_m^2} (Q_1 + Q_2 + C_{g1}V_{g1} + C_{g2}V_{g2})^2 \\ &= \frac{1}{2} \frac{C_m}{(C_m + \tilde{C}_1)(C_m + \tilde{C}_2) - C_m^2} (Q_1 + Q_2 + C_{g1}V_{g1} + C_{g2}V_{g2})^2 \\ &\approx \frac{1}{2} \frac{1}{\tilde{C}_1 + \tilde{C}_2} (Q_1 + Q_2 + C_{g1}V_{g1} + C_{g2}V_{g2})^2 \\ &= \frac{[-(N_1 + N_2)e + C_{g1}V_{g1} + C_{g2}V_{g2}]^2}{2(\tilde{C}_1 + \tilde{C}_2)} \end{aligned} \tag{2.46}$$

where $\tilde{C}_{1(2)} = C_{1(2)} - C_m$. This means that a large inter-dot capacitance C_m leads to one big dot, and high conductance peaks are shown as diagonal lines, as in Fig. 2.9(c).

When the double dot system is in an intermediate inter-dot coupling, the stability diagram shows a honeycomb structure [Fig. 2.9(b)]. In the linear regime of conductance, i.e. the difference of the electrochemical potentials of the left and right leads $\mu_L - \mu_R = -eV_{ds} \approx 0$, electrons can only tunnel through the double dot at the vertices of the hexagonal domains – "triple points" without co-tunneling, through the following sequences: $(N_1, N_2) \to (N_1 + 1, N_2) \to (N_1, N_2 + 1) \to (N_1, N_2)$ and $(N_1 + 1, N_2 + 1) \to (N_1 + 1, N_2) \to (N_1, N_2 + 1) \to (N_1 + 1, N_2 + 1)$, as shown in Fig. 2.9(d).

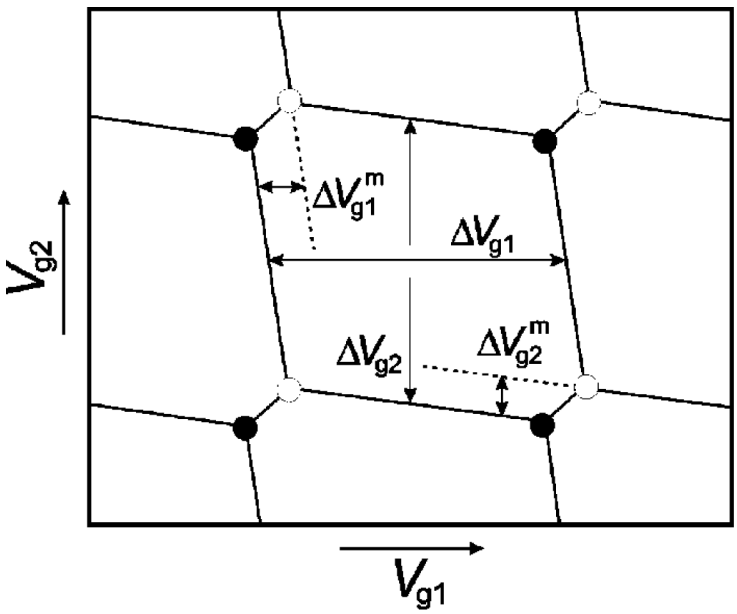

Figure 2.10: Schematic stability diagram of a double quantum dot system, showing the Coulomb peak spacings in Eq. 2.47–Eq. 2.50. (After W. G. van der Wiel et al., Ref. [36].)

The nearly vertical edges of the hexagonal domains are defined by equation $\mu_1(N_1, N_2; V_{g1}, V_{g2}) = 0$; the nearly horizontal edges are defined by equation $\mu_2(N_1, N_2; V_{g1}, V_{g2}) = 0$. So the dimensions of the hexagon [Fig. 2.10] can be expressed in terms of the capacitances from Eq. 2.42 and Eq. 2.43, and ΔV_{g1},ΔV_{g2}, ΔV_{g1}^{m} and ΔV_{g2}^{m} are defined in Fig. 2.10.

$$\mu_1(N_1, N_2; V_{g1}, V_{g2}) = \mu_1(N_1 + 1, N_2; V_{g1} + \Delta V_{g1}, V_{g2}),$$

$$\Rightarrow \Delta V_{g1} = \frac{e}{C_{g1}}. \tag{2.47}$$

Similarly,

$$\Delta V_{g2} = \frac{e}{C_{g2}}. \tag{2.48}$$

$$\mu_1(N_1, N_2; V_{g1}, V_{g2}) = \mu_1(N_1, N_2 + 1; V_{g1} + \Delta V_{g1}^{m}, V_{g2}),$$

$$\Rightarrow \Delta V_{g1}^{m} = \frac{eC_m}{C_{g1}C_2} = \Delta V_{g1}\frac{C_m}{C_2}. \tag{2.49}$$

Correspondingly,

$$\Delta V_{g2}^{m} = \frac{eC_m}{C_{g2}C_1} = \Delta V_{g2}\frac{C_m}{C_1}. \tag{2.50}$$

It can be proved that Eq. 2.49 and Eq. 2.50 are correct even if there are cross capacitances between V_{g1} and dot2, and between V_{g2} and dot1.[37]

2.2.2 Tunnel coupling between two dots

We have only discussed the regime of weak tunnel coupling between the two dots, i.e. the tunnel resistance $R_m >> 25.8$ kΩ. When the tunnel coupling (the tunnel conductance $G_m = 1/R_m$) increases, the charge quantization in dot 1 and dot 2 is gradually destroyed. Fig. 2.11 (A) to (F) show the evolution of the stability chart from the characteristic honeycomb structure to diagonal lines (one large dot) when the inter-dot tunnel conductance G_m increases from $0.22G_0$ to $0.98G_0$, where $G_0 = 2e^2/h$.[38]

In (A), the inter-dot conductance is small, and the pattern is a hexagonal array of points, with two triple points split due to the inter-dot capacitance C_m. There are two distinct effects due to the tunnel coupling. In (B)–(F), when the inter-dot conductance increases, the triple points develop into crescents, and the splitting between two nearby crescents increases; the edges of the honeycomb cells become visible due to the co-tunneling process. (Please note that the inter-dot capacitance C_m usually increases with the increasing inter-dot conductance, and also contributes to the splitting.) The condition that both N_1 and N_2 are quantized, are relaxed into

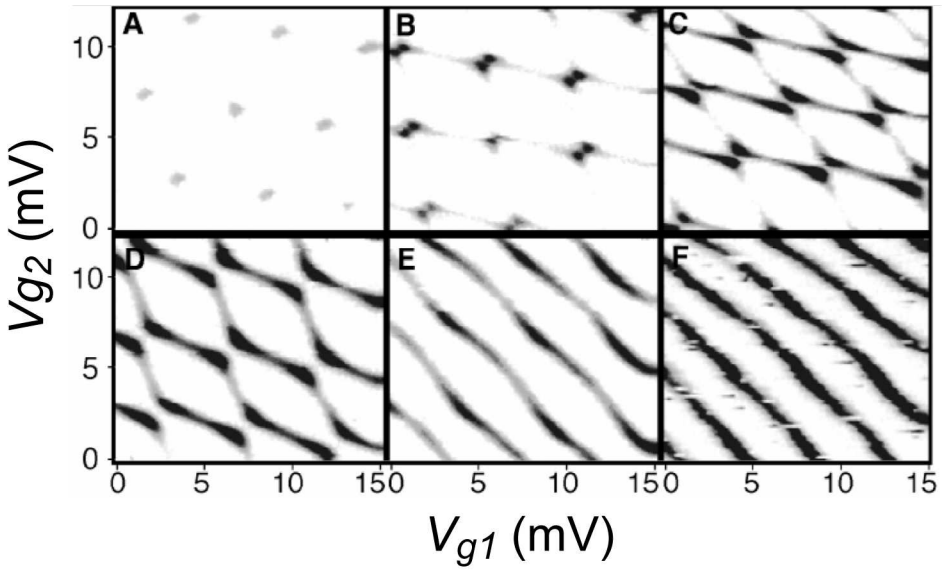

Figure 2.11: Logarithm of double dot conductance as a function of gate voltages V_{g1} and V_{g2}, which are offset to zero. Dark indicates high conductance; white regions represent low conductance. Inter-dot conductances are (A) $G_m = 0.22G_0$, (B) $G_m = 0.40G_0$, (C) $G_m = 0.65G_0$, (D) $G_m = 0.78G_0$, (E) $G_m = 0.96G_0$, and (F) $G_m = 0.98G_0$ (where $G_0 = 2e^2/h$); (F) is thresholded to a higher value of conductance to accommodate a higher background conductance.(After C. Livermore et al., Ref. [38].)

a single condition that the total charge $N_1 + N_2$ is quantized.

Excluding the splitting due to C_m, the splitting between two crescents measures the inter-dot interaction energy, which can be modeled by a two-level system $(M, N+1)$ and $(M+1, N)$, where the pair of integers refers to the numbers of electrons in dot 1 and dot 2, as discussed in Ref. [36]. The inter-dot interaction energy is an analog of the molecular binding energy. The double dot becomes an artificial covalent molecule.

Chapter 3

Device structure and fabrication process

3.1 Device structure

Our silicon single electron transistor (SET) is based on an enhancement-mode metal-oxide-semiconductor (MOS) structure. Figures 3.1(a) and (b) show the schematic top view and the cross-sectional view of the device. The device is fabricated on an N-type, high purity silicon (100) wafer with a resistivity of $3 \sim 5$ kΩ·cm. At cryogenic temperatures, a high purity wafer is non-conductive due to the lack of thermally generated carriers (dopant freeze-out). When silicon is heavily doped, it undergoes the metal-insulator transition and becomes conductive even at low temperatures.[39] So the heavily phosphorous-doped regions are used as the source (1) and the drain (2). A 27 nm thick thermal oxide is grown on the wafer by dry oxidation. Six side gates (labeled as A-F in Fig. 3.1(a)) are deposited on the thermal oxide and buried in the second dielectric layer, SiO_2, which is 400 nm in thickness and grown by high-density plasma-enhanced chemical-vapor-deposition (HDPECVD). Finally the top gate (G1) is deposited on the HDPECVD oxide. The top gate also laterally overlaps with the Ohmic source and drain regions, as shown in Fig. 3.1.

The top gate is positively biased to induce 2D electron gas (2DEG) at the Si/SiO_2 interface, similar to an enhancement-mode field-effect transistor. In addi-

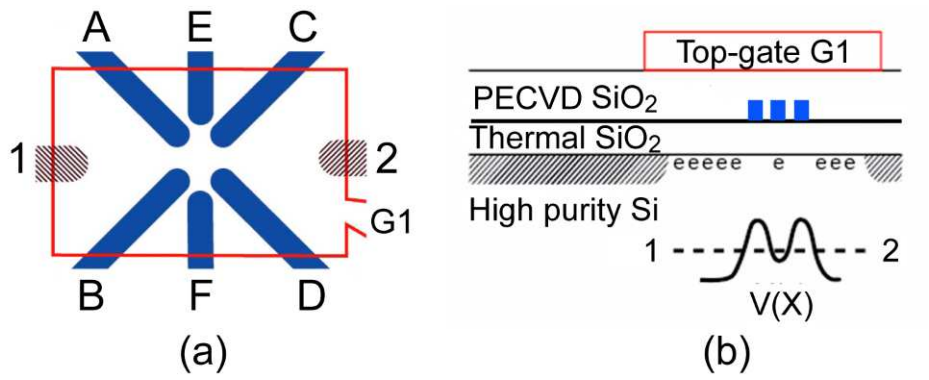

Figure 3.1: (a) Schematic top view and (b) cross sectional view of an MOS-SET. In (a), the shaded regions, the solid lines, and the rectangular area, depict the heavily phosphorus-doped Ohmic source (1) and drain (2) leads, the 6 side gates (A-F), and the top gate (G1), respectively. The schematic potential profile along the 1-2 direction under the single electron tunneling condition is shown as the inset (lower right) in (b).

tion, 6 side gates, located below the top gate and above the thermal oxide, screen the electric field from the top gate, depleting only the electrons below the side gates. Side gates A, B and side gates C, D are used to define the quantum dot and the two tunneling barriers. Side gates E, F, which are less effective to change the tunneling barriers, are used to squeeze the electrons out of the quantum dot, so they are called plunger gates. As shown in Fig. 3.2 (c) and (d), blue areas (low potential energy for electrons) are the source and drain regions, and the center pit defines the quantum dot. As the negative depletion voltages on plunger gates E, F are increased, the size of the quantum dot decreases, which will reduce the number of electrons in the dot. This bi-layer design offers the flexibility in the device layout and allows independent control over the 2D electron density, the tunneling conductance and the electron

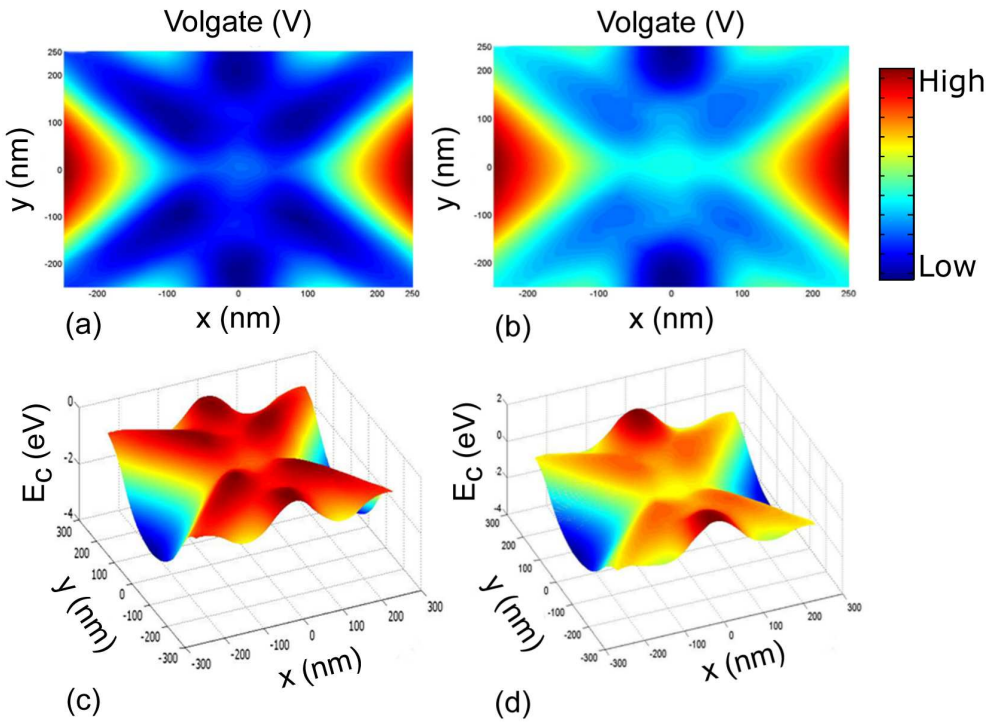

Figure 3.2: Electric potential distributions of the 2D electron gas (a), (b), and corresponding profiles of the electric potential energy for an electron (c), (d), simulated by Femlab. In (c), (d), blue areas (low potential energy for an electron) are the source and drain regions, and the center pit defines the quantum dot. As we increase the negative depletion voltage on plunger gates E, F from (c) to (d), the size of the quantum dot decreases. In the simulation, we assume that silicon is a insulator for simplicity, so these are only schematic graphs to show the basic operating principle of the device, and scales are arbitrary.

population in the quantum dot.

3.2 Fabrication process

The basic fabrication process of Si MOS-SETs is as follows. First, Ion Implantation is used to define the source and drain regions, and then thermal oxide is grown on it. After that, e-beam lithography and reactive ion etching (RIE) are used to make trenches to cut the leakage path from the top gate to the 2DEG at the active region (Fig. 3.5). Next, 6 side gates are defined by e-beam lithography. The second silicon oxide layer is grown on the top of the device using HDPECVD. Photolithography is used to define the top gate, and open via holes to make contacts to the side gates, the source and the drain. The devices are annealed in forming gas at 420 °C for 30 minutes as the last step. The fabrication recipe was first developed by Greg Jones, et al.,[13] and further improved by me. A complete recipe can be found in the Appendix C. Some important steps are discussed here.

3.2.1 Ohmic contacts by ion implantation

In order to be conductive at cryogenic temperatures, silicon must be heavily doped. The metal-insulator transition happens at 3.5×10^{18} dopants/cm^3 for phosphorus in doped silicon.[39] When ions implant into single-crystal material along a major crystal orientation, channeling can occur. So a 27 nm thick thermal oxide (amorphous) was grown before ion implantation, and the implantation was done off axis by 7°.[41] The conditions necessary to produce a degenerately doped Ohmic

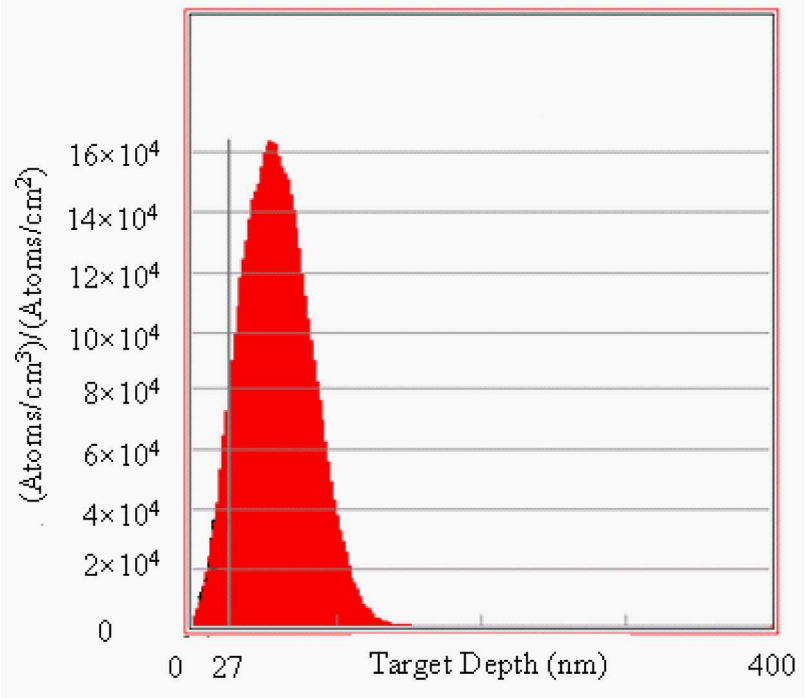

Figure 3.3: The concentration of phosphorous dopants as a function of depth in silicon simulated by SRIM. The phosphorous implantation energy is 40 KeV, and the surface has a 27 nm SiO_2 implant mask. (After Greg Jones, Ref.[40])

contact were simulated using a program called SRIM with different implantation energies, so that the peak ion concentration is located in silicon near the Si/SiO_2 interface. Figure 3.3 shows the 40 keV SRIM simulation result. In addition, these different implantation energies were tested using real devices, by implanting phosphorus into hall bar patterns with implantation energies of 20, 30, 40 keV, and the dose of 7×10^{14} dopants/cm^2. The recipe targets a dopant concentration above 7×10^{18} dopants/cm^3, which is well above the metal-insulator transition. The dopants were activated at 1000 °C. The devices were characterized at 4.2 K. It was determined that the 40 KeV recipe was the best with the lowest sheet resistivity of 63 Ω/square.[40]

After ion implantation, the 27 nm sacrificial oxide layer was removed and a fresh 27 nm oxide was re-grown at 1000 °C for 20 minutes. The thermal oxide regrowth is the only step which will cause appreciable dopant diffusion in the following process. The SUPREM IV is used to investigate the diffusion in the process. From the simulation (Fig. 3.4), the contour line for 1×10^{18} dopants/cm^3 is located at about 0.38 μm from the ion implantation edge ($x = 0$). Here we have already considered the oxidation enhanced diffusion. The minimum distance between two ion implantation regions is about 2 μm. It looks as if we are safe, but it is still possible that some phosphorus ions exist near the quantum dot. There is really no constraint on the minimum distance between two ion implantation regions. The minimum distance should be increased if we want to improve the device quality. Also the peak phosphorus concentration is about 1.86×10^{19} dopants/cm^3 after the thermal oxide regrowth, which is still well above the critical dopant concentration

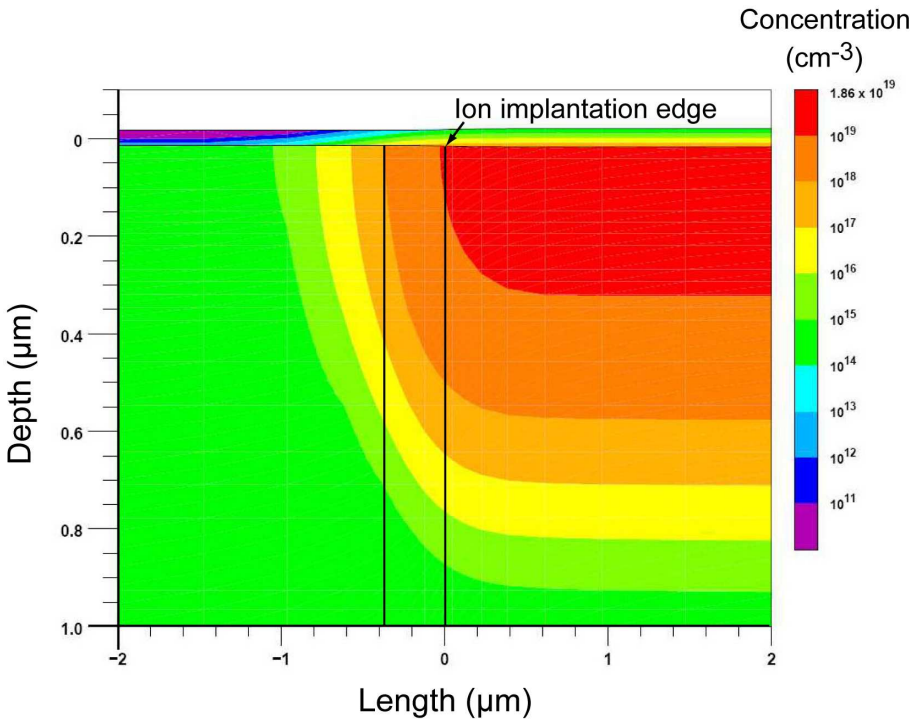

Figure 3.4: The distribution of phosphorous dopants after 1000 °C, 20 min annealing, simulated by SUPREM IV. The phosphorous implantation energy is 40 KeV, and the dose is 7×10^{14} dopants/cm^2. The initial phosphorous concentration is 1×10^{14} cm^{-3}, and the surface has a 27 nm SiO_2 implant mask.

of the metal-insulator transition.

3.2.2 Trench isolation

The device active region should be isolated from the environment to avoid any potential undesirable side effect. There are two methods. One is the pn junction isolation, which uses ion implantation to form a reverse biased pn junction structure. The other is the dielectric isolation, which uses a thick dielectric structure to separate the device from the environment. The trench isolation is one kind of the dielectric isolation.

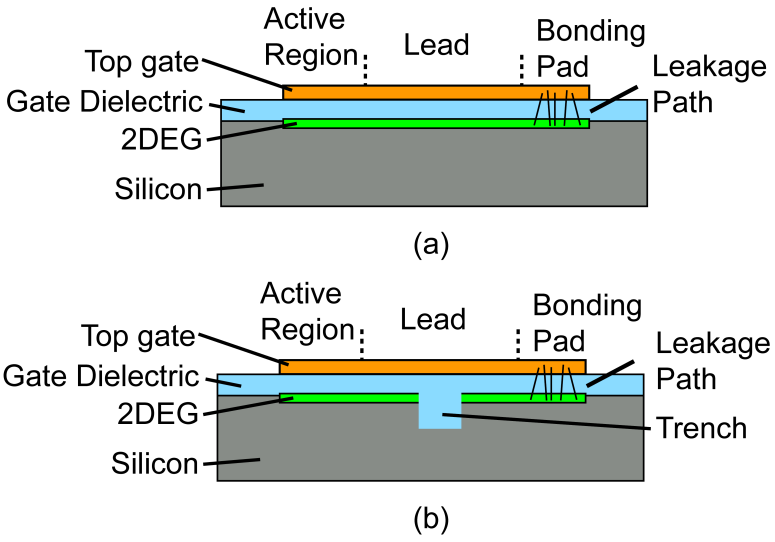

Figure 3.5: (a) Soldering a gold wire to the bonding pad introduces the damage of the oxide underneath the bonding pad. This forms a leakage path, and causes the gate leakage problem. (b)Shallow trench isolation can effectively cut the leakage path, and stop the gate leakage.

As illustrated in Fig. 3.5(a), when the positive-biased top gate induces the

Table 3.1: Etch rates of PMMA, SiO_2, Si using SI1DRH

PMMA	SiO_2	Si
335 nm/min	54 nm/min	610 nm/min

2DEG at the Si/SiO_2 interface, the 2DEG not only exists at the device active region, but also exists at the gate lead region and all way to the bonding pad region. When we solder a gold wire to the bonding pad, it causes the damage of the oxide underneath the bonding pad, and forms a leakage path. This results in the gate leakage from the top gate to the 2DEG at the device active region. The leakage path can be effectively cut by using shallow trench isolation underneath the gate lead as shown in Fig. 3.5(b). Greg first introduced this method in our SET device.[40]

But at that time, the results were still inconsistent. It turned out that the process to create the shallow trench had a problem. We use e-beam lithography to define the trenches, and then use reactive ion etch (RIE) (Recipe name: SI1DRH) for 1 minute to create them. PMMA, used in e-bream lithography as a resist, has poor resistance to plasma etching. I have measured the etch rates of PMMA, SiO_2, and Si using recipe SI1DRH, and they are shown in Table. 3.1.

In the previous recipe, the thickness of the PMMA layer was 180 nm, and the thickness of the thermal oxide was 27 nm. So it is clear that after the RIE etch for 1 minute, both the PMMA and the thermal oxide were etched away at the resist PMMA mask area, and the trench depth was about 0.3 μm at the trenches. Since the thermal oxide was etched away, this caused a serious leakage problem. After

this finding, I changed the PMMA 950 A4 (4% 950 PMMA solid in anisole solvent) into the PMMA 950 A8 (8% 950 PMMA solid in anisole solvent), and the thickness of the PMMA layer became 800 nm. The gate leakage problem was finally solved. We have got consistent results since then.

3.2.3 High density PECVD oxide

We have tried benzocyclobutene (BCB) as the second dielectric layer, and successfully fabricated some devices using it.[13] But there was leakage current between the top gate and the source, drain leads in most devices. So we switched to HDPECVD oxide since HDPECVD oxide is stable. We fabricated a MOS structure with a 27 nm thermal oxide and a 400 nm HDPECVD dioxide without side gates, and applied up to 155 V ($\sim$ 4 MV/cm) on this double-oxide layer without measurable leakage. So the quality of HDPECVD oxide is good. But we still suffered the leakage problem. Later, as I have discussed in the previous section, the problem was identified due to the wrong trench process used at that time. Now when I look back to the leakage problem with BCB, I suspect that this problem was due to the same cause.

The other problem I encountered is the positive charge in the HDPECVD oxide. As I will discussed in Sec. 4.3.2, once I found that the dielectric constant for this HDPECVD oxide became about 5.5 after some repair was done on the HDPECVD machine, which is larger than 3.9 for pure silicon oxide. The flatband voltage was $V_{fb} = -4.03$ V, and the threshold voltage was $V_t = -4.85$ V for the device. The

effective oxide charge concentration was $N_{ox} = 2.85 \times 10^{11}$ /cm^2. This positive HDPECVD oxide charge caused a problem, because it induced 2DEG all over the wafer. Fortunately, after changing the flow ratio of SiH_4 and N_2O from (4 sccm : 20 sccm) to (4 sccm : 40 sccm), (sccm: Standard Cubic Centimeters per Minute) the positive oxide charge inside the HDPECVD oxide was reduced, because HDPECVD oxide was changed from silicon-rich to oxygen rich. The dielectric constant also became about 3.9.

3.2.4 All e-beam defined side gates

The previous side-gate material Al/Ti/Au cannot sustain 400 °C annealing. Only 2 of 12 devices worked. So we switched to Al as the side-gate material. But our previous side-gate process required two steps: e-beam lithography was used to define the inner part of side gates (the smaller features), as shown in Fig. 3.6(a), and photo lithography was used to define the side gates' leads (the larger features), as shown in Fig. 3.6(b). Since we used Al as the inner part material, after Al was deposited, it immediately grew a native oxide layer, which is very difficult to penetrate. We used high temperature annealing (450 °C for 45 minutes) to penetrate the native oxide layer, but suffered serious side gates leakage. Finally we switched to one step process – using e-beam lithography to define the whole side gates, as shown in Figure 3.6(c). It turned out to be good, although it takes 3 hours to finish 12 devices.

After these changes, we have developed a reliable process to fabricate silicon SETs.

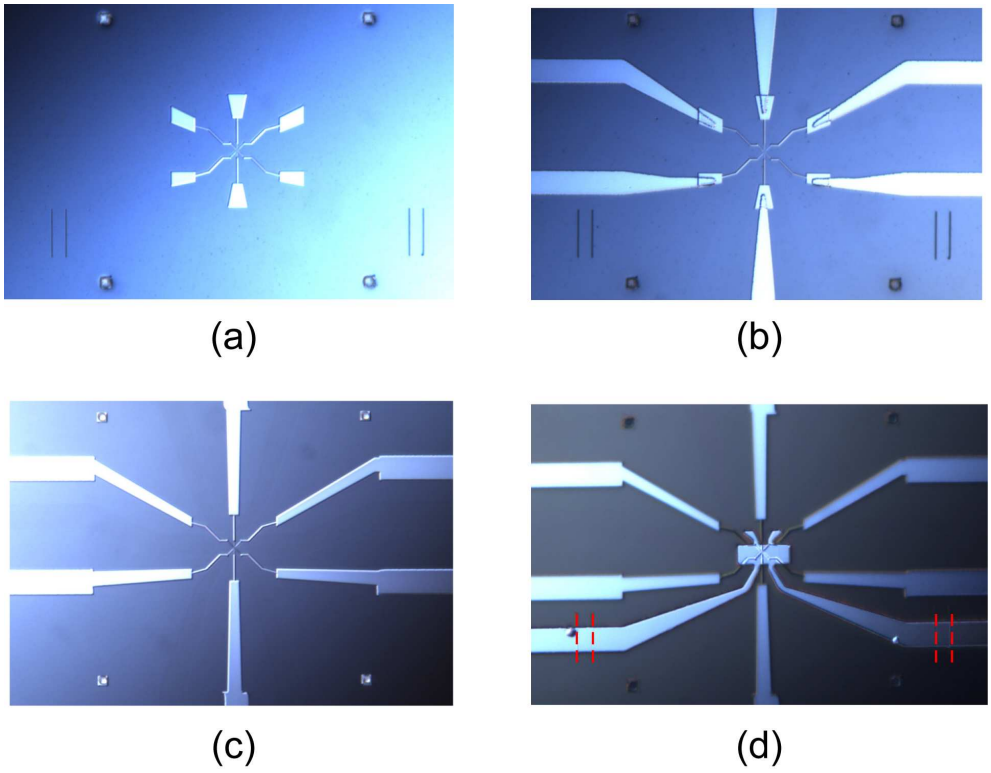

Figure 3.6: Side gate fabrication process. (a) and (b) are previous two-step process. (c) is the current all e-beam process. (d) is the final device. In (d), the dashed lines indicate the locations of the isolation trenches.

Chapter 4

Silicon metal-oxide-semiconductor field-effect transistors

4.1 Introduction

For a standard n-channel enhancement-mode metal-oxide-semiconductor field-effect transistor (MOSFET), the device is fabricated on a p-type substrate, with two heavily doped n-type regions, i.e. the source and the drain regions. A thin layer of silicon dioxide (SiO_2) is grown on the surface of the substrate, and then metal is deposited on the top of the oxide layer to form the gate. The device is normally off. When the top gate is sufficiently positively biased, above a threshold voltage, an electron channel forms at the silicon and silicon oxide interface, so electrons can flow from the source to the drain, thus the device is on. In addition, the electrons are strongly confined at the Si/SiO_2 interface, so they can only move freely along the interface and form a 2-dimensional electron gas (2DEG). The 2DEG can be depleted into a quantum dot using lateral electrostatical gates, and an MOS single electron transistor (SET) is created.

An MOS capacitor is first discussed in Sec. 4.2. For a practical MOS structure, interface traps and oxide charges always exist and affect the ideal MOS characteristics. We have used two methods, i.e. the capacitance-voltage(C-V) method and the conductance method, to measure the interface trap density in Sec. 4.3. 2D electron concentration and electron mobility are another two important parameters

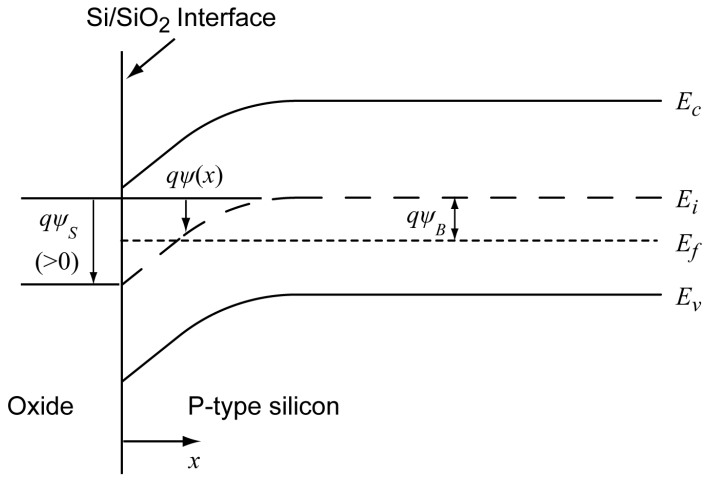

Figure 4.1: Energy band diagram near the surface of a p-type MOS capacitor. E_c, E_v, E_i and E_f are the energy levels of the conduction band edge, the valence band edge, the intrinsic Fermi level and the Fermi level, respectively. The band bending ψ is defined as positive when the bands bend downward with respect to the bulk.

to characterize the 2DEG. Hall measurements and Shubnikov-de Haas oscillation measurements are used to determine these two parameters in Sec. 4.4.

4.2 Metal-oxide-semiconductor(MOS) capacitors

4.2.1 MOS capacitors at room temperature

First we consider an ideal MOS capacitor without interface traps and oxide charges. The relations among the surface potential, charge distribution and electric field in silicon are derived by solving the Poisson's equation.[42] The band diagram near the surface of a p-type silicon MOS capacitor is shown in Fig. 4.1. Here we defined the band bending $\psi(x) = \psi_i(x) - \psi_i(x = \infty)$, where $x = 0$ is at the silicon

surface, $\psi_i(x)$ and $\psi_i(x = \infty)$ are the intrinsic potential at position x and in the bulk silicon respectively. (*Please note that ψ represents the electric potential, not a wavefunction in this chapter.*) So the boundary conditions are $\psi(x = \infty) = 0$ and $\psi(x = 0) = \psi_s$. The surface potential ψ_s depends on the applied gate voltage. Poisson's equation is

$$\frac{d^2\psi}{dx^2} = -\frac{dE}{dx} = -\frac{q}{\epsilon_{si}}[p(x) - n(x) + N_d^+(x) - N_a^-(x)] \tag{4.1}$$

where E is the electric field, q is the electron charge, ϵ_{si} is the permittivity of silicon, $p(x)$ is the hole concentration, $n(x)$ is the electron concentration, $N_d^+(x)$ and $N_a^-(x)$ are the densities of the ionized donors and acceptors respectively.

In bulk silicon, charge neutrality condition for a uniformly doped p-type silicon requires

$$N_d^+(x) - N_a^-(x) = -N_a + \frac{n_i^2}{N_a} \tag{4.2}$$

where N_a is the density of the acceptors and n_i is the intrinsic carrier density.

In the surface region,

$$p(x) = N_a e^{-q\psi/kT} \tag{4.3}$$

and

$$n(x) = \frac{n_i^2}{N_a} e^{q\psi/kT} \tag{4.4}$$

So we get

$$\frac{d^2\psi}{dx^2} = -\frac{q}{\epsilon_{si}}[N_a(e^{-q\psi/kT} - 1) - \frac{n_i^2}{N_a}(e^{q\psi/kT} - 1)] \tag{4.5}$$

From Eq. 4.5, we get

$$E^2(x) = (\frac{d\psi}{dx})^2 = \frac{2kTN_a}{\epsilon_{si}}[(e^{-q\psi/kT} + \frac{q\psi}{kT} - 1) + \frac{n_i^2}{N_a^2}(e^{q\psi/kT} - \frac{q\psi}{kT} - 1)] \quad (4.6)$$

At $x = 0$, let $E = E_s$. By Gauss's law, the space charge per unit area in silicon is

$$Q_s = -\epsilon_{si}E_s = \pm\sqrt{2\epsilon_{si}kTN_a}[(e^{-q\psi_s/kT} + \frac{q\psi_s}{kT} - 1) + \frac{n_i^2}{N_a^2}(e^{q\psi_s/kT} - \frac{q\psi_s}{kT} - 1)]^{1/2} \quad (4.7)$$

For a p-type silicon, when $\psi_s < 0$, the device is in the accumulation region with $Q_s \sim exp(q|\psi_s|/2kT)$. When $\psi_s = 0$, the device is in the flat-band condition with $Q_s = 0$. For $0 < \psi_s < \psi_B$, the device is in the depletion region with $Q_s \sim \sqrt{\psi_s}$. For $\psi_B < \psi_s$, the device is in the inversion region. For $\psi_B << \psi_s$, $Q_s \sim -exp(q\psi_s/2kT)$. Here ψ_B is the difference between the Fermi level and the intrinsic Fermi level in the bulk silicon, and is given by

$$\psi_B = \frac{kT}{q}ln(\frac{N_a}{n_i}). \quad (4.8)$$

The onset of the strong inversion is usually defined as $\psi_s(inv) = 2\psi_B$. The threshold voltage is given by

$$V_t = V_{fb} + 2\psi_B + \frac{\sqrt{4\epsilon_{si}qN_a\psi_B}}{C_{ox}}. \quad (4.9)$$

where V_{fb} is the flat-band voltage and C_{ox} is the oxide capacitance per unit area given by

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}, \quad (4.10)$$

where ϵ_{ox} is the permittivity of the oxide, and t_{ox} is the oxide thickness.

The differential capacitance in silicon is

$$C_{si} = -\frac{\partial Q_s}{\partial \psi_s}. \tag{4.11}$$

The total differential capacitance of the MOS capacitor at low frequency is

$$C = \frac{C_{ox} C_{si}}{C_{ox} + C_{si}}. \tag{4.12}$$

4.2.2 MOS capacitors at low temperatures

At low temperatures, the Fermi level moves from near the middle of the band gap to the band edge.

For an n-type silicon with only donors, charge neutrality condition requires

$$n = N_d^+ + p, \tag{4.13}$$

where n is the electron concentration, p is the hole concentration and N_d^+ is the concentration of ionized donors.

$$n = N_c e^{-(E_c - E_f)/kT},\ p = N_v e^{-(E_f - E_v)/kT} \tag{4.14}$$

and

$$N_d^+ = N_d[1 - f(E_d)] = N_d(1 - \frac{1}{1 + \frac{1}{2} e^{(E_d - E_f)/kT}}). \tag{4.15}$$

where N_c and N_v are the effective densities of states at the conduction band and the valence band respectively, E_c, E_v, E_f and E_d are the energy levels of the conduction band edge, the valence band edge, the Fermi level and the donor respectively.

At low temperatures, $p \sim 0$, so

$$\begin{aligned} n &\approx N_d^+ \\ N_c e^{-(E_c - E_f)/kT} &\approx \frac{N_d}{1 + 2e^{(E_f - E_d)/kT}} \end{aligned} \tag{4.16}$$

This is an algebra equation about E_f. We can get

$$E_f = E_d + kTln[-\frac{1}{4} + \frac{1}{4}\sqrt{1 + 8\frac{N_d}{N_c}e^{(E_c - E_d)/kT}}] \tag{4.17}$$

Since $N_c \sim T^{3/2}$, at very low temperatures, $N_d/N_c >> 1$, Eq.2.16 becomes

$$E_f \approx \frac{E_c + E_d}{2} + \frac{kT}{2}ln(\frac{N_d}{2N_c}) \tag{4.18}$$

At $T = 0K$, $E_f = \frac{E_c + E_d}{2}$.

In the case of a partially compensated semiconductor with $N_d >> N_a$, the approximate expression for the electron density is [43]

$$n \approx (\frac{N_d - N_a}{2N_a})N_c e^{(E_c - E_d)/kT} \tag{4.19}$$

for

$$N_a >> \frac{1}{2}N_c e^{(E_c - E_d)/kT}.$$

So

$$E_f \approx E_d + kTln(\frac{N_d - N_a}{2N_a}). \tag{4.20}$$

And at $T = 0K$, $E_f = E_d$.

In practice, there is always a small amount of acceptors (donors) in an n-type (p-type) silicon, so the Fermi level is located at the ionized-donor (ionized-acceptor) energy level.

At cryogenic temperatures, normal silicon wafers are nonconductive. Both n-type silicon and p-type silicon can be used as the substrate to form the 2DEG. The threshold voltage is defined as the gate voltage at which the conduction band edge is aligned with the Fermi level in the bulk.

For a p-type silicon (Fig. 4.1), the band bending is $\psi_s = E_g - (E_a - E_v)$, where E_g is the energy gap in silicon and E_a is the acceptor energy level. The threshold voltage is

$$V_t = V_{fb} + \psi_s + \frac{\sqrt{2\epsilon_{si} q N_a \psi_s}}{C_{ox}}. \tag{4.21}$$

For an n-type silicon, the band bending is $\psi_s = E_c - E_d$, and $V_t = V_{fb} + \psi_s$, since the silicon is in neutral.

For a real MOS capacitor, the threshold voltage is also affected by interface traps and oxide charges. But the MOS capacitor can still be approximately modeled as a parallel-plate capacitor

$$n_{2d} = \frac{C_{ox}}{q}(V_g - V_t) \tag{4.22}$$

where V_g is the top gate voltage, and $V_g > V_t$.

In this thesis, we mainly use n-type high purity silicon wafers to fabricate devices, since the threshold voltage in an n-type silicon is lower than that in a p-type silicon. But it also causes some undesirable effect. When the second SiO_2 layer, deposited by high-density plasma-enhanced chemical-vapor-deposition (HDPECVD), contains some positive charges, a 2DEG is induced all over the wafer, and the devices do not work properly.

For a silicon MOS-SET, the device works at V_g around 16 V with $n_{2d} =$

8.43×10^{11} cm^{-2}. These electrons are confined in a potential well at the Si/SiO_2 interface. The potential well can be approximated as a triangle potential

$$V(x) = \begin{cases} \infty & \text{if } x < 0 \\ eE_x x & \text{if } x \geq 0 \end{cases} \tag{4.23}$$

where E_x is the average electric field perpendicular to the interface, and x is the distance from the interface into the silicon.

The solutions of the Schrödinger equation with the triangle potential are Airy functions with eigenvalues E_j given by [44]

$$E_j = \left[\frac{3hqE_x}{4\sqrt{2m_x}}\left(j + \frac{3}{4}\right)\right]^{2/3}, j = 0, 1, 2, ... \tag{4.24}$$

where h is Planck's constant, and m_x is the effective mass of electrons perpendicular to the surface which is $0.98m_e$ for silicon (100).

For $V_g = 16$ V,

$$E_x \approx \frac{qn_{2d}}{2\epsilon_{si}} = 6.46 \times 10^4 \text{ V/cm}.$$

So $E_0 = 27.4$ meV, $E_1 = 47.9$ meV, and $\Delta E = E_1 - E_0 = 20.5$ meV.

On the other hand, for the 2DEG, the Fermi level is located at

$$E_f - E_0 = \frac{h^2 n_{2d}}{2\pi m^* g} = 5.3 \text{ meV}, \tag{4.25}$$

where the degeneracy g is 4 for silicon (100). So at a temperature $T \leq 4.2$ K (0.36 meV) the electrons only occupy the lowest subband of energy E_0, and form a 2DEG.

Similarly, for $V_g = 55$ V with $n_{2d} = 2.8 \times 10^{12}$ cm^{-2}, $E_0 = 61$ meV, $E_1 = 106.7$ meV, and $\Delta E = E_1 - E_0 = 45.7$ meV, whereas $E_f - E_0 = 17.5$ meV. So the electrons also occupy the lowest subband only. This is confirmed by the Shubnikov-de Haas

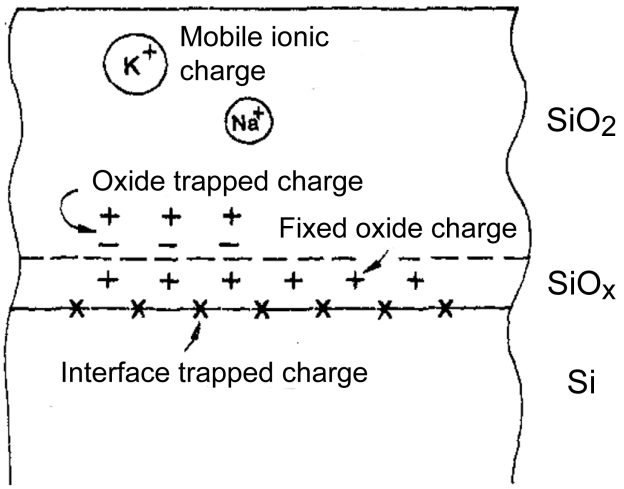

Figure 4.2: Charges and their locations in thermally oxidized silicon. (After Deal, Ref. [45].)

oscillations measurements. The Fourier transformation of the data only shows one major peak, which corresponds to the lowest subband with the degeneracy $g = 4$ at $V_g = 55$ V.

4.3 Oxide charges and interface trapped charges

Oxide charges and interface trapped charges play a significant role in the MOSFET transport properties. As shown in Fig. 4.2, they are classified as (1)interface trapped charge Q_{it}, (2)fixed oxide charge Q_f, (3)oxide trapped charge Q_{ot} and (4) mobile ionic charge Q_m.[45] (Here Q denotes the charge per unit area.) At cryogenic temperatures, mobile ionic charges are almost fixed. We can group the last three types of charges into one, i.e. the effective net oxide charge Q_{ox}, because they can not exchange charges with the 2DEG at the Si/SiO_2 interface. They only cause the

threshold voltage to shift by

$$\Delta V_t = -\frac{Q_{ox}}{C_{ox}}, \tag{4.26}$$

and

$$V_{fb} = \phi_{ms} - \frac{Q_{ox}}{C_{ox}}, \tag{4.27}$$

where ϕ_{ms} is the work function difference between the gate and the silicon substrate.

On the other hand, interface trapped charges are electrons and holes trapped in the Si/SiO_2 surface states. These surface states can exchange electrons with the 2DEG, and reduce the conduction current by trapping them. These trapped electrons and holes can also act as charged scattering centers at the interface, thus lowering the electron mobility. So the density of surface states or the density of interface traps is an important parameter in determining the quality of the Si/SiO_2 interface.

We have used two different methods to estimate the density of interface traps. One is the capacitance method (C-V method), and the other is the conductance method.

4.3.1 Capacitance method (C-V method)

The most convenient C-V method is the Terman method, a room-temperature, high-frequency (HF) capacitance method.[46] It assumes that the measurement frequency is sufficiently high that interface traps do not respond to the AC probe frequency. They only respond to the slowly varying dc gate voltage and stretch out the C-V curve. The interface-trap density D_{it} is determined by comparing the

experimental curve with the ideal curve

$$D_{it} = \frac{C_{ox}}{q^2}\frac{d\Delta V_g}{d\psi_s} \tag{4.28}$$

where $\Delta V_g = V_g - V_g(ideal)$ is the voltage shift of the experimental curve from the ideal curve. This method is considered to be useful for measuring interface trap densities of 10^{10} $\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ and above. The measurement frequency is usually 1 MHz.

Our silicon wafers are uniformly doped Si(100), and can be described by the doping concentration N_a for p-type or N_d for n-type.

In order to use the Terman method, we need to calculate the theoretical high frequency curves. First we consider a p-type uniformly doped silicon without oxide charges and interface trapped charges.

$$V_g = V_{ox} + \psi_s = -\frac{Q_s}{C_{ox}} + \psi_s, \tag{4.29}$$

where V_{ox} is the potential drop across the oxide, and Q_s is the space charge per unit area.

The total MOS capacitance is define as

$$C = \frac{d(-Q_s)}{dV_g} \tag{4.30}$$

From Eq. 4.11, 4.29 and 4.30,

$$C = \frac{C_{ox}C_{si}}{C_{ox} + C_{si}}. \tag{4.31}$$

When the MOS capacitor is in the inversion region, Eq. 4.11 is valid only at low-frequency, since it assumes that the minority carriers (the inversion charges) are

able to follow the applied AC probe signal. In order for the inversion charges to respond, the space-charge region current $J_{scr}(= qn_iW/\tau_g)$ must be able to supply the required displacement current $J_d(= CdV_g/dt \approx C_{ox}dV_g/dt)$, where W is the depletion layer width and τ_g is the electron-hole pair generation lifetime. So

$$\frac{dV_g}{dt} \leq \frac{qn_iW}{\tau_g C_{ox}}. \tag{4.32}$$

For oxide thickness $t_{ox} = 400$ nm, $W = 1$ μm, and $\tau_g = 10$ μs, $dV_g/dt = 2.7$ V/s, so the frequency should be less than $dV_g/dt/(2\pi V_{ac}) = 8.5$ Hz for AC excitation voltage $V_{ac} = 50$ mV. This is much lower than 1 MHz. So we need to use a different formula to calculate the high-frequency C-V curve in the inversion region.

When measuring the high-frequency C-V curves, the dc voltage sweep rate must be sufficiently low to generate the necessary inversion charges, according to Eq. 4.32.

Here we use the formula derived by J.R. Brews.[47] The basic assumptions are

1. The total number of minority carriers in the inversion layer is fixed by the dc gate bias and doesn't respond to the AC probe voltage.

2. The minority carriers can move spatially in the inversion layer in response to the high frequency probe voltage. These minority carriers are governed by a constant quasi-Fermi level.

Under these assumptions, the high-frequency capacitance in inversion is

$$\begin{aligned} C_{si} &= C_{FB}\left[1 - e^{-u_s} + \left(\frac{n_i}{N}\right)^2 (e^{u_s} - 1)\frac{\Delta}{1+\Delta}\right. \\ &\quad \left. + \left(\frac{n_i}{N}\right)^2 u_s \frac{1}{1+\Delta}\right] [F(u_s, u_B)]^{-1}, \end{aligned} \tag{4.33}$$

where

$$\Delta = \frac{F(u_s, u_B)}{(e^{u_s} - 1)} \left[\int_0^{u_s} du \left(\frac{e^u - e^{-u} - 2u}{F(u, u_B)^3} \right) - 1 \right], \tag{4.34}$$

$$F(u_s, u_B) = \sqrt{2} \left[e^{-u_s} + u_s - 1 + e^{-2u_B}(e^{u_s} - u_s - 1) \right]^{1/2}, \tag{4.35}$$

$$u_s = \frac{q\psi_s}{kT},$$

and

$$u_B = ln\frac{N_a}{n_i}.$$

Other equations used in analyzing HF C-V curves are summarized here.

1. **Extracting the doping concentration** N_a

   The maximum depletion width

   $$W_{dm} = \sqrt{\frac{4\epsilon_{si} kT ln(N_a/n_i)}{q^2 N_a}}, \tag{4.36}$$

   and the minimum HF capacitance

   $$\frac{1}{C_{min}} = \frac{1}{C_{ox}} + \frac{1}{\epsilon_{si}/W_{dm}} = \frac{1}{C_{ox}} + \sqrt{\frac{4kTln(N_a/n_i)}{\epsilon_{si} q^2 N_a}}. \tag{4.37}$$

   The gate area s of the MOS capacitor is known. From the measured HF C-V curve, we can get the maximum capacitance $C_{m\_max}(\approx C_{ox}s)$ and the minimum capacitance $C_{m\_min}(\approx C_{min}s)$. So

   $$\frac{C_{ox}}{C_{min}} = 1 + C_{ox}\sqrt{\frac{4kTln(N_a/n_i)}{\epsilon_{si} q^2 N_a}} \tag{4.38}$$

   $$\frac{C_{m\_max}}{C_{m\_min}} \approx 1 + \frac{C_{m\_max}}{s}\sqrt{\frac{4kTln(N_a/n_i)}{\epsilon_{si} q^2 N_a}} \tag{4.39}$$

   From Eq. (4.39), N_a is approximately determined.

2. **Extracting the oxide capacitance C_{ox}**

C_{ox} is extracted when the MOS capacitor is in strong accumulation. In strong accumulation, $Q_s \propto exp(-q\psi_s/2kT)$, so $C_{si} = -dQ_s/d\psi_s = (q/2kT)Q_s = (q/2kT)C_{ox}|V_g - V_{fb} - \psi_s|$, and the MOS capacitance is

$$\frac{1}{C} = \frac{1}{C_{ox}} + \frac{2kT}{qC_{ox}} \frac{1}{|V_g - V_{fb} - \psi_s|}. \tag{4.40}$$

where the flatband voltage V_{fb} is the gate voltage at the flatband condition.

Since $2kT/q \approx 0.052$ V, ψ_s is limited to 0.1 to 0.3 in accumulation, and $V_g - V_{fb} \gg \psi_s$ in strong accumulation.

At the flat band condition, the Debye length

$$L_D = \sqrt{\frac{\epsilon_{si} kT}{q^2 N_a}}, \tag{4.41}$$

and the flatband capacitance per unit area

$$\begin{aligned} \frac{1}{C_{fb}} &= \frac{1}{C_{ox}} + \frac{L_D}{\epsilon_{si}} \\ &\approx \frac{s}{C_{m\_max}} + \frac{L_D}{\epsilon_{si}}. \end{aligned} \tag{4.42}$$

The flatband capacitance is $C_{m\_fb} = C_{fb}s$. The flatband voltage is determined from the measured HF C-V curve using the flatband capacitance.

So in strong accumulation, the measured MOS capacitance

$$\frac{1}{C_m} \approx \frac{1}{C_{ox}s} + \frac{2kT}{qC_{ox}s} \frac{1}{|V_g - V_{fb}|}. \tag{4.43}$$

where s is the gate area.

After Linear fitting the measured data $1/C_m$ v.s. $1/|V_g - V_{fb}|$ in strong accumulation, the intercept is $1/C_{ox}s$. Thus C_{ox} can be accurately determined.

In addition, N_a can be more accurately calculated using Eq. (4.38) with this C_{ox}.

4.3.2 Experimental MOS C-V curves

C-V curves are measured using an HP 4194A impedance analyzer in series capacitance and series resistance mode ($C_s - R_s$ mode) while sweeping the bias voltage. The AC excitation is 50 mV, 1 Mhz. We have used these C-V curves to extract device parameters, and to estimate the interface-trap density. We have also investigated the effect of fabrication steps on the interface-trap density.

The thermal oxide is grown using an MOS-grade furnace in NIST. Sample (P_thermal_oxide) is an MOS capacitor on a p-type device wafer with 27 nm thermal oxide, and an aluminum top gate using e-beam evaporation. Fig. 4.3(a) and (c) show the C-V curves before and after forming gas annealing. The forming gas annealing effectively reduces the interface-trap density D_{it} from 10^{12} cm^{-2}eV^{-1} to less than 10^{11} cm^{-2}eV^{-1}, the limit of the Terman method.

We also find that the interface-trap density of $\sim 10^{12}$ cm^{-2}eV^{-1} is due to the e-beam evaporation process for the top gate metallization, since the device shows the similar C-V curves (Data not show here.) with $D_{it} \sim 10^{12}$ cm^{-2}eV^{-1}, if it is forming gas annealed first, and then the aluminum top gate is deposited using a CHA e-beam evaporator as the last step. (CHA Industries is a company name.) The similarity is not a surprise, because when the oxide was thermally grown at NIST, the forming gas annealing was the last step. The Si/SiO_2 interface is damaged by the

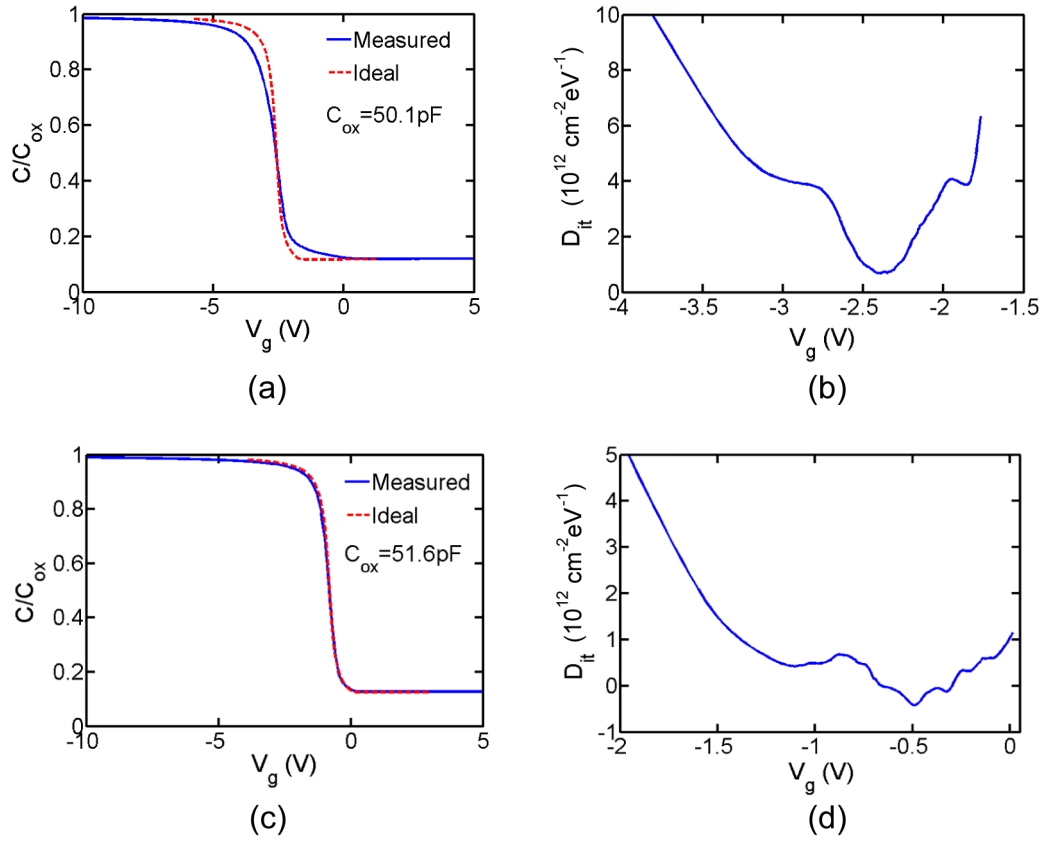

Figure 4.3: Characterization of the p-type device wafer using Sample (P_thermal_oxide). (a) and (c) are high frequency C-V curves before and after forming gas annealing. The blue curves are measured data, and the red curves are from theoretical simulation using J.R. Brews' formula. (b) and (d) show the interface-trap densities D_{it} extracted from (a) and (c) using the Terman method. The limitation of the Terman method is $\sim 10^{10}$ $\mathrm{cm}^{-2}\mathrm{eV}^{-1}$.

X-Ray radiation from the e-beam evaporator, and a thermal evaporation process is preferred, because the resistive thermal evaporation doesn't have X-Ray radiation.

From these C-V curves and the gate area $s = (200\,\mu\text{m})^2 = 4 \times 10^{-4}\,\text{cm}^2$, the device parameters can be extracted. From Fig. 4.3(c), the doping concentration N_a is $2.75 \times 10^{15}\,\text{cm}^{-3}$, which is consistent with the wafer's resistivity $1 \sim 10\,\Omega\text{cm}$ in the specification from the manufacturer. The oxide thickness d is 26.7 nm, which is also consistent with the target thermal oxide thickness 27 nm. The flatband voltage is $V_{fb} = -0.813$ V and the threshold voltage is $V_t = 0.0033$ V for the device. The effective oxide charge is $Q_{ox} = C_{ox}(\phi_{ms} - V_{fb}) = -1.5 \times 10^{-9}$ C/cm^2, and the effective oxide charge concentration is $N_{ox} = Q_{ox}/q = -9.4 \times 10^9$ /cm^2, which is already in the limit of the Terman method, and thus not a valid result, where the work function difference between the aluminum gate and the p-type silicon substrate ϕ_{ms} is -0.8247 V here.

For an n-type device wafer, the HF C-V curve after forming gas annealing is shown in Fig. 4.4(a). From this C-V curve and the gate area $s = 4 \times 10^{-4}$ cm^2, the doping concentration is $N_d = 9.23 \times 10^{13}\,\text{cm}^{-3}$ and the oxide thickness d is 26.2 nm. The flatband voltage is $V_{fb} = -0.287$ V and the threshold voltage is $V_t = -0.769$ V for the device. Because ϕ_{ms} is -0.283 V here, the effective oxide charge concentration is less than 1×10^{10} /cm^2. It is also in the limit of the Terman method, and not a valid number.

MOS capacitors with oxide deposited by high-density plasma-enhanced chemical-vapor-deposition (HDPECVD) are also investigated. Sample (N_HDPECVD_oxide) is an MOS capacitor on the n-type device wafer with 27 nm thermal oxide and 400

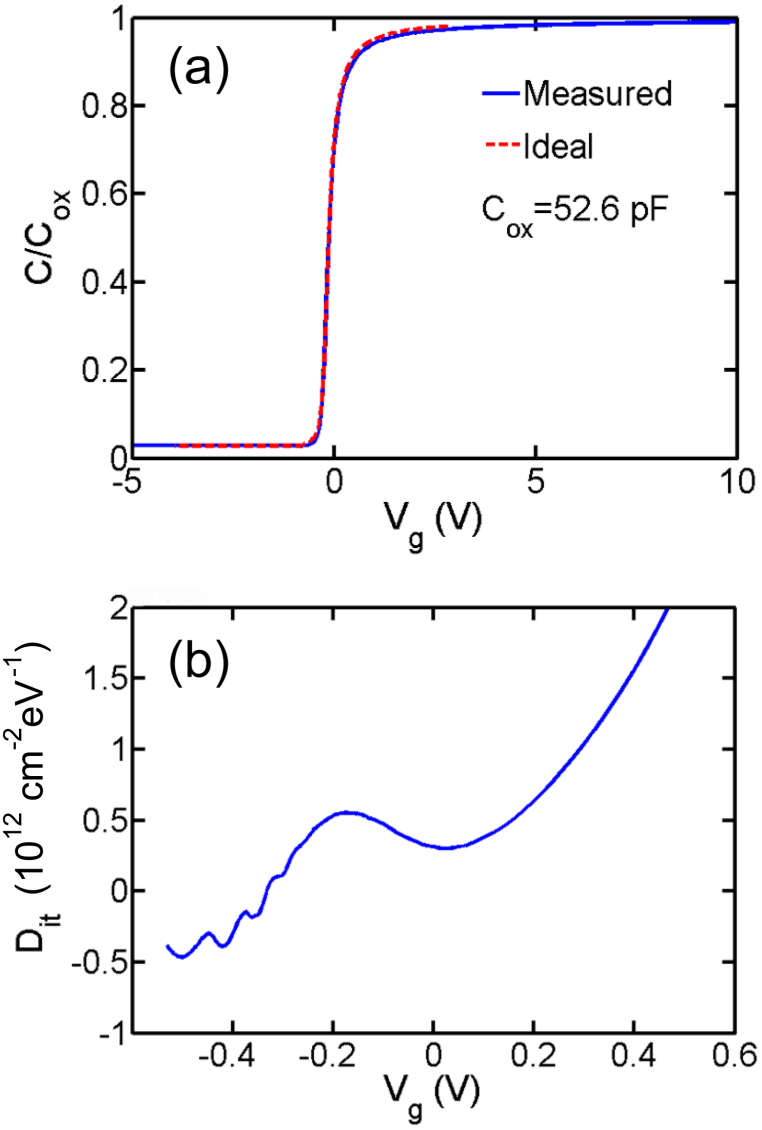

Figure 4.4: Characterization of the n-type device wafer using Sample (N_thermal_oxide). (a) High frequency C-V curve after forming gas annealing. The blue curve is measured data, and the red curve is from theoretical simulation. (b) Interface-trap density D_{it} extracted from (a) using the Terman method. The limitation of the Terman method is $\sim 10^{10}$ cm^{-2}eV^{-1}.

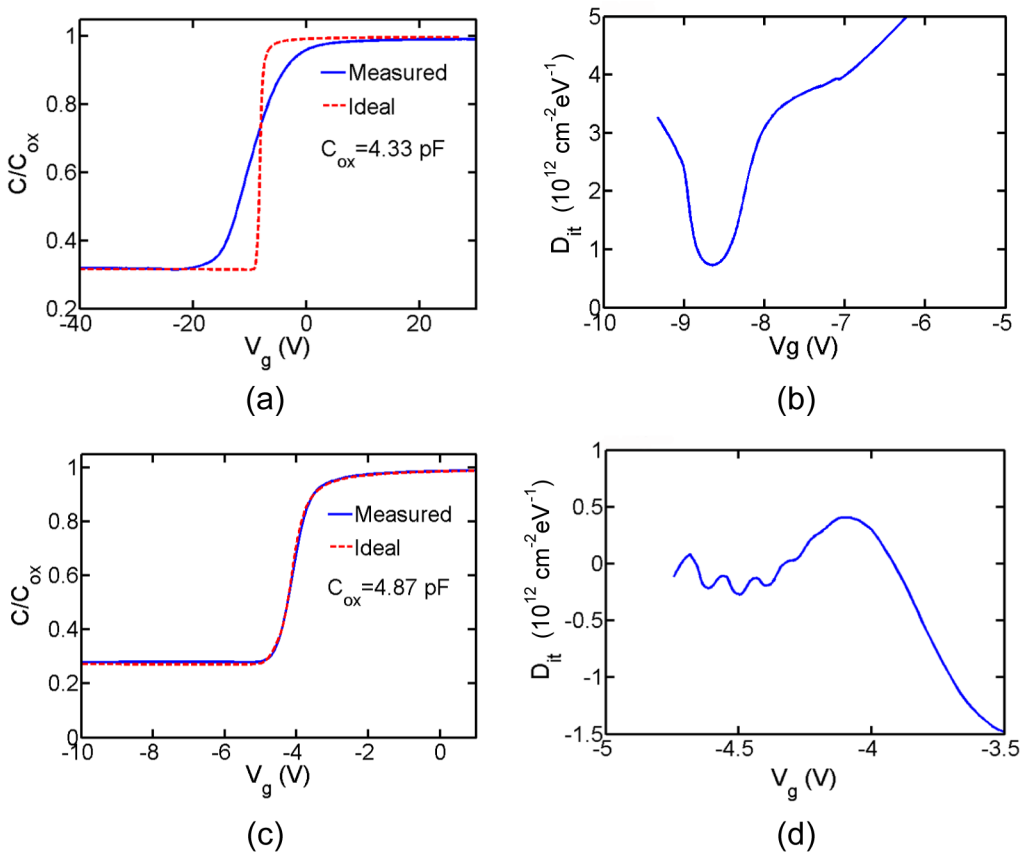

Figure 4.5: Characterization of the n-type device wafer with oxide by high density chemical vapor deposition(HDPECVD) using Sample (N_HDPECVD_oxide). (a) and (c) are high frequency C-V curves before and after forming gas annealing. The blue curves are measured data, and the red curves are from theoretical simulation as before. (b) and (d) show the interface-trap densities D_{it} extracted from (a) and (c) using the Terman method. The limitation of the Terman method is $\sim 10^{10}$ cm^{-2}eV^{-1}.

nm HDPECVD oxide, and an aluminum top gate. As shown in Fig. 4.5(b) and (d), the interface-trap density (about 10^{12} $cm^{-2}eV^{-1}$) reduces to less than 10^{11} $cm^{-2}eV^{-1}$ after forming gas annealing. Since e-beam vapor deposition process can introduce the interface-trap density of $\sim 10^{12}$ $cm^{-2}eV^{-1}$, it is unclear whether the HDPECVD process will cause further damage to the Si/SiO_2 interface or not. Because there is a 27 nm thermal oxide protecting the interface, the damage from the HDPECVD process could be small.

From the HF C-V curve in Fig. 4.5(c) and the gate area $s = 4 \times 10^{-4}$ cm^2, the doping concentration is $N_d = 1.20 \times 10^{14}\,cm^{-3}$. The HDPECVD oxide thickness is about 400 nm measured by an N&K thin film analyzer, and the capacitance is $C_{ox} = 4.87$ pF, so the dielectric constant for this HDPECVD oxide is about 5.5, which is larger than 3.9 for pure silicon oxide. The flatband voltage is $V_{fb} = -4.03$ V and the threshold voltage is $V_t = -4.85$ V for the device. Because ϕ_{ms} is -0.276 V here, the effective oxide charge concentration is $N_{ox} = C_{ox}(\phi_{ms} - V_{fb})/q = 2.85 \times 10^{11}$ $/cm^2$. This positive HDPECVD oxide charge causes a problem later, because it induces 2DEG all over the wafer. Fortunately, after changing the flow ratio of SiH_4 and N_2O from (4 sccm : 20 sccm) to (4 sccm : 40 sccm), (sccm: Standard Cubic Centimeters per Minute) the positive oxide charge inside the HDPECVD oxide is reduced, because HDPECVD oxide is changed from silicon-rich to oxygen rich.

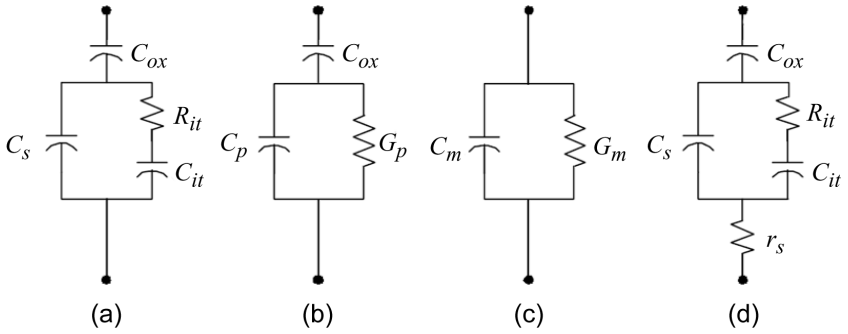

Figure 4.6: Equivalent circuits for the conductance method. (a) An MOS capacitor with interface traps. (b) Simplified circuit of (a). (c) Equivalent circuit in measurement. (d) Including series resistance r_s for a real device. (After D. K. Schroder, Ref. [48].)

4.3.3 Conductance method

As we can see, the HF C-V method is useful to extract device parameters, and to estimate the interface-trap density above 10^{11} $\text{cm}^{-2}\text{eV}^{-1}$, but it is not adequate to estimate the interface-trap density less than 10^{11} $\text{cm}^{-2}\text{eV}^{-1}$ after forming gas annealing. In order to estimate the interface-trap density in the final device, the conductance method is used.

The conductance method was introduced by Nicollian and Goetzberger.[49] It can detect the interface-trap densities of 10^{9} $\text{cm}^{-2}\text{eV}^{-1}$. The technique is based on measuring the equivalent parallel conductance G_p as a function of frequency in an MOS capacitor. The conductance comes from the lossy nature of interface-trap capture and emission of carriers. So it can be used to determine the interface-trap density.

Fig. 4.6(a) shows the equivalent circuit. It consists of the oxide capacitance

C_{ox}, the silicon capacitance C_s, and the interface-trap capacitance C_{it}. The capture and emission of carriers by D_{it} is a lossy process, and represented by the resistance R_{it}. So

$$C_p = C_s + \frac{C_{it}}{1 + (\omega\tau_{it})^2} \tag{4.44}$$

$$\frac{G_p}{\omega} = \frac{q\omega\tau_{it}D_{it}}{1 + (\omega\tau_{it})^2} \tag{4.45}$$

where $C_{it} = qD_{it}$, ω is the angular frequency and $\tau_{it} = R_{it}C_{it}$, the interface trap time constant.

Eq. 4.44 and 4.45 are for interface traps with a single energy level. However, interface traps are continuously distributed in the energy band gap, and the equation changes to [49]

$$\frac{G_p}{\omega} = \frac{qD_{it}}{2\omega\tau_{it}} ln[1 + (\omega\tau_{it})^2]. \tag{4.46}$$

Due to interface-trap time constant dispersion caused by surface potential fluctuations, the analysis is further complicated. As a simple estimation, the Eq. 4.46 is used here. An approximate expression giving the interface trap density in terms of the maximum conductance is

$$D_{it} \approx \frac{2.5}{qs} \left(\frac{G_p}{\omega}\right)_{max} \tag{4.47}$$

when $\omega\tau_{it} \approx 2$, where s is the gate area.[48]

In order to successfully use the conductance method, the series resistance should be minimized both in the preparation of samples and in the measurement; otherwise, there will be no conductance peak due to the large series resistance.

For the n-type high-purity device wafer, the series resistance is more then 10 kΩ in the samples, and the measured conductance curves do not show peaks even after correction. So we only use the p-type device wafer to estimate the interface-trap density in the device here.

In strong accumulation, C_{ox} is extracted as in Sec. 4.3.1, and the series resistance r_s is determined when the HP 4194A impedance analyzer is in $C_s - R_s$ mode.

The conductance curve is measured using the HP 4194A impedance analyzer in $C_p - G_p$ mode while sweeping the frequency at a fixed bias voltage. The measured data are denoted as C_m and G_m. When the series resistance r_s is taken into account,

$$\frac{G_p}{\omega} = \frac{\omega G_c C_{ox}^2}{G_c^2 + \omega^2 (C_{ox} - C_c)^2}, \tag{4.48}$$

where

$$C_c = \frac{C_m}{(1 - r_s G_m)^2 + (\omega r_s C_m)^2}, \tag{4.49}$$

and

$$G_c = \frac{\omega^2 r_s C_m C_c - G_m}{r_s G_m - 1}. \tag{4.50}$$

Fig. 4.7 shows the measured data G_p/ω v.s. f. The devices are biased near the middle band gap. Fig. 4.7(a) and (b) are measured on Sample (P_thermal_oxide) before and after forming gas annealing. The corresponding interface-trap densities are 1.09×10^{12} cm^{-2}eV^{-1} for the device before forming gas annealing, which is consistent with the result in Fig. 4.3(b), and 8.32×10^{10} cm^{-2}eV^{-1} after forming gas annealing. Fig. 4.7(c) and (d) are measured on Sample (P_HDPECVD_oxide) before and after forming gas annealing. The corresponding interface-trap densities

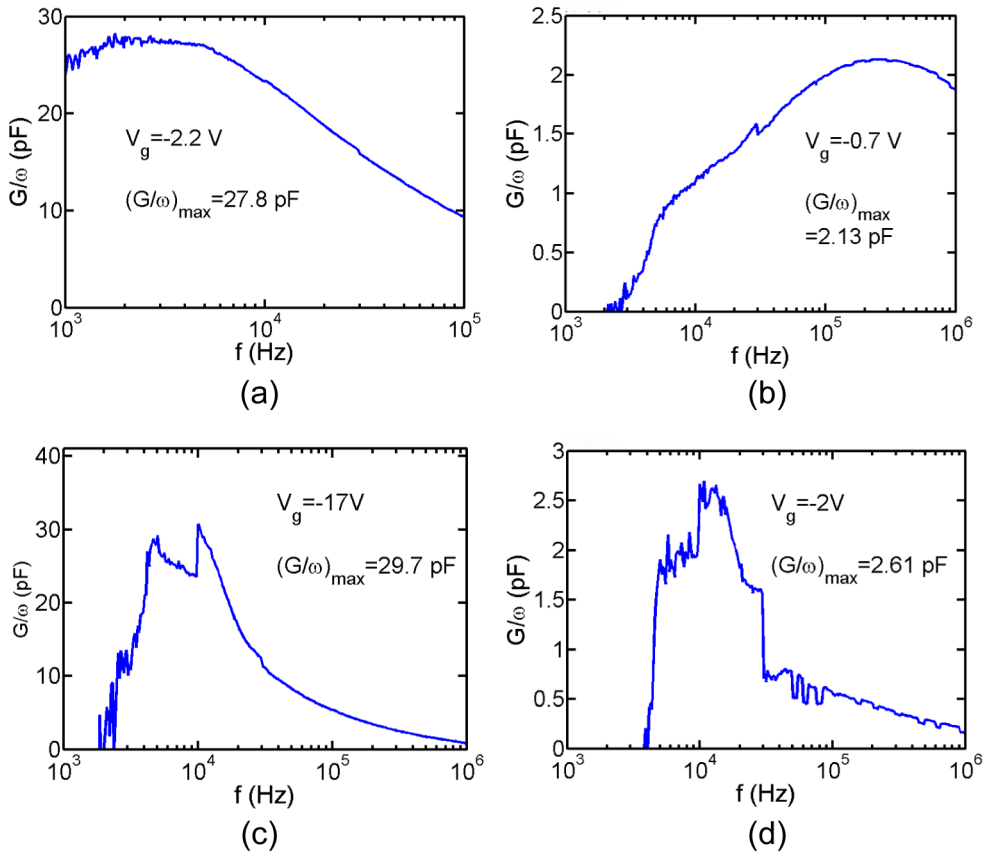

Figure 4.7: Experimental G_p/ω v.s. f curves. (a) and (b) are conductance curves on Sample (P_thermal_oxide) before and after forming gas annealing. (c) and (d) are conductance curves on Sample (P_HDPECVD_oxide) before and after forming gas annealing. For all devices, the gate area $s = 4 \times 10^{-8}$ cm^2.

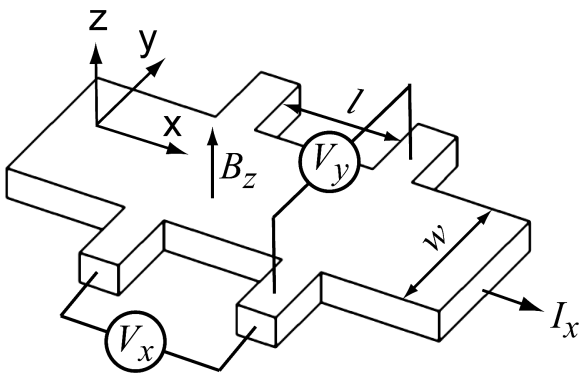

Figure 4.8: Basic setup for hall measurements.

are 1.16×10^{12} cm^{-2}eV^{-1} before forming gas annealing, and 1.02×10^{11} cm^{-2}eV^{-1} after forming gas annealing.

So the interface-trap density in the final device is estimated to be about 1×10^{11} cm^{-2}eV^{-1}.

4.4 2D electron concentration and electron mobility

The 2D electron concentration as a function of top gate voltage is determined using hall measurements[50] and Shubnikov-de Haas oscillation measurements[51].

The characterized device has the same structure as the final MOS-SET, which is on an n-type high purity Si(100) wafer, but without 6 side gates. The basic setup for hall measurements is shown in Fig. 4.8. 1 μA 37 Hz AC current I_x runs through the hall bar pattern. The longitudinal potential difference V_x and the transverse potential difference V_y are measured using lock-in technique.

$$\rho_{xy} = \frac{E_y}{J_x} = \frac{E_y w}{J_x w} = \frac{V_y}{I_x} \tag{4.51}$$

$$J_x = qnv \tag{4.52}$$

$$qE_y = qvB_z \tag{4.53}$$

where E_y is the transverse electric field, J_x is the longitudinal current density, n is the electron concentration, v is the electron drift velocity, and B_z is the perpendicular magnetic field.

So

$$\frac{V_y}{I_x} = \frac{B_z}{qn} \tag{4.54}$$

A more detailed analysis shows that [43]

$$\frac{V_y}{I_x} = r\frac{B_z}{qn} \tag{4.55}$$

where r is the hall scattering factor, which is between 1 and 2. At high magnetic field, $r \sim 1$. In our hall measurements, $r \approx 1$ after comparing the hall measurement results with the result from Shubnikov-de Haas oscillation measurements, as we will discuss later.

The measured V_y/I_x is linearly fitted with B_z, and the 2D electron concentration is determined using Eq. 4.54. The final results are shown in Fig. 4.9. The solid line is a linear fit of n_{2d} with V_g using a parallel capacitor model, and

$$n_{2d} = 3.376 \times 10^{10} + 5.059 \times 10^{10} V_g \ \ \text{cm}^{-2}. \tag{4.56}$$

At $V_g = 55$ V, ρ_{xx} is measured while magnetic field B is swept from 0 to 9 T, as shown in Fig. 4.10(a). Shubnikov-de Haas oscillations are observed. After the

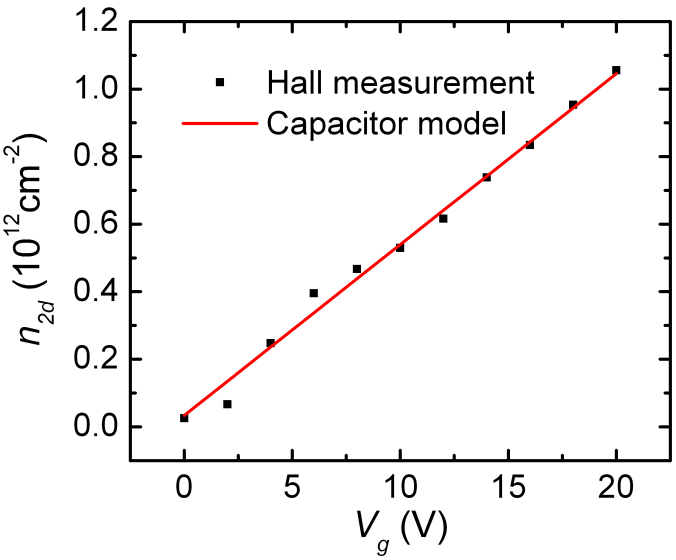

Figure 4.9: 2D electron concentration n_{2d} as a function of top gate voltage V_g using hall measurements, assuming the hall scattering factor $r = 1$. The solid line is a linear fit of n_{2d} with V_g using a parallel capacitor model.

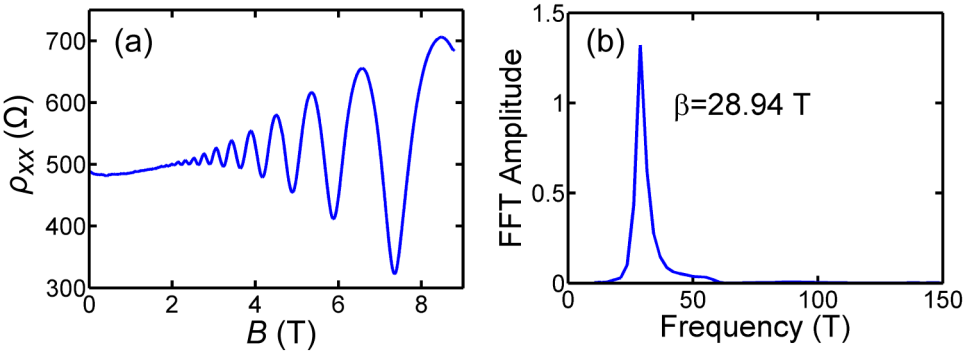

Figure 4.10: (a) Shubnikov-de Haas oscillations at $V_g = 55$ V. (b) The Fourier transformation of the data in (a).

Fourier transformation of ρ_{xx} as a function of $1/B$, the peak is located at $\beta = 28.94$ T. The 2D electron concentration is [52]

$$n_{2d} = g\frac{q\beta}{h} = 2.807 \times 10^{12} \text{ cm}^{-2}, \tag{4.57}$$

where the degeneracy g is 4, including the spin degeneracy and the 2 fold valley degeneracy for silicon(100), and h is the Planck constant. This is consistent with the result 2.816×10^{12} cm^{-2} from Eq. 4.56. So $r \approx 1$ here. The Fourier transformation of the Shubnikov-de Haas oscillations only shows one major peak, corresponding to the lowest subband with the degeneracy $g = 4$, which means that the electrons only occupy the lowest subband, otherwise there will be multiple peaks at different frequencies.

Compared Eq. 4.56 with the parallel capacitor model $n_{2d} = C_{ox}(V_g - V_t)/q$, $C_{ox} = 8.09$ nC/cm^2 and $V_t = -0.667$ V. The HDPECVD oxide thickness is about 400 nm measured by an N&K thin film analyzer, so the dielectric constant is about 3.91 in this final device, which is consistent with 3.9 for pure silicon oxide. So changing the flow rate of SiH_4 and N_2O not only reduces the positive oxide charge inside the HDPECVD oxide, but also changes the dielectric constant from 5.5 to 3.91, as previously discussed in Sec. 4.3.2.

At $B = 0$ T,

$$\rho_{xx} = \frac{E_x}{J_x} = \frac{E_x}{qn\mu E_x} = \frac{1}{qn\mu},$$

where μ is the conductivity mobility, and n is the electron concentration from previous measurements. So

$$\mu = \frac{1}{qn\rho_{xx}}, \tag{4.58}$$

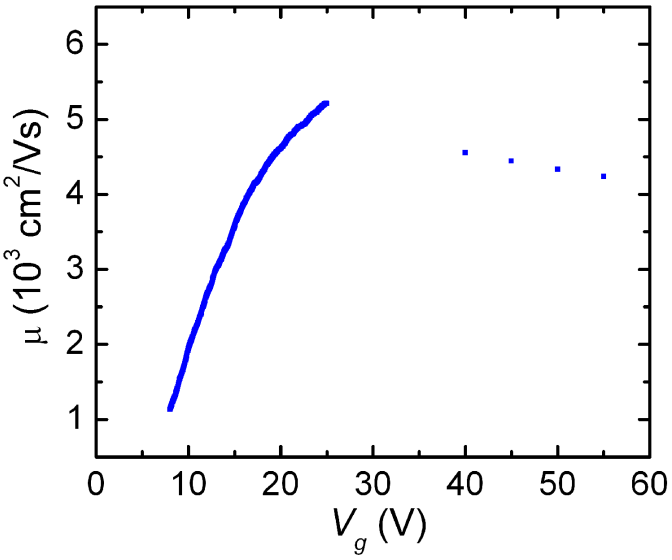

Figure 4.11: The electron mobility of the hall bar device.

The experimental results are shown in Fig. 4.11. The peak mobility is about 5200 cm^2/Vs at V_g=30 V, which is similar to 5000 cm^2/Vs in Ref.[14] and 5600 cm^2/Vs in Ref.[53]. Because the electron mobility provides information about the Si/SiO_2 interface quality, and the hall bar device went through the same fabrication process as the final MOS-SETs, we believe that the quality of our MOS-SETs are comparable with the similar devices from other gourps.

4.5 Summary

In this chapter, we have discussed the physics of the MOS capacitors at room temperature and at cryogenic temperatures. The device parameters are extracted from the HF C-V measurement. Oxide charges and interface trapped charges affect the MOS device operation. The Terman method reveals that the e-beam evaporation

process introduces the interface-trap density $\sim 10^{12}$ $cm^{-2}eV^{-1}$, although it can be reduced to less than 10^{11} $cm^{-2}eV^{-1}$ after forming gas annealing. The conductance method is used to estimate the interface-trap densities in the final devices, and they are about 1×10^{11} $cm^{-2}eV^{-1}$. The 2D electron concentration is determined by hall measurements and Shubnikov-de Haas oscillation measurements. The results show that data can be well explained by a parallel capacitor model. The conductivity mobility is also determined by the hall measurements. The peak mobility is about 5200 cm^2/Vs, which is similar to the data from other groups.[14][53] So the quality of our device is comparable with the quality of the devices from these groups, although it can be further improved.

Chapter 5

Silicon SET transport characteristics

5.1 Introduction

Silicon SETs were systematically investigated both at 4.2 K and in a dilution refrigerator (DR) with a base temperature of 10 mK. The unique side gate configuration was used to verify the formation of quantum dots (QDs) and to determine their locations, which are largely neglected in some early works. The results show that besides an electrostatically defined QD, there is an intrinsic QD at each point-contact channel. These intrinsic QDs are due to the potential fluctuations at the Si/SiO_2 interface. The flexibility of the system enables us to investigate a single intrinsic QD. This will be discussed in Chapter 6 in detail.

Because of the relatively low electron mobility in silicon, in order to remove the intrinsic QD in an point-contact channel, the device dimensions need to be scaled down. When the gap between two neighboring side gates of a point-contact channel was reduced from ~ 160 nm to ~ 90 nm, the point contact can turn off the device without Coulomb oscillations in some devices, which means that there is no intrinsic QD inside the point-contact channel. In these devices, the electrostatically defined QD dominates the device behavior.

5.2 Measurement setup

In simple device characterization, an HP 4142B semiconductor parameter analyzer (including 8 DC Source/Monitor units (SMUs)) is used to do the DC measurement. However, while the source-drain, the side gates and the top gate are DC biased by the HP 4142B, the source-drain differential conductance is measured by the standard AC lock-in technique in most experiments, since the AC lock-in technique can reduce the deteriorating effect from noise, and makes the high precision measurement possible.

The measurement circuit is shown schematically in Fig. 5.1. A lock-in amplifier PAR 124A generates a 37 Hz sine wave signal with $V_{rms} = 1$ V. The signal goes through a 1000:1 voltage divider(100 KΩ : 10 Ω), and becomes $V_{rms} = 0.1$ mV. Since the source-drain DC bias in our experiment is about 25 mV, and the minimum voltage range for the HP 4142B is 2 V, the DC voltage from channel 1 goes through a 86:1 voltage divider (4.22 KΩ : 50 Ω). Then the AC and the DC voltages are added together by an AC+DC adder box, and are applied between the source and the drain of an SET. The source-drain current is amplified by a transimpedance amplifier (pre-amp) with a gain of $10 - 100$ MΩ, and measured by the lock-in amplifier PAR 124A. The monitor signal and the reference signal from the PAR 124A are fed into another lock-in amplifier 7265. Then data from the lock-in amplifier 7265 are collected by computer programs through GPIB interface. The measurement system and techniques are discussed in detail in Appendix A.

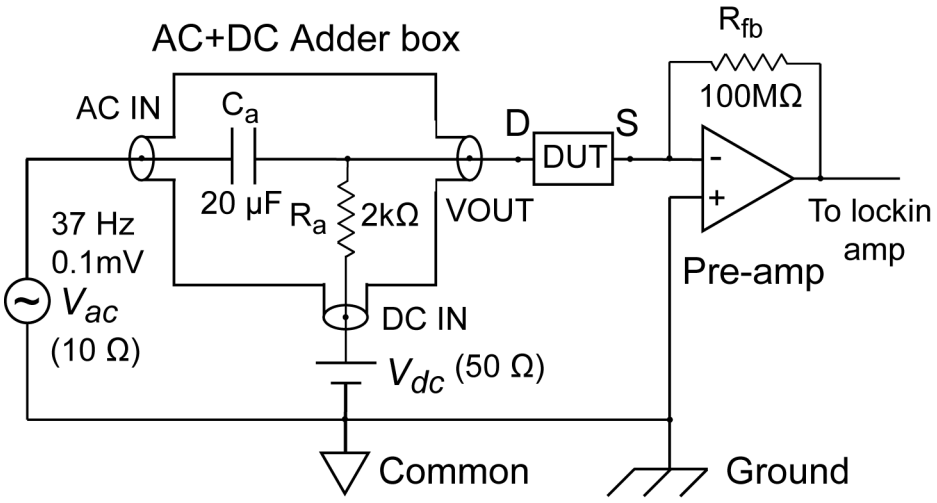

Figure 5.1: Schematic of the AC measurement system. The lock-in amplifier PAR 124A generates a 37 Hz sine wave signal with $V_{rms} = 1$ V, which goes through a 1000:1 voltage divider(100 KΩ:10 Ω), and becomes $V_{ac} = 0.1$ mV. The DC voltage from channel 1 of the HP 4142B goes through an 86:1 voltage divider (4.22 KΩ:50 Ω) and becomes V_{dc}. Then the AC and the DC voltages are added together by an AC+DC adder box, and are applied between the source and the drain of the device under test (DUT). The source-drain current is amplified by a transimpedance amplifier (pre-amp), and measured by the lock-in amplifier PAR 124A.

5.3 Transport characteristics at 4.2 K

5.3.1 Functions of side gates

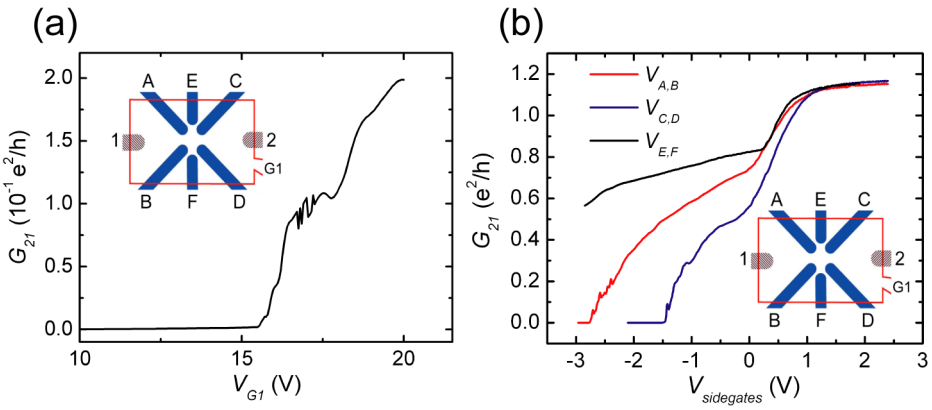

Figure 5.2: (a) The source-drain conductance G_{21} vs. the top gate voltage V_{G1} with 6 side gates (A-F) kept at 0 V. The conductance is in the unit of $e^2/h = 1/(25.8\ \mathrm{k}\Omega)$. The device is SiHB6L20070102(2,1). (b) The source-drain conductance G_{21} vs. side gate voltages. The top gate voltage is 22.5 V. When one pair of side gates are tested, the other two pairs of side gates are set to about 2 V. The inset shows the schematic top view of the MOS-SET with 6 side gates (A-F), the top gate G1, and the source 1, drain 2 regions.

The device is SiHB6L20070102(2,1), and it turns on at the top gate voltage $V_{G1} = 15.5$ V (Fig. 5.2(a)). When V_{G1} is set at 22.5 V, the functions of the side gates are examined. Each pair of side gates A,B, side gates C,D, and side gates E,F independently define one point-contact channel. Before one pair of side gates are tested, about 2 V DC bias is applied on the other two pairs, so that the point-contact channels defined by these two pairs can be fully conductive. Fig. 5.2(b) shows that side gates A,B and side gates C,D can turn off the device, but side gates E,F can

not do it up to -3 V. This is because the gap between side gates E,F ($\sim$ 330 nm) is much larger than that between side gates A,B ($\sim$ 160 nm), and the same is true for side gate C,D.

5.3.2 Stability charts of silicon SETs

First I adjusted side gate voltages $V_{A,B}(= V_A = V_B)$ and $V_{C,D}(= V_C = V_D)$, and then swept side gate voltages $V_{E,F}(= V_E = V_F)$ as plunger gates. Coulomb blockade oscillations are observed, when the source-drain conductance G_{21} versus $V_{E,F}$ is measured at $V_{A,B} = 0.23$ V, $V_{C,D} = -0.5$ V and $V_{ds} = 0.1$ mV (Fig. 5.3(a)), but the result is rather complicated. Fig. 5.3(b) shows the stability chart at the same gate bias conditions. There are some diamond structures, but they are distorted.

Later I found that even using only one pair of side gates, we can still get diamond-shaped stability charts. Fig. 5.4(a) shows a clean Coulomb blockade oscillations, when G_{21} versus $V_{A,B}$ is measured at $V_{C,D} = 2$ V, $V_{E,F} = 2$ V and $V_{G1} = 22.5$ V. (When a pair of side gates are applied by 2 V DC bias, the defined point-contact channel is fully conductive, as shown in (Fig. 5.2(b)).) The corresponding stability chart also shows a simple diamond structure (Fig. 5.4(b)). Further investigation revealed that the complex structure in Fig. 5.3(b) is because each pair of side gates A,B and side gates C,D independently confine one intrinsic QD in their point-contact channels.

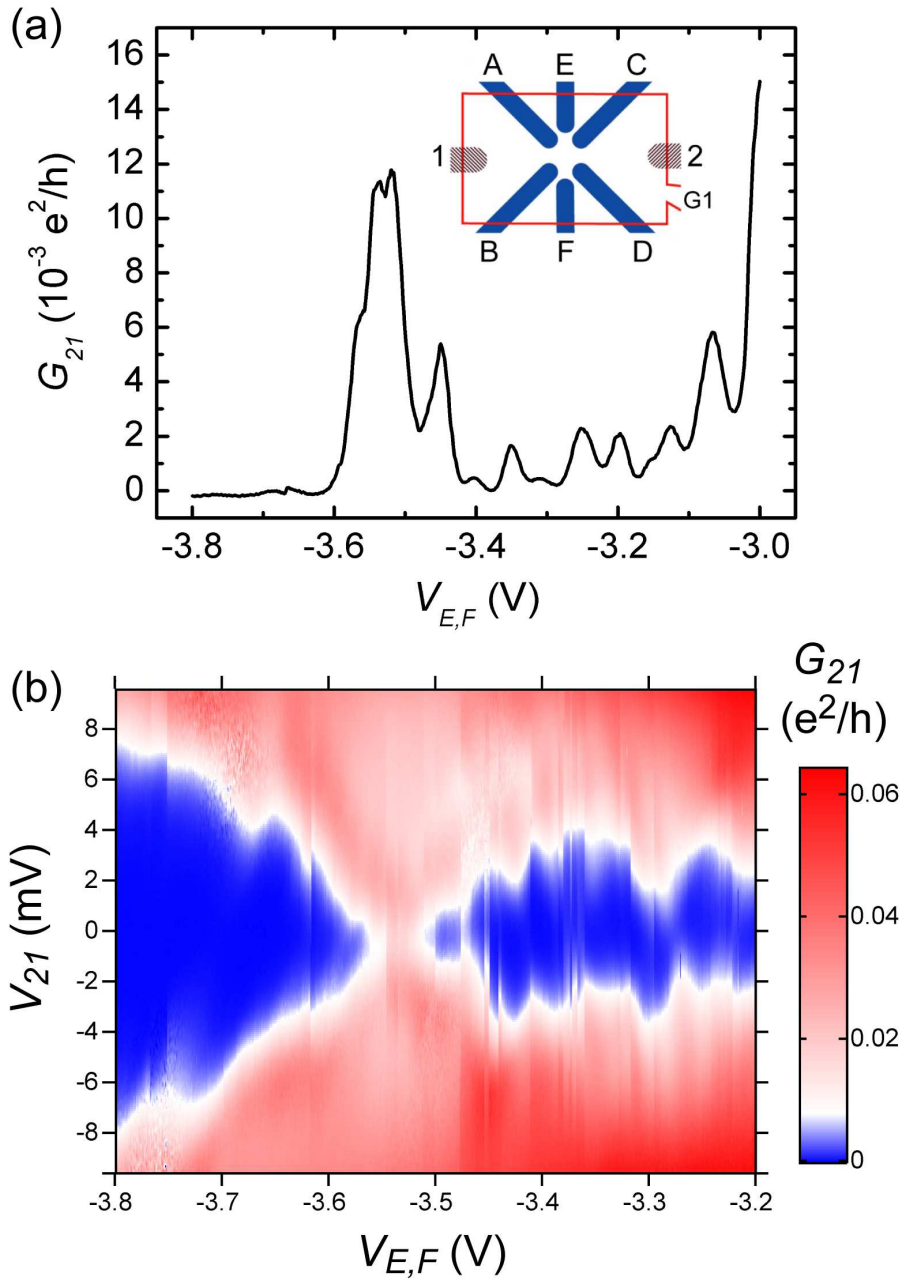

Figure 5.3: Coulomb blockade oscillations (a) and the stability chart (b) of the device, when sweeping the side gate voltages $V_{E,F}(= V_E = V_F)$ as plunger gates. The bias conditions are $V_{G1} = 22.5$ V, $V_{A,B} = 0.23$ V, and $V_{C,D} = -0.5$ V.

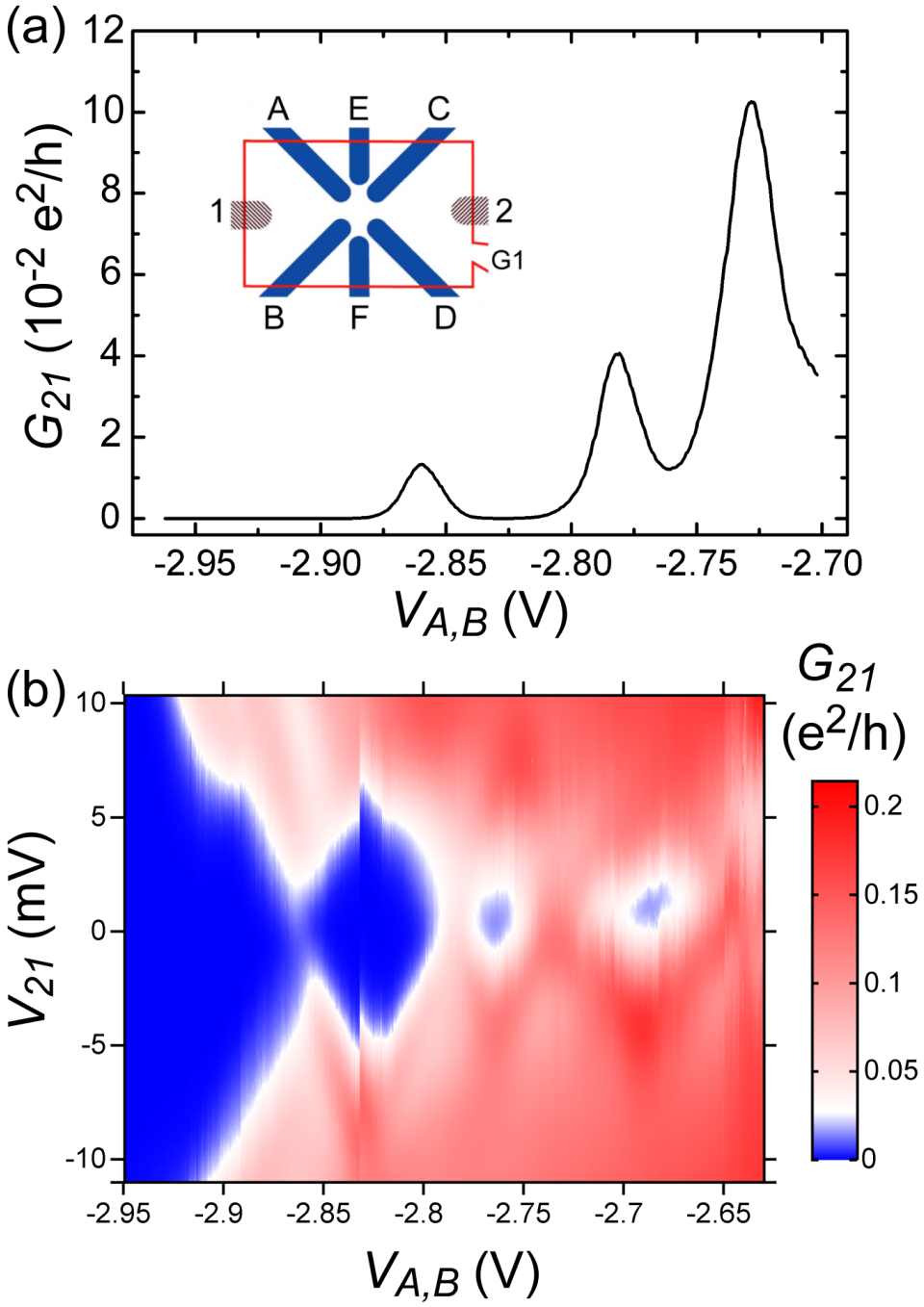

Figure 5.4: Coulomb blockade oscillations (a) and the stability chart (b) of the device, when sweeping side gate voltages $V_{A,B}(= V_A = V_B)$. The bias conditions are $V_{G1} = 22.5$ V, $V_{C,D} = 2$ V, and $V_{E,F} = 2$ V.

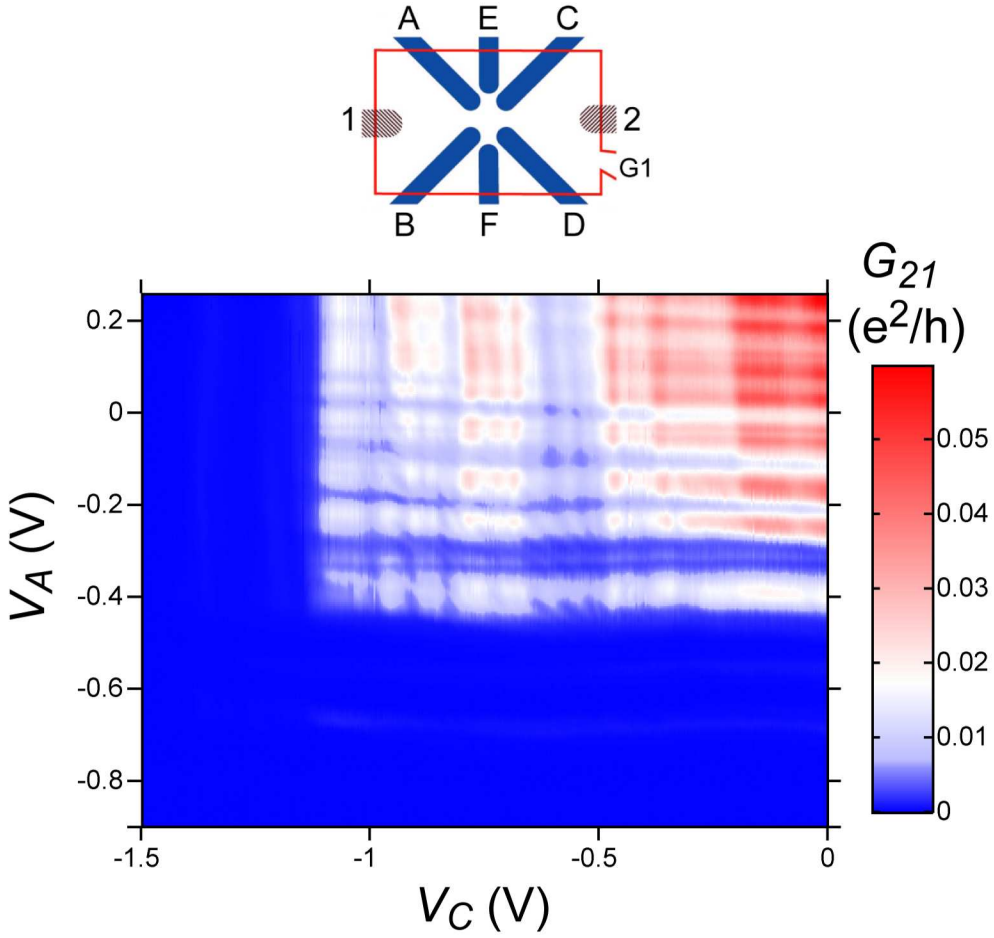

Figure 5.5: The source-drain conductance G_{21} as a function of side gate voltages V_A and V_C shows double QD like behavior, when $V_{G1} = 11$ V, $V_{ds} = 0$ V, $V_B = V_D = V_F = 0$ V, and $V_E = 1$ V. The device is SiHB6L20070102(3,2), and it turns on at $V_{G1} = 9.5$ V.

5.3.3 Double quantum dot like behavior

Since the capacitance between the QD and the side gate implies the distance between them, capacitive coupling strength can be used to determine the location of the QD. When $V_{ds} = 0$ V, G_{21} is measured as a function of side gate voltages V_A and V_C. We get double QD like behavior. As discussed in Sec. 2.2.1, Fig. 5.5 shows that there are two weakly coupled QDs in the SET, since the high conductance regions are only located at the intersections of nearly horizontal and vertical lines, which are the degenerate triple points at $C_m \sim 0$. The nearly horizontal and vertical lines mean that these two QDs are spatially separated; one is near side gate A, and the other is near side gate C. A silicon point-contact showing single electron transistor behavior is not unusual, as previously demonstrated by Hiroki Ishikuro.[54] The possible sources can be some impurities, local potential fluctuations or Si/SiO_2 interface roughness, etc, which will be further discussed later.

5.4 Transport characteristics in a dilution refrigerator

These SETs were further characterized in a dilution refrigerator with a base temperature of 10 mK. The basic characteristics are similar to those at 4.2 K, but more details are revealed with much better resolution. The following data are from another device SiHB6L20070102(2,4).

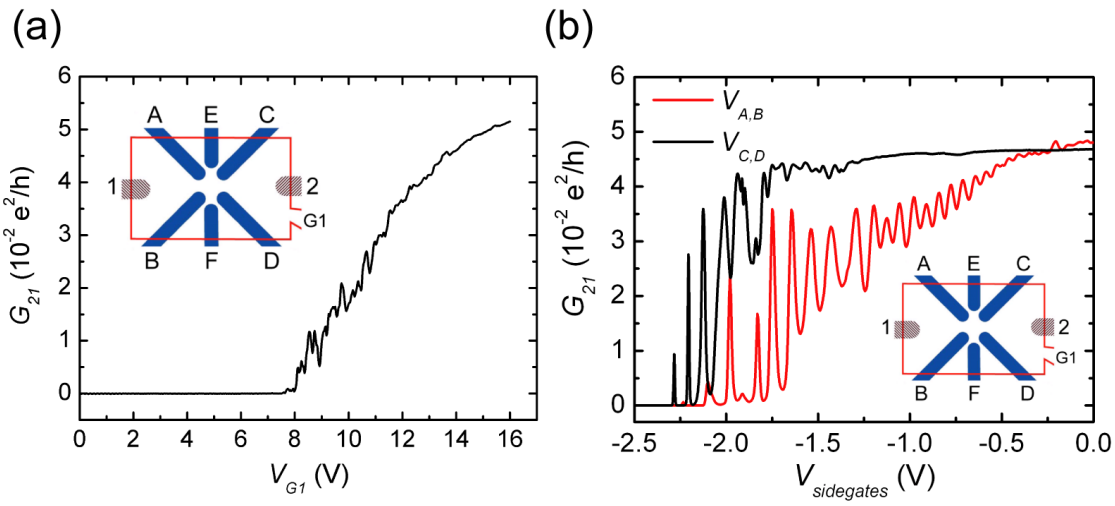

Figure 5.6: (a) The source-drain conductance G_{21} vs. the top gate voltage V_{G1} with 6 side gates (A-F) kept at 0 V. The device is SiHB6L20070102(2,4). (b) The source-drain conductance G_{21} vs. side gate voltages. The top gate voltage is 16 V, and $V_E = V_F = 1$ V. When one pair of side gates is tested, the other pair of side gates is kept at 0 V. The inset shows the schematic top view of an MOS-SET with 6 side gates (A-F), the top gate G1, and the source 1, drain 2 regions.

5.4.1 Functions of side gates

The device turns on at about $V_{G1} = 8$ V (Fig. 5.6(a)). When $V_{G1} = 16$ V and $V_E = V_F = 1$ V, the functions of side gates A,B and C,D are tested. Each pair of side gates A,B and side gates C,D independently define one point-contact channel. When one pair of side gates are tested, the other pair are kept at 0 V, so that the point-contact channel defined by them can be fully conductive. Fig. 5.6(b) shows that side gates A,B and side gates C,D can turn off the device, when more than -2.5 V DC bias is applied on them. The source-drain conductance G_{21} also shows Coulomb blockade oscillations before turning off.

5.4.2 Multiple quantum dot behavior

In a dilution refrigerator, we investigated the formation of QDs in detail, and determined their locations using the specific 6 side gate configuration, because capacitive coupling strength between a QD and a side gate indicates the distance between them.

G_{21} as a function of $V_{A,B}$ and $V_{C,D}$ is measured, when $V_{G1} = 16$ V, $V_E = V_F = 1$ V, and $V_{ds} = 0$ V. The result is shown in Fig. 5.7(a), a portion of which is enlarged and shown in Fig. 5.7(b). The observed characteristics demonstrate double QD behavior, as at 4.2 K. When $V_{A,B}$ and $V_{C,D}$ are both less than about -1.2 V, the device is in the weakly coupled double QD regime. In this weakly coupled regime, two triple points merge into ones, i.e. the high conductance spots as clearly shown in Fig. 5.7(b), and are located at at the intersections of nearly horizontal and vertical

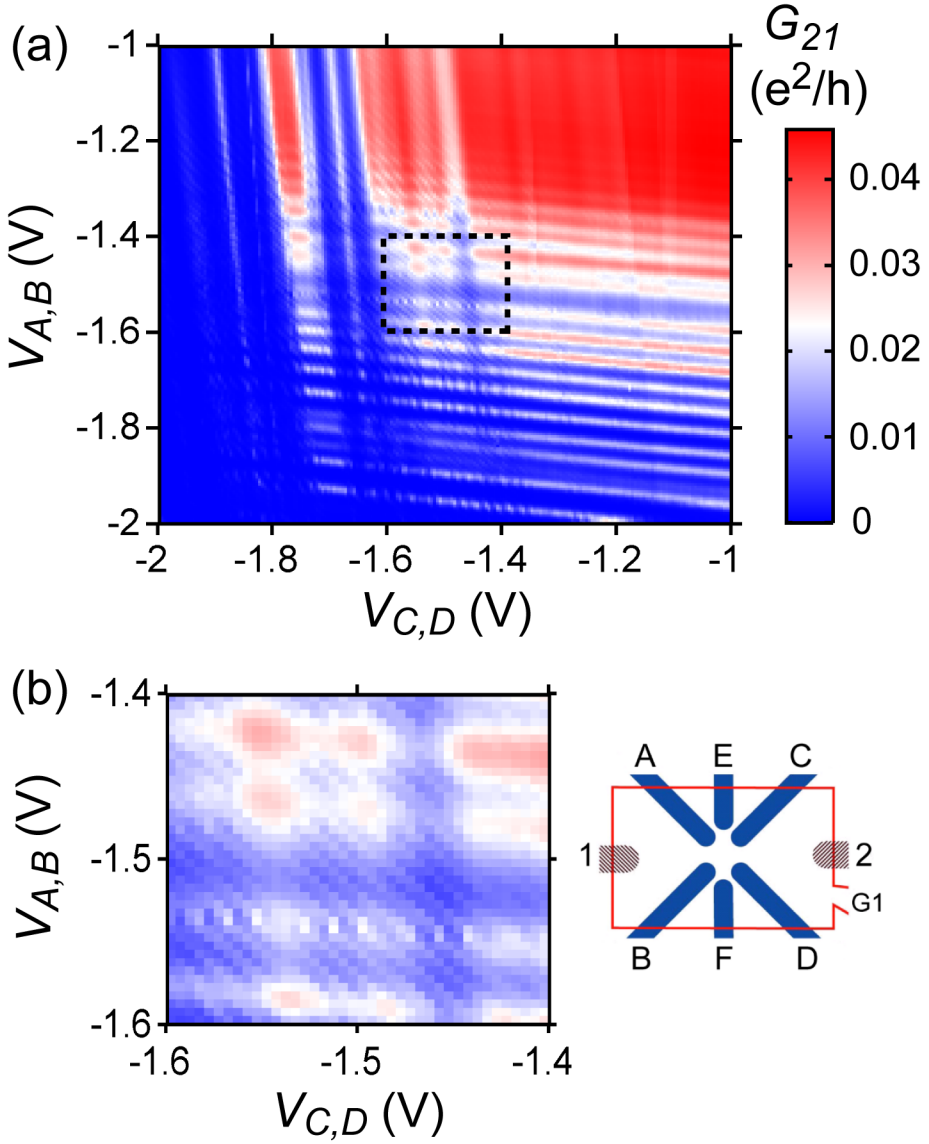

Figure 5.7: (a) The source-drain conductance G_{21} as a function of side gate voltages $V_{A,B}$ and $V_{C,D}$ shows double QD like behavior, when $V_{G1} = 16$ V, $V_E = V_F = 1$ V, and $V_{ds} = 0$ V. (b) Enlarged section of (a) shows the underlining diagonal lines from the center large QD.

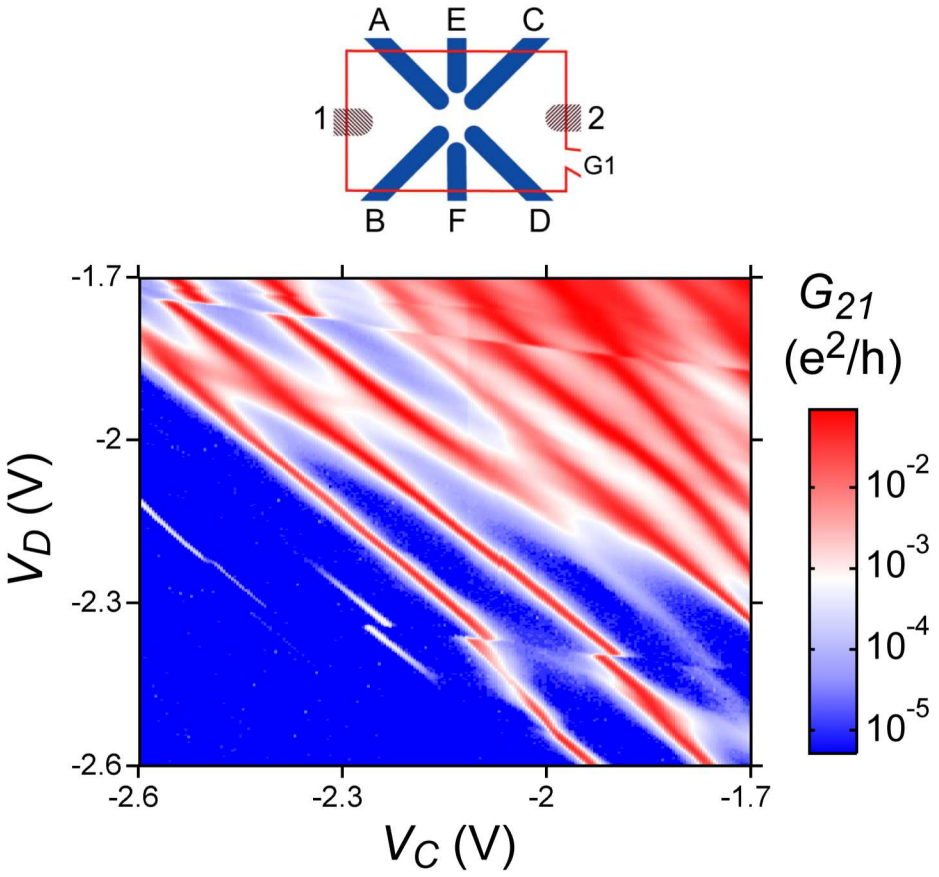

Figure 5.8: The source-drain conductance G_{21} as a function of side gate voltages V_C and V_D, when $V_{G1} = 16$ V, $V_A = V_B = 0$ V, $V_E = V_F = 1$ V, and $V_{ds} = 0$ V. Diagonal lines suggest that side gate C and side gate D are coupled to the small QD with the same capacitance.

lines. These nearly horizontal and vertical lines mean that there are two spatially separated QDs. One dot is strongly capacitively coupled to side gates A and B, but only weakly coupled to gates C and D; the other, however, is strongly coupled to gates C and D, and only weakly coupled to gates A and B. That is, one dot is physically near side gates A and B, whereas the other dot is located close to side gates C and D. In addition to these two small QDs discussed above, side gates A, B, C, D, E, and F, can also electrostatically define a (large) QD. A closer look at Fig. 5.7(a) indeed reveals the underlining diagonal lines from this dot, as shown in Fig. 5.7(b). So when $V_{A,B} < -1.2$ V and $V_{C,D} < -1.2$ V, the SET is under multiple QD bias conditions, and the equivalent circuit consists of a large QD at the center in series with two small QDs, as illustrated in Fig. 5.9(a).

When -1.2 V $< V_{C,D} < 1$ V, sweeping $V_{A,B}$ reproduces the single electron tunneling features similar to that shown in Fig. 5.6(b). The same is observed when sweeping $V_{C,D}$ while keeping -1.2 V $< V_{A,B} < 1$ V. So when -1.2 V $< V_{A,B} < 1$ V (or -1.2 V $< V_{C,D} < 1$ V), the SET is under single QD bias conditions, and only the small QD defined by side gates C,D (or A,B) is left, as shown in Fig. 5.9(b).

Furthermore, data shown in Fig. 5.8 suggest that the QD is located in the narrow point contact channel. Keeping the other two point contacts fully conductive ($V_{G1} = 16$ V, $V_E = V_F = 1$ V, and $V_A = V_B = 0$ V) and sweeping the voltages of side gates C and D in the range of interest, the source-drain conductance shows clear single electron tunneling characteristics. Because the capacitances between the dot and, separately, side gates C and D are the same, the peak positions in Fig. 5.8 display diagonal dependence. To be more specific, the QD is physically

located in the point-contact channel, at equal distances to side gates C and D. The discontinuity of the diagonal lines is caused by a mere one electron charge change at a nearby charge trap. Similar characteristics are also observed for the dot defined by side gates A and B. These observations are consistent with the picture that there are two weakly-coupled QDs, located in the point-contact channels defined by side gates A,B, and C,D, respectively.

These pictures capture the major characteristics of the device. However, there is not always a QD in a point contact. As we will discuss later, when the gap between two neighboring side gates was reduced from $\sim$ 160 nm to $\sim$ 90 nm, the defined point contact can be smoothly turned off without Coulomb oscillations in some of our samples. In these devices, the electrostatically defined dot is the only feature.

5.4.3 Stability chart of an intrinsic quantum dot in silicon

Fig. 5.10(b) shows the stability chart of the SET, while V_{G1} = 16 V, V_A = V_B = 0 V and V_E = V_F = 0 V (single QD bias conditions). The observation of Coulomb diamonds confirms the formation of a single QD in the point-contact channel. Within the orthodox theory, the QD is modeled by a disc with a diameter d, and the total capacitance C_Σ is $4\varepsilon d$, where ε (= 11.9 in silicon) is the dielectric constant. From the diamond shown in Fig. 5.10(b), we can directly measure the half height ($V_{21} = e/C_\Sigma$), and the obtained charging energy E_c (= e^2/C_Σ) is about 6 meV. So C_Σ is approximately 27 aF, which suggests a disc diameter of about 60

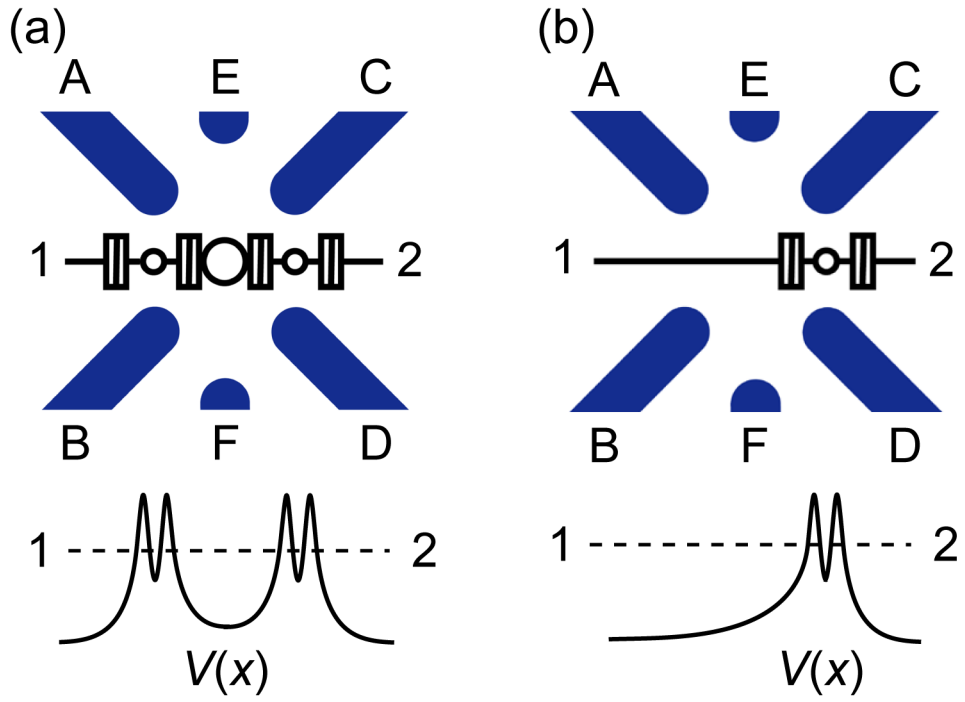

Figure 5.9: (a) Equivalent circuit and schematic potential profile when the SET is under multiple QD bias conditions. (b) Equivalent circuit and schematic potential profile when the SET is under single QD bias conditions.

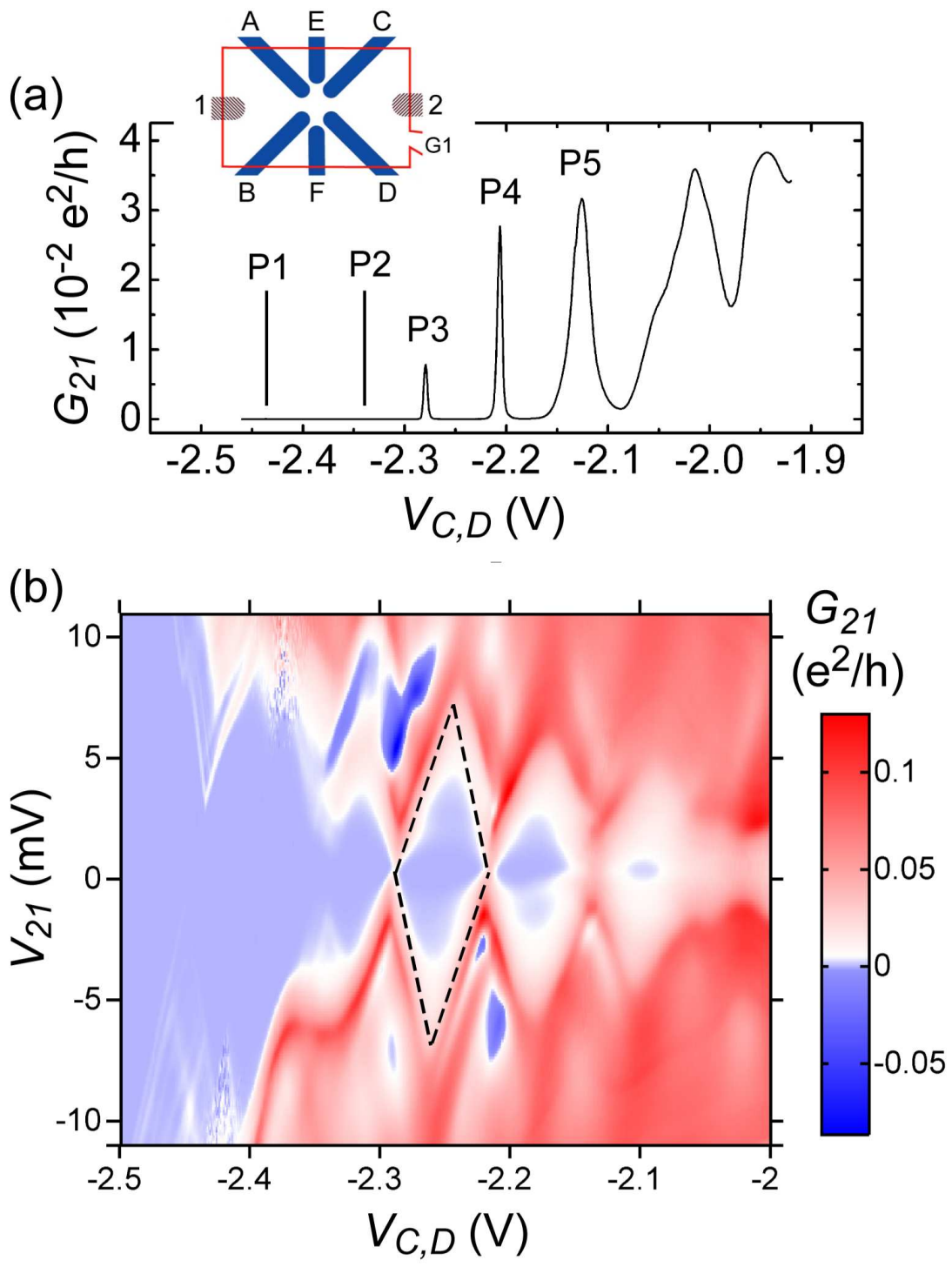

Figure 5.10: (a) The Coulomb blockade oscillations with the first 5 peaks, labeled as P1–P5, when $V_{G1} = 16$ V, $V_A = V_B = 0$ V, $V_E = V_F = 1$ V (single QD bias conditions). (b) Stability chart, taken under the same bias conditions as that in (a).

nm. Because the top gate threshold voltage is measured to be ~ 8 V for the SET, using a parallel capacitor model and that the quantum disc had a diameter of 60 nm, at $V_{G1} = 16$ V the number of electrons in the QD is estimated to be at most ~ 10. As the source-drain dc bias V_{21} is fixed at 0 V, the SET displays the Coulomb blockade oscillations, as shown in Fig. 5.10(a).

5.5 Towards an electrostatically well defined quantum dot in silicon

Electrostatically well defined lateral QDs with great versatility and controllability are a must in order to realize spin qubit devices in Si. As we have seen, however, intrinsic QDs are easily formed at point-contact channels in Si, due to the inherent material properties, especially the potential fluctuations at the Si/SiO_2 interface. There are two ways to remove the intrinsic QD in a point-contact channel. One is to improve the quality of the Si/SiO_2 interface, so that the interface potential fluctuations can be reduced, and the electron localization length can be increased. The other is to scale down the device, so that within the device dimensions, the potential fluctuations are not able to confine an intrinsic QD in a point-contact channel. Here we demonstrate that the intrinsic QD can be removed from the point-contact channel, when the device is scaled down.

Due to the limitation of our e-beam lithography recipe, the minimum line width is about 50 nm, and the standard line width is about 70 nm. After the device was scaled down, the gap between two neighboring side gates reduced from ~ 160 nm to ~ 90 nm, as shown in Fig. 5.11.

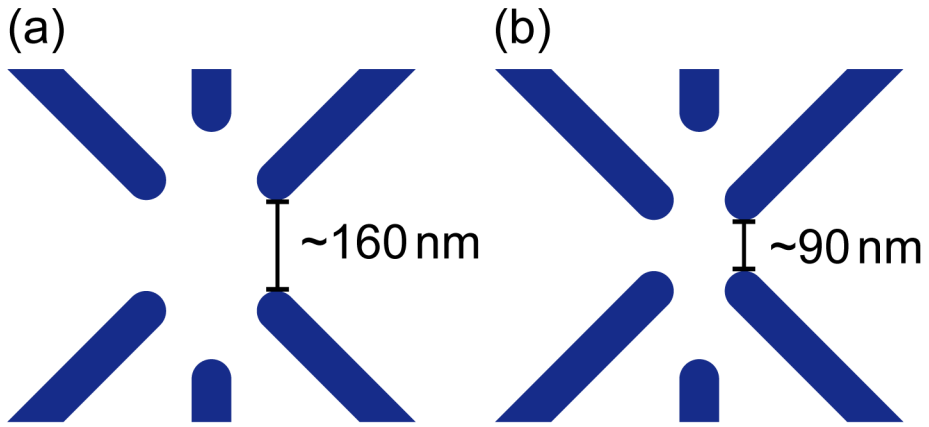

Figure 5.11: Schematic of the device scale-down. After the device scale-down, the gap between two neighboring side gates reduces from $\sim$ 160 nm in (a) to $\sim$ 90 nm in (b).

5.5.1 Removal of intrinsic quantum dots

The basic idea to remove the intrinsic QD in a point-contact channel is to reduce the effective channel length. Fig. 5.12 shows the cross sections of the potential profiles along the point-contact channel. At a lower top gate voltage and a lower side gate voltage, the potential barrier has a wider peak (Fig. 5.12(a)); at a higher top gate voltage and a higher side gate voltage, the potential barrier becomes sharper (Fig. 5.12(b)). Thus the effective channel length is shortened. When a sharper potential barrier is combined with the intrinsic potential fluctuations, the intrinsic potential fluctuations can be effectively reduced. This lowers the potential barriers of the intrinsic QD, as shown in Fig. 5.12(c)–(f). As we continue to push the potential barrier sharper, the intrinsic potential fluctuations are no longer able to confine electrons in it. That is, we effectively remove the intrinsic QD in the point-contact channel.

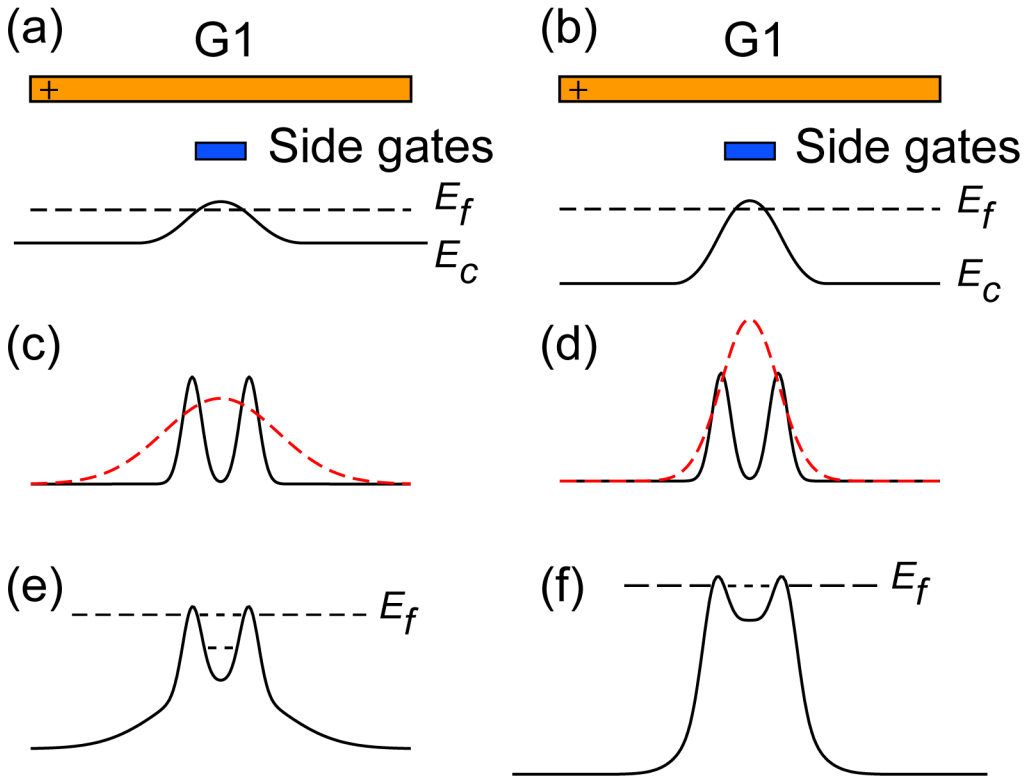

Figure 5.12: Schematic potential profiles at different gate bias conditions. (a) At a lower top gate voltage and a lower side gate voltage, the potential barrier has a wider peak; (b) at a higher top gate voltage and a higher side gate voltage, the potential barrier is sharper. The dark lines in (c) and (d) illustrate the interface potential fluctuations, and the dash lines are the electrostatically defined barriers. (e) and (f) show the resultant potential profiles after combining these two together. For the sharper potential barrier, the intrinsic potential fluctuations are effectively reduced.

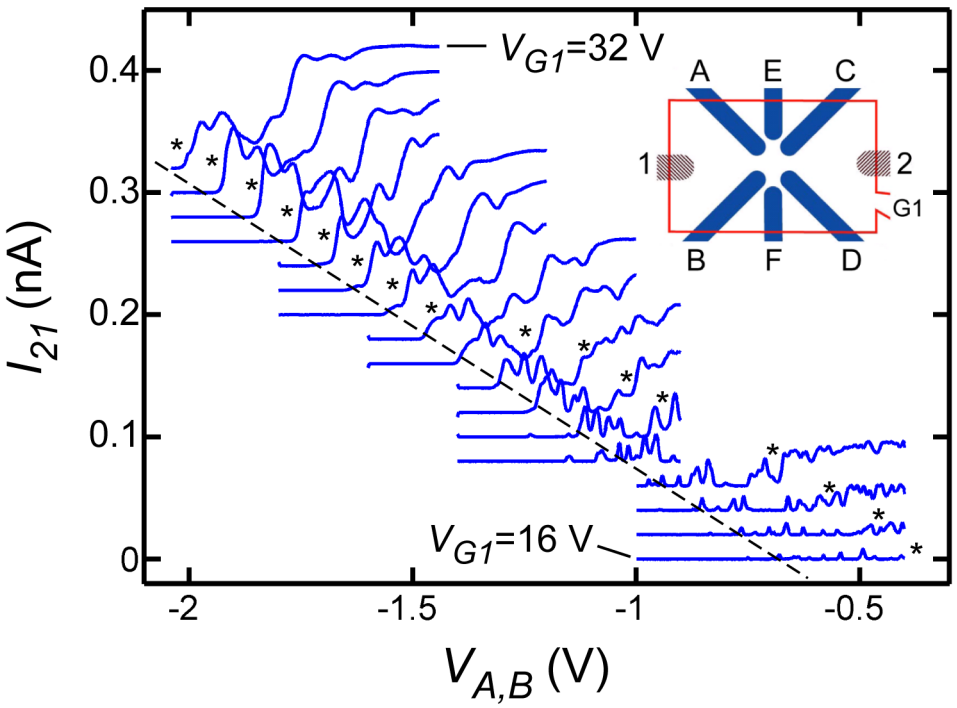

Figure 5.13: The source-drain current I_{21} as a function of side gate voltages $V_{A,B}$ when the top gate voltage V_{G1} increases from 16 V to 32 V. Each Star labels the position (in $V_{A,B}$) where the intrinsic QD transits from an open dot into a closed dot. The dash line shows the linear dependence between V_{G1} and $V_{A,B}$. Each curve is offset linearly with V_{G1} for clarity.

One experimental result is shown in Fig. 5.13, where the source-drain current I_{21} is measured against $V_{A,B}$ when the top gate voltage V_{G1} is changed from 16 V to 32 V. At lower top gate voltages, there are some Coulomb blockade oscillations, which means that the device is in the closed dot regime. When the top gate voltage increases, the Coulomb blockade is lifted, which means that the device enters the open dot regime. At $V_{G1} = 32$ V, there is no Coulomb blockade oscillations any more, which means that the intrinsic QD is removed from the point-contact channel.

5.5.2 Functions of side gates

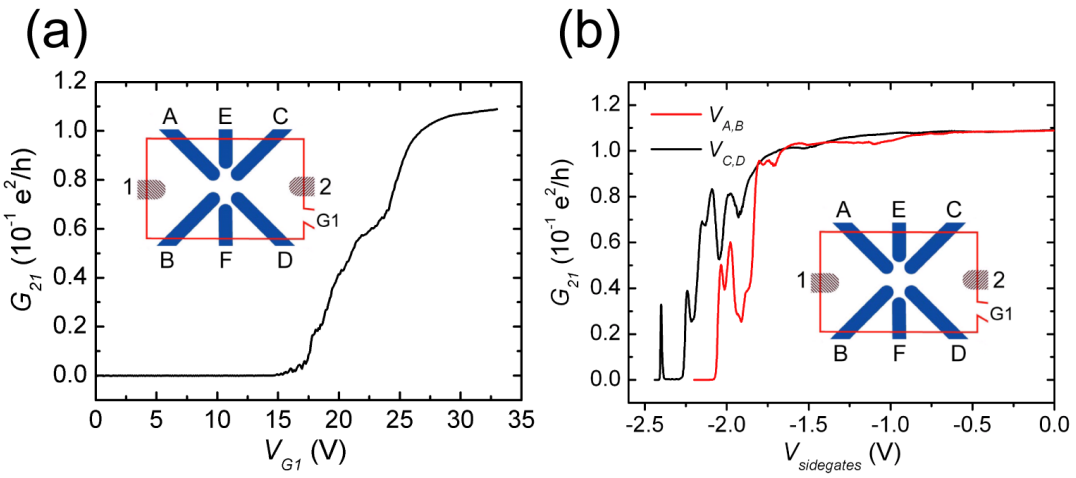

Figure 5.14: Electrical characteristics of a scaled-down device, measured in a dilution refrigerator with 10 mK base temperature. (a) The device turns on at about $V_{G1} = 15$ V, when 6 side gates (A-F) are kept at 0 V. The device is SiHB6L20080820(2,4). (b) The source-drain conductance G_{21} vs. side gate voltages, when $V_{G1} = 33$ V, and $V_E = V_F = 0$ V. When one pair of side gates are tested, the other pair of side gates are kept at 0 V.

The scaled-down devices were also characterized in the dilution refrigerator.

Fig. 5.14(a) shows that the device turns on at about $V_{G1} = 15$ V, when all of side gates A-F are kept at 0V. At $V_{G1} = 33$ V, $V_{E,F} = 0$ V and $V_{C,D} = 0$ V, side gates A,B can smoothly turn off the device without any Coulomb blockade oscillations, so there is no QD in this point-contact channel; However, there is one Coulomb blockade oscillation, when side gates C,D turn off the device, as shown in Fig. 5.14(b).

5.5.3 The electrostatically defined quantum dot

Since the electrostatically defined QD is located at the center of the 6 side gates, the capacitances between the QD and, separately, side gates A,B and side gates C,D should be the same. Indeed, Fig. 5.15 shows that the Coulomb oscillation peaks are located at the diagonal lines of $V_{A,B}$ and $V_{C,D}$, which means that side gates A,B and side gates C,D are equally capacitively coupled to the QD. After comparing Fig. 5.15 with Fig. 5.7, we conclude that the electrostatically defined QD dominates the device behavior.

Fig. 5.16(b) shows the stability chart of the electrostatically defined QD, where diamond-shaped Coulomb blockade regions are clearly visible. When $V_{C,D} <$ 2.05 V, Coulomb blockade oscillations show closely adjacent peaks (double peaks) (Fig. 5.16(a)), which means that there were two tunnel-coupled QDs at low electron concentration.

We can use the orthodox theory to estimate the size of the dominant QD by modeling it as a disc. For a disc with a diameter d, the total capacitance C_Σ is $4\varepsilon d$,

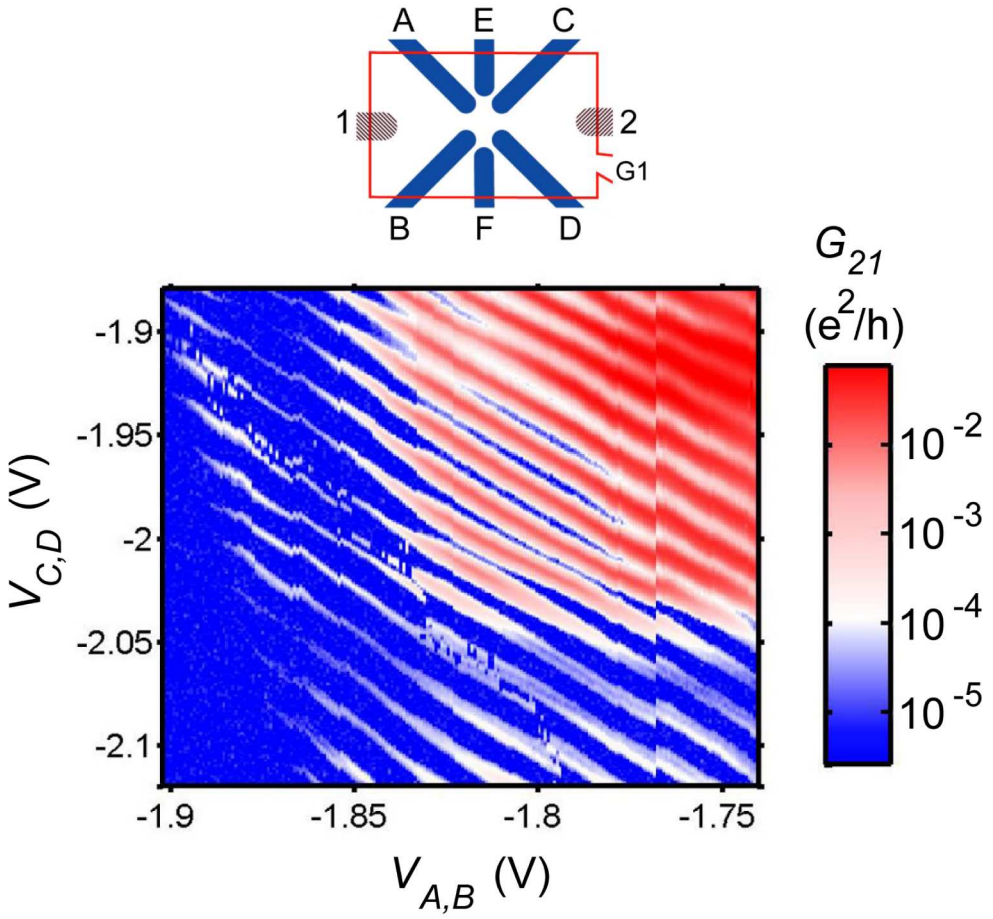

Figure 5.15: (a) The source-drain conductance G_{21} as a function of the side gate voltages $V_{A,B}$ and $V_{C,D}$, when $V_{G1} = 33$ V, $V_E = V_F = 0$ V, and $V_{ds} = 0$ V. Diagonal lines suggest that side gates A,B and side gates C,D are coupled to the QD with the same capacitance.

where ε (= 11.9 in silicon) is the dielectric constant. From the diamond shown in Fig. 5.16(b), we can get the half height ($V_{21} = e/C_\Sigma$), and the obtained charging energy E_c ($= e^2/C_\Sigma$) is about 2 meV. So C_Σ is approximately 80 aF. This suggests a disc diameter of about 190 nm, which is consistent with the gap between two diagonal side gates ($\sim$ 160 nm) in Fig. 5.11(b). Because the top gate threshold voltage is measured to be $\sim$ 15 V for the SET, using a parallel capacitor model

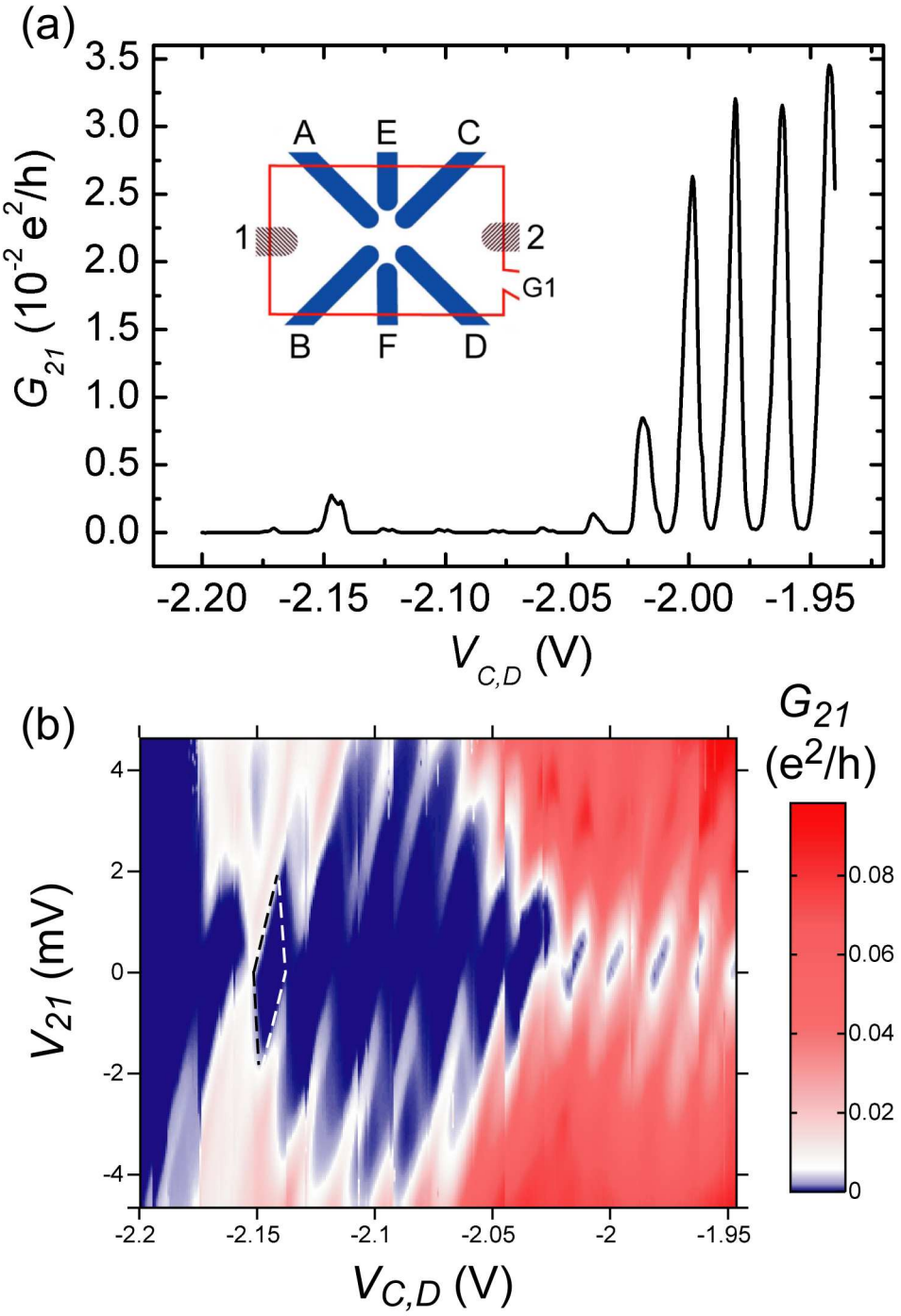

Figure 5.16: The Coulomb blockade oscillations of the electrostatically defined QD, when $V_{G1} = 33$ V, $V_A = V_B = -1.85$ V, and $V_E = V_F = 0.4$ V. (b) Stability chart under the same bias conditions as that in (a).

and that the quantum disc has a diameter of 190 nm, at $V_{G1} = 33$ V the number of electrons in the QD is estimated to be about ~ 250. So the device is in the many-electron regime. When many electrons are in the QD, they can smooth out the potential fluctuations, and one electrostatically defined QD can be formed.

5.6 Origin of intrinsic quantum dots

It is important to find out the origin of intrinsic QDs in our devices, so that we can better understand them and improve the device quality in the future. There are at least three different theories in the literature. H. Ishikuro and T. Hiramoto think that QDs are probably from Si nanoparticles in their point-contact SOI devices.[54] This is not likely in our devices, since in our devices, the QDs are located at the silicon/thermal oxide interface, and there are no nanoparticles. H. Sellier et al., demonstrated that QDs can be from single dopants such as arsenic or phosphorous atoms in their silicon SOI SETs.[55] The major characteristic of a single-dopant QD is the large charging energy, at the order of the first binding energy, which is about 25 meV in silicon. Our observed charging energy (~ 6 meV in Fig. 5.10(b)) is much less than that using the single dopant picture. In addition, the single-dopant QD can only trap a few electrons, 2 for phosphorous. There are at least 5 peaks in Fig. 5.10(b). In another sample, there are 10 peaks with similar addition energies between $4 \sim 8$ meV, as shown in Fig. 5.17(b). It is impossible for a single-dopant QD to trap so many electrons. So the single-dopant quantum-dot picture alone cannot explain the intrinsic QDs in our devices. They also have different filling pattern.[55]

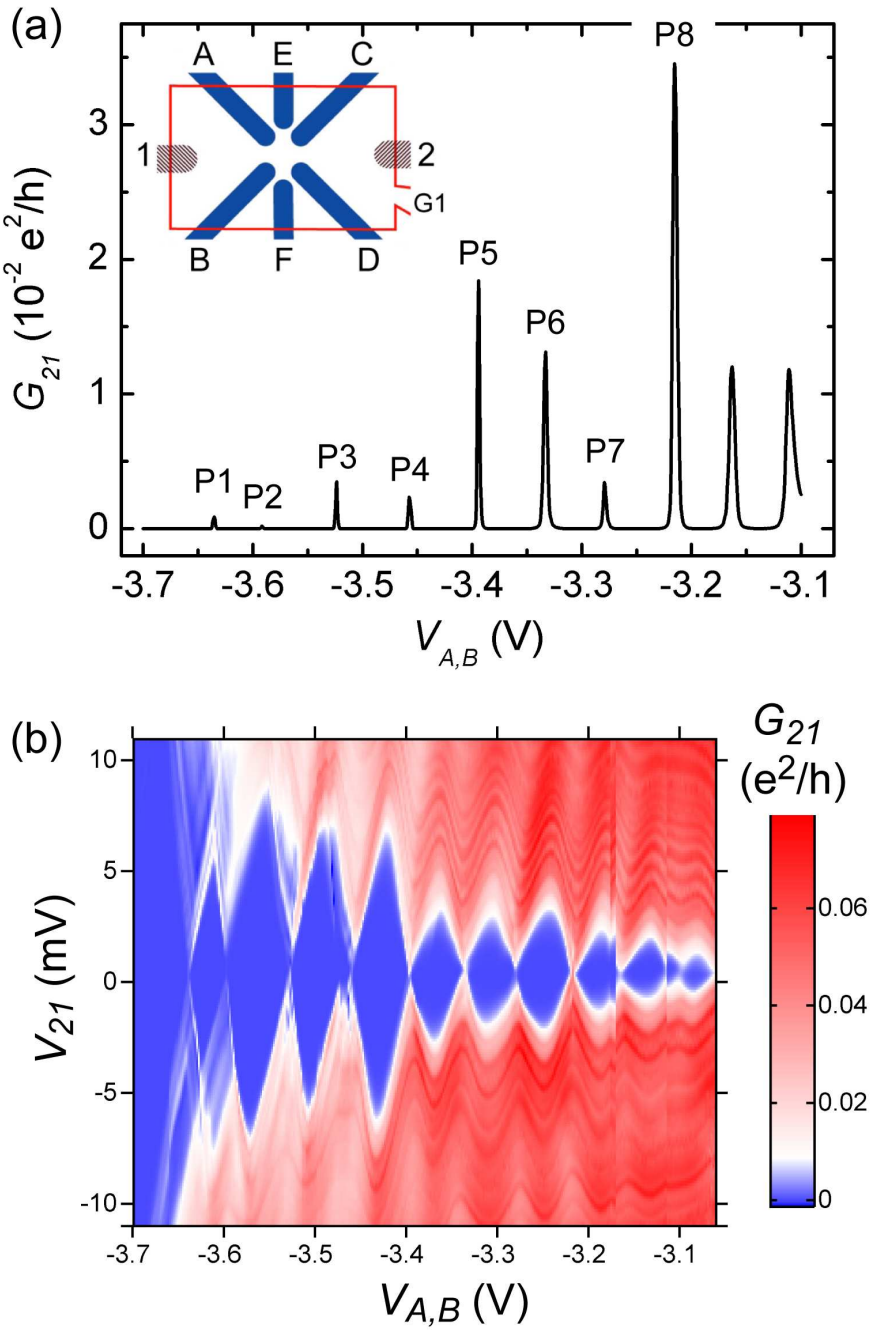

Figure 5.17: (a) The Coulomb blockade oscillations with the first 8 peaks labeled as P1–P8, when $V_{G1} = 25$ V, $V_C = V_D = 2$ V, $V_E = V_F = 0.8$ V. The device is SiHB6L20070102(2,3). (b) Stability chart, taken under the same bias conditions as that in (a).

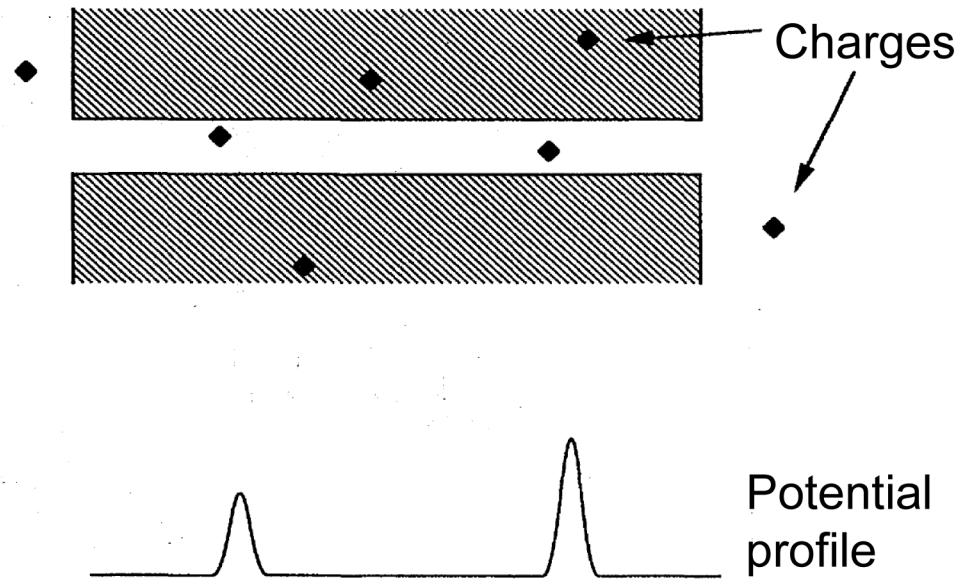

Figure 5.18: Schematics of a narrow channel with random trapped charges. The channel is defined by a metal gate (the shaded area) with a gap. The black diamonds are randomly distributed charges near the interface. The corresponding electrostatic potential along the channel is also shown in the figure. (After M. A. Kastner, Ref.[17])

We believe that the intrinsic QDs are due to the inherent potential fluctuations at the Si/SiO_2 interface, caused by the localized nature of interface trapped charge, oxide fixed charge, and chemical inhomogeneity (stretched, bent, or broken bonds, and oxide compositional variations).[56] M.A. Kastner proposed that QDs are from the negatively charged electron traps.[17] As shown in Fig. 5.18, the negative interface-trapped charges create potential barriers along the one-dimensional (1D) point-contact channel and form a QD. (Since the oxide fixed charges are predominantly positive, they cannot create potential barriers.) The interface trapped charges are always present at the Si/SiO_2 interface. As we have measured by the conductance method using large area devices at room temperature in Sec. 4.3.3, The interface trap density is about $\sim 1 \times 10^{11}$ $cm^{-2}eV^{-1}$, corresponding to an average spacing between charges to be ~ 30 nm. This is close to the diameter of the intrinsic QD (~ 60 nm) in Sec. 5.4.3, considering the overestimation nature of the simple disc model. At this time, the exact origin of intrinsic QDs is unknown. But the interface trap density could be a good indicator of the localization length, thus it can be related to the size of the intrinsic QDs. Nevertheless, it is still a guess. It needs a systematic experiment to verify the relationship between intrinsic QDs and the interface trap densities.

5.7 Coupled double quantum dots in silicon

A coupled double QD system with a tunable inter-dot potential barrier is necessary to realize the CNOT operation, one of the ingredients of a universal operation

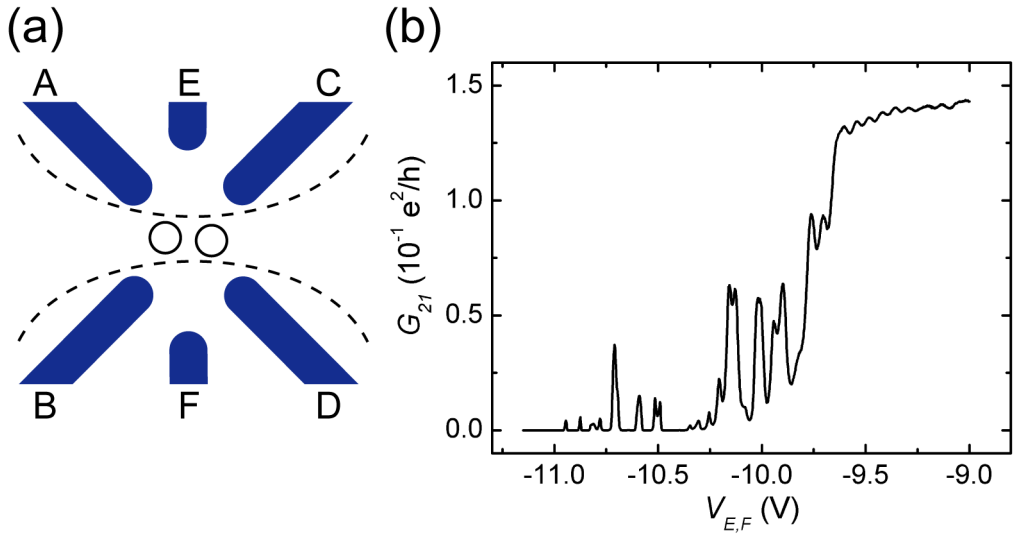

Figure 5.19: (a) Schematic of coupled double dots in silicon. Side gates E,F define a point-contact channel. Side gates A,B and side gates C,D are used to probe the double dots. (b) The source-drain conductance G_{21} vs. side gate voltages $V_{E,F}$, when $V_{G1} = 33$ V, and $V_A = V_B = V_C = V_D = 0$ V. In contrast to side gates A,B and C,D, more then -10 V is applied on side gates E,F to turn off the device. The device is SiHB6L20080820(2,4).

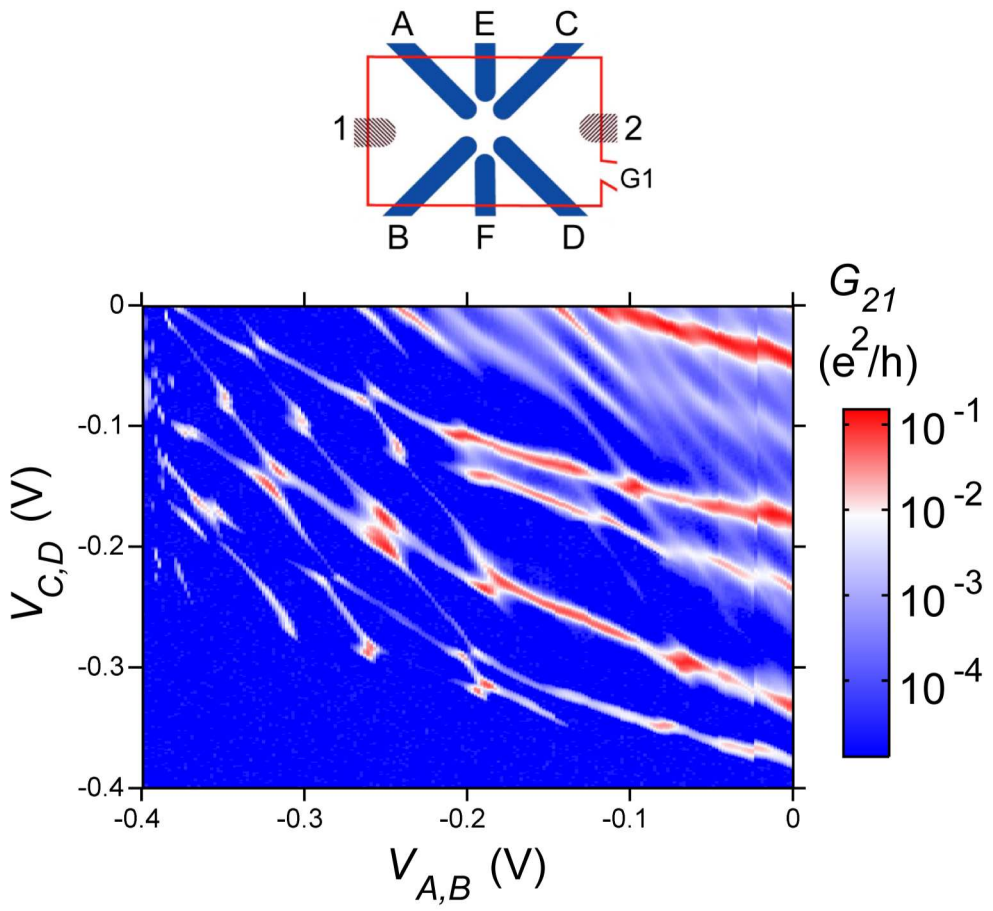

Figure 5.20: The stability chart of the device with $V_{G1} = 33$ V, $V_{E,F} = -9.8$ V, which shows the typical double dot honeycomb structure.

set for quantum computation. Our 6 side gate Silicon SET structure is flexible. As shown in Sec. 5.5.3, when side gates E,F are biased at 0 V, side gates A,B and side gates C,D can electrostatically define one large QD. On the other hand, if side gates E,F are used to define one point-contact channel, and side gates A,B and side gates C,D are used to define the other two barriers, two coupled QDs can be formed (Fig. 5.19(a)).

The measured stability chart shows the typical double dot honeycomb struc-

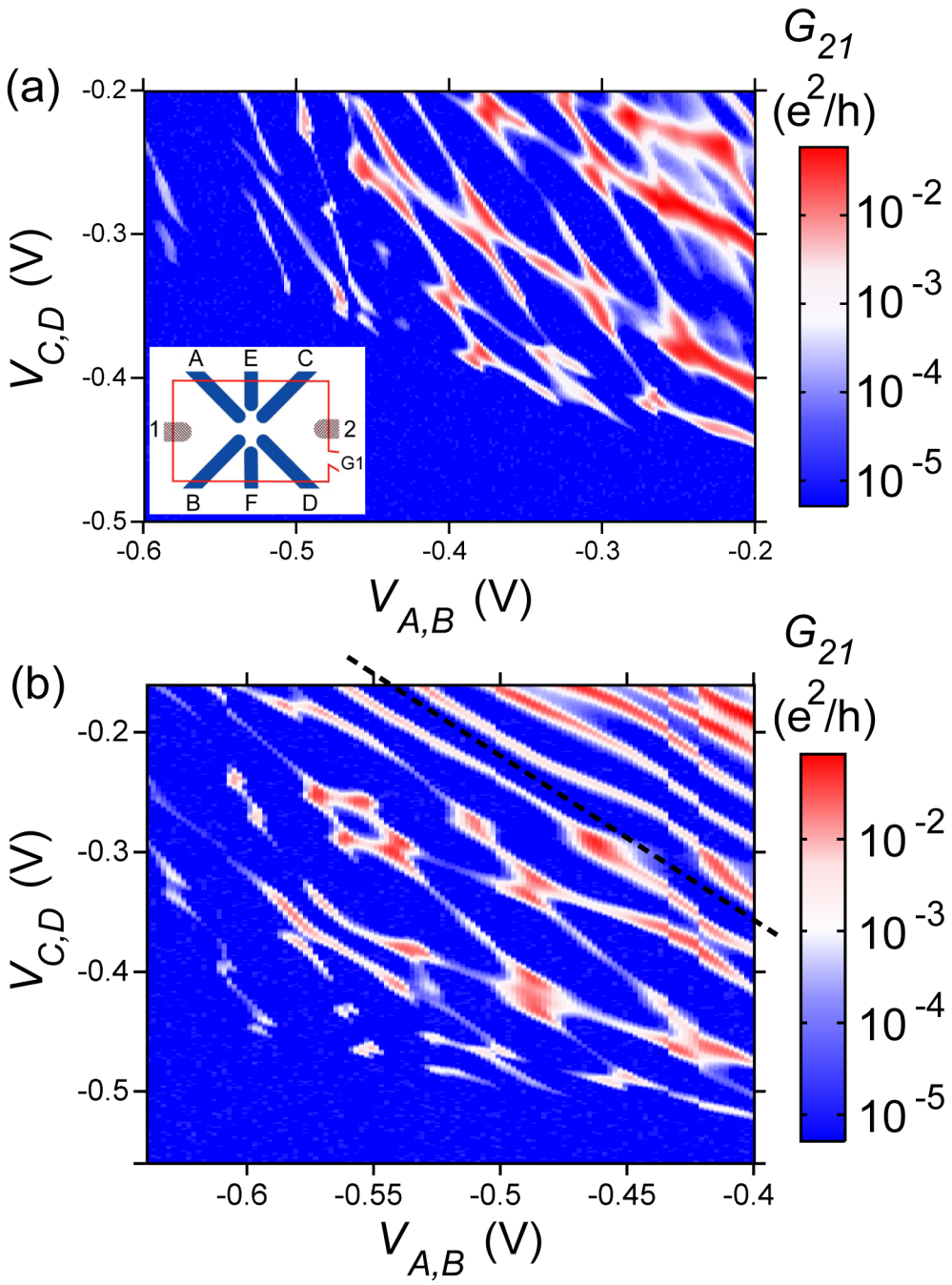

Figure 5.21: The stability charts of the device with $V_{G1} = 33$ V, and different $V_{E,F}$: (a) $V_{E,F} = -9.4$ V, (b) $V_{E,F} = -9.0$ V. In (b), the dash line separates the figure into two parts. The lower left part shows the double dot honeycomb structure; the upper right part shows the evolution into a single dot.

ture (Fig. 5.20). I have tried to adjust the coupling between these two dots by changing $V_{E,F}$. Although there are some small changes when $V_{E,F}$ decreases from -9.8 V to -9.4 V, and to -9.0 V, over all results are similar (Fig. 5.20, and Fig. 5.21). Also the tunneling rate between these two dots are substantial, so that the triple points are not well defined points but regions. This means that these two dots are not well separated, but form a molecule. We also notice that there are two distinct regimes in Fig. 5.21(b), separated by a dash line. The lower left part shows the double dot honeycomb structure, and the upper right part shows the evolution into a single dot. But whether these two dots are truly electrostatically defined, or mainly due to the interface potential fluctuations cannot be distinguished well. The later one is most likely. Considering the size of our current device, it is quite possible that these two dots are mainly from the interface potential fluctuations. Further scaling-down is necessary to create an electrostatically defined double dots.

5.8 Summary

In this chapter, we have presented the transport results of the silicon SETs both at 4.2 K and in a dilution refrigerator. We have discovered that each pair of side gates A,B and C,D can confine one intrinsic QD in its point-contact channel due to the interface potential fluctuations. In the dilution refrigerator, we have found that in addition to these two smaller intrinsic QDs, there is also one larger electrostatically defined QD. Thus we are able to establish equivalent circuits to describe the observed device characteristics.

But the electrostatically well defined QD is necessary to realize spin qubit devices. After scaling down the device and reducing the effective channel length of the point-contact channel by increasing the top gate voltage and the side gate voltages, we have demonstrated that the intrinsic QD can be removed from the point-contact channel. Thus the electrostatically defined QD dominates the device behavior. However, the device only works in the many-electron regime. At low electron concentration, the electrostatically defined QD separates into multiple dots.

We have discussed the origin of intrinsic QDs. We believe that they are not due to Si nano-particles or single dopants, but from the intrinsic potential fluctuations at the Si/SiO_2 interface. There may be some signals from single dopants, but most of Coulomb blockade oscillations cannot be from the single dopants.

We have also demonstrated that our 6 side gate structure can form a coupled double dot system, although the inter-dot coupling cannot be tuned efficiently. The interface potential fluctuations probably play a major role in forming the inter-dot tunnel barrier.

In order to realize electrostatically well defined QD in silicon, we need to continue scaling down the device and improving the interface quality, so that the device dimensions can be smaller than the localization length. Someone may argue that the physical dimensions are not necessary to be small, because the size of the QD can be electrostatically confined small. But we should keep in mind that an SET includes three inseparable parts: two tunnel barriers and a QD. Each of them can be electrostatically confined small, but the overall size is determined by the physical dimensions of the device. From the size of the intrinsic QD, we can estimate that the

device dimension should be less than 60 nm for our current interface quality, which is reachable. The current industrial standard provides an interface trap density of approximately 1×10^{10} cm^{-2}, so the device dimension can be ~ 100 nm.

Chapter 6

Excited states and magnetic field spectra of a single quantum dot

6.1 Introduction

For the purpose of investigating single spin in silicon, we focus on a single intrinsic quantum dot confined in a point-contact channel in this chapter. Due to their small size and large quantization energy, intrinsic quantum dots provide us an important tool to explore quantum dots in silicon beyond the limitation of our lithography system.

Under single dot bias conditions, the investigated SET is comprised of a single intrinsic quantum dot, as discussed in Chapter 5. In addition to single electron tunneling behavior, we also measured its magnetic field dependence. In a magnetic field perpendicular to the sample surface, the observed magnetic field dependence of the ground-state and excited-state energy levels mostly can be attributed to the Zeeman effect. Furthermore, the spectrum enables us to directly observe the singlet-triplet (ST) transition. In this two-electron ST transition, electron-electron Coulomb interaction plays a significant role. The evolution of Coulomb blockade peaks with magnetic field was also measured. The data suggest that the ground state energy levels also shift with the applied magnetic field owing to the Zeeman effect. Up to 9 T, there is no obvious orbital effect. The evolution of peak amplitudes can be

explained by the spin blockade effect. When the two-electron system forms a singlet state at low magnetic fields, and the injection current from the lead becomes spin-down polarized, the tunneling conductance is reduced by a factor of 8. At higher magnetic fields, due to the ST transition, the spin blockade effect is lifted, and the conductance is fully recovered. In a magnetic field parallel to the source-drain current direction, the shifts of the ground state energy levels can also be attributed to the Zeeman effect.

The stability chart near the first Coulomb blockade peak shows some fine features. When the magnetic field sweeps from 0 T to 9 T, the energy differences between the ground state and some "excited states" are barely changed. We suspect that these "excited states" are false. It is expected that the electron concentration in the source and drain regions is lower near the first Coulomb blockade peak than that near other peaks, so the local potential fluctuations are rougher, and the quantum interference effect is more pronounced. Using this quantum interference picture, we are able to qualitatively explain some observed phenomena.

6.2 Singlet-triplet transition in a silicon SET

Here we continue our investigation on device SiHB6L20070102(2,4). Under single quantum dot bias conditions, the device has a single quantum dot in the channel defined by side gates C,D (Fig. 6.1). Fig. 6.2 shows the stability chart of the SET, when $V_{G1} = 16$ V, $V_A = V_B = 0$ V, $V_E = V_F = 1$ V. In this stability chart, there is an excited state near peak 4 (the dashed line). We first focus on the

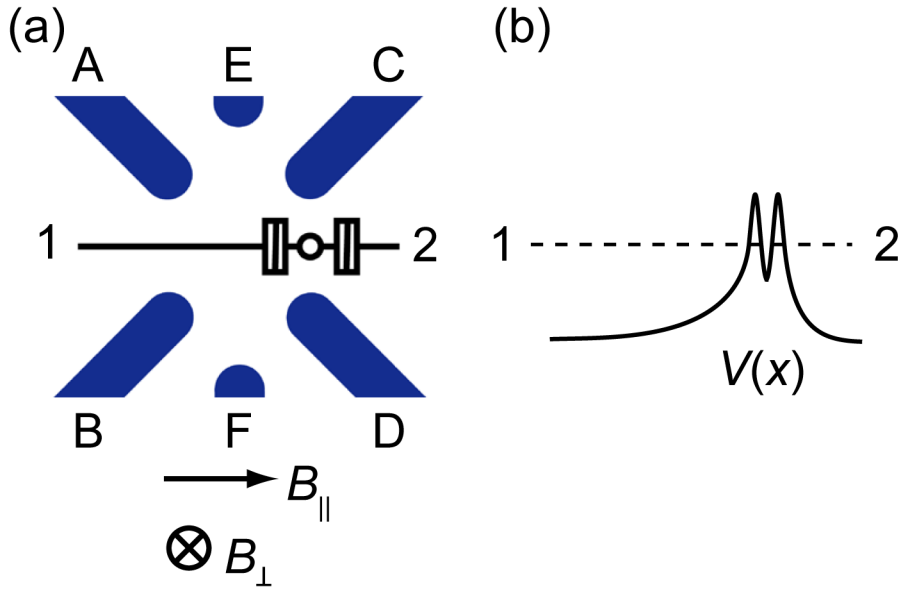

Figure 6.1: (a) The equivalent circuit and (b) schematic potential profile under single quantum dot bias conditions. It also shows two different directions of magnetic field B. One ($B_{||}$) is parallel with the source-drain current direction, and the other($B_{\perp}$) is perpendicular to the sample surface.

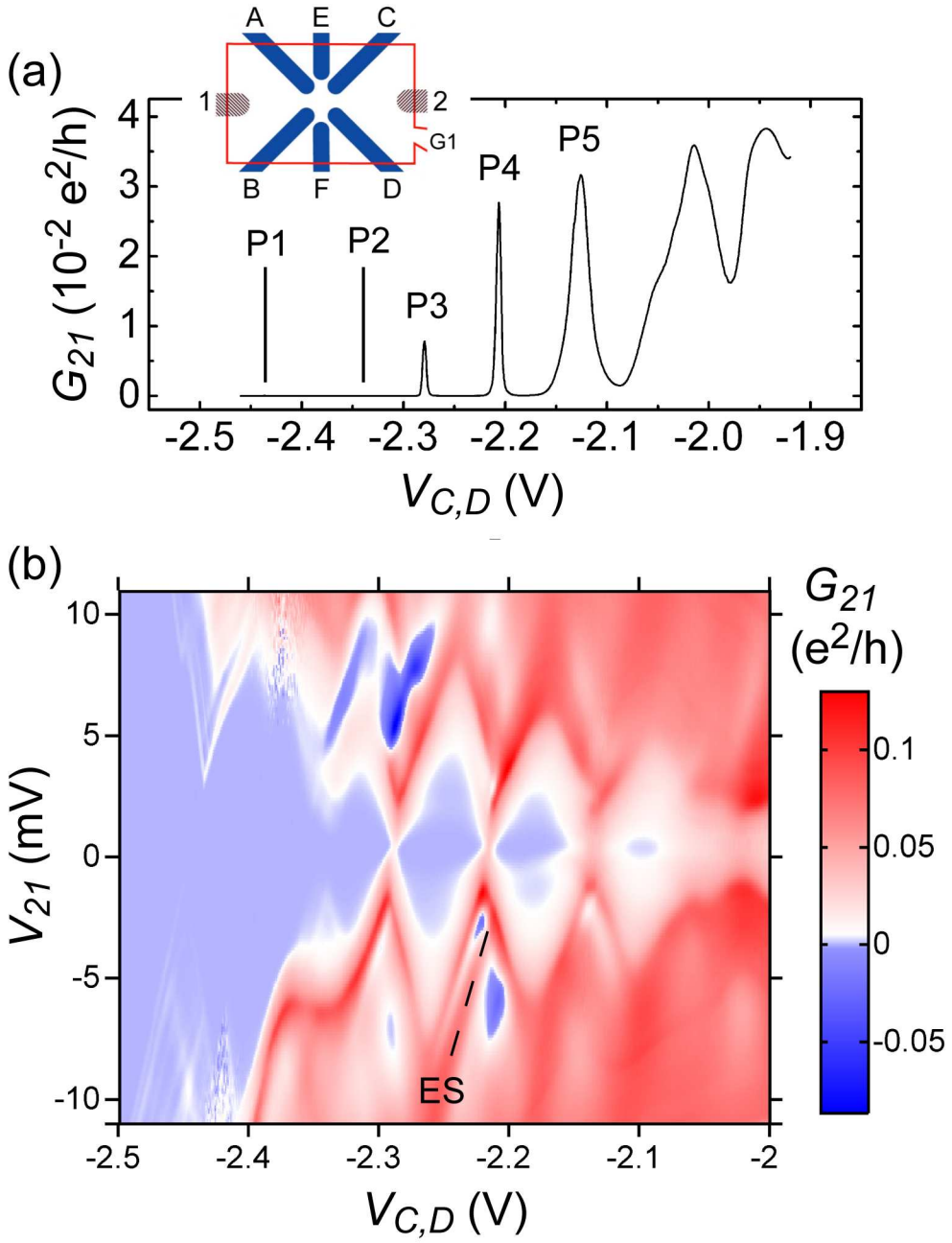

Figure 6.2: (a) The Coulomb blockade oscillations with the first 5 peaks, labeled as P1–P5, when $V_{G1} = 16$ V, $V_A = V_B = 0$ V, $V_E = V_F = 1$ V (single quantum dot bias conditions). (b) Stability chart, taken under the same bias conditions as those in (a). The dashed line indicates an excited state.

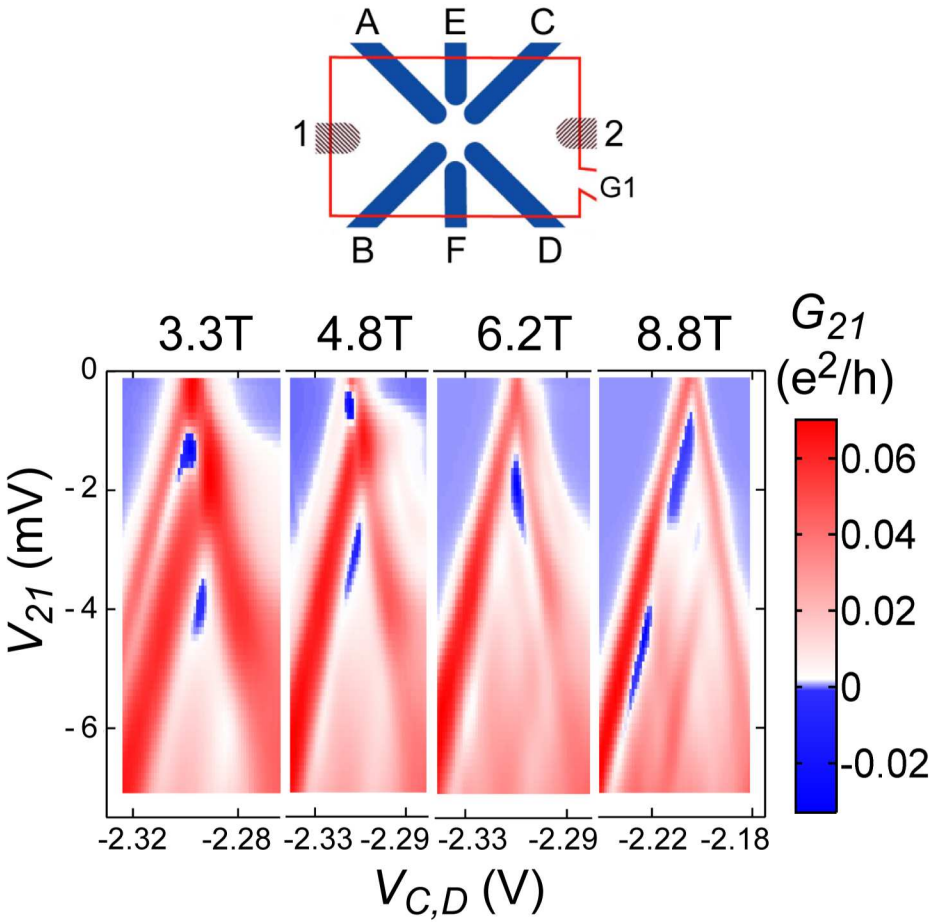

Figure 6.3: Source-drain differential conductance G_{21} as a function of V_{21} and $V_{C,D}$, near peak 4, at the magnetic fields of 3.3, 4.8, 6.2, and 8.8 T, respectively. The excited state moves toward the ground state as the magnetic field increases.

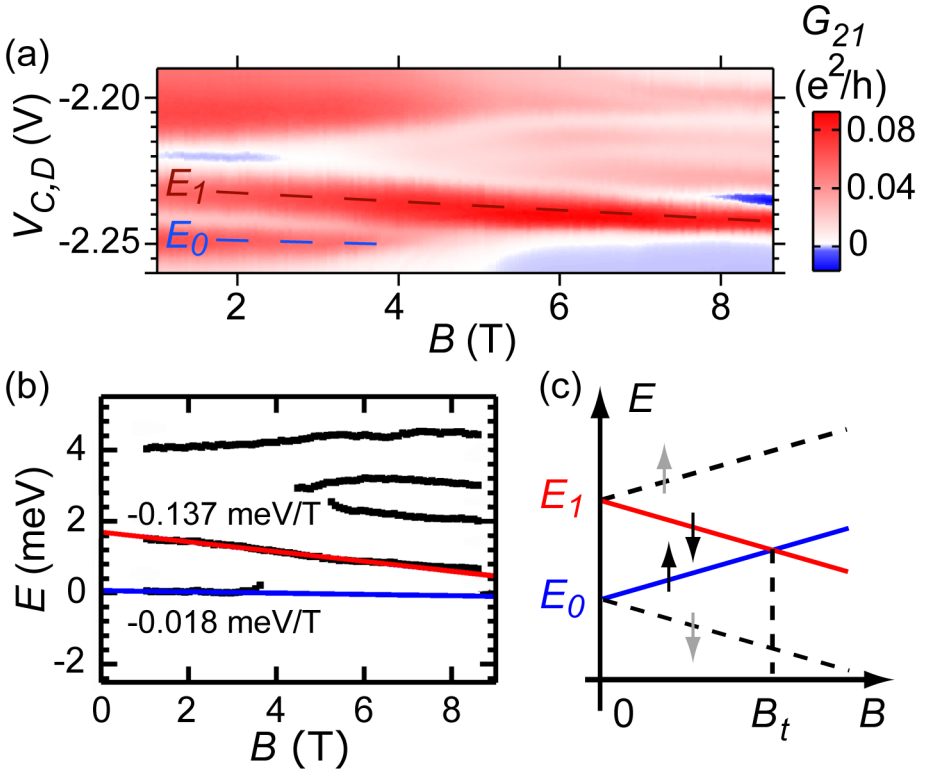

Figure 6.4: (a) The excitation stripe taken at $V_{21} = -4.7$ mV with B between 1 and 9 T. (b) Peak positions as a function of B are extracted from the raw data in (a) using Gaussian fitting, and then converted into electrochemical potential by $E = e\alpha V_{C,D}$ with an arbitrary offset. The straight lines are from linear fitting of the data. The ground state E_0 is only fitted from 1 to 3 T. Their slopes are labeled in the graph. (c) Schematic showing the evolution of single-particle energy levels E_0 and E_1 driven by the Zeeman effect, where E_0 is the ground state with a spin-up electron, and E_1 is the first excited state with a spin-down electron, depicted in solid lines. In a two-electron system, the B dependence of the singlet and the triplet states follows $E_0(B)$ and $E_1(B)$, respectively. Therefore we also label the singlet state as E_0 and the triplet state as E_1. The crossing between E_0 and E_1, the singlet-triplet transition, occurs at B_t.

evolution of the corresponding ground state and the excited state in magnetic field. For the data presented in this section and Sec. 6.3, the field is applied perpendicular to the sample surface. Figure 6.3 shows that this excited state moves toward the ground state as the magnetic field increases. When the source-drain DC bias V_{21} is fixed at -4.7 mV, and the source-drain conductance G_{21} is measured against B and $V_{C,D}$, the excitation stripe near peak 4 is shown in Fig. 6.4(a). With the increase of magnetic field ($0 < B < 6$ T), the distance between the first excited state E_1 and the ground state E_0 decreases. As $B > 6$ T, the state E_1 becomes the new ground state. The maxima in Fig. 6.4(a) are fitted with Gaussians (as functions of $V_{C,D}$), and $V_{C,D}$ can be converted into electrochemical potential, using $E = e\alpha V_{C,D}$, where the ratio of the gate capacitance to the total capacitance $\alpha = 0.09$, extracted from the slopes of nearby diamonds in the stability chart.[22] After an arbitrary offset, the resulting peak positions (in electrochemical potential) are shown in Fig. 6.4(b). The absolute energy position of a peak in the excitation spectrum is determined by practical experimental parameters; we therefore focus on the difference between the most important features, that is, the energy difference between the E_0 state and the E_1 state. To first order, both E_0 and E_1 show apparent linear dependence on magnetic field B. We therefore fit $E_0(B)$ (in the range 1 T $< B <$ 3 T) and $E_1(B)$ (in the range 1 T $< B <$ 9 T) by linear lines. Straight lines for the guidance of the eye are shown in Fig. 6.4(b). The energy difference, $E_1(B) - E_0(B)$, $(= [-0.137 - (-0.018)]$(meV/T)$\cdot B)$, is -0.119 meV/T. The Zeeman splitting in bulk silicon is 0.116 meV/T ($= g\mu_B B$, with $g = 2$, μ_B is the Bohr magneton.). It is clear that the energy difference mostly can be attributed to the Zeeman effect.

Figure 6.4(c) illustrates the Zeeman effect and the ST transition. Here we only consider the two electrons in the outmost shell. The lower lying electrons, if any, are ignored, because they have less influence on the spin dynamics. For a one-electron system, the ground state is a spin-down state in a magnetic field. The Zeeman splitting has the linear magnetic field dependence, $\pm g\mu_B B/2$. In Fig. 6.4(c), the spin of the second electron is depicted by solid black arrows. When this second electron is added to the quantum dot, it would interact with the electron already in the dot and form either singlet or triplet states. As shown in Fig. 6.4(c), quantum selection rule dictates the spin of incoming electrons, that is, a spin-up electron for the singlet state E_0, and a spin-down electron for the triplet state E_1. At low magnetic fields, $0 < B < B_t$, the singlet state has a lower energy than that of the triplet state. While at a higher magnetic field ($B > B_t$), the triplet state E_1 becomes the ground state. The spin configuration of the ground state of the two-electron system is therefore controlled by the magnetic field. Were the Zeeman splitting the only effect to be considered, the magnetic field induced ST transition should occur at about 14 T. This estimate is based on the data shown in Fig. 6.4(b), and there is ~ 1.6 meV energy difference between E_1 and E_0 at zero field. But the observed ST transition happens at about 5 T. This discrepancy can be explained by the direct Coulomb interaction and the exchange interaction using a two electron model.

First, we introduce two single particle spatial wave functions: ψ_0 with energy E_0' and ψ_1 with energy E_1'. Because of the Zeeman effect, energy levels E_0' and E_1' will split in a magnetic field (the same as E_0 and E_1 in Fig. 6.4(c)). Then we

use these two single particle wave functions to construct the singlet and the triplet states of the two electron system

$$\Phi_S = \psi_0(\vec{r}_1)\psi_0(\vec{r}_2)|00\rangle, \tag{6.1}$$

and

$$\Phi_T = \frac{1}{\sqrt{2}}[\psi_0(\vec{r}_1)\psi_1(\vec{r}_2) - \psi_0(\vec{r}_2)\psi_1(\vec{r}_1)]|1-1\rangle, \tag{6.2}$$

where $\vec{r}_1$ and $\vec{r}_2$ label the two electrons, and $|00\rangle$ and $|1-1\rangle$ are the spin parts of the wave functions.

When the electron wave functions shrink with increasing magnetic field B, the interdependence of Coulomb interaction and single-particle states should be taken into account. If we consider the first-order corrections due to the direct Coulomb interaction and the exchange interaction, the total energy of the two electron singlet state is

$$U_S(2) = E'_{0\uparrow} + E'_{0\downarrow} + C_{00}(B), \tag{6.3}$$

and the total energy of the two electron triplet state is

$$U_T(2) = E'_{0\downarrow} + E'_{1\downarrow} + C_{01}(B) - K_{01}(B), \tag{6.4}$$

where $\uparrow$ and $\downarrow$ are spin-up and spin-down of an electron, $C_{00}(B)$ is the direct Coulomb interaction when both electrons are in the ψ_0 state, and $C_{01}(B)$ and $K_{01}(B)$ are the direct Coulomb interaction and the exchange interaction respectively when one electron occupies the ψ_0 state and the other occupies the ψ_1 state. For the wave functions ψ_i and ψ_j, the direct Coulomb interaction C_{ij} and the exchange interaction

K_{ij} are

$$C_{ij} = \iint |\psi_i(\vec{r}_1)|^2[e^2/4\pi\varepsilon|\vec{r}_1 - \vec{r}_2|]|\psi_j(\vec{r}_2)|^2 d\vec{r}_1 d\vec{r}_2$$
$$K_{ij} = \iint \psi_i^*(\vec{r}_1)\psi_j^*(\vec{r}_2)[e^2/4\pi\varepsilon|\vec{r}_1 - \vec{r}_2|]\psi_j(\vec{r}_1)\psi_i(\vec{r}_2) d\vec{r}_1 d\vec{r}_2.$$

The energy of the one electron ground state is

$$U(1) = E'_{0\downarrow}. \tag{6.5}$$

So the electrochemical potential is $\mu_S(B) = U_S(2) - U(1) = E'_0 + g\mu_B B/2 + C_{00}(B) = E_0(B) + \Delta C_{00}(B)$ for the singlet state, where $E'_0 + C_{00}(0) = E_0(0)$, and $\Delta C_{00}(B) = C_{00}(B) - C_{00}(0)$. For the triplet state, the electrochemical potential is $\mu_T(B) = U_T(2) - U(1) = E'_1 - g\mu_B B/2 + C_{01}(B) - K_{01}(B) = E_1(B) + \Delta C_{01}(B) - \Delta K_{01}(B)$, where $E'_1 + C_{01}(0) - K_{01}(0) = E_1(0)$, $\Delta C_{01}(B) = C_{01}(B) - C_{01}(0)$ and $\Delta K_{01}(B) = K_{01}(B) - K_{01}(0)$.[22] Both the direct Coulomb interactions and the exchange interaction increase with increasing B, since the size of the wave functions decreases. For the triplet state, we hypothesize that $\Delta C_{01}(B)$ is largely compensated by $\Delta K_{01}(B)$, due to the apparent linear B dependence of the triplet state; while the singlet electrochemical potential $\mu_S(B)$ rapidly increases with increasing B, because the two electrons both occupy the same single particle spatial wave function ψ_0, and the $\Delta C_{00}(B)$ term can increase fast near the ST transition. So the electron-electron Coulomb interaction (including the direct Coulomb interaction and the exchange interaction) plays a significant role here and drives the singlet-triplet transition at a much lower magnetic field between 4 and 6 T.

There are some minor deviations from the linear B dependence of the triplet state, which can be from the $\Delta C_{01}(B) - \Delta K_{01}(B)$ term, and also can be from other

effects, for example the corrections from a more accurate self-consistent Hartree-Fock treatment or the electron correlation effect beyond the Hartree-Fock approximation, and the possible orbital effect in a high magnetic field, etc.[22] A detailed analysis requires the precise information of the electrostatic potential in the quantum dot, which is beyond the scope of this dissertation.

There is another interesting phenomena in Fig.6.4(a). The conductance of the singlet state drops when $B > 4$ T, and it becomes invisible after the ST transition, whereas the conductance of the triplet state doesn't change much. This can be understood by the transport blockade effect, which will be discussed in the next section.

6.3 Electron spin blockade in a silicon SET

The magnetic field dependence of the first five Coulomb blockade peaks is shown in Fig. 6.5. Peak position and peak amplitude are extracted from this raw data by fitting each peak with equation, [25]

$$G/G_{max} = \cosh^{-2}(e\alpha(V_g - V_{g0})/2k_BT), \tag{6.6}$$

where G is the source-drain conductance, G_{max} is the maximum conductance at the peak, V_{g0} is the gate voltage at resonance, k_B is the Boltzmann constant, T is the electron temperature, and $\alpha = C_g/C_\Sigma$, the ratio of the gate capacitance to the total capacitance. The shift of a peak position (in $V_{C,D}$) can be converted into the change in electrochemical potential using the formula:$\Delta E(B) = e\alpha[V_{C,D}^{Peak}(B) - V_{C,D}^{Peak}(0)]$. In the stability chart [Fig. 6.2(b)], from the slopes of Coulomb diamonds, α is about

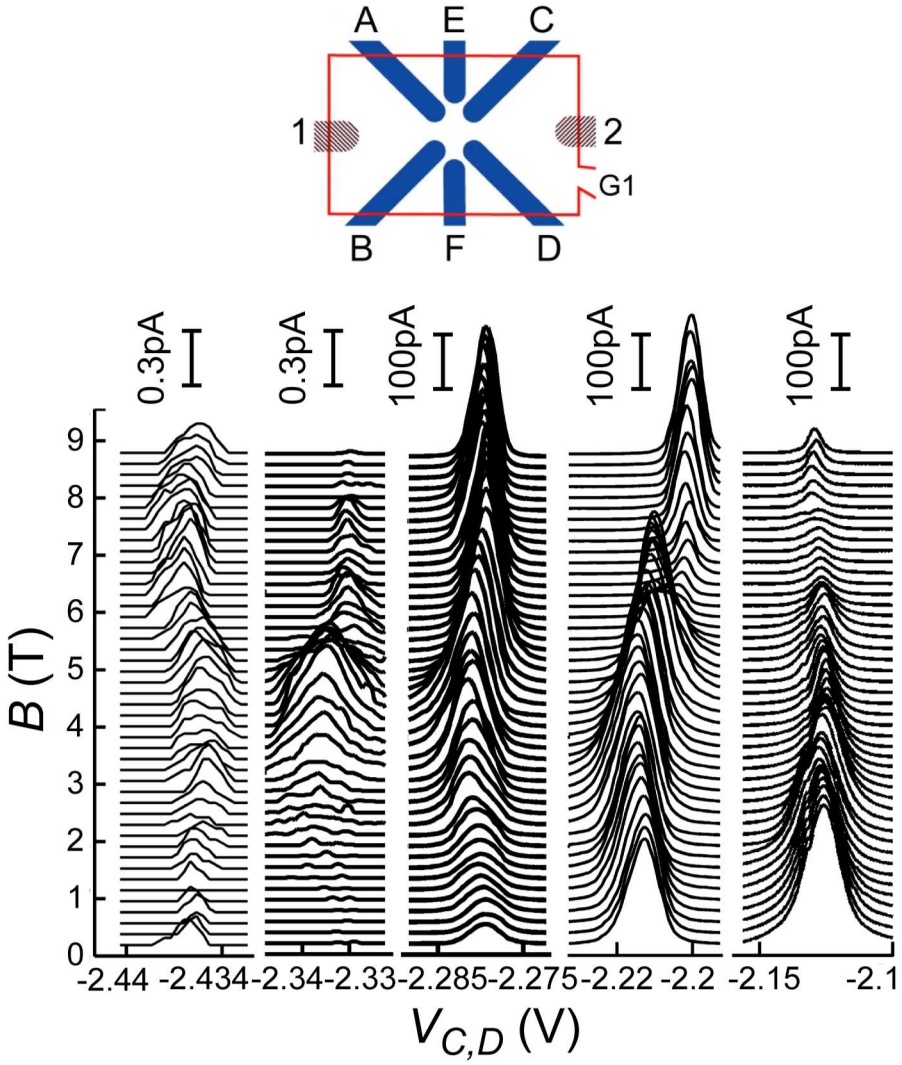

Figure 6.5: Evolution of Coulomb blockade peaks with magnetic field B. Drain current of the first 5 peaks (P1–P5) is measured with $V_{ac} = 0.1$ mV, $V_{G1} = 16$ V, $V_A = V_B = 0$ V, $V_E = V_F = 1$. Each trace is offset linearly with B.

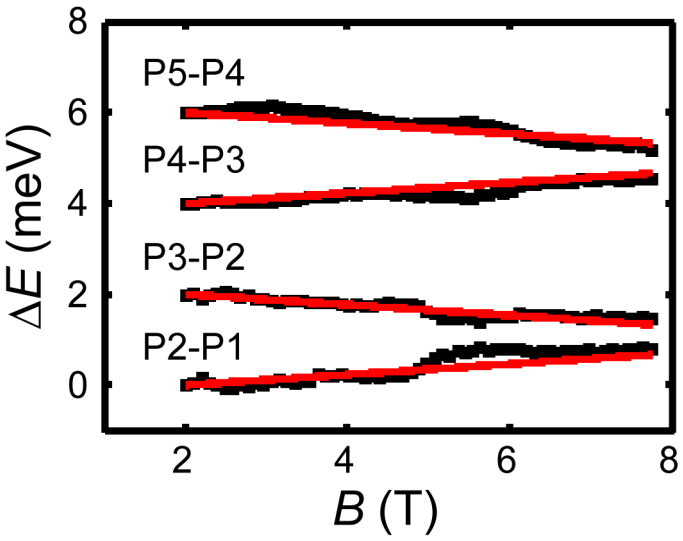

Figure 6.6: Peak spacing between successive peaks as a function of B, extracted from the raw data in Fig. 6.5 by fitting each peak with Eq. (6.6). The peak spacing is in electrochemical potential and is offset for clarity. The straight lines have slopes of $g\mu_B B$ or $-g\mu_B B$, assuming the Zeeman splitting with $g = 2$.

0.15, 0.088, 0.090, 0.090 and 0.061 for peaks 1 to 5, respectively.[22] These changes in electrochemical potential reflect the evolution of the ground state energy levels. Figure 6.6 shows the peak spacing (in electrochemical potential) between successive peaks as a function of magnetic field. Using the energy difference in our analysis not only minimizes the uncertainty due to charge fluctuations and the long term drift in analog electronics, but also helps in identifying the role of spin in the addition energy spectrum. Based on the Zeeman splitting, four straight lines with slopes of either $g\mu_B B$ or $-g\mu_B B$, with g=2, are plotted. Since the second peak only appears between 2 T and 8 T, we compare the first 5 peaks in this field range. Each curve is arbitrarily offset for clarity. Because the electrochemical potential differences between successive peaks show the alternate slopes of $g\mu_B$ and $-g\mu_B$, it is clear that the shifts of the ground states show an spin-down spin-up filling pattern and are dominated by the Zeeman effect at low magnetic field ($B < 4$ T). The fact that the data are well explained by the Zeeman splitting is consistent with our earlier finding that the magnetic field dependence of energy levels (both the ground state and the excited state) is owing to the Zeeman effect.

For the data reported in this work, the magnetic field is applied perpendicular to the sample surface, therefore, the orbital effect might be expected. However, all of the observed magnetic field dependence can be well explained by the Zeeman effect with a g-factor of ~ 2, leaving orbital effect not a significant factor in our system. This is probably caused by the specific shape of our quantum dot. When we squeeze side gates C and D, the electron wave function is elongated along the source-drain direction, the orbital effect becomes not important. This leaves the

B dependence of the peak positions dominated by spin effects.[57] The spin-down spin-up filling pattern requires that the two-fold valley degeneracy is lifted,[57] and the excitation spectrum suggests that the valley splitting should be larger than 1.6 meV. This is probably due to the strong lateral confinement in our quantum dot. The valley splitting in Si quantum dots has not been systematically investigated in experiments, presumably due to the lack of the excitation spectrum data. Some speculate that it is $0.35 \sim 0.46$ meV.[51, 58] Others indicate that it is of order of a few meV and greater even at $B = 0$ T.[59, 18] Further investigation is necessary to reach an affirmative conclusion.

Finally, we turn our attention to the tunneling peak amplitudes in Fig. 6.5. In a lateral quantum dot device, an electron tunnels into and out of the quantum dot from the two-dimensional electron gas (2DEG) leads. As illustrated in Fig. 6.7(b), the electron concentration varies from the bulk value to zero near the tunnel barriers; the Zeeman splitting effect causes the separation of the conduction band minima of the spin-up and spin-down electrons in a magnetic field, which results in the different populations of the spin-up and spin-down subbands. In GaAs, spin polarized injection is due to edge states in a magnetic field perpendicular to the 2DEG.[60] This edge-state picture is probably not valid in our silicon devices, because the Shubnikov-de Haas oscillation minima are not zero up to 9 T in the test hall bar device. However, magnetic field can fully spin polarize a 2DEG with $n_{2d} = 1.5 \times 10^{11}$ cm^{-2} for $B = 5$ T solely with spin effects in a high mobility Si MOSFET.[61, 62] In the test hall bar device, the Shubnikov-de Haas oscillations show that the 2DEG is spin polarized at least at $B = 9$ T for $n_{2d} = 7.12 \times 10^{11}$ cm^{-2}, as we will discuss later.

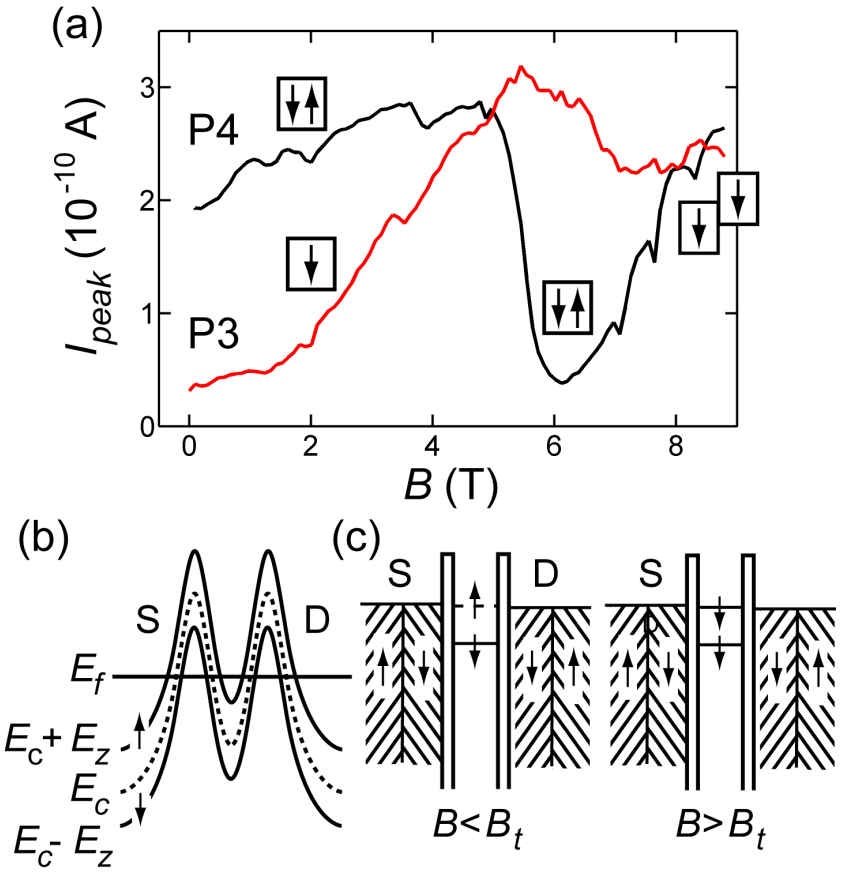

Figure 6.7: (a) Peak amplitude as a function of B. The arrows in the squares indicate the spin configurations of peak 3 (P3) and peak 4 (P4). (b) In a magnetic field, the conduction band minima of spin-down and spin-up electrons split due to the Zeeman effect with $E_z = g\mu_B B/2$. Electrons are fully spin-down polarized in the source and the drain leads near the quantum dot. (c) Schematic showing the transport of a two-electron system with spin-down polarized leads. When $B < B_t$, the ground state is a singlet, which only allows a spin-up electron to tunnel through, i.e., the spin blockade state. When $B > B_t$, the ground state becomes a triplet, which permits a spin-down electron to tunneling through. The spin blockade is thus lifted.

So it is reasonable to expect that electrons will be polarized into the spin-down state in the 2DEG leads near the tunnel barriers above some magnetic field, as illustrated in Fig. 6.7(b). This effectively makes the spin-up electrons tunnel through a thicker barrier, so the injection current is dominated by spin-down electrons, assuming there is no effective spin scattering. Due to the low spin-orbit scattering and lattice inversion symmetry,[63] spin-polarized electrons have long lifetime and transport length in silicon. It has been demonstrated that conduction-band spin-polarized electrons can coherently transport across 10 μm undoped Si.[63] So the assumption should hold well.

In theory, peak amplitude is proportional to the tunneling probability and depends exponentially on how much wave functions in the dot and in the contacts overlap with each other.[57] There are two different mechanisms governing the amplitude modulation of Coulomb blockade peaks in a magnetic field. One is the spatial mechanism when an electron tunnels into different spatial wave functions in a quantum dot, not related to spin; the other is the spin-blockade mechanism due to spin-polarized injection.[64] The lower edge of the excitation stripe in Fig. 6.4(a) reflects the evolution of peak 4 as a function of magnetic field in Fig. 6.5. The conductance of the singlet state drops dramatically when $B > 4$ T, while the conductance of the triplet state doesn't change much, as shown in Fig. 6.4(a) and Fig. 6.5. This difference can be explained by both the spatial effect and the spin-blockade effect, since the singlet state requires a spin-up electron which occupies a smaller wave function at the center, while the triplet state requires a spin-down electron which occupies a larger wave function. However, the electron states corresponding

to peak 3 (a spin-down single electron state) and peak 4 (a singlet state) should have similar (single-particle) spatial wave functions; their peak amplitude should have similar dependence on the magnetic field if there is no spin blockade effect. In Fig. 6.7(a), the amplitude of peak 4 decreases much faster than peak 3 when the magnetic field is larger than 4 T. A reasonable explanation is the spin polarized injection from the leads.[64] For 0.1 mV small ac excitation between the source and the drain, only the ground state can lie in the transport window. For a singlet state, only a spin-up electron can tunnel into and out of the quantum dot, the amplitude of peak 4 decreases dramatically owing to this spin polarized injection, and as shown in Fig. 6.7(a), it occurs at $B > 5$ T. This forms the spin blockade. After $B > 6$ T, the amplitude of peak 4 increases with the magnetic field, because the ground state corresponding to peak 4 changes from a singlet state to a triplet state. For the triplet state, the incoming spin-down electron experiences lower and thinner barriers, thus the higher conductance. Since the amplitude of Coulomb blockade peaks is determined by the two-electron spin configuration in the dot with spin polarized leads, it can be used to distinguish a singlet state from a triplet one. Monitoring the amplitude is thus a form of spin blockade spectroscopy.[60]

6.4 Further discussion

6.4.1 Orbital effect

To further verify the orbital effect, I have repeated the peak position as a function of magnetic field B measurement at two different orientations of B, i.e.

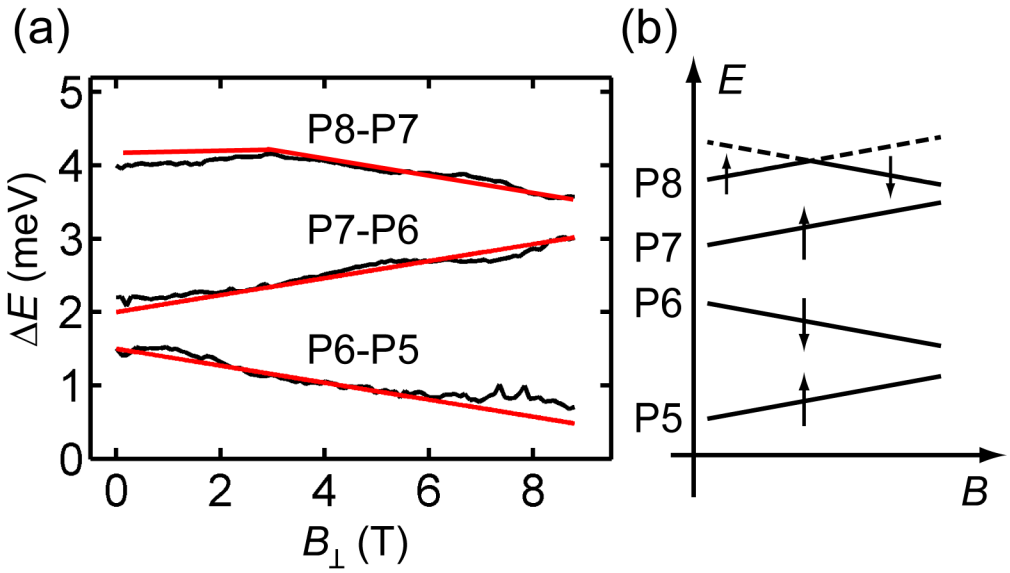

Figure 6.8: (a) Peak spacings between successive peaks as a function of $B_\perp$. Since only peaks 5, 6, 7 and 8 are well visible in the magnetic field range between 0 and 9 T, they are analyzed. The peak spacings are in electrochemical potential and offset for clarity. The straight lines have slopes of $\pm g\mu_B B$ or 0, assuming the Zeeman splitting with $g = 2$. The device is SiHB6L20070102(2,3). The Coulomb blockade oscillations and the stability chart are shown in in Fig. 5.17 in Chapter 5. (b) Possible spin configurations for peaks 5, 6, 7 and 8, corresponding to the straight lines in (a), assuming only the Zeeman splitting effect.

$B_{||}$ and $B_{\perp}$, as shown in Fig. 6.1. Because a non-rotating sample holder was used, I had to cool down the device twice. After the first measurement (in $B_{\perp}$), the device was de-soldered from the sampler holder, mounted parallel to the magnetic field, and then measured in $B_{||}$. There is no direct comparison between these two measurements because of the change of device characteristics, but we still can reach some basic conclusions.

The investigated device is SiHB6L20070102(2,3). The Coulomb blockade oscillations and the stability chart in $B_{\perp}$ are shown in Fig. 5.17 in Chapter 5. Because peak 2 disappears between 4 and 6 T, and peak 4 becomes a double peak in some B range, I only compare peaks 5, 6, 7 and 8. The peak spacings between successive peaks (in electrochemical potential) are shown in Fig. 6.8(a). The straight lines have slopes of $\pm g\mu_B B$ or 0, assuming the Zeeman splitting with $g = 2$. So the shifts of the ground state energy levels as a function of B mostly can be attributed to the Zeeman effect. The corresponding spin configurations are shown in Fig. 6.8(b).

Figure 6.9 shows the Coulomb blockade oscillations and the stability chart in a magnetic field parallel to the direction of the source-drain current. Because peaks 2 and 3 are not visible between 4 and 6 T, only peaks 4, 5, 6, and 7 are analyzed. The peak spacings between successive peaks (in electrochemical potential) as a function of B are shown in Fig. 6.10(a). It is expected that there is no orbital effect in a magnetic field parallel to the direction of the source-drain current. As shown in Fig. 6.10, the shifts of the ground state energy levels as a function of B can also be explained by the Zeeman effect.

For a quantum dot with a two-dimensional (2D) parabolic confining potential

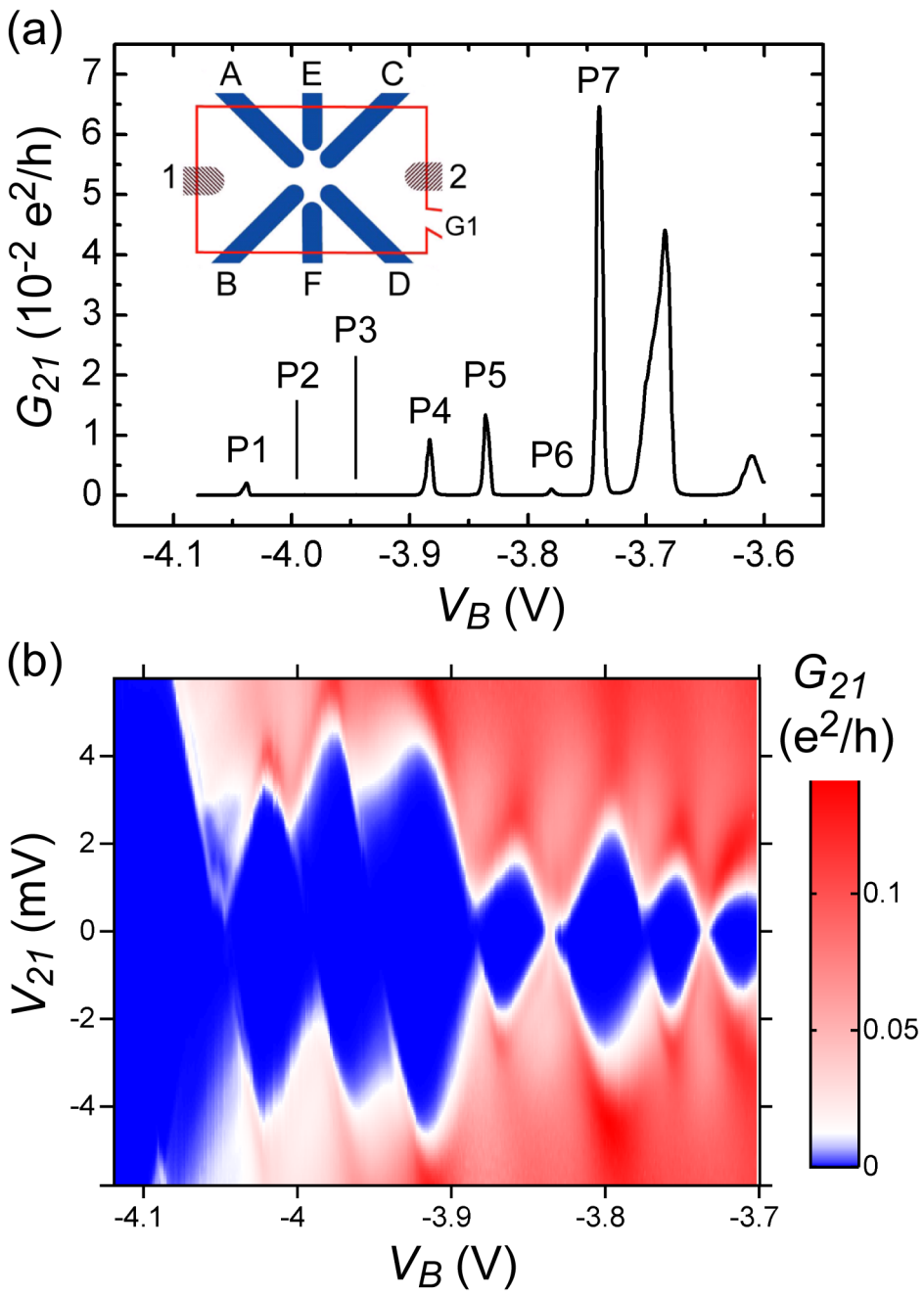

Figure 6.9: (a) The Coulomb blockade oscillations with the first 7 peaks, labeled as P1–P7, when $V_{G1} = 25$ V, $V_A = -3.4$ V, $V_C = V_D = 1$ V, $V_E = V_F = 0.8$ V. The device is SiHB6L20070102(2,3). (b) Stability chart, taken under the same bias conditions as those in (a).

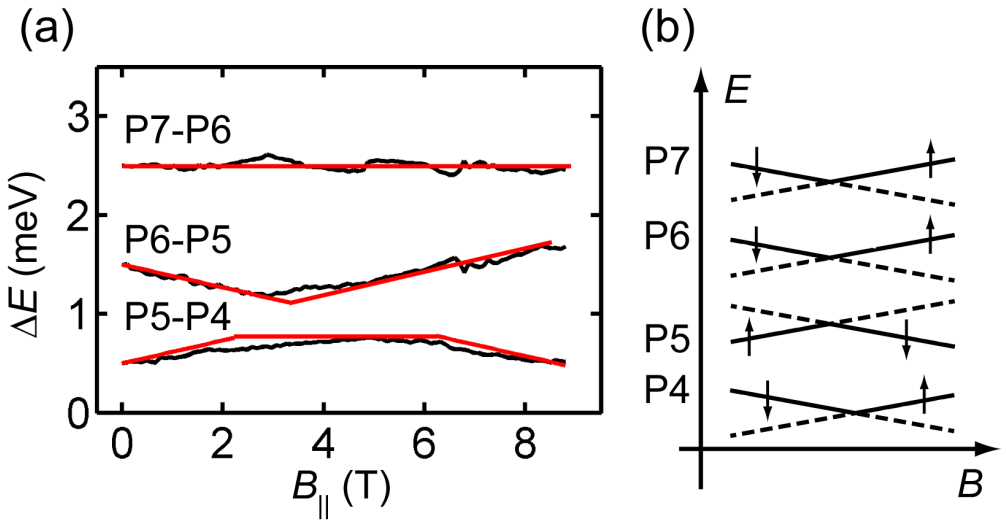

Figure 6.10: (a) Peak spacings between successive peaks as a function of $B_{||}$. Since only peaks 4, 5, 6, and 7 are well distinguishable in the magnetic field range between 0 and 9 T, they are analyzed. The peak spacings are in electrochemical potential and offset for clarity. The straight lines have slopes of $\pm g\mu_B B$ or 0, assuming the Zeeman splitting with $g = 2$. (b) Possible spin configurations for peaks 4, 5, 6, and 7, corresponding to the straight lines in (a), assuming only the Zeeman splitting effect.

and cylindrical symmetry, when the magnetic field is applied perpendicular to the 2D electron gas, the orbital effect causes the shift of peak spacing between two successive peaks by $\sim \hbar\omega_c = \hbar qB/m^* = 0.6$ (meV/T)$\cdot B$ in silicon, where ω_c is the cyclotron frequency, and m^* is the electron effective mass.[22] (In contrast, the shift of peak spacing between two successive peaks is about 0.116 meV/T due to the Zeeman effect.) Both Fig. 6.6 and Fig. 6.8(a) do not show such large shift of peak spacing when the magnetic field is swept. When the quantum dot doesn't have cylindrical symmetry, and the 2D parabolic confining potential is elongated along one direction and strongly confined along the other direction, the orbital effect becomes not important, as shown in Sec. 2.1.5.[65]

Here I want to point out that the spin configuration analysis based on the peak position as a function of magnetic field B measurement is not always reliable, because such measurement is not quite accurate, and the contribution from electron-electron Coulomb interaction and other effects can further complicate the result. For example, in Fig. 6.6, the peak spacing (P4-P3) is well aligned with the straight line with the slope of $g\mu_B B$ between 2 and 8 T, but we know from the excitation spectrum that the electron for peak 4 is spin-up between 2 and 4 T, and spin-down between 6 and 8 T, while the electron for peak 3 is spin-down in the whole range, so it should be a horizontal line between 6 and 8 T.

Comparing Fig. 6.8(a) to Fig. 6.10(a), we don't find any obvious difference when the device is in a perpendicular magnetic field or in a parallel magnetic field. We conclude that there is no obvious orbital effect within the accuracy of the measurement. If there is any orbital effect, it is less than the Zeeman effect in the

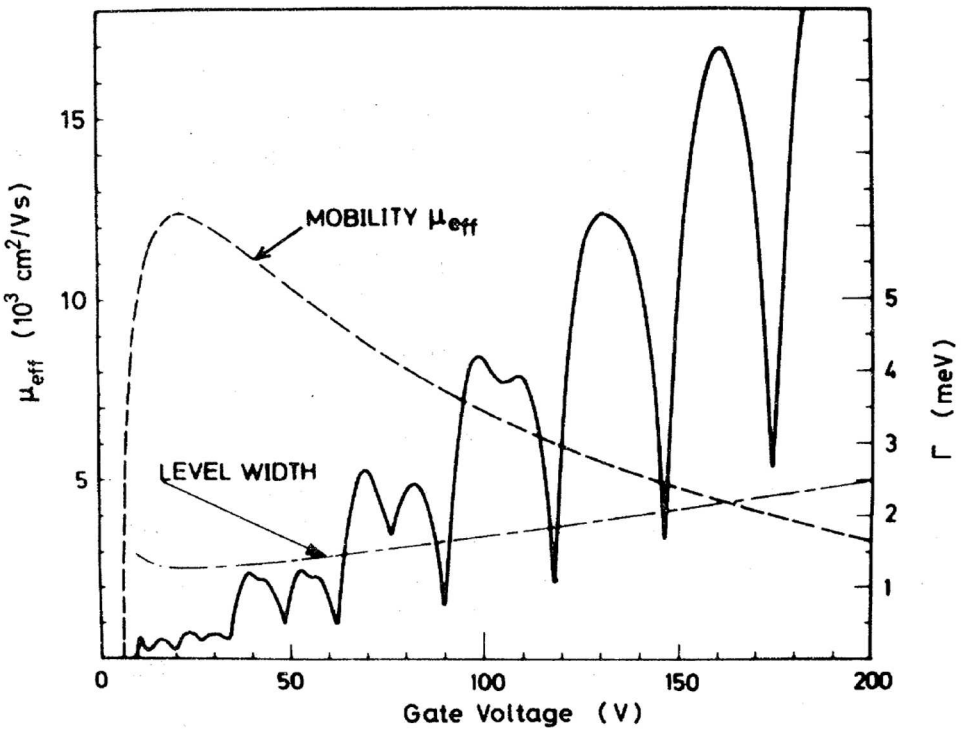

Figure 6.11: An example of the oscillations of the conductivity observed by Kobayashi and Komatsubara (Ando et al., 1972a; Komatsubara et al., 1974) in an n-channel inversion layer on a Si (100) surface at $B = 9.5$ T. The effective mobility and the corresponding level width Γ are also shown. After Ando and Uemura (1974a). (p546 in Ref.[51])

discussed magnetic field range.

6.4.2 Spin polarization in silicon

In a Si(100) inversion layer, the spin splitting and the valley splitting are observed in the Shubnikov-de Haas oscillations in strong magnetic fields and at relatively low electron concentrations, as shown in Fig. 6.11. This is due to the Landau level broadening. In the presence of scatterers, each Landau level is broadened.

Within the self-consistent Born approximation, the Landau level width[51]

$$\Gamma = \sqrt{\frac{2}{\pi}\hbar\omega_c\frac{\hbar}{\tau_f}} = \frac{q\hbar}{m^*}\sqrt{\frac{2B}{\pi\mu}} \sim \sqrt{\frac{B}{\mu}}, \tag{6.7}$$

in the case of short-range scatterers ($d < l/\sqrt{2\nu+1}$), where the cyclotron frequency $\omega_c = qB/m^*$, the relaxation time $\tau_f = \mu m^*/e$, m^* is the electron effective mass, μ is the electron mobility, d is the order of the range, $l = \sqrt{\hbar/eB}$ is the magnetic length, ν is the Landau level index. The Zeeman splitting is $E_z = g\mu_B B \sim B$, so the relative resolution $E_z/\Gamma \sim \sqrt{B}$ increases with an increasing B. The valley splitting also increases with an increasing magnetic field, although a detailed relationship is not available yet. The mobility peak is located at a relatively low electron concentration, where the width of the Landau level also reaches its minimum. This improves the resolution of the spin splitting and the valley splitting. It is believed that the enhancement of the spin splitting and the valley splitting also improves the resolution at low electron concentrations.[51]

For our test hall bar device, the longitudinal conductivity σ_{xx} was measured at $B = 8.8$ T (Fig. 6.12). Backscattering is strongly suppressed, when the Fermi level is located at the middle of two landau levels. This results in the conductivity peaks in σ_{xx}, which coincide with the hall conductivity σ_{xy} plateaus for 2D electron gas. The conductivity peaks corresponding to $\nu = 8$, 12 can be clearly identified. Although it is not quite clear, the features near $V_{G1} = 15$ V should correspond to $\nu = 3$, 4. The observation of a conductance peak corresponding to $\nu = 3$ requires both the valley splitting and the spin splitting. In this test hall bar device, $\mu = 5000$ cm^2/Vs and $B = 8.8$ T, so $\Gamma = 2.1$ meV (using Eq. 6.7). On the other hand, the

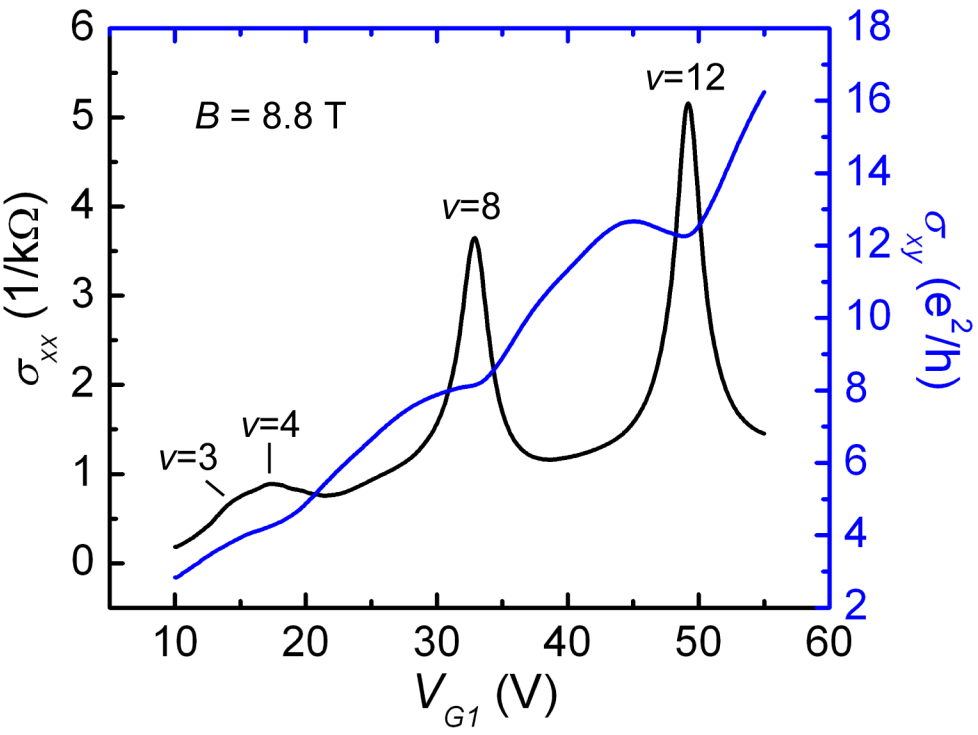

Figure 6.12: The longitudinal and the Hall conductivities of the test hall bar device at B = 8.8 T. ν is the Landau level index.

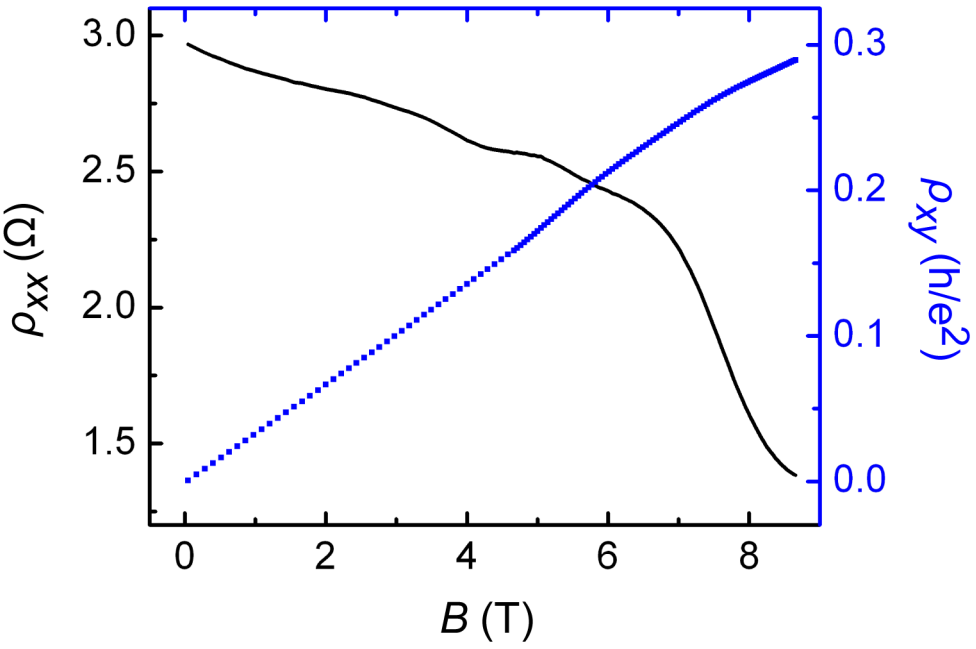

Figure 6.13: The longitudinal and the Hall resistances of the test hall bar device at at $V_{G1} = 14$ V with $n_{2d} = 7.12 \times 10^{11}$ cm^{-2}.

Zeeman splitting $E_z = g\mu_B B = 1$ meV with $g = 2$. Even with the enhancement of the spin splitting and the valley splitting, the resolution still can be not great, but there should be some signature.

In addition, the Shubnikov-de Haas oscillations were measured at $V_{G1} = 14$ V with $n_{2d} = 7.12 \times 10^{11}$ cm^{-2}. The hall resistance ρ_{xy} begins to nonlinearly depend on magnetic field when $B = 4.5$ T, as in Fig. 6.13. If we fit ρ_{xx} in 8 T$< B <$ 8.8 T with a parabolic function, the ρ_{xx} reaches its minimum at $B = 9$ T, with a plateau $\rho_{xy} \approx 0.3h/e^2$. The filling factor is about 3, since $n_{2d}/(eB/h) = 7.12 \times 10^{11}\text{cm}^{-2}/2.2 \times 10^{11}\text{cm}^{-2} = 3.2$. The hall resistance plateau corresponding to an odd number of filling factor only can be explained by the different populations at spin-up and spin-down subbands and the valley splitting. Although it is difficult

to determine the ratio of the densities of the spin-down and spin-up electrons due to the limited data, the spin polarization of electrons is resolvable at least at $B = 9$ T with $n_{2d} = 7.12 \times 10^{11}$ cm^{-2} in this test hall bar device.

6.5 Quantum interference and false excited states

In Fig. 6.2(b), when we look closely on the stability chart near the first peak, some fine features are revealed, i.e. the parallel lines in the conducting region (Fig. 6.14(b)). We believe that these "excited states" are false, based on two facts. First, the energy differences between these states are only about 0.5 meV, which are much less than 1.6 meV that we observed in the stability chart near the fourth peak. Second, the excitation spectrum shows that the difference between the first "excited state" and the ground state doesn't change much (less than 0.5 meV) when the magnetic field is swept from 0 T to 9 T, as shown in Fig. 6.16. This is in contradiction with the Zeeman effect. For real excited states, we would expect some obvious motion of the energy levels and energy level crossing when the magnetic field is swept from 0 T to 9 T, since the Zeeman splitting energy is about 1 meV at $B = 9$ T. These "excited states" are most likely from the quantum interference effect at the source/drain leads.[66][67]

It has long been known that conductance fluctuations occur in a mesoscopic system, when its dimensions are smaller than the phase coherence length l_φ (limited by inelastic scattering). Electron waves travel coherently along different paths in the device. Because these paths have different phase shifts, quantum interference

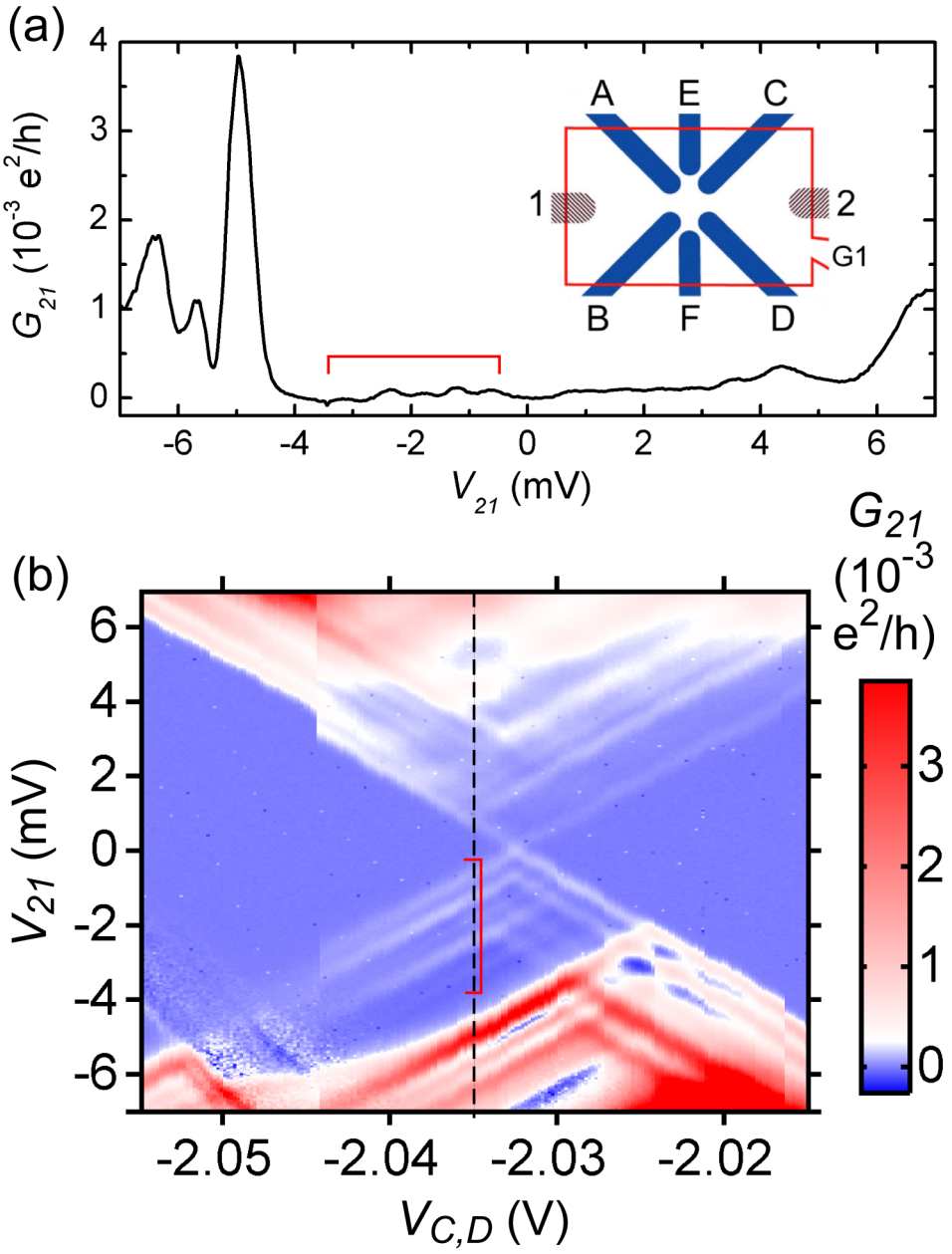

Figure 6.14: (a) Cross section at $V_{C,D} = -2.035$ V. (b) Stability chart at $B = 0$ T. The dash line is the location of the cross section. The device is SiHB6L20070102(2,4). The data were measured in the first cool-down.

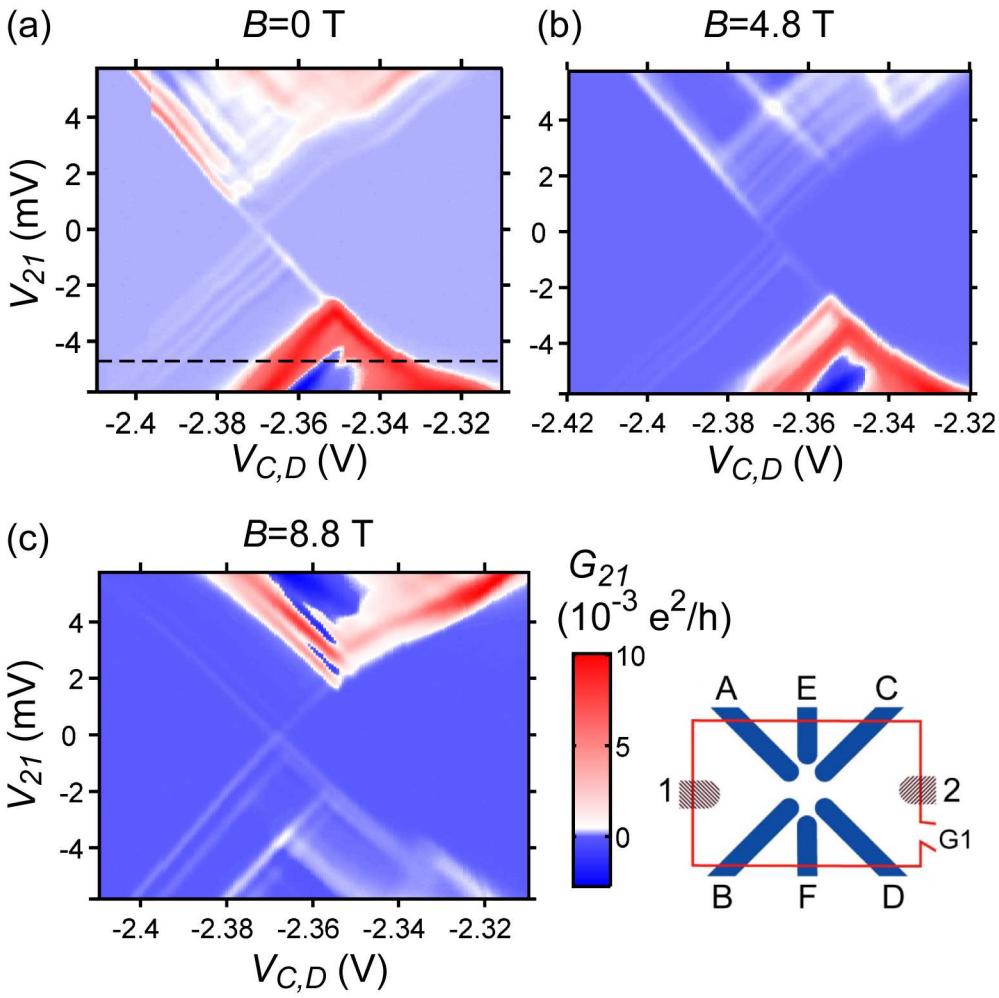

Figure 6.15: Stability charts at (a) $B = 0$ T, (b) $B = 4.8$ T, and (c) $B = 8.8$ T.The dash line in (a) is at $V_{21} = -4.7$ mV, where the excitation spectrum is measured. The device is SiHB6L20070102(2,4). The data were measured in the second cool-down.

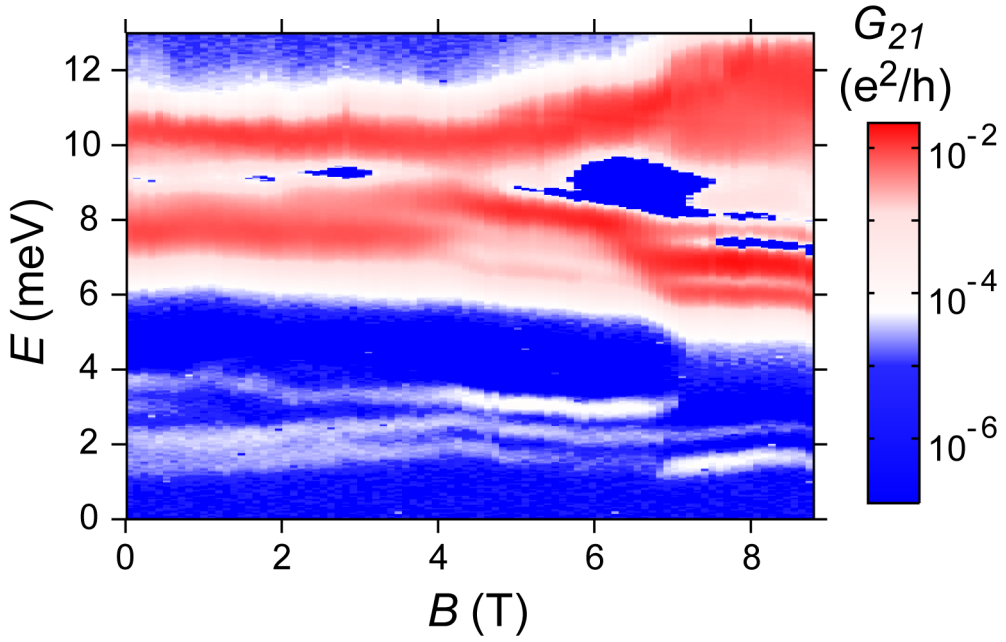

Figure 6.16: The excitation spectrum near the first peak at $V_{21} = -4.7$ mV. The peak positions are in electrochemical potential, $E = e\alpha V_{C,D}$, with an arbitrary offset. The energy difference between the ground state and the first "excited state" is barely changed.

causes the conductance fluctuations. In previous experimental studies, the quantum interference pattern depends on the configuration of disorders. The phase shifts of different paths are sensitive to changes in magnetic field, electric field, and the configuration of scatterers. Any changes of these parameters will induce variations in the phase shifts, so the conductance fluctuations exhibit the random nature, as shown in the universal conductance fluctuations (UCF).[68] In contrast, some "excited states" are nearly equally spaced in Fig. 6.14(b) and Fig. 6.15, and keep unchanged when the magnetic field is swept from 0 T to 9 T (Fig. 6.16). These features are also stable after thermal cycling (Fig. 6.14 and Fig. 6.15). Since the magnetic field is applied perpendicular to the sample surface, and there is little motion of some features when the magnetic field is swept, these fine "excited states" are probably from an open quantum dot defined electrostatically in the point-contact channel and by local potential fluctuations, similar to the previous intrinsic quantum dot. Thus the possible paths that electron waves travel through are quasi one-dimensional, and the phase shifts of these paths are not sensitive to the changes of the magnetic field.

Here I discuss a simple one-dimensional example to illustrate that the quantum interference can introduce these fine features. It includes an SET with 2 rectangular potential barriers, and a parabolic potential well at the drain (Fig. 6.17). In the quantum Coulomb blockade regime, the conductance is given by[25]

$$G \propto \frac{\Gamma^l \Gamma^r}{\Gamma^l + \Gamma^r}, \tag{6.8}$$

where Γ^l and Γ^r are the tunneling rates from the quantum dot to the left and

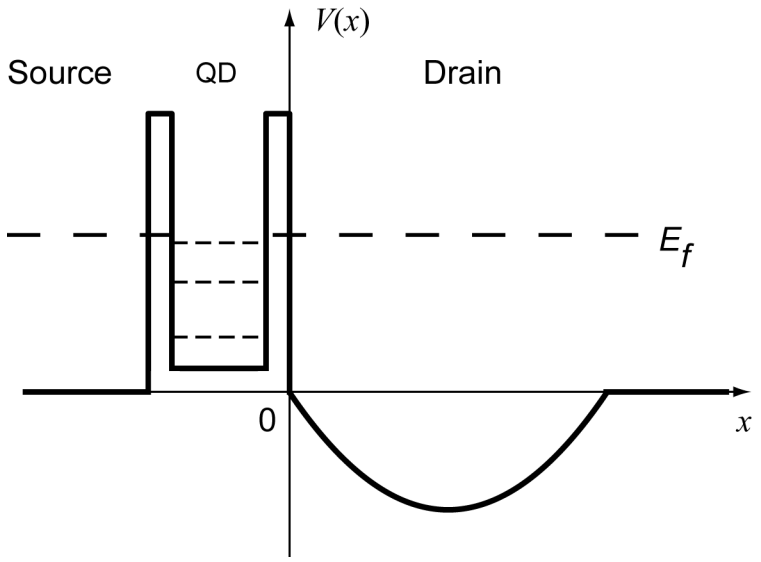

Figure 6.17: The one-dimensional potential profile of an SET with a parabolic potential well at the drain lead.

the right leads respectively. So the conductance of the SET is modulated by the transmission coefficients of the source and drain potential barriers. Here it is also modulated by the transmission coefficient of the parabolic potential well. I have numerically calculated the transmission coefficient of the parabolic potential well using the Numerov algorithm.[69] The result is shown in Fig. 6.18. There are transmission coefficient peaks with approximately equal distances, which are consistent with the results in the excitation spectrum. As shown in Fig. 6.19, the quantum state (E_{QD}) inside the quantum dot acts as a probe to detect the transmission coefficient of an electron (or the local density of states).[66]

If we consider the spin blockade effect, only spin-down electrons can effectively tunnel through the quantum dot. Because the electron energy is relative to the conduction band minima in the relationship between the electron energy and the

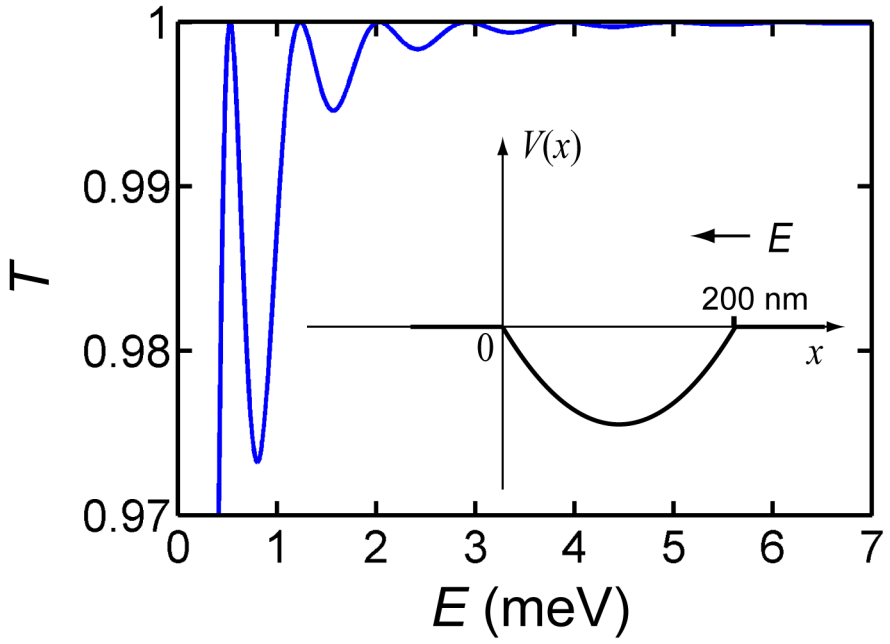

Figure 6.18: The transmission coefficient of an electron with energy E, flying over a parabolic potential well with $\hbar\omega_0 = 0.3$ meV. The potential width is 200 nm. The inset shows the potential profile.

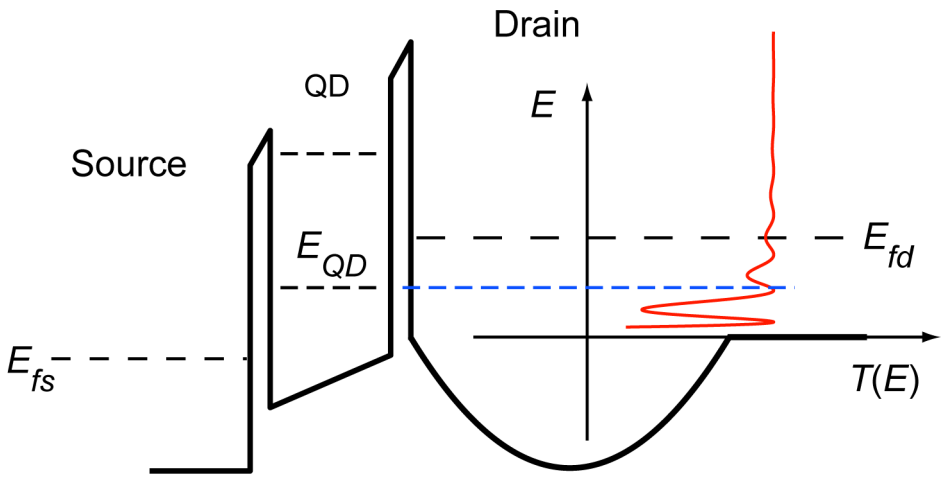

Figure 6.19: Schematic of the conductance fluctuations caused by the quantum interference in the drain lead ($V_{21} < 0$). The quantum state (E_{QD}) inside the quantum dot acts as a probe to detect the transmission coefficient of an electron (or the local density of states).

transmission coefficient, and the magnetic field induced Zeeman splitting only causes the common motion of the potential profile, the energy difference between the first "excited state" and the ground state will not change even when the magnetic field is swept.

However, there are some more features in the data that I cannot explain. The quantum interference effect at the source/drain leads provides a general explanation. Many-body effects can also play a role here. For example, the peak spacing can also include the electron-electron interaction (charging energy) when electrons are partially localized in the parabolic potential well, as we discussed in Chapter 2.

6.6 Summary

In this chapter, we have investigated a single intrinsic quantum dot electrostatically defined in a point-contact channel and by the local potential fluctuations both in a perpendicular magnetic field and in a parallel magnetic field.

In the perpendicular magnetic field, excitation spectroscopy was used to probe the excited states as well as the ground state, so the ST transition was observed unambiguously and analyzed in detail. The results show that the magnetic field dependence of the excitation spectrum mostly can be attributed to the Zeeman effect; however, the electron-electron Coulomb interaction plays a significant role in the ST transition, and thus it happens at a much lower magnetic field.

In addition, the evolution of Coulomb blockade peaks with the magnetic field B was also measured. The data suggest that the ground state energy levels also shift with the applied magnetic field owing to the Zeeman effect. However, such measurement is not quite accurate. And the spin configuration analysis based on these data is not always reliable. So the excitation spectroscopy has inherent advantage. The evolution of peak amplitudes illustrates the spin blockade effect, which is also confirmed by the excitation spectrum.

After comparing the peak position shifts of Coulomb blockade oscillations in a perpendicular magnetic field and in a parallel magnetic field, we didn't find any obvious difference. So we conclude that the orbital effect is less than the Zeeman effect in the investigated magnetic field range.

Later, we have demonstrated that the quantum interference effect at the

source/drain leads can introduce some excited-state like signals in the excitation spectrum. This makes the excitation spectrum complicate. Fortunately the ground states are unique, because they are the boundaries of Coulomb blockade diamonds, and can not be false. So the excitation spectrum near the fourth peak is still valid.

Chapter 7

Conclusion and future work

7.1 Conclusion

When we design the metal-oxide-semiconductor single electron transistor (MOS-SET), we would expect that the device consists of an electrostatically defined quantum dot (QD). It is always interesting to find out the difference between reality and expectation. In this case, we have systematically investigated the MOS-SETs. Our specific 6 side-gate structure enables us to verify the formation of QDs and to determine their locations, because the capacitive coupling strength between the dot and the gate implies the distance between them. This approach provides us much more insight to understand the device operation than the traditional method in the literature, i.e. Coulomb blockade oscillations and diamond-shaped stability chart method. The traditional method only proves the existence of a quantum dot, but doesn't provide any information about its location. However, some authors presumed that their quantum dots are completely electrostatically defined with little direct evidence.[70][13] Here we have demonstrated that intrinsic QDs can easily form in point-contact channels, and QDs can be from the potential fluctuations. This can explain the discrepancy between the size of observed dots and the dimensions of the devices in early works.[70][13] We believe that intrinsic QDs are due to the potential fluctuations at the Si/SiO_2 interface, an inherent property of the inter-

face. We have also demonstrated that intrinsic QDs can be removed. One possible explanation of the origin of intrinsic QDs is from negatively charged interface traps, as has been proposed by M. A. Kastner.[17] But there can be other explanations. The exact origin of the intrinsic QDs is not clear yet. We believe that intrinsic QDs are caused by the localized nature of interface trapped charge, oxide fixed charge, and chemical inhomogeneity.[56] The interface quality can be characterized by the interface trap density, so the size of intrinsic QDs can also relate to the interface trap density.

One interesting question is the strength of lateral confinement of a point-contact channel in our devices. Due to the difficulty of the self-consistent simulation at low temperatures, I only simulated it at room temperature, and found out that the parabolic potential can be represented as an harmonic oscillator with $\hbar\omega_0 \approx$ 7 meV, which is larger than $\hbar\omega_c = 5.5$ meV at 9 T in silicon, where ω_c is the cyclotron frequency. Another interesting thought is the explanation of Fig. 5.8. When $V_C \neq V_D$, the center of the point-contact channel can move away from the middle of the side gates C and D, and signals in the graph can be related to the different electrostatic potential profiles at different locations (Fig. 7.1). Can we get any information of the potential fluctuations from this? There are some interesting questions here.

On the other hand, we have taken advantage of our understanding on the MOSFETs, and isolated out one single intrinsic QD. We have investigated the quantum dot in a magnetic field, and observed the singlet-triplet transition and spin blockade effect, which had not been reported previously. The magnetic field is applied

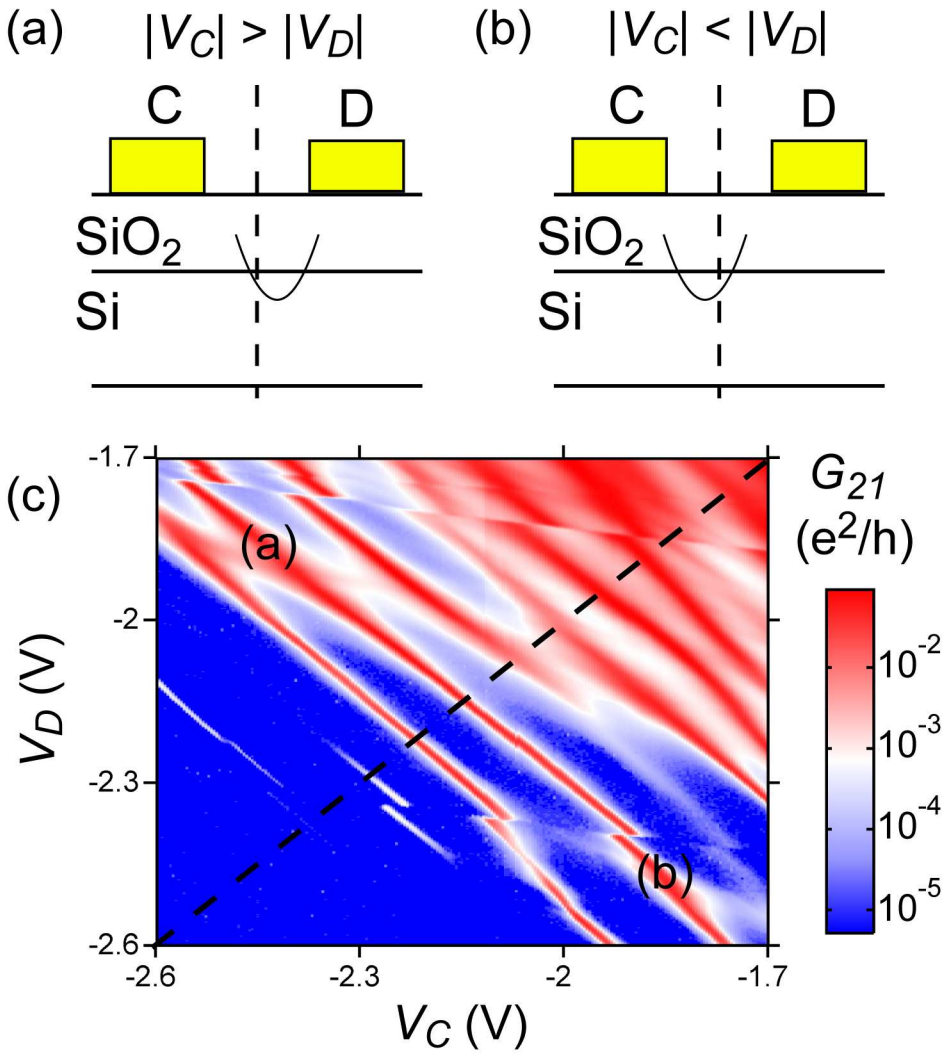

Figure 7.1: In (a) and (b), side gates C and D are at the top of SiO_2 and Si layers. The parabolic lines are the lateral confinement of the point-contact channels. In (a), when $|V_C| > |V_D|$, the center of the point contact channel moves towards side gate D. In (b), when $|V_C| < |V_D|$, the center moves towards side gate C. (c) The source-drain conductance G_{21} as a function of side gate voltages V_C and V_D, when $V_{G1} = 16$ V, $V_A = V_B = 0$ V, $V_E = V_F = 1$ V, and $V_{ds} = 0$ V. Signals in the graph can be related to the different electrostatic potential profiles at different locations (a) or (b).

perpendicular to the sample surface. However, the magnetic field dependence of the ground-state and excited-state energy levels of the QD mostly can be attributed to the Zeeman effect, with no obvious orbital effect up to 9 T. We believe that it is due to the strong lateral confinement in the point-contact channel. But it is still open for debate.

Finally, we return to the electrostatically defined silicon quantum dots. We have demonstrated that intrinsic QDs can be removed, and an electrostatically defined dot can be created by scaling down the devices. Now the size of the observed dot and the dimensions of the device are consistent. But these devices currently only work in the many-electron regime.

7.2 Future work

Single electron confinement in an electrostatically well defined silicon quantum dot is our goal. At this point, it is still an open question whether this goal can be reached or not. In order to realize single electron confinement, substantial improvements are necessary. As we have discussed in Sec. 5.5, there are two length scales — the device dimensions and the localization length. The device dimensions should be smaller than the localization length. So there are two basic approaches. One is to improve the quality of the Si/SiO_2 interface, so that the localization length can be increased. The other is to scale down the device, so that the device dimensions are reduced.

The fabrication process should be further optimized. For example, probably we

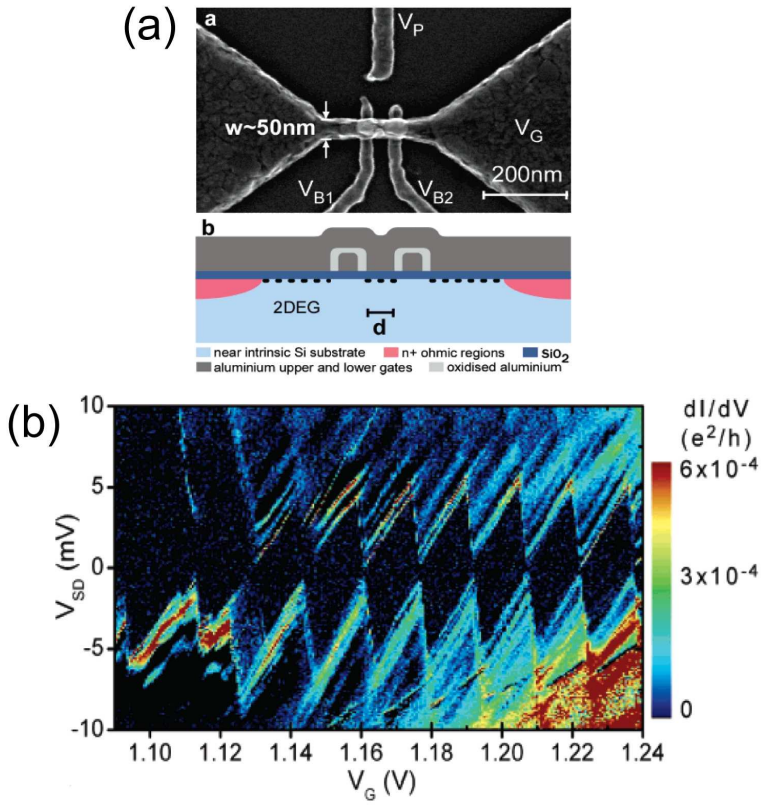

Figure 7.2: (a) Device structure of a small Si SET with a 40 nm quantum dot ($d < 40$ nm). (b) Stability chart of the SET, working in the few electron regime. (After S. J. Angus et al., Ref.[14])

need better thermal oxide. It has been shown that the e-beam evaporation process used in the metal gate deposition can introduce many interface traps because of the x-ray radiation (Sec. 4.3.2), which can be replaced by the resistive thermal evaporation. It is possible that e-beam lithography can also damage the interface. UV-cured Nanoimprint lithography (NIL) can be a solution, or high temperature post annealing can also help to solve the issue. People at Sandia National Lab are working on optimizing the fabricaion process.[20]

Device down-scaling is also necessary. From the size of the intrinsic QD, we

can estimate that the device dimension should be less than 60 nm for our current interface quality, which is not difficult to reach. An Australian group demonstrated a small device with a 40 nm quantum dot(Fig. 7.2), which works in the few electron regime, although the device quality must still be improved.[14]

With continuous down-scaling and improvement, it is quite possible to reach the few-electron regime, and even realize single electron confinement.

Appendix A

Measurement system and techniques

A.1 Introduction

In this appendix, we will discuss the measurement system and techniques in detail. There are 20 pairs of twisted wires which run from room temperature to cryogenic temperatures, and each device needs 9 signal wires, so 4 devices are installed on the sample holder in each cool-down.

The whole measurement system includes three parts: a dilution refrigerator, a superconducting magnet and its power supply, and measurement electronics, as shown in Fig. A.1.

A.2 Dilution refrigerator

In order to resolve small energy differences, such as excited states and the Zeeman splitting, devices should be measured at low temperatures as possible. We use a powerful MNK 126-700 dilution refrigerator from Leiden Cryogenics B.V.. Its base temperature is about 10 mK, and the cooling power at 120 mK is about 700 μW. In our measurement, the electron temperature is estimated about 400 mK.

There is a superconducting magnet inside the liquid helium dewar, which can be used to apply a magnetic field up to 9 T on the sample. As shown in Fig. A.2(b), samples can be mounted vertically or horizontally, so that the sample surface can

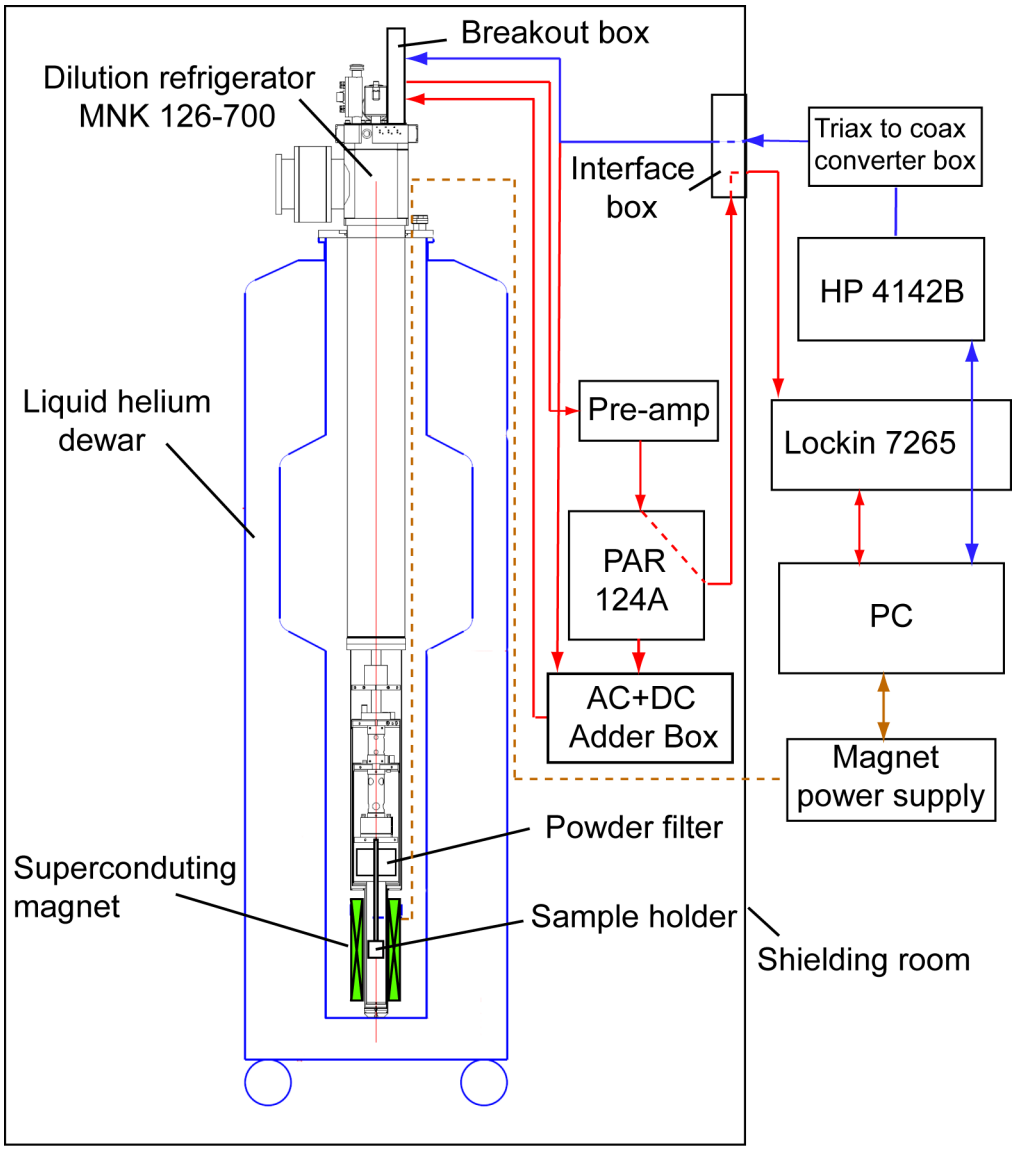

Figure A.1: Schematic of the measurement system, including a dilution refrigerator, a superconducting magnet and its power supply, and measurement electronics. The measurement electronics and the superconducting magnet power supply are controlled by computer programs. Blue lines form the DC measurement signal paths. Red lines form the AC measurement signal path; the AC signal from the PAR 124A and the DC signal from the interface box are divided by voltage dividers before entering the AC+DC adder box.

be parallel or perpendicular to the magnetic field.

A.3 DC measurement system

The DC measurement system consists of an HP 4142B semiconductor parameter analyzer (including 8 DC Source/Monitor units (SMUs)), a triax to coax converter box, and a shielding room interface box. Inside the shielding room, DC signals are connected from the interface box to the dilution refrigerator breakout box. The DC signal paths are shown as blue lines in Fig. A.1. The HP 4142B is controlled by computer programs through GPIB interface, and data are collected at the same time. Since the HP 4142B has 8 SMUs, it can independently control 6 side gates, the top gate, and the source-drain voltage. At the same time, it can measure the gate leakages and the source-drain current. This is quite convenient.

A detailed wiring for each channel is shown in Fig. A.3. Since we short the sense and the force terminals of each channel in the triax-to-coax converter box, and the HP4142B is not well calibrated, there is an offset voltage between $-5 \sim 5$ mV when the SMU is set to 0 V. So the source-drain voltage is set at 10 mV in the DC measurement.

A.4 AC measurement system

A.4.1 Basic setup of the AC measurement system

In the AC measurement system, the source-drain conductance is measured by the standard AC lock-in technique using a 37 Hz, 0.1 mV excitation voltage, since

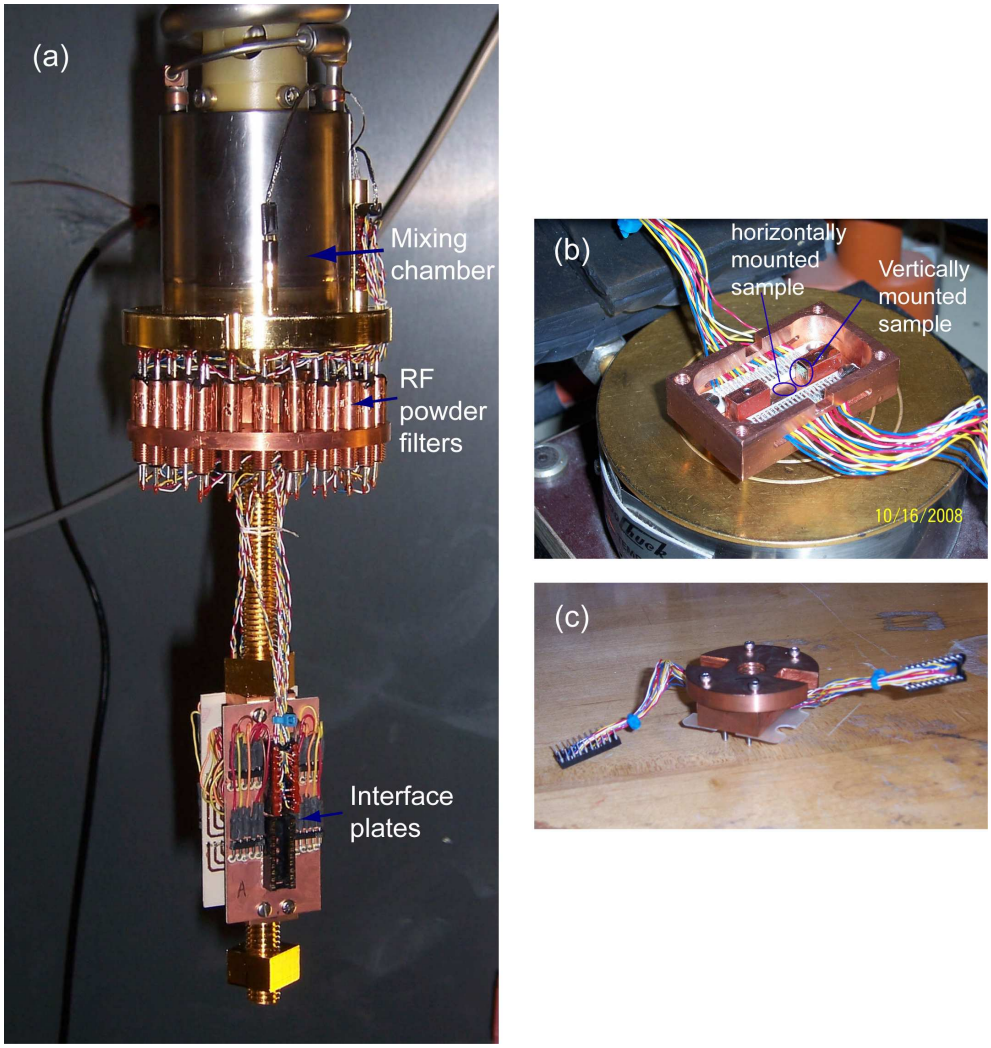

Figure A.2: Sample wiring near the mixing chamber. (a) RF powder filters are installed underneath the mixing chamber, and all signal wires are filtered by them. Then signal wires are connected to the interface plates. (b) Samples can be mounted vertically or horizontally. Horizontally mounted samples are not shown here. (c) After samples are installed on the sample holder, a copper plate is attached at the top of the sample holder. Then the copper plate is screwed on the lower end of the threaded rod in (a). Two 20-pin DIP heads of the sample holder are plugged in the DIP sockets on the interface plates.

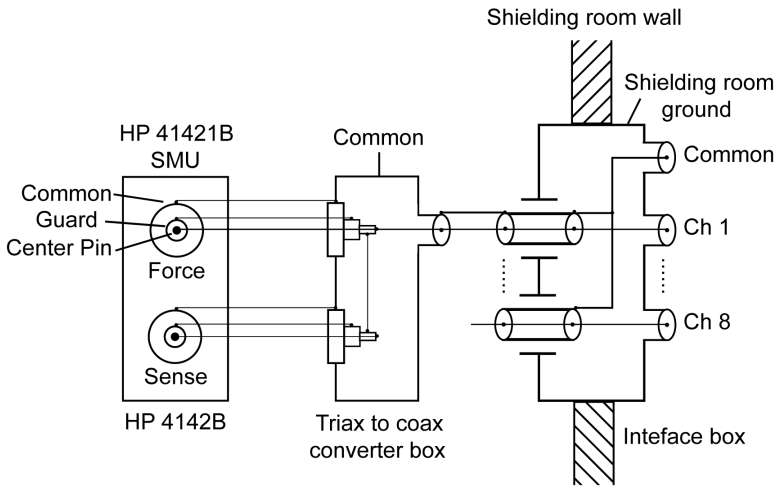

Figure A.3: Detailed wiring of the DC measurement system, from the HP 4142B, the triax to coax converter box, to the shielding room interface box.

AC lock-in technique can effectively filter out noise, and makes the high precision measurement possible. As shown in Fig. A.1, the AC measurement system includes an analog lock-in amplifier PAR 124A, some voltage dividers, an AC+DC adder box, a pre-amplifier, and another lock-in amplifer SIGNAL RECOVERY 7265.

The measurement circuit is shown schematically in Fig. A.4. The PAR 124A generates a 37 Hz sine wave signal with $V_{rms} = 1$ V. The signal goes through a 1000:1 voltage divider(100 KΩ : 10 Ω), and becomes $V_{rms} = 0.1$ mV. Since the required source-drain DC voltage in our experiment is about 25 mV, and the minimum voltage range for the HP 4142B is 2 V, the DC voltage from channel 1 goes through a 86:1 voltage divider (4.22 KΩ : 50 Ω). Then the AC and the DC voltages are added together by an AC+DC adder box, and are applied between the source and the drain of an SET. The source-drain current is amplified by a transimpedance amplifier (pre-amp), and measured by the lock-in amplifier PAR 124A. The monitor

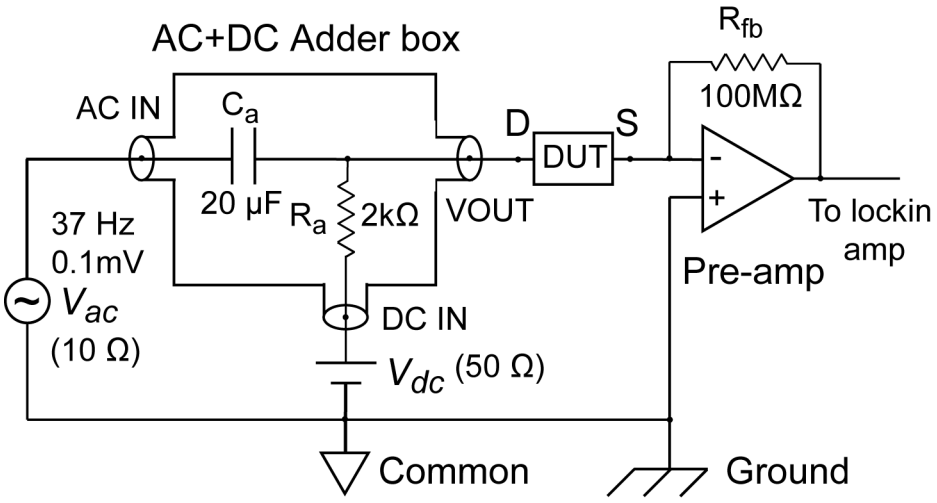

Figure A.4: Schematic of the AC measurement system. The lock-in amplifier PAR 124A generates a 37 Hz sine wave signal(V_{rms} = 1 V), which goes through a 1000:1 voltage divider(100 KΩ:10 Ω), and becomes V_{ac} = 0.1 mV. The DC voltage from channel 1 of the HP 4142B goes through an 86:1 voltage divider (4.22 KΩ:50 Ω) and becomes V_{dc}. Then the AC and the DC voltages are added together by an AC+DC adder box, and are applied between the source and the drain of the device under test (DUT). The source-drain current is amplified by a transimpedance amplifier (pre-amp), and measured by the lock-in amplifier PAR 124A.

signal and the reference signal from the PAR 124A are fed into the lock-in amplifier 7265. Then data from lock-in amplifier 7265 are collected by computer programs through GPIB interface.

The AC+DC adder box is essentially a high-pass filter with $f_{3dB} = 1/(2\pi RC) =$ 4 Hz. From Fig. A.4, the transfer function T is $0.99\angle 6°$, and the output impedance is about 213 Ω at 37 Hz. In the AC measurement, the source-drain impedance is usually above 25.8 kΩ, so the output impedance of the adder box can be neglected. The frequency response of the adder box is also measured by an HP 41941A gain-phase analyzer with output channel impedance 50 Ω and test channel impedance 1 MΩ, while the DC IN input of the adder box is shorted by a 50 Ω terminator. The result is shown in Fig. A.5.

In order to minimize the ground loop issue and noise, care has been taken to separate the AC power line ground, the shielding room ground, and the electronic circuit common. The electronic circuit common is shorted to the shielding room ground in the lock-in amplifier PAR 124A. The pre-amp common can also be shorted to the shielding room ground directly or through 50 Ω terminator (the semi-floating ground). In principle, there should be only one ground point, but I have found that these two grounding configurations both work fine.

A.4.2 Filtering

When wires run from room temperature to the device in the dilution refrigerator, they need to be carefully filtered to reduce the injection of radio frequency

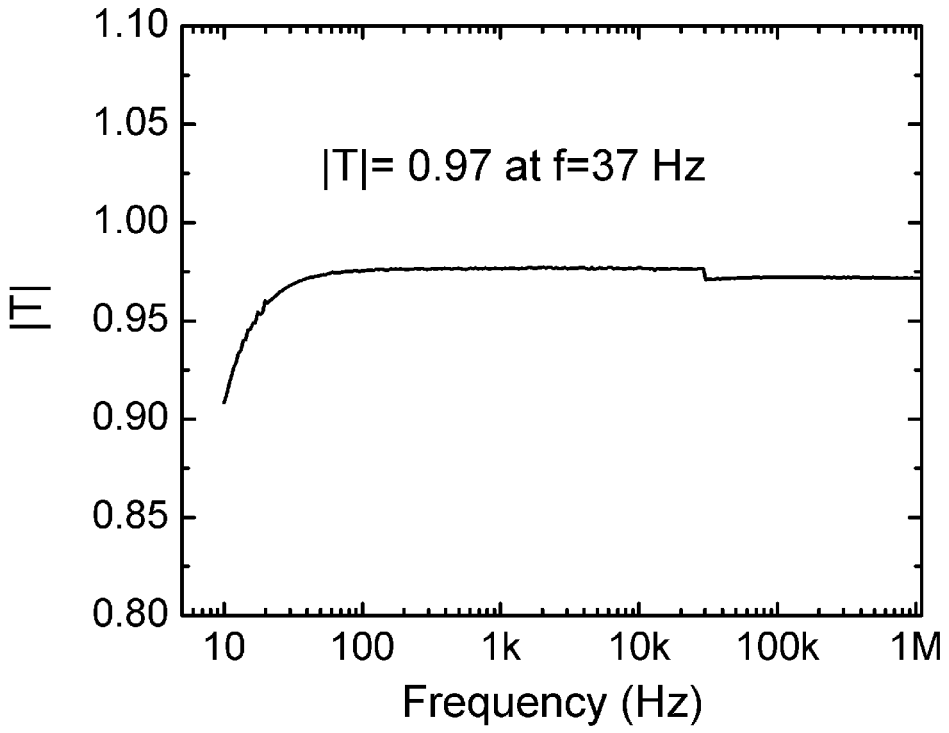

Figure A.5: The frequency response of the AC+DC adder box with $f_{3dB} = 4$ Hz, measured by an HP 41941A gain-phase analyzer with output channel impedance 50 Ω and test channel impedance 1 MΩ, while the DC IN input of the adder box is shorted by a 50 Ω terminator.

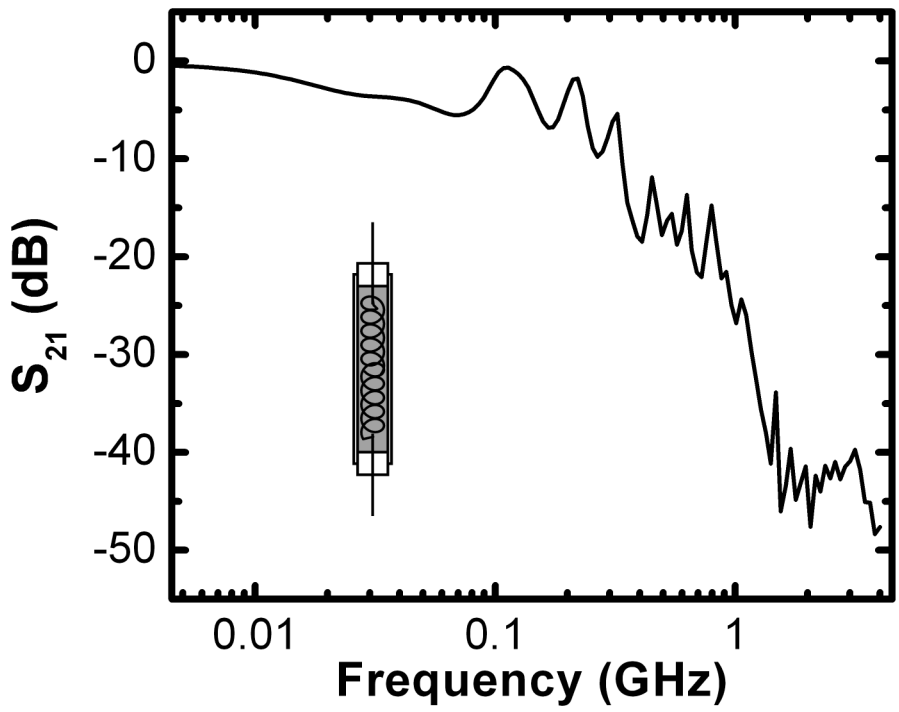

Figure A.6: The frequency response of an RF powder filter, measured in a 50 Ω environment with a Rohde & Schwarz vector network analyzer. The inset shows the cross sectional view of the RF powder filter. A copper wire spiral is inside a copper tube. The stainless steel powder is packed inside the tube. One DC feedthrough is soldered at each end of the spiral, and is sealed on the tube.

energy into the cryostat. In addition to the shielding room, three different filtering stages are used to cover the different frequency ranges.

At base temperature, all signal wires are filtered by RF powder filters.[71][72] We use stainless steel powder instead of copper powder, since stainless steel powder is more efficient than copper powder.[71] The stainless steel powder is from Alfa Aesar with Stock #88390 (type 316-L, 325 mesh, nominal 44um). For each RF powder filter, one 0.6 m gauge 36 copper wire (coated with heavy formvar) is wound

into a 3 mm diameter 70 turn spiral, and half of the spiral is counterwound to reduce magnetic field pick up. This spiral is fitted in a 4 mm diameter, 25 mm long cavity in a copper tube. We solder one DC feedthrough at the end of the spiral, and use stycast 2850FT epoxy to seal this feedthrough on the tube. After the spiral goes through the tube, another DC feedthrough is soldered on it. The stainless steel powder is packed inside, and the second DC feedthrough is sealed on the tube. The cross sectional view of an RF powder filter is shown in Fig. A.6 as an inset. The powder filter can effectively absorb the high-frequency noise, and attenuate signal above 1 Ghz by 40 dB, as shown in Fig. A.6.

Commercial low-pass filters (BLP-1.9+ from Mini-Circuits) are installed on the breakout box, before signals enter the dilution refrigerator. The 3 dB frequency of the BLP-1.9+ is 2.5 MHz. The attenuation is better than 20 dB above 3.4 MHz, and 40 dB above 4.7 MHz.

In addition, homemade low-pass (LP) filters are used after the shielding room interface box. The schematic of an LP filter is shown as an inset in Fig. A.7. It is a second-order RLC LP filter with $R = 100\ \Omega$, $L = 47$ mH, and $C = 10\ \mu$F. The designed bandwidth is about 200 Hz with $Q \approx 0.707$ (A 2nd-order Butterworth filter with maximally flat response). But due to the inherent resistance in the inductor (about $25 \sim 30\ \Omega$), the measured bandwidth is about $f_{3dB} = 160$ Hz. The frequency response of the LP filter is measured by an HP 41941A gain-phase analyzer with output channel impedance 50 Ω and test channel impedance 1 MΩ. The result is shown in Fig. A.7. The attenuation is better than 20 dB above 700 Hz, and 50 dB between 5 KHz and 10 MHz. Due to the parasite capacitance between the input

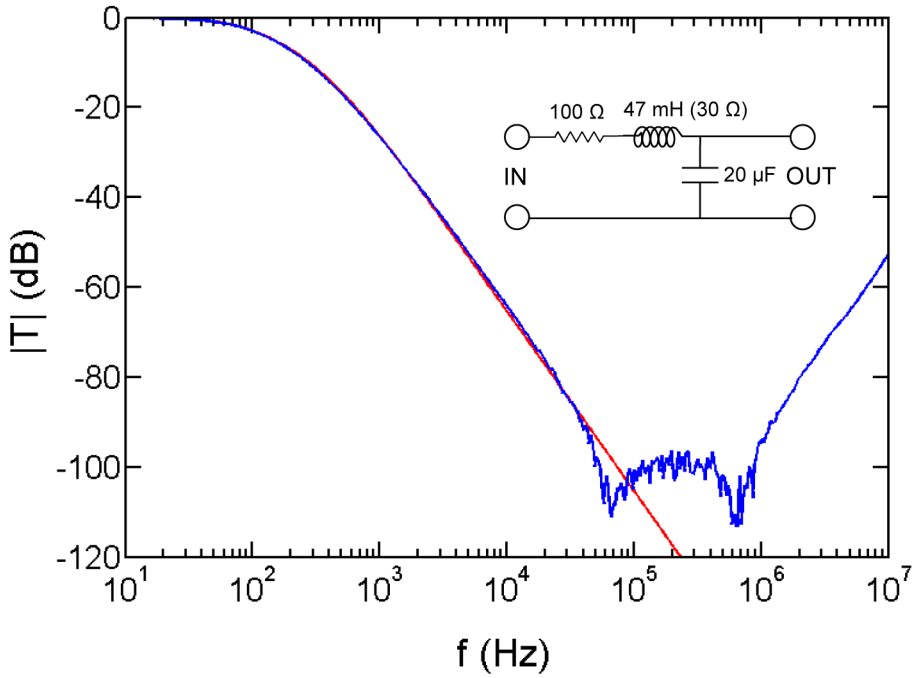

Figure A.7: The frequency response of a homemade 2nd-order low-pass (LP) filter with $f_{3dB} = 160$ Hz. The blue line is the result, measured by an HP 4194A gain-phase analyzer with output channel impedance 50 Ω and test channel impedance 1 MΩ. The red line is the simulation result. The inset shows the schematic of the LP filter. Due to the parasite capacitance between the input and the output, the attenuation decreases above 1 MHz.

and the output, the attenuation decreases above 1 MHz. The LP filter increases the output impedance of each SMU to about 130 Ω. But the load of each channel is much larger than that, so it is not a issue here.

After combining all these filters, the heating from RF noise can be effectively reduced.

A.4.3 Transimpedance amplifier (pre-amp)

The transimpedance amplifier (current amplifier) is an essential component in the AC measurement system. They are well discussed by C. Julian Chen in Ref. [73]. Because of the different center frequency and bandwidth requirements between the STM application and the AC lock-in technique, the noise from each component is reevaluated here.

In AC measurement, V_{ac} is a 37 Hz, 0.1 mV sine wave, and the feedback resistor of the amplifier is 100 MΩ. The measurement resolution is limited by two components: the 100 MΩ feedback resistor and the input capacitance at -IN node of the op-amp from the coaxial cable.

The noise from V_{ac}, V_{dc} are measured both less than 1×10^{-15} V^2/Hz. In the AC+DC adder box, the thermal noise of R_a is $4k_BTR_a = 3.3 \times 10^{-17}$ V^2/Hz, where k_B is Boltzmann's constant, $T = 300$ K is the room temperature, and $R_a = 2$ kΩ is the resistor value. An OPA602 from TI is used as the op-amp. The voltage noise at -IN node of the op-amp is less than 5.3×10^{-16} V^2/Hz. (All noise data of the op-amp is from the OPA602 datasheet.) As shown in Fig. A.8, the

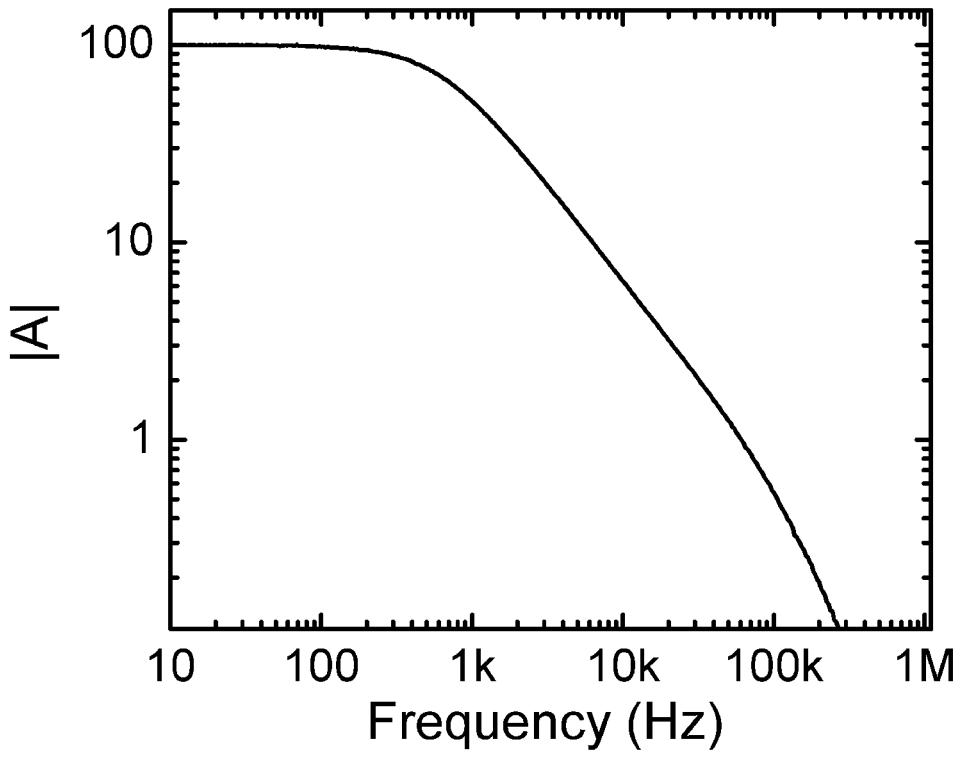

Figure A.8: The frequency response of the transimpedance amplifier with $f_{3dB} = 578$ Hz, while the input of the amplifier is first connected to a 1 MΩ resistor in series, then connected to the test channel of the HP 4194A gain-phase analyzer. Here an OPA602 from TI is used as the op-amp.

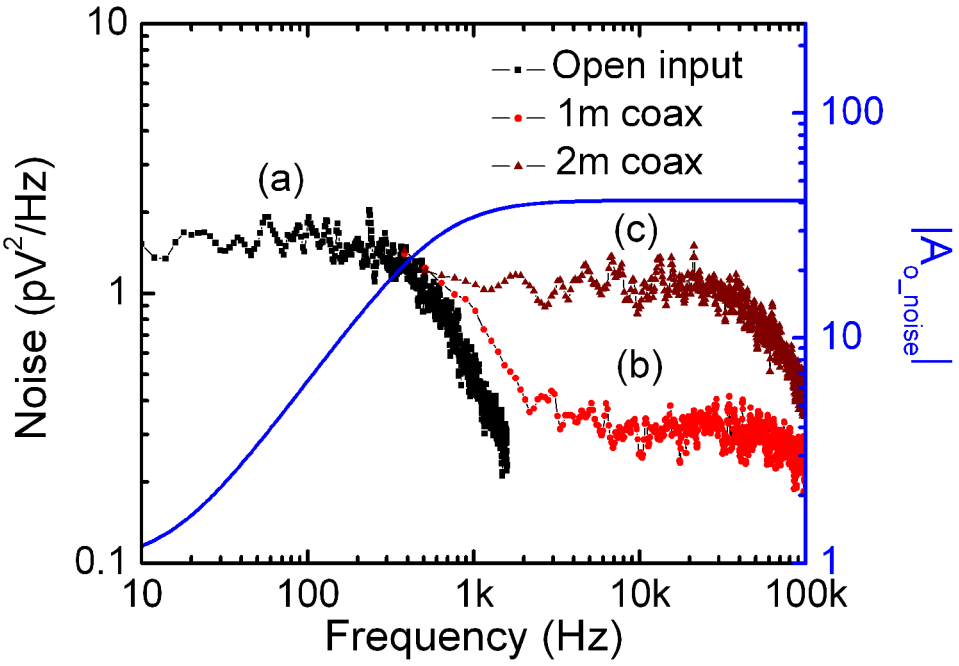

Figure A.9: Noise of the transimpedance amplifier with three different input configurations: (a) the input of the amplifier is open, (b) 1 m coaxial cable is connected to the input of the amplifier, and (c) 2 m coaxial cable is connected to the the input of the amplifier. Data are measured by an HP 35665A dynamic signal analyzer. Noise gain A_{o_noise} due to the 100 pf input capacitance is also shown in the graph.

amplifier bandwidth is measured to be $f_{3dB} = 578$ Hz, and the equivalent noise bandwidth is about $1.57 \times f_{3dB} = 907$ Hz.[74] So the total voltage noise is less than $\sqrt{(10 + 10 + 0.33 + 5.3) \times 10^{-16} \times 907} = 1.6$ μV, which is applied on the device under test (DUT), and also appears at the amplifier output. Since V_{ac} is 0.1 mV, and much larger than 1.6 μV, the voltage noise is not a limiting factor.

The current noise from the feedback resistor is one of the limiting factors. The current noise is $\sqrt{4k_BT/R_{fb}} = 12.9 \times 10^{-15}$ A/$\sqrt{\text{Hz}}$ and the corresponding voltage

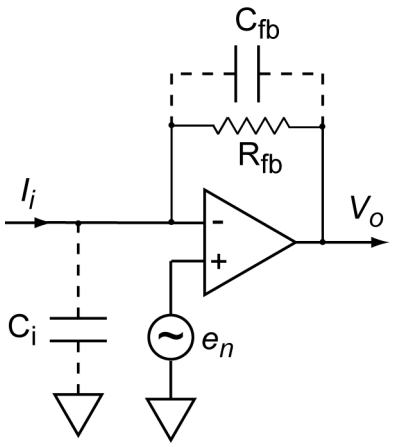

Figure A.10: Output noise due to the input capacitance. The smaller the input impedance, the larger the noise at the output. So the input capacitance generates a large high-frequency noise, and the noise gain is only limited by C_{in}/C_{fb}. Since $R_{fb} \sim 100$ MΩ and $C_i \sim 100$ pF, the 3dB noise gain is at 16 Hz.

noise is $4k_BTR_{fb} = 1.66 \times 10^{-12}$ V^2/Hz. The noise spectrum of the amplifier is measured by an HP 35665A dynamic signal analyzer, while the input of the amplifier is open. The result is shown as (a) in Fig. A.9. The noise at $f < 8$ Hz is artificial from the measurement system, since it exists even the HP 35665A input is shorted by a 50 Ω terminator, and is not shown here. The voltage noise is indeed about 1.6×10^{-12} V^2/Hz in the amplifier bandwidth. The roll-off of the noise is due to the parasite capacitance C_{fb} in the amplifier. At 37 Hz, the noise is about 1.6×10^{-12} V^2/Hz, which is corresponding to current noise $\sqrt{1.6 \times 10^{-12}}/R_{fb} = 12.6 \times 10^{-15}$ A/$\sqrt{\text{Hz}}$.

The input capacitance C_i at -IN node of the op-amp is another limiting factor. Because we use a coaxial cable to connect the DUT to the amplifier, and the length

of the coaxial cable is about 1 m, the input capacitance C_i is about 100 pF. The acoustic noise in the shielding room deforms the coaxial cable, and changes the capacitance. This is called microphone effect.[73] The current is

$$I = \frac{dQ}{dt} = C\frac{dV}{dt} + V\frac{dC}{dt}. \tag{A.1}$$

Since the voltage on the coaxial cable is almost zero, this is not an important effect. But the cables should avoid large motion when the measurement is in progress. I usually tape the cables on the rigid frame.

The input capacitance at the -IN node has another important effect. At the input of the op-amp, there is always a small voltage noise. As shown in Fig. A.10, this small voltage noise will be amplified by the op-amp. Here the input noise is modeled as an AC source e_n at the +IN node of the amplifier. The output noise voltage is

$$V_{o\_noise} = e_n(1 + \frac{Z_{fb}}{Z_{in}}), \tag{A.2}$$

and the noise gain is

$$A_{o\_noise} = 1 + \frac{Z_{fb}}{Z_{in}}, \tag{A.3}$$

where $Z_{fb} = R_{fb}/(1 + j\omega C_{fb}R_{fb})$ and $Z_{in} = 1/(j\omega C_{in})$, and ω is the angular frequency. First, the input noise e_n should be minimized. In our previous measurement system, Burr-Brown ISO100 optically-coupled linear isolation amplifier was used to to drive the +IN node. Since the voltage noise from ISO100 is quite large, the measurement resolution was limited at ~ 1 pA. Now the +IN node is grounded, so the input noise is less than 5.3×10^{-16} V^2/Hz, the intrinsic noise of the OPA602. Because the bandwidth of the amplifier is $f_{3dB} = 578$ Hz, C_{fb} is estimated to be

$1/(2\pi f_{3dB} R_{fb}) = 2.5$ pF. C_{fb} about 0.5 pF is common,[73] and the larger C_{fb} helps reduce the noise, but it also reduces the bandwidth. Fortunately $f_{3dB} = 578$ Hz is still sufficient for the AC lock-in measurement. One important difference between the STM application and the AC lock-in technique is the center frequency and bandwidth requirements. In the AC lock-in technique, the center frequency is normally less than 100 Hz, and the bandwidth of the low-pass filter in the lock-in amplifier is about 1 Hz. These two factors almost eliminate the deleterious effect of the noise due to the input capacitance. In our system, 37 Hz sine wave is used, so $V_{o\_noise} = 2.58 e_n = 2.58 \times 23 = 59.3$ nV/$\sqrt{\text{Hz}}$ or 3.52×10^{-15} V^2/Hz, which is much less than the thermal noise (1.6×10^{-12} V^2/Hz) from the feedback resistor. However, the noise can overload the lock-in amplifier, when high sensitivity is used. As shown in Fig. A.9, the noise gain A_{o_noise} increases with an increasing frequency. So the input capacitance generates a large high-frequency noise, and the noise gain is only limited by C_i/C_{fb}. For example, at 10 kHz, if a 1 m coaxial cable is connected to the input of the amplifier, the noise is about 0.3 pV^2/Hz; if it is a 2 m coaxial cable, the noise is about 1.1 pV^2/Hz, as shown in Fig. A.9. Thus the input capacitance effectively increases the noise at high frequency, and makes the output of the amplifier noisy. The band-pass filter in the lock-in PAR 124A helps reduce the overloading problem.

Here an FET-input OPA602 is used, because FET-input op-amps have smaller bias current and lower current noise than bipolar-input op-amps. For OPA602, the current noise is about 0.6×10^{-15} A/$\sqrt{\text{Hz}}$, which is also much smaller than the thermal noise (12.6×10^{-15} A/$\sqrt{\text{Hz}}$) from the feedback resistor.

After the careful design of the AC measurement system, the system noise is only dominated by the thermal noise of the feedback resistor. The measurement resolution is also related to the bandwidth of the LP filter in the lock-in amplifier. The smaller the bandwidth is, the higher the measurement resolution is, but the measurement time increases. A good trade-off is that the time constant (TC) of the LP filter is set at 100 ms, and the equivalent noise bandwidth is $1/(8TC) = 1.25$ Hz for the 12 dB/octave two section filter.[75] So the current resolution is

$$I_{rms} = 12.6 \times 10^{-15} A/\sqrt{\text{Hz}} \times \sqrt{1.25 \text{ Hz}} = 14.1 \text{ fA}. \tag{A.4}$$

A measurement example is shown in Fig. A.11. The current measurement resolution is about 25 fA, which is consistent with $I_{peak} = \sqrt{2} \times 14.1$ fA $= 20$ fA.

To further increase the current resolution, a 1 GΩ resistor can be used as the feedback resistor. I found that it does help improve the current resolution, which also means that the thermal noise of the feedback resistor is the main source of the system noise, but the system response becomes slow. The 100 MΩ feedback resistor is used in most measurements here.

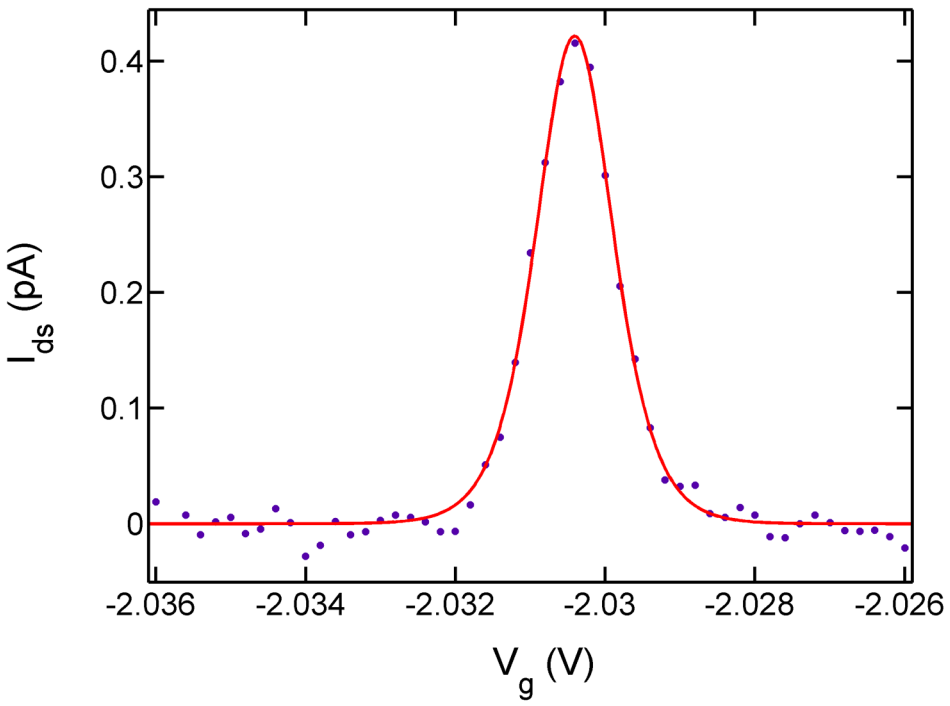

Figure A.11: A Coulomb oscillation peak measured by the AC lock-in technique, shown as the dots. The measurement resolution is about 25 fA. The electron temperature is about 400 mK by fitting the Coulomb blockade peaks with the equation, $I/I_{max} = cosh^{-2}(e\alpha(V_g - V_{g0})/2k_BT)$, where I is the source-drain current, I_{max} is the maximum current at the peak, V_{g0} is the gate voltage at resonance, k_B is the Boltzmann constant, T is the electron temperature, and the ratio of the gate capacitance to the total capacitance $\alpha = C_g/C_\Sigma$.[25]

Appendix B

AC measurement system

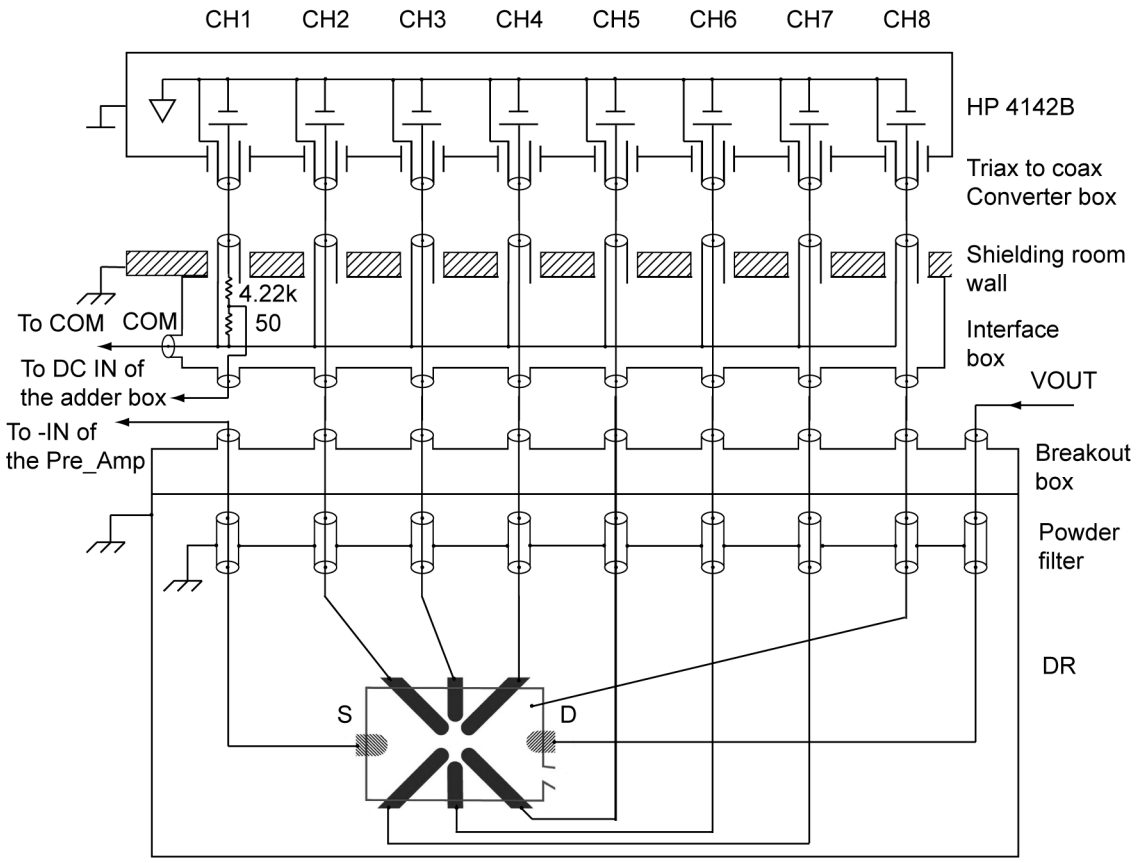

Figure B.1: Detailed wiring of the AC measurement system part 1/2

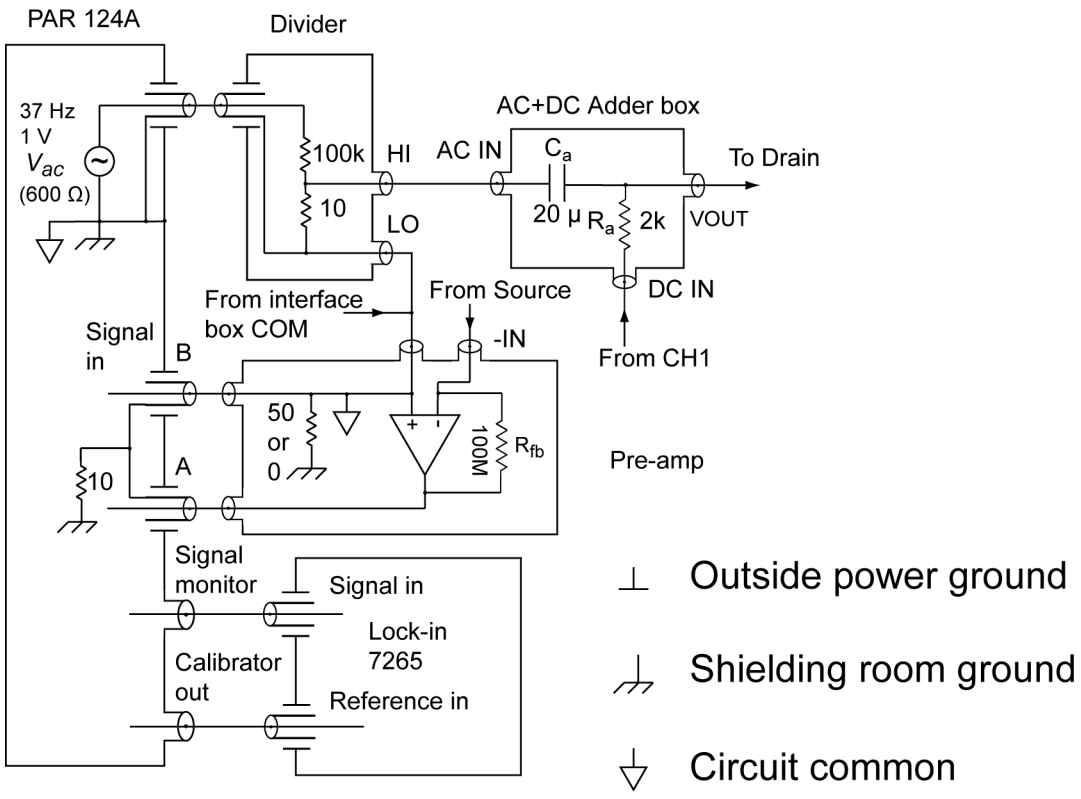

Figure B.2: Detailed wiring of the AC measurement system part 2/2

Appendix C

Recipe for silicon MOS-SETs

Summary: Thermal oxide + HDPECVD oxide and 6 side gates with trenches under top gate.

Name:_____ **Sample #:**_____ **Run Name:** _____

Step 1: Top Gate Trenches (isolation)

I. Pirahanna Clean: 4 H_2SO_4: 1 H_20_2

- ☐ H_20_2, 15ml in the beaker, add H_2SO_4 60ml slowly with stirring. Put the samples in and continue heating at 100°C (not higher than 120 °C) for 15 min. Rinse with DI thoroughly.

II. Sample Preparation Time: _____

- ☐ Solder In dots for secondary electron emission collection and check resistance
- ☐ PMMA A8 5 krpm 60 sec (expected 8000 Å)
- ☐ Pre-bake: 180°C, 3min. w/vac. (Actual Temp = _____°C, Time = _____s)

III. E-Beam Writing Time: _____

1. 4pt focusing :

- ☐ Check resistance between spring and other In dot (0.5MΩ typical)
- ☐ Drop one drop gold solution at each corner of the sample
- ☐ Global correction : _____ degree
- ☐ Focus at each point and run 4pt2.exe to get the fitting plane function

2. E-beam Writing : Runfile name : _____ *.rf6 :* _____ *.rf6 :* _____ *.rf6 :* _____ *.rf6 :* _____ *.rf6 (marker and alpha step)*

Line dose : 5nC/cm, c-t-c = 43 Å, current ∼ 12 pA
Area dose : 500uC/cm^2, c-t-c = 101Å, line spacing = 101 Å, current ∼ 12 pA
Actual dose : _____

- ☐ Pinhole current : _____ pA for CC = ____

  _____ pA for CC = ____ before writing

☐ Pinhole current : _____ pA for CC = ____

_____ pA for CC = ____ after writing

IV. Development Time: _____

☐ Develop: MIBK: IPA(1:1) 80 seconds

☐ Overlap Rinse: IPA 80 sec

☐ Overlap Rinse: DI water 80 sec ;

☐ Blow dry

V. Dry Etch Time: _____

☐ Process "SI1DRH" 1 minute target 0.3um deep

VI. Clean:

☐ Using ACE 30min (Actual dip time: _____) + IPA +DI

☐ DI Rinse.

Step 2: Surface Gates

I. Sample Preparation Time: _____

☐ Solder In dots for secondary electron emission collection and check resistance

☐ MMA EL11 5 krpm 60 sec (expected 3800-4200Å)

☐ Pre-bake: 150°C, 60 sec. w/vac. (Actual Temp = _____°C, Time = _____s)

☐ PMMA A4 5 krpm 60 sec (expected 1800 Å)

☐ Pre-bake: 180°C, 60 sec. w/vac. (Actual Temp = _____°C, Time = _____s)

II. E-Beam Writing Time: _____

1. 4pt focusing :

☐ Check resistance between spring and other In dot (0.5MΩ typical)

☐ Drop one drop gold solution at each corner of the sample

☐ Global correction : _____ degree

☐ Focus at each point and run 4pt2.exe to get the fitting plane function

2. E-beam Writing : Runfile name : _____ *.rf6 :* _____ *.rf6 :* _____ *.rf6 :* _____ *.rf6 :* _____ *.rf6 (marker and alpha step)*

Small area: (1000X) Line dose : 2.5nC/cm, c-t-c = 43 Å, current ∼12 pA
Area dose : 250uC/cm^2, c-t-c = 101Å, line spacing = 101 Å, current ∼12 pA
Actual dose : _____

☐ Pinhole current : _____ pA for CC = ____ before writing

☐ Pinhole current : _____ pA for CC = ____ after writing

Lead area:
(100X) Area dose : 300uC/cm^2, c-t-c = 800Å, line spacing=800Å, current $\sim$ 0.8 nA
(30X) Area dose: 300uC/cm^2, c-t-c=3000 Å, line spacing=3000Å, current $\sim$ 7nA

☐ Pinhole current : _____ pA for CC = 8

_____ pA for CC = 5 before writing

☐ Pinhole current : _____ pA for CC = 8

_____ pA for CC = 5 after writing

III. Development Time: _____

☐ Develop: MIBK: IPA(1:1) 80 seconds (1:20)

☐ Overlap Rinse: IPA 20 sec (1:40)

☐ Develop: MIBK: IPA(1:2) 90 seconds (3:10)

☐ Overlap Rinse: IPA 80 sec (4:30)

☐ Overlap Rinse: DI water 30 sec (5:00)

☐ Blow dry

V. Evaporation Time: _____

☐ Evaporate: Process # (50) Al_____(1000) Å

VI. Clean: Using heated ACE (Actual dip time: _____) + IPA +DI ...
Alpha step

Step 3: HDPECVD SiO2 deposition **Time:** _____

☐ Recipe Name_____

Pressure_____(1mTorr), Temperature _____(300 °C),
SiH_4 flow rate_____(4sccm) N_2O flow rate _____(20sccm)
Deposition rate_____ (312 Å /m), time _____
Thickness 4000 Å, actual_____

Step 4: Top Gate

I. Sample Preparation Time: _____

☐ NR-7 1500PY 4000 RPM

☐ Pre-bake: 120°C, 60 sec. w/vac. (Actual Temp = _____°C, Time = _____s)

II. Photolithography Time: _____

☐ UV exposure using 5x stepper with vacuum.

Mask: HB/QPC/Cross"Top Gate ***4of3***"
exposure time **4.2 sec** (check)
focal number **0**
XL=XR=10, YF=10, YR=10

☐ Post-bake: 110°C, 60 sec. w/vac. (Actual Temp = _____°C, Time = _____s)

☐ Develop in **RD-6** 3 sec (exactly)

☐ DI water rinse 1-2 min

III. Evaporation Time: _____

☐ Evaporate: Process # ( ) (Al with 2% Si)_____(3600) Å

IV. Liftoff: RR2 Hot plate heated to 100C for 10 min Remove metal with DI spray, Soak another 5 min, Rinse 3 min with DI (hall bars may take up to 30 min to liftoff properly)

Step 5: Bonding Pad Via Holes Date:_____

Via to ion implantation:
I. Sample Preparation Time: _____

☐ Spin HMDS adhesion promoter at 3 krpm, 60 sec

☐ Spin positive PR **OiR 908-35** at 3 krpm, 60 sec (4μm thick)

☐ Bake wafer at 90°C, 3 min

☐ UV exposure using 5x stepper with vacuum.

Mask No: 2 (2865030A00)
exposure time **1.8 sec** (check) (normal 1.6 sec, overexpose for thick edge)
focal number **0**

☐ Develop in **OPD 4262** 90 sec (above 1 Hz stirring)

☐ DI water rinse 3 min

☐ Alpha step: thickness:_____

II. Dry Etch Time: _____

☐ SIO2PT 12min, target 390nm, (32nm/min),

power_____(175W), Pressure _____(100mTorr)
Gas_____(CHF_3 18sccm, O_2 2sccm)

III. Wet Etch for ohmic

☐ Wet Etch BOE 6:1, 90sec@25°C (<2 min), Make sure Air bubbles are gone from surface at features (∼100nm/min).

☐ Verify by probe station.

IV. Liftoff: Using ACE 60min (Actual dip time: _____ hr) + IPA + DI

☐ Alpha step: thickness:_____

Via to side gates:
I. Sample Preparation Time: _____

☐ Spin HMDS adhesion promoter at 3 krpm, 60 sec

☐ Spin positive PR **OiR 908-35** at 3 krpm, 60 sec (4um thick)

☐ Bake wafer at 90°C, 3 min

☐ UV exposure using 5x stepper with vacuum.
Mask No: 2 (2865030A00)
exposure time **1.8 sec** (check)) (normal 0.37 sec, overexpose for thick edge)
focal number **0**

☐ Develop in **OPD 4262** 90 sec (above 1 Hz stirring)

☐ DI water rinse 3 min

☐ Alpha step: thickness:_____

II. Dry Etch Time: _____

☐ SIO2PT 15min 05sec, target 450 nm and over etch 1 min, (32nm/min),

power_____(175W), Pressure _____(100mTorr)
Gas_____(CHF_3 18sccm, O_2 2sccm)

☐ Verify by probe station

IV. Liftoff: Using ACE 60min (Actual dip time: _____ hr) + IPA + DI

☐ Alpha step: thickness:_____

Step 6: Annealing Date:_____

Forming gas (H_2+N_2) annealing, 420°C for 30 minutes.
Switch to forming gas at 250°C
Switch back to N_2 after devices cool down under 300°C.

Bibliography

[1] P. Benioff, J. Stat. Phys. **22**, 563 (1980).

[2] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).

[3] D. Deutsch, Proc. Roy. Soc. London A **400**, 97 (1985).

[4] P. W. Shor, in *Proc. 35th Annu. Symp. Foundations of Computer Science*, edited by S. Goldwasser (IEEE Computer Society Press, 1994), p. 124.

[5] L. K. Grover, *Proc. 28th Annu. ACM Symp. on the Theory of Computing* (ACM, 1996), p. 212.

[6] *A quantum information science and technology roadmap*, by Los Alamos National Security, LLC, (2004), Website, `http://qist.lanl.gov/qcomp_map.shtml`.

[7] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).

[8] V. N. Golovach and D. Loss, Semiconductor Science and Technology **17**, 355 (2002).

[9] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Science **309**, 2180 (2005).

[10] F. H. L. Koppens, C. Buizert, K. J. Tielrooij, I. T. Vink, K. C. Nowack, T. Meunier, L. P. Kouwenhoven, and L. M. K. Vandersypen, Nature (London) **442**, 766 (2006).

[11] C. B. Simmons, M. Thalakulam, N. Shaji, L. J. Klein, H. Qin, R. H. Blick, D. E. Savage, M. G. Lagally, S. N. Coppersmith, and M. A. Eriksson, Appl. Phys. Lett. **91**, 213103 (2007).

[12] A. Fujiwara, H. Inokawa, K. Yamazaki, H. Namatsu, Y. Takahashi, N. M. Zimmerman, and S. B. Martin, Appl. Phys. Lett. **88**, 053121 (2006).

[13] G. M. Jones, B. H. Hu, C. H. Yang, M. J. Yang, R. Hajdaj, and G. Hehein, Appl. Phys. Lett. **89**, 073106 (2006).

[14] S. J. Angus, A. J. Ferguson, A. S. Dzurak, and R. G. Clark, Nano Lett. **7,7**, 2051 (2007).

[15] C. Tahan and R. Joynt, Phys. Rev. B **71**, 075315 (2005).

[16] T. Meunier, I. T. Vink, L. H. W. van Beveren, F. H. L. Koppens, H. P. Tranitz, W. Wegscheider, L. P. Kouwenhoven, and L. M. K. Vandersypen, Phys. Rev. B **74**, 195303 (2006).

[17] M. A. Kastner, Rev. Mod. Phys. **64**, 849 (1992).

[18] H. W. Liu, T. Fujisawa, Y. Ono, H. Inokawa, A. Fujiwara, K. Takashina, and Y. Hirayama, Phys. Rev. B **77**, 073310 (2008).

[19] F. Boeuf, X. Jehl, M. Sanquer, and T. Skotnicki, IEEE Trans. Nanotechnol. **2**, 144 (2003).

[20] E. P. Nordberg, G. A. T. Eyck, H. L. Stalford, R. P. Muller, R. W. Young, K. Eng, L. A. Tracy, K. D. Childs, J. R. Wendt, R. K. Grubbs, et al. (2009), `http://arxiv.org/abs/0906.3748`.

[21] *Coulomb blockade*, from Wikipedia, the free encyclopedia, Website, `http://en.wikipedia.org/wiki/Coulomb_blockade`.

[22] L. P. Kouwenhoven, D. G. Austing, and S. Tarucha, Rep. Prog. Phys. **64**, 701 (2001).

[23] L. P. Kouwenhoven, C. M. Marcus, P. L. McEuen, S. Tarucha, R. M.Westervelt, and N. S. Wingreen, *Electron transport in quantum dots*, in *Mesoscopic Electron Transport*, NATO Science Series, edited by L. L. Sohn, L. P. Kouwenhoven and G. Schon (Kluwer, Dordrecht, 1997).

[24] K. Likharev, Proceedings of the IEEE **87**, 606 (1999), ISSN 0018-9219.

[25] C. W. J. Beenakker, Phys. Rev. B **44**, 1646 (1991).

[26] E. B. Foxman, P. L. McEuen, U. Meirav, N. S. Wingreen, Y. Meir, P. A. Belk, N. R. Belk, M. A. Kastner, and S. J. Wind, Phys. Rev. B **47**, 10020 (1993).

[27] R. Akis, J. P. Bird, D. Vasileska, D. K. Ferry, A. P. S. de Moura, and Y.-C. Lai, *On the influence of resonant states on ballistic transport in open quantum dots: Spectroscopy and tunneling in the presence of multiple conducting channels*, in *Electron transport in quantum dots*, edited by J. P. Bird (Kluwer Academic Publishers, 2003).

[28] H. Yamamoto and K. Miyamoto, Phys. Stat. Sol. (b) **200**, 89 (1997).

[29] L. I. Schiff, *Quantum mechanics* (McGraw-Hill, Inc., 1971), p. 101, 3rd ed.

[30] L. I. Schiff, *Quantum mechanics* (McGraw-Hill, Inc., 1971), p. 39, 3rd ed.

[31] S. M. Sze, *Modern Semiconductor Device Physics* (Wiley-Interscience, 1997), p. 259, 1st ed.

[32] R. Landauer, IBM J. Res. Develop. **1**, 233 (1957).

[33] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen, Reviews of Modern Physics **79**, 1217 (pages 49) (2007).

[34] B. Schuh, J. Phys. A **18**, 803 (1985).

[35] O. Dippel, P. Schmelcher, and L. S. Cederbaum, Phys. Rev. A **49**, 4415 (1994).

[36] W. G. van der Wiel, S. D. Franceschi, J. M. Elzerman, T. Fujisawa, S. Tarucha, and L. P. Kouwenhoven, Rev. Mod. Phys. **75**, 1 (2003).

[37] F. Hofmann, T. Heinzel, D. Wharam, J. Kotthaus, G. Böhm, W. Klein, G. Tränkle, and G. Weimann, Phys. Rev. B **51**, 13872 (1995).

[38] C. Livermore, C. H. Crouch, R. M. Westervelt, K. L. Campman, and A. C. Gossard, Science **274**, 1332 (1996).

[39] A. Ferreira da Silva, Phys. Rev. B **37**, 4799 (1988).

[40] G. M. Jones, *Quantum transport in nanoscale semiconductors*, Ph.D. thesis (2006).

[41] S. A. Campbell, *The Science and Engineering of Microelectronic Fabrication* (Oxford University Press, Inc., 2001), p. 110, 2nd ed.

[42] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices* (Cambridge Univesity Press, 1998), pp. 58–90.

[43] R. A. Smith, *Semiconductors* (Cambridge Univesity Press, 1979), 2nd ed.

[44] F. Stern, Phys. Rev. B **5**, 4891 (1972).

[45] B. E. Deal, IEEE Trans. Electron Devices. **27**, 606 (1980).

[46] L. M. Terman, Solid-State Electron. **5**, 285 (1962).

[47] J. R. Brews, J. Appl. Phys. **45**, 1276 (1974).

[48] D. K. Schroder, *Semiconductor Material and Device Characterization* (John Wiley & Sons, Inc., New Jersey, 2006), pp. 347–350, 3rd ed.

[49] E. H. Nicollian and A. Goetzberger, Bell Syst. Tech. J. **46**, 1055 (1967).

[50] E. H. Hall, Amer. J. Math. **2**, 287 (1879).

[51] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

[52] C. Kittel, *Introduction to Solid State Physics* (John Wiley & Sons, Inc., New York, 1986), p. 554, 6th ed.

[53] F. Simmel, D. Abusch-Magder, D. A. Wharam, M. A. Kastner, and J. P. Kotthaus, Phys. Rev. B **59**, R10441 (1999).

[54] H. Ishikuro and T. Hiramoto, Appl. Phys. Lett. **71**, 3691 (1997).

[55] H. Sellier, G. P. Lansbergen, J. Caro, S. Rogge, N. Collaert, I. Ferain, M. Jurczak, and S. Biesemans, Phys. Rev. Lett. **97**, 206805 (2006).

[56] E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology* (John Wiley & Sons, Inc., 1982), p. 237, 1st ed.

[57] L. P. Rokhinson, L. J. Guo, S. Y. Chou, and D. C. Tsui, Phys. Rev. B **63**, 035321 (2001).

[58] S. Goswami, K. A. Slinker, M. Friesen, L. M. McGuire, J. L. Truitt, C. Tahan, L. J. Klein, J. O. Chu, P. M. Mooney, D. W. van der Weide, et al., Nature Physics **3**, 41 (2007).

[59] K. Takashina, Y. Ono, A. Fujiwara, Y. Takahashi, and Y. Hirayama, Phys. Rev. Lett. **96**, 236801 (2006).

[60] M. Ciorga, A. S. Sachrajda, P. Hawrylaka, C. Gould, P. Zawadzki, S. Jullian, Y. Feng, and Z. Wasilewski, Phys. Rev. B **61**, 16315 (2000).

[61] S. A. Vitkalov, M. P. Sarachik, and T. M. Klapwijk, Phys. Rev. B **64**, 073101 (2001).

[62] A. A. Shashkin, S. V. Kravchenko, V. T. Dolgopolov, and T. M. Klapwijk, Phys. Rev. Lett. **87**, 086801 (2001).

[63] I. Appelbaum, B. Huang, and D. J. Monsma, Nature (London) **447**, 295 (2007).

[64] A. S. Sachrajda, P. Hawrylaka, M. Ciorga, C. Gould, and P. Zawadzki, Physica E **10**, 493 (2001).

[65] P. A. Maksym, Physica B **249-251**, 233 (1998).

[66] T. Schmidt, R. J. Haug, V. I. Fal'ko, K. v. Klitzing, A. Forster, and H. Luth, Europhys. Lett. **36**, 61 (1996).

[67] M. T. Bjork, C. Thelander, A. E. Hansen, L. E. Jensen, M. W. Larsson, L. R. Wallenberg, and L. Samuelson, Nano Lett. **4(9)**, 1621 (2004).

[68] S. Feng, P. A. Lee, and A. D. Stone, Phys. Rev. Lett. **56**, 1960 (1986).

[69] B. R. Johnson, J. Chem. Phys. **67**, 4086 (1977).

[70] M. Khoury, M. J. Rack, A. Gunther, and D. K. Ferry, Appl. Phys. Lett. **74**, 1576 (1999).

[71] K. Bladh, D. Gunnarsson, E. Hurfeld, S. Devi, C. Kristoffersson, B. Smalander, S. Pehrson, T. Claeson, P. Delsing, and M. Taslakov, Rev. Sci. Instrum. **74**, 1323 (2003).

[72] A. Fukushima, A. Sato, A. Iwasa, Y. Nakamura, T. Komatsuzaki, and Y. Sakamoto, IEEE Trans. Instrum. Meas. **46**, 289 (1997).

[73] C. J. Chen, *Introduction to Scanning Tunneling Microscopy* (Oxford University Press, Inc., 1993), p. 251, 1st ed.

[74] P. Horowitz and W. Hill, *The Art of Electronics* (Cambridge University Press, 1989), p. 453, 2nd ed.

[75] *Model 124a lock-in amplifier instruction manual*, (EG&G Instruments Corporation, 1983-1992), p31.