

DEVELOPMENT OF HIGH-BRIGHTNESS ULTRAFAST ELECTRON
MICROSCOPE FOR STUDYING NANOSCALE DYNAMICS ASSOCIATED
WITH STRONGLY CORRELATED MATERIALS

By

Zhensheng Tao

A DISSERTATION

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

Physics — Doctor of Philosophy

2014

ABSTRACT

DEVELOPMENT OF HIGH-BRIGHTNESS ULTRAFAST ELECTRON MICROSCOPE FOR STUDYING NANOSCALE DYNAMICS ASSOCIATED WITH STRONGLY CORRELATED MATERIALS

By

Zhensheng Tao

Strongly correlated-electron materials are a class of materials that exhibit numerous intriguing emergent phenomena, including metal-to-insulator transition, colossal magnetoresistance, high-temperature superconductivity, etc. These phenomena are beyond the reach of the conventional solid state physics, which is based on the band theory. Instead, strong electron-electron correlations are found to play important roles, which leads to complicated interplay between different degrees of freedoms (charge, lattice, spins ...). In this thesis, ultrafast electron diffraction (UED) is used to investigate the photo-induced ultrafast structural dynamics of strongly correlated materials, among which VO_2 is taken as an exemplar system, one that reveals the fundamental physics behind photo-induced phase transitions, electron-electron correlation on nanometer scales, and the electron-phonon coupling in this exotic class of materials. The phenomena presented here are expected to have more general significance as they may reflect the physics to which other strongly correlated materials also conform.

In polycrystalline VO_2 thin films, the structural changes resulting from photoexcitation with femtosecond laser pulses with different wavelengths are observed to lead to non-thermal phase transitions, which require less energy compared to the phase transitions induced by thermal excitation. The details of the structural change are extracted from the UED results revealing stepwise atomic movements after photoexcitation, which suggests the phase transition starts with a dilation of the correlated d electrons. On the other hand, the structural phase transition is found to be decoupled from the metal-to-insulator transition when the sample dimension is reduced to the

sub-micrometer scale, which is attributed to the interface charge doping effects from different substrates. A new phase (M_3 , monoclinic metallic phase) is distinguished, which has not been discussed by the existing theoretical investigations. The reduction of the optical spectral weight and the anisotropic phononic response is revealed by the UED measurements in the noncooperative phase-transition region, suggesting intriguing interplay between the Mott-Hubbard correlated electrons and the Peierls lattice distortion.

The first-generation UED system is found to be limited by its brightness when high spatiotemporal resolution is required for the studies on nanometer-scale materials. The major constrain on the brightness is the space charge effect, which affects the phase space of the electron pulses. Using the projection-shadow-image technique, the space charge effects in the near-cathode-surface region are investigated. The results suggest a strong space-charge-led perturbation on the electrons' spatial and momentum distributions in the early stage of the short-pulse generation, and the performance with possible corrections in the drift region is discussed under the framework of a mean-field theory.

In laboratory, a radio-frequency (RF) cavity is implemented as a longitudinal focusing lens in the ultrafast electron microscope (UEM). The RF compression together with several magnetic lenses in the beam path, reshapes the electrons' phase space to achieve high brightness and high temporal resolution at the same time. High precision phase-lock between the electron pulses and the RF electric field timing is achieved by implementing a low-level RF phase-locked loop (PLL), an RF-amplifier station and a cavity PLL. The details of these RF systems are introduced, with characterization results presented. The RF-compression UEM is preliminarily characterized, which demonstrates the feasibility of using RF compression to generate high-brightness electron pulses. Future improvements and prospects for the system are also discussed.

To *my dear mother and father*

ACKNOWLEDGEMENTS

My six years at Michigan State University were a quite exciting, rewarding and enjoyable period time in my life. I can feel myself learning and achieving new things everyday, thanks to many people's helps on both my academic and personal lives. For the work presented in this thesis, I owe my utmost thanks to my advisor, Dr. Chong-yu Ruan, for his encouragements, guidance and (sometimes) tolerance. He gave me many great opportunities on scientific research. He also taught me how to 'do' science, how to think independently, how to be a mature researcher, etc. Without his help, I would not have had such a fruitful period of time.

I acknowledge all the group members of our UED lab who have contributed to our two ultrafast electron systems. First of all, many of my research results came from the UED instrument at the basement lab, where Dr. Ryan Murdick, Dr. Ramani K. Raman and Dr. Yoshie Murooka contributed most of their Ph.D time building it. Their well-organized work makes the interesting physics we discovered afterwards become possible. Tzong-Ru T. Han (Terry), Kiseok Chang and I am the second generation students of our lab. In Terry, I feel really fortunate to have a colleague, who is so careful and so skillful in research. He has a pair of magic hands that I can rely on in many of the experiments. In Kiseok, I want to acknowledge his utmost contributions to the ultrafast electron microscope in the ground-floor lab. Kiseok is a very hardworking, gentle and cheerful person. Without him, we would not be able to move that fast in the UEM project. My grad school experience would not have been half as enjoyable without them. I thank them for all the helps they gave me both inside the lab and outside of it. At the time of writing this thesis, Kiseok and Terry are also preparing for their defenses. I wish them success in their final defenses. In addition, I also want to thank the third-generation students. Fei Yuan has helped me on C++ programming, even though he switched to the nuclear physics. Faran Zhou works very hard and we worked together on some of the VO_2 and TaS_2 experiments. Nan Du is the new star in

the ground-floor lab, who workes very hard on characterizing the new system. Didi Luo is relatively young in research and starts to catch up. All of them are very promising students and I wish them good luck in their future career.

I acknowledge all my committee members for their guidance in these years. In Dr. Phil. Duxbury, we had collaborations on photoemission and VO_2 projects. In Dr. Martin Berz, his knowledge on beam dynamics are precious when developing the new system. Dr. Martin Crimp and Dr. Marcos Dantus are also my collaborations on the UEM projects. Thanks to Dr. John McGuire, for the helpful, friendly and inspiring discussions we had, which I appreciate very much. All my committee members are not only my collaborators, but also mentors, who taught me what is scientific research and how to be a good scientist, which I believe I will cherish throughout my career.

All of my projects were finished with collaborations. Without my collaborators, I would not imagine that I could have finished any of them. In addition to the group members and my committee members, I want to thank the following collaborators. Thanks to Dr. Junqiao Wu and Dr. Kevin Wang from UC Berkeley for providing us with the VO_2 nanobeam samples. Dr. Nelson Sepulveda and his students, David Torres and Tony Wang, from Material Science and Engineering Department of MSU helped us growing the VO_2 thin-film membrane samples. Dr. Mercouri G. Kanatzidis and Dr. Christos Malliakas for providing us with CDW materials CeTe_3 and TaS_2 (even though I could not introduce these materials in my thesis). Thanks to Dr. Jenni Portmann, Dr. He Zhang and Dr. Kyoko Makino for electron beam dynamics and photoemission simulations. Thanks to Dr. Subhendra D. Mahanti for his theoretical support on my material science researches. When developing the RF system for the UEM, I got helps from Dr. Valery Dolgashev from SLAC, Dr. Qiang Zhao and Dr. Dan Morris from NSCL at MSU. Our design of RF amplifiers were suggested by Dr. Ti Ruan from SOLEIL. Also, I want to thank Dr. Jun Ye from JILA, who I did not have a chance to discuss with personally. But his suggestion initiates all the work I

have done on the RF systems for UEM. The discussions with Dr. Norman Birge and Dr. Chih-Wei Lai were also very helpful for me to understand more on the academic life. I want to especially thank to Dr. David Tománek for his outstanding lectures on condensed matter physics.

The machine shop staff Tom Palazzolo, Tom Hudson, Jim Muns and Rob Bennett are the finest machinists I have even seen. They contribute tremendously in building various parts for the UEM system. Their suggestions and skills are something indispensable for this project. Special thanks to Barry Tigner, who is the electronic shop staff. I started the RF system design with very little knowledge in electronics. Barry is the one who taught me and helped me along the way. Every time when I have questions in electronics, I know who I should talk to. Reza Loloee and Baokang Bi helped me in great measure to utilize the various experimental methods in the basement. Xudong Fan at the MSU Center for Electron Microscopy taught me how to use the TEM for the VO₂ projects. Thanks to Cathy Cords, Debbie Barratt and (cheerful) Kim Crosslan for their fantastic administrative support throughout. Especially Cathy, who has been very patient to take care so many orders (from both Kiseok and me) for building the UEM system. From all these people, I can feel the friendship they shared with me, which will become my cherished memory in my life.

On the personal front, I have been fortunate to befriend some genuine individuals who have made my stay at MSU cherished. Chen Qiu, Iwan Setiewan, Kaijie Xu, Chen Wang, Xu Lu, Bai Nie, Zhen Zhu, Xukun Xiang, and many many more (Sorry, the list is just too long to finish...) - thanks for all the great times and wonderful memories. Also thanks to Salmon Rachel, Joe Williams, Thomas McCoy and Eric Gingrich for proofreading this long (and tedious) thesis.

Thanks to MSU Physics Department, NSF and DOE for supporting my research financially in these years.

Finally, I would not, and could not have undertaken this arduous journey without

the support and unconditional love from my parents. I love you.

TABLE OF CONTENTS

LIST OF TABLES	xii
LIST OF FIGURES	xiv
Chapter 1 Introduction and Methodology	1
1.1 Ultrafast Electron Diffraction	3
1.2 Diffraction Theory	5
1.2.1 Bragg's Law	6
1.2.2 Laue Diffraction Condition	7
1.2.3 Ewald Sphere Construction	10
1.2.4 Scattering from Isolated Atoms	12
1.2.5 Structure Factor $F(\theta)$ and Intensity of Bragg Reflections	17
1.3 Current Status of Ultrafast Electron Diffraction and Microscopy Techniques	18
1.3.1 Compact Geometry of Electron Gun	20
1.3.2 UEM with Single Electron Per Pulse	21
1.3.3 Dynamic TEM	22
1.3.4 UEM with MeV Energy	23
1.4 RF-Compression Ultrafast Electron Microscope at Michigan State University	24
Chapter 2 Correlated-Electron Materials	26
2.1 Band Theory of Solids	27
2.1.1 Bloch's Theorem	27
2.1.2 Nearly-Free Electron Model and Tight-Binding Approximation	28
2.1.3 Hartree-Fock Theory and Fermi Liquid Theory	33
2.1.4 Density Functional Theory	35
2.2 Strongly Correlated Electrons	37
2.3 Theoretical Treatment	44
2.3.1 DFT LDA+ U	45
2.3.2 Dynamical Mean Field Theory	47
2.4 Emergent Phenomena of Strongly Correlated Materials	50
2.4.1 Metal-to-insulator Transition	50
2.4.2 Colossal Magnetoresistance	51
2.4.3 High-temperature Superconductivity	53
Chapter 3 Photo-induced Phase Transitions in VO₂: an Investigation on Thin-film Membrane VO₂	5
3.1 General Properties of VO ₂ Metal-to-Insulator Transition and Structural Phase Transition	57
3.2 General Characteristics of VO ₂ Thin-film Membrane	59
3.3 Photo-induced Nonthermal Phase Transition	64
3.3.1 Suppression of Threshold Energy Density with Long Wavelength	

Pump	65
3.3.2 Penetration Depth of VO ₂ Thin Film	67
3.3.3 Temperature Dependence of Critical Energy Density	69
3.4 Atomic Pathway of Photo-induced Structural Phase Transition in VO ₂	75
3.4.1 Comparing Atomic Movements with Pumps of Different Wave-lengths	75
3.4.2 Atomic Pathways for VO ₂ Structural Phase Transition	76
3.5 Discussion and Summary	83

Chapter 4 Decoupling of Electronic and Structural Phase Transitions

in VO₂: an Investigation on Single-crystal Nanobeam VO₂	5
4.1 VO ₂ : Peierls Insulator or Mott Insulator?	86
4.2 Decoupling of Metal-to-Insulator Transition and Structural Phase Transition Observed in Steady-state Measurements	94
4.3 Decoupling of Metal-to-Insulator Transition and Structural Phase Transition in VO ₂ in Other Experiments	99
4.4 Ultrafast Electron Diffraction Studies of As-grown VO ₂ Nanobeams and VO ₂ Nanobeams on Gold Surfaces	100
4.4.1 As-grown VO ₂ Nanobeams on SiO ₂ Substrate	101
4.4.2 VO ₂ Nanobeams Suspended on Au Substrate	102
4.5 Summary	106

Chapter 5 RF-compression Ultrafast Electron Microscope Development:

Photoemission and Space Charge Effects	108
5.1 Different Methods of Characterizing the Electron-pulse Duration	111
5.2 Experimental Setup and Analytical Description of Projection-Shadow-Imaging Technique	116
5.2.1 Experimental Setup	116
5.2.2 Description of the Shadow Images	117
5.3 Measurement of Spatial and Temporal Evolution of Photo-electron Pulses from Gold Photocathode	121
5.4 Space Charge Effects and Modeling	123
5.4.1 Fractional Power-law Dependence in the Space-charge-led Pulse Lengthening	124
5.4.2 Density-Dependent Broadening of Initial Electron Velocity Distribution	125
5.4.3 Three-step Model of Photoemission	129
5.4.4 Space-Charge Limitation of Pulsed Photoemission Quantum Yield	134
5.4.5 The Effect of Multi-Photon Photoemission	136
5.5 Spacial and Temporal Resolution of Ultrafast Electron Diffraction and Imaging Systems: Analytical Gaussian Model Simulation	137
5.6 Conclusions	145

Chapter 6 Low-level Radio-frequency Synchronization System 148

6.1 Optical Frequency Comb and Comb Selection	150
---	-----

6.1.1 Timing Jitter and Temporal Resolution in UEM	152
6.2 An Introduction to Phase-Locked Loop	155
6.2.1 Static Phase Error	156
6.2.2 Linear Model of the PLL	158
6.2.3 PLL Transfer Function	159
6.3 Phase Noise and Static Phase Shift	161
6.3.1 Phase Noise	162
6.3.2 Shift of Static Phase Error	165
6.4 Loop Filter Design	168
6.5 Realization of PLL	172
6.6 Characterization of PLL	175
6.6.1 Spectrum Analysis	175
6.6.2 Phase-locked Loop Method	178
6.7 Temperature-Stabilized PLL	182
6.8 Conclusion	184
Chapter 7 RF Power Amplification and RF-cavity Phase-lock	185
7.1 RF Cavity Design	186
7.2 RF-amplifier Station	189
7.2.1 Power Requirement	189
7.2.2 RF Amplifiers	192
7.2.3 Structure of the RF-amplifier Station	193
7.2.4 Characterization of RF-amplifier Station	196
7.3 Cavity Phase for the Resonant Condition	198
7.3.1 Equivalent Circuit Model	198
7.3.2 Effects of Frequency Detuning	200
7.4 RF Cavity Phase-locked Loop	205
7.5 Conclusion	209
Chapter 8 Preliminary Characterizations of the RF-compression UEM and Conclusions	212
8.1 Preliminary Characterizations for RF-compression UEM	212
8.1.1 Sensitivity of Temporal Resolution to the Timing Jitter . . .	213
8.1.2 Characterizing the Pulse Compression	217
8.2 Conclusions	220
APPENDICES	225
Appendix A VO ₂ Structures and Diffraction Simulation	226
Appendix B Optical Constants of VO ₂ M ₁ Phase	235
Appendix C Supplementary Materials for VO ₂ Nanobeam Experiments . .	238
Appendix D Supplementary Materials for Projection-shadow Imaging Technique	247
REFERENCES	256

LIST OF TABLES

Table 3.1	List of orthogonal atomic coordinates of VO ₂ monoclinic (M ₁) structure. Crystal symmetry: P21/c; Unit cell definition: $a=5.743 \text{ \AA}$, $b=4.517 \text{ \AA}$, $c=5.375 \text{ \AA}$, $\beta=122.61^\circ$ [122]	63
Table 3.2	List of orthogonal atomic coordinates of VO ₂ rutile (R) structure. Crystal symmetry: P4 ₂ /mmn; Unit cell definition: $a=4.530 \text{ \AA}$, $c=2.869 \text{ \AA}$ [123] . .	64
Table 3.3	List of optical constants of VO ₂ M ₁ phase at different wavelengths	72
Table 5.1	Location of electron optical components in the pulse-compressed UEM column.	149
Table 6.1	List of RF components in the PLL with critical specifications. Some of the specifications, which are not available from manufacturers, are filled with ‘-’. For LNA, usually only the noise figure (NF) is available. ‘+0.5 dBc/Hz’ means additional 0.5 dBc/Hz phase noise is added to the signal after amplification.	184
Table A.1	List of orthogonal atomic coordinates and internal fractional coordinates of VO ₂ Rutile (R) structure	236
Table A.2	Crystal structure parameters of M ₁ -VO ₂ as given by Anderson [122] . .	237
Table A.3	List of the orthogonal atomic coordinates and the internal fractional coordinates of M ₁ -VO ₂ structure	238
Table A.4	List of the orthogonal atomic coordinates of the M ₁ -VO ₂ with the supercell definition. The ‘Unit cell shift’ shows the movements of the atom from its original atomic coordinates in Table A.3 to the new coordinate. For example, ‘($a + 1$, $b - 1$, 0)’ means the new atomic coordinates are derived by moving the atom along + a_{M1} axis by one unit cell and along the $-b_{M1}$ axis by one unit cell from its original coordinates	240
Table A.5	List of the orthogonal atomic coordinates of R-VO ₂ with the supercell definition. The ‘Unit cell shift’ shows the movements of the atom from its original atomic coordinates in Table A.1 to the new coordinates. The list is ordered in a one-to-one correspondence with the M ₁ structure atomic coordinates in Table A.4	242
Table A.6	Four pairs of the V-V bonds in the supercell. $\Delta x_i = x_{i(m)} - x_{i(n)}$ for the bond V _{n} -V _{m} , where $i = \{x, y, z\}$	242
Table A.7	List of the orthogonal atomic coordinates of the rutile and the M ₁ structures	

in the supercell coordinates. The atoms are moved that one-to-one correspondence on the atomic coordinates is formed between two struc- tures.	244
--	-----

Table C.1 Various critical temperatures obtained during heating (T_c^H) and cooling periods (T_c^C) in the metal-insulator transition (MIT) and the structural phase transition (SPT) characterizations of VO ₂ nano-beams supported on different substrates. The supercooling effect can be gauged by the downshift of the critical temperature obtained during cooling compared to that during heating period (ΔT)	250
--	-----

LIST OF FIGURES

Figure 1.1: Ultrafast electron diffraction setup. An amplified laser pulse (red) is split into two parts. One part goes through a delay stage and into the UHV chamber as pump to excite the sample. The other part is frequency-tripled to 266 nm (blue) and sent into the chamber to excite the photocathode. The photo-generated electron beam (pink) is accelerated to 30 keV and diffracted from the sample. An intensified CCD is used to record the diffraction pattern at different time delays	4
Figure 1.2: Bragg diffraction. Two rays of electrons approach the lattice with same wavelength, same phase and same incident angle. They are scattered by two different atoms in the same set of crystal planes. The lower beam travels an extra distance $2d \sin(\theta/2)$. Constructive interference occurs when this distance equals to an integer multiple of the wavelength.	7
Figure 1.3: Laue diffraction conditions. Two rays of electrons with the same phase and wavevector approach the crystal with incident angle θ_1 . They are scattered by two atoms separated by distance a . The angle after scattering is θ_2 . The path difference of the two rays is $a(\cos \theta_1 - \cos \theta_2)$, which equals to an integer multiple of the wavelength, when the constructive interference occurs	8
Figure 1.4: The Ewald sphere is shown intersecting the reciprocal lattice points. The reciprocal lattice vectors are $\vec{a}^*$ and $\vec{b}^*$. $\vec{k}_i$ is the wavevector of incident electrons, pointing to the origin of the reciprocal lattice O . The wavevector of scattered electrons is $\vec{k}_f$. The vector of momentum transfer is $\vec{s}$	12
Figure 1.5: Scattering of the electron wave with an isolated atom. The incident electron wavevector is $\vec{k}_i$ and after scattering the wavevector becomes $\vec{k}_f$. k is the scalar wavevector of the scattered spherical wave. The circles represent the wave-front of scattered spherical waves	14
Figure 1.6: Change of form factor $f(\theta)$ as a function of scattering angle θ for several different metals. The strength of $f(\theta)$ decays with the increase of θ and increases with the atomic number Z . Figure from Ref. [7]	16
Figure 1.7: Illustration of crystal unit cell with two atoms in the basis, at_1 and at_i . The unit cell vectors are $\vec{a}$, $\vec{b}$ and $\vec{c}$. The origin is placed at at_1 and the coordinates of at_i are defined by the fractional coordinates x_i , y_i and z_i .	17
Figure 1.8: Illustration of dynamic scattering of electrons. The electron beam originally scattered from planes $[hkl]$ is scattered for the second time from planes $[h'k'l']$, which causes redistribution of intensities of Bragg reflections.	20

Figure 2.1: Free electron levels in a periodic potential depicted in a repeated-zone scheme	30
Figure 2.2: a) Schematic representation of non-degenerate electronic levels in an atomic potential; b) Energy levels of N such atoms in a periodic array, plotted as a function of mean inverse interatomic spacing. When the atoms are far apart (small overlap integrals), the levels are nearly degenerate, but when the atoms are closer together (larger overlap integrals), the levels broaden into bands. Figure from Ref. [45]	33
Figure 2.3: Flow chart of a typical DFT algorithm.	40
Figure 2.4: Comparison of the energies of the itinerant ground state and the localized ground state of the half-filled system. A phase transition happens at the critical Hubbard U_c , which separates the metallic and Mott-insulating states	43
Figure 2.5: Band structures comparison between a band insulator and a Mott insulator. In a band insulator, as illustrated in a), the valence band is half-filled. For N sites on the lattice, there are totally $2N$ electrons filled in the valence band. The factor of 2 accounts for electron spins. The $2N$ empty states are located in the conduction band. b) in a Mott insulator, there are N sites on the lattice and only N electrons are filled. The half-filled band (in band-theory sense), is separated into two narrow bands: the lower Hubbard band and the upper Hubbard band. The two bands are separated by an energy $U - t$. Each Hubbard band contains N states and the lower Hubbard band is fully filled. When an electron moves from one site to the neighboring site, additional energy due to the on-site Coulomb interaction needs to be paid, which lifts its energy level, as illustrated by the real-space atomic energy levels [56]	43
Figure 2.6: A schematic band structure of VO_2 in a) the metallic-rutile phase and b) the insulating-monoclinic phase, according to Ref. [57] and [58] . .	44
Figure 2.7: Band structure of insulating monoclinic M_1 phase of VO_2 calculated a) using DFT LDA method. Figure from Ref. [60]. b) using DFT LDA+ U method. Figure from Ref. [61]	48
Figure 2.8: DMFT replaces the full lattice and electrons with a dimer cluster (yellow spheres) and a bath of non-interacting electrons. The model captures snapshots of the electronic state fluctuation on the dimer cluster. The population of the dimer cluster is determined by the quantum mechanical probability for hopping of electrons from bath to the cluster, which is quantified by parameter V_v	49
Figure 3.1: The structures of VO_2 rutile (R) and monoclinic (M_1) phases	58

Figure 3.2: SEM image of VO ₂ thin film. Nanogranular feature of the thin film is observed with grain size ~50 nm	62
Figure 3.3: a) Ground-state powder diffraction pattern from VO ₂ thin-film membrane. b) 1D curve extracted from experimental diffraction pattern (black squares) with the diffraction pattern calculated using the kinematic diffraction theory (red curve). Miller indices are labeled for reference.	63
Figure 3.4: Resistance change (red line) and integrated intensity change of the Bragg peak [30 $\bar{2}$] (black symbols) as a function of sample temperature. ΔI is the intensity change of the Bragg peak and I_0 the intensity of the Bragg peak at the room-temperature ground state	64
Figure 3.5: Absorbed-energy-density dependence of the integrated peak intensity of [30 $\bar{2}$] at 295 K excited by pump photons of different wavelengths. The integrated intensity is selected at 150 ps after pump excitation. Solid lines are the error-function fittings of the experimental data. Inset: Room-temperature transient change of the integrated intensity of peak [30 $\bar{2}$], after the excitation by a 50 fs laser pulse centered at 800nm with absorbed fluence 7.8 mJ/cm ²	66
Figure 3.6: Transmittance measurement of 800 nm (a), 1319 nm (b) and 2000 nm (c) on 1) reference sample with supporting silicon membrane (dashed lines); 2) VO ₂ thin-film membrane (solid lines). (d) Illustration of the optical reflection and transmission when the VO ₂ thin-film membrane is measured. n_{VO_2} and n_{Si} are the refractive indices of VO ₂ and amorphous silicon, respectively. Reflectance on the interface is R_{Int}	70
Figure 3.7: Dependence of the critical energy density E_c on the sample temperature T_B . The triangle symbols represent the width of the energy-density window for the phase transitions, which is extracted by the width from the error-function fittings in Fig. 3.5. The orange line is the thermodynamic energy requirement for the phase transition calculated according to Ref. [133].	72
Figure 3.8: Electronic band structure of VO ₂ M ₁ phase [57, 115]. With 2000 nm (0.62 eV) pump excitation, the excited electrons can efficiently fill in the π^* band across the band gap, while the photon energy of 800 nm (1.55 eV) pump is so high that the electrons are directly excited to the high-energy $d_{ }^*$ band. An additional scattering process (white line) is required to transfer these electrons to the π^* band and contribute to screening the strong electron correlation.	74
Figure 3.9: Comparison of Bragg peak intensity dynamics of different Miller indices for 800 nm pump and 2000 nm pump excitations.	75

- Figure 3.10: Structures of VO_2 monoclinic and rutile structures. The red dashed lines define the unit cell of rutile structure. The blue parallelogram defines the unit cell of monoclinic structure. The blue dashed lines represent how the vanadium atoms are distorted from rutile structure to monoclinic structure. 77
- Figure 3.11: Schematic diagram of the atomic movements in the case a) directly move from the monoclinic positions to the rutile positions; b) first tilt perpendicular to the c_R axis, then stretch along it. Half of the supercell is plotted, and the V atoms in the other half move following their structurally equivalent atoms. The blue arrows represent the directions V atoms move. c) Comparison of the simulated intensities with the experimental results in the case of direct movement. The experimental results are obtained by pumping with 800 nm laser at the fluence 7.8 mJ/cm^2 . Dashed lines are the simulation results. Symbols are the experimental results. The intensity changes of the peaks $[30\bar{2}]$ and $[31\bar{3}]$ are scaled by 0.375 in the plot. d) Comparison of the calculated intensities with the experimental results in the case of tilting prior to stretching scheme. The change of the movement directions happens at 2 ps, which is marked by the vertical dashed lines. 81
- Figure 3.12: a) Schematic diagram of a three-step movement scenario. The V atoms first move away from each other along the dimerization direction. The second step happens at ~ 1 ps after pump excitation, in which the V atoms start turning into the direction perpendicular to the dimerization. At ~ 2.5 ps time delay, the movement direction turns for the second time that the V atoms move towards the rutile structure positions. b) Comparison of the calculated intensity changes (dashed lines) for the scenario provided in a) with the experimental results (symbols). . . . 82
- Figure 4.1: Molecular orbital diagram depicting the electronic structure of the monoclinic and rutile phases of VO_2 . The left MO diagram corresponds to the undistorted metallic phase of VO_2 , whereas the diagram on the right shows the altered MO diagram upon transition to the distorted insulating phase of VO_2 [57, 139]. 88
- Figure 4.2: a) Monoclinic unit cell of VO_2 M_2 phase. The displacement of vanadium atoms from equilibrium position of its rutile structure is presented. b) Phase diagram of $\text{V}_{1-x}\text{Cr}_x\text{O}_2$ [143]. 91
- Figure 4.3: Comparison of experimentally measured photoemission spectra with the calculation results of c-DMFT. Top: Bulk-sensitive photoemission spectra of VO_2 taken at 300K (insulating phase) and 348 K (metallic phase) [58]. Bottom: c-DMFT results of V $3d(t_{2g})$ spectral weights from Biermann et al. [69]. 92

Figure 4.4:	(a) Optical image of a VO ₂ nano-beam suspended on a TEM window. (b) The normalized color-contrast change in the optical image of VO ₂ associated with a first-order insulator-to-metal transition at 339 K. The color contrast is quantified by the intensities of the three primary colors (blue, green and red) of the VO ₂ nanobeams in the optical images. (c) TEM diffraction patterns obtained from a VO ₂ nano-beam suspended over a Ni TEM grid showing a monoclinic-to-rutile structural transition at 344 K. (d) Various metal-to-insulator transitions (MIT) and structural phase transitions (SPT) characterized over different metal TEM grids (Ni, Au, Cu), on a Si window, and in the as-grown condition where the nano-beam is anchored on a Si substrate.	95
Figure 4.5:	a) VO ₂ phase diagram as a function of the interface charge doping. The vertical axis is the noncooperativity (η), which measures the upshift of the structural critical temperature (T_{SPT}) from the electronic one (T_{MIT}). b) Band structures of VO ₂ in different phases. The left and middle panels are schematically drawn following Ref. [57, 106]. The right panel describes the proposed band structure for the monoclinic metal (M_3).	97
Figure 4.6:	Optical images of VO ₂ nano-beam supported on a Cu TEM grid. A shift in color depth as a result of increase in optical reflectivity occurs at ≈ 339.3 K, observed under a color filter centered at RGB values: (185, 159, 179) to enhance the contrast.	98
Figure 4.7:	(a) SEM image of VO ₂ nano-beams grown on a SiO ₂ substrate. (b) The critical fluence (F_c) for inducing the monoclinic-to-rutile transition as a function of base temperature T_B . The insets show the corresponding UED patterns of the two phases. (c) The threshold behavior of laser-Induced phase transition, as characterized by the disappearance of the dimer reflections ($S1$), $[51\bar{2}]$, $[30\bar{1}]$ and $[\bar{1}10]$ in (b).	103
Figure 4.8:	(a) UED pattern of a single suspended VO ₂ nano-beam, showing the reciprocal lattice of the ac-plane of the monoclinic structure. (b) The corresponding real-space arrangement of the V sub-lattice. The fs electron beam is directed along b_M . (c) Lattice fluctuational dynamics in M_1^* and M_3 states are measured by the projected atomic displacement variance: Δu_{out}^2 (along a_M^* , out-of- bc -plane vibration), and Δu_{in}^2 (in bc -plane vibration). The corresponding Δu^2 and temperature increases expected for the M_1 and R phases are plotted for comparison.	105
Figure 5.1:	Dynamics of charge-density-wave (CDW) satellite reflections as the function of delay time, plotted for different 266 nm pulse energies focused on the photocathode for photoelectron generation. The solid lines are the error-function fitting to extract the pulse duration for each laser pulse energy. Inset: UED electron pulse duration as a function of the 266 nm laser-pulse energy.	113

- Figure 5.2: a) Geometry of the all-optical electron-pulse duration measurement. Electron-beam cross sections are shown at various positions after the interaction region for a 600-fs electron pulse; b) Electron-laser pulse cross-correlation. The reduction in electron-beam intensity transmitted through aperture A_2 is sensitive to τ and the duration of the electron pulse at the interaction point (see the legend). The impulse response of the measurement (inset) is ~ 300 fs e^{-2} full width (FW) for the parameters $A_1=30$ μm , $t_l=200$ fs (dashed curve), but only, 80 fs e^{-2} FW for $A_1=10$ μm , $t_l=60$ fs (solid curve). Figure from Ref. [167]. 115
- Figure 5.3: Schematic of the projection-shadow-imaging experiment. For illustration purposes, the angular span of the shadow is significantly increased from typical values ≤ 1 mrad. 116
- Figure 5.4: a) Snapshot shadow images of photoemitted electron bunch obtained with $x_0=4.3$ mm, which gives magnification $M \sim 38$. The cross-sectional line profiles (black curves) are extracted along the dashed line in the second panel. b) Fitting curves (lines) of cross-sectional line-profile data (circles) using Eqn. 5.2.1. Figure from Ref. [158]. 118
- Figure 5.5: Geometry of projection-shadow-imaging technique, where the photoelectron pulse is fitted to a three-dimensional Gaussian function, $f(x, y, z)$. The profile of the shadow image on the camera is described as an analytic function $F(d)$, where d is the position on the camera screen. 118
- Figure 5.6: The snap-shots of the normalized shadow images at selected times. The magnification of the projection imaging is ≈ 33 123
- Figure 5.7: The results of electron point-projection imaging. (a) The COM-frame expansion dynamics of electron pulses with longitudinal pulse length σ_z . The solid lines through the experimental data (symbols) are from N-particle simulations. (b) Dependence of pulse length on the electron sheet density at time 100 ps for various F and F_a settings, showing a universal power-law increase with exponent roughly $\gamma \approx 0.5$. Predictions based on mean-field [14] and one-dimensional fluid [175] models are also presented. 124
- Figure 5.8: (a) The symbols represent the values of γ found at different delay times. The solid line represents results from an analytical Gaussian model (AG-M) simulation using broadened initial longitudinal velocity spread due to near cathode space charge effect extracted by fitting the early time σ_z trajectory, which is presented in Fig. 5.9(c) and the text. (b) The symbols are the linescans of the shadow images recorded on the phosphor/CCD screen produced by photoelectrons ($\Sigma = 7.12 \times 10^{13}$ e/m²) at different times. The linescans are fitted with a Gaussian profile (solid lines). A top-hat profile (dashed line), convoluted with the projection

geometry, is also drawn for comparison. The magnification of the projection imaging is ≈ 33 126

Figure 5.9: (a) The normalized initial longitudinal phase space employed in the N-particle simulations, which is parameterized by a thermal parameter a and scaled by the initial length σ_{zi} and momentum spread σ_{pz} . (b) The percentage change in the pulse length σ_z obtained at 100 ps as a function of a . (c) The initial longitudinal velocity spread, $\Delta v_z(t=0)$, obtained by fitting N-particle trajectories to the imaging data depicted in Fig. 5.7(a). The extraction field F_a applied is 0 at the five lowest Σ s, 0.32 MV/m at 40×10^{12} e/m² and at 70×10^{12} e/m². 128

Figure 5.10: The symbols represent the experimental longitudinal pulse length (σ_z) of the photoelectron pulses with different densities tracked at different times. The solid line represents the analytical Gaussian model (AGM) simulation using the different initial longitudinal velocity spread specified by Fig. 5.9(c) and the initial thermal (chirp) parameter, a , set to 0. The dotted line represents the analytical Gaussian model simulation using a constant initial longitudinal velocity spread $\Delta v_z(t=0)=0.084$ $\mu\text{m/ps}$ 130

Figure 5.11: Momentum cut model for selecting electrons for photoemission used in the three step model. (a) The electron energy distribution before and after photo-excitation. In this model, the electrons are assumed to escape the cathode surface before thermalization and thus have the same Fermi-Dirac (FD) energy spread (exaggerated for illustration purpose) defined at the temperature prior to photoemission. The shaded area ($E \geq E_F + \Phi_{eff}$) represents the electron population that is qualified for photoemission. (b) The selection of electrons in the three-dimensional momentum phase space that are qualified for photoemission. The electron must have an energy larger than $EF + \Phi_{eff}$, as described in (a); it must also have a minimum longitudinal momentum (p_z^0) to overcome the work function in order to escape the surface, as specified by the shaded area. 132

Figure 5.12: (a) The photoelectron velocity spread in the longitudinal (Δv_L) and transverse (Δv_T) directions. The solid symbols represent results obtained using N-particle simulation, and the hollow symbols represent results obtained by integrating the analytical equation reported in Ref. [184]. (b) The longitudinal velocity distribution of photoelectrons generated by N-particle simulation based on the three step model. . . 133

Figure 5.13: (a) Schematic of an extended three step model for intense photoemission. The net work function is the sum of the Schottky potential Φ_{Sch} , the surface dipole potential Φ_{dp} , and the intrinsic work function Φ_w . (b) The quantum efficiency derived based on modified TSM (dashed line) for $\Phi_w=4.26$ eV and the experimental data with a fit to a constant

behavior at low field and a linear behavior at high field. (Inset) The electron escape ratio, R_{esp} , is calculated, including the dipole field of the virtual cathode and its image for $\Phi_w=4.26$ eV. (c) Electron pulse sheet density as a function of fluence for high applied field and for zero applied field (inset). 135

Figure 5.14: A schematic depiction showing the AGM parameters in the representation of the usual position-momentum phase-space plots: $\sqrt{\sigma_i}$ and $\sqrt{\eta_i}$ are the variances of the i direction spatial and momentum distributions, respectively; and the slope, given by γ_i/σ_i , is related to the (linear) momentum chirp of the pulse. 138

Figure 5.15: (a) Space-charge-limited temporal resolutions in ultrafast electron diffraction (UED) and microscopy (UEM) systems. Solid squares and hollow diamonds show the Coulomb-explosion-led pulse lengthening calculated for 100 keV UEM system (squares) with cathode-to-sample distance of 70 cm and 30 keV UED system (diamonds) with cathode-to-sample distance of 5 cm. The rectangular shaded areas depict experimental resolutions reported in current UED and UEM systems. In comparison, the solid stars and circles show the improvements in temporal resolution by employing an RF recompression in the UEM beam column (see Fig. 5.16a) optimized for nano-area diffractive imaging (stars) and for single-shot UED (circles). 141

Figure 5.16: (a) Photoelectron pulse evolution along an RF-enabled UEM column. The shaded regions represent the locations of the electron optical elements. (b) A scale-up view of the pulse profiles near the sample plane for nano-area diffractive imaging containing 10^5 electrons/pulse. An aperture with radius of $15\text{ }\mu\text{m}$ is employed to thin out the peripheral electrons to achieve a divergence angle $\alpha \leq 1.7$ mrad. The minimum transverse radius σ_T is $0.5\text{ }\mu\text{m}$, and the minimum longitudinal pulse length σ_L is $1.8\text{ }\mu\text{m}$. (c) A scale-up view of the pulse profiles near the sample plane for an ultrafast single-shot UED containing 10^8 electrons/pulse. The minimum transverse radius σ_T is $51\text{ }\mu\text{m}$, and the minimum longitudinal pulse length σ_L is $0.88\text{ }\mu\text{m}$ 143

Figure 5.17: Structure of RF-compression UEM constructed at Michigan State University, with important components labeled. 147

Figure 6.1: Concept of an electron beam injector column for ultrafast electron microscope demonstrated using an N-particle simulation of electron pulse propagation in an electron beam column with an RF cavity. (a) Phase space adjustment before and after the RF cavity. (b) The corresponding real space pulse profile at each electron optical components. 149

Figure 6.2: Summary of the time-frequency correspondence for a pulse train [196]. . 151

Figure 6.3: a) Optical frequency comb measured using a high-speed photodetector and a spectrum analyzer. Each equally spaced line represents a comb frequency. The spacing between combs equals to the repetition rate of laser pulse train. b) Experimental setup for single-harmonic extraction. The RF bandpass filter is used to select a single harmonic from the comb structure.	153
Figure 6.4: a) Perfect electron pulse timing matches with the RF cavity field. No COM velocity change. b) and c) are the situations where the electron pulse timing is shifted relatively to the RF-field phase to decelerate the COM in b) or accelerate the COM in c), respectively.	154
Figure 6.5: Block diagram of a general phase-locked loop (PLL).	156
Figure 6.6: Signal-flow graph of a PLL under linear model. Loop filter is assumed to be unitary in this model.	159
Figure 6.7: Frequency response of a PLL.	160
Figure 6.8: a) Signal flow diagram of a PLL with the loop filter, K_h . b) Simple voltage attenuator design of the loop filter to reduce the PLL bandwidth.	162
Figure 6.9: a) Noise as a function of time; b) Mean-square noise in the time domain; c) Power spectral density of the noise [203].	163
Figure 6.10: a) Phase noise definition in the typical spectrum of a non-ideal oscillator b) Single-sideband Phase noise spectrum plotted with offset frequency in log-scale	164
Figure 6.11: a) Power fluctuation of the 12 th harmonic from the laser oscillator; b) Shift of static-phase-error caused by the shift of input reference frequency. An RF signal generator was used in this measurement to replace laser oscillator, which gives more flexible control of reference frequency. Black symbols represent the results with the passive loop filter and red symbols with the active filter. Design of active filter is shown in Fig. 6.15. c) Two period monitoring of the reference frequency drift from laser oscillator.	166
Figure 6.12: a) Active loop filter design with ~100 dc gain for the RF-compression UEM. b) Frequency response of the loop filter simulated using QUCS.	169
Figure 6.13: Experimental setup to measure the bode plots of the PLL	170
Figure 6.14: a) Design of the phase-lag compensator b) Bode plots for the PLL with the original loop filter design and a loop filter with a phase-lag compensator.	173

Figure 6.15: Overall loop filter design with ~ 100 dc gain and lag-compensator to improve its stability	174
Figure 6.16: a) Signal flow diagram with detailed setup of RF components b) Picture of the hardware setup of the PLL	176
Figure 6.17: Phase noise spectra of the PLL output and reference input measured using spectrum analyzer(FSP13). Dashed lines are the fittings of the noise spectrum using power-law relations with different exponents. . .	179
Figure 6.18: a) Experimental setup of timing jitter measurement using the phase-locked loop method. b) Timing jitter of the PLL for ~ 10 hours' period. c) Room-temperature measured in the same period of time as b). . . .	181
Figure 6.19: a) Schematic diagram of temperature feedback control of PLL electronic box. b) Timing jitter measurements of temperature stabilized (black) and unstablized (red) PLL with similar room temperature fluctuations. . .	183
Figure 7.1: a) Simulation result of the RF cavity field distribution with Superfish code, plotted in cylindrical coordinates. Shaded area is the cavity chamber. Solid lines are the equipotential lines. Inset: Zoom-in view of the electric-field distribution of the TM_{010} mode on the axis, along which the electron pulses propagate. b) Picture of the copper RF cavity for the RF-compression UEM.	188
Figure 7.2: a) Circuit layout of a solid-state amplifier b) Typical gain measurement for the solid-state amplifier. Blue solid line is the fitting with 16.2 dB gain. Red dashed line shows the 1 dB compression power.	194
Figure 7.3: Schematic diagram of RF amplifier station. The power level is marked for the situation, when the amplifier station is running at its maximum power output.	195
Figure 7.4: Output power measurement for RF amplifier station in 1) Pulsed mode (black square) and 2) Continuous mode (red triangle). The solid line is a linear fit on the low-power data.	197
Figure 7.5: Equivalent circuit model of RF cavity	199
Figure 7.6: Power (a) and relative phase (b) response, when the frequency is detuned from CRF by Δf	201
Figure 7.7: a) Experimental setup for measuring the cavity frequency detuning effects b) A typical power and phase response to the frequency detuning, Δf . Cavity power (black) reaches the maximum at the resonant frequency ($\Delta f = 0$), while the reflected power (red) drops to a minimum. The relative phase (blue) shifts with the frequency detuning Δf	203

Figure 7.8: a) Cavity resonant frequency measurements with different angles of the external tuner; Inset: setup of the cavity external tuner b) Reflection power as a function of relative phase with different external tuner angles, ϕ ; Inset: the shift of cavity resonant frequency (CRF) as a function of external tuner angle. CRF at 0° tuner angle is taken as the reference.	204
Figure 7.9: Block diagram of the setup for cavity phase-lock .	206
Figure 7.10: Cavity-level phase jitter measurement with a) input RF power ~460 W and cavity phase-lock working; b) input RF power ~40 W and cavity phase-lock working; c) input RF power ~40 W and ~460 W, but cavity phase-lock not working .	207
Figure 7.11: Cavity-level phase jitter measurement with a) input RF power ~50 W and detuning frequency, $\Delta f = +3$ kHz; b) input RF power ~50 W and $\Delta f = -4$ kHz; c) input RF power ~460 W and $\Delta f = +3.2$ kHz. (positive Δf means increase of input signal frequency; negative Δf means decrease) .	209
Figure 7.12: Schematic diagram for overall RF system setup.	211
Figure 8.1: a) Optical microscope image of VO ₂ thin-film membrane sample. The VO ₂ is directly deposited on the silicon membrane in the square opening windows. The window size is 100×100 μm^2 ; b) Powder diffraction pattern from thin-film membrane VO ₂ ; c) 1D curve reduced from the diffraction pattern shown in b), which shows the intensity distribution of Bragg peaks as a function of momentum transfer [44].	214
Figure 8.2: Experimental setup for measuring the cavity electric field strength using electron pulses. The optical delay line is installed in the path of 266 nm laser, which adjusts the relative timing difference between the electron pulses and the RF field. The electron energy is probed using the diffraction patterns from the sample collected with an intensified CCD camera.	216
Figure 8.3: Experimental results for the electron energy as a function of the delay of the 266 nm laser pulses. The electron energy is extracted from the diffraction pattern using the Bragg's law (Eqn. 1.2.1). The red curve is the theoretical fitting of the experimental data using Eqn. 7.2.1, and the relativistic effects are considered. According to the fitting, the electric field strength E_0 is 2.43 MV/m and the RF field frequency is 1.0125 GHz. Inset: electron velocity as a function of timing jitter.	218
Figure 8.4: a) Intensity dynamics of the Bragg reflection $[30\bar{2}]$ with different RF powers in the cavity. The solid lines are the fitting of the dynamics with error functions, whose width yields the temporal resolution of the RF-compression UEM. The total number of electrons are all $\sim 10^6$ e ⁻ /pulse	

before the RF cavity. Due to the aperture cutting, $\sim 10^5$ e⁻/pulse for 0 W and 210 W RF power and $\sim 10^4$ e⁻/pulse for 300 W power are used for diffraction. b) Temporal resolution, which is extracted from the error-function fitting in (a), as a function of RF power. 221

Figure A.1: V-V bond distortion illustrated under supercell definition. The dashed red circles represent the rutile structure, in which the V atoms are equally spacing along rutile c_R axis. The blue solid circles represent the M_1 structure. The blue dashed lines show the atomic distortion from rutile position to monoclinic position in y-z plane. The cross means atom moves along -x direction, while the dot +x direction. The displacements of V atoms are exaggerated for the purpose of illustration. Blue solid line encircles the unit cell of M_1 -VO₂. 233

Figure B.1: Refractive index n and extinction coefficient k as a function of the photon energy, derived using the data from Ref. [126]. 236

Figure B.2: Reflectivity R and optical penetration depth δ as the function of the wavelength λ 237

Figure C.1: a. The SEM image of VO₂ nano-beams grown on a SiO₂ substrate. The electron beam cross-section is presented in the circular area. b. Diffraction pattern of VO₂ nano-beam as-grown on SiO₂ substrate at room temperature 241

Figure C.2: a) Electron diffraction pattern with Miller indices of VO₂ single crystal nanobeam measured using UED. Inset: Crystal-maker simulation of the diffraction pattern. 243

Figure C.3: Principle of projection of atomic variance with arbitrary vector q^* and $q_{\perp}^*$ onto $q_{[100]}^*$ 244

Figure C.4: (a) The projected atomic displacement variance u^2 , compiled based on data extracted from X-ray diffraction experiments (see Marezio et al. [143] and McWhan et al. [149]), and X-ray near-edge absorption spectroscopy experiment (see Yao et al. [217]). Generally, u^2 is larger along the rutile c-axis (c_R) than along the rutile a-axis (a_R) and b-axis (b_R , not shown, but similar to a_R values). A significant rise of u^2 following insulator-metal transition is clearly visible. (b) The correlations between the temperature rise (ΔT) and the increase in atomic displacement variance (Δu^2) in the transition region are shown. 246

Figure D.1: a) Analytical description of the shadow images (Eqn. 5.2.5) on the camera, with $\lambda\zeta\Sigma = 1.0$, $\sigma_z = 10$ μ m, $L = 165$ mm, $z_0 = 20$ μ m and varying PPD ($x_0 = 5$ (black), 4(red), 2(green), 1(blue) mm). The curves have the same intensity, but different integrated intensities. (b) Analytical function after normalization, $M(d)$ (Eqn. D.1.2), with the

same set of parameters used in (a). The integrated intensity is a constant with varying x_0 249

Figure D.2: a) Side-view of the experimental setup. Pump laser is aimed at a fixed position on HOPG surface with fluence 54 mJ/cm^2 , to ensure the reproducible photoemission conditions, when x_0 is changing. b) Center of mass (COM) location above HOPG surface as the function time, which is obtained from analytical-model analysis (Eqn. 5.2.6). c) Electron pulse longitudinal width, σ_z , evolves as the function of time. 250

Figure D.3: a) Geometry of shadow image simulation. The source point P is at the same height as the target electron pulse. b) Comparison of simulated shadow image (red open triangle) with the experimental results (black solid square) with $x_0 = 5.48 \text{ mm}$. The data is taken with delay time $\Delta t = 100 \text{ ps}$ after photoemission. The parameters in the shadow-imaging simulation are $\sigma_x = 370 \text{ }\mu\text{m}$, $\sigma_y = 260 \text{ }\mu\text{m}$, $\sigma_z = 40 \text{ }\mu\text{m}$ and $\Sigma = 8.7 \times 10^{12} \text{ m}^{-2}$. c) Comparison of simulated shadow image (red open triangle) with the experimental results (black solid square) with $x_0 = 1.71 \text{ mm}$. The dimensional parameters are the same as b), while $\Sigma = 8.5 \times 10^{12} \text{ m}^{-2}$. . . 253

Figure D.4: a) Schematics of the PSImg technique [158]. Bragg peaks from HOPG sample cross the photoelectron pulse above the HOPG surface, as shown in the inset. The dashed arrows illustrate the effects of the photoelectron pulse applying on the Bragg beams. b) Bragg beams' movements as a function of time delays. The symbols are the experimental results. The solid lines are the results of the Bragg-beam deflection simulation with the electron sheet density $\Sigma = 8.0 \times 10^{12} \text{ m}^{-2}$ 255

Chapter 1

Introduction and Methodology

To have a camera with the ability to record the atoms in motion is a scientist's wildest dream. This 'dream camera' will enable us to resolve the fundamental physics behind many fascinating phenomena across different disciplines.[1] In condensed matter physics, resolving the atomic motions throughout a phase transition leads to the understanding of the phase-transition mechanism; in chemistry, the ability to directly observe the process of making and breaking chemical bonds in a reaction is the key to finding out the reaction pathway, and thus, to ultimately control the products of chemical reactions; in structural biology, knowing the structural dynamics of the protein is very important to figure out its functionality in a biological process, etc. To fully describe these processes, a 'camera' which can follow a sequence of steps in real-time and characterize the given process is required. More technically, such a 'camera' needs to satisfy two major requirements:

1. sensitivity to the atomic structure, (sub-Angström resolution)
2. temporal resolution of tens to hundreds of femtoseconds. ($1 \text{ fs} = 10^{-15} \text{ s}$).

Electron diffraction and microscopy has a long history as powerful techniques to study the static structures of materials. Because the de Broglie wavelength of high-energy electrons reaches the picometre scale, sub-Angström resolution can be achieved

in real-space images using an electron microscope. In addition, electron diffraction, which utilizes the interference effects from elastically scattered electrons, also provides important characterization of material structures. So, equipping the ‘camera’ with electrons as the probe allows it to satisfy the first requirement on the spatial resolution.

In solids, the speed of sound is usually 300 - 1000 m/s. This determines the timescale with which the atoms move from one structural phase to another by a displacement ranging from Angströms to microns. So, the time-scale of atomic motions in solids ranges from femtoseconds (fs) to nanoseconds(ns) [2]. To clearly resolve such movements, a temporal resolution of tens to hundreds of femtoseconds is a necessity. Unfortunately, this is beyond the speed of any mechanical or electronic camera shutters. To resolve the dynamics on fs time scales, the pump-probe method based on ultrafast lasers has been routinely used [3]. Temporal resolution as short as 10 fs can be achieved, which is only limited by the laser pulse duration. The details of the pump-probe method are included in Section 1.1. So, by combining electron diffraction (microscopy) and the pump-probe method, the ‘dream camera’ can be realized, which enables us to ‘see’ the atomic motions in real-time throughout a phase transition, a chemical reaction or a biological process. This ‘camera’ is the technique of *ultrafast electron diffraction/microscopy*, which is one of the main topics of this thesis.

This chapter introduces the technique of ultrafast electron diffraction (UED) and the fundamental diffraction theories, which is necessary for understanding the discussions in the following chapters. The description of the pump-probe method and the instrumental setup of the UED system is first given in Section 1.1. The theories for diffraction have been well established, but they are very important for understanding the various results to be discussed in the following chapters. An attempt has been made in Section 1.2 to provide an introduction to diffraction and the interpretation of the diffraction patterns. The next generation of ultrafast electron diffraction and microscopy systems are being actively developed throughout the world. In Section 1.3,

the current status of these systems is reviewed and the advantages and disadvantages of these techniques are also discussed. In Section 1.4, the RF-compression ultrafast electron microscope currently under development at Michigan State University is introduced.

1.1 Ultrafast Electron Diffraction

Ultrafast electron diffraction (UED) is a pump-probe experimental method that uses a fs pump laser pulse to excite the sample, while its consequent structural dynamics are probed by a short electron probe pulse through electron diffraction.

The pump-probe experimental method was first applied in time-resolved optical experiments. In these experiments, an ultrashort laser pulse is split into two portions: a strong pump pulse and a weak probe pulse, both of which are focused and overlapped at the same position on the sample (*spatial overlap*). The relative time delay between the pump and probe pulses can be adjusted in fine steps by changing the optical path length difference. The purpose of the pump pulse is to prepare an excited state of the sample. Its optical constants (optical reflectivity, transmittance, spectrum, etc.) at a certain *delay* time are probed through the weak probe pulse, which arrives at a precise temporal delay. Clearly, when the probe pulse arrives earlier than the pump pulse, it sees the sample before excitation, which we call the *ground state*. When the optical path difference between pump and probe is adjusted so that the probe pulse arrives later than the pump, some changes induced by the pump pulse are ‘seen’ by the probe pulse as the change of the values of the optical constants compared to the ground state. The changes of optical constants can be stamped precisely in time, by knowing the *zero-of-time* (ZOT). Technically, ZOT is defined as the *temporal overlap* of pump and probe at the sample. The temporal resolution of a pump-probe experiment is usually limited by the pulse duration of either pump pulse or probe pulse. The

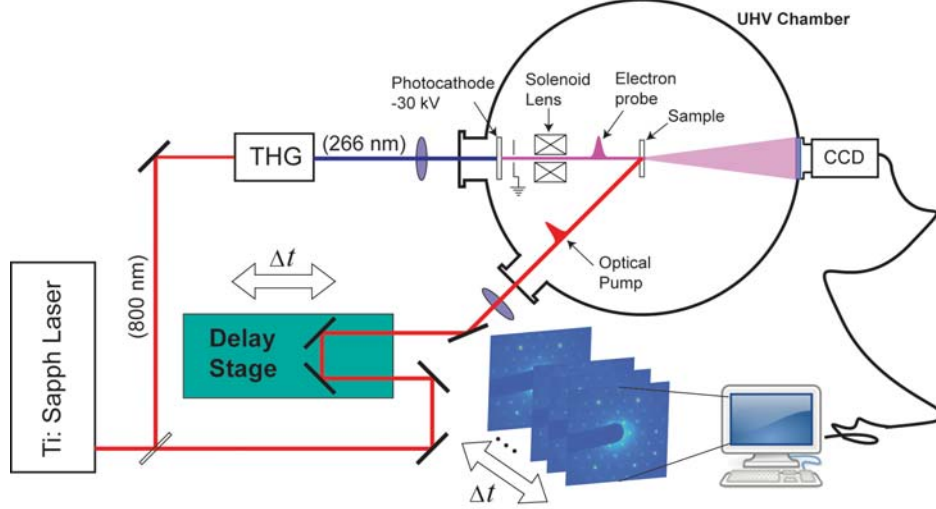


Figure 1.1: Ultrafast electron diffraction setup. An amplified laser pulse (red) is split into two parts. One part goes through a delay stage and into the UHV chamber as pump to excite the sample. The other part is frequency-tripled to 266 nm (blue) and sent into the chamber to excite the photocathode. The photo-generated electron beam (pink) is accelerated to 30 keV and diffracted from the sample. An intensified CCD is used to record the diffraction pattern at different time delays.

temporal resolution in an optical pump-probe experiment can be as short as 10 fs. Even though they provide necessary temporal resolution, optical techniques mostly reflect the dynamics of the electronic degrees of freedom, rather than the structural motion, because optical probes do not couple to the lattice directly.

The difference in an UED experiment is that short, high-energy electron pulses, instead of fs laser pulses, are used as the prob. Because the wavelength of high-energy electrons is comparable to the lattice spacing in solids, they are sensitive to structure through diffraction effects. The schematic setup of UED experiments is depicted in Fig. 1.1. Similar to the optical pump-probe technique, the fs laser pulse is split into two portions and the first portion is used as the pump to excite the sample in the ultrahigh-vacuum (UHV) chamber. The second portion is frequency-tripled and focused onto the photocathode to generate ultrashort electron pulses through the photoelectric effect. The frequency tripling is realized through the third-harmonic-generation of the originally 800 nm (1.55 eV) laser pulses and is necessary to overcome

the work function of the photocathode material (silver, work function ~ 4.26 eV [4]). The relative time delay between the pump laser pulse and the electron probe pulse is adjusted by an optical delay stage as shown in Fig. 1.1. The structural change induced by the pump laser is hence probed by high-energy electron pulses through electron diffraction. The diffraction patterns are recorded by an intensified CCD camera [5].

In the first-generation UED system, the electron pulses are accelerated to 30 keV in the electron gun through a DC diode geometry [6], and the number of electrons per pulse is typically 800 - 1000. In the newly developed RF-compression ultrafast electron microscope (UEM), pulses with 10^5 - 10^7 electrons are available and accelerated to 100 keV through the Pierce-gun geometry, which facilitates high-current generation. Usually a solenoid magnetic lens sits behind the anode to focus the electron beam in the transverse direction. Because the electrons cannot travel in the air for a long distance, the sample holder, high-voltage electron gun and the imaging system are sealed inside a vacuum chamber, whose pressure is maintained at $\sim 10^{-7}$ - 10^{-9} torr. More details about the electron-gun geometry, vacuum and imaging systems of UED are available in Ref. [5]. The temporal resolution of the UED system is limited by the pulse duration of probe electron pulses. Due to strong Coulomb repulsion between electrons (space charge effects), the pulse duration of probe electron pulses broadens as they travel in free space. There are several ways to quantitatively measure the electron pulse duration. More details about the space charge effects in short electron pulses and the methods of measuring electron pulse-width are introduced in Chapter 5.

1.2 Diffraction Theory

To interpret electron diffraction phenomena, it is best to think of electrons as waves. In an optical double-slit or multi-slit experiment, the interference pattern of light can be observed in the far field when the slit size is comparable to the wavelength of light.

According to the Huygens-Fresnel principle, the center of each slit acts as a new source of a wave of the same wavelength. The sum of these secondary waves determines the form of the wave at the subsequent time and gives the interference effects. Similar to light, the propagating electron waves should also exhibit constructive and destructive interference, but only if slits can be found with size comparable to its wavelength. In nature, the lattice spacing in solids is typically $\sim 5 \text{ \AA}$, which serves as the ideal slit for electron waves.

Momentum conservation is a manifestation of translational symmetry. In the case of a periodic lattice, the translational invariance is discrete. So, the momentum of the electron waves and the atoms in the lattice must be conditional conserved. The discrete translational invariance is captured by the *reciprocal lattice* as the set of points that indicate the directions and spatial period of the translational invariance. When electron waves are scattered from the periodic lattice, the momentum conservation yields the result that the change of the electrons' momentum $\vec{s} = \vec{k}_f - \vec{k}_i$ should equal any discrete reciprocal lattice vector $\vec{G}$, which is also called the crystal momentum. This momentum conservation is mathematized in several different ways: Bragg's law, Laue diffraction conditions and Ewald sphere construction.

1.2.1 Bragg's Law

When the electron wave is scattered from a set of periodic lattice planes, the constructive interference condition is quantitatively described with Bragg's law,

$$2d \sin\left(\frac{\theta}{2}\right) = n\lambda, \quad (1.2.1)$$

where d is the interplanar distance, θ the scattering angle, n the index of diffraction order and λ the wavelength of electrons. The illustration of Bragg's law is depicted

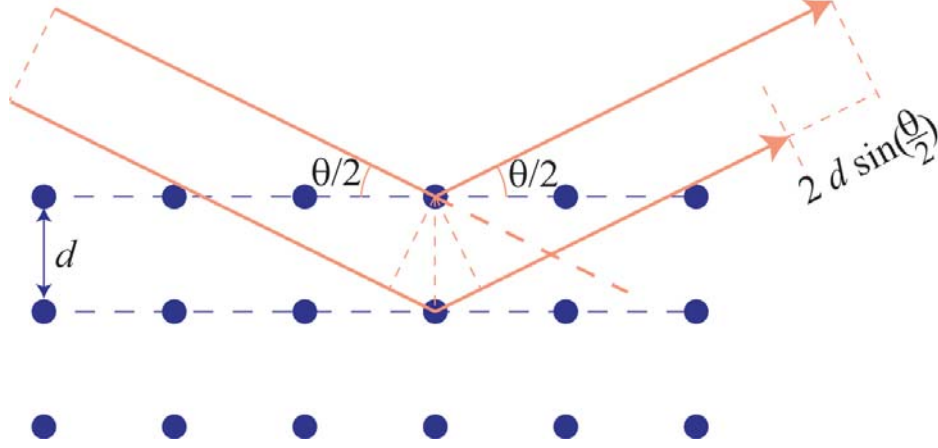


Figure 1.2: Bragg diffraction. Two rays of electrons approach the lattice with same wavelength, same phase and same incident angle. They are scattered by two different atoms in the same set of crystal planes. The lower beam travels an extra distance $2d \sin(\theta/2)$. Constructive interference occurs when this distance equals to an integer multiple of the wavelength.

in Fig. 1.2. The idea of Bragg's law is that the scattered electron waves interfere constructively, when the path difference between two waves equals an integer multiple of the wavelength.

1.2.2 Laue Diffraction Condition

Bragg's law is a very concise description of the diffraction condition and is widely used in the diffraction community. However, its idea of reflected electrons is physically wrong. More rigorously, Von Laue derived the diffraction conditions in three dimensions using the well-known light-optics approach,

$$a(\cos \theta_1 - \cos \theta_2) = h\lambda, \quad (1.2.2)$$

where a is the projected atomic distance in one direction, θ_1 the incident angle, θ_2 the exit angle and h an integer number. The cross-section view of the scattering condition in one dimension is plotted in Fig. 1.3(a). Similar to Bragg's law, the idea of Laue's equations is that the path difference between the electron waves scattered from

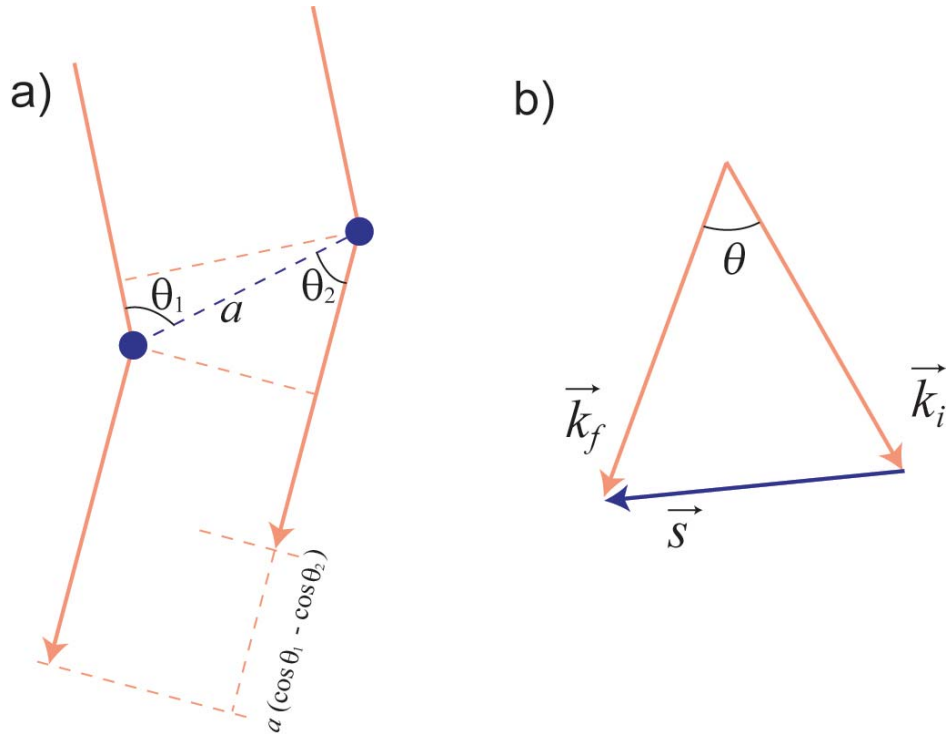


Figure 1.3: Laue diffraction conditions. Two rays of electrons with the same phase and wavevector approach the crystal with incident angle θ_1 . They are scattered by two atoms separated by distance a . The angle after scattering is θ_2 . The path difference of the two rays is $a(\cos\theta_1 - \cos\theta_2)$, which equals to an integer multiple of the wavelength, when the constructive interference occurs.

neighboring atoms is an integer multiple of the electron wavelength when a constructive interference is formed. Eqn. 1.2.2 can be generalized to three dimensions with two more distances, b and c , and appropriate angles θ_n ,

$$\begin{aligned} b(\cos \theta_3 - \cos \theta_4) &= k\lambda, \\ c(\cos \theta_5 - \cos \theta_6) &= l\lambda. \end{aligned} \tag{1.2.3}$$

Eqn. 1.2.2 and 1.2.3 are called *Laue diffraction conditions*. More usefully, the Laue conditions can be rewritten in a vector form:

$$\begin{aligned} \vec{a} \cdot \vec{s} &= 2\pi h, \\ \vec{b} \cdot \vec{s} &= 2\pi k, \\ \vec{c} \cdot \vec{s} &= 2\pi l, \end{aligned} \tag{1.2.4}$$

where $\vec{a}$, $\vec{b}$, $\vec{c}$ are the primitive vectors of the lattice, $\vec{s}$ the vectorial change of wavevectors of the incident beam ($\vec{k}_i$) and the scattered beam ($\vec{k}_f$), as shown in Fig. 1.3(b). We also call $\vec{s}$ the *momentum transfer* under some circumstances.

For a periodic lattice in real space, the primitive vectors in the reciprocal space are given by

$$\begin{aligned} \vec{a}^* &= 2\pi \frac{\vec{b} \times \vec{c}}{\vec{a} \cdot (\vec{b} \times \vec{c})}, \\ \vec{b}^* &= 2\pi \frac{\vec{c} \times \vec{a}}{\vec{b} \cdot (\vec{c} \times \vec{a})}, \\ \vec{c}^* &= 2\pi \frac{\vec{a} \times \vec{b}}{\vec{c} \cdot (\vec{a} \times \vec{b})}. \end{aligned} \tag{1.2.5}$$

So, we have

$$\begin{aligned} \vec{a} \cdot (h\vec{a}^*) &= 2\pi h, \\ \vec{b} \cdot (k\vec{b}^*) &= 2\pi k, \\ \vec{c} \cdot (l\vec{c}^*) &= 2\pi l. \end{aligned} \tag{1.2.6}$$

Comparing Eqn. 1.2.4 and 1.2.6, we can get

$$\vec{s} = h\vec{a}^* + k\vec{b}^* + l\vec{c}^* = \vec{G}, \quad (1.2.7)$$

where $\vec{G}$ is the linear combination of the primitive vectors of a reciprocal lattice and is called the ‘*reciprocal lattice vector*’. Here, diffraction-order indices h , k and l are equivalent to the *Miller indices* of the reciprocal lattice vector. Eqn. 1.2.7 is a direct depiction of the momentum conservation of the electron waves and the lattice atoms and it directly connects the diffraction pattern to the reciprocal lattice of a solid. In practice, the momentum transfer $\vec{s}$ is not available for direct measurement. However, its value can be measured through the scattering angle θ , as shown in Fig. 1.3(b). With the knowledge of θ , $\vec{s}$ is given by

$$\begin{aligned} |\vec{s}| &= 2k \sin\left(\frac{\theta}{2}\right) \\ &= \left(\frac{4\pi}{\lambda}\right) \sin\left(\frac{\theta}{2}\right). \end{aligned} \quad (1.2.8)$$

Here we used the fact that electron diffraction is an elastic scattering effect of electrons, thus, $|\vec{k}_f| = |\vec{k}_i| = k = 2\pi/\lambda$. Mathematically, it can be proven that the Laue diffraction conditions are equivalent to the Bragg’s law, which is a trivial extension and will not be discussed here.

1.2.3 Ewald Sphere Construction

With Bragg’s law (Eqn. 1.2.1) and Laue diffraction conditions (Eqn. 1.2.7), it is clear that a diffraction pattern is a projection of three-dimensional reciprocal lattice onto the direction normal to the incident electron beam. Thus, it can be used to derive the lattice structure in real space. It seems redundant to introduce a third equivalent formulation of diffraction. However, the concept of *Ewald sphere construction* is very powerful in many situations, especially for the solids with complex structures.

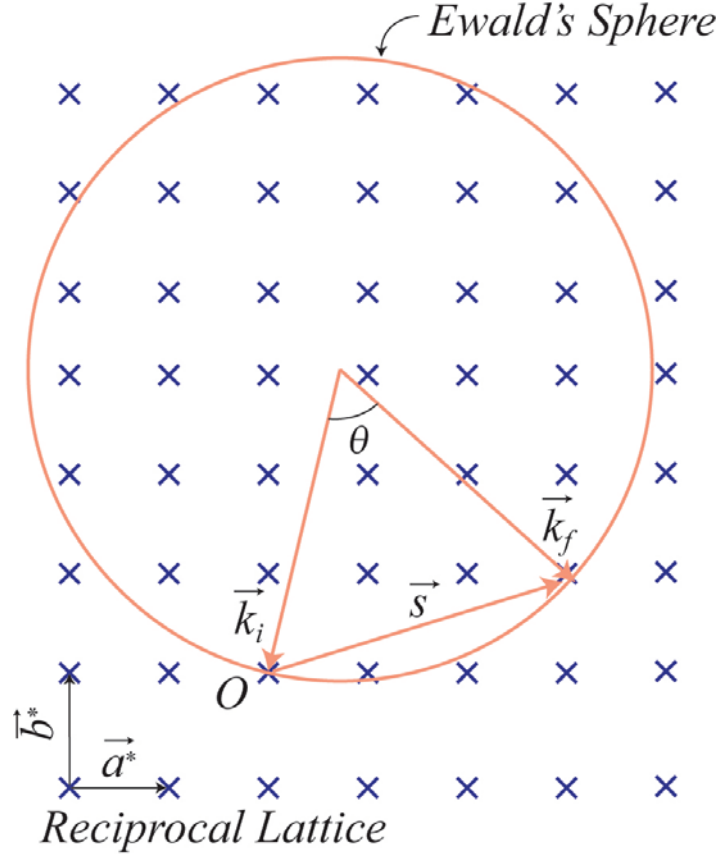


Figure 1.4: The Ewald sphere is shown intersecting the reciprocal lattice points. The reciprocal lattice vectors are $\vec{a}^*$ and $\vec{b}^*$. $\vec{k}_i$ is the wavevector of incident electrons, pointing to the origin of the reciprocal lattice O . The wavevector of scattered electrons is $\vec{k}_f$. The vector of momentum transfer is $\vec{s}$.

The Ewald sphere is constructed as follows:

1. Draw the incident wavevector in the reciprocal space, ending at the origin.
2. Construct a circle with radius $2\pi/\lambda$, which passes through the origin.
3. Whenever a reciprocal lattice point is located on the circle, it leaves a bright spot as the Bragg reflection.

In Fig. 1.4, the illustration of the Ewald sphere is depicted. $\vec{k}_i$ and $\vec{k}_f$ are the wavevectors of incident and diffracted wavevectors, respectively. For the reason of elastic scattering, $|\vec{k}_i| = |\vec{k}_f| = k$. Angle θ_B is the Bragg angle of diffraction, which holds the relation $2\theta_B = \theta$. Obviously, any reciprocal lattice vector sitting on the

Ewald sphere satisfies the Laue diffraction conditions in Eqn. 1.2.7. So, the Ewald sphere construction is equivalent to the Laue conditions and Bragg's law. Note that Fig. 1.4 is only a two-dimensional representation of the Ewald sphere. In practice, the Ewald sphere is a three-dimensional sphere, with its center on the origin of the incident wavevector and with a radius of $2\pi/\lambda$. From Fig. 1.4, it is also shown that the origin of the reciprocal lattice sits on the image of incident electron beam, which we call 'direct beam' under some circumstances.

1.2.4 Scattering from Isolated Atoms

From the formulations of Bragg's law, Laue diffraction conditions and Ewald sphere construction, the meaning of the diffraction pattern is clear and directly connected to the lattice structure in real space. In fact, the intensities of Bragg reflections in the diffraction pattern also carry important information regarding the atomic structure within the lattice unit cell, which is conventionally termed the *basis*.

To start with, let us first take a look at the case when electrons are scattered from isolated atoms. Because the wave-nature of electrons is important for considering interference effects, we use Huygens-Fresnel principle which takes the scattering atom as the source of the scattered spherical wave, as shown in Fig. 1.5. For simplicity, we first formulate the situation in one dimension where the scattered wave propagates along the same direction as the incident wave. The incident wave has the form $\Psi_i = \psi_0 \sin(kr)$. Since the scattering is elastic, we would expect that after scattering the amplitude of wave ψ_0 does not change, while a phase shift ϕ is induced. Hence, the scattered wave is given by

$$\begin{aligned}\Psi_{Tot} &= \psi_0 \sin(kr + \phi) \\ &= \psi_0 [\sin(kr) \cos(\phi) + \cos(kr) \sin(\phi)],\end{aligned}\tag{1.2.9}$$

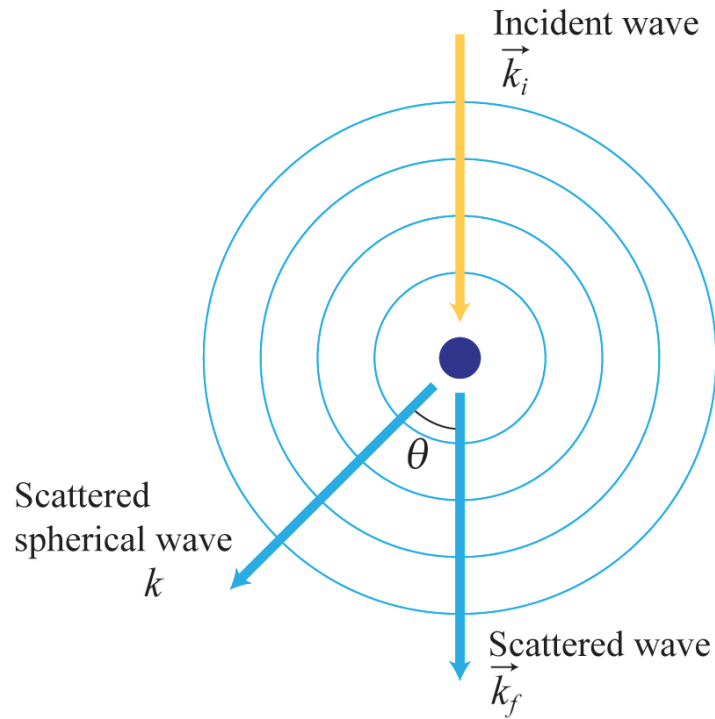


Figure 1.5: Scattering of the electron wave with an isolated atom. The incident electron wavevector is $\vec{k}_i$ and after scattering the wavevector becomes $\vec{k}_f$. k is the scalar wavevector of the scattered spherical wave. The circles represent the wave-front of scattered spherical waves.

where k is the wavevector, r the atomic coordinate. If ϕ is small, then

$$\Psi_{Tot} \approx \psi_0 \sin(kr) + \phi\psi_0 \sin(kr + \frac{\pi}{2}). \quad (1.2.10)$$

From Eqn. 1.2.10, we can find the physical meaning of Ψ_{Tot} . The first term on the right side of the Eqn. 1.2.10 is identical to the incident wave, while the second term has a reduced amplitude and a phase shift of $\pi/2$. In the weak-scattering limit, it can be expected that the majority of the incident wave continues to propagate along the original direction, while the remaining part contributes to the scattered wave. Ψ_{Tot} is simply the sum of the two. Therefore, the scattered wave is

$$\Psi_{Sc} = \phi\psi_0 \sin(kr + \frac{\pi}{2}). \quad (1.2.11)$$

The $\pi/2$ phase shift would arise if the imaginary identity i is multiplied to the exponential form. In addition, Eqn. 1.2.10 can be extended into a 3D spherical wave by considering scattered waves along different directions. When the scattering wave is not along the original propagating direction, Eqn. 1.2.11 still holds for the weak scattering limit. However, two notes need to be taken:

1. Momentum needs be conserved, so a momentum change of the atom can be expected;
2. The phase shift ϕ is expected to be dependent on the scattering angle θ , so the amplitude of scattered wave is also θ dependent.

For the first note, because of the huge mass difference between electrons and atoms, we would not expect significant movement of the atoms. For the second note, an atomic-scattering factor, or the *form factor*, $f(\theta)$, is defined to represent the angularly dependent ratio between the scattered-wave amplitude and the incident-wave ampli-

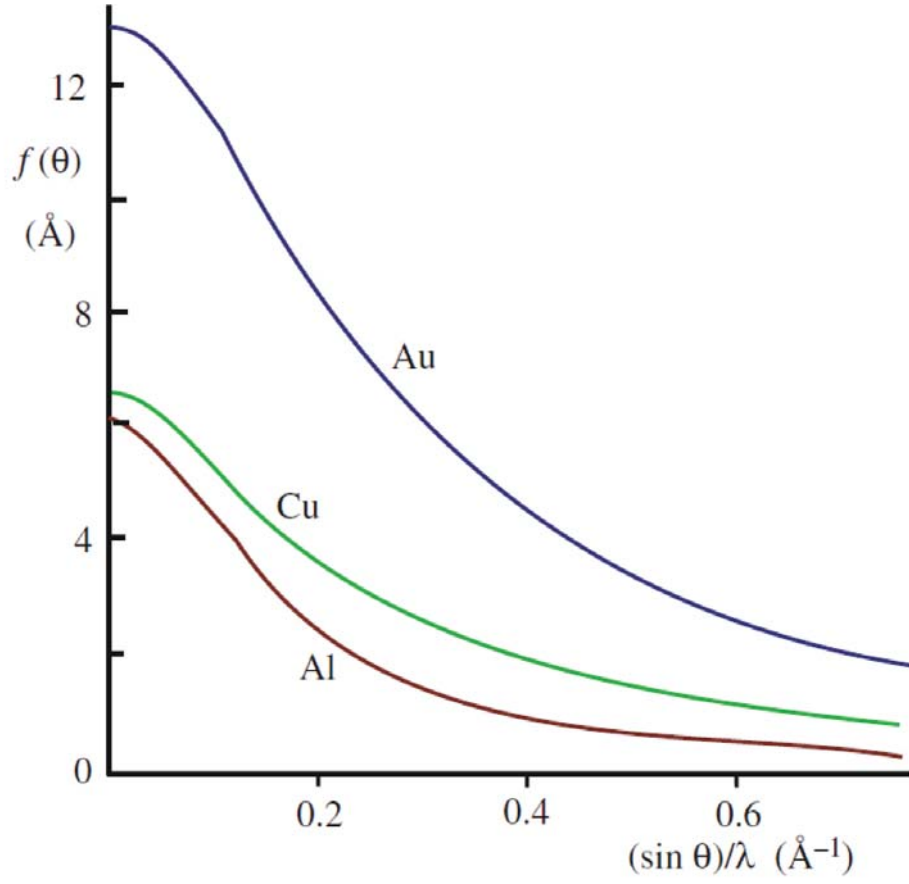


Figure 1.6: Change of form factor $f(\theta)$ as a function of scattering angle θ for several different metals. The strength of $f(\theta)$ decays with the increase of θ and increases with the atomic number Z . Figure from Ref. [7].

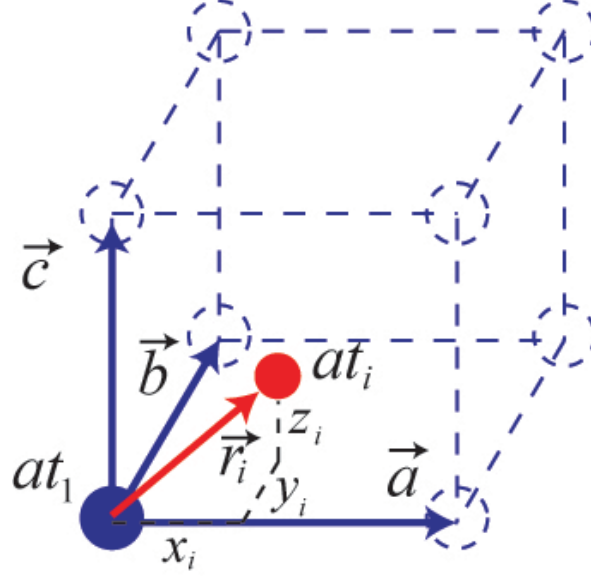


Figure 1.7: Illustration of crystal unit cell with two atoms in the basis, at_1 and at_i . The unit cell vectors are $\vec{a}$, $\vec{b}$ and $\vec{c}$. The origin is placed at at_1 and the coordinates of at_i are defined by the fractional coordinates x_i , y_i and z_i .

tude. Thus, the scattered wave is given by

$$\Psi_{Sc} = f(\theta)\psi_0 \frac{e^{ikr}}{r}, \quad (1.2.12)$$

and the total wave after scattering is

$$\Psi_{Tot} = \Psi_i + i\Psi_{Sc}. \quad (1.2.13)$$

According to Rutherford scattering, the amplitude of a scattered wave decays with increase of the scattering angle. Fig. 1.6 shows a typical decaying trend of $f(\theta)$, which is calculated using the Rutherford model [7]. For other atoms, the form factor can be found in Ref. [8] with a clear definition.

1.2.5 Structure Factor $F(\theta)$ and Intensity of Bragg Reflections

When the electrons are scattered from a lattice with a basis, the interference of electron waves scattered from different atoms in the basis becomes very important. For the discussion, we define a lattice with a simple basis of two atoms, at_1 and at_i , as shown in Fig. 1.7. The origin of the unit cell is placed at at_1 , and the unit cell vectors are defined as vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$. The location of the second atom at_i is defined by $\vec{r}_i = x_i\vec{a} + y_i\vec{b} + z_i\vec{c}$, where x_i , y_i and z_i are the fractional coordinates for at_i . Using the Laue diffraction conditions, the phase difference $\Delta\phi_i$ between the electron waves scattered on at_i and those scattered on at_1 can be calculated as

$$\begin{aligned}\Delta\phi_i &= \vec{r}_i \cdot \vec{s} \\ &= (x_i\vec{a} + y_i\vec{b} + z_i\vec{c}) \cdot \vec{s} \\ &= 2\pi(hx_i + ly_i + kz_i),\end{aligned}\tag{1.2.14}$$

where h, k, l are the Miller indices for a specific Bragg reflection. Thus, the contribution of at_i to the amplitude of the total scattered wave is

$$\psi_i(\theta_{hkl}) = f_i(\theta_{hkl})e^{i\Delta\phi_i} = f_i(\theta_{hkl})e^{2\pi i(hx_i + ly_i + kz_i)}.\tag{1.2.15}$$

Here, $f_i(\theta_{hkl})$ is the form factor of at_i , which was defined in the previous section. Eqn. 1.2.15 can be extended to the situation for a lattice with a basis of an arbitrary number of atoms by simply summing the contributions from all the atoms, giving the definition of the *structure factor* $F(\theta)$

$$\begin{aligned}F(\theta_{hkl}) &= \sum_i \psi_i(\theta_{hkl}) \\ &= \sum_i f_i(\theta_{hkl})e^{2\pi i(hx_i + ly_i + kz_i)}.\end{aligned}\tag{1.2.16}$$

The intensity of the Bragg reflection with Miller indices $[hkl]$ is given by

$$I_{hkl} = |F(\theta_{hkl})|^2. \quad (1.2.17)$$

We need to note that the calculation of intensity ignores the dynamical scattering effects, so it is also called the ‘*kinematic theory of diffraction.*’ Because electrons strongly interact with materials, the dynamical scattering effects can be very important. An illustration of dynamical scattering effects is depicted in Fig. 1.8. The diffracted electron beam from crystal plane $[hkl]$ can be re-diffracted by a second set of planes $[h'k'l']$ in the same specimen, which causes redistribution of intensities among Bragg reflections with different Miller indices. So, using the kinematic diffraction theory, only qualitative agreement with experimental data can be expected. The dynamical scattering effect is a very complicated problem in electron diffraction. More rigorous mathematical treatments can be found in Refs. [7, 9–12].

1.3 Current Status of Ultrafast Electron Diffraction and Microscopy Techniques

With the recent commissioning of the first hard X-ray free electron laser (X-FEL) at LCLS single-shot, femtosecond X-ray diffraction experiments are now in principle possible [13]. Compared to X-ray techniques, electron techniques own unique advantages:

1. Electrons interact with materials ~ 1000 times stronger than X-ray photons;
2. UED (or UEM) instruments are of table-top scale;
3. Electron beam optics are easier to implement and more flexible compared to X-ray;

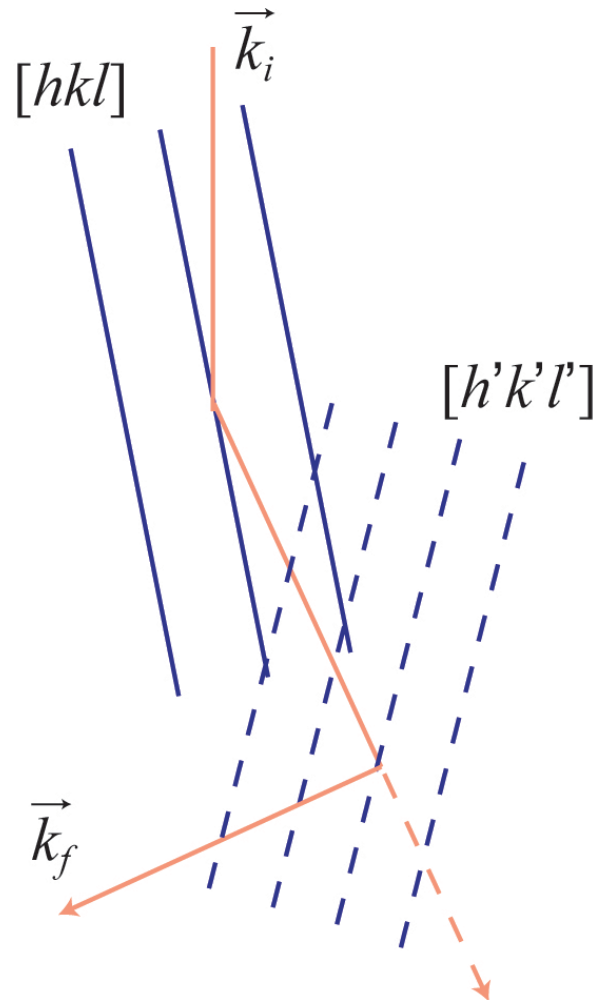


Figure 1.8: Illustration of dynamic scattering of electrons. The electron beam originally scattered from planes $[hkl]$ is scattered for the second time from planes $[h'k'l']$, which causes redistribution of intensities of Bragg reflections.

4. Penetration depth of high-energy electrons better matches with that of optical pumps.

For these reasons, ultrafast electron diffraction and microscopy is fast growing in recent years. Technology is no longer limited to the first-generation UED systems. Systems with better temporal resolution and higher brightness are preferred. The fundamental limit of the temporal resolution of the UED or UEM with high-brightness sources is caused by the strong Coulomb repulsion between electrons (*space charge effects*) inside the high-density pulses. To circumvent the space charge effects, several approaches have been attempted at different places of the world, which are going to be introduced in this section.

1.3.1 Compact Geometry of Electron Gun

The first approach to avoid space-charge broadening is to design a compact electron-gun geometry and reduce the distance from sample to photocathode. By reducing the sample-to-cathode distance, the space-charge effects on temporal resolution at sample location can be significantly reduced [14]. The sample-to-cathode distance in a typical UED system is usually limited by the size of magnetic lens mounted behind the electron-gun anode. M. Harb, et al. from Dwayne Miller's group at University of Toronto designed the electron-gun geometry with cathode-anode assembly enclosed within the magnetic lens iron shielding, which reduces the sample-to-cathode distance ~ 3 cm [15]. An optimized temporal resolution ~ 300 fs with 10^4 electrons per pulse is achieved [15, 16]. In parallel, C. Gerbig et al. at University of Oldenburg designed the UED system with a magnetic lens located behind the sample to focus the diffracted electron beams [17]. By doing so, the sample-to-cathode distance is further reduced to 8.5 mm and a temporal resolution ~ 150 fs with an available 5000 electrons-per-pulse.

By applying the compact electron-gun geometry, a very good temporal resolution can be achieved with a moderately high amount of electrons in individual pulses. The

advantage of this approach is that it is easy to implement. However, it also suffers from the limited space between the electron gun and the sample. First, the strength of the magnetic lens, if it is in between the anode and the sample, is quite limited, which makes the electron beam size not less than $100\text{ }\mu\text{m}$ on sample. Its application to nanometer-scale materials is not possible. Secondly, the sample surface, where the electron probe is incident, becomes inaccessible to the optical pump, due to the short distance between the sample surface and the anode. Usually a back-illumination geometry of the pump laser has to be implemented, which can cause problems in the penetration depth and the velocity mismatch between the pump laser and the probe electron pulses.

1.3.2 UEM with Single Electron Per Pulse

The second approach to avoid the space-charge expansion is to reduce the number of electrons in a pulse to approximately one electron, while increasing the repetition frequency to tens of MHz [18]. Based on this idea, Zewail’s group at Caltech first built an ultrafast electron microscope (UEM) with single electron per pulse [19, 20]. In their system, the temporal resolution is fundamentally limited by the timing jitter of the photoelectric effect. It is shown that the temporal resolution can be as short as $\sim 70\text{ fs}$, by optimizing the excitation photon energies [21]. Simulations also show that by incorporating an RF-gun for single electron photoemission, the individual electrons can arrive at the sample plane within a time-window less than a fs [22].

There are several advantages to this approach. First, with a single electron per pulse, the conventional electron optics for electron microscopes can be directly used. The UEM system of Caltech is based on a conventional transmission electron microscope (120-kV G2 12 TWIN Tecnai TEM) [19], which is modified for the optical accesses of the optical pump laser and the UV laser. Minor efforts on electron beam alignment are required to achieve diffraction and image functions with original TEM

electron optics. Thus, a spatial resolution of Angström scale can be achieved. The second advantage, which is equally important, is the very small energy spread of single electron pulses. In a conventional TEM, the energy spread of the electron beam is typically in the range 0.3 - 3 eV (at 100 keV total energy), which is very important for the application of electron energy spectroscopy, for example the electron energy-loss spectroscopy (EELS). With multiple electrons per pulse, the space charge effects not only broaden the pulse width, but also increase the energy spread of the electron pulses. The energy spread $\Delta E \sim 90$ eV is expected for 10^4 electrons per pulse [14]. In the single electron UEM, the energy spread reaches the fundamental limit of photoelectric effect and the energy-bandwidth of optical excitation. The energy spread of 0.87 eV has been reported, with which the ultrafast energy-loss spectroscopy becomes possible [23, 24]. However, this technique requires that the sample can be reproducibly pumped and probed for $\sim 10^6$ times to obtain diffraction patterns with sufficient quality, which limits the material systems available for studies.

1.3.3 Dynamic TEM

Another development is focused on obtaining electron diffraction or image patterns using a single shot of electron pulse. According to the Rose criteria [25], at least 100 electrons per pixel on the camera are required to achieve a sufficient signal-to-noise ratio (SNR) for an image or a diffraction pattern. For a typical 1000×1000 pixel camera, it means at least 10^8 electrons in a single pulse for a full image and $\sim 10^6$ electrons for a diffraction pattern. With so many electrons in a single pulse, the pulse duration is broadened to 30 - 100 ns after traveling through a conventional TEM column, which is the temporal resolution for its application. The single-shot approach was pioneered by Dömer and Bostanjoglo [26]. They demonstrated a spatial resolution of ~ 200 nm and a temporal resolution of ~ 10 ns with $\sim 10^8$ electrons per pulse. The spatial resolution was further improved by the research group from Lawrence Livermore

National Laboratory to ~ 1 nm with 10^7 - 10^8 electrons per pulse and the temporal resolution of ~ 10 ns [27–29]. More recently, a ‘movie mode’ has been implemented in their dynamic TEM (DTEM). In the movie mode, a single pump excitation is followed by a sequence of electron probes with well-defined time delay. After being scattered from the sample, the electrons are deflected and project images with sequential time delays on the CCD camera at the same time [30]. Movie-mode DTEM enables research on irreversible processes, which is not available with UED techniques, because it does not require repeating pump-probe cycles to get structural dynamics at different time delays. However, its time resolution is limited to 10 ns. Thus, the initial structural dynamics in ps or even fs time scale are not able to be studied, which is especially important for the materials whose electronic degrees of freedom play important roles in structural dynamics [31, 32].

1.3.4 UEM with MeV Energy

In traditional accelerator physics, it is known that the space charge effects are significantly reduced, when charged particles are accelerated to relativistic velocity. So, another way to circumvent the space-charge broadening is to increase electron energy to MeV scale [33]. In a MeV UED system, a photocathode RF gun is equipped as the source, which is able to provide an extraction field up to 100 MV/m. The photoelectrons are hence accelerated to 1 - 3 MeV energy in a short acceleration gap. Currently, several groups in the world are making experimental efforts to demonstrate and improve the MeV UED systems, including SLAC [34], Pohang Accelerator Lab at POSTECH [35], Pegasus Lab at UCLA [36, 37], Brookhaven National Laboratory [38], Tsinghua University of China [39], and Osaka University [40]. A typical temporal resolution of ~ 200 fs can be achieved with 10^7 to 10^8 electrons per pulse.

Space charge effects can be significantly reduced with relativistic electron pulses. The pulse duration can be preserved at the sub-ps scale, even when there are 10^8

electrons in the single pulse. At the same time, the shorter-wavelength electrons can, in principle, lead to better spatial resolution. All these are the advantages of MeV UED or UEM systems. However, the very high energy of electrons causes several problems when they are used for material science research, the most serious of which is the knock-on damage to the sample [7]. The energy that high energy electrons transfer to atoms increases with the increase of incident electron energy. For most light elements, like C, N, O, etc., the transferred energy to the atoms can be as high as 100 eV, when the incident electron energy is higher than 400 keV. This value is higher than the displacement energy E_d of the atoms, which is typically 25 eV only. In other words, the atoms can be knocked out of the lattice by the high-energy electrons. That is also the reason that material scientists prefer typical 100 - 200 keV electrons for an electron microscope. The second limitation is the electron optics. When the electron energy is very high, the electron optics have to provide much higher focusing power, which leads to bulky optical components with less flexibility.

1.4 RF-Compression Ultrafast Electron Microscope at Michigan State University

At Michigan State University, the second-generation of the ultrafast electron diffraction and microscopy system is under construction and characterization. The approach we implemented is to re-compress the electron pulses, which already suffer from the space-charge broadening, back to short pulses at the sample plane using an RF cavity as the compressor. The technique of RF compression is widely applied in accelerator physics [41], and was previously applied to UED systems as well [42, 43]. In our system, there is an additional strong magnetic lens after the RF cavity, which focuses the electron beam after temporal compression, to achieve the spatial and temporal focusing at the same time. Another set of projection lenses are going to be equipped after the sample

in the future, for the purpose of image formation.

The mechanism of RF compression is to decelerate the fast moving electrons, while accelerating the slow moving electrons, by applying an oscillating longitudinal electric field on the propagating electron pulse and reversing its momentum-space correlation in the phase space. More details about the mechanism of RF compression are introduced in Chapter 6. In this thesis, the major problems of RF synchronization and RF signal amplification will be introduced, while the details for the vacuum systems, the high-voltage electron gun and the electrons optics can be found in Ref. [44].

Chapter 2

Correlated-Electron Materials

Materials with correlated electrons exhibit some most intriguing phenomena in condensed matter physics. In this chapter, a general introduction on the strongly correlated materials is given in both theoretical and experimental aspects. In Section 2.1, we start with the fundamental theory of condensed matter physics, band theory, which is based on the single-electron approximation. The band theory fails, when electron correlation is too strong to take as a perturbation. In Section 2.2, we explain the notion of strongly correlated electrons and why the band theory fails. A simple, but profound, model, the one-band Hubbard model, is introduced as well. Many numerical methods based on the Hubbard model have been developed to treat the strongly correlated electrons. Two methods, DFT LDA+ U and DMFT are briefly introduced in Section 2.3, which helps in understanding some of the discussions in the following chapters. Strongly correlated materials are not only interesting due to the fundamental physical questions they pose to condensed matter physicists, but also many exotic emergent phenomena they exhibit. In Section 2.4, some of these phenomena are introduced, including metal-to-insulator transition, colossal magnetoresistance, high-temperature superconductivity.

2.1 Band Theory of Solids

Band theory is one of the most successful theories of condensed matter physics because of its breadth and subtlety of the phenomena it explains. The fundamental assumption of band theory is the *independent electron (quasiparticle) approximation*, in which the electron wavefunction obeys the single-electron Schrödinger equation,

$$H\psi = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\right)\psi = E\psi, \quad (2.1.1)$$

where V is the potential energy, which carries the same periodicity as the underlying Bravais lattice.

2.1.1 Bloch's Theorem

Bloch's Theorem tells us that when V is a periodic potential of an ideal lattice, the eigenstate ψ has the form of a plane wave with a periodic amplitude modulation, and its periodicity is the same as that of the Bravais lattice

$$\psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}). \quad (2.1.2)$$

The periodic amplitude modulation $u_{n\vec{k}}(\vec{r})$ is given by

$$u_{n\vec{k}}(\vec{r} + \vec{R}) = u_{n\vec{k}}(\vec{r}), \quad (2.1.3)$$

where $\vec{R}$ is the Bravais lattice vector, n and $\vec{k}$ stand for band index and wave vector, respectively. Eqn. 2.1.2 and 2.1.3 imply that

$$\psi_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}} \psi_{n\vec{k}}(\vec{r}). \quad (2.1.4)$$

So, the wave vector $\vec{k}$ in Bloch's theorem can always be confined to the *first Brillouin zone* (BZ) in reciprocal space, because for any $\vec{k}'$ not in the first Brillouin zone,

$$\vec{k}' = \vec{k} + \vec{G}, \quad (2.1.5)$$

where $\vec{G}$ is the reciprocal lattice vector and $\vec{k}$ lies in the 1st BZ. Since $e^{i\vec{G}\cdot\vec{R}} = 1$, the Bloch wave of $\vec{k}'$ has the same form as that of $\vec{k}$.

The potential V in Eqn. 2.1.1 is attributed to the electron-ion interaction. Depending on the strength of V , Eqn. 2.1.1 can be solved with two different methods.

2.1.2 Nearly-Free Electron Model and Tight-Binding Approximation

Here, general features and important conclusions from these two methods are going to be introduced. More detailed derivations can be found in Ref. [45–47].

When the electron-ion interaction is very weak, the eigenstate ψ of Eqn. 2.1.1 can be expanded with plane waves, which are the solutions to Schrödinger equation of free electrons. Since Bloch's theorem tells us that the Bloch state at crystal momentum $\vec{k} + \vec{G}$ is a replica of Bloch state at $\vec{k}$, where $\vec{k}$ lies in the 1st BZ, we would expect the parabolic energy-momentum dependence to repeat in the reciprocal space at every $\vec{G}$ point. The second-order perturbation theory opens band gaps at the boundaries of BZs, because of the degeneracies. In Fig. 2.1, the free electron E vs. k parabola in one dimension is plotted in a repeated-zone scheme. The distortions of the parabola can be observed in the neighborhood of the BZ boundaries, which break the degeneracies. We call the energy continuum between gaps the *energy bands*. $2n|V_G|$ is the size of the band gap between n^{th} and $(n+1)^{\text{th}}$ bands, where

$$V_G = \langle \vec{k} | H | \vec{k} - \vec{G} \rangle. \quad (2.1.6)$$

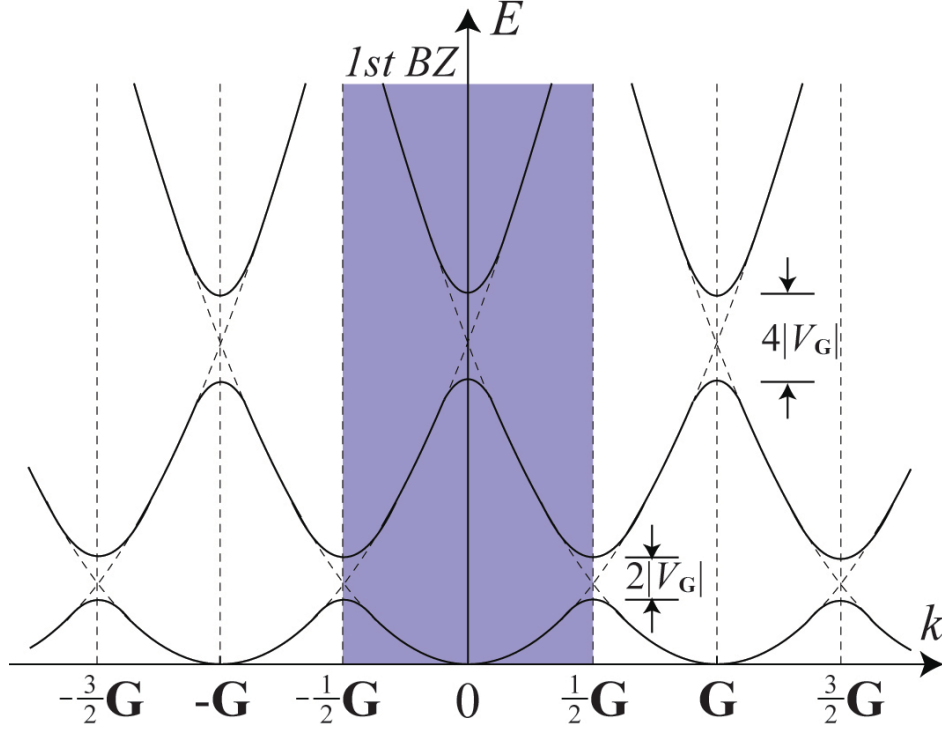


Figure 2.1: Free electron levels in a periodic potential depicted in a repeated-zone scheme.

For three-dimensional crystals, the electronic structure is more complicated, but the general features are captured by Fig. 2.1. In each band, the number of available electron states is $2N$, where N is the number of atoms in the solid, and the factor 2 counts for the two spin states of electrons. Due to the Pauli exclusion principle, electrons cannot exist in identical energy states. As the result, at absolute zero, when they are packed into the lowest available energy states, the energy of the highest occupied states is finite and is defined as the *Fermi surface*. We can imagine this process as filling balls (electrons) into drawers (states), and every drawer can contain two balls. When the Fermi surface lies in the middle of a band, which means this band is half-filled, the electrons at the Fermi surface can be driven to other empty states in the neighboring sites by applying an electric field, so the material appears to be a metal. On the other hand, when the Fermi surface lies in a gap, the bands are either fully filled or completely empty, and electrons do not have empty sites to move to, so the

material is an insulator or semiconductor. The semimetal happens when the Fermi surface lies where the density of states is zero, but there is no energy gap. A good example is at the crossing point of the Dirac-cone in graphene.

The properties of metals are well described by the nearly free electron model, including elements in groups I, II, III and IV of the periodic table [45]. Part of the reason is that the electron-ion interaction is screened by redundant conduction electrons, which legitimizes the condition of a weak ionic potential V . For describing the electronic structure of insulators, V can no longer be taken as a perturbation, so that another approach, the *tight-binding approximation* is applied. In contrast to the nearly-free electron model, the tight-binding approximation starts with localized atomic wavefunctions, $\varphi_n(\vec{r})$, where the ionic potential is fully considered in the Hamiltonian, namely $\varphi_n(\vec{r})$ is the eigenstate of the atomic Hamiltonian H_{at}

$$H_{at}\varphi_n(\vec{r}) = E_n\varphi_n(\vec{r}). \quad (2.1.7)$$

Considering the interaction of neighboring atoms, corrections need to be made on H_{at} to give the correct crystal Hamiltonian

$$H_{cryst} = H_{at} + \Delta V, \quad (2.1.8)$$

where ΔV is the correction term. Thus, the crystal Schrödinger equation becomes

$$H_{cryst}\psi_{\vec{k}}(\vec{r}) = E_{\vec{k}}\psi_{\vec{k}}(\vec{r}). \quad (2.1.9)$$

Different from the nearly-free electron model, the eigenstates are expanded using localized atomic wavefunctions in the tight-binding approximation. Because of this, the tight-binding approximation is also called ‘*Linear combination of atomic orbitals*’

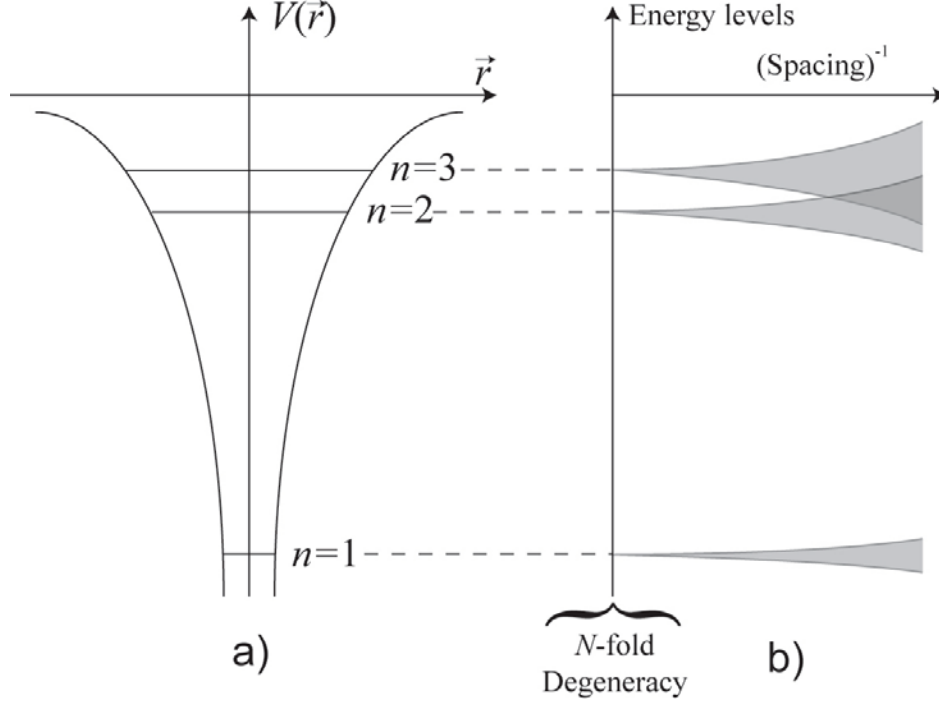


Figure 2.2: a) Schematic representation of non-degenerate electronic levels in an atomic potential; b) Energy levels of N such atoms in a periodic array, plotted as a function of mean inverse interatomic spacing. When the atoms are far apart (small overlap integrals), the levels are nearly degenerate, but when the atoms are closer together (larger overlap integrals), the levels broaden into bands. Figure from Ref. [45].

(LCAO) theory.

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \phi(\vec{r} - \vec{R}), \quad (2.1.10)$$

where $\phi(\vec{r} - \vec{R})$ is called *Wannier function*. It can be proven that the eigenstates of Eqn. 2.1.10 satisfy the Bloch's theorem. The Wannier function can be further expanded with the localized atomic wavefunctions $\varphi_n(\vec{r})$,

$$\phi(\vec{r}) = \sum_n b_n \varphi_n(\vec{r}). \quad (2.1.11)$$

In a diatomic molecule, like H_2 molecules, due to the linear combination of the atomic orbitals, the original atomic orbitals are split into *bonding* and *antibonding* orbitals. This picture can be generalized for the case of a crystal with N atoms. When

the N atoms are far apart, the electrons simply reside on any one of the N isolated atoms. We would expect the energies of electrons to be N -fold degenerate, if the atoms are identical. When the atoms are brought closer, interactions from the neighboring atoms start to split the electronic energy levels and break the degeneracy, which is similar to what happens in a diatomic molecule. Here, the total number of states are conserved, but their energies spread into broad bands as shown in Fig. 2.2. This is a very crude description, but it catches the essential features of the tight-binding energy bands. Obviously, the width of energy bands depends on the extent of overlap between atoms. For the case where only the s orbitals of atoms are considered, the energy-momentum relation can be quantitatively given by [45]

$$E_{\vec{k}} = E_s - \frac{\beta + \sum \gamma(\vec{R}) e^{i\vec{k} \cdot \vec{R}}}{1 + \sum \alpha(\vec{R}) e^{i\vec{k} \cdot \vec{R}}}, \quad (2.1.12)$$

where E_s is the original energy of atomic s orbital, and factors β , $\alpha(\vec{R})$ and $\gamma(\vec{R})$ are defined as follows

$$\beta = - \int \varphi_s^*(\vec{r}) \Delta V(\vec{r}) \varphi_s(\vec{r}) d\vec{r}, \quad (2.1.13)$$

$$\alpha(\vec{R}) = \int \varphi_s^*(\vec{r}) \varphi_s(\vec{r} - \vec{R}) d\vec{r}, \quad (2.1.14)$$

and

$$\gamma(\vec{R}) = \int \varphi_s^*(\vec{r}) \Delta V(\vec{r}) \varphi_s(\vec{r} - \vec{R}) d\vec{r}. \quad (2.1.15)$$

Because α is usually much less than 1, Eqn. 2.1.12 can be further simplified to

$$E_{\vec{k}} = E_s - (\beta + \sum \gamma(\vec{R}) e^{i\vec{k} \cdot \vec{R}}). \quad (2.1.16)$$

Eqn. 2.1.16 reveals that the single atomic energy level E_s is expanded into a band, when the periodic ionic potential strongly affects the electronic states. The bandwidth - i.e.

the spread between the minimum and maximum energies in the band - is proportional to the overlap integral γ , which we call the ‘*hopping integral*’. Since only the s orbitals are considered, there is only a single s band in the derivation above. We would expect similar situations happen to other atomic orbitals, when the hopping integral $\gamma(\vec{R})$ is nonzero. Thus, multiple bands in the electronic structure can be obtained with the tight-binding approximation, which is similar to the picture provided by the nearly-free electron model.

The Fermi surface and electron filling picture discussed for the nearly-free electron model is also valid for the tight-binding theory. Furthermore, the tight-binding approximation provides us with clearer understanding of the origin of each energy band. A simple argument about whether the material is a conductor or an insulator can be made based on the number of valence electrons per unit cell. If this number is even, every orbital, if it is filled by electrons, is going to be fully filled. The same is true for the energy bands. Therefore, the material is expected to be an insulator if the bands do not overlap. On the other hand, if this number is odd, there will always be a half-filled band and the material is always a metal.

2.1.3 Hartree-Fock Theory and Fermi Liquid Theory

In both the nearly-free electron model and the tight-binding approximation, the electrons are treated as independent particles and only the ionic potential is considered in the Hamiltonians. Obviously, it is a very crude approximation, because the electron-electron interaction is completely ignored. Beyond the independent electron approximation, many efforts have been made to include the electron-electron interaction into the model, among which the Hartree-Fock approximation is a very important theory.

The Hartree-Fock equation is given by

$$\begin{aligned}
H_{\text{HF}}\psi_i(\vec{r}) = & \\
& -\frac{\hbar^2}{2m}\nabla^2\psi_i(\vec{r}) + V^{\text{ion}}(\vec{r})\psi_i(\vec{r}) + V^{\text{el}}(\vec{r})\psi_i(\vec{r}) \\
& -\sum_j \int d\vec{r}' \frac{e^2}{|\vec{r}-\vec{r}'|} \psi_j^*(\vec{r}')\psi_i(\vec{r}')\psi_j(\vec{r})\delta_{s_i s_j} = E_i\psi_i(\vec{r}),
\end{aligned} \tag{2.1.17}$$

where H_{HF} is the Hartree-Fock Hamiltonian, V^{ion} the ionic potential

$$V^{\text{ion}}(\vec{r}) = -Ze^2 \sum_R \frac{1}{|\vec{r}-\vec{R}|}, \tag{2.1.18}$$

and V^{el} represents the potential that electrons experience from all the other electrons. Here, the rest of electrons are treated as a smooth distribution of negative charges with density ρ , so the potential energy is given by

$$V^{\text{el}} = -e \int d\vec{r}' \rho(\vec{r}') \frac{1}{|\vec{r}-\vec{r}'|}, \tag{2.1.19}$$

where $\rho(\vec{r}) = -e \sum_i |\psi_i(\vec{r})|^2$. In addition, the third term in Eqn. 2.1.17 is the *exchange energy*, which directly takes into account the antisymmetry of the electron wavefunction. In the Hartree-Fock approximation, instead of dealing with the interactions between single electrons, the interaction potential is treated in a mean-field manner, as if the electron of interest travels in the averaged field from other electrons with a smooth density $\rho(\vec{r})$. Thus, the Hartree-Fock equation is fundamentally a single-electron Schrödinger equation. In practice, Eqn. 2.1.17 can be solved in an iterative way to find the self-consistent solution of the wavefunction $\psi_i(\vec{r})$.

The single-electron approximation is further generalized by Landau through Fermi Liquid Theory [48]. Fermi Liquid Theory is a phenomenological model about the qualitative effects of electron-electron interaction on the electronic structure. In Landau's theory, even though the electron-electron interaction is strong and the independent

electron approximation is, in principle, far from a good approximation for the problem, the electrons can still be considered as Fermions *dressed* with the electron-electron interaction field, which are termed ‘*quasiparticles*’. If the quasiparticles are Fermions and obey the Pauli exclusion principle, the arguments used in discussing the independent electron approximations still work on them. Landau’s theory implies that the excitation spectrum (electronic band structure) for the interacting system, though numerically different from the non-interacting system, nevertheless has a structure of free-electron type.

2.1.4 Density Functional Theory

Density Functional Theory (DFT) is one of the most widely used methods to calculate band structures of solids [49]. Compared to the Hartree-Fock method, it is able to treat many problems in high accuracy, while it requires less-expensive calculations. In the Hartree-Fock theory, the total averaged energy of the state is determined by its wavefunction

$$E(\Psi) = \langle \Psi | H_{\text{HF}} | \Psi \rangle, \quad (2.1.20)$$

where $\Psi = \Psi(\vec{r}_1, \vec{r}_2, \vec{r}_3 \dots \vec{r}_N)$ is the N -electron wavefunction. Here, Eqn. 2.1.20 implies that the energy E is a *functional* of the wavefunction Ψ . According to the *variational theorem*, the energy is higher than ground state energy E_0 ,

$$E(\Psi) \geq E_0, \quad (2.1.21)$$

unless the wavefunction corresponds to the ground state wavefunction $\Psi = \Psi_0$. In principle, the ground-state wavefunction can be found by starting with a set of trial bases and iteratively searching for the possible combinations that minimizes the total energy. However, to describe the spatial variation of the wavefunction, a large set of bases is usually required, which makes its computational cost formidably high, even

for a moderately complex problem.

Hohenberg and Kohn first put forward and proved two theorems (Hohenberg-Kohn Theorem) [50], which makes it possible to solve this difficult problem. The Hohenberg-Kohn Theorems are stated as follows:

1. **Theorem I:** For any system of interacting particles in an external potential $V_{\text{ext}}(\vec{r})$, the density is uniquely determined (in other words, the external potential is a unique functional of the density);
2. **Theorem II:** A universal functional for the energy $E(n)$ can be defined in terms of the density. The exact ground state is the global minimum value of this functional.

Theorem I states that the energy is a functional of the charge density - $E[\rho]$. Here, notation ‘[]’ means the relation of being a functional. **Theorem II** establishes the variational principle: $E[\rho] \geq E_0$. Utilizing the Hohenberg-Kohn Theorem, the ground state energy can be found through minimizing the total energy by changing the spatially dependent electron density, instead of changing a complex N -electron wavefunction. This method saves dramatically on its computational cost. The starting point for DFT is the Kohn-Sham equations [51]

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{KS}}\right)\psi_i(\vec{r}) = E_i\psi_i(\vec{r}), \quad (2.1.22)$$

which has the form of the one-particle Schrödinger equations. V_{KS} is the Kohn-Sham potential, which is a functional of the electron density $\rho(\vec{r})$

$$\begin{aligned} V_{\text{KS}}[\rho(\vec{r})] &= V_{\text{ion}} + V_{\text{ee}}[\rho(\vec{r})] + V_{\text{Ex}}[\rho(\vec{r})] \\ &= V_{\text{ion}} + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \frac{\delta E_{\text{Ex}}[\rho(\vec{r})]}{\delta[\rho(\vec{r})]}. \end{aligned} \quad (2.1.23)$$

The density of electrons is determined by the wavefunction $\psi_i(\vec{r})$,

$$\rho(\vec{r}) = \sum_i |\psi_i(\vec{r})|^2. \quad (2.1.24)$$

Compared to the Hartree-Fock equation (Eqn. 2.1.17), the Kohn-Sham equations have exactly the same structure. V_{ee} stands for the electron-electron Coulomb interaction, while V_{Ex} is the exchange-correlation energy. For V_{ee} , a similar mean-field approximation as in the Hartree-Fock equation is applied, assuming that the electron of interest travels in the averaged field generated by a smoothly distributed charge density. The functional-relation of the exchange-correlation energy V_{Ex} is not explicitly clear, but approximations can be made to get the functional relation, based on Thomas-Fermi's early studies on homogeneous electron gases, which are called the *local density approximations* (LDA).

Using DFT, the electronic structure is calculated iteratively to achieve self-consistent convergence on the electron density and the total energy. It starts from a test electron density $\rho_t(\vec{r})$ and the corresponding Kohn-Sham potential is calculated using the functionals in Eqn. 2.1.23. The Kohn-Sham equation is then diagonalized to get the electron wavefunction $\psi_i(\vec{r})$. The electron density distribution $\rho(\vec{r})$ is recalculated using Eqn. 2.1.24. The calculation will continue using the new density $\rho(\vec{r})$, until convergence is achieved for the electron density. Fig. 2.3 shows the flow chart of a typical DFT algorithm.

2.2 Strongly Correlated Electrons

In the Hartree-Fock approximation and DFT, the electron-electron interactions are included in a mean-field and perturbative manner to calculate the band structure of solids. These two theories have achieved quantitative agreement with experiments on many materials, especially using DFT, which has been the most favored calculation

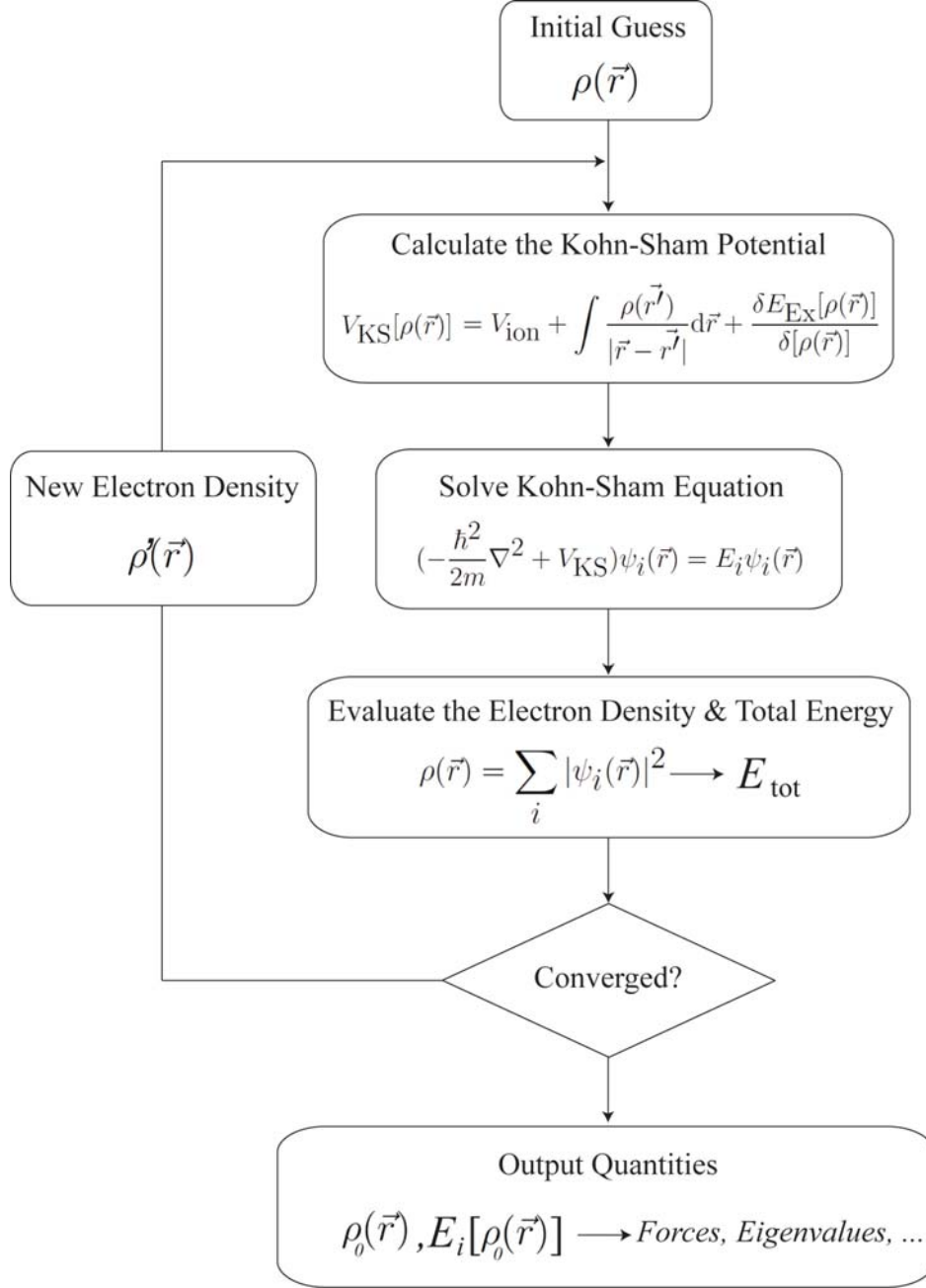


Figure 2.3: Flow chart of a typical DFT algorithm.

method by condensed matter physicists. As we have introduced, these methods are still based on the independent-electron approximations, which is legitimized by Landau's Fermi Liquid Theory. In his theory, the strong electron-electron interaction only modifies the properties of *quasiparticles*, including their energy, effective mass, etc., but the fundamental free-electron-like band structure is not changed.

However, the independent-electron approximation and DFT method are not accurate enough, when they are applied to strongly correlated materials. The failure of the band theory was first noticed in insulators, such as nickel oxide. These materials have an odd number of electrons per unit cell, which are always metals in band theory (see the discussion in Section 2.1.2) [52]. Mott and Peierls first pointed out the importance of the electron-electron correlation in these materials: strong Coulomb repulsion between electrons could be the origin of the insulating behavior [53]. In the discussion, Peierls noted the following:

‘It is quite possible that the electrostatic interaction between the electrons prevents them from moving at all. At low temperatures the majority of the electrons are in their proper places in the ions. The minority which have happened to cross the potential barrier find therefore all the other atoms occupied, and in order to get through the lattice have to spend a long time in ions already occupied by other electrons. This needs a considerable addition of energy and so is extremely improbable at low temperatures.’

His argument explained essentially the mechanism of strong electron correlation and how this correlation affects the transport properties of materials. The difficulty for band theory here is that the electron-electron interaction can no longer be taken as a perturbation to the system, but a *non-perturbative* many-body theory has to be applied. A prototype of theoretical understanding for the strong electron correlation is given by the Hubbard model [54, 55]. The single-band Hubbard model Hamiltonian

is given by

$$H_{Hub} = - \sum_{i,j,\sigma} t_{ij} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \quad (2.2.1)$$

where i and j label the Wannier states on the lattice sites, t_{ij} is the amplitude of hopping integral, $\hat{c}_{i\sigma}^\dagger$ ($\hat{c}_{j\sigma}$) is the creation (annihilation) operator of electrons with spin σ , U is the celebrated *Hubbard* U , which represents the energy of on-site Coulomb repulsion, $\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}$ is the number operator for spin σ on site i .

In the Hubbard model, the first term, $H_t = - \sum_{i,j,\sigma} t_{ij} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma}$, is usually called the *hopping Hamiltonian*. t_{ij} is the hopping integral, which was described in the tight-binding approximation as well (see Section 2.1.2). In fact, the hopping term is responsible for the band structure and can be diagonalized with Bloch waves. So, this part of the Hamiltonian is completely uncorrelated, and what it describes is simply the Fermi liquid.

Now let us take a look at the second term, $H_U = U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}$, which is called the *interaction Hamiltonian*. From its form, only the on-site Coulomb interaction is considered, while the long-range interaction is ignored. This approximation is legitimized by two facts: 1) the Coulomb interaction decays with the distance between sites; 2) the Coulomb potential is effectively screened by electrons between sites. If the system has one electron per site (half-filled case), the ground state must be that every site is occupied by one electron. Doubly occupied sites are excluded, because they lead to higher energy. So, the electrons are localized on the atomic sites in the interaction Hamiltonian.

Since the half-filled Hubbard-model Hamiltonian, H_{Hub} , is the sum of the ‘hopping Hamiltonian’, whose ground state is an *itinerant* ground state, and the ‘interaction Hamiltonian’, whose ground state is a *localized* ground state, competition between these two states can be expected.

The competition can be illustrated by comparing the energy of two ground states. In the itinerant ground state, the hopping energy is approximately $-t$, which is propor-

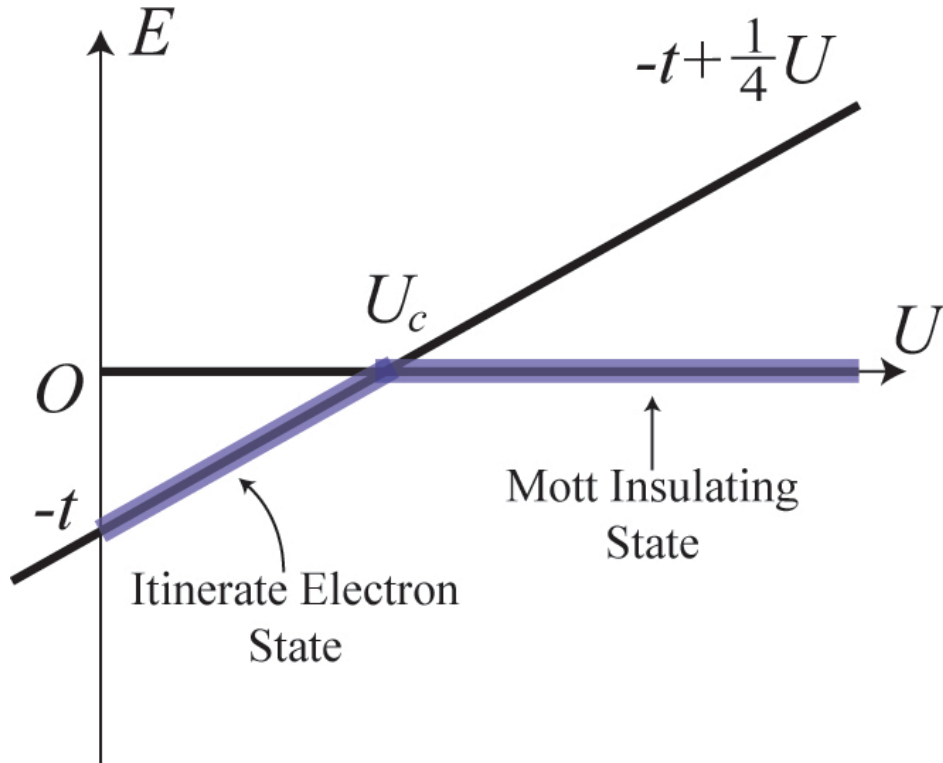


Figure 2.4: Comparison of the energies of the itinerant ground state and the localized ground state of the half-filled system. A phase transition happens at the critical Hubbard U_c , which separates the metallic and Mott-insulating states.

tional to the band width, according to the tight-binding approximation. To estimate the interaction energy in the itinerant state, we assume the probability of finding a spin-up electron is independent of the probability of finding a spin-down electron on a site. Since the number of electrons with spin-up is half of the number of sites, the probability of finding a spin-up electron on site i is $\frac{1}{2}$. The same argument is also applied to the spin-down electrons. So, the probability to find a site with both spin-up and spin-down electrons is simply $p_{double} = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. Because only the doubly occupied sites contribute to the interaction energy, the interaction energy is $\frac{1}{4}U$. Thus, the energy of the itinerant state is $E_{iti} = -t + \frac{1}{4}U$, which is plotted in Fig. 2.4. The energy of the localized ground state can be calculated in a very straightforward way. The hopping energy is completely killed in the localized ground state, because any movement of electrons will cause doubly occupied sites. The interaction energy is also simply zero, because every site is half-filled and either $n_{i\uparrow}$ or $n_{i\downarrow}$ equals zero. Thus, the ground state of the system alternates with the relative strength of t and U , as shown in Fig. 2.4. When the itinerant ground state is changed to the localized ground state, since electrons are forbidden to hop between sites, a metal-to-insulator transition (MIT) happens at the critical Hubbard U_c . This transition is also called a Mott transition.

In a band insulator, there must be an energy band gap between the lowest empty band and the highest filled band, and the Fermi surface sits inside the band gap. A similar situation can be expected for the Mott-insulating state. Fig. 2.5 shows the difference between a band insulator and a Mott insulator in the electron distribution in real space and the electronic energy spectrum. In a Mott insulator, even when the energy band is half-filled, moving an electron from one site to a neighboring site requires additional energy $\sim U$, due to the on-site Coulomb repulsion between two electrons after their movement. In the electronic band structure, the originally half-filled energy band in the semiconductor band structure is split into two narrow bands. The band gap equals $U - t$, as shown in Fig. 2.5(b). The higher energy band is called the *upper*

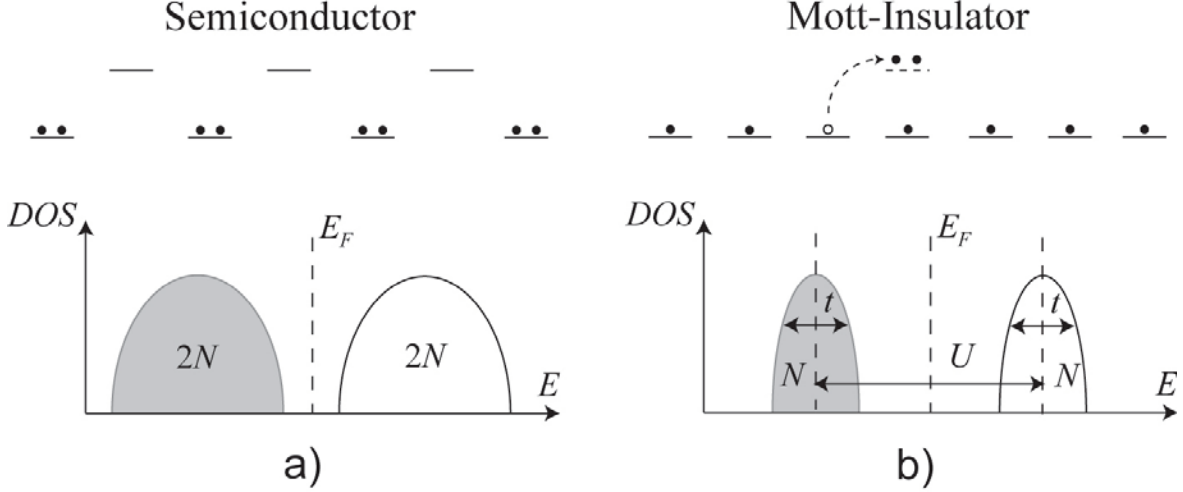


Figure 2.5: Band structures comparison between a band insulator and a Mott insulator. In a band insulator, as illustrated in a), the valence band is half-filled. For N sites on the lattice, there are totally $2N$ electrons filled in the valence band. The factor of 2 accounts for electron spins. The $2N$ empty states are located in the conduction band. b) in a Mott insulator, there are N sites on the lattice and only N electrons are filled. The half-filled band (in band-theory sense), is separated into two narrow bands: the lower Hubbard band and the upper Hubbard band. The two bands are separated by an energy $U - t$. Each Hubbard band contains N states and the lower Hubbard band is fully filled. When an electron moves from one site to the neighboring site, additional energy due to the on-site Coulomb interaction needs to be paid, which lifts its energy level, as illustrated by the real-space atomic energy levels [56].

Hubbard band, while the one with a lower energy is called the *lower Hubbard band*.

From the Hubbard model, can we predict in what kinds of materials, electron correlation effects will be important? As shown in Fig. 2.4, the system stays in a localized ground state, when the Hubbard U is larger than the band width (hopping integral amplitude) t . In experiments, the electron correlation effects are often observed in materials containing transition metal elements (V, Cu, Ni, etc.), because the open valence bands in these atoms consist of the d and f orbitals, which are usually very narrow. Vanadium dioxide (VO_2), which will be introduced in the next chapters, serves as a very good example. The d electronic states of vanadium contribute to a very narrow half-filled d band at the Fermi surface at its high-temperature metallic state. And this d band is split by a band gap into an upper Hubbard band and a lower

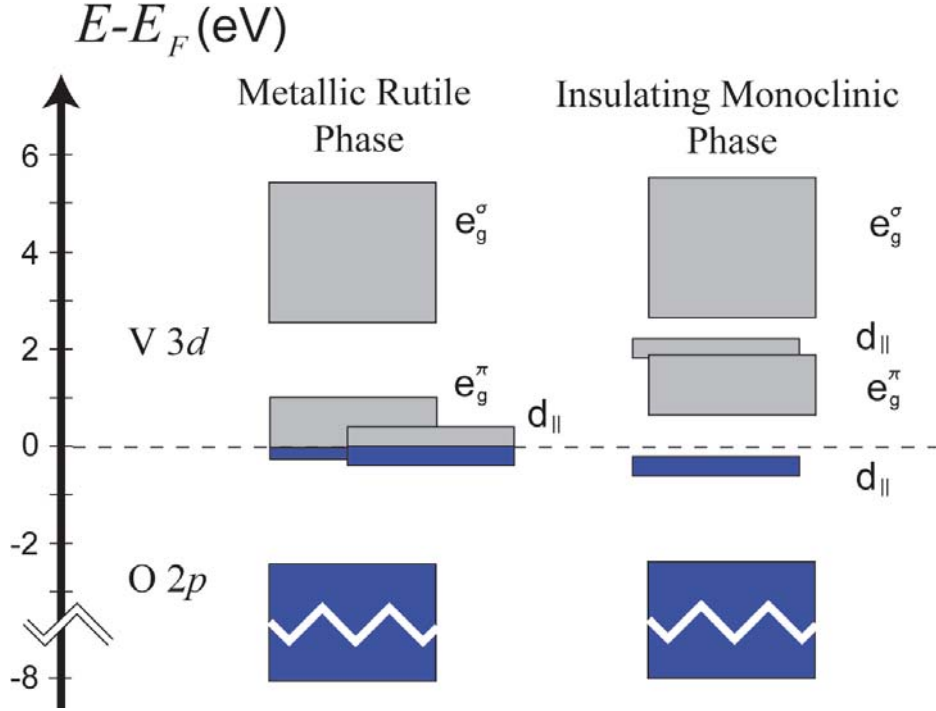


Figure 2.6: A schematic band structure of VO₂ in a) the metallic-rutile phase and b) the insulating-monoclinic phase, according to Ref. [57] and [58].

Hubbard band, when the system is cooled down to its low-temperature insulating state through a metal-to-insulator transition, as shown in Fig. 2.6. However, the situation of VO₂ is more complicated, due to a spontaneous lattice distortion, which will be introduced in more detail in Chapter 3.

2.3 Theoretical Treatment

The Hubbard model contains rich physics in its simple form. However, even the single-band Hubbard model cannot be affordably used to calculate a three-dimensional lattice, not to mention the extension to a multi-band model. Some perturbative treatments have been applied to reduce the computational cost, but these treatments fail in the most interesting region, where U and t are comparable.

2.3.1 DFT LDA+ U

As has been introduced in previous sections, Density Functional Theory (DFT) is the most favorable tool to calculate the electronic structure for systems of realistic complexity. Its simplicity and efficiency stems from the possibility of expressing all the ground-state properties as functionals of the charge density and the existence of a variational principle for the total energy functional. Because of its low computational cost, it would be favorable, if the electron correlation energy can also be functionalized and included into the Kohn-Sham equations (Eqn. 2.1.22 - 2.1.24). In recent years, DFT LDA+ U method is extensively developed for this purpose. Here, the general formulation and its advantages and disadvantages will be introduced. More detailed information is available in review articles (e.g. Ref. [59] and the references therein).

In the standard DFT, the functional of exchange-correlation energy is unknown. The forms of the exchange-correlation functionals are approximately calculated assuming a homogeneous free-electron gas, which is the local-density approximation (LDA). Since LDA is based on a free-electron model, it is not hard to understand its intrinsic difficulty in capturing the properties of strongly correlated materials, whose ground state is characterized by the localized electrons. The idea of LDA+ U is quite simple. It describes the correlated electrons of a system (typically, the localized d and f orbitals) using the Hubbard model (Eqn. 2.2.1), while the remaining uncorrelated electrons are treated using the standard approximation of DFT functionals. The total energy within LDA+ U framework is given by

$$H_{LDA+U}[\rho(\vec{r})] = H_{LDA}[\rho(\vec{r})] + H_{Hub}[\{n_{i\sigma}^{mm'}\}] - H_{dc}[\{n_{i\sigma}\}], \quad (2.3.1)$$

$$H_{Hub}[\{n_{i\sigma}^{mm'}\}] = \sum_i \frac{U_i}{2} \sum_{m,\sigma \neq m'\sigma'} n_{i\sigma}^m n_{i\sigma'}^{m'}, \quad (2.3.2)$$

$$H_{dc}[\{n_{i\sigma}\}] = \frac{U_i}{2}n_i(n_i - 1). \quad (2.3.3)$$

In Eqn. 2.3.1, H_{LDA} represents the standard DFT Hamiltonian with LDA approximation. The second term on the right side, H_{Hub} , is the same as in the Hubbard-model Hamiltonian, which is defined in Eqn. 2.3.2. Compared with the Hubbard model, Eqn. 2.3.2 has a very similar form. Only the Coulomb interaction on a single site i is included and U_i represents the strength of interaction. The difference here is the summation over m , which is the index of atomic states. So, a multi-band Hubbard model is incorporated in the LDA+ U method. In the standard LDA approximation, part of the electron-correlation energy is already included in H_{LDA} . So, it is necessary to eliminate it. This task is accomplished by subtracting the third term, H_{dc} . Here, ‘ dc ’ stands for ‘double counting’. Unfortunately, the functional of the double counting term is unknown, and different formulations have been implemented. The two most popular formulations lead to the ‘around mean-field’ (AMF) and ‘fully localized limit’ (FLL) implementations of LDA+ U . As the name suggests, the AMF LDA+ U is a better treatment for ‘weakly correlated electrons’, while the FLL is more appropriate for materials, whose electrons are strongly localized. In Eqn. 2.3.3, the FLL implementation is used.

Taking VO_2 in the insulating M_1 phase as an example, the electronic band structure calculated using the standard LDA is plotted in Fig. 2.7(a) [60]. The density of states at the Fermi level ($E = 0$) is nonzero, which means the standard DFT LDA calculation yields a metallic band structure for the low-temperature monoclinic VO_2 . In Fig. 2.7(b), the same calculation is repeated using the DFT LDA+ U method [61]. From the results, we find that the local density of states exhibit a gap of about 0.7 eV, which is in agreement with the optical measurements [62]. A Hubbard U of ≈ 4.2 eV is applied in the calculation. So, it is essential to consider electron-correlation energy to predict the correct ground state for VO_2 . Unfortunately, using the same Hubbard

U , LDA+ U also yields an insulating ground state for VO_2 's high-temperature rutile phase [63]. This dilemma reveals the limitation of the LDA+ U method in treating the problems of strongly correlated materials.

2.3.2 Dynamical Mean Field Theory

Dynamical mean field theory (DMFT) is a state-of-art technique to calculate the properties of strongly correlated materials. In DMFT, the Hubbard model is mapped onto a self-consistent quantum impurity model, i.e. a set of local quantum mechanical degrees of freedom that interacts with a bath consisting of a continuum of noninteracting excitations. This approximation is referred as *local impurity self-consistent approximation* (LISA) [64–66].

In the Hubbard model (Eqn. 2.2.1), the hopping integral t_{ij} describes the hopping of electrons between sites i and j . Inspired by Weiss mean-field theory [67], the DMFT separates a single lattice site's atomic degrees of freedom, treated as H_{atom} , from the remaining degrees of freedom, treated as a bath of electrons with energy E_{ν}^{bath} . Thus, the many-body problem is effectively mapped to the *Anderson impurity model* [68]. The Hamiltonian is given by [66]

$$H_{\text{AIM}} = H_{\text{atom}} + \sum_{\nu, \sigma} E_{\nu}^{\text{bath}} n_{\nu, \sigma}^{\text{bath}} + \sum_{\nu, \sigma} (V_{\nu} c_{0, \sigma}^{\dagger} a_{\nu, \sigma}^{\text{bath}} + h.c.). \quad (2.3.4)$$

The subscript ‘AIM’ denotes the Anderson impurity model, ν is the index of site in the bath, V_{ν} represents the hybridization between impurity and bath electrons, $c_{0, \sigma}^{\dagger}$ is the creation operator for an electron with spin σ on the impurity and $a_{\nu, \sigma}^{\text{bath}}$ is the annihilation operator of the bath. Fig. 2.8 shows the physical picture of the Eqn. 2.3.4. V_{ν} and E_{ν}^{bath} are the parameters to be defined. The conduction bath density of states is given by

$$\Delta(\omega) = \sum_{\nu} \frac{|V_{\nu}|^2}{\omega - E_{\nu}^{\text{bath}}}. \quad (2.3.5)$$

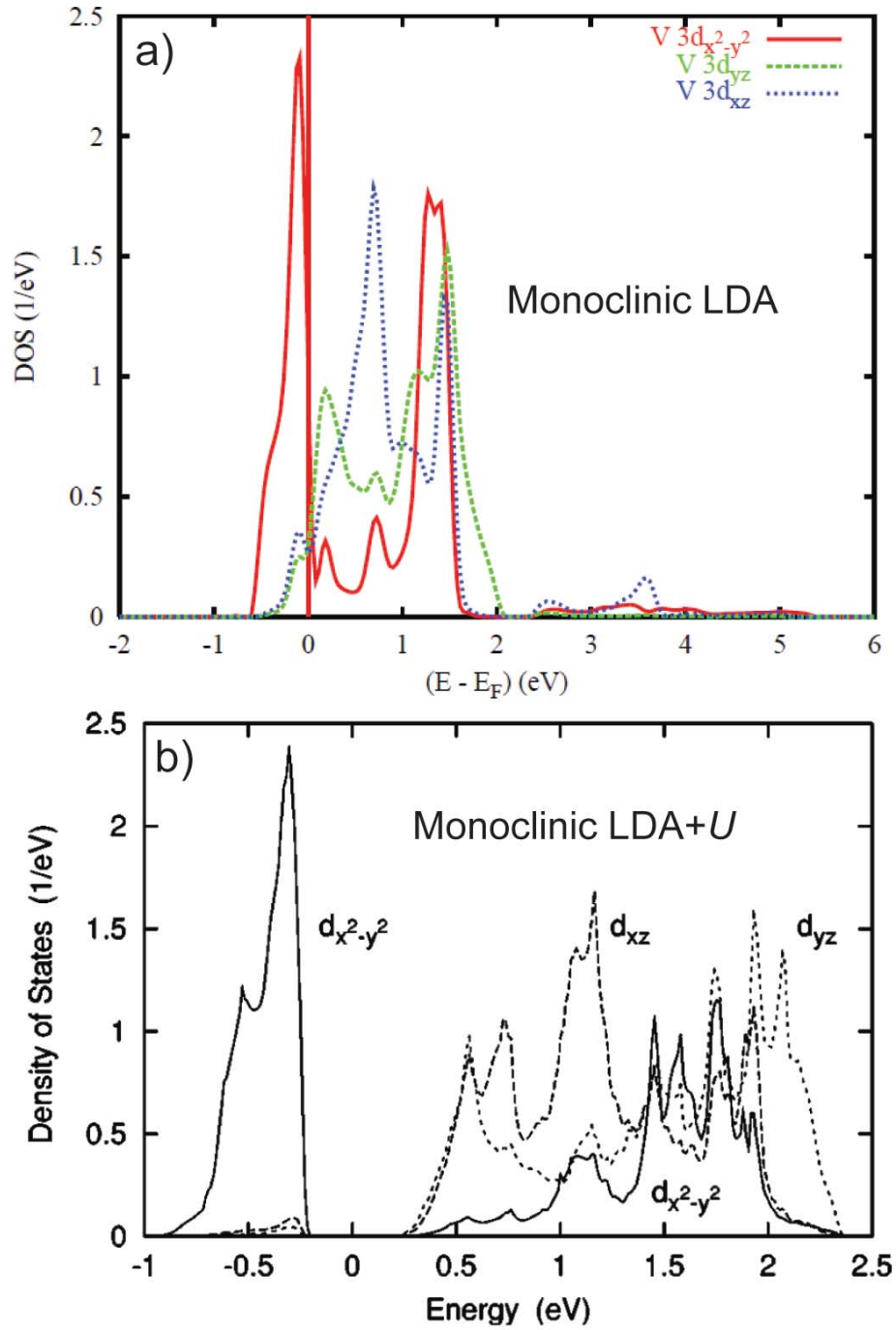


Figure 2.7: Band structure of insulating monoclinic M_1 phase of VO_2 calculated a) using DFT LDA method. Figure from Ref. [60]. b) using DFT LDA+ U method. Figure from Ref. [61]

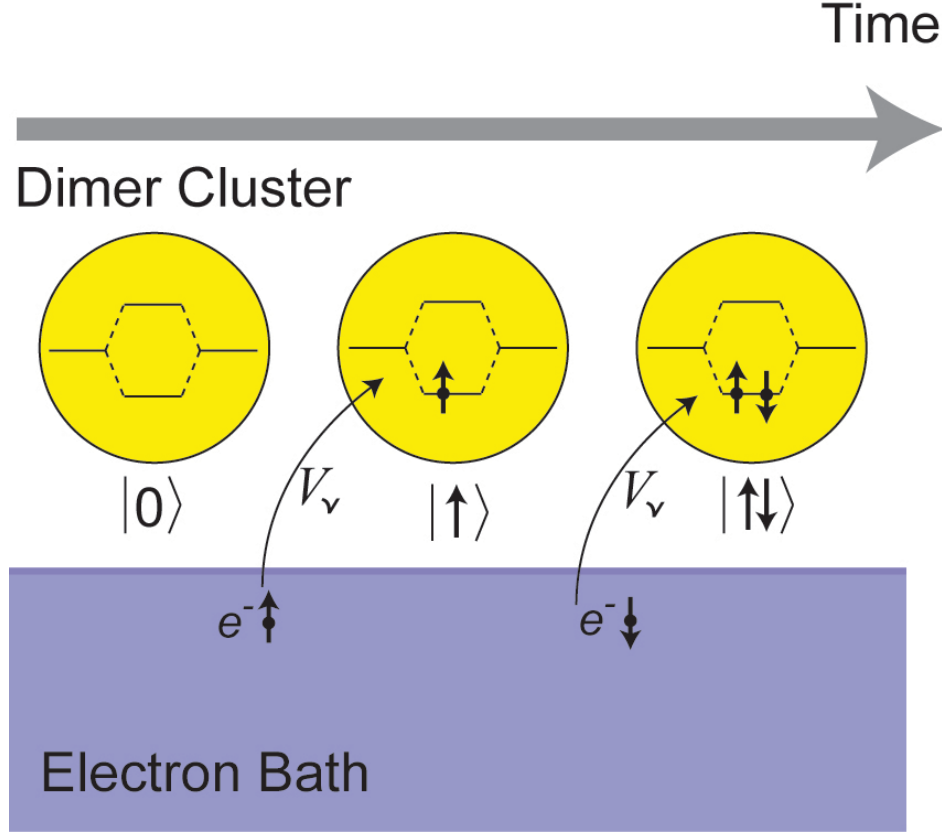


Figure 2.8: DMFT replaces the full lattice and electrons with a dimer cluster (yellow spheres) and a bath of non-interacting electrons. The model captures snapshots of the electronic state fluctuation on the dimer cluster. The population of the dimer cluster is determined by the quantum mechanical probability for hopping of electrons from bath to the cluster, which is quantified by parameter V_ν .

Because the bath describes the same electrons as those on the local site, $\Delta(\omega)$ can be resolved in a self-consistent way using the local Green function

$$G[\Delta(\omega)] = \sum_{\vec{k}} \{\omega - \Sigma[\Delta(\omega)] - t_{\vec{k}}\}^{-1}, \quad (2.3.6)$$

where the self-energy term $\Sigma[\Delta(\omega)] \equiv \Delta(\omega) - 1/G[\Delta(\omega)] + \omega$ takes the meaning of the frequency-dependent potential, and $t_{\vec{k}}$ is the Fourier transform of the hopping integral t_{ij} . Detailed derivations are available in Ref. [64, 65].

The major advantage of DMFT is that it takes into account the dynamical (frequency-dependent) effects of electronic correlations, which the LDA+ U method cannot do. Un-

fortunately, DMFT is significantly more computationally demanding than DFT and requires a significant effort to be implemented in or to be interfaced with existing DFT codes. Again, for the two different phases of VO_2 , the standard DMFT yields incorrect metallic solutions for VO_2 M_1 insulating phase [61]. With *cluster* extension of DMFT (c-DMFT), which takes V-V dimer as the key unit of calculation, a very good agreement with the experimental results has been achieved [58, 69, 70]. These results underscore the importance of the intradimer correlation, which cannot be correctly captured by the LDA+ U method and the standard DMFT.

2.4 Emergent Phenomena of Strongly Correlated Materials

The effects of electron correlation are profound and impact the material properties significantly. Because the interaction among electrons is so strong, the perturbative understanding based on the band theory fails on these strongly correlated materials and many phenomena occurring become unpredictable for us. Here, several emergent phenomena of strongly correlated materials are going to be briefly introduced.

2.4.1 Metal-to-insulator Transition

According to the Hubbard model, metal-to-insulator transition (MIT) is a natural consequence of the two competing mechanisms in strongly correlated materials: the itinerant electrons and the localized electrons. In experiments, a conductivity or resistivity change by several orders of magnitude can be observed when the sample's temperature crosses the critical temperature (T_c). The phase transition is usually first order, which requires latent heat throughout the phase transition and defines a sharp temperature boundary between insulating and metallic phase. More intriguingly, the phase boundary (T_c) is very susceptible to the internal and external parameters, includ-

ing doping, pressure, field, etc. A detailed summary of the MIT properties of strongly correlated materials can be found in Ref. [52]. VO_2 serves as a prototypical material presenting all these properties and we will discuss its phase diagram in more detail in Chapter 3.

In addition to the MIT itself, a more challenging subject is to describe and understand the metallic phase near T_c . In this regime, fluctuations of spin, charge and orbitals are strongly enhanced. The metallic phases of strongly correlated materials are usually not well conducting like the normal metals, so these phases are also called ‘bad metals’. A typical anomalous fluctuation is the electron mass enhancement in the metallic phase near the critical temperature, which has been observed in VO_2 through the infrared micro-spectroscopy [71]. At the same time, the metallic and insulating domains form around the critical temperature [71], which is found to be susceptible to the local strain in the sample [72]. Around the critical temperature, lattice distortions happen accompanied by an electronic spectra-weight transfer [58]. The orbital switch happens either to facilitate the MIT or as a consequence of it, which is still an important question related to the electronic correlation effects. In recent years, the interest in VO_2 ’s MIT problem was refueled, because the phase transition was found to be optically induced on an ultrafast timescale [73, 74]. Many other experiments have been done afterwards using various techniques. On the ultrafast timescale, VO_2 is driven into a strongly nonequilibrium state, which reveals important information relating to electronic correlations.

2.4.2 Colossal Magnetoresistance

Colossal magnetoresistance (CMR) is a property of some materials, mostly manganese-based perovskite oxides that enables a change in the electrical resistance by several orders of magnitude in the presence of a magnetic field. These materials are also called ‘*manganites*’. The general chemical formula for manganites is $\text{Re}_{1-x}\text{A}_x\text{MnO}_3$, which is

described by the seminal paper of Jonker and Van Santen [75]. Re is a trivalent rare earth or Bi^{3+} cation, and A is a divalent alkaline or Pb^{2+} cation. Oxygen is in the O^{2-} state, and the relative fraction of Mn^{3+} and Mn^{4+} is regulated by ‘x’ in the formula.

Even though large magnetoresistance (MR) effects had been discovered in 1989 [76], manganites were mainly considered as an interesting physics problem, instead of being used for real applications, because of their requirements on low Curie temperature and high critical-field strength [77]. The big boost of the field was later produced by the discovery of the ‘colossal magnetoresistance’ (CMR) effect. The MR measurements on the thin films of manganites exhibited more than 1000-fold change in resistivity [78, 79], which is much higher than what observed in the giant magnetoresistance (GMR) effect [80]. Especially in the intermediate doping region $0.2 \leq x \leq 0.5$, the metal-to-insulator transition happens from a low-temperature ferromagnetic metal to a high-temperature paramagnetic semiconductor. The critical temperature (T_c) of CMR is very close to room temperature. CMR effect is observed mainly in the region very close to T_c [78]. Accompanied by the change in magnetoresistance, a Jahn-Teller type lattice distortion also occurs, which suggests a complicated mechanism of the CMR effect.

The double-exchange (DE) mechanism was proposed many years ago to account for the formation of ferromagnetic metallic state in manganites [81, 82]. However, Millis pointed out that the DE mechanism alone is not sufficient to explain the very large resistivity of the $T > T_c$ phase [83]. It was later proposed that the strong electron-phonon coupling due to the Jahn-Teller distortion localizes the conduction electrons as polarons at high temperature, and the polaron effect regulates the competition between electronic itineracy and self-trapping throughout the phase transition [84]. This theoretical description suggests that the interplay between correlated electrons and other degrees of freedom (spins, orbital ordering, ...) is the key to understand the CMR problems. More recently, studies also suggest that the ground state of manganite tends to be intrinsically inhomogeneous due to a phase separation. Calculations of

the resistivity versus temperature using mixed states lead to a good agreement with experiments. A quantitative understanding is still elusive, which is an active subject for many theorists and experimentalists.

2.4.3 High-temperature Superconductivity

High-temperature superconductivity (high- T_c) probably is the most significant and famous emergent phenomenon which strongly correlated materials bring us. When Bardeen, Cooper and Schrieffer proposed their celebrated BCS theory [85–87] and were awarded 1972 Nobel Prize in Physics, most people probably believed that the problem of superconductivity is well understood and resolved, until the materials with abnormally high critical temperature for superconductivity, which we know as the high- T_c materials, were for the first time discovered in 1986 by G. Bednorz and K. Alex Müller [88], who were also awarded the Nobel Prize in Physics one year later.

The first high- T_c materials found are the layered copper oxides, which are called *cuprate superconductors*. At room temperature, these materials are mostly ceramics, which are insulating. However, when the temperature is cooled down below the critical temperature T_c , the resistivity disappears in these materials as the superconductivity is formed, and T_c of these materials are usually much higher than conventional (*s*-wave) superconductors. The superconductivity in cuprates is very robust against disorders, which is another difference compared to the conventional superconductors. Structurally, very similar to the CMR materials and many other strongly correlated materials, high- T_c cuprate superconductors are related to the perovskite structure, but with distortion and oxygen deficiency. Various phases with different charge-, spin-, and orbital orderings have been observed in high- T_c materials, which are strongly affected by charge density (doping), external pressure and other factors. In addition to a very complicated phase diagram, coexistence of phases is also observed, similar to the observations in the MIT and CMR effects of other strongly correlated materials.

Because the pairing electrons are believed to be d -orbital electrons (d -wave superconductor), which is confirmed by various experiments, the strong electronic correlation is considered central to the problem of the high-temperature superconductivity in cuprates [89–91]. Numerous theories have been proposed to explain the physics of high- T_c materials. Nice references to the experimental and theoretical approaches can be found in Refs. [92] and [93]. The pairing mechanism of the Cooper pairs in high-temperature superconductors is so far still under debate, mainly because of the complexity of the materials and the indirect nature of experimental results.

The cuprate with the highest critical temperature is so far the mercury barium calcium copper oxide ($\text{HgBa}_2\text{Ca}_2\text{Cu}_3\text{O}_8$) at ~ 133 K [94]. In 2006, the field of high-temperature superconductors was boosted by the discovery of another family of high- T_c materials, iron-based superconductors [95]. Iron-based superconductors contain layers of iron and a pnictogen, so they are also called iron-pnictide superconductors. The highest critical temperature in this family is so far ~ 55 K. There is still a long way to go to realize the dream of the room-temperature superconductivity.

Chapter 3

Photo-induced Phase Transitions in VO₂: an Investigation on Thin-film Membrane VO₂

In Chapter 2, the general concepts and properties of strongly correlated materials were introduced and VO₂ was taken as an example of this family of materials. By directly measuring the conductance or resistivity, it is found that, when VO₂ is cooled down from high temperature and crosses its critical temperature T_c ($\sim 67^\circ\text{C}$), it undergoes a first-order metal-to-insulator transition (MIT). Accompanied by the MIT, there is a remarkable modification on the structure from the low-temperature monoclinic structure to its high-temperature rutile structure. Because its phase transition temperature is in the vicinity of ambient temperature and also, because of its large change in electrical conductivity and optical reflectance, VO₂ is considered as a promising candidate for several applications, including bolometers [96], memristors [97], tunable-frequency metamaterials [98] and data storage [99]. Researches on further decreasing the critical temperature down to room temperature are still going on, including using tungsten doping [100–102] or applying external strain [103, 104].

Besides its prosperity in applications, VO_2 is also considered as an important scientific problem in studies of strongly correlated materials for decades. Because of its simple atomic composition and structure, VO_2 provides a benchmark for testing different theoretical models for strongly correlated materials, like density functional theory (DFT) [105], dynamical mean field theory (DMFT) [69, 106], etc. In recent years, more and more time-resolved experiments were applied on VO_2 using various techniques, including time-resolved optical spectroscopy [73, 74, 107, 108], time-resolved X-ray diffraction and spectroscopy [109–113], time-resolved Terahertz spectroscopy [114–117], and ultrafast electron diffraction [19, 20, 118, 119]. In these experiments, VO_2 is first excited by a pump laser pulse, and a highly non-equilibrium state between electrons and lattice is thus generated. Its consequent dynamics are probed by the later arriving optical or electron probes. With time-resolved techniques, the electronic and lattice responses to the perturbation of strongly correlated electrons become more clear, which helps in understanding the fundamental relationships of the electronic correlation and the structural transition in VO_2 .

In this chapter, the experimental results on ultrafast electron diffraction (UED) on VO_2 thin-film membrane will be presented. Here, we want to address the following two questions: 1) Is the photo-induced phase transition in VO_2 a charge-driven or thermally driven process? 2) How does the structural phase transition happen, after photoexcitation? UED method is an ideal tool to investigate these two questions, because it can capture the atomic dynamics, after fs laser excitation.

The chapter starts with an introduction of the VO_2 metal-to-insulator transition (MIT) and structural phase transition (SPT) in Section 3.1. The general characterization of the VO_2 thin-film membrane samples for the UED experiments is presented in Section 3.2, which gives us a fundamental understanding of the sample properties. The thresholds of pump fluence to induced the phase transitions with different pump wavelengths are first investigated using the UED method, which gives the direct evidence

to resolve the first question, and the results are presented in Section 3.3. In Section 3.4, we will take a close look at the atomic dynamics of the photo-induced VO₂ phase transitions within the first several picoseconds timescale after photoexcitation. A brief summary is given in Section 3.5.

3.1 General Properties of VO₂ Metal-to-Insulator Transition and Structural Phase Transition

In 1959, F. J. Morin first found VO₂ undergoes a first-order metal-to-insulator phase transition at the critical temperature $T_c \sim 340$ K [120]. In his experiment, a conductivity discontinuity of two orders of magnitude was observed, but subsequently, a discontinuity of as much as five orders of magnitude has been observed [121], due to the advancement in synthesis techniques, which leads to better sample quality.

Accompanied by the metal-to-insulator, a structural phase transition also occurs in VO₂ at the critical temperature. High-temperature VO₂ is identified to have a tetragonal (P4₂/mm) rutile structure, while the structure of the low-temperature VO₂ is a monoclinic (P2₁/c) structure. The unit cell structures of the insulating monoclinic (M₁) phase and the metallic rutile (R) phase are depicted in Fig. 3.1. A striking feature of the monoclinic phase is that the V atoms alternatively move in the direction normal to the $a_M(2c_R)$ axis and dimerize with each other along the $a_M(c_R)$ axis. This pairing of V atoms breaks the symmetry in VO₂ and doubles the unit-cell size along the a_M direction in its low-temperature monoclinic phase, as shown in Fig. 3.1. Both the monoclinic and rutile structure have been investigated by X-ray crystallography [122–124] and the atomic coordinates of V and O atoms can be derived from the X-ray diffraction data. Table 3.1 and Table 3.2 list the orthogonal atomic coordinates in a VO₂ unit cell for the low-temperature monoclinic structure and high-temperature rutile structure, respectively.

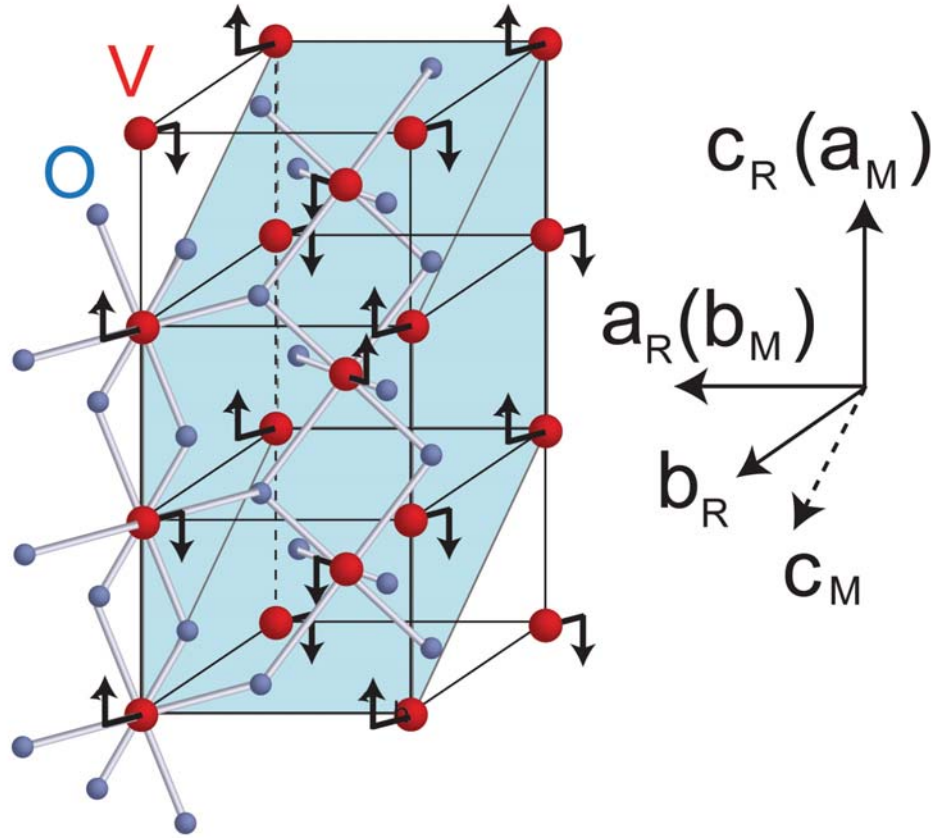


Figure 3.1: The structures of VO_2 rutile (R) and monoclinic (M_1) phases.

The structural dynamics throughout the photo-induced phase transition in VO_2 were investigated in previous time-resolved X-ray [109] and electron crystallography [118] studies and important information relating to the timescale and atomic movements has been obtained with these techniques. However, in these experiments, only the near-infrared 800 nm pump was used to induce the phase transitions and questions relating to the nature of photo-induced phase transition in VO_2 were not addressed.

3.2 General Characteristics of VO_2 Thin-film Membrane

In this set of experiments, VO_2 samples in the form of polycrystalline thin-film membranes are used. Here, the VO_2 thin film is homogeneously deposited onto a TEM grid with silicon windows, which have 9 nm thick silicon membrane as the supporting membrane. The window is $100 \times 100 \mu\text{m}^2$ in size, which is large enough to contain the whole electron beam in our UED system ($\sim 40 \mu\text{m}$ FWHM), which helps to efficiently use available electrons for diffraction, thus improves the signal-to-noise ratio (SNR). This is essential for resolving the detailed structural dynamics. The deposition is finished using the pulsed laser deposition (PLD) method [125]. The VO_2 thin-film membrane samples are prepared by Dr. Sepulveda's group from the Material Science and Engineering department of Michigan State University. The thickness of the thin film is ~ 50 nm. Under a scanning electron microscope (SEM), the nanogranular morphology of the film is evident. The average grain size is ~ 50 nm, as shown in Fig. 3.2.

The thin-film-membrane sample is transferred into the UHV chamber and first probed with the 30 keV electron beam first without pump excitation. The ground-state diffraction pattern is recorded. Because of the polycrystalline nature of the thin film, the diffraction pattern consists of a group of Debye-Scherrer rings, as shown in Fig. 3.3(a). The ring-shaped diffraction pattern is caused by the mixed orientation

Atom	$x(\text{\AA})$	$y(\text{\AA})$	$z(\text{\AA})$
V ₁	1.3174	4.4041	0.1132
V ₂	-0.1309	2.3714	2.3770
V ₃	2.9773	2.1456	2.1560
V ₄	1.5289	0.1129	4.4145
O ₁	-0.0050	0.9486	0.9053
O ₂	-1.4533	1.3099	3.1693
O ₃	-0.0485	3.6588	3.5768
O ₄	1.4465	1.4003	3.2147
O ₅	1.3998	3.1167	1.3130
O ₆	2.8949	0.8582	0.9508
O ₇	2.8513	3.5684	3.6222
O ₈	4.2997	3.2071	1.3583

Table 3.1: List of orthogonal atomic coordinates of VO₂ monoclinic (M₁) structure. Crystal symmetry: P2₁/c; Unit cell definition: $a=5.743 \text{ \AA}$, $b=4.517 \text{ \AA}$, $c=5.375 \text{ \AA}$, $\beta=122.61^\circ$ [122]

Atom	$x(\text{\AA})$	$y(\text{\AA})$	$z(\text{\AA})$
V ₁	0.4530	0.4530	0.2869
V ₂	0.4530	0.4530	1.7214
V ₃	2.7180	2.7180	0.2869
V ₄	2.7180	2.7180	1.7214
O ₁	1.3363	1.3363	0.2869
O ₂	1.3363	4.0997	1.7214
O ₃	1.8347	1.8347	0.2869
O ₄	1.8347	3.6013	1.7214
O ₅	3.6013	1.8347	1.7214
O ₆	3.6013	3.6013	0.2869
O ₇	4.0997	1.3363	1.7214
O ₈	4.0997	4.0997	0.2869

Table 3.2: List of orthogonal atomic coordinates of VO₂ rutile (R) structure. Crystal symmetry: P4₂/mm; Unit cell definition: $a=4.530 \text{ \AA}$, $c=2.869 \text{ \AA}$ [123]

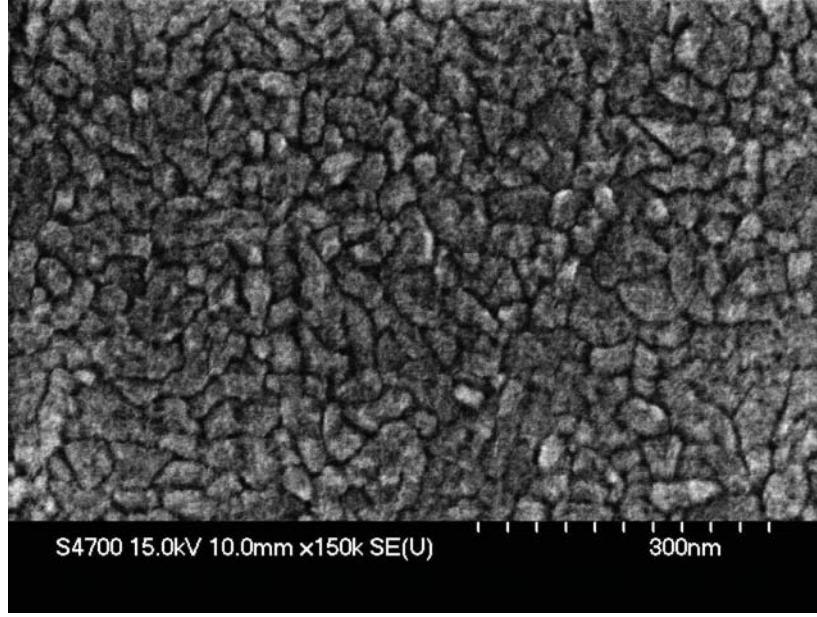


Figure 3.2: SEM image of VO_2 thin film. Nanogranular feature of the thin film is observed with grain size ~ 50 nm.

of the nano-grains, which scatter electrons homogeneously around the radial direction. Because of the radial symmetry, the diffraction pattern can be radially averaged to obtain the one-dimensional curve as shown in Fig. 3.3(b) (black symbols). The radius of the rings are converted to the momentum transfer s , with the unit [$\text{\AA}^{-1}$], as shown in Fig. 3.3(b). The vertical axis is the averaged intensity of the rings.

In powder diffraction, the Laue diffraction conditions are satisfied. Each Bragg peak in Fig. 3.3(b) is contributed from electrons scattered from a set of crystal planes with a specific Miller index. To correctly index the Bragg peaks, the momentum transfer and intensity of Bragg reflections can be calculated using the kinematic diffraction theory, which was described in detail in Chapter 1, using the atomic coordinates of M_1 phase listed in Table 3.1. The theoretical curve is plotted as the red line in Fig. 3.3(b). Compared with the experimental result, a good agreement is achieved on peak positions, which are determined by the dimension of the real-space unit cell, with a reasonable agreement on the peak intensities. The discrepancy in intensities can be caused by the dynamical scattering effect (see Chapter 1), which redistributes

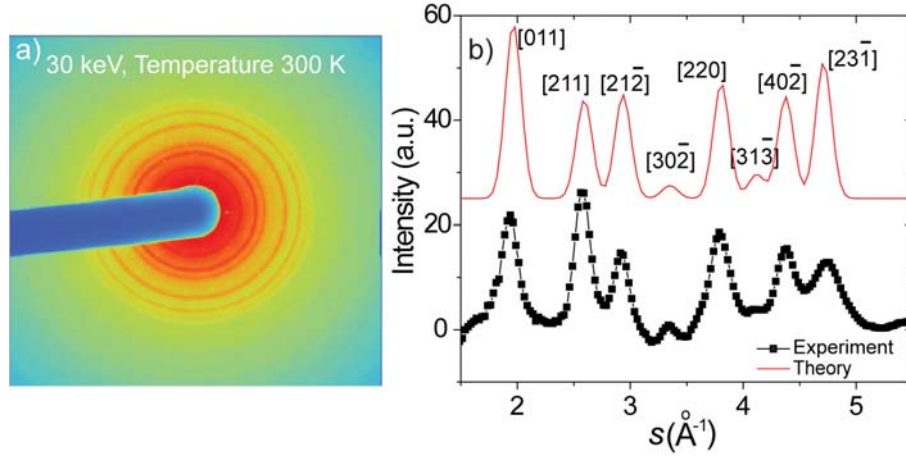


Figure 3.3: a) Ground-state powder diffraction pattern from VO_2 thin-film membrane. b) 1D curve extracted from experimental diffraction pattern (black squares) with the diffraction pattern calculated using the kinematic diffraction theory (red curve). Miller indices are labeled for reference.

the scattered electrons among different Bragg reflections. However, we will find the absolute intensity of Bragg peaks is not important in the following analysis, because only the relative intensity change of each Bragg peak, which is subtracted from the ground-state intensity, is of interest. Here, we assume that the photo-induced structural change is not large enough to affect the conditions of the dynamical scattering of electrons.

In the UED system, the sample temperature can be changed across the critical temperature, $T_c \sim 340$ K. The integrated intensity of the Bragg peak $[30\bar{2}]$ decreases to zero when the sample temperature is raised higher than T_c , which is shown in Fig. 3.4. From the kinematic diffraction calculation, it can be confirmed that the intensity of the peak $[30\bar{2}]$ is exclusively contributed by the distortions of the V atoms. As the V-V dimerization dissolved in the rutile structure, the peak intensity decays to zero. Here, the peak intensity reduction induced by the Debye-Waller effect is estimated to be $< 5\%$, so the thermally induced fluctuations contribute much less to the intensity reduction compared to the atomic movements throughout the phase transition. The intensity of peak $[30\bar{2}]$ can be taken as the signature of the monoclinic-

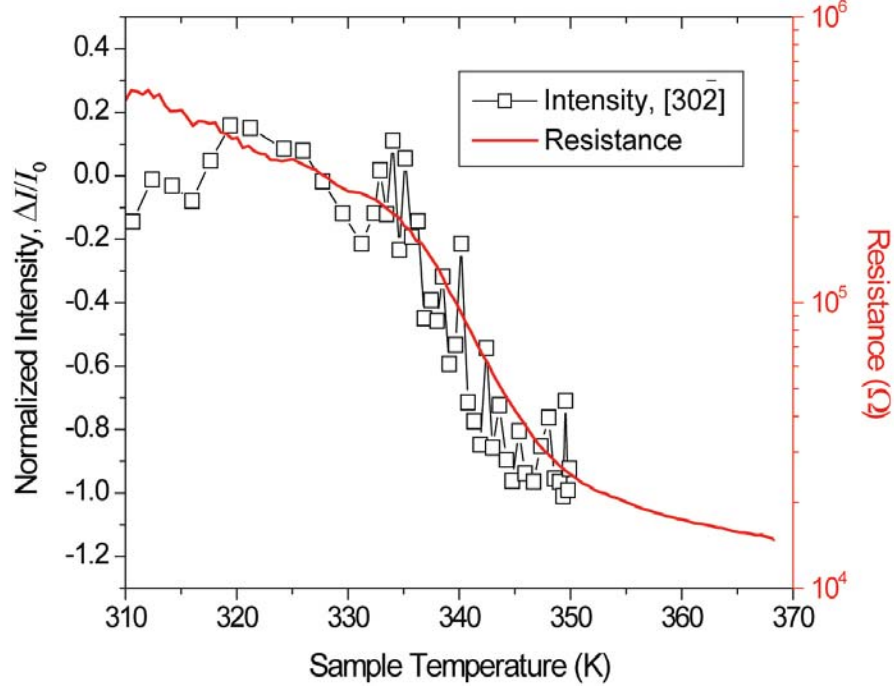


Figure 3.4: Resistance change (red line) and integrated intensity change of the Bragg peak $[30\bar{2}]$ (black symbols) as a function of sample temperature. ΔI is the intensity change of the Bragg peak and I_0 the intensity of the Bragg peak at the room-temperature ground state.

to-rutile structural phase transition in VO_2 . On the other hand, the resistance of the sample is also reduced by approximately one order of magnitude at the similar critical temperature, as shown in Fig. 3.4. The resistance is measured by our collaborators using the standard four-probe method on a reference thin film deposited on a silicon chip with the same growth condition.

3.3 Photo-induced Nonthermal Phase Transition

In addition to thermal excitation, the application of femtosecond (fs) laser excitation can also trigger the phase transitions in VO_2 [74, 114, 118]. The non-equilibrium nature of the photo-induced phase transition was discussed in previous works [114, 115], but it is still an unsolved question whether the photo-induced phase transition is a thermally driven process.

3.3.1 Suppression of Threshold Energy Density with Long Wavelength Pump

Here, we use the UED method to address this question. The UED experiment is operated in the pump-probe manner. The experimental setup is introduced in detail in Sec. 1.1. In addition, an optical parametric amplifier (OPA) is implemented in the optical-pump path, which enables us to adjust the optical pump wavelength λ_p from 300 nm to 2500 nm. The temperature of the sample, T_B , is, on the other hand, controlled and measured through a commercial PID temperature controller (Lakeshore 335 temperature controller).

A typical structural dynamics is shown in the inset of Fig. 3.5 pumped with 800 nm laser with fluence 7.8 mJ/cm². The structural phase transition is marked by the decrease of the integrated intensity of the Bragg peak $[30\bar{2}]$, as we have seen in the thermodynamic process. The structural change completes within 2 ps and reaches equilibrium after ~ 5 ps. The recovery of the structural change is very slow (>1 ms). In the experiments, the laser repetition rate is reduced to 500 Hz to make sure the sample is fully recovered to its ground state before the next pump-probe cycle. In Fig. 3.5, the intensity of the peak $[30\bar{2}]$ is plotted as a function of the absorbed energy density E for three different wavelengths: 800 nm(1.55 eV), 1319 nm(0.94 eV) and 2000 nm (0.62 eV). The intensity is selected at the time delay of 150 ps after the zero-of-time. Here, we make two observations:

1. For all different wavelengths, the intensity holds for small fluences, while it decreases rapidly above the threshold. The solid lines in Fig. 3.5 represent error-function fittings of the experimental data. We define the center of fitting as the critical energy density, E_c , and the width, σ_D .
2. With lower photon energy (longer wavelength) of the pump, $E_c(\lambda)$ decreases.

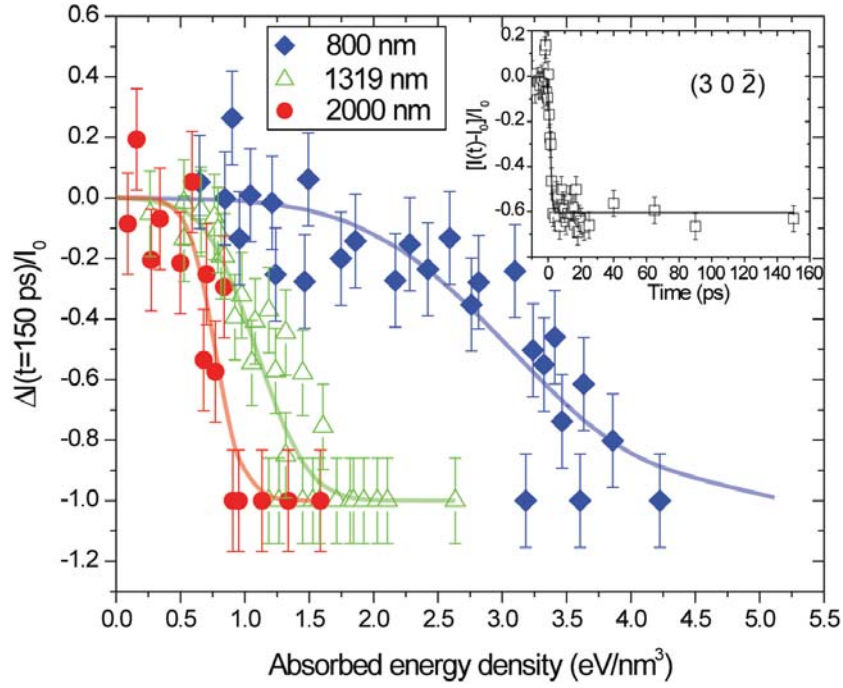


Figure 3.5: Absorbed-energy-density dependence of the integrated peak intensity of $[3 \ 0 \ \bar{2}]$ at 295 K excited by pump photons of different wavelengths. The integrated intensity is selected at 150 ps after pump excitation. Solid lines are the error-function fittings of the experimental data. Inset: Room-temperature transient change of the integrated intensity of peak $[3 \ 0 \ \bar{2}]$, after the excitation by a 50 fs laser pulse centered at 800nm with absorbed fluence 7.8 mJ/cm^2 .

Here, the absorbed energy density is defined as

$$E = \frac{(1 - R)F}{\delta}, \quad (3.3.1)$$

where R is the reflectivity, F the applied laser fluence, and δ the penetration depth. In the equation, $(1 - R)F$ gives the absorbed laser fluence.

3.3.2 Penetration Depth of VO₂ Thin Film

Obviously, the results in Fig. 3.5 can be affected by the estimation of optical constants, R and δ , of VO₂ thin films. The optical constants are available in Ref. [126]. (See Appendix B for detailed derivations.) The values of R and δ for three wavelengths in the experiment are listed in Table 3.3. The reflectivity R is almost a constant for different pump wavelengths, while the penetration depths differ by almost an order of magnitude. This difference is confirmed by the transmittance measurements, as shown in Fig. 3.6. In the measurement, the strength of optical pulses is measured by the photodiodes (Thorlab DET10A for 800 nm, DET10D for 1319 and 2000 nm) and the data are recorded by an oscilloscope (Tektronix TDS3054B). The silicon TEM grid without VO₂ coating is taken as the reference sample. The results, when the optical pulses transmit through a single window of the reference sample, are shown as the dashed lines in Fig. 3.6, while the solid lines are the pulse traces when the optical pulses transmit through the single window with VO₂ thin film. The difference between solid and dashed lines is the contribution from the VO₂ thin film solely. The different time-domain shapes of the pulses in Fig. 3.6(a) and (b), (c) are caused by the different response times of the photodiodes used in measurements.

In Fig. 3.6, it is obvious that a larger portion of incident photons are able to penetrate through the VO₂ thin film, when the wavelength increases from 800 nm to 2000 nm, which is consistent with the increases of penetration depth δ as a function

of wavelength, as shown in Fig. B.2. It is difficult to quantitatively calculate R and δ using the measurement results above, because the optical constants of silicon membrane are unknown. However, we can do a rough estimation as the follows. In Ref. [127], the penetration depth of amorphous silicon is reported to be larger than $1 \mu\text{m}$ in the wavelength range of interest, which means the 9 nm thick silicon membrane makes almost no contribution to the penetration depth. However, its reflectivity R_{Si} counts, because the laser is reflected directly from the silicon membrane when measuring the reference sample. When the reference sample is measured, the transmitted laser pulse intensity is given by

$$I_{\text{Si}} = I_0(1 - R_{\text{Si}}), \quad (3.3.2)$$

where I_0 is the incident intensity. With VO_2 thin film coated, the reflectivity at the interface R_{Inter} needs be taken into consideration, which is given by [128]

$$R_{\text{Int}} = \left(\frac{n_{\text{VO}_2} - n_{\text{Si}}}{n_{\text{VO}_2} + n_{\text{Si}}} \right)^2, \quad (3.3.3)$$

where n_{VO_2} is the refractive index of VO_2 , which is ~ 3.1 for the wavelengths of interest (see Appendix B), n_{Si} the refractive index of amorphous silicon, ~ 3.7 [129]. So, $R_{\text{Inter}} \sim 0.01$, which is negligible. The transmitted pulse intensity through VO_2 coated membrane is then given by

$$I_{\text{VO}_2} = I_0(1 - R_{\text{VO}_2})(1 - R_{\text{Inter}})e^{-\frac{d_{\text{VO}_2}}{\delta}}. \quad (3.3.4)$$

Thus, using the reflectivity of VO_2 , $R \approx 0.24$, and of amorphous Si, $R_{\text{Si}} \approx 0.32$ [127], the penetration depth of the VO_2 film is given by

$$\frac{d_{\text{VO}_2}}{\delta} = -\ln\left[\frac{1 - R_{\text{Si}}}{(1 - R)} \frac{I_{\text{VO}_2}}{I_{\text{Si}}}\right], \quad (3.3.5)$$

where d_{VO_2} is the thickness of VO_2 film, $I_{\text{VO}_2}/I_{\text{Si}}$ the ratio of pulse intensities for the

Wavelength (nm)	Photon energy (eV)	Reflectivity R	Penetration depth, δ (nm)
800	1.55	0.24	156
1319	0.94	0.28	270
2000	0.62	0.28	1036

Table 3.3: List of optical constants of VO₂ M₁ phase at different wavelengths

samples with VO₂ thin film and the reference samples, which is shown in Fig. 3.6. The penetration depth is estimated as follows: $\delta(800\text{nm}) \approx 71$ nm, $\delta(1319\text{nm}) \approx 125$ nm and $\delta(2000\text{nm}) \approx 603$ nm. The values here are approximately a factor of two less compared to the literature values listed in Table 3.3, that the filling factor of the VO₂ film can be the reason. It was found in similar VO₂ films that the nanograins in the film do not cover the whole substrate, but leave vacancies in the film [130, 131]. However, the rough estimation confirms that the penetration depth for 2000 nm photon is approximately one order longer than 800 nm photon, which is robust against the variations of optical constants of the silicon membrane. In addition, since the volume optical energy density is used in the analysis, the filling factor of the thin film does not affect the following analysis. For the purpose of consistency, in the following analysis, the optical constants in Table 3.3 will be used.

3.3.3 Temperature Dependence of Critical Energy Density

The critical energy-density measurements are repeated at different sample base temperature, T_B , for 800 nm and 2000 nm pump wavelengths. Raising the sample temperature close to T_c leads to a decrease of the critical energy density E_c for both wavelengths, as shown in Fig. 3.7. On the other hand, the photo-induced phase transition is suppressed, when the VO₂ is cooled down to cryogenic temperature. Over a wide range of

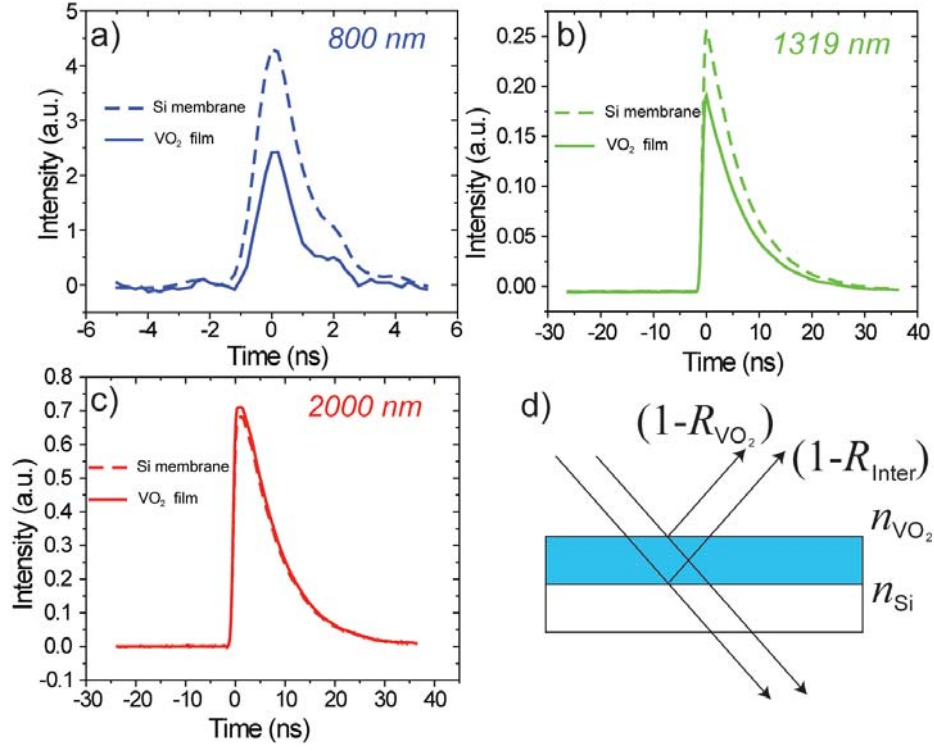


Figure 3.6: Transmittance measurement of 800 nm (a), 1319 nm (b) and 2000 nm (c) on 1) reference sample with supporting silicon membrane (dashed lines); 2) VO₂ thin-film membrane (solid lines). (d) Illustration of the optical reflection and transmission when the VO₂ thin-film membrane is measured. n_{VO_2} and n_{Si} are the refractive indices of VO₂ and amorphous silicon, respectively. Reflectance on the interface is R_{Int} .

temperature, the critical energy density using 800 nm photoexcitation is always higher than the value when using 2000 nm photoexcitation. The value of E_c we report in Fig. 3.7 for 800 nm is in good agreement with previous time-resolved electron diffraction [118] and THz spectroscopy [114, 115, 130] measurements, if the optical constants in Table 3.3 are used in conversion. Since E_c is the energy required to induce the phase transition using photoexcitation, it is instructive to compare it with the thermodynamic requirement, E_{th} . E_{th} can be calculated as the integral of the lattice heat capacity between a given temperature and T_c plus the latent heat, H_L

$$E_{th}(T_B) = \int_{T_B}^{T_c} c_v(T) dT + H_L, \quad (3.3.6)$$

where $c_v(T)$ is the heat capacity of VO_2 . Based on the 750 K Debye temperature [132] and room temperature heat capacity $3.0 \times 10^6 \text{ J/m}^3 \text{ K}$ [133], the heat capacity $c_v(T)$ of insulating VO_2 can be calculated at all temperatures. The latent heat is also known to be $H_L = 235 \times 10^6 \text{ J/m}^3$ [133]. Thus, the minimum required energy density, in the unit of $[\text{eV/nm}^3]$, is calculated and plotted in Fig. 3.7 as the orange line. The thickness of the line is determined by considering the finite temperature window for the phase transition around T_c . Compared with the experimental results, the thermodynamic requirement E_{th} approximately agrees with the room-temperature and low-temperature E_c when 800 nm pump is applied. This result is also consistent with previous studies [114, 115, 130]. However, this agreement deviates, when T_B is raised to close to T_c . Especially, the experimental measured E_c is even less than the latent heat H_L , when T_B is very close to T_c . Moreover, when the 2000 nm pump is used for excitation, the critical energy density E_c is always less than the thermodynamic requirement E_{th} over the temperature range in our experiments.

The results in Fig. 3.7 indicate that the photo-induced phase transition is different from the thermodynamic process, since it does not require as much as the thermody-

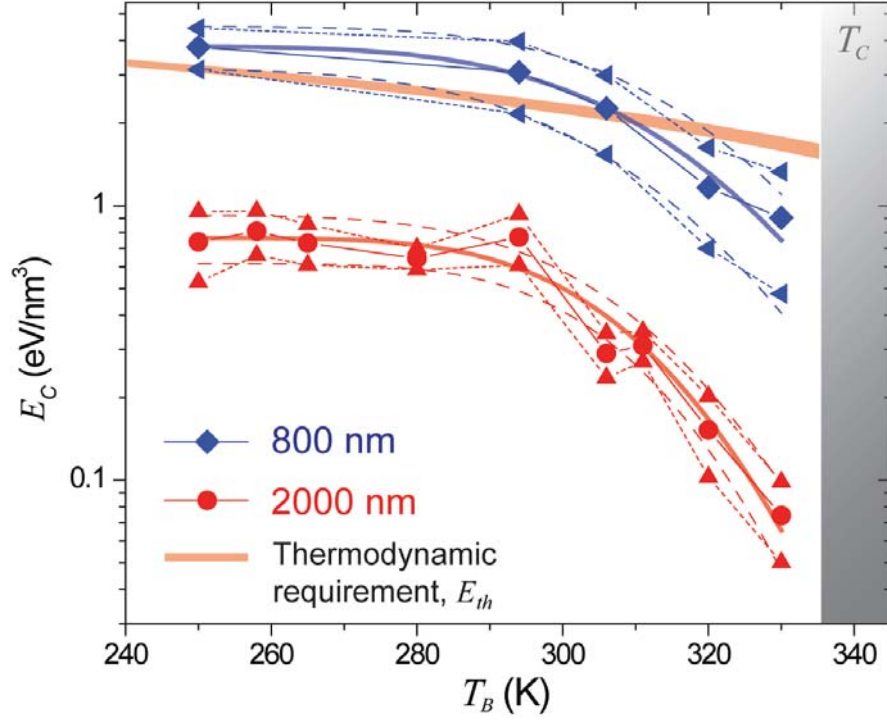


Figure 3.7: Dependence of the critical energy density E_c on the sample temperature T_B . The triangle symbols represent the width of the energy-density window for the phase transitions, which is extracted by the width from the error-function fittings in Fig. 3.5. The orange line is the thermodynamic energy requirement for the phase transition calculated according to Ref. [133].

namic energy to induce the phase transitions. In Fig. 3.7, the critical energy density E_c for 800 nm pump wavelength is, in general, ~ 4 times larger compared to the values for 2000 nm wavelength. Considering the photon energies of the two wavelengths, the critical photon density N_c can be converted from E_c . We find that N_c is approximately twice as high as the values with 2000 nm pump. So, from the experimental results, it is not conclusive to claim that the photo-induced phase transition is driven by the density of photo-excited charges. But we can make the following arguments. First, at low temperature, where the thermal excitation of carriers is not significant, the critical (photo-induced) carrier density asymptotically approaches $2.3 \text{ e}^-/\text{nm}^3$ and $1.2 \text{ e}^-/\text{nm}^3$ for 800 nm and 2000 nm pumps, respectively, assuming each incident photon yields one electron. The critical carrier density here agrees well with the carrier density in the high-temperature metallic phase of VO_2 [131, 134], which implies the requirement of photo-induced phase transition is to achieve the critical charge density equaling to the metallic free carrier density in VO_2 . Secondly, we can assume only those itinerant electrons in the π^* band contribute to screening the strong electron correlations. With 800 nm pump excitation, the photon energy is so high that most of the electrons are directly pumped to the upper $d_{\parallel}^*$ band, and are required to be scattered to the π^* band with lower energy to contribute to the screening, as shown in Fig. 3.8. Even though these two bands slightly overlap in energy scale, the scattering process might require the involvement of lattice phonons due to different orientations of the two orbitals. The difference of N_c between 800 nm pump and 2000 nm pump might be attributed to the efficiency of this scattering process.

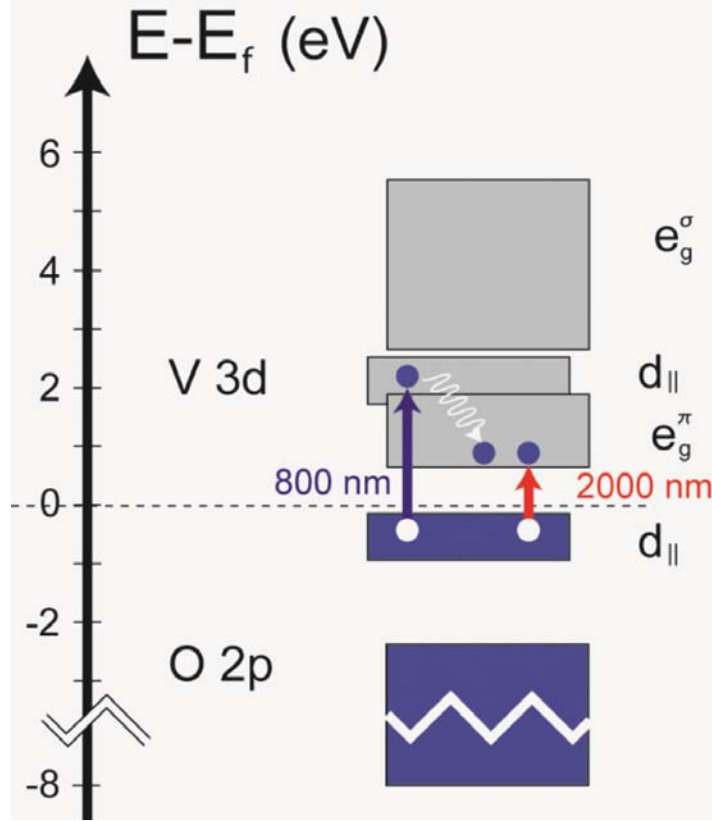


Figure 3.8: Electronic band structure of VO_2 M_1 phase [57, 115]. With 2000 nm (0.62 eV) pump excitation, the excited electrons can efficiently fill in the π^* band across the band gap, while the photon energy of 800 nm (1.55 eV) pump is so high that the electrons are directly excited to the high-energy $d_{||}^*$ band. An additional scattering process (white line) is required to transfer these electrons to the π^* band and contribute to screening the strong electron correlation.

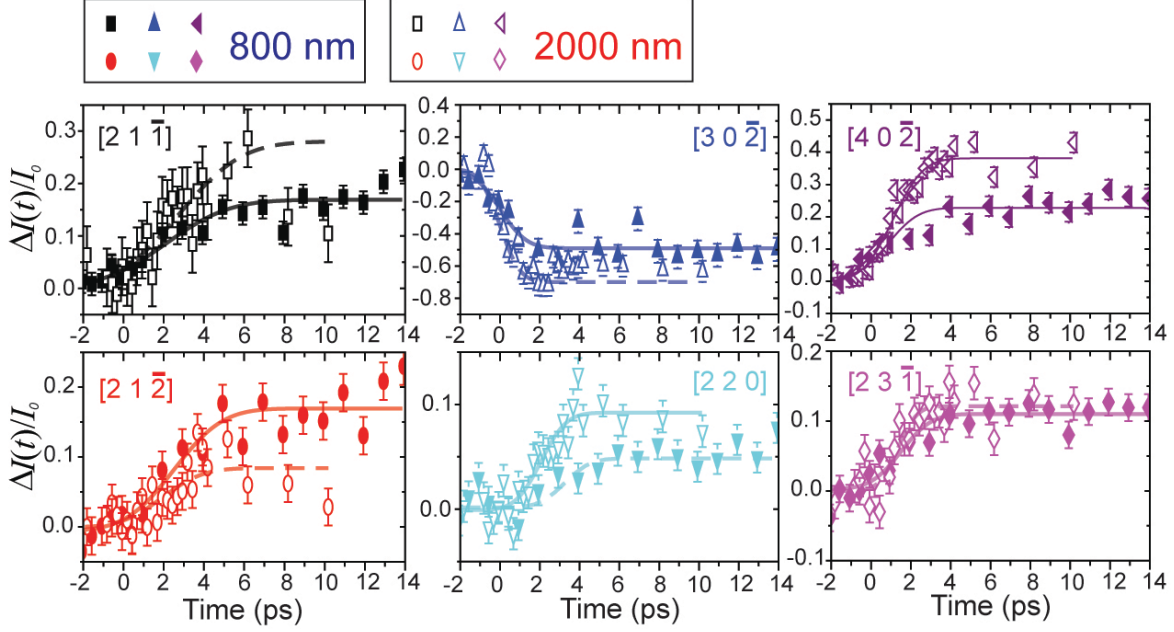


Figure 3.9: Comparison of Bragg peak intensity dynamics of different Miller indices for 800 nm pump and 2000 nm pump excitations.

3.4 Atomic Pathway of Photo-induced Structural Phase Transition in VO₂

3.4.1 Comparing Atomic Movements with Pumps of Different Wavelengths

In Fig. 3.7, we found that the critical energy density E_c for 800 nm photons is close to the thermodynamic requirement E_{th} , while it is much smaller for 2000 nm photoexcitation. This result brings up a question whether the photo-induced VO₂ phase transition undergoes different pathways with these two wavelengths. UED provides us with a useful tool to resolve this question. If the phase transition pathways are different, we would expect the atoms in VO₂ to move differently after pump excitation.

According to the Bragg's law and the Laue diffraction conditions, every Bragg reflection corresponds to a reciprocal lattice point, whose intensity represents the co-

herence of a set of crystal planes along a specific direction, and this direction is specified by the Miller indices. The atomic movements in the direction normal to the crystal plane will break the coherence of the planes, hence, affect the intensity of the Bragg reflections, while the intensity is less affected if the movements happen in the plane. With the powder diffraction pattern as shown in Fig. 3.3(a), movements happening in multiple sets of crystal planes can be recorded at the same time. It is like monitoring the structural dynamics in various directions, which gives a very stringent confinement on the problem. In Fig. 3.9, the changes of peak intensity of 6 Bragg reflections as a function of the time delay are plotted for both 800 nm and 2000 nm pump excitations. The pump fluences are chosen to be close to the critical energy density for both wavelengths. If some atoms move in different directions for one of the wavelengths, we would expect the intensity change of a specific Bragg peak to be different for this pump wavelength compared to the other one. However, the results in Fig. 3.9 clearly show that the polarities of the changes of Bragg peak intensity are all the same for all 6 Bragg peaks, even though the magnitudes of change are slightly different. From this result, we can conclude that the phase-transition pathways for 800 nm excitation and 2000 nm excitation are very similar within the resolution of our experiments.

3.4.2 Atomic Pathways for VO₂ Structural Phase Transition

According to the kinematic diffraction theory, which has been described in Sec. 1.2, the intensities of Bragg peaks are determined by the (fractional) atomic coordinates in the unit cell. Therefore, the sequence of atomic movements can be extracted from the experimental data, by fitting the intensity changes of multiple Bragg reflections using the kinematic diffraction theory.

First, the unit cells of VO₂ M₁ structure and rutile structure need to be defined. Fig. 3.10 illustrates the difference in V atoms' positions in VO₂ monoclinic and rutile structures. In a single unit cell of the rutile structure, two V and four O atoms are

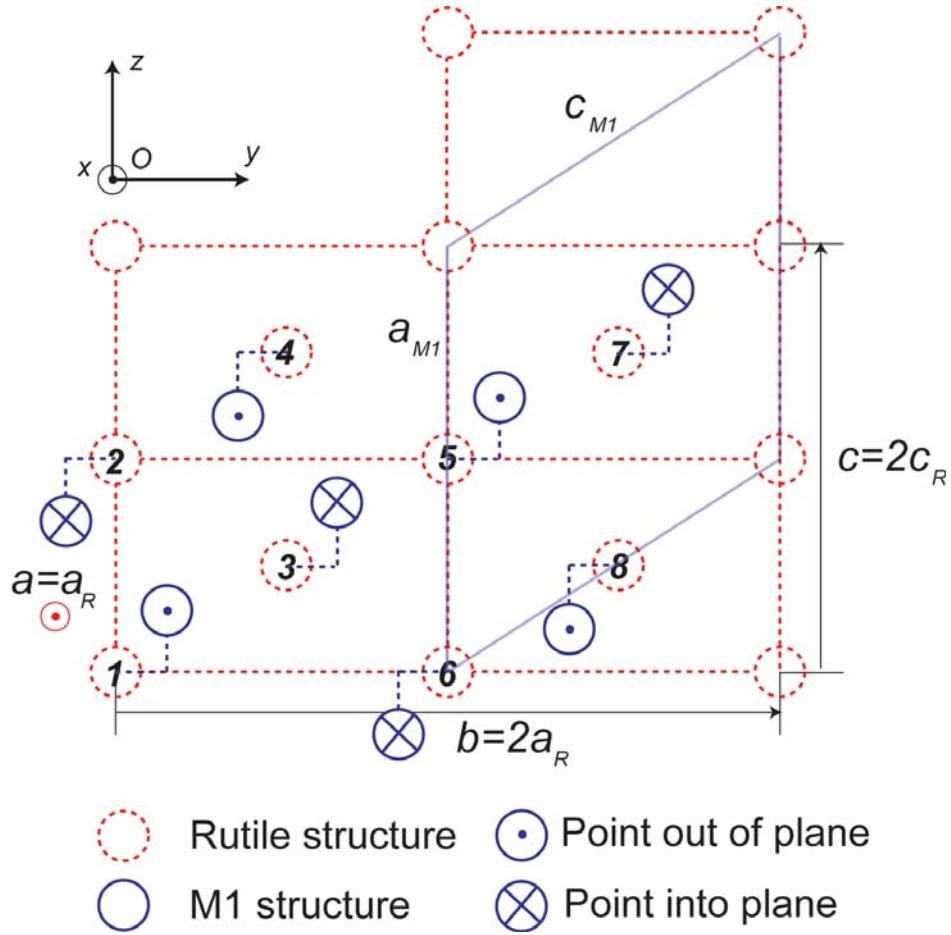


Figure 3.10: Structures of VO_2 monoclinic and rutile structures. The red dashed lines define the unit cell of rutile structure. The blue parallelogram defines the unit cell of monoclinic structure. The blue dashed lines represent how the vanadium atoms are distorted from rutile structure to monoclinic structure.

contained. The unit cell is defined by the red dashed rectangular boxes in Fig. 3.10. Because of the V-V dimerization, hence, the symmetry breaking, a M_1 unit cell contains two rutile unit cells, which are stacked along the c_R (z) axis. Thus, four V atoms and eight O atoms are contained. Its unit cell is defined by the blue parallelogram in the figure. When the structures changes from M_1 to R, there exists a rotation of lattice vector, which is not very convenient to deal with in the simulation. So, we define a supercell, which is constructed by stacking 2 rutile unit cells along both y and z directions, as shown in Fig. 3.10. The supercell, thus, contains eight V atoms and sixteen O atoms. The supercell vectors are defined as follows

$$\begin{aligned}\vec{a} &= \vec{a}_R, \\ \vec{b} &= 2 \times \vec{b}_R, \\ \vec{c} &= 2 \times \vec{c}_R.\end{aligned}\tag{3.4.1}$$

The atomic coordinates inside the supercell are listed for both M_1 and R structures in Table A.7. With the supercell definition, the atoms are allowed to be moved within the same cell structure continuously and convert the VO_2 monoclinic structure to its rutile structure.

In a supercell, there are in total 24 atoms allowed to move, which is a formidable problem due to the huge amount of the degrees of freedom. So, additional confinements have to be placed in the simulation by the following approximations:

1. The supercell size does not change throughout the structural phase transition.
2. Only the V atoms move, while the O atoms are fixed.
3. The movements of the equivalent V atoms in M_1 unit cell are the same.
4. The V atoms move in pairs and their center-of-mass does not move.

It was reported in previous X-ray crystallography studies that there is a lattice

strain through the VO_2 's structural phase transition, and it leads to the consequence that the unit-cell size along the c_R axis of the M_1 structure is not exactly twice the size of the rutile structure ($a_{M1}/(2c_R) = 1.0070$) [122]. The approximation 1. means the lattice strain is neglected and this ratio is set to be unity. By doing so, we basically ignore the shifts of the Bragg peaks' positions in the simulation, and only their intensities are of interest. Approximation 2. is validated, because 1) the O atoms are relatively lighter electron scatterers compared to the V atoms, and contribute less to the diffraction intensity; 2) the displacements of the O atoms are almost an order of magnitude less compared to the displacements of the V atoms. Approximation 4. can be validated by comparing the atomic coordinates of M_1 and R structure in Table A.7. It can be found that, if the change of unit-cell size is ignored, the V atoms move in pairs within the V-V dimers, and their center-of-mass position is approximately fixed. Approximation 3. is a condition we imposed, by assuming the super-cell symmetry is not changed throughout the structural phase transition. In principle, the atoms can move along any arbitrary paths from the M_1 structure to the rutile structure. But when the atoms, which are equivalent in M_1 unit cell, move differently, the unit cell of the structure will have to be defined larger due to the symmetry breaking. Here, the approximation 3. is applied to eliminate this situation. In Fig. 3.10, the atoms labeled with numbers 1-8 are pairwise equivalent: "1"- "5", "2"- "6", "3"- "7" and "4"- "8".

In addition to the four approximations, the observation that the Bragg peaks $[30\bar{2}]$ and $[31\bar{3}]$ are solely contributed by the V-V dimers in the monoclinic structure also helps us to simulate the atomic movements after photoexcitation. In the simulation, the following strategy is applied. The V-V dimer pairs are moved along the predefined paths and the Bragg peaks' intensities are calculated after each step of movement. The step size of movement is adjusted to fit the intensity change of the peaks $[30\bar{2}]$ and $[31\bar{3}]$ recorded in experiments. Then, by comparing the simulation results with the dynamics of other Bragg reflections, a judgement is made on whether the predefined path is a

good fit to the experimental results.

Here, several possible paths of V-V dimers are investigated and compared with experimental data:

1. Directly move from the positions in the M_1 structure to the positions in the rutile structure.
2. First tilt in the direction perpendicular to the c_R axis, then stretch along the c_R axis.
3. First stretch along the direction of V-V dimers, then tilt to the positions in the rutile structure.

The schematics of the first and second pathways are plotted in Fig. 3.11(a) and (b), respectively. Half of the supercell is plotted, while the atoms in the other half supercell shall move following their structurally equivalent atoms in the first half due to Approximation 3. Fig. 3.11(a) and (b) shows the comparison of the simulation results with the experimental data. The step size of each movement is adjusted to get the best fit to the intensity changes of the peaks $[30\bar{2}]$ and $[31\bar{3}]$. Clearly, neither of the situations agrees with the experimental observations. In Fig. 3.11(a) and (b), it is found that the movements in the direction normal to the c_R axis quickly drive the intensity increase of the peaks $[220]$, $[23\bar{1}]$ and $[40\bar{2}]$, which does not agree with the experimental results. Especially for peak $[220]$, which contains little contributions from the atomic movements in the c_R direction, the intensity change stays very little for the first ~ 3 ps. This observation indicates that the movement direction right after the pump excitation must deviate from the c_R direction.

For the third possible path under consideration, a three-step scenario is proposed to optimize the fitting to the experimental results. The three-step scenario is plotted in Fig. 3.12(a), where the V atoms first move away from each other along the axis of V-V dimers. At time delay ~ 1 ps, the movement direction is switched to a direction

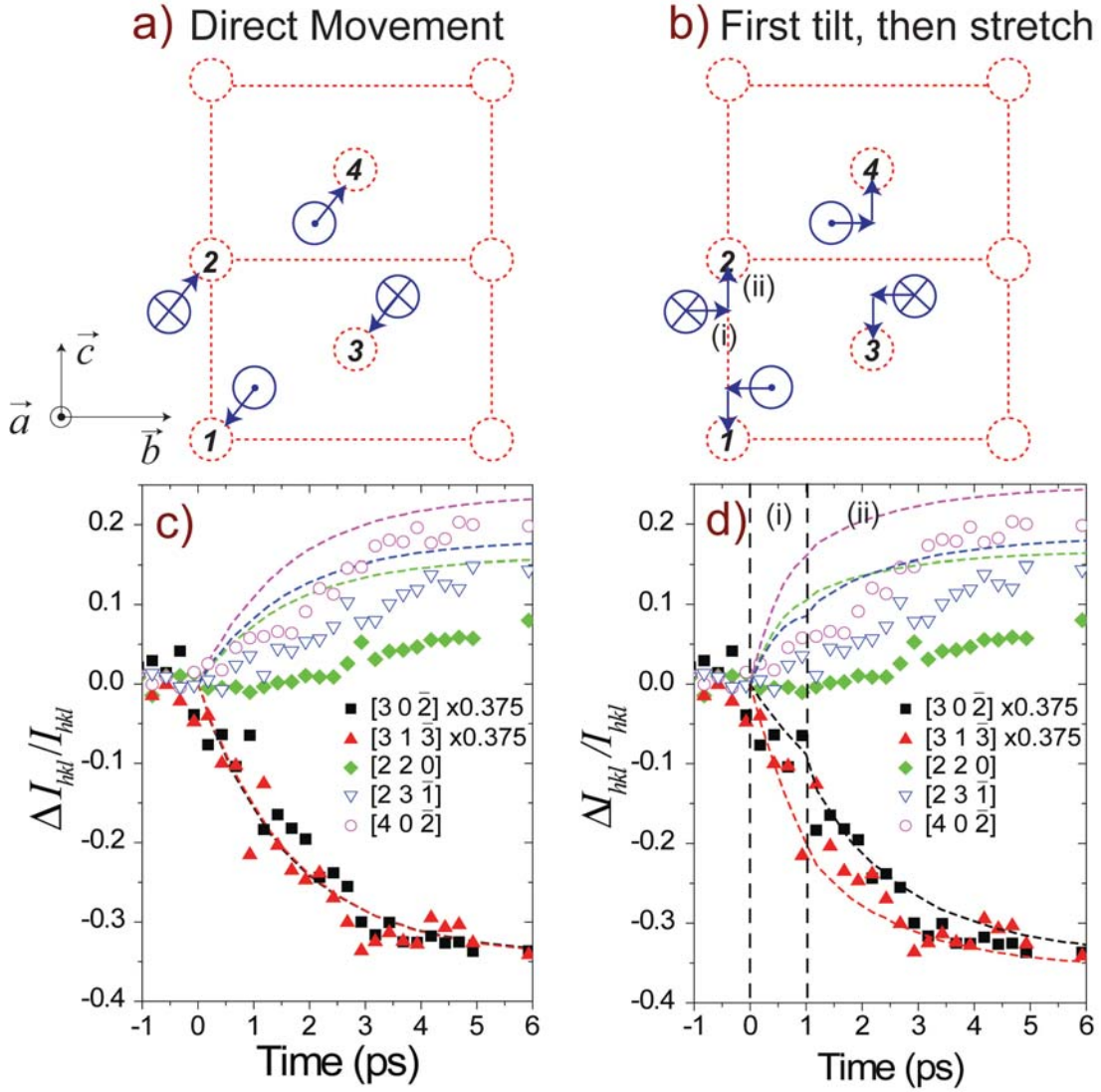


Figure 3.11: Schematic diagram of the atomic movements in the case a) directly move from the monoclinic positions to the rutile positions; b) first tilt perpendicular to the c_R axis, then stretch along it. Half of the supercell is plotted, and the V atoms in the other half move following their structurally equivalent atoms. The blue arrows represent the directions V atoms move. c) Comparison of the simulated intensities with the experimental results in the case of direct movement. The experimental results are obtained by pumping with 800 nm laser at the fluence 7.8 mJ/cm². Dashed lines are the simulation results. Symbols are the experimental results. The intensity changes of the peaks $[30\bar{2}]$ and $[31\bar{3}]$ are scaled by 0.375 in the plot. d) Comparison of the calculated intensities with the experimental results in the case of tilting prior to stretching scheme. The change of the movement directions happens at 2 ps, which is marked by the vertical dashed lines.

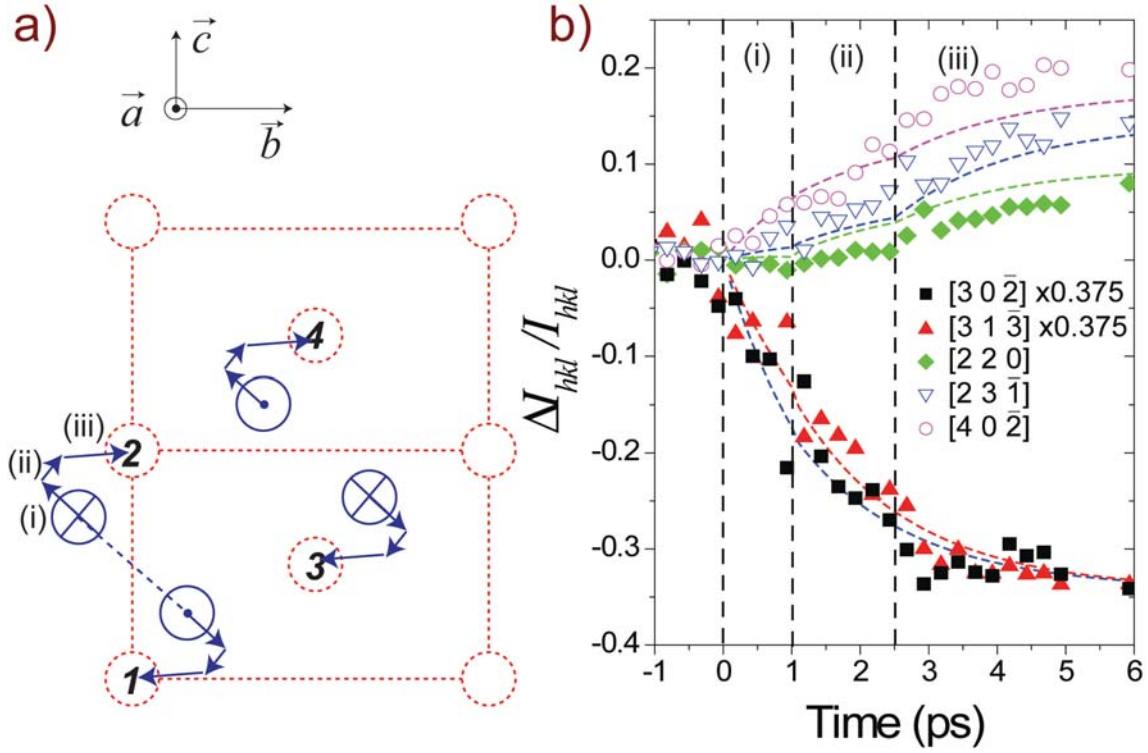


Figure 3.12: a) Schematic diagram of a three-step movement scenario. The V atoms first move away from each other along the dimerization direction. The second step happens at ~ 1 ps after pump excitation, in which the V atoms start turning into the direction perpendicular to the dimerization. At ~ 2.5 ps time delay, the movement direction turns for the second time that the V atoms move towards the rutile structure positions. b) Comparison of the calculated intensity changes (dashed lines) for the scenario provided in a) with the experimental results (symbols).

approximately perpendicular to the original direction. And after another ~ 1.5 ps, the movement direction changes for the second time and ultimately drives the V atoms to rutile structure positions. Again, the movement step size is optimized to fit the intensity changes of the peaks $[30\bar{2}]$ and $[31\bar{3}]$. With this three-step scenario, a reasonable agreement between the experimental results and the simulation results can be achieved.

3.5 Discussion and Summary

In this chapter, VO₂ thin-film membrane has been investigated using the UED method. By using optical pumps with different wavelengths, we try to address the question whether photo-induced VO₂ phase transitions are the same as the thermodynamic process. From the results of fluence dependence, a threshold behavior is identified, where a critical energy density E_c is required to induce the structural phase transitions in VO₂. In addition, the critical energy density is suppressed when the pump wavelength gets longer. It is found that E_c for 2000 nm pump is always lower than the 800 nm over a wide range of temperature. More importantly, it is lower than the thermodynamic energy requirement, E_{th} . This result indicates that the photo-induced phase transitions in VO₂ are very different from the ones with thermal excitation. Previous studies already indicate that in a short period after pump excitation, the system is in a strongly non-equilibrium state, where the electrons move regardless of the position of ions and the Born-Oppenheimer approximation is no longer valid [114, 115]. In addition, the results obtained in the time-resolved optical and photoemission experiments also indicate that the metal-to-insulator transition occurs in a period as short as 80 fs, after pump excitation [74, 135]. However, the first step of the atomic movements does not finish in ~ 1 ps, which is indicated by our UED measurements. The difference in timescales tells us that the insulating band gap closes a way before the structural changes are finished. So, what we observed here is a decoupling of the metal-to-insulator transition from the structural phase transition in an ultrafast timescale. This result also indicates that the photo-induced band-gap collapsing is more likely driven by the photo-excited electrons, which screen the electronic correlations. Moreover, with the ultrafast collapsing of the insulating band gap, there are a huge number of itinerant electrons traveling in the space among atoms, which can be orders of magnitude higher compared to the number in the insulating phase of VO₂. The energy landscape of atoms in the structure can be

strongly modified. The modification leads to a decreased energy requirements for the phase transitions compared to the thermodynamic process. So, we propose that the photo-induced phase transition in VO_2 is a non-equilibrium charge-driven process.

Comparing the atomic dynamics induced by pumps of 2000 nm and 800 nm wavelengths, we found they go through similar processes. Attempts are made to fit the experimental results with the simulation results using the kinematic diffraction theory, by comparing several possible atomic pathways. It is found that the V atoms first move along the their dimerization directions in the first 1 ps. The stepwise movements were also addressed in pioneer UED measurements on VO_2 single-crystal film [118]. However, there exist subtle difference in the atomic dynamics, which requires further efforts to elucidate. One possible reason to the difference might be because of the sample differences, where single-crystal films grown on sapphire substrates were studied in their experiments.

Here, one more note on our results. Even though a reasonable agreement has been achieved with a three-step movement scenario, we cannot completely exclude other possible pathways, if the structurally equivalent atoms are allowed to move in different ways. UED experiments on single-crystal samples will be more preferable to extract more information on the structural dynamics.

Chapter 4

Decoupling of Electronic and Structural Phase Transitions in VO₂: an Investigation on Single-crystal Nanobeam VO₂

In Chapter 3, the photo-induced phase transitions in VO₂ thin-film membrane were investigated using the UED. It is found that the photo-induced phase transitions are different from those induced by thermodynamic process as a highly non-equilibrium state is found to be excited by a pump laser, which reduces the energy barrier for the VO₂ phase transition. To study the relationship between electronic and lattice degrees of freedom, the thin film sample with nanogranular morphology is found to be a limitation, due to the complicated strain and composition environments of individual nano-grains, which leads to convoluted observations in the experiments [71, 72, 104]. So, single-crystal samples with well-controlled properties are preferred to study the fundamental mechanisms of phase transitions in VO₂.

Only in recent years, the single-crystal nanobeams became available due to the

advancement of synthesis techniques [72, 136, 137]. These VO₂ nanobeams not only have high-quality crystalline structure, but also provide a way to actively control the VO₂ phase transitions by applying external strain [104, 138], which have been a difficult problem on VO₂ single-crystal or poly-crystalline thin-film samples. It has been experimentally demonstrated that, by tuning the applied strain, VO₂'s phase diagram can be extended to a range that was unavailable to the film samples [104].

In this Chapter, VO₂ single-crystal nanobeams are investigated using combined experimental methods, including optical microscopy, electron microscopy and ultrafast electron diffraction (UED). We will start with an introduction to the questions concerning the mechanisms of VO₂'s phase transition in Section 4.1, which has been a debate for decades in the history of research on strongly correlated materials. In Section 4.2, we will present the experimental evidence of the decoupling of the metal-to-insulator transition (MIT) and the structural phase transition (SPT) in VO₂ single-crystal nanobeams on different substrates. More recently, more and more experiments have shown that the MIT and SPT can be decoupled at ultrafast timescales or microscopic scales. In Section 4.3, these experiments will be briefly discussed. Further investigations of the nanobeams on different substrates using UED are presented in Section 4.4, which provides us with some important understanding about the phase transitions in VO₂. A summary is given in Section 4.5.

4.1 VO₂: Peierls Insulator or Mott Insulator?

It has been a long debate whether VO₂ is a *Peierls insulator* or a *Mott insulator*. In 1971, Goodenough first proposed that the MIT of VO₂ is induced by the antiferroelectric distortion of vanadium ions [57]. The changes of the electronic band structure, based on the crystal-symmetry considerations, are plotted in Fig. 4.1. In the high-temperature rutile phase, due to the crystal field from oxide ligands surrounding

the vanadium ions, the d orbital of the vanadium ion is split into two sets: twofold-degenerate states of e_g symmetry and threefold-degenerate states of t_{2g} symmetry. Because the VO_6 octahedral sites share common edges only along the c_R axis, there is an orthorhombic component of the crystal field that completely removes the d orbital degeneracy. The consequence is that the e_g orbital set is split into d_{z^2} and $d_{x^2-y^2}$ orbitals, which point to the oxide ligands and form strong σ bonding and antibonding in the form of a $\text{V}_d\text{-O}_p$ molecular orbital. At the same time, the t_{2g} set is split and gives rise to a sideways π bonding between the oxide ligands and d_{xz} , d_{yz} orbitals, whereas the d_{xy} orbital is projected along the c_R axis, connects adjacent vanadium atoms at an equal distance 2.85 Å on either side, and gives rise to the $d_{\parallel}$ band in the band structure. In the metallic state, the $d_{\parallel}$ level is non-bonding, and the electrons occupying $d_{\parallel}$ level are considered as itinerant electrons in metallic VO_2 . When VO_2 is cooled through the critical temperature, subtle changes in lattice structure happen, which leads to dramatic modifications of the VO_2 band structure. First, the vanadium ion in the oxide octahedral cage moves away from the center towards one edge. This change helps the π bonding through overlaps between V t_{2g} and O $2p$ orbitals and raises the antibonding π^* band. More importantly, the anti-parallel displacement of the vanadium ions along the c_R axis doubles the unit-cell size in that direction. The broken translational symmetry splits the $d_{\parallel}$ band into two sub-bands across the Fermi level (*Peierls instability*). Thus, a band gap is open between the highest occupied states, $d_{\parallel}$, and the lowest unoccupied states, π^* , as shown in the band structure of the low-temperature insulating VO_2 in Fig. 4.1.

Goodenough's elegant theory can qualitatively explain the mechanism of VO_2 's MIT, but it was soon challenged by new theoretical and experimental discoveries. First, density functional theory (DFT) calculations in the local density approximation (LDA) confirmed that monoclinic structure is the true ground state structure of VO_2 , however, they failed to open a band gap for the insulating M_1 phase without exagger-

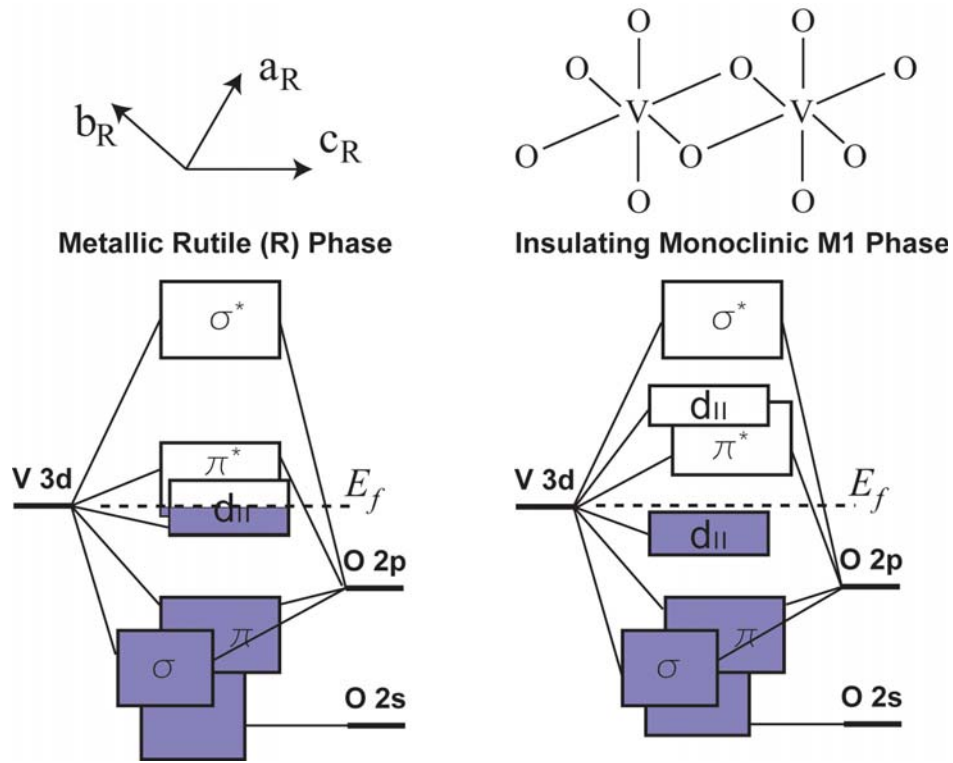


Figure 4.1: Molecular orbital diagram depicting the electronic structure of the monoclinic and rutile phases of VO_2 . The left MO diagram corresponds to the undistorted metallic phase of VO_2 , whereas the diagram on the right shows the altered MO diagram upon transition to the distorted insulating phase of VO_2 [57, 139].

ating the lattice distortions [140]. On the experimental side, the M_2 phase was found by J. P. Pouget, et al., in Cr-doped VO_2 compounds ($\text{V}_{1-x}\text{Cr}_x\text{O}_2$) [141] and was further stabilized in pure VO_2 samples by applying an uniaxial strain [103]. The phase diagram of VO_2 is shown in Fig. 4.2(b). The M_2 phase has a bizarre structure, in which the V atoms are paired through alternate displacement along the rutile c_R axis (no tilting of the pairs) on one set of the vanadium chains,, while, on the other set the V atoms are placed in a zig-zag fashion through alternate displacements perpendicular to the c_R axis (no pairing). The resulting structure is presented in Fig. 4.2(a). In the physical picture provided by Goodenough, the M_2 phase is supposed to be conducting, because of the undimerized vanadium chains. However, in the M_2 phase, VO_2 is an insulator . Moreover, DFT-LDA calculation on VO_2 's M_2 phase also fails to yield a band gap. As has been discussed in Sec. 2.1.4, when dealing with materials with strong electronic correlations, DFT usually fails in making precise predictions. All these observations suggest that the absence of Peierls-type structural dimerization does not necessarily predicate the induction of metallic transport in VO_2 . This situation inspired the invocation of a Mott-Hubbard picture, in which the strong electron correlation plays an important role in opening the insulating gap [142]. The Hubbard model was discussed in detail in Sec. 2.2. Because every two V atoms contain one d valence electron in VO_2 , if the Coulomb energy of the nearest neighbor is considered, the Hubbard model predicts a superstructure in the real-space electronic distribution to keep the total energy low in which the electrons occupy every other V atom. This prediction has two consequences: 1) the narrow d band is split into the upper Hubbard band and the lower Hubbard band with a band gap of size $\sim \text{Hubbard } U$; 2) the translational symmetry is broken along that direction. And very interestingly, both of these have been observed in VO_2 . Moreover, the competition between its iterative ground state and localized ground state naturally gives rise to the MIT in VO_2 . In an DFT LDA+ U calculation, a typical Hubbard U of ~ 4 eV is required to open an insulting band gap

in VO₂'s monoclinic phase [69].

In addition, the influence of the strong electronic correlations is underlined by several experimental observations. First, the metallic phase of VO₂ exhibits relatively low conductivity and this characteristic ‘bad metal’ behavior suggests the heavier fermion mass due to the electron-electron correlation [71]. Next, in the time-resolved THz or ultrafast diffraction experiments with near IR optical pump excitation, it has been shown that a critical threshold flux is required to trigger the phase transitions [114, 118, 144], which is consistent with the Mott-Hubbard picture, in which critical density of free carriers are needed to screen the on-site electronic correlation for the phase transitions to happen.

For decades, theorists spent most of their time developing a theoretical model that can capture the nature of phase transitions in VO₂: closure of the ~ 0.6 eV insulating gap and dissolving of V-V dimer clusters. As mentioned above, the attempts with DFT were not very successful in that DFT LDA failed to explain the insulating gap of VO₂ in its monoclinic structure. The failure was cured by the involvement of DFT LDA+ U but, ironically, it resulted in an insulating state for the rutile structural [63, 105] as well. It is believed that its limited ability to treat the frequency-dependent electronic interaction is the reason for the failures, when treating many strongly correlated materials using the DFT LDA+ U method [59]. In this context, the cluster dynamical mean field theory (c-DMFT) calculation by Biermann et al. makes a major advance in that it explicitly takes into consideration the k -dependent self-energy correction due to the Coulomb interaction of electrons on the V-V dimer sites [69]. His predictions on the orbital occupancy in both metallic and insulating phases were further confirmed by studies using polarization-dependent X-ray absorption spectroscopy (XAS) [70] and bulk-sensitive valence band photoemission spectroscopy (PES) [58]. In these two experiments important features of VO₂'s MIT were captured: 1) The band gap was observed in the insulating state, while a strong quasiparticle peak exists around

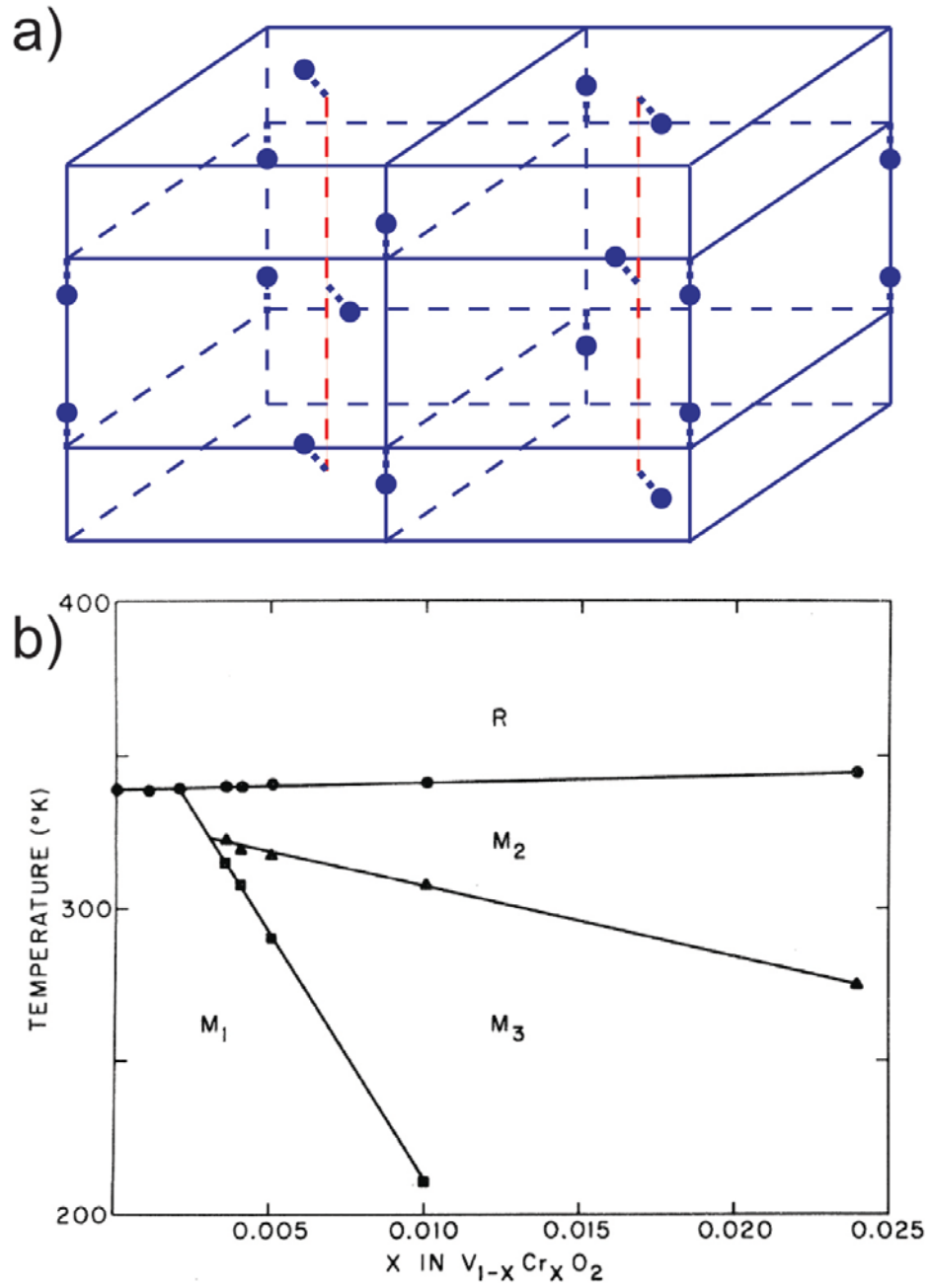


Figure 4.2: a) Monoclinic unit cell of VO_2 M_2 phase. The displacement of vanadium atoms from equilibrium position of its rutile structure is presented. b) Phase diagram of $V_{1-x}Cr_xO_2$ [143].

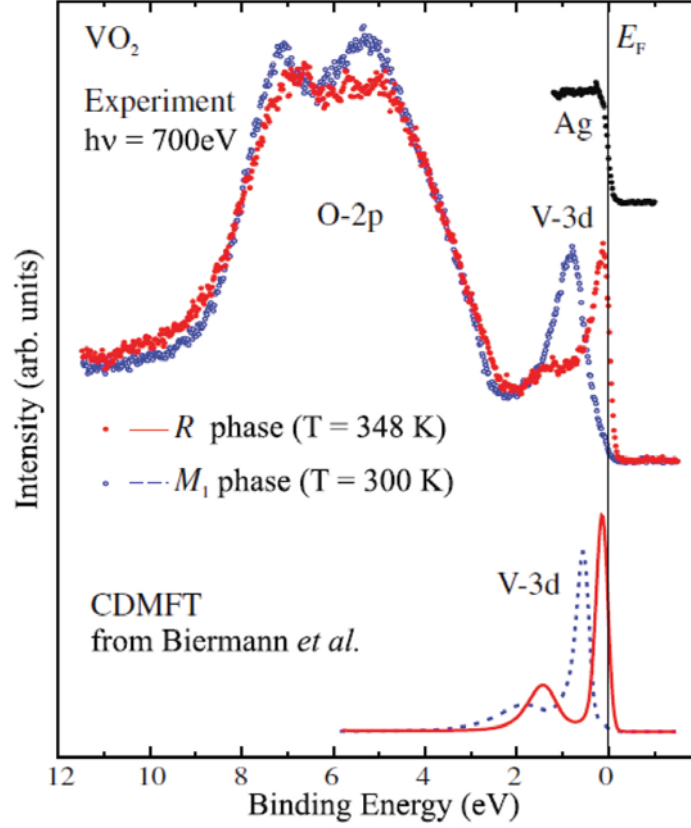


Figure 4.3: Comparison of experimentally measured photoemission spectra with the calculation results of c-DMFT. Top: Bulk-sensitive photoemission spectra of VO_2 taken at 300K (insulating phase) and 348 K (metallic phase) [58]. Bottom: c-DMFT results of V $3d(t_{2g})$ spectral weights from Biermann *et al* [69].

the Fermi surface in VO_2 's metallic state; 2) the insulator peak's binding energy is lower than the 1.3 eV satellite peak in the metallic spectrum, as shown in Fig. 4.3; 3) at the critical temperature, the orbital occupancy changes so that the VO_2 's electronic structure is effectively switched from a three-dimensional structure in its metallic state to a one-dimensional structure in the insulating state. The quasiparticle peak in the metallic state in feature 1) reinforces the importance of strong electronic correlation in VO_2 's metal-to-insulator transition [52, 145]. Nevertheless, feature 2) indicates that VO_2 cannot be simply classified as a single-band Mott-insulator, because the satellite (the lower Hubbard band) in metallic state is usually at equal or slightly higher energies. Feature 3) provides us with a feasible physical picture for VO_2 's MIT, in which the system becomes more susceptible to the Peierls instability after the orbital distribution switches from three-dimensional to one-dimensional structure. Hence, the insulating state of VO_2 was classified as a 'renormalized Peierls insulator' and the bonding-antibonding splitting is renormalized due to the strong electronic correlations [69].

The contexts above introduced the theoretical and experimental efforts to resolve the mechanisms of the metal-to-insulator transition in VO_2 . It has been pre-assumed that MIT and SPT occur simultaneously in the theoretical investigations. Here, we want to pose the question in another way: Is the SPT from M_1 to rutile a pre-requisite for VO_2 's MIT?

4.2 Decoupling of Metal-to-Insulator Transition and Structural Phase Transition Observed in Steady-state Measurements

Single-crystal VO_2 nano-beam samples with very few defects are used in the experiments. The nanobeams are grown using vapor transport deposition [104, 138]. The samples are provided by Dr. Wu's group from the University of California, Berkeley. In the experiments, to minimize the strain effects, the nanobeams (~ 100 to 500 nm laterally) are freely suspended on a silicon window or on transmission electron microscope (TEM) grids of different metals. These sample geometries yield strongly first-order phase transitions as revealed by optical microscopy, where the color contrast tracks the optical reflectivity that varies drastically during the MIT [104, 138]. With the reflectivity changing, the color of the nanobeams turns to a darker color, which can be quantified by the intensity changes of the three primary colors (blue, green and red) of the VO_2 nanobeams in the images and extracted using the Image-J program. As shown in Fig. 4.4(b), the first-order MIT at $T_{\text{MIT}}=339$ K leads to a sharp transition in the color contrast as the sample is heated through the MIT critical temperature. Very differently from reported observations, where substrate strain exists on VO_2 nanobeams [72], no domains are observed during the MIT on these suspended nano-beams, which indicates the suspended nanobeams are in a strain-free condition. As shown in Fig. 4.4(d), the change of substrates has little influence on the nanobeams' MIT critical temperature, T_{MIT} .

Observation of the diffraction pattern of the microbeam under TEM (JEOL 2200FS) as the sample is heated yields the results of Fig. 4.4(c), showing that the dimer reflections (S1), such as $[101]$, $[100]$, and $[10\bar{1}]$, disappear beyond T_{SPT} (344 K on a Ni TEM grid), when monoclinic VO_2 turns into the more symmetric rutile structure through un-

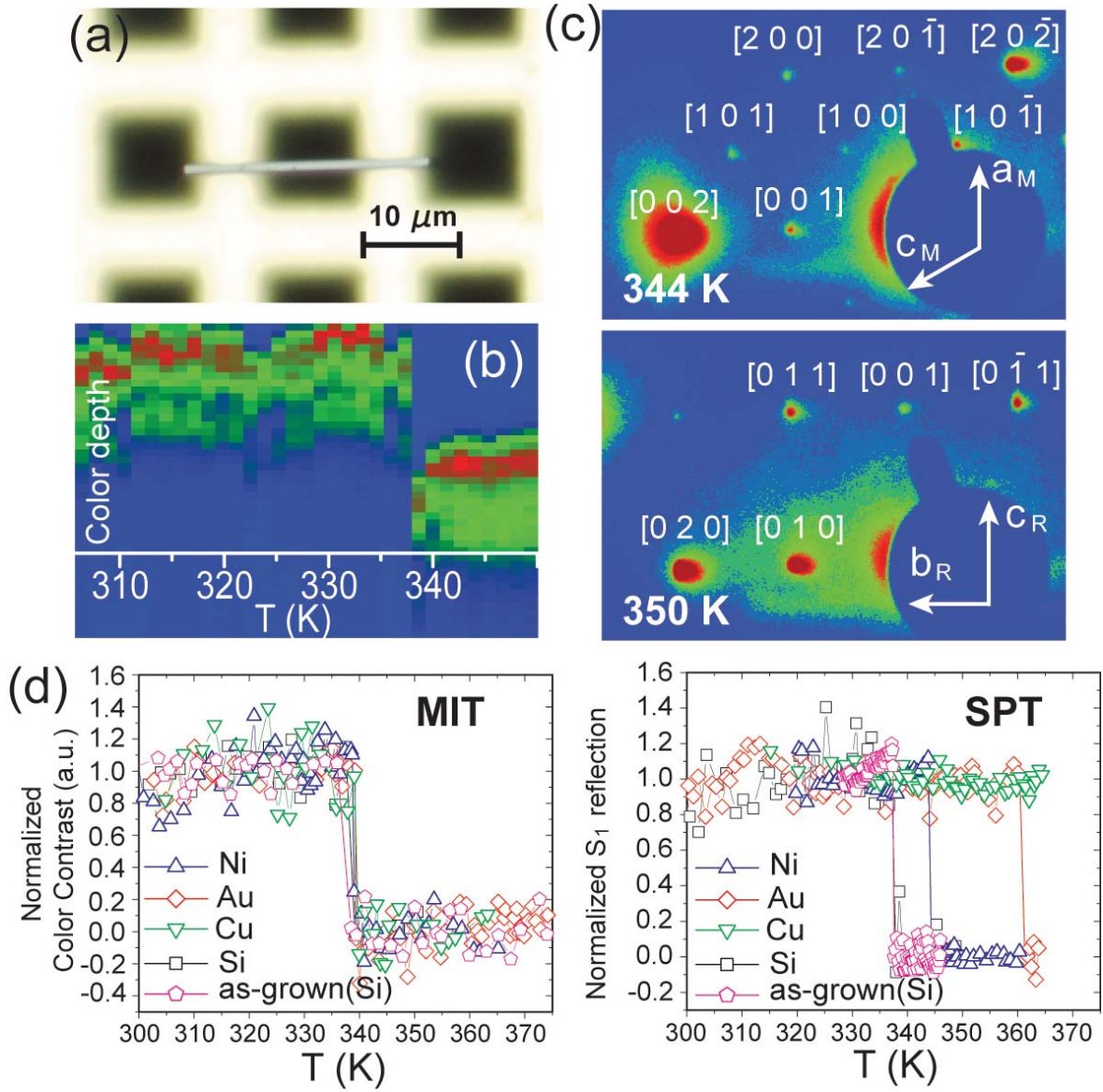


Figure 4.4: (a) Optical image of a VO₂ nano-beam suspended on a TEM window. (b) The normalized color-contrast change in the optical image of VO₂ associated with a first-order insulator-to-metal transition at 339 K. The color contrast is quantified by the intensities of the three primary colors (blue, green and red) of the VO₂ nanobeams in the optical images. (c) TEM diffraction patterns obtained from a VO₂ nano-beam suspended over a Ni TEM grid showing a monoclinic-to-rutile structural transition at 344 K. (d) Various metal-to-insulator transitions (MIT) and structural phase transitions (SPT) characterized over different metal TEM grids (Ni, Au, Cu), on a Si window, and in the as-grown condition where the nano-beam is anchored on a Si substrate.

pairing of V atoms along the c_R axis. Similar observations have also been made in the powder diffraction patterns of VO_2 polycrystalline thin-film membrane samples, which were discussed in detail in Sec. 3.3.1. The experimental geometry determines which sets of the dimer reflections appear in the diffraction pattern for the low-temperature monoclinic phase. Very differently from the MIT observed under the optical microscope, the T_{SPT} can be significantly affected by the type of substrates. Measurements of MIT and SPT for several systems are presented in Fig. 4.4(d), showing that T_{SPT} upshifts from T_{MIT} by different amounts on different metallic grids, whereas, on silicon, the SPT and MIT critical temperatures are approximately equal. The results in Fig. 4.4 were confirmed by repeating the measurements on 11 samples with different substrates. The extended measurement results are presented in Appendix C.1.

Here, we define noncooperativity η as the upshift of the critical temperature of the SPT (T_{SPT}) from the MIT (T_{MIT}). The observations in Fig. 4.4 allow us to define a new phase diagram of VO_2 , which is shown in Fig. 4.5(a). In the phase diagram, there exists a monoclinic metallic phase (M_3) when VO_2 single-crystal nanobeams are placed on metallic TEM grids. We need to note that the new M_3 phase is different from the previously reported M_2 phase, because the M_3 phase is a metallic phase, while the M_2 phase is insulating [103, 141]. We can also mostly exclude stress as the reason for stabilizing the M_3 phase, since very abrupt single-domain first-order electronic phase transitions have been observed on the suspended samples as illustrated by the optical contrast change in Fig. 4.6, which is very different from the MIT on strained nanobeams [72]. Here, we attribute the upshift of T_{SPT} on metal substrates to a charge doping effect mediated by charge injection through the metal contact into VO_2 . In a rigid band picture, initially these charges will go to the π^* band, making the interface metallic. To screen the injected charges a Debye sheath will be introduced to prevent further charge injection, so the charge doping effect will be limited in the interface region ($\sim 10\text{-}100$ nm) [146]. In the light of the current theoretical picture where both Mott and Peierls

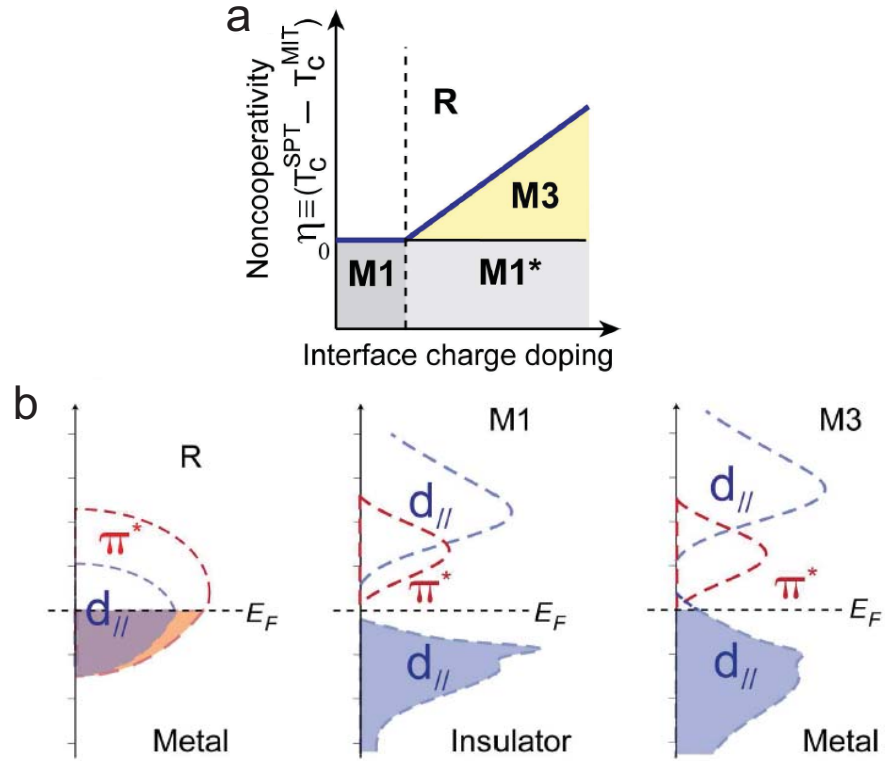


Figure 4.5: a) VO_2 phase diagram as a function of the interface charge doping. The vertical axis is the noncooperativity (η), which measures the upshift of the structural critical temperature (T_{SPT}) from the electronic one (T_{MIT}). b) Band structures of VO_2 in different phases. The left and middle panels are schematically drawn following Ref. [57, 106]. The right panel describes the proposed band structure for the monoclinic metal (M_3).

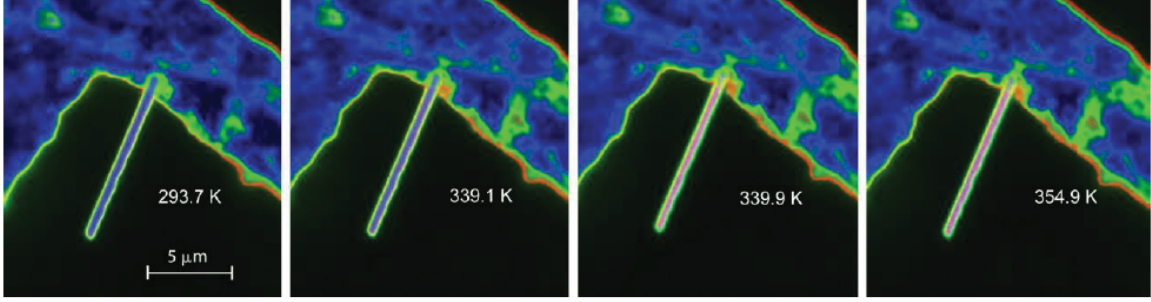


Figure 4.6: Optical images of VO_2 nano-beam supported on a Cu TEM grid. A shift in color depth as a result of increase in optical reflectivity occurs at ≈ 339.3 K, observed under a color filter centered at RGB values: (185, 159, 179) to enhance the contrast.

physics are involved [105, 106], we suggest that these injected charges into the π^* band renormalize the spectral function [58, 70] through local shifts in the Hartree energy (~ 0.5 eV) such that the d -component of the spectral character of the upper valence band increases to accommodate these extra electrons. In this process the bonding d -character is increased, which leads to an enhanced splitting between the bonding-antibonding Peierls gap, thus increases the Peierls-transition temperature (compared to the M_1 phase). The corresponding band structure is shown in Fig. 4.5(b). According to this argument, it is obvious that the low-temperature phase when VO_2 nanobeams are placed on metallic surfaces is already different from the normal M_1 phase due to the interface charge doping effect, even though no discernable structural differences are observed. We denote as the M_1^* phase as shown in the phase diagram Fig. 4.5(a). The observed decoupling of SPT and MIT of VO_2 indicates that the existence of the M_3 phase and the transition from M_1^* to M_3 is primarily driven by Mott physics.

4.3 Decoupling of Metal-to-Insulator Transition and Structural Phase Transition in VO₂ in Other Experiments

In fact, more and more experimental evidence points to the decoupling between VO₂'s MIT and SPT. First is that the laser-induced MIT occurs within <100 fs in VO₂ [74], while the observed monoclinic-to-rutile structural transition occurs in the timescale of picoseconds [109, 118], which indicates that VO₂ is transformed to the metallic state before the structural phase transition is complete. In another set of time-resolved experiments, C. Kübler et al. used a THz probe to investigate the ultrafast dynamics induced by the near-IR pump laser. Because the broad-band THz probe they used covers a wide range of frequency, which couples with both the optical conductivity and coherent phonon modes, they could track the lattice dynamics at the same time as the electronic dynamics. Remarkably, above the threshold pump fluence, a coherent phonon mode is excited and persists for several periods of oscillation, whereas the electronic conductivity settles at a maximum, suggesting a metallic phase, after only a single oscillation cycle coherent with the phonon oscillations. This result implies that the metallic state is established in VO₂, well before the lattice gets a chance to settle into a new structure [114].

In addition to the time-resolved pump-probe experiments, the decoupling of MIT from SPT in VO₂ was also observed in steady-state experiments. B. J. Kim et al., who induced VO₂ MIT through applying dc electric field, directly observed that the metallic phase formed in advance of the monoclinic-to-rutile structural phase transition by simultaneously measuring the electrical conductivity and the micro-x-ray diffraction patterns [147]. Very recently, M. K. Liu et al. also observed a metallic monoclinic feature on the strained VO₂ thin films. In this experiment, the authors used the

scanning near-field infrared nano-imaging spectroscopy to study the VO_2 film grown on a TiO_2 substrate, which is known to induce a large uniaxial strain on the VO_2 film. They found that the metallic phase appears on some domains, where the surface feature with the monoclinic structure still persists [148]. Including our experimental results, the decoupling of MIT and SPT in VO_2 is found to happen at 1) the ultrafast timescales and 2) the microscopic scales.

Whether VO_2 's MIT is decoupled from SPT is an important question in the following sense: 1) If the monoclinic-to-rutile SPT is not a pre-requisite for MIT, the Mott nature of the insulating gap is confirmed without doubt and the theories about VO_2 will have to be revised for the existence of the metallic monoclinic phase; the physics of phase transitions under certain conditions (pressure, charge doping, etc.) has to be taken into consideration. 2) It raises possibilities for high-speed electronic and photonic applications of VO_2 , including high-speed switches, transistors and optical modulators, because the speed-bottleneck and hysteresis imposed by repeating SPT will no longer limit the switching speed; and, at the same time, the problems of damage and fatigue of the devices due to repeated structural transitions can be overcome.

4.4 Ultrafast Electron Diffraction Studies of As-grown VO_2 Nanobeams and VO_2 Nanobeams on Gold Surfaces

Further information concerning the underlying physics of the MIT and SPT can be revealed by ultrafast pump-probe studies, where the transient nonadiabatic insulator-metal transition within the monoclinic structure can be stimulated. The application of femtosecond (fs) laser pulses allows rapid excitation of electron-hole pairs from the insulating ground state, effectively driving a transition to the metallic state while the

Peierls pairing is still intact.

4.4.1 As-grown VO₂ Nanobeams on SiO₂ Substrate

We used UED experiments to investigate optically driven dynamics of single-crystal VO₂ nanobeam in different fluences and temperature regimes. We started with studies of the as-grown VO₂ nanobeams, where we applied a 50 fs near-infrared (800 nm) pulse to illuminate the full nanobeam, followed by a sequence of probe electron pulses focused near the tip of the beam to track the lattice transformations that are expected to strongly couple to carrier excitations. The side view of as-grown VO₂ nanobeams on SiO₂ substrate is shown in Fig. 4.7(a). A full scale M₁-to-R transition can be stimulated by fs laser pulses, as shown in the insets of Fig. 4.7(b). The SPT is tracked from the disappearance of dimer reflections (S1): $[51\bar{2}]$, $[30\bar{1}]$, $[1\bar{1}0]$, which occurs when the applied fluence (F) exceeds a critical value (F_c). Fig. 4.7(c) shows a laser-induced first-order SPT occurring at $F_c = 0.59 \text{ mJ cm}^{-2}$ with VO₂ held at a base temperature $T_B=325 \text{ K}$. It should be noted that the photo-induced SPT is much sharper compared to the transitions on VO₂ thin-film membrane samples (See Fig.3.5). It is because each unstrained nanobeam contains a single domain, while the observations on the thin-film membranes suffered by the convolution from the properties of different nano-grains. We also examine F_c at different base temperatures, starting from around $T_c \sim 339 \text{ K}$ where SPT occurs spontaneously. We observe a linear increase of F_c as T_B decreases from T_c as shown in Fig. 4.7(b), suggesting that the laser-induced SPT can be described by crossing of the free energies from M₁ to R. To confirm this, we equate the absorbed optical energy density

$$E_{ph} = F_c/\delta, \quad (4.4.1)$$

where δ is the optical penetration depth, which can be calculated using the optical constants in Appendix B. To match Eqn. 4.4.1 with the thermodynamic energy, we

can get the following equation

$$E_{ph} = -C_v(T_B - T_c) + H_L, \quad (4.4.2)$$

where C_v and H_L are the specific and latent heats, respectively. From our fit, as shown by the solid line in Fig. 4.7(b), we find $C_v = 3.2 \pm 0.2 \text{ J K}^{-1} \text{ cm}^{-3}$, consistent with the previously established $C_v = 3.1 \text{ J K}^{-1} \text{ cm}^{-3}$ [133] for M_1 -VO₂ and $H_L \sim 0$. We find that below F_c the lattice fluctuations as characterized by the atomic displacement variance are reversible, whereas, above F_c , the R state did not revert to M_1 within the pump-probe cycle of 1 ms. This intriguing result may be explained based on a Mott-driven, Peierls-limited structural phase transition consistent with a recent near-infrared optical microscopy experiment, which indicates that the transition occurs via nucleation of the nanoscale correlated metallic puddles within the microcrystalline insulating VO₂ hosts [71]. Observed here is a temporal hysteresis (supercooling) known to be associated with the rutile metallic phase that at $F > F_c$ grows from nanoscale puddles to produce a SPT for the full nanobeam over many cycles, removing H_L from the threshold requirement. We also conducted UED on VO₂ beams suspended over a Silicon TEM window and found similar critical fluences, albeit the suspended VO₂ occasionally reorients itself after the SPT. Since enthalpy is dominated by the lattice component, these results support the hypothesis that the cooperative (SPT and MIT at the same temperature) phase transition has a strong Peierls character.

4.4.2 VO₂ Nanobeams Suspended on Au Substrate

The Au-supported nanobeams, which show a decoupling of the MIT and SPT in steady-state measurements, are also investigated using UED. The fs electron beam is directed along the monoclinic b -axis (b_M), as shown in Fig. 4.8(b). The optically stimulated response of the M_1^* phase is studied by setting $T_B = 337 \text{ K}$, which is just 2 K below T_{MIT} .

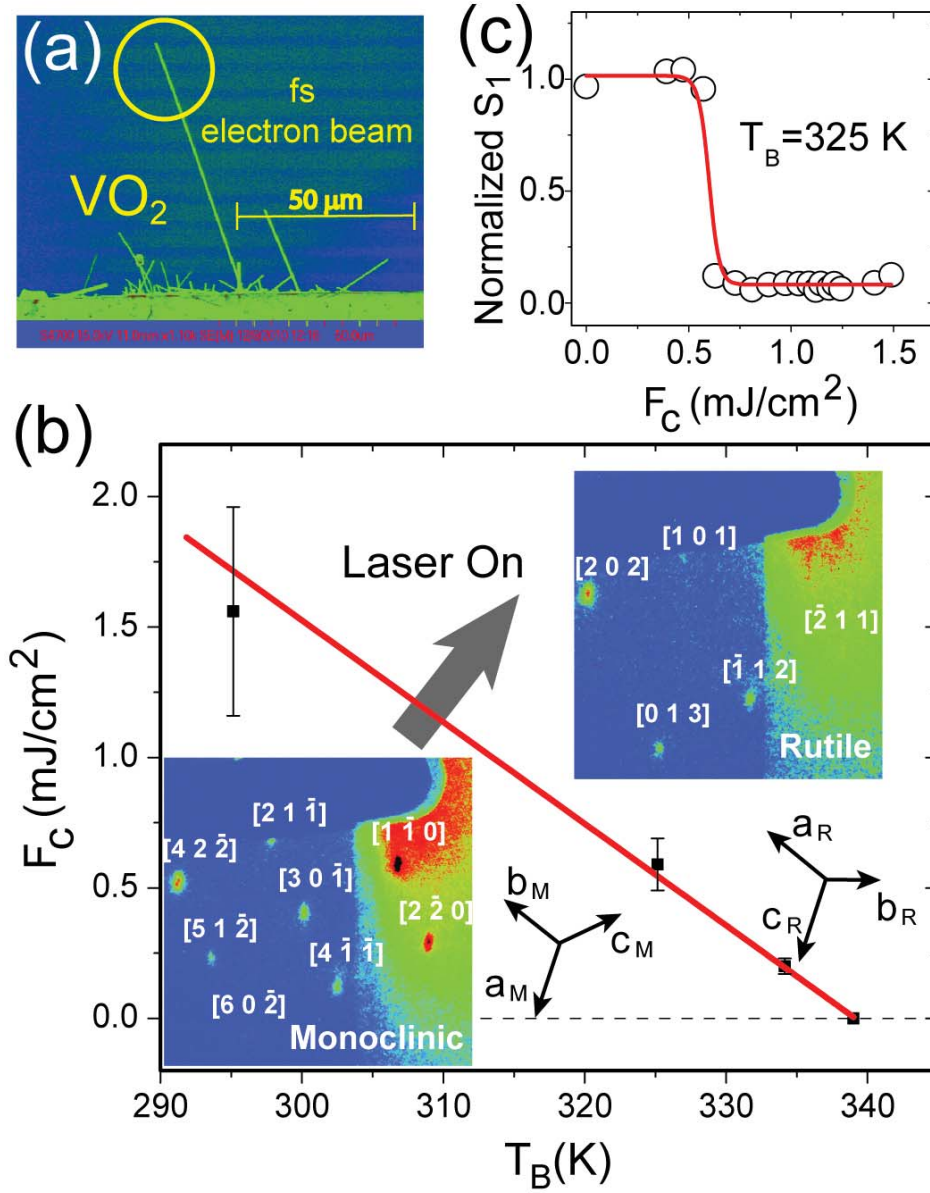


Figure 4.7: (a) SEM image of VO₂ nano-beams grown on a SiO₂ substrate. (b) The critical fluence (F_c) for inducing the monoclinic-to-rutile transition as a function of base temperature T_B . The insets show the corresponding UED patterns of the two phases. (c) The threshold behavior of laser-induced phase transition, as characterized by the disappearance of the dimer reflections (S_1), $[51\bar{2}]$, $[30\bar{1}]$ and $[1\bar{1}0]$ in (b).

In stark contrast to the observations made on the as-grown nanobeams, no SPT into the R phase is observed with the fluence up to 7 mJ/cm^2 , which has already exceeded the critical fluence of $\sim 3 \text{ mJ/cm}^2$ reported in other experiments at room temperature [111, 114]. Further increasing the laser fluence is limited by the damage-threshold of gold TEM grids. This suppression of the monoclinic-to-rutile transition is consistent with the upshift of the critical temperature of SPT, both of which indicate that the Peierls distortion is more stable when the nanobeams are placed on the metallic surface. From time-dependent Debye-Waller (DW) analysis, we deduce that the induced mean atomic displacement variance (Δu^2) is less than $0.0015 \text{ \AA}^2$ when $F = 7 \text{ mJ/cm}^2$, as shown in Fig. 4.8(c). The Debye-Waller analysis is commonly used in the community of crystallography to describe the attenuation of coherent X-ray, electron and neutron scattering caused by thermal motion of atoms. Here, taking advantage of the completeness of sample's momentum space obtained by single-crystal electron diffraction, we extend the normal Debye-Waller analysis to the anisotropic Debye-Waller analysis, which takes into account the anisotropic responses of lattice planes with different orientations. The details of the analysis are described in Appendix C.3.

In comparison, our investigation of the M_3 phase, conducted by setting $T_B = 341 \text{ K}$, shows a sizeable increase in the fluctuational response (also at $F = 7 \text{ mJ/cm}^2$). This increase is reminiscent of the anomalously large DW factors identified in the metallic rutile phase in the earlier X-ray experiments, suggesting a lattice softening in the metallic state [149]. Using the DW metric deduced from X-ray experiments the Δu^2 increase at M_3 corresponds to a temperature rise (ΔT) of no more than 20 K , which is significantly less than the theoretical predictions for M_1 and R phases, which are made using the optical and thermal constants of VO_2 , as shown by the dashed lines in Fig. 4.8(c). The strong reduction in the absorbed energy can be understood based on an upshift in optical band gap, which, consistent with the increase of Peierls transition temperature, leads to strong reduction in optically accessible spectral weight compared

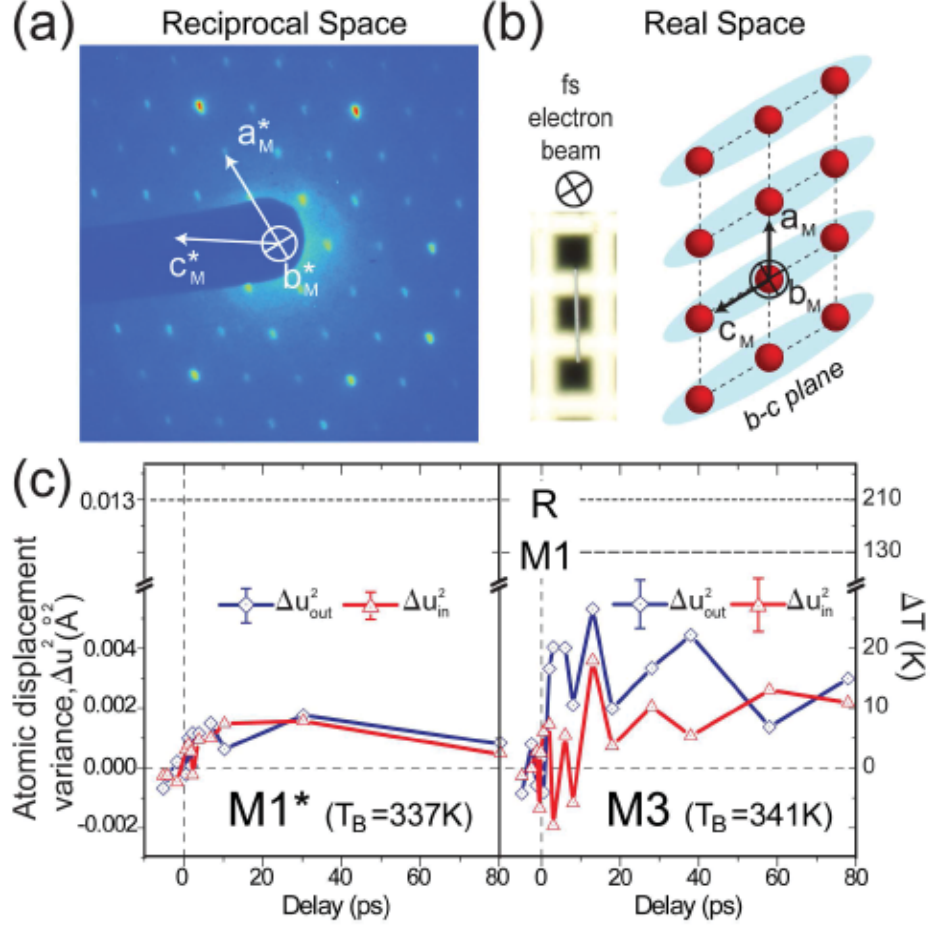


Figure 4.8: (a) UED pattern of a single suspended VO_2 nano-beam, showing the reciprocal lattice of the ac -plane of the monoclinic structure. (b) The corresponding real-space arrangement of the V sub-lattice. The fs electron beam is directed along b_M . (c) Lattice fluctuational dynamics in M_1^* and M_3 states are measured by the projected atomic displacement variance: Δu_{out}^2 (along a_M^* , out-of- bc -plane vibration), and Δu_{in}^2 (in- bc -plane vibration). The corresponding Δu^2 and temperature increases expected for the M_1 and R phases are plotted for comparison.

to that in the cooperative M_1 -R transition. The persistence of a large Peierls gap in M_3 is a clear evidence that insulator-metal transition involves only a small portion of the spectral weight and the low energy spectra near the Fermi surface (Fig. 4.5(b)). We also establish that the spectral function adjustment associated with the MIT may be coupled to a soft phonon mode that favors the lattice distortion along a_M shown by the suppression of Δu_{in}^2 (in- bc -plane vibration) relative to Δu_{out}^2 (out-of- bc -plane vibration) in Fig. 4.8(c).

4.5 Summary

We have shown clear evidence that the critical temperature associated with the Mott instability is lower than the structural phase transition temperature in VO_2 nano-beams. We also observe a new monoclinic metallic phase (M_3), which we show may be stabilized by interface charge doping effects [58, 147]. Informed by the ultrafast single-beam measurements, the metal-insulator transition in the undoped case can be classified as Mott-driven, Peierls-limited, where the imposed cooperativity is mainly governed by the large spectral weight and the high energy scale inherent to the Peierls transition and that such stabilizing features are subject to the strong Coulomb interaction predominant only in the insulating state without doping [106]. We also uncover the unique delocalized nature of the interface-mediated charge doping effect that can be rationalized with an active spectral weight transfer from the unfilled band to the Peierls band in the monoclinic structure, which enhances the Peierls interaction supported by its selective effect on the structure phase transition temperature. While the results presented here are very specific, they are generally compatible with the Peierls-Mott scenario proposed by most recent theories [69, 106], and amazingly reconcilable with an array of different experiments supportive of either Peierls or Mott scenarios. We expect the findings presented here will have important ramifications in bringing forth

a unified understanding of general MIT problems in strongly correlated materials.

Chapter 5

RF-compression Ultrafast Electron Microscope Development: Photoemission and Space Charge Effects.

In previous chapters, we presented a UED study on strongly correlated material VO_2 with thin-film membrane and single-crystal nanobeam samples. To avoid convolution from the property variance of multiple nano-grains in thin-film membrane, single-crystal nanobeam samples were studied, which provided us with the single-crystal diffraction patterns mapping a large momentum space, as shown in Fig. 4.8(a). However, the signal-to-noise ratio (SNR) of the data is not high enough so that the detailed structural dynamics cannot be extracted. The dynamics of Bragg peaks' intensity shown in Fig. 4.8 is obtained by averaging more than 30 peaks to improve the SNR. (For details, see Appendix C.3.) The main reason for this problem is 1) the number of electron is limited to 800 - 1000 electrons per pulse; 2) the electron beam's footprint on the sample is $\sim 40 \mu\text{m}$ in FWHM. So, on average, there is only ~ 1 electron per

μm^2 per pulse available for diffraction. With such a low electron flux, the SNR is not sufficient to study samples at the micrometer or nanometer scales, where the strongly correlated materials are found to exhibit the most exotic phenomena. Driven by the need to improve the brightness of the UED instruments, an RF-compression ultrafast electron microscope (UEM) is being designed and developed at Michigan State University. The mechanism of RF compression is to compress the electron pulse, which is already broadened by space charge effects, by reversing its longitudinal momentum-space correlation in the phase space using an oscillating electric field. After the RF cavity, a strong magnetic lens is implemented to focus the longitudinally compressed electron pulse in its transverse direction to achieve high beam brightness. For clarity, we define the ‘*longitudinal*’ direction as the electron beam propagation direction, while the ‘*transverse*’ directions as the directions perpendicular to the propagation direction.

In the physics of charged particle beams, the brightness (or the directional beam intensity) of an electron beam at a given point in a given direction is defined as the current per unit area normal to the given direction, per unit solid angle, which is given by the reduced brightness B_r [150, 151],

$$B_r \equiv \frac{1}{E_e} \frac{\partial^2 I}{\partial A \partial \Omega} = \frac{I m c^2}{2 \pi^2 \varepsilon_x \varepsilon_y}, \quad (5.0.1)$$

where E_e is the beam energy, I the beam current, A the beam cross-section area, Ω the beam solid angle, ε_x and ε_y are the beam transverse *normalized emittances*. Normalized emittance is an important concept in particle beam physics, because it quantifies the beam coherence. The normalized emittance in the transverse direction ε_i is defined as [151]

$$\varepsilon_i = \frac{1}{m c} \sqrt{\langle x_i^2 \rangle \langle p_i^2 \rangle - \langle x_i p_i \rangle^2}, \quad (5.0.2)$$

where $i = \{x, y\}$ denotes the transverse directions, x_i and p_i are the position and momentum of electrons, the brackets $\langle \dots \rangle$ stands for an average over the ensemble of

electrons in the bunch. Eqn. 5.0.2 can be reduced to $\varepsilon_i = 1/(mc)\sigma_i\sigma_{p_i}$ at the beam waist, where σ_i is the rms bunch radius and σ_{p_i} is the rms of the momentum distribution. The normalized emittance can also be understood as the area the electron pulse occupies in its phase space. According to Liouville's theorem, this area should be conserved throughout the beam propagation. In other words, the ultimate beam brightness on the sample is limited by the initial electron-beam emittance when it is generated on the photocathode in an ideal situation. So, it is a necessity for us to understand the photoemission process, especially the situation of short-pulse generation.

Research on photoemission can be traced back to 1887, when Hertz originally observed the photoelectric effect [152]. Later, this effect was explained by Einstein as the manifestation of the quantum nature of light [153]. When light is incident on a material surface, an electron absorbs a photon and escapes out of the surface with its maximum energy $\hbar\omega - \phi$, where $\hbar\omega$ is the incident photon energy and ϕ the work function of the material. Within the single-electron limit, the photoemission process is understood so well that photoemission spectroscopy is recognized as the best technique to study the intrinsic electronic structure of materials with very high energy resolution (< 2 meV) in the last twenty years [154]. With the emergence of high-power fs laser systems and the demand for high-brightness electron beams as probes [32, 118, 155] or synchrotron sources [156, 157], research on photoemission is extended to the region of ultrashort, ultrahigh brightness electron pulses. In this region, the strong Coulomb repulsion between electrons, which is known as the ‘*space charge effect*’, has strong effects on the pulse duration and the energy distribution.

In this chapter, I will introduce our recently developed technique, the ‘*projection-shadow-imaging (PSImg) technique*’ [158, 159], and use it to measure the electron-pulse duration and study the space charge effects in the near-cathode-surface region. We will start with a brief introduction to the methods to characterize the electron-pulse duration in Section 5.1. The experimental setup for the PSImg technique and the

analytical description of the shadow images are described in Section 5.2. In Section 5.3 and 5.4, we present the experimental results on the gold photocathode investigated with the PSImg method, which highlights the strong effects of electron-electron repulsion at the beginning of the photoemission process. Based on the results in these two sections, a useful mean-field model will be introduced in Section 5.5, which helps to simulate and define the geometry of the RF-compression UEM system. The limitations on its combined spatial and temporal resolution in different operational regimes will be discussed as well. This chapter will conclude with an introduction to the structure of the UEM system in Section 5.6.

5.1 Different Methods of Characterizing the Electron-pulse Duration

Characterizing the temporal resolution of electron pulses is very important for experiments like UED, free-electron lasers, and for accelerator physics in general. In time-resolved optical experiments, the laser pulse duration can be characterized by the optical cross-correlation or auto-correlation methods, in which the electric field of one pulse becomes convolved with another pulse through the instantaneous polarization of a nonlinear medium. Similar ideas can be applied in a UED experiment, where the electron pulses are used as probes. When characterizing the UED temporal resolution using the cross-correlation method, a near-infrared fs laser pulse are used to induce the structural change in the sample, and the late-coming electron pulse, whose relative delay is adjusted by an optical delay stage, probes the induced structural changes at a specific time delay through the electron diffractions. If the structural change of the sample is fast enough, the observed structural dynamics will be limited by the electron-pulse duration, thus, its value can be deconvoluted from the experimental results. In Fig. 5.1, we plot the characterization of the temporal resolution of our first-generation

UED system [160]. The charge-density-wave (CDW) dynamics of 1T-TaS₂ after 800 nm, 50 fs pump-laser excitation is used as the reference. It was found that the response time of CDW in 1T-TaS₂ can be as fast as 220 fs [32]. In the experiment, by attenuating the pulse energy of the 266 nm fs laser, which is focused on the photocathode, hence, reducing the number of electrons in individual pulses, the observed intensity decay of the CDW satellite peaks appears faster, due to reduced electron-pulse duration with less significant space charge effect. The pulse duration as a function of the 266 nm laser pulse energy is plotted in the inset of Fig. 5.1. To the point where the response time reaches the minimum, the response time is ultimately limited by the speed of the structural dynamics of the material. The cross-correlation method is very useful for *in-situ* characterization of the UED temporal resolution [6, 161], but it is limited by the speed of materials' structural response to the fs laser pulse, which is usually hundreds of fs at least.

Another conventional way of measuring the electron-pulse duration is to use a streak camera [14, 162, 163]. In a streak camera, the electron pulse is sent through a pair of deflection plates with a ramped voltage applied. When the electron pulse passes through the plates, the leading electrons and late-arriving electrons are deflected in the opposite direction transversely. Thus, the electron distribution along the longitudinal direction can be mapped transversely, which gives a measurement of the electron-pulse duration. The sweeping voltage applied to the deflection plates is synchronized through a pair of GaAs photoconductive switches, which are triggered by the laser pulses split from the same laser source used to generate photo-electrons on the photocathode. The temporal resolution of the streak camera is limited by the sweeping speed of the applied high voltage and the sensitivity of the imaging system. In a single shot measurement (>1000 electrons per pulse), the temporal resolution of a streak camera is ~ 1.2 ps [161, 164], even if the highest sweeping speed available in current technology is applied. In a multi-shot, low-flux measurement, the amplitude jitter in the switching

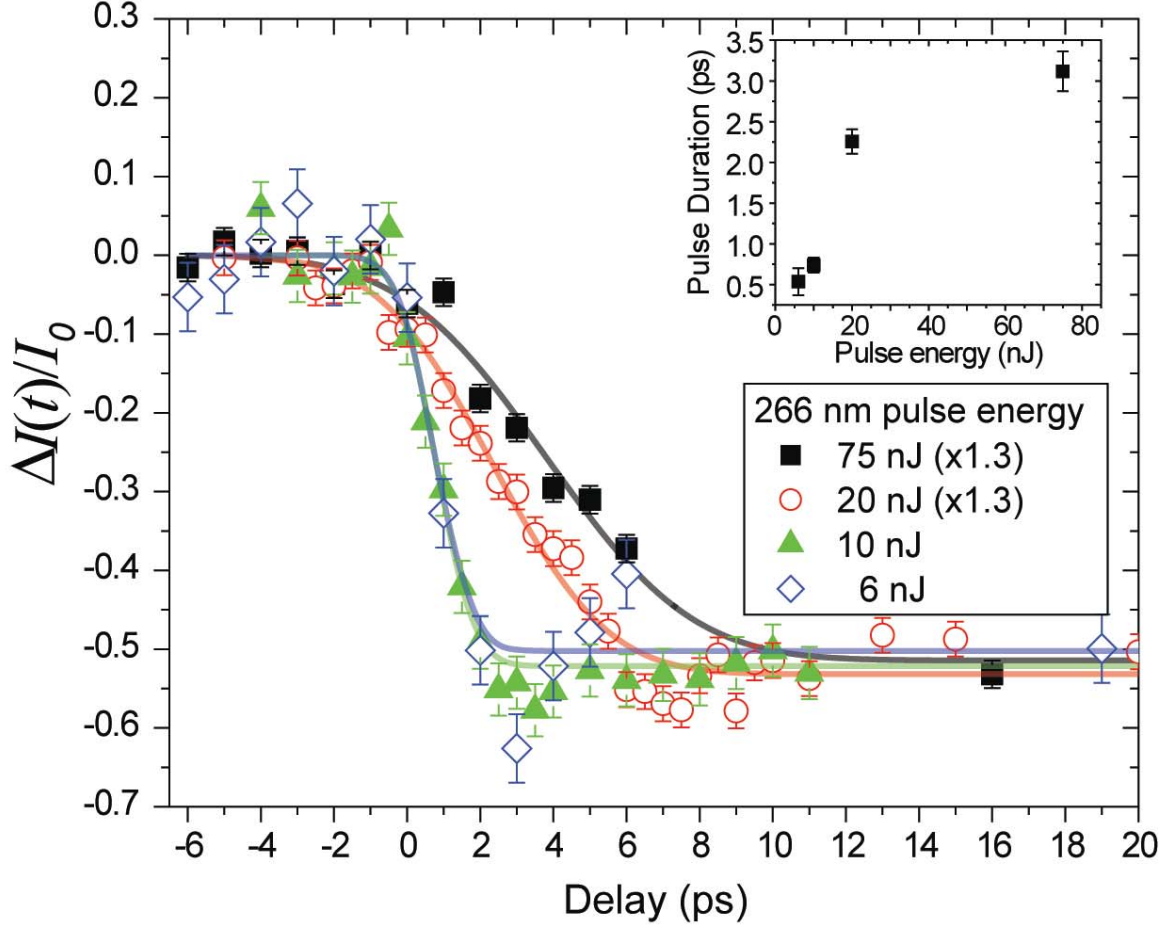


Figure 5.1: Dynamics of charge-density-wave (CDW) satellite reflections as the function of delay time, plotted for different 266 nm pulse energies focused on the photocathode for photoelectron generation. The solid lines are the error-function fitting to extract the pulse duration for each laser pulse energy. Inset: UED electron pulse duration as a function of the 266 nm laser-pulse energy.

process transfers into the timing jitter, which affects the temporal resolution of the streak camera. Temporal resolution less than 1 ps has been reported [162, 165] with deconvolution for low flux electron pulses, but the details of deconvolution are not straightforward, as it requires enough electrons above the background in the pixellated image and the knowledge of the pulse shape. Recently, the RF-cavity based streak camera has been developed and employed to measure the electron pulse durations less than 100 fs [36, 166]. The advantage of using an RF-cavity is its available electric field strength and sweeping speed can be orders of magnitude higher compared to the conventional dc-voltage deflection plates.

Recently, Siwick, et al. proposed a new method of characterizing the electron-pulse duration in free space by using the electron-laser pulse direct cross-correlation [167]. In their work, they made use of the laser pulse ponderomotive potential, U_p , to scatter electron pulse and measure the transmitted electron-pulse intensity as a function of the relative delay between the electron and laser pulses. The ponderomotive potential is given by

$$U_p(\vec{r}, t) = \frac{e^2 \lambda^2}{8\pi^2 m \epsilon_0 c^3} I(\vec{r}, t), \quad (5.1.1)$$

where λ and I are the laser pulse wavelength and intensity, respectively. Fig. 5.2 shows the geometry of the cross-correlation measurement and the electron beam cross-section images before and after laser pulse scattering. The temporal resolution of this technique can be less than 100 fs in an optimum condition [167]. In a recent characterization of their pulse-compressed UED system, the authors got the temporal resolution ~ 300 fs for the electron pulse [43].

The methods introduced above are very useful in characterizing the electron-pulse duration in free space. However, to study space charge effects in the near-cathode-surface region ($\sim 100 \mu\text{m}$ above the cathode surface), these methods are no longer appropriate. Using the general setup of a UED experiment, we developed a method, which can measure the electron-pulse duration and its dynamics in the region very

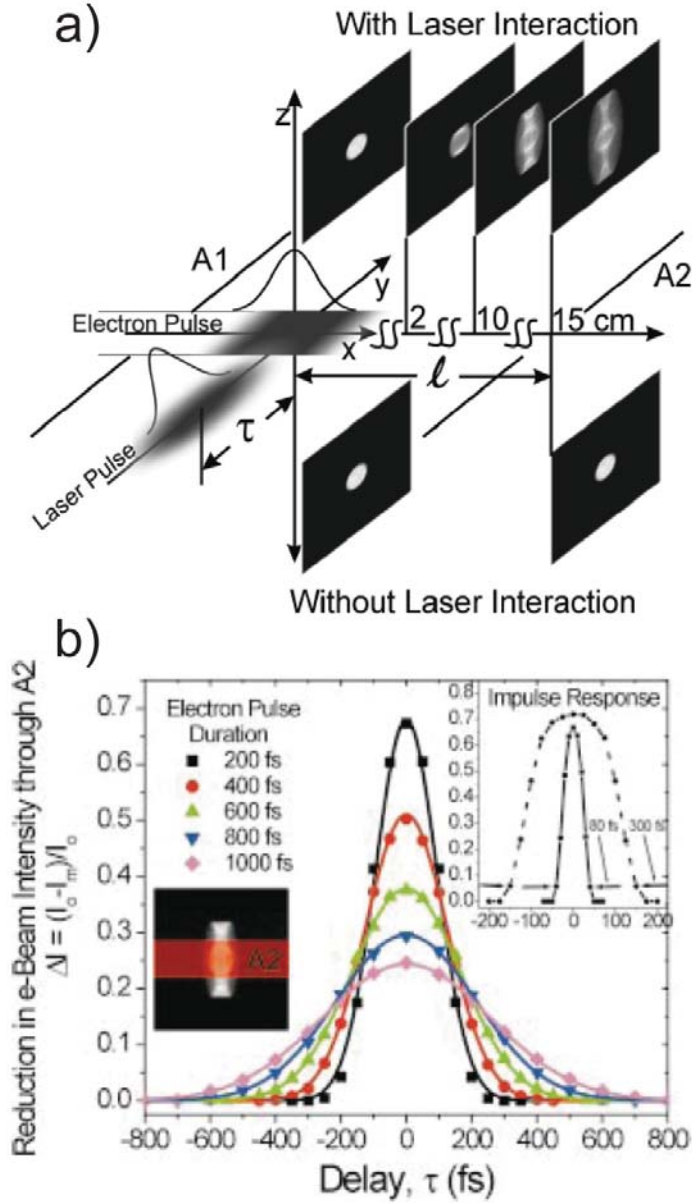


Figure 5.2: a) Geometry of the all-optical electron-pulse duration measurement. Electron-beam cross sections are shown at various positions after the interaction region for a 600-fs electron pulse; b) Electron-laser pulse cross-correlation. The reduction in electron-beam intensity transmitted through aperture A2 is sensitive to τ and the duration of the electron pulse at the interaction point (see the legend). The impulse response of the measurement (inset) is ~ 300 fs e^{-2} full width (FW) for the parameters $A1=30 \mu\text{m}$, $t_l=200$ fs (dashed curve), but only, 80 fs e^{-2} FW for $A1=10 \mu\text{m}$, $t_l=60$ fs (solid curve). Figure from Ref. [167].

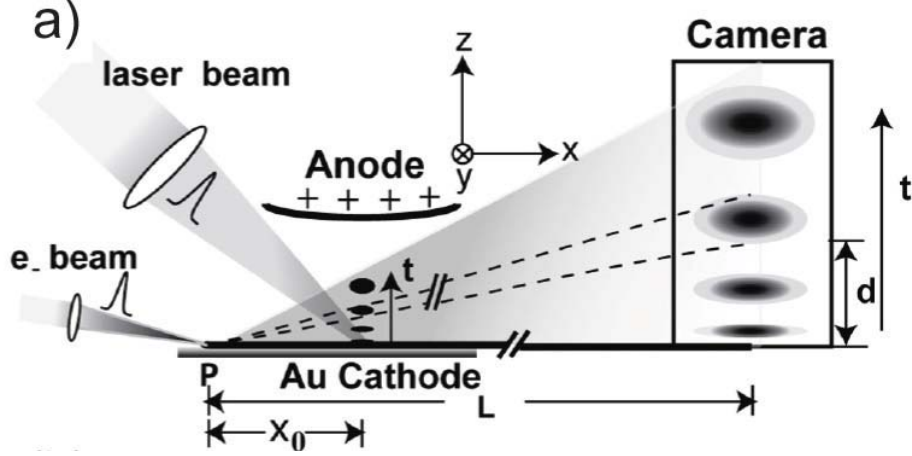


Figure 5.3: Schematic of the projection-shadow-imaging experiment. For illustration purposes, the angular span of the shadow is significantly increased from typical values ≤ 1 mrad.

close to the cathode surface, thus providing us with information relating to the space charge effects at the very beginning of the photoemission process.

5.2 Experimental Setup and Analytical Description of Projection-Shadow-Imaging Technique

5.2.1 Experimental Setup

The general experimental setup for the PSImg experiment is shown in Fig. 5.3 [158, 159]. The experiment is conducted in a pump-probe arrangement, in which the fs laser (266 nm or 800 nm, 50 fs) serves as the pump to generate the short-pulse photoemission from the material surface. The photoemitted electron pulse is subsequently probed by the delayed surface-scattered electrons. The incident pump beam is displaced from the source of scattered electron by a distance of x_0 (typically a few millimeters) to establish a projection imaging geometry. The presence of the photoelectron pulse diminishes the diffusive surface-scattered electrons generated at P , thus casts a shadow on the screen, which is at a distance $L = 165$ mm away from P . The magnification of this projection

imaging is $M \approx L/x$. By adjusting the arriving time of the probing electrons relative to that of the pump laser, the evolution of the shadow images can be recorded. Fig. 5.4(a) shows the images of photoelectron pulses emitted from a highly ordered pyrolytic graphite (HOPG) surface as a function of time [158]. As shown in Fig. 5.4(b), the cross-sectional line-scans obtained from these shadow images reveal an accumulation peak near the shadow edge, caused by surface scattering, and this peak decays sharply into the vacuum with a Gaussian depletion profile, which evolves in space and time. To extract the spatiotemporal dynamics of the electron pulses, the line-scans of the shadow images are fitted with the following analytical form, which takes into account the geometry of the projection imaging.

$$F(d) = \frac{B\Sigma}{\sigma_z} \frac{\exp\left[-\frac{(dx_0-Lz_0)^2}{2(d^2\sigma_x^2+L^2\sigma_z^2)}\right]}{\sqrt{\frac{1}{\sigma_x^2} + \frac{d^2}{L^2\sigma_z^2}}}, \quad (5.2.1)$$

where d is the position on the camera screen, Σ is the charge-sheet density of the electron pulse, x_0 is the source-to-beam distance, z_0 is the electron pulse's center-of-mass (COM) position, σ_x and σ_z are the widths in transverse and longitudinal directions, and L is the camera distance. The PSImg experimental geometry is plotted in Fig. 5.5, in which the dimensional parameters in Eqn. 5.2.1 are defined. The values of the key parameters, such as Σ , z_0 , σ_z and σ_y can be determined by fitting the profiles of the longitudinal and transverse shadow images.

5.2.2 Description of the Shadow Images

Assuming the intensity of the shadow image at the location d on camera is proportional to the scattering events along the probing electron path, as shown in Fig. 5.5, we can derive Eqn. 5.2.1 purely based on the ‘scattering model’. Here, the charge distribution of the photoelectron pulse is taken to be a three-dimensional Gaussian function, which

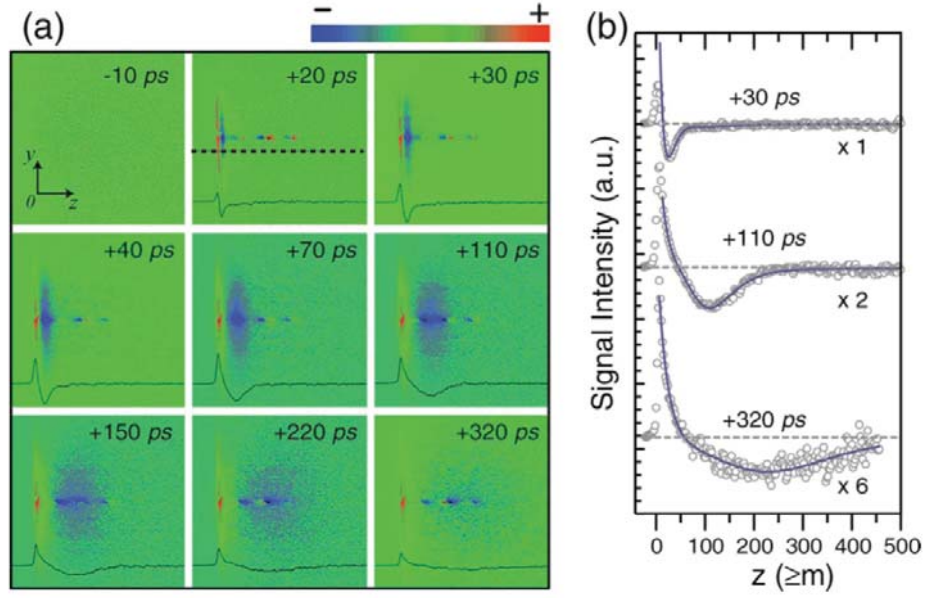


Figure 5.4: a) Snapshot shadow images of photoemitted electron bunch obtained with $x_0=4.3$ mm, which gives magnification $M \sim 38$. The cross-sectional line profiles (black curves) are extracted along the dashed line in the second panel. b) Fitting curves (lines) of cross-sectional line-profile data (circles) using Eqn. 5.2.1. Figure from Ref. [158].

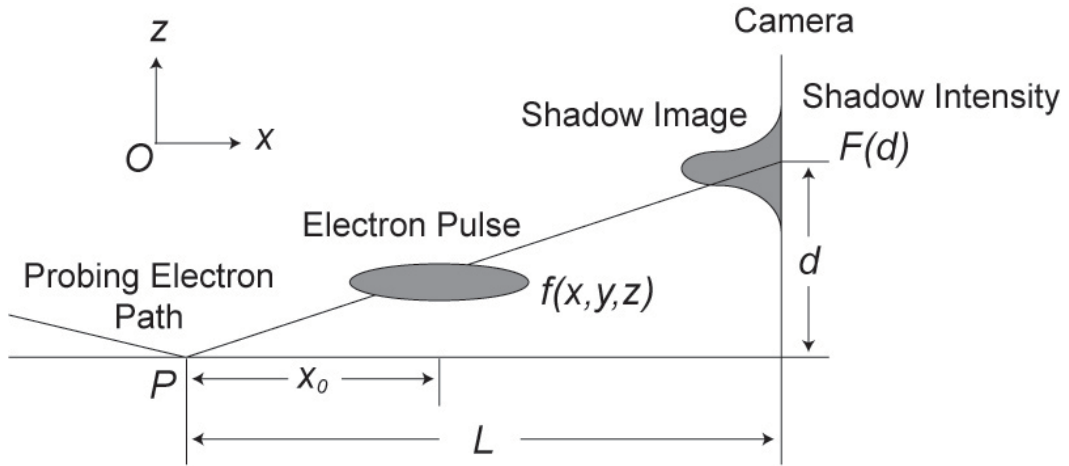


Figure 5.5: Geometry of projection-shadow-imaging technique, where the photoelectron pulse is fitted to a three-dimensional gaussian function, $f(x, y, z)$. The profile of the shadow image on the camera is described as an analytic function $F(d)$, where d is the position on the camera screen.

is justified by the Gaussian distribution of the pump laser pulse:

$$f(x, y, z) = \frac{N_e}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} \exp\left[-\frac{x^2}{2\sigma_x^2}\right] \exp\left[-\frac{y^2}{2\sigma_y^2}\right] \exp\left[-\frac{(z-z_0)^2}{2\sigma_z^2}\right], \quad (5.2.2)$$

where the center of the electron pulse is located at $(0, 0, z_0)$ above the cathode surface. In the linescan analysis, the first step is to select a region-of-interest (ROI) at the center of electron pulse along the y direction. The 1D-linescan profile is thus the averaged intensity in the ROI, which is given by

$$\begin{aligned} g(x, z) &= \frac{\int_{-l}^l f(x, y, z) dy}{2l} \\ &= \frac{N_e}{(2\pi)\sigma_x\sigma_z} \exp\left[-\frac{x^2}{2\sigma_x^2}\right] \exp\left[-\frac{(z-z_0)^2}{2\sigma_z^2}\right] \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-l}^l \exp\left[-\frac{y^2}{2\sigma_y^2}\right] dy \\ &= \frac{N_e}{\pi\sqrt{2\pi}\sigma_x\sigma_y\sigma_z} \xi \exp\left[-\frac{x^2}{2\sigma_x^2}\right] \exp\left[-\frac{(z-z_0)^2}{2\sigma_z^2}\right], \end{aligned} \quad (5.2.3)$$

where l is half width of the ROI along the y direction and ξ is defined as

$$\xi = \frac{1}{4l} \int_{-l}^l \exp\left[-\frac{y^2}{2\sigma_y^2}\right] dy. \quad (5.2.4)$$

When $l \ll \sigma_y$, $\xi \approx 1/2$. In the ‘scattering model’, the intensity on the camera position d is attenuated by the scattering events of probing electrons with the photoelectron pulse of interest along the beam path: $z = d/L$, as shown in Fig. 5.5, where P is defined as the origin of coordinates. Here, we can assume that the scattering probability is proportional to the integrated number of photoelectrons along the path, which is given by

$$\begin{aligned} F(d) &= \lambda \int_{-\infty}^{\infty} g(x, \frac{d}{L}x) dx \\ &= \lambda \frac{\xi \Sigma}{\sigma_z} \frac{\exp\left[-\frac{(dx_0 - Lz_0)^2}{2(d^2\sigma_x^2 + L^2\sigma_z^2)}\right]}{\sqrt{\frac{1}{\sigma_x^2} + \frac{d^2}{L^2\sigma_z^2}}}, \end{aligned} \quad (5.2.5)$$

where λ represents the electron-electron scattering cross-section, $\Sigma = \frac{N_e}{\pi\sigma_x\sigma_y}$ is the electron pulse transverse sheet density. $\lambda\xi\Sigma$, as a whole, serves as the fitting parameter $B\Sigma$ in Eqn. 5.2.1. Eqn. 5.2.5 requires further normalization to the total number of electrons, because its integration over space yields a varying value with some of the experimental parameters, including x_0 , L and σ_x . To normalize $F(d)$ to a conserved total number of electrons, the pre-factor $x_0/(L\sigma_x)$ is applied. Details for deriving the pre-factor can be found in Appendix D.1. The normalized analytical function is thus given by

$$M(d) = \frac{x_0}{L\sigma_x} F(d). \quad (5.2.6)$$

The fundamental assumption of the analytical description in Eqn. 5.2.5 and 5.2.6 is that the intensity of shadow images is proportional to the number of scattering events, which ignores the Coulomb force between the photoelectron pulse and the probing electrons. Physically, this analytical description cannot be correct. However, in practice, we will find that the transverse and longitudinal sizes (σ_x , σ_z) of the electron pulse can be reliably extracted from experimental results using this analytical description, when the charge volume density is not very high.

Except for those dimensional parameters, the electron density cannot be directly extracted from the shadow images using the analytical description. To characterize the PSImg method, a series of experiments have been carried out on HOPG, which are described in Appendix D.2. Combining with the ‘ray-tracing simulation’ and ‘the Bragg-beam deflection’ experiments, we can extract the charge-sheet density from the shadow images in the experiments as well. (See Appendix D.2 for details.) In practice, we take the strategy to first extract the pulse dimensions using the analytical fitting which is less computational expensive. Then, the electron density can be calibrated and extracted using the combined methods described in Appendix D.2. With the complete understanding of the PSImg technique, we will study the photoelectron pulses from a gold photocathode, which is of practical interest for our UEM designs.

5.3 Measurement of Spatial and Temporal Evolution of Photo-electron Pulses from Gold Photocathode

Here, the PSImg technique is used to measure the spatial and temporal evolution of photoelectron pulses from a gold cathode in the near-cathode region. The advantages of using gold as the photocathode material include its high quantum efficiency, easy preparation and inertness to environment. The gold photocathode is prepared via vapor deposition on a quartz substrate with a homogeneous film thickness of 30 nm. We investigate photoelectron pulse dynamics in a dc photo-gun arrangement, where a positive electrode (anode) is separated by 5 mm from the grounded cathode surface, providing an applied field F_a to facilitate photoemission. The schematic of the dc photo-gun geometry is shown in Fig. 5.3. Instead of using the 800 nm fs laser as the pump, the 266nm ultraviolet (UV) fs laser pulses (50 fs, $\hbar\omega = 4.66$ eV) are used as pump to trigger photoemission from the gold cathode. The fs UV laser pulses arrive at the cathode surface at 45° incidence and define the zero-of-time (ZOT) of photoemission. The photon energy is slightly higher than the reported work function (Φ_w) of a polycrystalline gold film that ranges from 4.0 - 4.6 eV [21, 168–170], allowing photoemission with a small energy spread. High-intensity pulses (10^5 - 10^7 electrons per pulse) are generated, which is relevant to the high-brightness implementations of ultrafast electron diffraction [14, 163, 171] and imaging [19, 27]. As has been described in previous sections, the dynamics of the electron pulses can be extracted from the dynamics of the shadow images cast on the metalized phosphor screen, which is connected to an intensified CCD camera. The point source for PSImg is generated by focusing an independent low-density, fs, 30 keV photoelectron beam, containing ~ 800 electrons per pulse with a beam waist of ~ 40 μm [6]. The projection imaging is performed at

a repetition rate of 1 kHz, and the snap-shot images at each delay Δt are averaged over $10^5 - 10^6$ repetitions of pump-probe cycles to obtain sufficient SNR and to average out the pulse-to-pulse fluctuations of pump and probe pulses. The shadow patterns, obtained by taking the difference between the positive frames ($\Delta t > 0$) and the negative frame ($\Delta t < 0$), are normalized to the negative frame image to cancel out the angular dependence of the illumination from the point source.

Fig. 5.6 shows the shadow profiles of photoelectron pulses generated from the gold surface. Quantitative results are obtained through fitting the linescan profiles with the analytical model describing the projection geometry in Eqn. 5.2.5 and Eqn. 5.2.6. The key pulse parameters, such as Σ , z_0 , σ_z , and σ_y can be determined by fitting its longitudinal and transverse shadow profiles, which evolve as a function of time. In this specific experiment, $x_0 = 5.0$ mm and $L = 16.5$ cm, giving a magnification of $M \approx 33$ for the images. In the front illumination geometry, the excitation laser has an elliptical footprint with $\sigma_x = 115$ μm and $\sigma_y = 81$ μm , determined *in situ* via examining surface voltammetry characterization [158, 172]. Characterization of the laser footprint on the photocathode yields an initial transverse profile of the photo-emitted electrons, which is consistent with the transverse electron bunch characterization from shadow imaging. The laser footprint characterization also yields quantitative measurements of the laser fluence (F) and the number of emitted electrons (N_e), combined with Σ obtained from imaging. Imaging experiments are performed using various excitation fluences $F = 1 - 10$ mJ/cm² and extraction field $F_a = 0 - 0.4$ MV/m, which together produce a wide range of values of N_e up to 10^7 electrons per pulse.

To quantify the rate of Coulomb explosion, the time-dependent electron-pulse length $\sigma_z(t)$ in the COM frame is extracted from the shadow profile as a function of electron density. Here, the longitudinal profile is observed to expand rapidly while the transverse profile increases little ($\leq 15\%$) over our tracked time window (0 - 120 ps). We extract the dependence of the electron-pulse longitudinal width $\sigma_z(t)$ on Σ from the

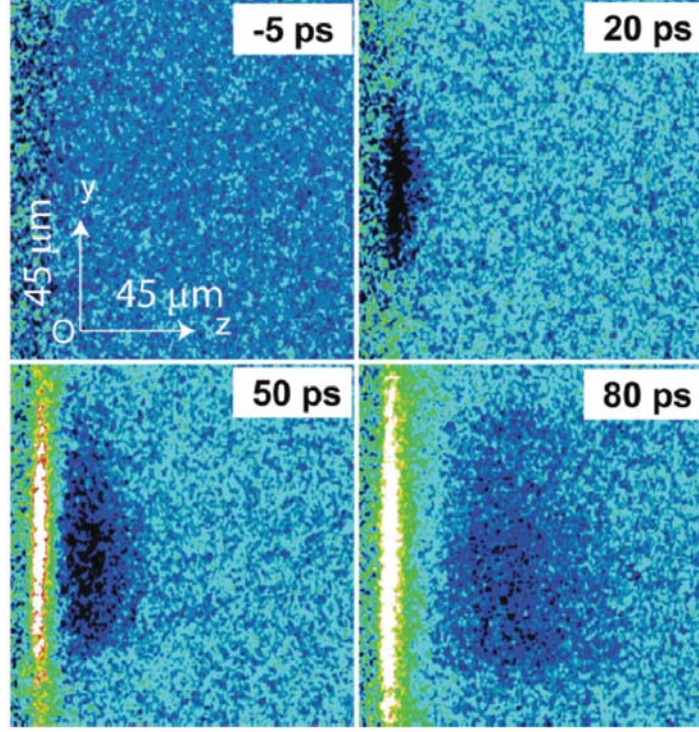


Figure 5.6: The snap-shots of the normalized shadow images at selected times. The magnification of the projection imaging is ≈ 33 .

shadow imaging data. Fig. 5.7(a) shows a super-linear increase of $\sigma_z(t)$ resulting from the strong space charge effects at high electron sheet density Σ . On the other hand, with low electron sheet density, $\sigma_z(t)$ changes linearly with a slope approaching $0.1 \mu\text{m}/\text{ps}$, which translates to a very small longitudinal photoelectron energy spread of $\sim 0.03 \text{ eV}$, supporting the over-the-barrier direct photo-ionization picture. As will be discussed later, in the weak emission limit, the initial momentum spread of the photoelectron pulse compares well with the theory of the three-step model, which is designed to model the photoemission at the single-electron (space-charge-free) limit [173, 174].

5.4 Space Charge Effects and Modeling

To elucidate the space charge effects in the observed photoelectron dynamics, we employ N-particle dynamical simulations to model the experimental results. In particular,

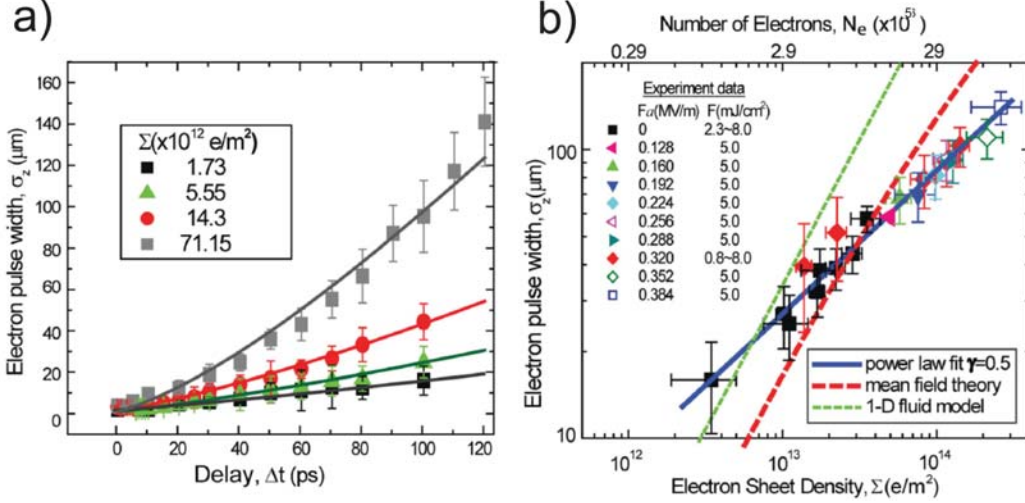


Figure 5.7: The results of electron point-projection imaging. (a) The COM-frame expansion dynamics of electron pulses with longitudinal pulse length σ_z . The solid lines through the experimental data (symbols) are from N-particle simulations. (b) Dependence of pulse length on the electron sheet density at time 100 ps for various F and F_a settings, showing a universal power-law increase with exponent roughly $\gamma = 0.5$. Predictions based on mean-field [14] and one-dimensional fluid [175] models are also presented.

we investigate how pulse expanding depends on electron density and the photoemission initial conditions.

5.4.1 Fractional Power-law Dependence in the Space-charge-led Pulse Lengthening

First, we investigate the scaling of the pulse spreading with respect to the electron density Σ . We extract the longitudinal pulse width σ_z at 100 ps from data sets using different extraction fields F_a or fluences (F) and plot σ_z as a function of Σ . Interestingly, we observe a power-law scaling of σ_z with Σ with an exponent that is significantly smaller than one, irrespective of the applied extraction field (F_a) or fluence (F). The results are presented in Fig. 5.7(b). Our results differ from the energy-dependent space-charge-led pulse lengthening measurements based on streaking technique, which determines σ_z at a fixed probe location [163]. Since ultrafast imaging tracks temporal

development of σ_z in the COM frame, this explains why there is no explicit energy (F_a) dependence in our data. Furthermore, since a high Σ can be achieved with either high F_a or F , the observation of $\sigma_z(t)$ being exclusively parameterized by Σ indicates the origin of the power-law scaling is mainly due to the space charge effects. Here, to a very good approximation, the power-law exponent γ is near 0.5. At earlier times, γ is smaller, as depicted in the inset of Fig. 5.8(a). For example, at 40 ps, $\gamma \approx 0.37$, and at 60 ps, $\gamma \approx 0.43$, while for later times, γ increases slowly with a terminal between 0.5 and 1. We expect that, at long times, the effect of internal space charge forces decreases and the lateral expansion is important, leading γ to saturate between 0.5 and 1. To understand the scaling law, envelope-based approaches for uniform charge distribution, such as mean-field [14], fluid [175], or self-similar [176] space charge models, are first applied to compare with the experimental data. When the charge distribution is non-uniform, the pulse length can be estimated by the mean-field theory with a scaling factor [14, 176], but the scale factor is different depending on the shape of the pulse [176]. From shadow images, we show that the longitudinal pulse profile remains largely near-Gaussian up to 120 ps, as shown in Fig. 5.8(b), which determines the scale factor for the mean-field theory. Both the one-dimensional fluid model (uniform charge distribution) and mean-field model (non-uniform distribution) predict the electron-pulse longitudinal width evolves self-similarly with $\gamma = 1$, which does not agree with the experimentally observed exponent $\gamma \approx 0.5$. We conjecture that this observation is a manifestation of a non-uniform electron pulse profile seeded with different initial phase-space conditions.

5.4.2 Density-Dependent Broadening of Initial Electron Velocity Distribution

The distinctive $\sigma_z(t)$ from several different Σ s, reported by $\sigma_z(t)$ in Fig. 5.7(a), can be used to investigate the initial phase space of photoelectron pulses by fitting them

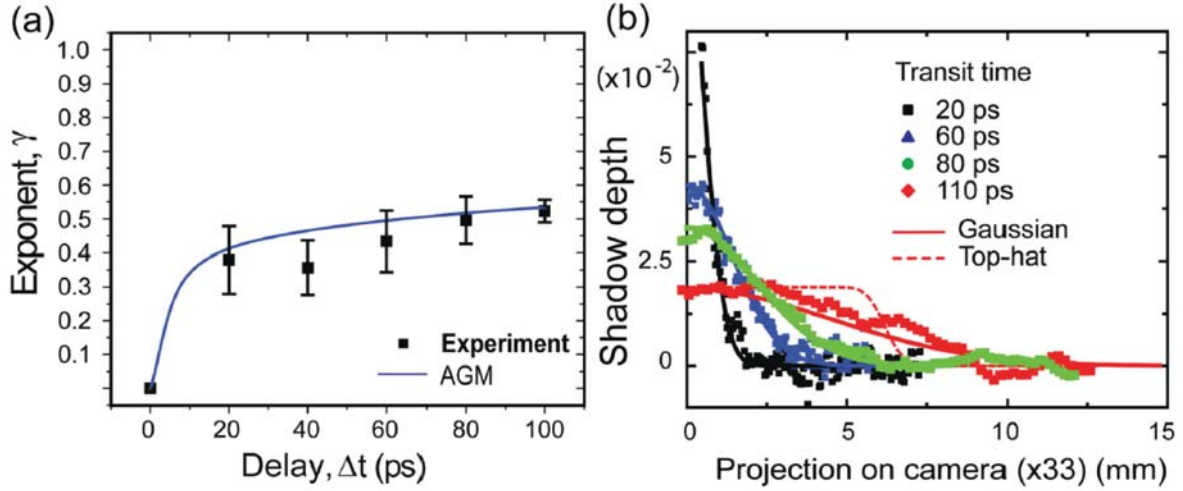


Figure 5.8: (a) The symbols represent the values of γ found at different delay times. The solid line represents results from an analytical Gaussian model (AGM) simulation using broadened initial longitudinal velocity spread due to near cathode space charge effect extracted by fitting the early time σ_z trajectory, which is presented in Fig. 5.9(c) and the text. (b) The symbols are the linescans of the shadow images recorded on the phosphor/CCD screen produced by photoelectrons ($\Sigma = 7.12 \times 10^{13} \text{ e/m}^2$) at different times. The linescans are fitted with a Gaussian profile (solid lines). A top-hat profile (dashed line), convoluted with the projection geometry, is also drawn for comparison. The magnification of the projection imaging is ≈ 33 .

with N-particle simulations. To parameterize the shape of the phase space, we define a chirp parameter a in the N-particle simulations. When $a = 0$, the electron pulse is fully thermal with no correlation between momentum and position, while, in the $a = 1$ limit, there is a perfect correlation between the position and momentum, as shown in Fig. 5.9(a). To compare with experimental results, the width of the momentum distribution σ_{p_z} and the chirp parameter a are varied to fit the trajectory $\sigma_z(t)$. The best fitting results are plotted as the solid lines in Fig. 5.7(a). The initial pulse length $\sigma_z(t=0)$, which is expected to be in the sub-micron range, has limited influence on the results of the simulations at later times ($t > 10$ ps), as the pulse length expands to μm scale within a few ps. Simulations have been performed at $\sigma_z(t=0) = 30$ nm, $1 \mu\text{m}$, and $2 \mu\text{m}$ to confirm this. Thus, $\sigma_z(t=0)$ is set to the resolution limit ($1 \mu\text{m}$) of the experiment to remove this as an adjustable parameter in comparing with the experimental data. At a given σ_{p_z} , we find that $a = 0$ gives the minimum σ_z , while $a = 1$ gives the highest σ_z , as shown by the range of σ_z as a function of a in Fig. 5.9(b). This sensitivity to the initial conditions allows us to parameterize the initial photoelectron phase space. The fitting results based on this parameterized N-particle simulation show a strong growth of the initial velocity spread obtained by $\Delta v_z(t=0) = \sigma_{p_z}/m_e$ when the charge density Σ increases, as depicted in Fig. 5.9(c). The observed significant increase of the initial velocity spread of electron-pulse generation can be associated with a near-cathode space charge effect at ultrashort timescales, which modifies the electron energy distribution from the single-electron emission regime. In contrast, the low-density photoelectron velocity spread at $\Sigma \approx 0$ can be extrapolated from the results in Fig. 5.9(c), and thus determined. $\Delta v_z(t=0) = 0.084 \pm 0.019 \mu\text{m}/\text{ps}$ represents the intrinsic, unperturbed initial velocity spread from photoemission. We note that the rapid increase of $\Delta v_z(t=0)$ cannot have come from the reduced effective work function, due to the applied F_a (the Schottky effect). In fact, for the lowest five Σ s, the applied field is the same ($F_a = 0$), and for the highest Σ , the corresponding energy spread approaches 1 eV, which cannot

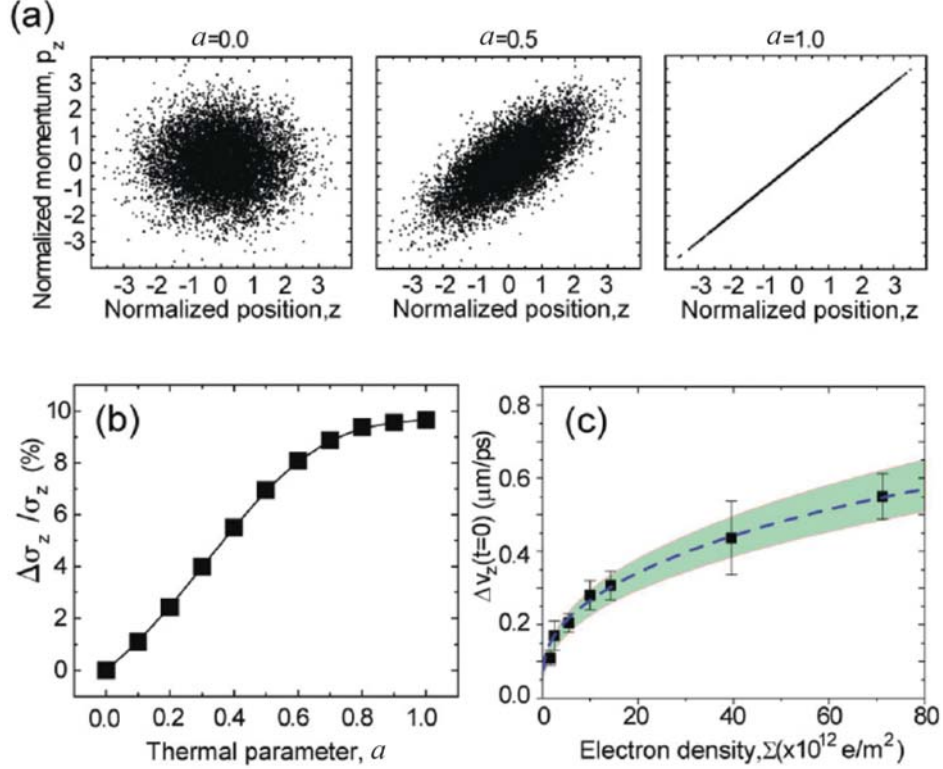


Figure 5.9: (a) The normalized initial longitudinal phase space employed in the N-particle simulations, which is parameterized by a thermal parameter a and scaled by the initial length σ_{zi} and momentum spread σ_{pz} . (b) The percentage change in the pulse length σ_z obtained at 100 ps as a function of a . (c) The initial longitudinal velocity spread, $\Delta v_z(t=0)$, obtained by fitting N-particle trajectories to the imaging data depicted in Fig. 5.7(a). The extraction field F_a applied is 0 at the five lowest Σ s, 0.32 MV/m at $40 \times 10^{12} \text{ e/m}^2$ and at $70 \times 10^{12} \text{ e/m}^2$.

be explained by F_a . Moreover, for all simulations, the strong thermal parameter range $a = 0.0-0.4$ gives a better fit to the data than the correlated limit, $a = 1$, indicating a near-stochastic initial photoelectron distribution.

Here, we are able to reproduce the N-particle results presented in Fig. 5.7(a) using the analytic Gaussian model (AGM) developed by Michalik and Sipe [177, 178] with the initial conditions according to those refined by N-particle fitting in Fig. 5.9(c). This is warranted as, within the observed timescale, the pulse profile resembles a Gaussian function, as shown in Fig. 5.8(b). Thus, the associated pulse spreading due to the nonlinear space-charge force can be properly modeled by AGM. In the following, we

substitute the N-particle simulation with the less computationally expensive AGM to investigate the cause of the scaling behavior identified in the space-charge-led pulse expansion with various initial conditions. Figure 5.10 shows the comparison of AGM simulation and the experimental results. The solid lines show the AGM simulations with the initial conditions given by the N-particle refinement, which correctly predict the trends of the experimental results. In particular, the increase of σ_z over Σ bears a resemblance to the power-law, with an exponent γ increasing from 0 to near 0.5. In contrast, with a constant initial velocity spread of $0.084 \mu\text{m}/\text{ps}$ (the single-electron limit), while also showing incremental increase of slope, AGM* results severely underestimate $\sigma_z(\Sigma)$, as shown in Fig. 5.10. The compliance of N-particle simulation and AGM with our experimental results indicates that the initial electron velocity spread is crucial for the transient power-law behavior observed experimentally. As to the cause of the initial velocity spread, we attribute it to the near-cathode space charge effect. Note that the space-charge-induced energy spread has previously been studied for photoemission experiments [179–181]. These studies mainly investigated the long-time limit of space charge effects, and they are different from the short-time behaviors, which embody the near-cathode space charge effect. The robust trend of $\Delta v_z(t=0)$ as a function of electron density in Fig. 5.9(c), obtained using a range of F_a and F , shows that neither laser heating (F effect) nor Schottky effect (F_a effect) is the main cause of photoelectrons gaining the initial velocity spread and indicates primarily space charge effect in its origin.

5.4.3 Three-step Model of Photoemission

To further understand the near-cathode space charge effect in the generation of photoelectrons, it is instructive to compare the observed photoelectron initial phase space at $\Sigma = 0$ limit with W. F. Krolikowski and W. E. Spicer's three-step model (TSM) [182], which has been further elucidated recently by K. L. Jensen, et al. [173, 174, 183] and

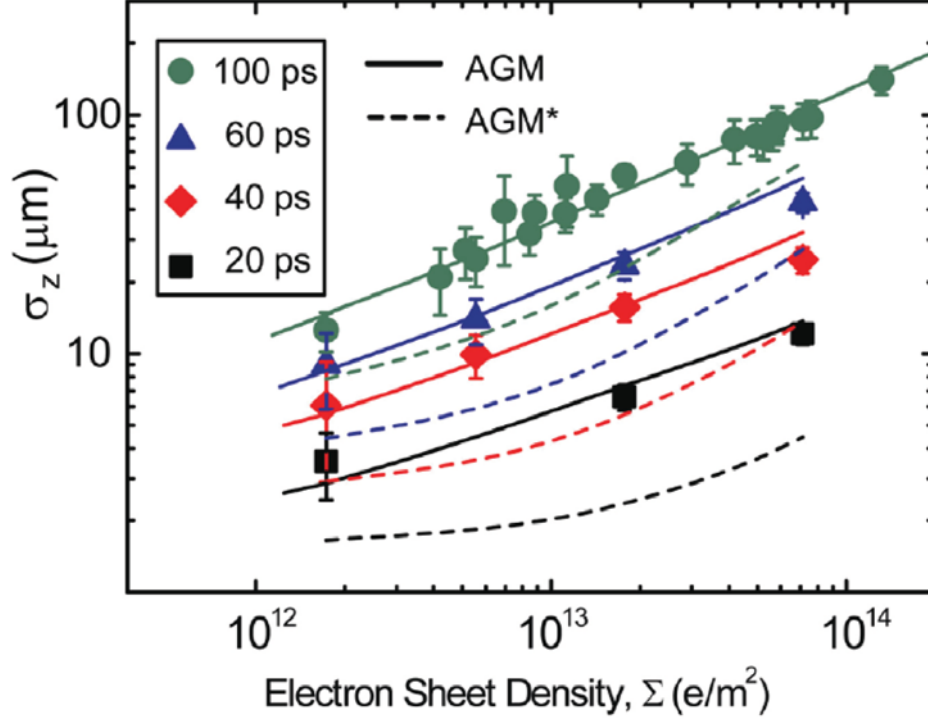


Figure 5.10: The symbols represent the experimental longitudinal pulse length (σ_z) of the photoelectron pulses with different densities tracked at different times. The solid line represents the analytical Gaussian model (AGM) simulation using the different initial longitudinal velocity spread specified by Fig. 5.9(c) and the initial thermal (chirp) parameter, a , set to 0. The dotted line represents the analytical Gaussian model simulation using a constant initial longitudinal velocity spread $\Delta v_z(t=0)=0.084 \mu\text{m}/\text{ps}$.

D. H. Dowell, et al. [184] to model experimental results primarily for the development of free electron lasers. In TSM, which is designed to treat the photoemission at the single-electron limit, photo-excitation gives a constant energy boost $\hbar\omega$ to the electrons, while maintaining a thermal energy distribution based on the initial Fermi-Dirac (FD) distribution, as shown in Fig. 5.11(a) (note FD width is exaggerated here for illustration purposes). To overcome the effective surface work function Φ_{eff} , particle selection is made based on a cut in momentum space at p_z^0 , which gives

$$\theta_{max}(E) = \cos^{-1} \sqrt{(E_F + \Phi_{eff})/(E + \hbar\omega)}, \quad (5.4.1)$$

as depicted in Fig. 5.11(b). Here, E is the total energy of the electron before photon absorption. The effective work function is given by $\Phi_{eff} = \Phi_w + \Phi_{Sch}$, where $\Phi_{Sch} = e\sqrt{eF_a/4\pi\epsilon_0}$ is the Schottky value [184]. The electron velocity distribution following photo-absorption can be modeled by

$$v_z^{in} = \sqrt{2(E + \hbar\omega)/m_e} \cdot \cos(\theta), \quad (5.4.2)$$

whose distribution is calculated using FD statistics and the momentum cut model. The shaded area marks those electrons with sufficient energy to escape the cathode surface in the absence of electronic thermalization. Here, θ is the internal emission angle approaching the surface, which falls in the range $[0, \theta_{max}]$, as depicted in Fig. 5.11(b). The external velocity distribution can then be calculated from v_z^{in} by considering the interface refraction:

$$v_z^{out} = \sqrt{2 \times [m_e(v_z^{in})^2 - E_F - \Phi_{eff}]/m_e}. \quad (5.4.3)$$

We want to note that TSM has been employed successfully to estimate the thermal emittance for metallic photocathodes to within a factor of 2 agreement with the ex-

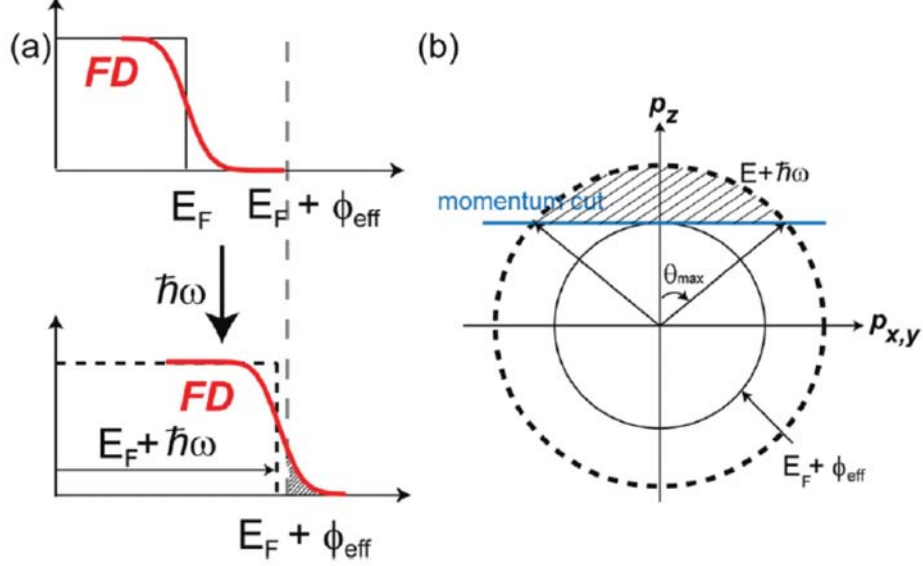


Figure 5.11: Momentum cut model for selecting electrons for photoemission used in the three step model. (a) The electron energy distribution before and after photoexcitation. In this model, the electrons are assumed to escape the cathode surface before thermalization and thus have the same Fermi-Dirac (FD) energy spread (exaggerated for illustration purpose) defined at the temperature prior to photoemission. The shaded area ($E \geq E_F + \Phi_{eff}$) represents the electron population that is qualified for photoemission. (b) The selection of electrons in the three-dimensional momentum phase space that are qualified for photoemission. The electron must have an energy larger than $E_F + \Phi_{eff}$, as described in (a); it must also have a minimum longitudinal momentum (p_z^0) to overcome the work function in order to escape the surface, as specified by the shaded area.

perimentally determined values [173, 184].

To provide a self-consistent, near-cathode space charge model, we first determine the work function Φ_w of our gold cathode by comparing the measured initial velocity spread $\Delta v_z(t=0)$ extrapolated to the single-electron limit ($\Sigma = 0$) with the TSM prediction. We use both N-particle simulation and the TSM analytical model (Eqn. (34) in Ref. [184]) to establish the relationship between $\Delta v_z(t=0)$ and Φ_w , as shown in Fig. 5.12(a). The agreements are generally good, with small deviations mainly in the low Φ_w regime, which are caused by sampling errors due to the limited number of electrons (10^4) used in simulations. We find that convoluting the 0.03 eV bandwidth of the 50 fs laser pulse to model the initial phase space of photoelectrons does not significantly

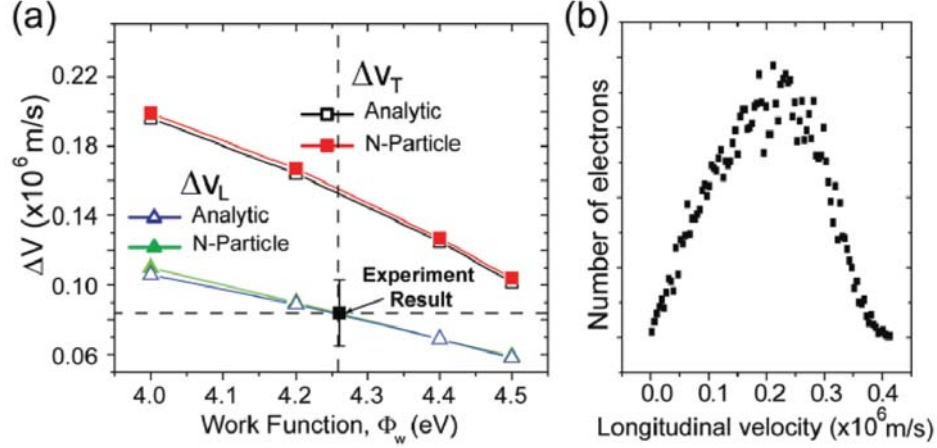


Figure 5.12: (a) The photoelectron velocity spread in the longitudinal (Δv_L) and transverse (Δv_T) directions. The solid symbols represent results obtained using N-particle simulation, and the hollow symbols represent results obtained by integrating the analytical equation reported in Ref. [184]. (b) The longitudinal velocity distribution of photoelectrons generated by N-particle simulation based on the three step model.

alter the result. Based on $\Delta v_z(t=0) = 0.084 \pm 0.019 \mu\text{m/ps}$, we deduce work function $\Phi_w = 4.26 \pm 0.16$ eV for the gold photocathode, which is consistent with literature values [21, 169]. We note that the initial velocity distribution predicted by the TSM is anisotropic, as the longitudinal velocity spread is roughly a factor of two smaller than the transverse velocity spread. This can be understood with the momentum cut model in selecting the electrons for photoemission, as shown in Fig. 5.11, where the extent of the eligible momentum phase space for photoemission along the transverse directions (p_{xy}) is larger than that along the longitudinal direction (p_z). The simulations validated here is further employed to model the photoelectron dynamics near the surface of photocathode to understand the near-cathode space charge effect on the photoemission quantum yield.

5.4.4 Space-Charge Limitation of Pulsed Photoemission Quantum Yield

The quantum efficiency (QE) of photoemission is defined as $QE = \Sigma / (F / \hbar\omega)$, where F is the incident laser fluence. The experimental measured QE is plotted in Fig. 5.13(b) as a function of the extraction fields F_a , which is found to be much smaller than the single-electron TSM limit (red dashed line in Fig. 5.13(b)). The TSM QE is calculated following Eqn. (3) in Ref. [184] using F and F_a employed in our experiment. The single-electron QE thus obtained is $\approx 3.0 \times 10^{-5}$ (based on $F = 5.0 \text{ mJ/cm}^2$, the reflectivity $R = 0.36$, the laser absorption depth $\lambda_{opt} = 9.5 \text{ nm}$ [185], and the mean free path of the photo-excited electrons $\lambda_{e-e} = 6.5 \text{ nm}$ [182]). It is found to be strongly dependent on Φ_w , but weakly on F_a .

In addition to the weak perturbation of F_a , the strong image charge potential Φ_{dp} associated with the fs electron bunch, which is a thin disk during this period, modifies the surface Schottky potential Φ_{Sch} for field-assisted emission, as illustrated in Fig. 5.13(a). To estimate the repulsive dipole potential Φ_{dp} , we use a multi-disk model with an initial velocity distribution calculated based on the TSM, as shown in Fig. 5.12(b). The electron pulse is divided into 5000 slices, leading to a set of 5000 discs of transverse radius $\sigma_{x,y} = 100 \text{ }\mu\text{m}$. Due to the pairwise repulsive forces between the slices, some electron slices return to the photocathode, so we obtain R_{esp} , which is defined as the ratio of the forward moving electrons relative to the total electrons emitted. In the inset to Fig. 5.13(b), we plot the results obtained for $\Phi_w = 4.26 \text{ eV}$, showing that R_{esp} drops significantly in the first 20 fs, indicating a rapid return of electrons to the surface. Multiplying the transient R_{esp} , evaluated at 5 ps, by the TSM QE yields a space-charge limited QE in semi-quantitative agreement with experiment.

While the QE can be improved by increasing the applied excitation photon flux or reducing Φ_w and/or increasing $\hbar\omega$ to expand the available phase space for photoe-

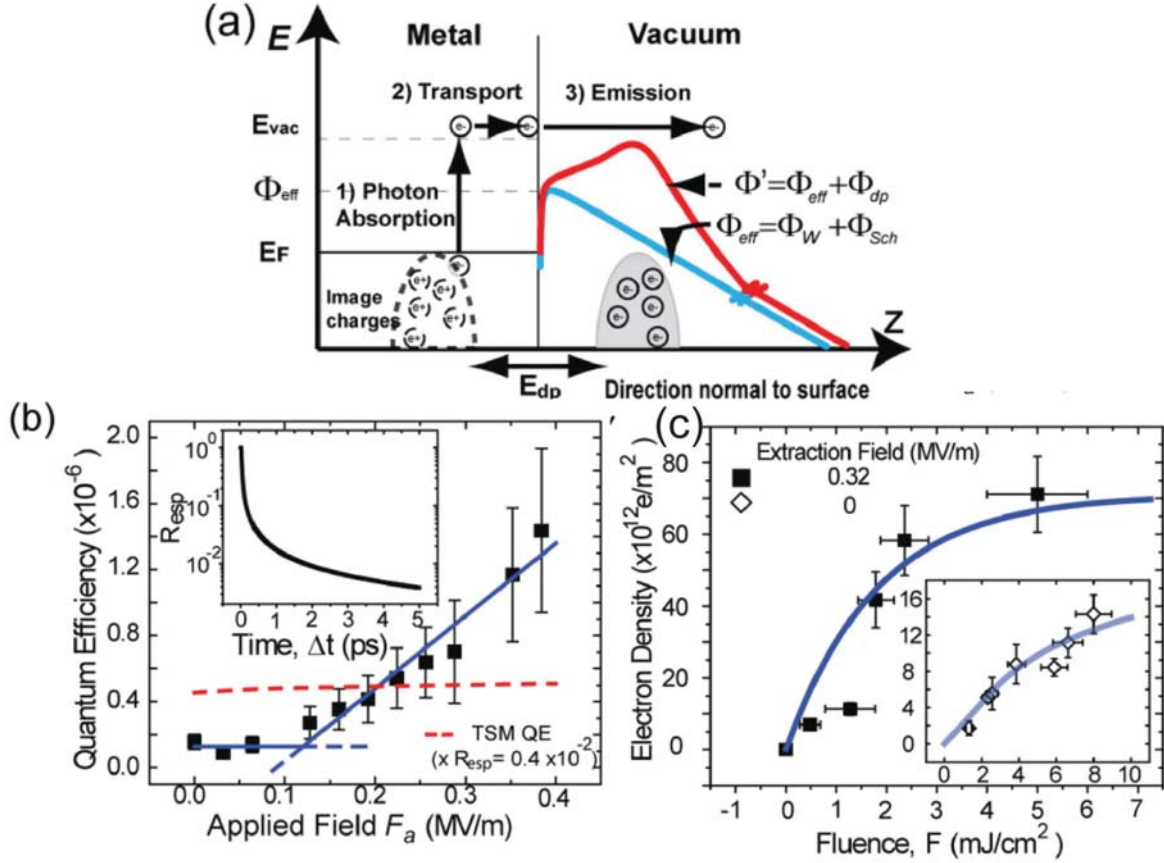


Figure 5.13: (a) Schematic of an extended three step model for intense photoemission. The net work function is the sum of the Schottky potential Φ_{Sch} , the surface dipole potential Φ_{dp} , and the intrinsic work function Φ_w . (b) The quantum efficiency derived based on modified TSM (dashed line) for $\Phi_w=4.26$ eV and the experimental data with a fit to a constant behavior at low field and a linear behavior at high field. (Inset) The electron escape ratio, R_{esp} , is calculated, including the dipole field of the virtual cathode and its image for $\Phi_w=4.26$ eV. (c) Electron pulse sheet density as a function of fluence for high applied field and for zero applied field (inset).

mission, nonetheless, a fundamental space charge limitation arises when the surface dipole field associated with emitted electrons exceeds the extraction field, as has been described by an analytical virtual cathode model [186]. The virtual cathode emission threshold is determined by treating the emitted electrons as a sheet of charge density Σ , which reduces the field at the surface of the cathode according to $F_s = F_a - \Sigma/\epsilon_0$. Photoemission ceases when the applied field F_a is completely screened ($F_s=0$), leading to a pulse-length independent emission threshold $\Sigma_C = \epsilon_0 F_a$, a threshold that depends linearly on the extraction field. This simple model predicts a threshold $\Sigma_C=1.77 \times 10^{13}$ e/m² at $F_a=0.32$ MV/m. The analytical virtual cathode model is tested by our experiment in the fs photoemission regime. We find a saturation of photoelectron flux appears nominally at $\Sigma=7 \times 10^{13}$ e/m² according to Fig. 5.13(c), which is nearly four times larger than the predicted threshold. The discrepancy may be attributed to the approximations inherent in the virtual cathode screening field and to the initial velocity effects that yield photoemission even when $F_a=0$, as shown in the inset of Fig. 5.13(c). Nevertheless discrepancies remain, and a more sophisticated N-particle photoemission model is required to replace the analytical approach to fully account for the virtual cathode effect in the fs regime.

5.4.5 The Effect of Multi-Photon Photoemission

It is prudent to examine whether or not the near-cathode energy spread might be affected by the multi-photon photoemission expected from fs photo-excitation. From the photoemission quantum efficiency study shown in Fig. 5.13(c), which is sub-linear with respect to fluence, we might conclude that the multi-photon photoemission is not yet prominent in the fluence regime investigated here. If the multi-photon photoemission is important, the initial energy spread will significantly increase, due to the generally much larger excess energy associated with multi-photon photoionization ($\Delta E=n\hbar\omega \geq 5$ eV). From the data presented in Fig. 5.7(b) and Fig. 5.9(c), we have not seen an

explicit fluence dependence, which seems to further confirm that the multi-photon ionization is a much less effective channel. Nonetheless, we cannot completely rule out its contribution, albeit being a minor channel, in increasing the initial energy spread through thermalization with the low energy electrons from the single-photon channel. Such an issue is best investigated in a multi-photon photo-gun configuration, where the photoemission is driven using a laser pulse with energy less than the work function of the cathode, where only multi-photon ionization contributes to the photoemission in the future.

5.5 Spacial and Temporal Resolution of Ultrafast Electron Diffraction and Imaging Systems: Analytical Gaussian Model Simulation

We have shown, from our ultrafast imaging of electron pulse generation, the important role of the near-cathode space charge effect in modifying the initial phase space of the photoelectron pulses. The energy spread associated with such an effect can easily exceed the intrinsic electron-energy spread predicted by the three-step model (TSM). This energy spread is different from a Coulomb-explosion-led one, as the near-cathode space charge effect increases the thermal emittance, whereas the energy spreading due to the Coulomb explosion is largely manifested through momentum chirping, not necessarily leading to the emittance growth [42, 187]. This recognition is important in addressing the space charge effects in the development of a high-brightness electron-beam systems for the next generation ultrafast electron diffraction (UED) and microscopy (UEM).

As we have discussed in previous sections, the pulse dynamics with space charge effects can be simulated using the analytical Gaussian model (AGM) [177, 178], in

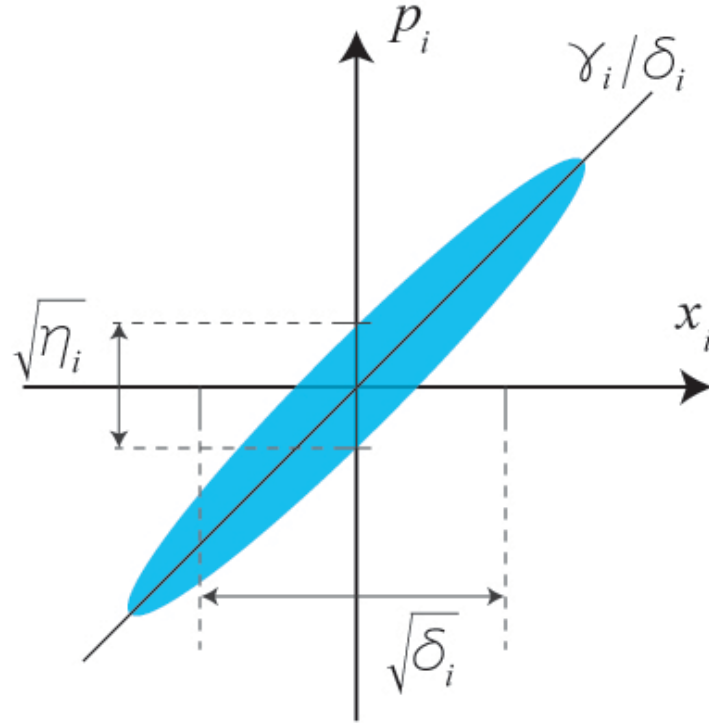


Figure 5.14: A schematic depiction showing the AGM parameters in the representation of the usual position-momentum phase-space plots: $\sqrt{\sigma_i}$ and $\sqrt{\eta_i}$ are the variances of the i direction spatial and momentum distributions, respectively; and the slope, given by γ_i/σ_i , is related to the (linear) momentum chirp of the pulse.

which the electron pulse's charge density is described as:

$$\begin{aligned}
f(x, y, z, p_x, p_y, p_z) &= \frac{N}{(2\pi)^3 \delta_T \eta_T \sqrt{\delta_z \eta_z}} \exp\left[-\left(\frac{x^2}{2\delta_T} + \frac{y^2}{2\delta_T} + \frac{z^2}{2\delta_z} \right. \right. \\
&\quad \left. \left. + \frac{(p_x - \frac{\gamma_T x}{\delta_T})^2}{2\eta_T} + \frac{(p_y - \frac{\delta_T y}{\delta_T})^2}{2\eta_T} + \frac{(p_z - \frac{\gamma_z z}{\delta_z})^2}{2\eta_z} \right)\right], \tag{5.5.1}
\end{aligned}$$

where $\delta_i = \sigma_i^2$ is the spatial variance, with σ_i the σ -width of the electron pulse, η_i is the local momentum variance and γ_i the momentum chirp, $i = \{T, z\}$ stands for the transverse and longitudinal directions, respectively. The meaning of the parameters in the electron pulse phase space is depicted in Fig. 5.14 [177, 178, 188]. Within the employed self-similar Gaussian pulse propagation approximation, A. M. Michalik and J. E. Sipe proved that the dynamics of these parameters can be determined by the following differential equations [177, 178]:

$$\begin{aligned}
\frac{d\delta_i}{dt} &= \frac{2}{m_e} \gamma_i, \\
\frac{d\gamma_i}{dt} &= \frac{1}{m_e} \left(\eta_i + \frac{\gamma_i^2}{\delta_i} \right) + \frac{1}{4\pi\epsilon_0} \frac{Ne^2}{6\sqrt{\delta_i\pi}} L_i(\chi), \\
\frac{d\eta_i}{dt} &= -\frac{2\gamma_i\eta_i}{m_e\delta_i}, \tag{5.5.2}
\end{aligned}$$

where N is the number of electrons, the parameter $\chi^2 \equiv \delta_z/\delta_T$ ($\chi > 0$) quantifies the ellipticity of the pulse and the function $L(\chi)$ is given by

$$L(\chi) = \begin{cases} \frac{\arcsin(\sqrt{1-\chi^2})}{\sqrt{1-\chi^2}} & 0 \leq \chi \leq 1 \\ \frac{\ln(\chi + \sqrt{\chi^2-1})}{\sqrt{\chi^2-1}} & \chi \geq 1. \end{cases}$$

We note that the AGM employs a mean-field approximation in the electron pulse's internal space-charge field. As indicated by Liouville's Theorem, Eqn. 5.5.2 can be reformulated to show that the product $\sqrt{\delta_i\eta_i}$ is conserved in both transverse ($i = T$) and longitudinal directions ($i = z$); that is, the pulse emittance, which is proportional to the

product of the position and momentum variances is conserved in the AGM formalism. We first simulate the temporal resolution of UED and UEM without incorporating RF compression. The initial electron velocity variance, which is measured by the PSImg method and mainly contributed by the near-surface space charge effect, is considered and serves as the initial condition in the simulation. Fig. 5.15 shows that the simulations are consistent with the reported range of temporal resolution in current UED [6, 163, 189, 190] and UEM [191, 192] systems. The difference in temporal resolutions between UEM (solid squares) and UED (hollow diamonds) comes from the difference in cathode-to-sample distance and different beam energy. For single-electron UEM systems, the space-charge induced temporal resolution deterioration is eliminated by very low electron density in individual pulses, and the SNR is, on the other hand, compensated by incorporating a high-repetition range (~ 100 MHz) fs laser trigger; thereby, fs temporal resolution and high coherence length can be achieved [191]. However, the temporal resolution for single electron UED and UEM systems is fundamentally limited by the shot-to-shot fluctuation of photoelectron's velocity [21, 173].

To simulate the electron pulse dynamics in the pulse-compressed UED and UEM apparatus, the ability to model the electron-optical elements, including photo-cathode, magnetic lens, RF compressor, etc., is required. Here, we conduct the AGM simulation for UEM system, which incorporates the external force extension developed by J. A. Berger and W. A. Schroeder [188]. The external force from the magnetic lenses and the RF cavity can be modeled as 'perfect lenses' with parabolic potential, which means the external force is linear with respect to the distance to the center of electron pulses. The external forces in the model is given by

$$\vec{F}_{ext} = -M_T r \hat{r} - M_z z \hat{z}, \quad (5.5.3)$$

where M_T and M_z are the parameters specifying the strength of a magnetic lens for

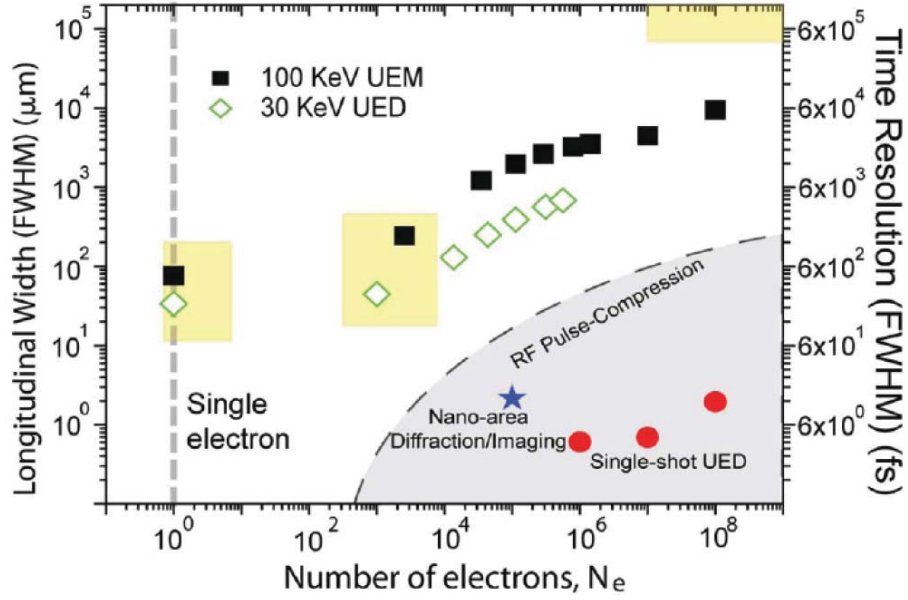


Figure 5.15: (a) Space-charge-limited temporal resolutions in ultrafast electron diffraction (UED) and microscopy (UEM) systems. Solid squares and hollow diamonds show the Coulomb-explosion-led pulse lengthening calculated for 100 keV UEM system (squares) with cathode-to-sample distance of 70 cm and 30 keV UED system (diamonds) with cathode-to-sample distance of 5 cm. The rectangular shaded areas depict experimental resolutions reported in current UED and UEM systems. In comparison, the solid stars and circles show the improvements in temporal resolution by employing an RF recompression in the UEM beam column (see Fig. 5.16a) optimized for nano-area diffractive imaging (stars) and for single-shot UED (circles).

transverse focusing and RF cavity for longitudinal focusing, respectively. When M_T and M_z are positive, focusing forces are considered. Including the external forces, Eqn. 5.5.2 can be reformulated as

$$\begin{aligned}\frac{d\delta_i}{dt} &= \frac{2}{m_e}\gamma_i, \\ \frac{d\gamma_i}{dt} &= \frac{1}{m_e}\left(\eta_i + \frac{\gamma_i^2}{\delta_i}\right) + \frac{1}{4\pi\epsilon_0}\frac{Ne^2}{6\sqrt{\delta_i}\pi}L_i(\chi) + M_i\delta_i, \\ \frac{d\eta_i}{dt} &= -\frac{2\gamma_i\eta_i}{m_e\delta_i}.\end{aligned}\tag{5.5.4}$$

In addition to the external field extension, the effect of acceleration gap is also considered, which arises from the difference in the traveling time of the fast electrons and slow electrons in a single pulse, which leads to an adjustment of momentum distribution in the propagation direction of the pulse when it crosses the anode pin-hole:

$$\delta p_z = \frac{\gamma_c m_e c}{2} \left(\sqrt{1 - \left(\frac{1}{\gamma_c + \Delta v_z^2 / 2c^2} \right)^2} - \sqrt{1 - \left(\frac{1}{\gamma_c - \Delta v_z^2 / 2c^2} \right)^2} \right), \tag{5.5.5}$$

where $\gamma_c = \frac{1}{\sqrt{1-v_e^2/c^2}}$, v_e is the electron pulse COM velocity, and Δv_z is the longitudinal velocity variance when the pulse arrives at the anode. The differential equations above are solved by the 4th-order Runge-Kutta method.

We consider an electron beam column consisting of a 100 keV dc gun, two pre-compressor magnetic lenses (condenser lenses), an RF compressor, an aperture, and a short-focal-distance (~ 2 cm) objective lens, as shown in Fig. 5.16(a) and Table 5.1.

The optimization of the beam quality delivered to the sample plane depends largely on overcoming the strong interplay between the longitudinal and transverse degrees of freedom due to the space charge effect. The strong transverse defocusing induced by the intense space charge force near the focal plane due to longitudinal focusing is compensated by higher strength of the magnetic lens to reach simultaneous focusing (longitudinal and transverse). Specifically here, we consider two regimes of operation. First, for nano-area diffractive imaging [193], the corresponding beam parameter re-

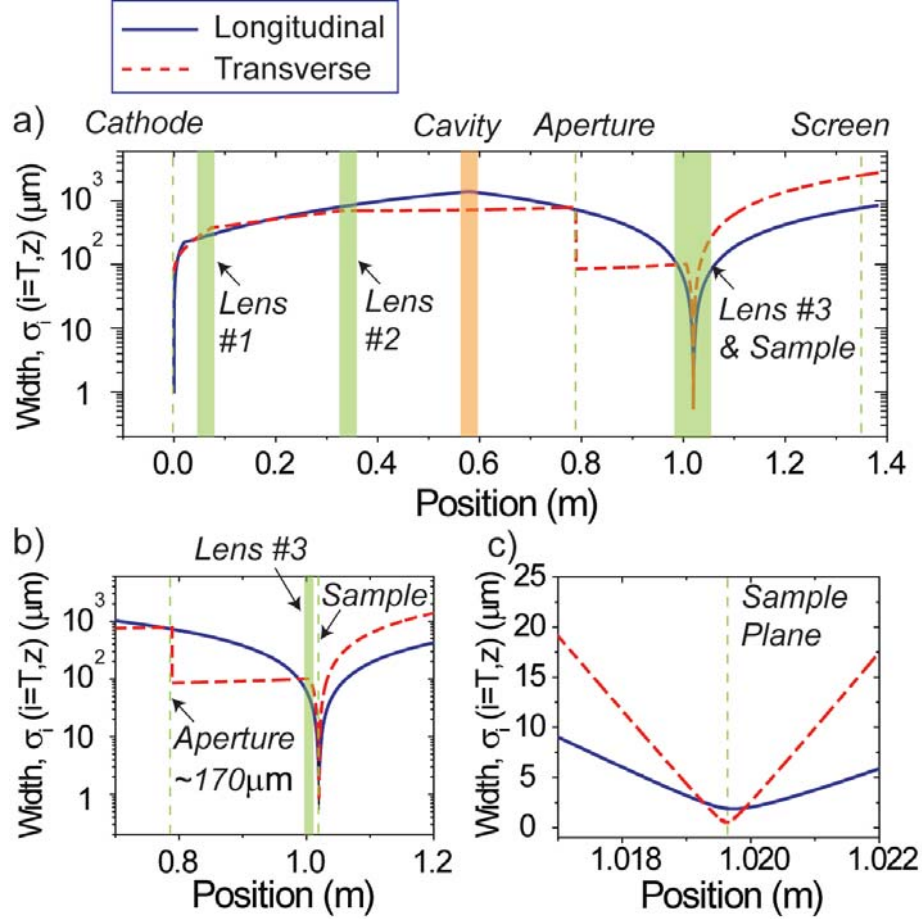


Figure 5.16: (a) Photoelectron pulse evolution along an RF-enabled UEM column. The shaded regions represent the locations of the electron optical elements. (b) A scale-up view of the pulse profiles near the sample plane for nano-area diffractive imaging containing 10^5 electrons/pulse. An aperture with radius of $15 \mu\text{m}$ is employed to thin out the peripheral electrons to achieve a divergence angle $\alpha \leq 1.7 \text{ mrad}$. The minimum transverse radius σ_T is $0.5 \mu\text{m}$, and the minimum longitudinal pulse length σ_L is $1.8 \mu\text{m}$. (c) A scale-up view of the pulse profiles near the sample plane for an ultrafast single-shot UED containing 10^8 electrons/pulse. The minimum transverse radius σ_T is $51 \mu\text{m}$, and the minimum longitudinal pulse length σ_L is $0.88 \mu\text{m}$.

Electron optical component	Position (m)
Cathode	0.000
Anode	0.020
Magnetic lens #1	0.075
Magnetic lens #2	0.329
RF cavity	0.581
Aperture	0.789
Magnetic lens #3	1.000
Sample plane	1.020
Camera screen	1.351

Table 5.1: Location of electron optical components in the pulse-compressed UEM column.

quirements are the transverse coherence length $L_T \geq 1$ nm and the electron probe size $\sigma_T \leq 500$ nm. Such beam characteristics will also be applied for UEM [189, 191]. Second, for single-shot ultrafast electron diffraction, we maintain the coherence length requirement, $L_T \geq 1$ nm, where the requirement of $\sigma_T \leq 1$ μm is relaxed, thereby allowing higher electron flux (10^6 electrons per pulse). Such beam characteristics are suitable for studying the irreversible process [27, 190]. An optimization to simultaneously achieve the high transverse coherence ($L_T \geq 1$ nm) and the high spatial resolution ($\sigma_T \leq 1$ μm) for UEM application is demonstrated in Fig. 5.16. Here, L_T is calculated by $L_T = \lambda_e/\alpha$, where $\lambda_e = 0.0037$ nm is the electron wavelength and α is the half divergence angle at the sample plane. α is determined by $\alpha = \Delta v_T/v_e$, where $v_e = 1.64 \times 10^8$ m/s is the relativistic electron-pulse COM velocity at 100 keV. To achieve $L_T \geq 1$ nm, the beam's half divergence angle at the sample plane, α , must satisfy $\alpha \leq 2$ mrad. Under these tight simultaneous spatial and temporal focusing requirements, the maximum number of available photoelectrons at the sample plane is reduced to $\sim 10^5$ electrons per pulse to maintain a long coherent length, which is selected by a small aperture in front of the objective lens. On the other hand, if the transverse confinement requirement is not required, the number of electrons can be significantly increased to $\sim 10^8$ electrons per pulse and, at the same time, the fs pulse length and high beam coherence are also maintained. In this operation mode, the pulse-compressed UED system can achieve single-shot, fs time-resolved diffraction, by removing the aperture and adjusting the focusing strength of the magnetic lenses (mainly the objective lens).

5.6 Conclusion

In this chapter, we introduced the projection-shadow-imaging (PSImg) technique and used it to investigate the space charge effects in the near-cathode-surface region. Three essential space-charge-led features have been elucidated: the pulse lengthening follow-

ing a power-law scaling, the broadening of the initial energy distribution, and the virtual cathode threshold. The impacts of these space charge effects on the performance of the ultrafast electron diffraction and microscopy have also been discussed. We used the analytical Gaussian model (AGM) to simulate the beam dynamics in the RF-compression ultrafast electron microscope (UEM). While the AGM model provides a rather promising outlook for the development of an RF-enabled high-brightness UEM, the technical challenges to achieve the prescribed spatial and temporal resolutions cannot be underestimated. The parameters for achieving the space-time focusing will likely vary from the optimization provided by AGM, when considering the inhomogeneous electron bunch and the nonlinear effects in the electron optics. Clocking the RF field with fs electron pulse to the precision demanded by the required temporal resolution is also required. Fortunately, these technologies do exist in the mature fields of electron microscopy [194] and precision RF-laser synchronization [195]. In Fig. 5.17, we plot the structure of the RF-compression UEM, which is constructed following the geometry reported in Fig. 5.16(a) and Table 5.1.

In particular, the ultrafast projection shadow imaging method introduced here can also be applied to different cathode designs and be incorporated at different stages along the column to characterize the transient beam dynamics. Different from other ultra-short pulse characterization methods [16, 36], the imaging probe provides simultaneous full-scale longitudinal and transverse pulse characterization, which is ideally suited to directly compare with beam dynamics simulations, thus forming an experimentally informed optimization scheme in designing electron optics for future high-brightness UEM systems.

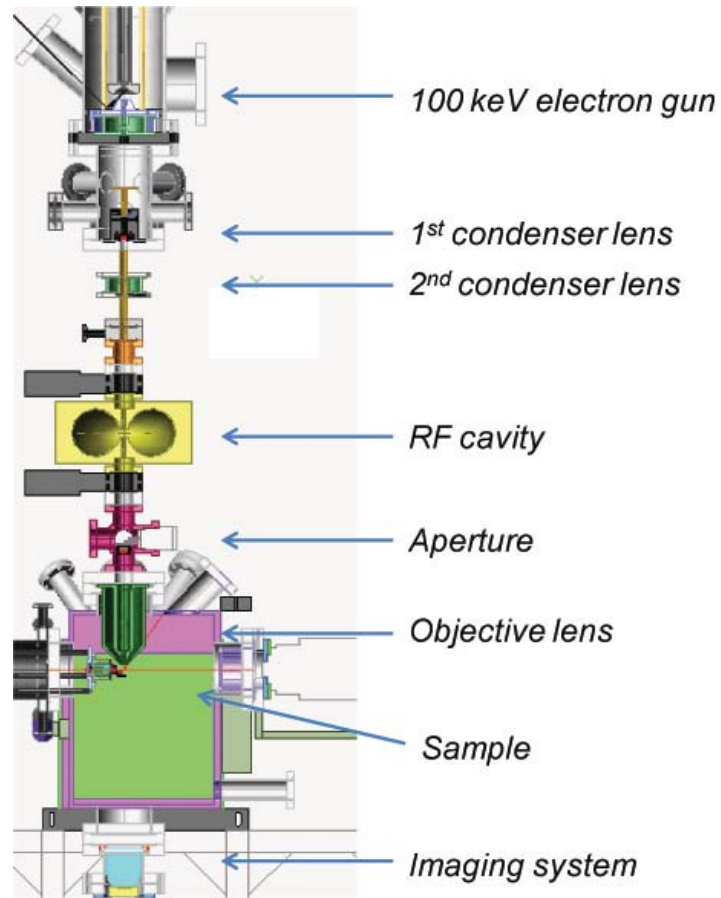


Figure 5.17: Structure of RF-compression UEM constructed at Michigan State University, with important components labeled.

Chapter 6

Low-level Radio-frequency Synchronization System

In the ultrafast electron microscope (UEM), the radio-frequency (RF) cavity is the ‘heart’ of the system, which compresses the electron pulse in its longitudinal direction and restores the temporal resolution. The general process which compresses the electron pulse with the RF field is depicted in Fig. 6.1. The idea is to apply a precisely timed oscillating electric field, such that the faster particles at the front of the bunch are decelerated, while the slower particles at the rear are accelerated, which leads to a ballistic compression in the subsequent drift space and reaches the minimum longitudinal size at the sample position. This oscillating electric field can be the TM_{010} mode in a cylindrical RF cavity, which is applied along the longitudinal direction. An RF cavity is a resonant component, which has its resonant frequency f_R . The f_R of our cavity is ~ 1 GHz. To compress the electron pulses, the switching time of the polarity of the electric field (zero-crossing point) has to coincide with the arrivals of the electron pulses. Because the electron pulses in the UEM system are originally triggered by femtosecond laser pulses generated by the laser systems, the synchronization between the RF field and the electron pulses is achieved, if the phase of the RF field and the

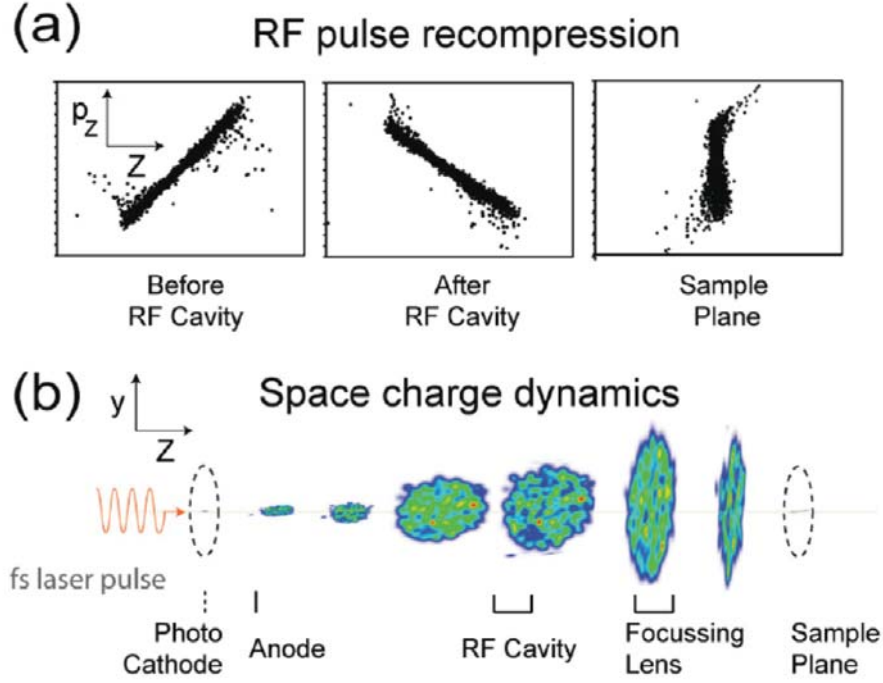


Figure 6.1: Concept of an electron beam injector column for ultrafast electron microscope demonstrated using an N-particle simulation of electron pulse propagation in an electron beam column with an RF cavity. (a) Phase space adjustment before and after the RF cavity. (b) The corresponding real space pulse profile at each electron optical components.

laser-pulse generation can be synchronized.

In this chapter, the low-level (low-power) synchronization system for the RF-compression UEM will be introduced. The optical frequency comb and how to do the comb selection will be first introduced in Section 6.1. It is followed by the introduction of phase-locked loop (PLL) and its linear theory in Section 6.2. The purpose of the PLL is to tightly lock the phase of fs laser pulse train with the RF electric field. Phase noise is a direct measurement of how good this lock is. The concept will be introduced in Section 6.3 and the effects of phase noise and static phase shift will be discussed as well. The loop filter design, which helps to minimize the timing jitter from phase noise and the shift of static phase error, will be introduced in Section 6.4. Section 6.5 contains the detailed setup of PLL, and the characterization of PLL is included in Section 6.6 and Section 6.7. At the end, a conclusion is given in Section 6.8.

6.1 Optical Frequency Comb and Comb Selection

Laser-RF synchronization via optical frequency combs is a technique now frequently used in optical frequency metrology and accelerator physics [196, 197]. The central concept of the optical frequency comb is that the pulse train generated by a mode-locked laser has a frequency spectrum that consists of a discrete, regularly spaced series of sharp lines, known as a frequency comb. In other words, a regularly spaced train of femtosecond laser pulses corresponds to a optical frequency comb in the frequency domain. This technique was realized more than 40 years ago by Hänsch [198] and Baklanov and Chebotayev [199]. In optical frequency metrology and synthesis, the carrier-envelope phase is also well controlled. However, it is not important for our applications.

For a single pulse, it has a spectrum that is its Fourier transform of its envelope function and is centered at its carrier frequency. Generally, for any pulse shape, the spectrum width is inversely proportional to its temporal duration of the envelope. If there is a train of identical pulses, which are separated by a fixed interval, the spectrum will be again the Fourier transform of the pulse train from the time domain to the frequency domain, which yields a comb of regularly spaced frequencies and the comb spacing is inversely proportional to the time between pulses, i.e. it is the repetition rate of the laser that is producing the pulses. Fig. 6.2 summarizes the time-frequency correspondence for a pulse train.

Physically, the optical frequency comb can be generated by simply collecting the laser pulse train using a high-speed photodetector (Newport 818-BB-45). With the Fourier transform from real-time to the frequency domain, the comb structure can be directly observed. An example of optical frequency comb measured using the spectrum analyzer (ROHDE&SCHWARZ FSP13) is plotted in Fig. 6.3(a). Because of its very narrow line width, each comb in the spectrum is a nearly ‘perfect’ sin-wave in the

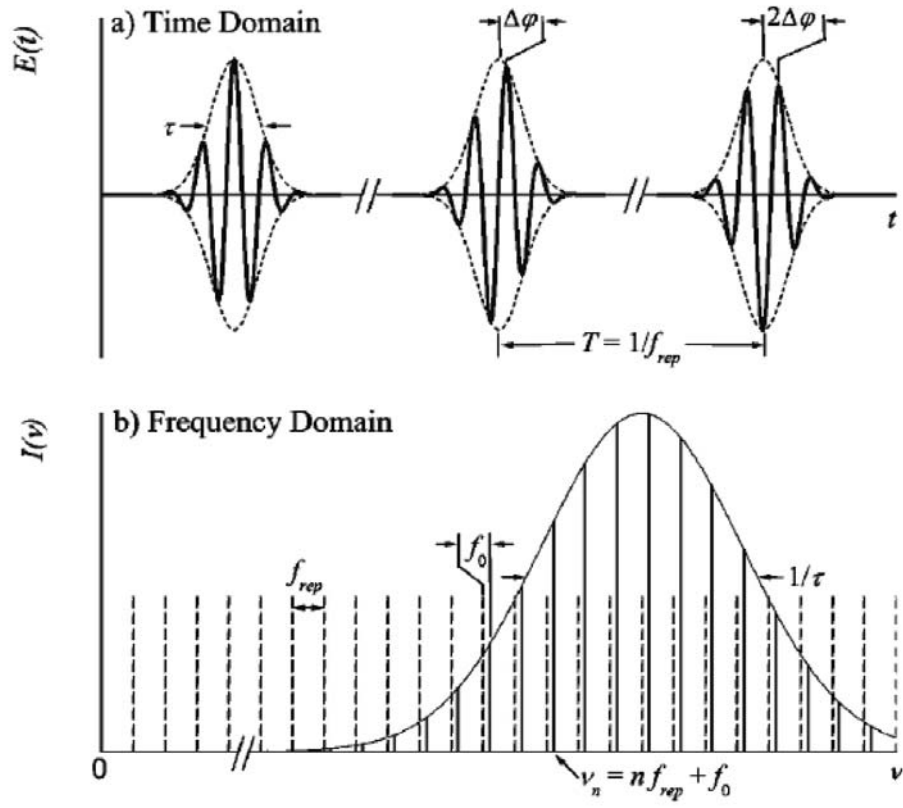


Figure 6.2: Summary of the time-frequency correspondence for a pulse train [196].

time domain, which can be selected by a narrow bandpass filter as shown in Fig. 6.3(b). In general, the harmonics of any order of the base repetition rate can be extracted, and used as the reference source for synchronization. We chose to use its 4th harmonic of frequency ~ 338 MHz, because of the availability of reliable and inexpensive components for constructing the phase-locked loop (PLL).

The major advantage of using a fs laser as the reference source is the very low phase noise and exceptional spectral purity it exhibits. It was demonstrated experimentally that the frequency stability of extracted microwave signal can reach single-sideband phase noise $L(f) = -104$ dBc Hz⁻¹ at 1 Hz offset from the carrier frequency, and decrease to near the photon shot-noise-limited floor of -157 dBc Hz⁻¹ at an offset of 1 MHz. The integrated timing jitter over this bandwidth is only 760 as [200].

6.1.1 Timing Jitter and Temporal Resolution in UEM

Before introducing the phase-locked loop, we need to ask two questions:

1. How does the phase noise or the timing jitter affect the temporal resolution of the UEM?
2. How much timing jitter can we tolerate?

In the RF-compression UEM system, the oscillating electric field in the RF cavity works in a way that the fast moving electrons in the leading half are decelerated, while the slow moving electrons in the tailing half are accelerated, and, hence, the polarity of the phase space chirping is reversed, as shown in Fig. 6.1. The most straightforward concern of the timing jitter would be that it might change the slope of the compressing force in time domain, by mapping electron pulse to a higher phase angle of the oscillating electric field, which reduces the compression strength of the RF field. However, it is found that, unless the phase jitter is as large as tens of picoseconds, this effect is not an important contribution to the RF compression. This is due to the

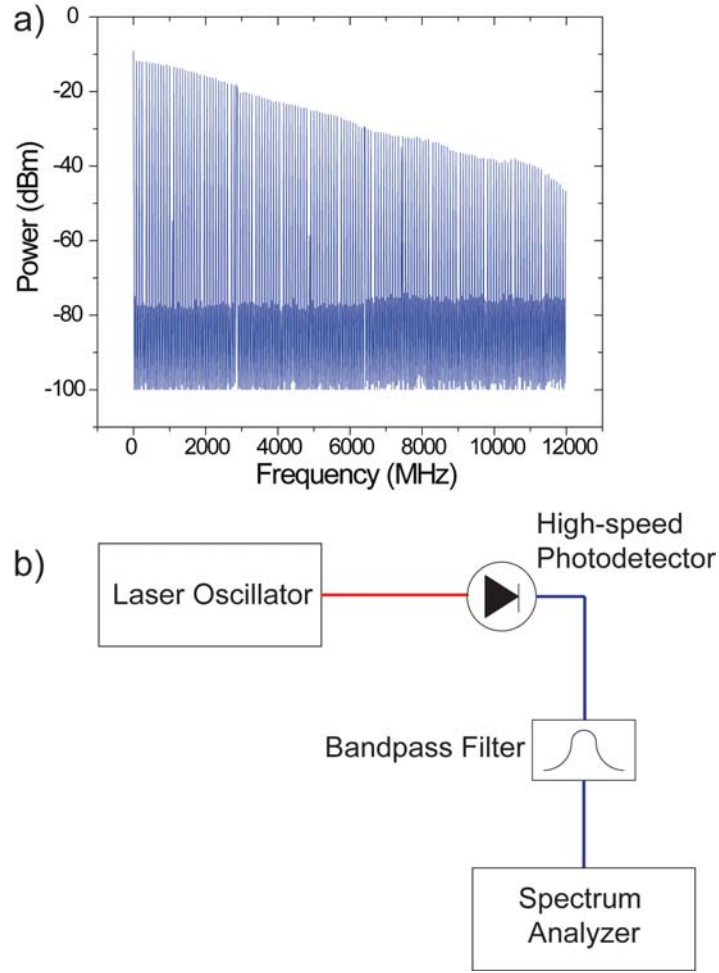


Figure 6.3: a) Optical frequency comb measured using a high-speed photodetector and a spectrum analyzer. Each equally spaced line represents a comb frequency. The spacing between combs equals to the repetition rate of laser pulse train. b) Experimental setup for single-harmonic extraction. The RF bandpass filter is used to select a single harmonic from the comb structure.

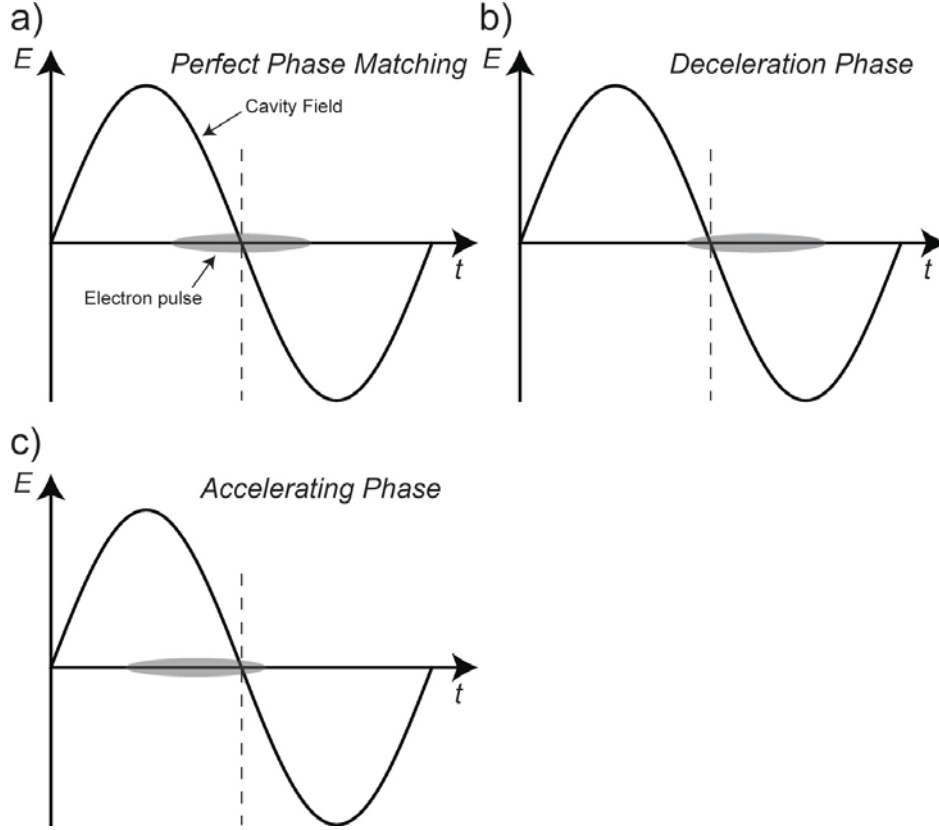


Figure 6.4: a) Perfect electron pulse timing matches with the RF cavity field. No COM velocity change. b) and c) are the situations where the electron pulse timing is shifted relatively to the RF-field phase to decelerate the COM in b) or accelerate the COM in c), respectively.

fact the temporal extension of an electron pulse is on the order of ten ps when it arrives in the cavity, while a half period of the RF field is ~ 500 ps (1 GHz), which is much larger than the pulse duration.

The temporal resolution can be degraded by the RF timing jitter in an indirect way as it varies the arrival time of the electron pulses to the sample plane, which determines the time delays in any time-resolved experiment.

In Fig. 6.4(a), we first plot the situation where the center-of-mass (COM) of the electron pulse arrives at the center of the cavity, when the RF field crosses zero. In this situation, the COM velocity stays the same, which is defined as a perfect phase matching situation. We assume the arrival time of the electron pulse at the sample

plane is t_0 in this situation. Now, let us take phase jitter of the RF field into consideration. In Fig. 6.4(b), the RF field is shifted in the time domain that the COM of the electron pulse falls into the region with decelerating fields. In this situation, the COM velocity will be decreased by the RF field, and its arrival time t_1 is longer than t_0 ($t_1 > t_0$). In the opposite situation, which is shown in Fig. 6.4(c), the COM velocity of the pulse is increased, and its arrival time t_2 is shorter than t_0 ($t_2 < t_0$). As we will see later, the phase jitter of the RF field is mostly random process. We can imagine that the phase of the RF field jitters back and forth around the perfect matching condition. The variation of the pulse arrival time at the sample plane will statistically broaden the pulse width and ultimately limits the temporal resolution of the system. Using a mean-field model simulation, it is found that the timing jitter for the RF-compression UEM has to be less than 0.87 ps to achieve sub-ps temporal resolution [159, 201], which corresponds to <0.3 Deg in the phase jitter for a frequency of ~ 1 GHz.

6.2 An Introduction to Phase-Locked Loop

The purpose of the low-level RF development is to tightly lock the fs laser pulse train with the RF electric signal feeding to the RF cavity. To achieve this, we designed and built a *Phase-locked Loop* (PLL). A PLL, by definition, is an electrical module (circuit) to lock the phase of the output to the input. This is a technology developed as early as the 1950s [202], and has been widely used in applications including electronics, communications, etc. It is also widely used in different fields of physical research, including optical metrology [196, 197] and accelerator physics (to synchronize multiple cavities and other RF components).

A PLL is constructed as a feedback control loop, as shown in Fig. 6.5, where the phase detector ϕ compares the phases of the *Variable* signal with the *Reference* signal, and a voltage V_d proportional to the phase difference is developed and serves as the

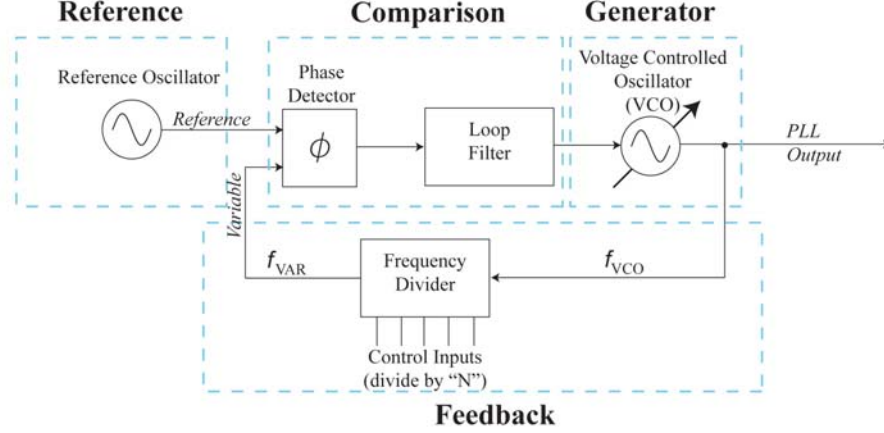


Figure 6.5: Block diagram of a general phase-locked loop (PLL).

control voltage V_c of the voltage-controlled oscillator (VCO) in the ‘Generator’ section. Through a negative feedback, the output frequency is locked to the reference frequency with a fractional number ‘ N ’, depending on the divider setup. The phase error between the reference and the generator signal is hence kept to a fixed value.

6.2.1 Static Phase Error

The phase detector (PD) and the voltage-controlled oscillator (VCO) are the two most critical components in a PLL. If θ_e represents the phase difference between the reference and the variable signals in Fig. 6.5, the voltage, V_d , generated by PD is usually proportional to θ_e ,

$$V_d = K_d \theta_e + V_{do}, \quad (6.2.1)$$

where the linear coefficient K_d is the PD gain and V_{do} the free-running detector offset voltage.

On the other hand, for a typical VCO, its output frequency ω_o is also linearly scaled with its control voltage V_c . The slope of the VCO characteristic in the vicinity of the lock-frequency is called the VCO gain K_o , where

$$K_o \equiv \frac{d\omega_o}{dV_c}. \quad (6.2.2)$$

With Eqn. 6.2.1 and Eqn. 6.2.2, an important concept of the PLL can be derived, which is called the *static phase error*. To simplify the discussion, we ignore the frequency divider in the feedback path in Fig. 6.5, and the loop filter gain is also fixed to be unity. The input signal and output signal are both assumed to be pure sinusoids, which is given by

$$\begin{aligned} v_i &= \sin(\omega_i t + \theta_i), \\ v_o &= \sin(\omega_o t + \theta_o). \end{aligned} \tag{6.2.3}$$

The phase error is given by $\theta_e = \theta_o - \theta_i$. After the phase error propagates through PD, PD generates a voltage $V_d = K_d \theta_e + V_{do}$, which is applied to VCO and changes VCO's output frequency. So, the output frequency of VCO is

$$\begin{aligned} \omega_o &= K_o V_d \\ &= K_o (K_d \theta_e + V_{do}). \end{aligned} \tag{6.2.4}$$

If we assume, when PLL is in-lock, the VCO's control voltage is V_{co} , then it is obvious that $\omega_i = \omega_o = K_o V_{co}$. Thus, the frequency deviation of output frequency to input reference is given by

$$\begin{aligned} \Delta\omega_o &= \omega_o - \omega_i \\ &= K_o (K_d \theta_e + V_{do}) - K_o V_{co} \\ &= K_o (K_d \theta_e + V_{do} - V_{co}). \end{aligned} \tag{6.2.5}$$

Again, we use the condition of equilibrium that $\langle \Delta\omega_o \rangle = 0$, where $\langle \dots \rangle$ stands for average over time. Thus, the phase error in equilibrium is $\theta_{eo} = \langle \theta_e \rangle = (V_{co} - V_{do})/K_d$. So, it is found that a static phase error is inevitable, if $V_{co} \neq V_{do}$. In practice, if the static error θ_{eo} is a fixed value over time, then it does not affect the phase stability between reference and output, but it adds an additional static phase difference. However, we will find that, when the reference frequency drifts over time, the static phase error

varies with it and the variation has to be minimized in our system, which will be discussed in Sec. 6.3.2 and Sec. 6.4.

6.2.2 Linear Model of the PLL

A system can be described by the linear model, if there are linear relations between the system excitations and responses. The linear responses of the PD (Eqn. 6.2.1) and the VCO (Eqn. 6.2.2) indicate that a PLL is a linear system, although this linearity holds only for a limited range of operation. So, the linear model discussion here is limited to small perturbations at the equilibrium state of a PLL, which do not push the system to its nonlinear region. From Eqn. 6.2.1 and Eqn. 6.2.2, it is found that V_d of the PD is in a linear relation with the phase error θ_e , while it is the frequency ω_o of the VCO, which is in a linear relation with V_c . So, to be consistent, a connection between the frequency and phase needs to be formulated to model the VCO. As the PLL is in-lock, it means that the output signal ω_o oscillates with the same frequency as the reference input ω_i . If there is a phase fluctuation on the θ_o , the output frequency shall also vary as the derivative of θ_o , which is given by

$$\Delta\omega_o = d\theta_o/dt \quad (6.2.6)$$

or

$$\theta_o = \int \Delta\omega_o dt, \quad (6.2.7)$$

where $\Delta\omega_o$ is the deviation of the output frequency to the lock frequency. So, when modeling VCO, the integration in Eqn. 6.2.7 has to be taken into account to convert its frequency response to the phase response. In general, the linear response of a system can be formulated through differential equations in the time domain. But a simpler description based on Laplace transform is widely used in the engineer community, which is called *transfer functions*. By applying the Laplace transform, the dc components

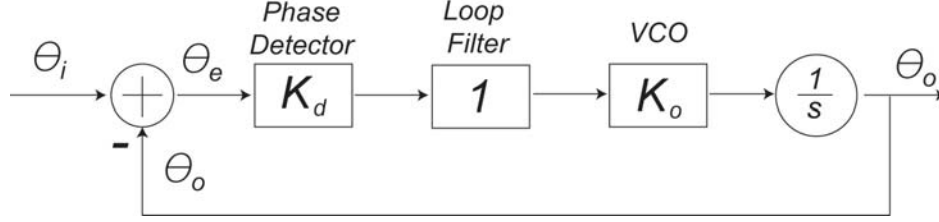


Figure 6.6: Signal-flow graph of a PLL under linear model. Loop filter is assumed to be unitary in this model.

of the PD and the VCO are basically ignored, and only the ac components, which represent the transient responses, are of interest. In other words, the static phase error is zero after the Laplace transform. Fig. 6.6 shows the signal-flow graph of this ac model of the PLL [203]. For the PD and the VCO, their transfer functions are simply their gains K_d and K_o , respectively. The block of $1/s$ is an integrator for the VCO, which is in its form after Laplace transform. s is the complex frequency, which can be replace by $j\omega$, when the details of the frequency responses are of interest.

6.2.3 PLL Transfer Function

With the linear model, several important concepts of PLL can be defined. First one is the *forward gain*, which represents the gain of a signal propagates rightward in Fig. 6.6. It is defined as

$$G(s) = K_d K_o / s. \quad (6.2.8)$$

The *(open) loop gain* is the gain, when the signal travels a full cycle in the PLL, but with the negative feedback path not connected. Since the divider is ignored, the gain on the negative feedback path is unity. So, the open loop gain in this situation equals to the forward gain in Eqn. 6.2.8.

The *transfer function* of the PLL is defined as the ratio between the output phase $\theta_o(s)$ and the input phase $\theta_i(s)$. It can be derived in a self-consistent way. Again, the phase error is $\theta_e(s) = \theta_o(s) - \theta_i(s)$, which becomes $G(s)\theta_e(s)$, after the forward gain is

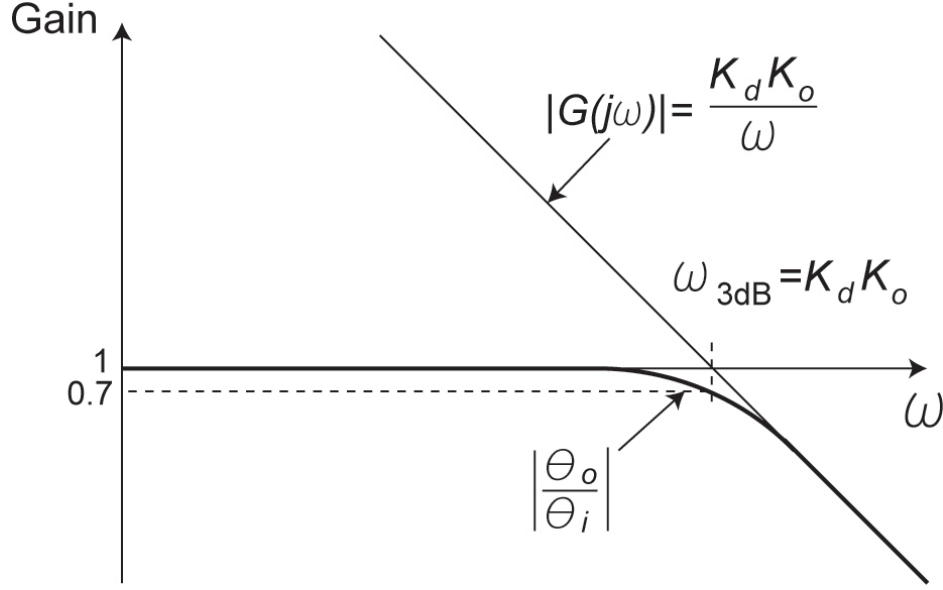


Figure 6.7: Frequency response of PLL.

applied. Obviously, this phase should be equivalent to the output phase $\theta_o(s)$ in the equilibrium state. So, we have the equation as follows

$$\theta_o(s) = G(s)(\theta_o - \theta_i). \quad (6.2.9)$$

Thus, $\theta_i(s) = [1 + G(s)]\theta_o(s)/G(s)$ and the transfer function of the PLL, $H(s)$, is given by

$$\begin{aligned} H(s) &= \frac{\theta_o(s)}{\theta_i(s)} \\ &= \frac{G(s)}{1 + G(s)} \\ &= \frac{K}{s + K}, \end{aligned} \quad (6.2.10)$$

where $K = K_d K_o$, which is sometimes called the *gain* of PLL. The PLL described by Eqn. 6.2.10 is also called a *first-order phase-locked loop*, because its transfer function has a first-order polynomial of s in the denominator.

From the transfer function of the PLL, its frequency-dependent performance can be extracted. In Fig. 6.7, the transfer function in Eqn. 6.2.10 is plotted as a function of the frequency ω ($s = j\omega$) in the log scale. At low frequency, when $|G(j\omega)| \gg 1$, $\theta_o(s)/\theta_i(s)$

is about unity. At high frequency, when $|G(j\omega)| \ll 1$, $\theta_o(s)/\theta_i(s)$ approximately equals to $|G(j\omega)|$. The *bandwidth* of the PLL is defined as the frequency, when $|G(j\omega)| = 1$, which gives $1 = K_d K_o / \omega_{3dB}$, or

$$\omega_{3dB} = K_d K_o. \quad (6.2.11)$$

Broad bandwidth of a PLL is not always good, since high-frequency noise can infiltrate into the PLL if attenuation is not sufficient. From Eqn. 6.2.11, it is obvious that the bandwidth can be reduced by reducing the PD gain K_d or the VCO gain K_o . However, in this way, the key components, such as VCOs and PDs, need to be replaced by new components with correct specifications, if there are any available in the market. A simpler way is to place another block with the gain K_h in the signal-flow diagram, as shown in Fig. 6.8(a). This block physically can be made by a voltage attenuator consisting of resistors R_0 and R_2 in this specific case, as shown in Fig. 6.8(b), and K_h thus equals to $R_2/(R_0 + R_2)$. Then, the new loop bandwidth is

$$\omega_{3dB} = K_d K_h K_o. \quad (6.2.12)$$

The block K_h in Fig. 6.8(a) is called a *loop filter*, and, more specifically, a *passive loop filter*, because there are no active components involved in it. As shown in the example above, the loop filter is a very powerful tool for engineers to control the PLL characteristics, without replacing the key components.

6.3 Phase Noise and Static Phase Shift

After we understand the general concepts of the PLL and available tools to model and control the performance of a PLL, let us take a close look at a PLL with undesirable phase variation, which can be caused either by phase noise or by the shift of static phase error.

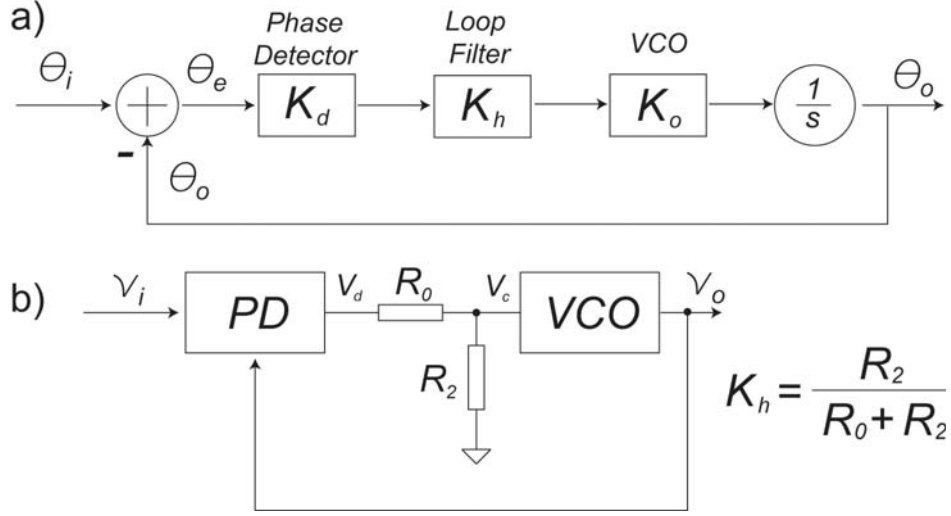


Figure 6.8: a) Signal flow diagram of a PLL with the loop filter, K_h . b) Simple voltage attenuator design of the loop filter to reduce the PLL bandwidth.

6.3.1 Phase Noise

In the linear model, the performance of phase-locked loops with perfect signals has been considered in the previous section. However, noise is always present in the real world. The noise of an oscillator can be defined by the *power spectral density*. An example of a random noise waveform $n(t)$ is shown in Fig. 6.9(a). A measure of its strength is the mean-square value $\overline{n^2}$, as shown in Fig. 6.9(b), also called its ‘power’.

$$\begin{aligned} \overline{n^2} &\equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T n^2(t) dt \\ &= \int_0^\infty \Phi_n(f) df, \end{aligned} \tag{6.3.1}$$

where $\Phi(f)$ is the power spectral density in the frequency domain and the integral of noise power is simply the area under the curve in Fig. 6.9(c). So, we can see that the phase noise can be measured through measuring the spectrum of a signal. The equivalence of the two areas in Fig. 6.9(b) and Fig. 6.9(c) is a convenient tool for finding the rms time jitter from the measured spectrum, which will be described in Sec. 6.6 in more detail. Here, a few of definitions are in order. For a typical spectrum

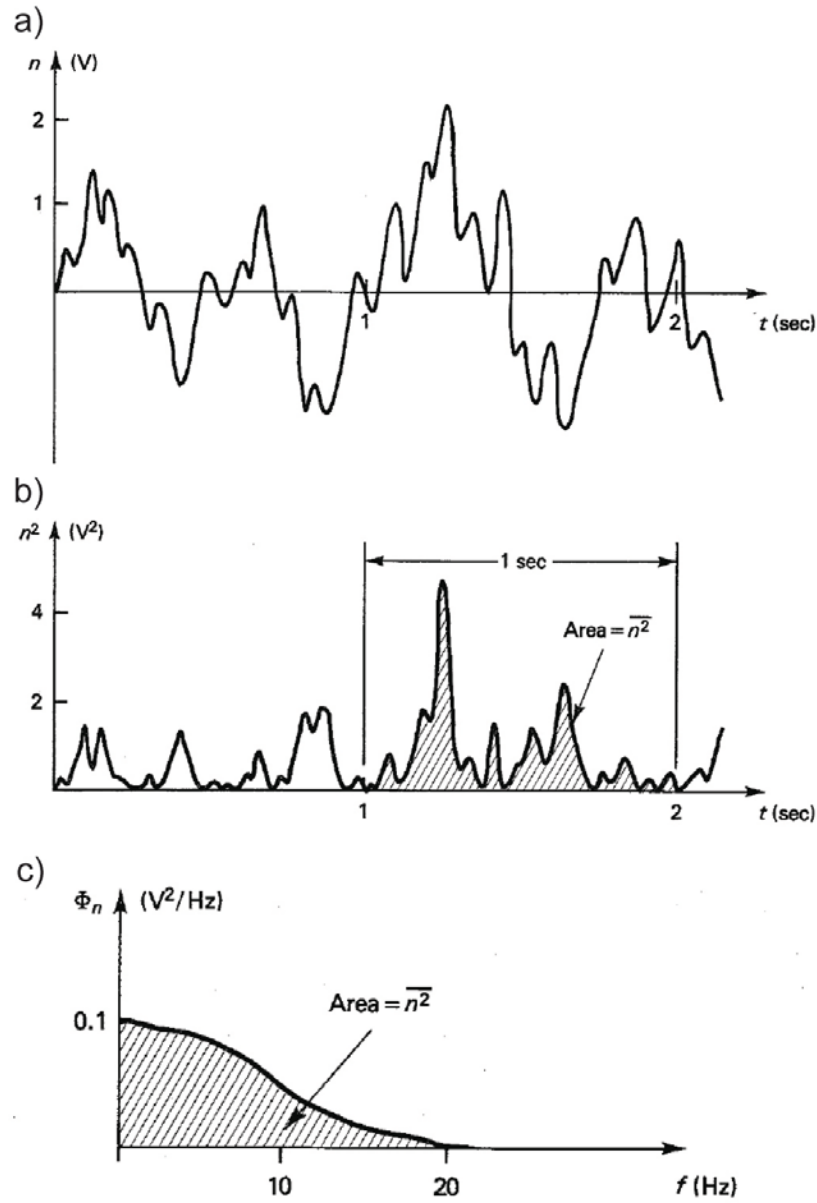


Figure 6.9: a) Noise as a function of time; b) Mean-square noise in the time domain; c) Power spectral density of the noise [203].

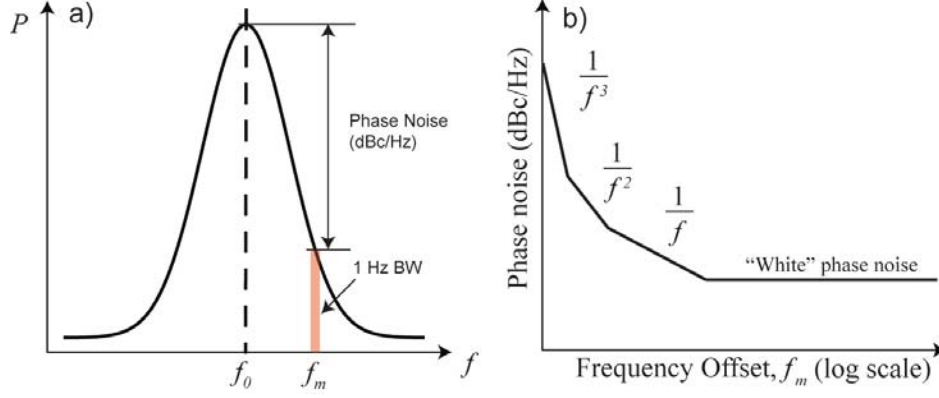


Figure 6.10: a) Phase noise definition in the typical spectrum of a non-ideal oscillator
b) Single-sideband Phase noise spectrum plotted with offset frequency in log-scale

of a non-ideal oscillator, as shown in Fig. 6.10(a), the phase noise is defined as the ratio of the noise in a 1-Hz bandwidth at a specified frequency offset, f_m , to the oscillator signal amplitude at carrier frequency, f_0 . It is usually customary to characterize the phase noise of an oscillator in its single-sideband phase noise spectrum. In Fig. 6.10(b), the phase noise (in the units of dBc/Hz) is plotted as a function of the offset frequency, f_m , with the axis on the log scale. A typical phase noise spectrum can be categorized by several different regions, which have different slopes in log-log plot ($\frac{1}{f^x}$, $x = 0, 1, 2, 3$), as shown in Fig. 6.10(b). These regions represent different types of the phase noise, which we will discuss in detail in Sec. 6.6.

Most sources of noise, such as shot noise and thermal noise, are white. They have spectral density essentially flat for all frequencies. According to Eqn. 6.3.1, white noise has infinite power. In practice, though, there is always some filtering function limiting the bandwidth and therefore the power of noise. In practice, it is especially useful to place a prefilter before the PLL to reduce the noise as much as possible without affecting the signal level.

In a PLL, the phase noise is most significantly contributed by several processes: 1) input noise from the reference; 2) phase noise of the voltage-controlled oscillator; 3) noise generated by loop filters and phase comparators. In the PLL we developed here,

the pulse train from the mode-locked laser serves as the reference. The noise of the pulse train from a mode-locked laser is found to be less than a femtosecond [200], which can be neglected in our discussion. However, more severe reference noise is found to be caused by amplitude-to-phase conversion in the photodetection process [204], and due to the fluctuation of laser pointing and optical power. The amplitude-to-phase conversion factor was measured to be 1 ps/mW for a 12 GHz InGaAs detector (with 6 V reverse bias) [205]. Fig. 6.11(a) shows the RF power fluctuation of laser oscillator, MaiTai (from Spectra Physics). We monitored the power fluctuation for more than 6 hours and found that the power fluctuation is $\sim 1.5\%$, which gives us 0.03 mW power fluctuation over 2 mW total optical power, thus, ~ 30 fs phase jitter. This value varies with the environmental conditions, such as temperature, humidity, etc. To achieve < 10 fs phase jitter, all optical synchronization scheme will need to be implemented [206, 207], which is out of the scale of our discussion here.

Other electrical components in the PLL can all cause additional noise, including voltage-controlled oscillators, phase comparators, amplifiers, power supplies, etc. Thus, carefully selecting the components with the right specifications and keeping them working in an optimum condition is a very important (but tedious) job. Unfortunately, most of these have to rely on the experience on electrical engineering and are very difficult to model and quantify. In addition, due to the vast amount of available RF components on the market, it is definitely impossible for us to try all the possibilities. So, what we can do here is to present the solution we have for the RF-compression UEM, while further improvements will be possible, if a better performance is desired.

6.3.2 Shift of Static Phase Error

In Sec. 6.2.1, the concept of static phase error was introduced. The static phase error is caused by the necessary phase difference between the VCO output and the reference signals for the PD to supply a correct control voltage to the VCO and sustain its

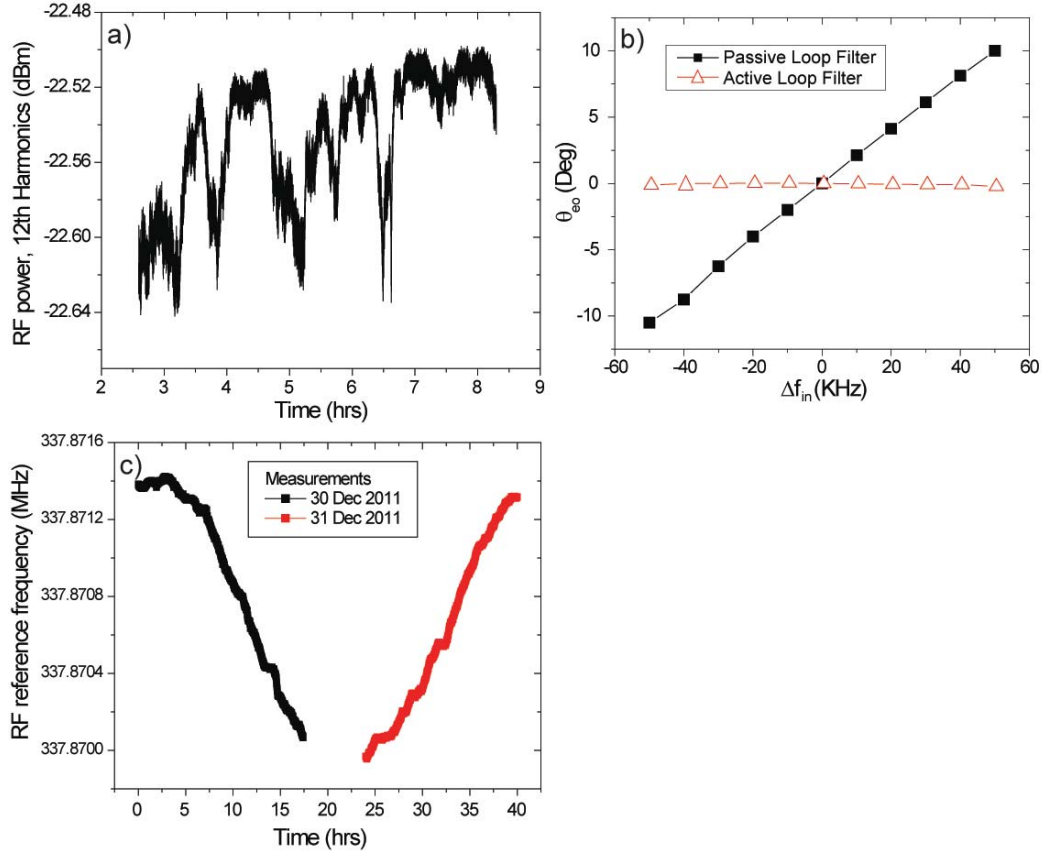


Figure 6.11: a) Power fluctuation of the 12th harmonic from the laser oscillator; b) Shift of static-phase-error caused by the shift of input reference frequency. An RF signal generator was used in this measurement to replace laser oscillator, which gives more flexible control of reference frequency. Black symbols represent the results with the passive loop filter and red symbols with the active filter. Design of active filter is shown in Fig. 6.15. c) Two period monitoring of the reference frequency drift from laser oscillator.

output frequency identical to the reference frequency. In an ideal situation, the static phase shift is a constant value determined by the zero-phase voltage of the PD V_{do} , in-lock control voltage of the VCO V_{co} and the PD gain K_d . However, we find that the static phase error shifts, when the reference frequency from the laser oscillator changes, which is a special situation in our application. Under the framework of the linear model, where the signal flow is described in Fig. 6.8(a) and the loop filter is taken into consideration, we have the following relation, when the PLL is in equilibrium ($\omega_o = \omega_i$),

$$\begin{aligned}\omega_o &= K_o V_{co} \\ &= K_o K_h (K_d \theta_{eo} + V_{do}) \\ &= \omega_i.\end{aligned}\tag{6.3.2}$$

Eqn. 6.3.2 gives the static phase error by

$$\theta_{eo} = \frac{\omega_i - K_o K_h V_{do}}{K_o K_h K_d}.\tag{6.3.3}$$

So, it is found that the static phase error θ_{eo} changes with the input reference frequency ω_i , and the sensitivity can be defined by the derivative:

$$\frac{d\theta_{eo}}{d\omega_i} = \frac{1}{K_o K_h K_d}.\tag{6.3.4}$$

The result of a direct measurement of the static phase shift as a function of the input reference frequency is plotted in Fig. 6.11. In this measurement, a low noise signal generator was used to supply the reference signal, which gives us more flexibility in controlling the reference frequency. The sensitivity of static phase error to the input reference frequency can be determined by the slope in Fig. 6.11, which is ~ 0.19 Deg/kHz, when a passive loop filter with unit dc gain is applied. Considering the PD gain $K_d = 340 \text{ mV}/\pi$, VCO gain $K_o = 4 \text{ MHz}/\text{V}$, which are determined by other independent calibrations, we would expect the slope to be $0.13 \text{ Deg}/\text{kHz}$, which is in

good agreement with the experimental results. In a long-term operation, the reference-frequency drift is expected to be ~ 2 kHz, as shown by the measurement in Fig. 6.11(c), which can cause as much as ~ 0.38 Deg phase shift and ~ 6.2 ps timing jitter, if we consider 338 MHz as the reference frequency of the PLL. Obviously, this value exceeds the tolerance on the timing jitter for the RF compression, which needs to be reduced. Fortunately, the solution to this problem is already implied by Eqn. 6.3.4. We can increase the PD gain, the VCO gain or the loop filter dc gain to reduce the susceptibility of the static phase error to the reference-frequency drift. Again, as we have discussed above, to increase the loop filter dc gain is the easiest way among all of the solutions. In practice, an operational-amplifier (op-amp) circuit can be implemented to push the loop-filter dc gain to ~ 100 . This loop filter design is called an *active loop filter*.

6.4 Loop Filter Design

An op-amp circuit as shown in Fig. 6.12(a) serves as a voltage amplifier and can provide enough dc gain to reduce the susceptibility of the static phase error to the reference-frequency drift. The dc gain is determined by $\frac{R_3}{R_1+R_2} \approx 100$. Before and after the op-amp, the resistor and capacitor network, R_1 and C_1 , provides the pre- and post-low-pass filtering of the signal to remove the undesired high-frequency noise. The overall bandwidth of the loop filter is determined by the values of R_1 , C_1 , R_4 and C_2 . The frequency response of the loop filter, which is simulated by QUCS, is plotted in Fig. 6.12(b). With dc input (0 Hz), the loop filter gain is ~ 100 (20 dB). With higher frequency input, the loop filter gain rolls down very fast and 3 dB bandwidth is about 25 Hz. At ~ 1400 Hz, the loop filter gain decays to unit (0 dB). However, because the control voltage of VCO is very low frequency (almost dc) when the PLL is in-lock, the gain of interest here is only its dc gain.

Everything seems simple and straightforward to this point. However, if you place

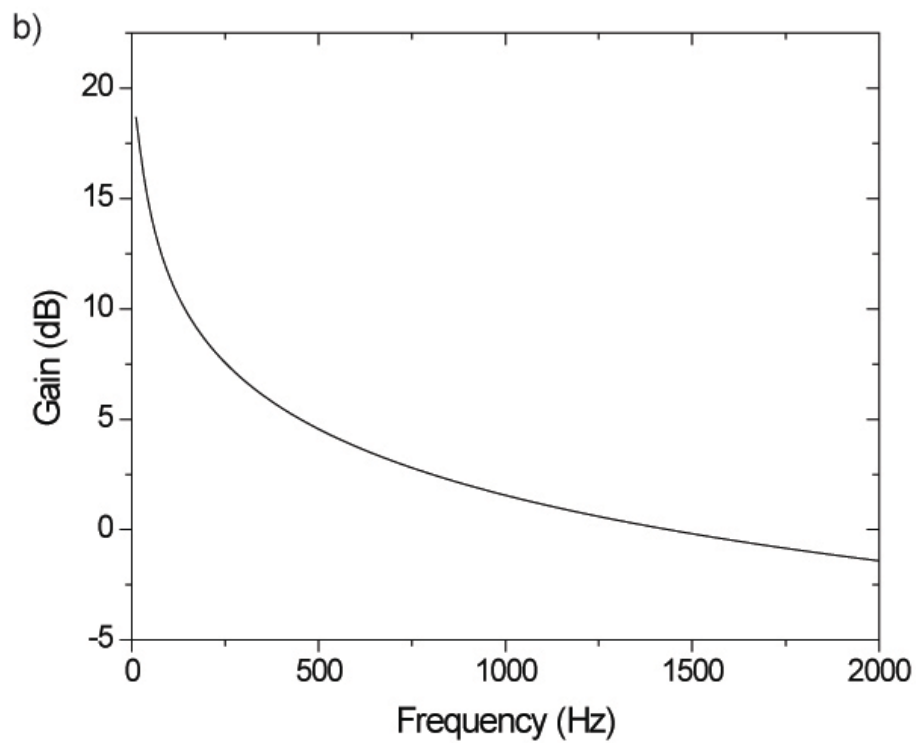
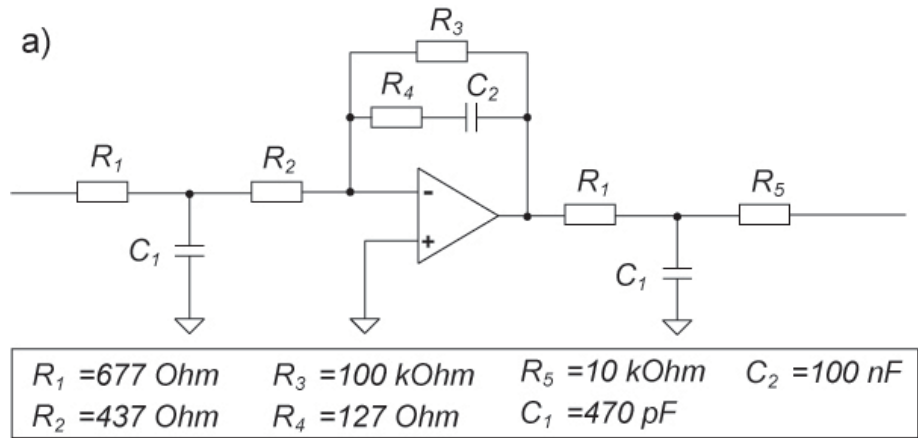


Figure 6.12: a) Active loop filter design with ~ 100 dc gain for the RF-compression UEM. b) Frequency response of the loop filter simulated using QUCS.

and phase difference between the open-loop response and input perturbation is then measured using the oscilloscope. The measured bode plots are plotted in Fig. 6.14(b). The results for the PLL with the original passive loop filter design are shown as the black curves. With increase of the modulation frequency, the open-loop gain rolls down and reaches unity at ~ 10 MHz. At the same time, the phase at the point B shifts relatively to the phase of modulation at the point A , as shown in the phase plot. Intuitively, if the phase is shifted more than -180° , before the gain rolls down below unity (0 dB), the signal fed back will enhance itself and ultimately cause oscillation [208]. Thus, two important crossing points can be defined in the bode plots. The *gain crossing point* ω_g is the frequency, where the gain crosses 0 dB, and the *phase crossing point*, ω_{ph} , is the frequency, where the phase crosses -180° phase difference. Then, the stability of a PLL can be quantified by the following two quantities:

1. Gain margin: Gain difference between the gain at phase crossing point to 0 dB.
2. Phase margin: Phase difference between the phase at gain crossing point to -180°

For a negative feedback control system, the system is expected to be stable, if these two margins are large. Extracted from Fig. 6.14(b), the phase margin Φ_1 with the original passive loop filter design is found to be $\sim 52^\circ$, which is enough for a stable PLL. However, the gain margin is no more than 15 dB, which is probably not sufficient. To improve the stability of the PLL, a lag-compensator, as shown in Fig. 6.14(a), is implemented and the bode plots of compensated system are shown as the red curves in Fig. 6.14(b). After compensation, the phase margin Φ_2 is increased to 92° , and, more significantly, the gain margin G_2 is also increased to more than 30 dB. (The phase crossing point does not appear in the measurement frequency range.) So, the lag-compensator successfully improve the stability of the PLL, which makes the implementation of a high-gain active loop filter possible. The overall loop filter design with op-amp circuit and the phase-lag compensator is plotted in Fig. 6.15. The susceptibility of the static

phase error to the reference frequency is measured again, after the active loop filter is implemented. As shown in Fig. 6.11(b), the sensitivity of static phase error to the reference frequency shift becomes very weak after the active loop filter is used in the PLL.

6.5 Realization of PLL

The PLL is constructed using RF components. In practice, it is important to first select the narrow-band-width components, because they give more restrictions on the design. Taking the design of this PLL as an example, the reason we select the 4th harmonic signal from laser oscillator for the phase-locked loop is mainly because the center frequency of the bandpass filter ‘BPF-A332+’, which is commercially available with an inexpensive price, is very close to it. The bandwidth of ‘BPF-A332+’ is very narrow (3 dB BW: 6MHz; 20dB BW: 60MHz), which is very important for rejecting other harmonics. On the other hand, if the narrow bandpass filter is custom-made, it will be much more expensive and take a very long lead time. (An example is the narrow bandpass filter to select 12th harmonics, ‘308-1013.61M-1.5P-A’ from TTE.) Except for the bandpass filter, usually other components, including low-noise amplifiers, phase detectors, power splitters, voltage-controlled oscillators, etc., allow a broad bandwidth.

Fig. 6.16(a) shows the signal-flow diagram in more detail, which shows different RF components. The part number and important specifications of these components are listed in Table 6.1. In the diagram, the reference RF signal is sent into the PLL from the right side. After the power splitter, its 4th harmonic is selected by BPF #2 and amplified by a low-noise amplifier, LNA. Then, another BPF #2 is applied to further clean up the signal and filter out some of the additional noise generated by the LNA. This signal is sent to the ‘LO’ port of the Mixer #2. Here, the Mixer #2 works as the phase detector in the PLL setup. This half of the PLL setup is called the

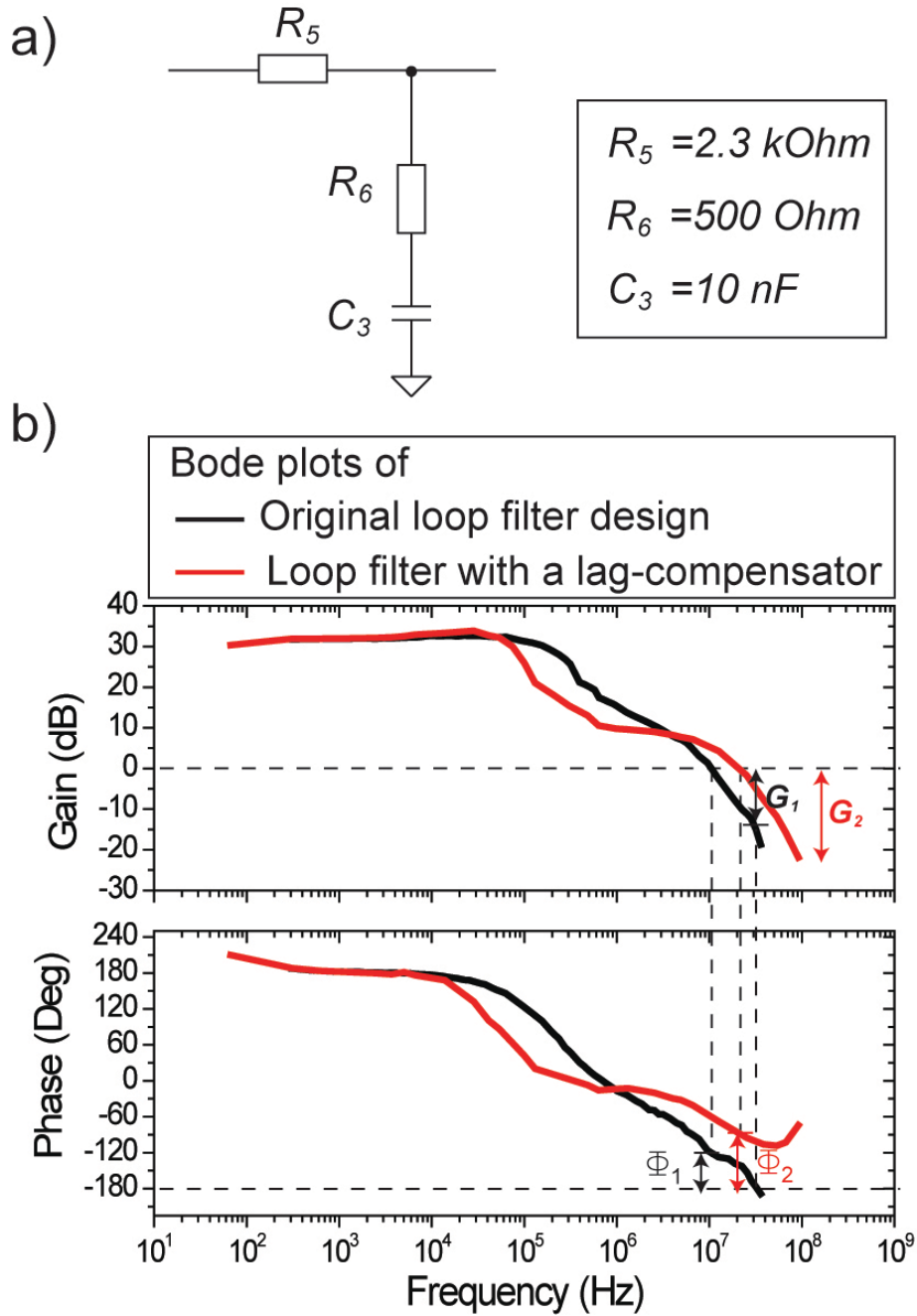


Figure 6.14: a) Design of the phase-lag compensator b) Bode plots for the PLL with the original loop filter design and a loop filter with a phase-lag compensator

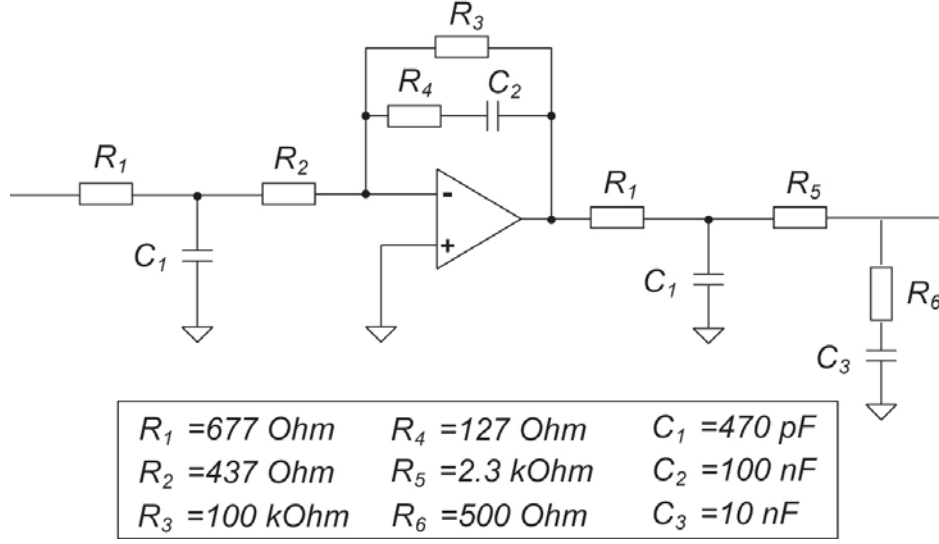


Figure 6.15: Overall loop filter design with ~ 100 dc gain and lag-compensator to improve its stability

Reference Arm. In the left half of the PLL setup, the RF signal from a RF synthesizer (Agilent E8247C) is sent to the Mixer #1 and the sum-frequency signal is generated with the VCO signal. The 1 GHz narrow bandpass filter, BPF #1, is used to select the sum-frequency signal and filter out other harmonics. The RF synthesizer in the setup is used to compensate the frequency difference between the VCO and the 12th harmonic of laser. After the BPF #1, the signal is split by a power splitter, in which half of the signal is sent out of the PLL as the PLL output. The other half goes through a low noise divider, whose ratio is set to be 3. Thus, the frequency after the divider is very close to the 4th harmonic of the laser oscillator. Because the divider contains an integrated amplifier, a proper attenuation on the RF signal after the divider is required, which is not shown in the diagram. Another BPF-LNA-BPF set is placed after the divider, which is identical to the one in the Reference Arm. With this setup, the thermal drift of BPFs and LNAs in two arms can cancel each other and improve the stability of the PLL. This half of the PLL is called the *Oscillator Arm*. The signal in the Oscillator Arm is ultimately fed into the ‘RF’ port of the Mixer #2. The phase error between signals is measured and fed back to the VCO through a low-pass filter

(LPF) and the loop filter. The bandwidth of the LPF is ~ 1.9 MHz and the schematic of the loop filter is already shown in Fig. 6.15. In Fig. 6.16(b), we show a picture of the real PLL setup with different components marked, which directly corresponds to the parts in Fig. 6.16(a).

6.6 Characterization of PLL

To characterize the performance of the PLL, the phase noise or the timing jitter needs to be measured. There is a vast amount of literature introducing how to do that. A good reference is Ref. [209]. In our experiments, two methods are used to characterize the phase noise:

1. Spectrum Analysis
2. Phase-locked loop method

6.6.1 Spectrum Analysis

In the spectrum analysis, the phase noise spectrum is directly measured using a spectrum analyzer. It is indicated by Eqn. 6.3.1 that the integrated area of the power spectra density is equivalent to the time-integral of the mean-square noise voltage, which is directly connected to the timing jitter in the time domain. The rms timing jitter can be derived using the following equation [210, 211]:

$$\Delta t_{\text{rms}} = \frac{\sqrt{2 \int_0^\infty 10^{\frac{L_c(f)}{10}} df}}{2\pi f_0}, \quad (6.6.1)$$

where $L_c(f)$ is the phase noise in the unit of [dBc/Hz], f_0 the center frequency. From Eqn. 6.6.1, the noise spectrum needs to be integrated in log-scale to calculate the timing jitter. As shown in Fig. 6.10(b), a noise process can usually be categorized into one of the power-law process with different exponents ($1/f^x$, $x = 0, 1, 2, 3$) [209]:

Label	From	Model Number	Center Frequency(MHz)	Bandwidth (MHz)	Gain(dB)	Phase noise (dBc/Hz)
RF Synthesizer	Agilent	E8247C	773.2	–	–	-163
Mixer #1	MiniCircuits	ZFM-4H-S+	–	5~1200	–	–
BPF #1	TTE	308-1013.61M-1.5P-A	1013.61	12	–	–
BPF #2	MiniCircuits	BPF-A332+	332	6	–	–
Power Splitter	MiniCircuits	ZFSC-2-11-S+	–	10 - 2000	–	–
Divider	Valon Technology	3008 Universal Divider	–	5~1600	–	-161
LNA	MiniCircuits	ZX60-P162LN+	–	700 - 1600	22.50	+0.5
Mixer #2	Holworth	HX3100	–	8~1000	–	-171
LPF	MiniCircuits	BLP-1.9+	–	DC~1.9	–	–
VCO	MiniCircuits	ZX95-258+	258	25	–	-160

Table 6.1: List of RF components in the PLL with critical specifications. Some of the specifications, which are not available from manufacturers, are filled with ‘–’. For LNA, usually only the noise figure (NF) is available. ‘+0.5 dBc/Hz’ means additional 0.5 dBc/Hz phase noise is added to the signal after amplification.

1. Random-walk frequency modulation (FM), phase noise $L_c(f)$ goes down with $\frac{1}{f^4}$;
2. Flicker FM, $L_c(f)$ goes down with $\frac{1}{f^3}$;
3. White FM, $L_c(f)$ goes down with $\frac{1}{f^2}$;
4. Flicker phase modulation (PM), $L_c(f)$ goes down with $\frac{1}{f}$;
5. White PM, $L_c(f)$ plot is flat.

The phase noise spectrum can thus be integrated in these regions separately. The phase noise spectra of the PLL output and the reference signal from the laser oscillator are plotted in Fig. 6.17. The dashed lines are the fittings of the noise spectrum with different power-law exponents. The integrated phase noise is calculated and converted to the timing jitter using Eqn. 6.6.1 in every region. The range of integration is limited to 25 kHz in a single sideband, since the bandwidth of the cavity is ~ 50 kHz. The total timing jitter is given by the sum of the jitters in all the regions, which is ~ 470 fs.

In Fig. 6.17, almost identical spectra are observed for both the PLL output and the reference signal, which tells us that the spectrum analysis is limited by several factors in this case. First, at high offset frequency, the measured phase noise is limited at ~ -110 dBc/Hz, which is the phase-noise floor of the spectrum analyzer. Secondly, when the offset frequency is very small, the model of power-law dependence fails. It is due to the limited frequency resolution (~ 100 Hz) of the spectrum analyzer. So, the measured timing jitter serves as an upper limit to the actual value. To measure the timing jitter more precisely, a spectrum analyzer with better frequency resolution and lower phase-noise floor is required.

6.6.2 Phase-locked Loop Method

In addition to the spectrum analysis, the timing jitter can also be characterized using the phase-locked loop (PLL) method [209]. In the PLL method, the oscillator of

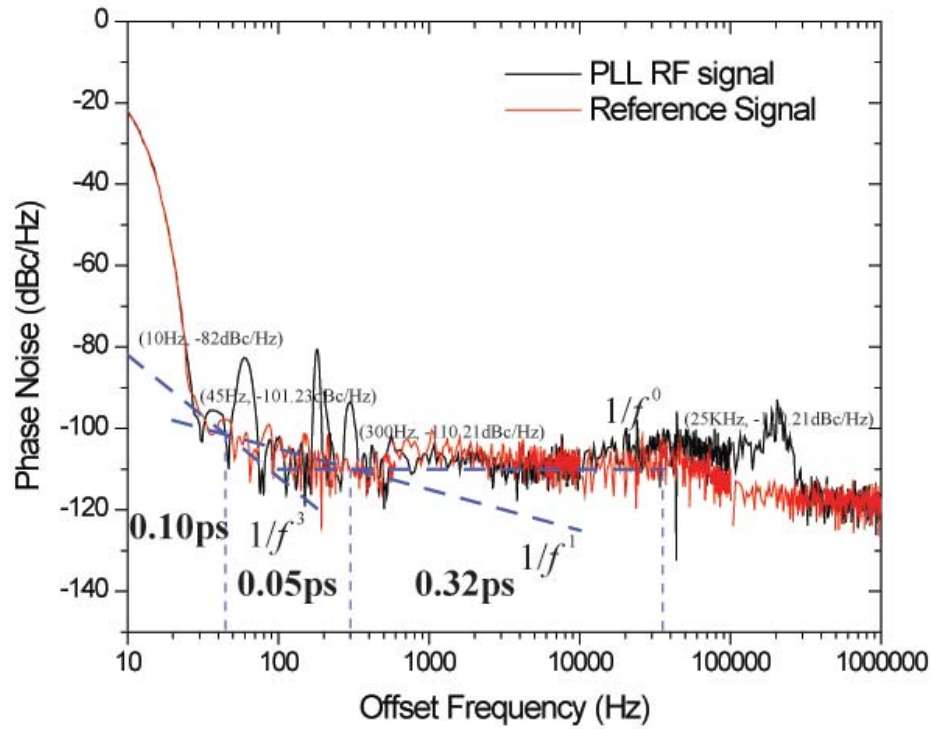


Figure 6.17: Phase noise spectra of the PLL output and reference input measured using spectrum analyzer(FSP13). Dashed lines are the fittings of the noise spectrum using power-law relations with different exponents.

interest is first phase-locked to a low-noise reference oscillator. And the phase (timing) jitter is directly measured by comparing the relative phase between them using a phase detector. It is very convenient for us to implement this method, since we already have a PLL set up. Fig. 6.18(a) presents the schematic diagram of the measurement setup, in which the phase difference is measured outside the loop between the reference and the PLL output. The phase detector we use is AD8302 from Analog Devices. The advantage of AD8302 is its high sensitivity and accurate voltage scaling (10 mV/Deg). The voltage output from the phase detector is then recorded by a multimeter, which represents the phase fluctuations between the two signals. Here, we can see that, using the spectrum analysis, the high-frequency (>100 Hz) phase noise can be characterized, while the PLL method is convenient for characterizing the long-term stability (from seconds to hours), which a single spectrum cannot tell. In principle, the high-frequency phase noise can also be measured using the PLL method, if the voltage sampling rate can be increased to kHz, or even MHz, by implementing a high-speed analog-to-digital convertor (ADC). The measurement bandwidth will eventually be limited by the speed of phase detector and ADC. Additional cares also need be taken in isolating the measurement noise at high frequency. The high-speed measurement using ADC is out of the scope of the discussion here. Here, we will use the PLL method only to investigate the long-term stability of the PLL output.

Fig. 6.18(b) shows a 10-hour monitoring of PLL timing jitter, together with the measurement of the room temperature in Fig. 6.18(c). Even though the spectrum analysis told us that the PLL timing jitter is less than 470 fs at high frequency, there exist ~ 10 ps timing drift over a 10-hour period. Comparing the features in Fig. 6.18(b) and Fig. 6.18(c), it is found that together with the temperature drop at ~ 16 hour and 20 hour, the PLL timing also has a major drift, which indicates a direct correlation between timing accuracy and the temperature of RF components.

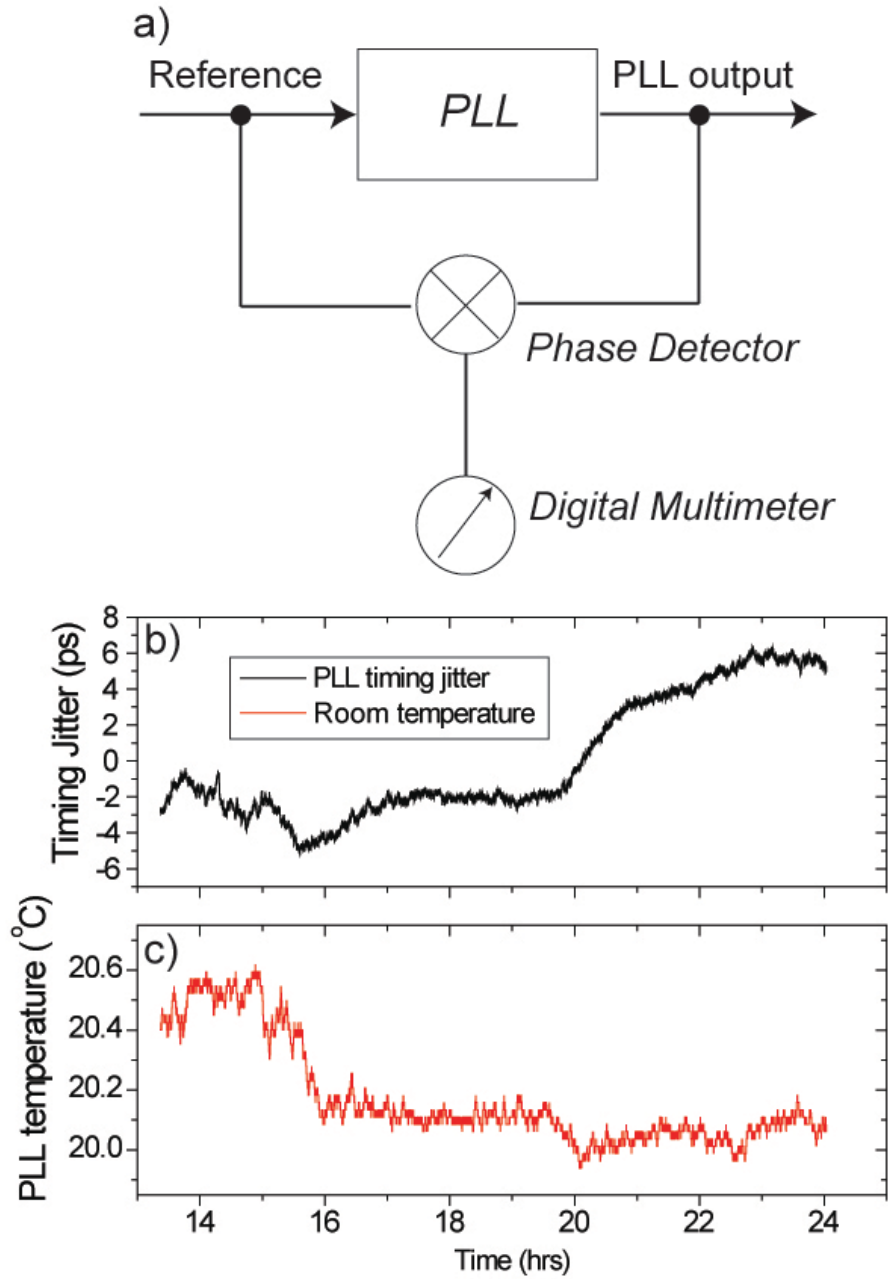


Figure 6.18: a) Experimental setup of timing jitter measurement using the phase-locked loop method. b) Timing jitter of the PLL for ~ 10 hours' period. c) Room-temperature measured in the same period of time as b).

6.7 Temperature-Stabilized PLL

It is not surprising that timing accuracy of the PLL is correlated to the room-temperature drift, since the properties of the most analog components drift with their temperature. For example, the resistance of a simple $1\text{ M}\Omega$ resistor can be changed by $\sim 50\text{ }\Omega$, when its temperature is changed by $\sim 1^\circ\text{C}$. The question is whether this temperature-induced change is significant to the PLL timing accuracy or not. With calibrations, it is found that different components have different temperature sensitivities, in which the VCO is the most sensitive component to the phase drift of the PLL output. The contributions from LNAs, BPFs and PDs are relatively small, and the contributions from power splitters are negligible. Due to the complication of this problem, in practice, we choose to seal the PLL electronics into a metal box and stabilize the temperature of the whole box.

A PID thermoelectric control scheme is implemented to stabilize the temperature of the PLL box. The bottom plate of the box is made with cast aluminum, which has better thermal conductivity compared to normal aluminum. The RF components are mounted on the top surface of the plate with thermal grease in the gap to improve the thermal contact. To actively stabilize the temperature, four thermoelectric modules are mounted on the back of bottom plate, as shown in Fig. 6.19(a). The temperature fluctuation is measured by an RTD temperature sensor mounted very close to the VCO. The sensor's resistance value, which varies with its temperature, is fed back to a PID controller (Thorlab ITC4005), which drives a bipolar current power supply and actively controls the thermoelectric modules to cancel the temperature fluctuations. The schematic diagram of the temperature feedback control setup is shown in Fig. 6.19(a).

In Fig. 6.19(b), we plot the timing jitter measurements for temperature-stabilized and unstabilized PLL. An optimum temperature stability of 10mK has been achieved,

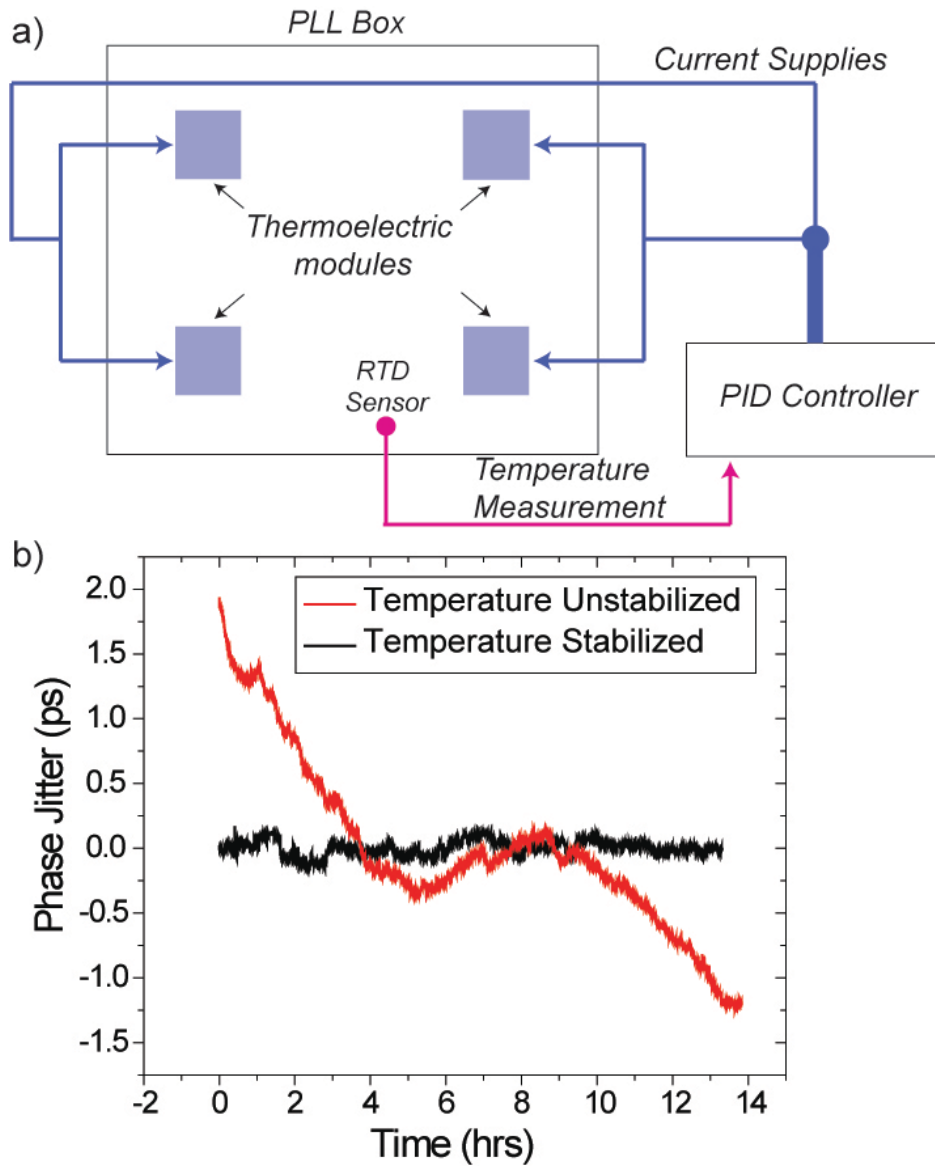


Figure 6.19: a) Schematic diagram of temperature feedback control of PLL electronic box. b) Timing jitter measurements of temperature stabilized (black) and unstabilized (red) PLL with similar room temperature fluctuations.

with the temperature feedback control. A significant improvement of the long-term stability can be observed by comparing the results under two different situations. We need to note that, in this measurement, a stable RF synthesizer is used as the reference, instead of directly locking the PLL to the laser oscillator, because only the effect of temperature drift to PLL timing drift is of interest here. Statistically, a rms timing jitter of ~ 50 fs is obtained with the temperature stabilization. However, when the PLL is directly locked to laser oscillator, the residual timing drift is observed, due to the long-term power and pointing drift of the laser oscillator, which converts to the phase drift by the PLL electronics.

6.8 Conclusion

In this chapter, the low-level phase-locked loop (PLL) for the RF-compression UEM is introduced. The purpose of the PLL is to synchronize the photoelectron pulses with the RF field, which is later delivered to the RF-amplifier station and RF cavity. The overall performance of the PLL electronics is characterized by phase noise and timing jitter measurements. Two different methods are implemented to measure the timing jitter, including the spectrum analysis and the PLL method. The spectrum analysis is used to characterize the high-frequency phase noise, which is found to be less than 470 fs in the bandwidth from 100 Hz to 25 kHz. The PLL method, on the other hand, is used to characterize the long-term stability. The timing of the PLL is found to drift when the environment temperature changes. This slow drift is further corrected by stabilizing the temperature of PLL electronics with a temperature feedback control scheme, and an optimum rms timing jitter of ~ 50 fs has been demonstrated. The ultimate timing jitter is found to be limited by the long-term power and pointing drift of the laser oscillator.

Chapter 7

RF Power Amplification and RF-cavity Phase-lock

In Chapter 6, the synchronization between the phase of the RF field and the electron-pulse generation in the RF-compression UEM by implementing a phase-locked loop has been introduced. The design and performance of the low-power phase-locked loop (PLL) were also described. The PLL electronics can generate an RF power at the level ~ 10 dBm, which is too low to generate any observable changes of the phase space of the electron pulses. To successfully compress the electron pulses, an electric field with strength of several MV/m is required [212]. Thus, a strong amplification of the RF signal after the PLL is necessary.

In this chapter, we are going to deal with the RF signal with high strength. At first, the structure of the RF cavity, which resonantly confines the RF field and executes the pulse compression, will be introduced in Section 7.1 and the requirement on the RF power will be estimated as well. In Section 7.2, the structure of the RF-amplifier station will be described, together with the characterization results of the station. The problems of the phase (timing) jitter and the solutions to it in the low-level PLL have already been extensively discussed in Chapter 6. Here, we are going to face the phase

jitter problem again at the high-power level. To discuss the phase jitter of the cavity field, the resonant condition of the RF cavity and the source of the cavity-field phase jitter will be introduced in Section 7.3. An active compensation scheme for the phase jitter is designed in Section 7.4, and its performance will also be characterized. This chapter will conclude with a description of the overall setup of the RF system for the UEM.

7.1 RF Cavity Design

An RF cavity is a hollow structure that confines the electromagnetic fields in the microwave frequency region. It has extremely low loss at its resonant frequency, due to its very high quality-factor (Q-factor). To construct the RF cavity, copper is usually the preferred material for room temperature application, due to its high electrical and thermal conductivity. The cavity is implemented in the RF-compression UEM to compress the electron pulses in the longitudinal direction, using its TM_{010} mode electric field.

The RF cavity is designed using the Superfish code [213]. The most important parameter to consider is its resonant frequency, which determines the volume of the cavity. For the UEM system, the PLL generates the RF signal, which is phase-locked to the 12th harmonic of the repetition rate of the laser oscillator pulse train. The frequency of the RF field is ~ 1013.4 MHz, which varies a little bit with the environmental fluctuations, as shown in Chapter 6. Fig. 7.1(a) shows the shape and electric field distribution in the 1 GHz RF cavity, which is calculated using the Superfish code. We choose to use the ‘donut’ shaped cavity, because it gives us $\sim 90\%$ reduction on the power dissipation compared to a ‘pillbox’ cavity, which has simpler structure. The inset shows the electric field distribution of the TM_{010} mode along the cavity axis, which is the electron-beam propagation direction. In Fig. 7.1(b), we show the picture

of the copper cavity, after its machining is finished.

The Superfish code not only calculates the resonant frequency of the cavity with given dimensions, but also provides us with other important parameters, which are briefly summarized as follows:

1) Quality factor, Q

Quality factor, Q , is the ratio of the energy stored in the cavity to the energy dissipated in the wall per RF cycle. At the same time, the Q factor is also related to the center frequency and bandwidth with the following equation:

$$Q = \frac{f_c}{\Delta f}, \quad (7.1.1)$$

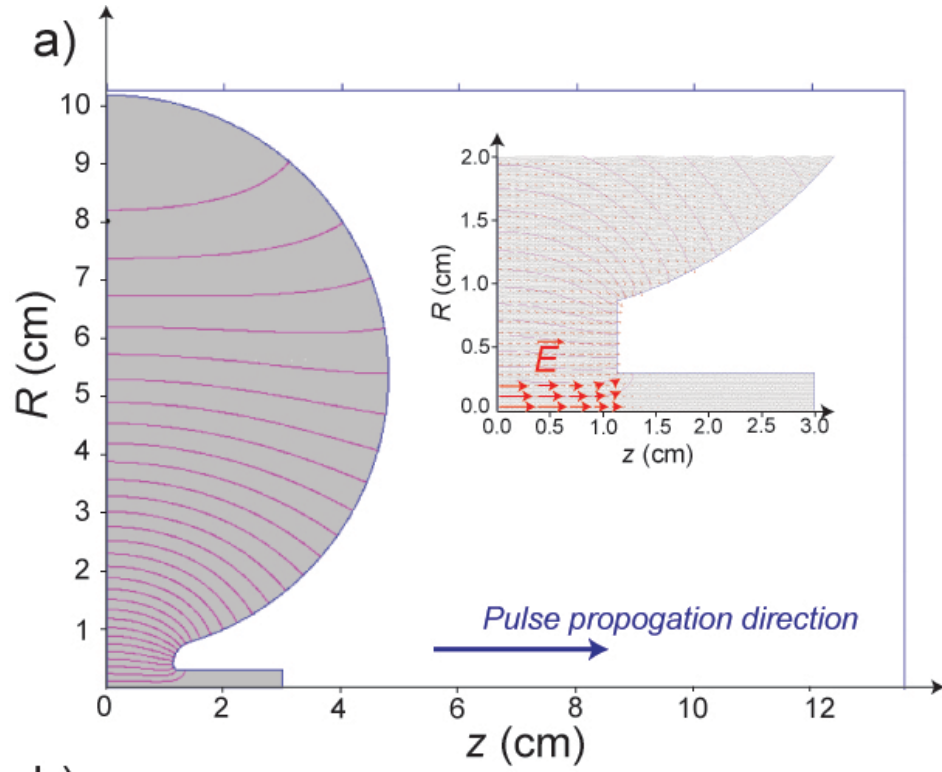
where f_c and Δf represent the center frequency and bandwidth of the RF cavity, respectively. A high Q is usually desirable, because it means low power dissipation and narrower noise bandwidth. For a cavity with a fixed geometry, the Q value is fixed. The Q -factor given by Superfish simulation is ~ 23200 . However, the experimentally measured Q -factor of our RF cavity is ~ 10000 . It is because there is loss on the RF coupler, which is not considered in the Superfish calculation. The structure and position of the coupler can directly affect the energy loss in the cavity.

2) Average compressing field, E_0

The average compressing field, E_0 , is defined as the integral of the z component of the field along the pulse longitudinal direction:

$$E_0 = \frac{1}{d} \int_{-d/2}^{d/2} E_z(R=0) dz, \quad (7.1.2)$$

where d is the length of RF cell, which is 2.25 cm in this case. The field, E_0 , executes the pulse compression, and its strength is related to the RF power stored in the cavity. The field strength inside the cavity can be directly measured by measuring the energy change of the high-energy electrons, after they pass through the cavity. This measure-



b)

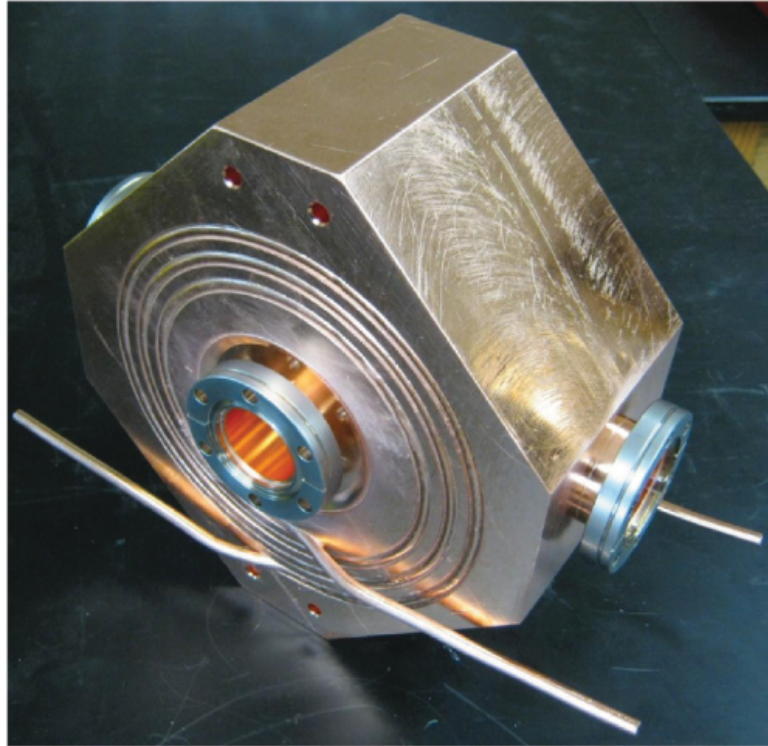


Figure 7.1: a) Simulation result of the RF cavity field distribution with Superfish code, plotted in cylindrical coordinates. Shaded area is the cavity chamber. Solid lines are the equipotential lines. Inset: Zoom-in view of the electric-field distribution of the TM_{010} mode on the axis, along which the electron pulses propagate. b) Picture of the copper RF cavity for the RF-compression UEM.

ment will be introduced in Sec. 8.1.2.

3) Power dissipation, P_{loss}

P_{loss} is the power dissipated in the cavity walls. Since the power dissipation causes cavity heating, this parameter can be used to estimate the power-rating of the chillers required in practice. Using the Superfish code, either the field strength, E_0 , or the power dissipation, P_{loss} , can be used as the input parameter to calculate the other. The relation of P_{loss} to E_0 is fixed for a cavity with a fixed geometry. However, note needs to be taken that enough safety margin needs to be left in practice when designing the cavity cooling system, because the Superfish code usually underestimates the power loss to the cavity walls. The safety margin sometimes can be as much as 50% of the maximum power.

4) Shunt impedance, Z

Shunt impedance Z has dimensions of Ω/m and serves as a figure of merit for the compressing efficiency. The larger the compressing field for a given power loss per unit length, the more efficient the cavity compression is. For a total power, P_T , the shunt impedance is given by

$$Z = \frac{E_0^2 d}{P_T}. \quad (7.1.3)$$

For a cavity with a fixed geometry, the shunt impedance is a fixed value as well. Our cavity has a shunt impedance $\sim 134 \text{ M}\Omega/\text{m}$.

7.2 RF-amplifier Station

7.2.1 Power Requirement

Before designing the RF amplifier station, we need to first estimate how much power is required for the pulse compression. An analytic model can be developed to make this estimation. The first step is to calculate how much the momentum changes for

a single electron with a given compressing field strength, E_0 . Here, we assume all the electrons initially have the same velocity v_c , before the RF cavity, and that the subsequent velocity changes are so small that the resulting changes in the transit time through the RF cavity are negligible. If the electron enters the cavity at time t_1 , its field-induced momentum change is given by the temporal integral of the electric field in the period it travels inside the RF cell

$$\Delta p_1 = - \int_{t_1}^{t_1+d/v_c} eE(t)dt, \quad (7.2.1)$$

where d is the cavity RF cell length. The electric field is an oscillating field with frequency ω , and is given by

$$E(t) = E_0 \sin(\omega t - (\omega d/2v_c) + \phi_0), \quad (7.2.2)$$

where E_0 is the amplitude, ϕ_0 the phase offset and v_c the center of mass (COM) velocity of the pulse. As we have discussed in Chapter 6, perfect phase matching happens when the electron at the pulse center experiences zero integrated field, after it crosses the RF cell. This means when it arrives at the cavity center, the electric field strength $E(t) = 0$. If we assume $t = 0$ is the time when the center of the electron pulse starts entering the cavity RF cell, then the time the pulse center arrives at the center of the cavity is simply $t = \frac{d}{2v_c}$. By substituting this expression into Eqn. 7.2.2 and equating the field strength to zero, we find that the perfect phase matching condition is simply $\phi_0 = 0$. Now we assume the electron pulse longitudinal width is L (half width) and its change is negligible throughout the period the electron pulse passes through the cavity. Thus, for an electron in the front of the bunch, the time it enters the cavity is $t_1 = -L/v_c$, whereas for an electron at the back, $t_2 = L/v_c$. The momentum changes of these two electrons can be calculated separately, using Eqn. 7.2.1. With the assumptions $\omega L/v_c \ll 1$ and $\omega d/2v_c \ll 1$, subtraction of these momenta is leading

to a momentum difference between an electron at the front and an electron at the back of the bunch, which is introduced by cavity field, and it is given by

$$\Delta p = \frac{eE_0\omega\tau d}{v_c} \cos(\phi_0). \quad (7.2.3)$$

Thus,

$$E_0 = \frac{v_c\Delta p}{e\omega\tau d}, \quad (7.2.4)$$

where $\phi_0 = 0$ and $\tau = 2L/v_c$ is the pulse duration. Eqn. 7.2.4 can be used to calculate the required electric field strength, when the momentum change within the electron pulse is known. Now, the question is how much momentum change is needed to longitudinally focus the electron pulse to the sample plane? This value is very difficult to estimate, because it requires the knowledge of space charge effects on the pulse width and momentum spread. Using the AGM simulation, it can be estimated that the longitudinal momentum spread before the cavity is $\delta p_z \approx 12.7 \times 10^{-24}$ kg m/s. Roughly speaking, the required momentum change is double this value, $\Delta p \approx 2\delta p_z$ [212]. Taking account of the pulse width $L \approx 1$ mm for 10^6 electrons per pulse, the required field strength can thus be estimated to be ~ 2.8 MV/m, which corresponds to ~ 500 W RF power according to the Superfish calculation.

We need to note that the estimation above is rough, because the electron optics are simply ignored, such as magnetic lenses, apertures. At the same time, it also ignores the space charge effect after the cavity. To overcome the additional pulse broadening in the free-drift region after the cavity, a stronger compressing field might be required. At the same time, the required RF power also varies with the number of electrons in the pulse. With larger number of electrons, obviously, more momentum change in the phase space will be needed from the cavity field, but, fortunately, the elongated pulse duration, τ , cancels some of this effect and the estimated electric field is found to be similar to the estimation above. At last, as have mentioned, the RF power is usually

underestimated from the Superfish code. For these reasons, enough margin for the RF power has to be made, when we are designing the RF-amplifier station. For a more precise estimation, multi-particle simulations, which take the electron optics and space charge effects into account, are needed.

7.2.2 RF Amplifiers

In the laboratory, an RF amplifier station with up to 2500 W (64 dBm) power delivery capability is built. From the low-level PLL, the RF power we can get is ~ -14 dBm, so the overall gain of the RF-amplifier station is $\sim +80$ dB, which is 10^8 times amplification in power!

In accelerator physics, the RF amplifiers employ either vacuum-tube or solid-state technology. The vacuum-tube-based amplifiers are still widely used in modern accelerators, including tetrodes, klystrons and magnetrons, to generate RF power up to 1 MW continuous wave (c.w.) and 150 MW pulsed mode [41]. In a vacuum tube, electrons are boiled off from a heated cathode, then pass through vacuum and many grids. Ultimately, they are absorbed by the anode plate and convert a small AC signal to a larger AC voltage. On the other hand, the solid-state amplifiers are based on the solid-state RF power transistors operating at voltages from tens to hundreds of volts. Since the electron mobility is much less in semiconductors than in vacuum, the device sizes are much smaller. Usually, a single solid-state transistor can generate power on the order of hundreds of watts continuous wave and up to 1 kW pulsed mode. Because of the relatively low power capability, a large number of solid-state transistors need to be operated in parallel to reach a power at the kilowatts level, where appreciable power is dissipated in the combining and provision circuits. However, compared to the vacuum-tube amplifiers, the solid-state amplifiers are claimed to be more reliable and stable by many groups [41]. They require less maintenance, less warm-up time and lower operation voltage. The supply voltage is usually tens of volts, which is low

enough to avoid the high-voltage and X-ray safety issues arising with the vacuum-tube amplifiers.

For the UEM system, because the power level perfectly fits into the range of the solid-state amplifiers, the RF-amplifier station is built with the solid-state technology. Our solid state RF amplifier system is resulted from a collaborator with Dr. Ti Ruan at SOLEIL. Fig. 7.2(a) shows the circuit layout of a solid-state amplifier. The key components are the two power transistors (Part NO. MRF8S9260H) and one drop-in circulator. The power transistors amplify the input RF signal. The drop-in circulator deflects the reflected RF power to a power resistor and protects the transistors from possible damage caused by the reflected power. Fig. 7.2(b) shows a typical gain measurement of a single solid-state amplifier. The blue solid line is the line for +16.2 dB gain, which fits the data at low input level. The red dashed line measures the 1 dB compression point of this amplifier. The 1 dB compression point is a measure of the output power level, at which the gain drops by 1 dB from its small signal value. It is an important specification of an amplifier, which indicates the limit of its capability for power amplification. From the results, 1 dB compression point is found to be ~ 55 dBm (~ 320 W) for a single solid-state amplifier.

7.2.3 Structure of the RF-amplifier Station

Since one solid-state amplifier can only supply power up to ~ 320 W, eight amplifiers in parallel will be needed to deliver power up to 2500 W. In Fig. 7.3, the schematic diagram for RF amplifier station is depicted. The gain of 8 amplifiers in parallel is the same of the single amplifier, +16 dB, which means the rest of gain, +64 dB, needs to be provided by a pre-amplifier set. As shown in Fig. 7.3, the power level before the 8-way splitter is only ~ 60 W, however, it provides most of the gain in the amplifier station (+66 dB). In the 8-way power splitter, the RF power is equally split into 8 paths, with the power of each path ~ 16 W (39 dBm). Then the 8 ways of RF power is amplified

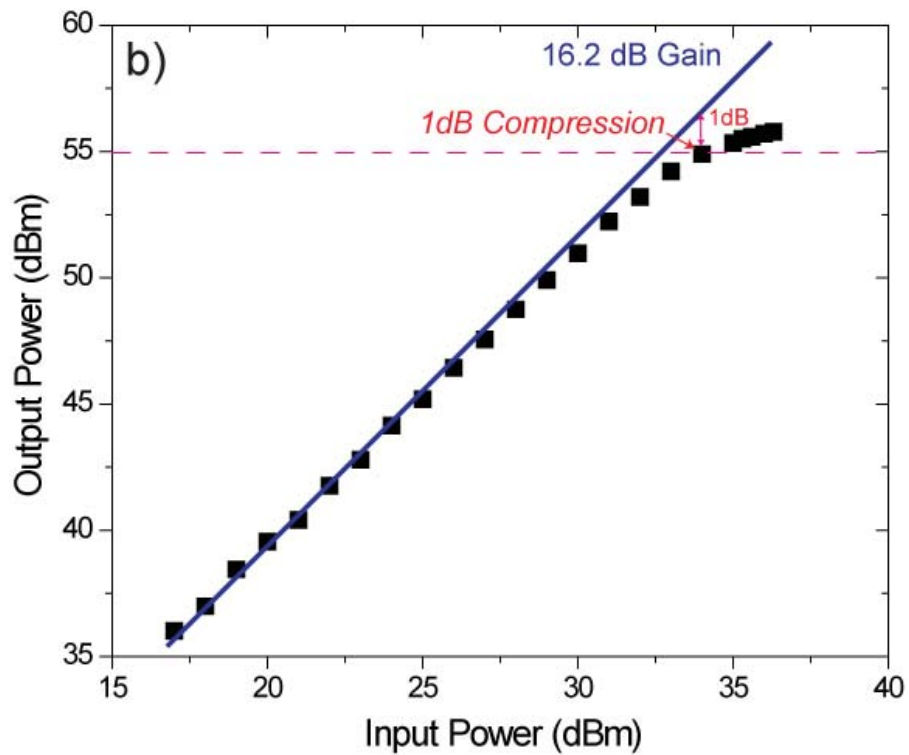
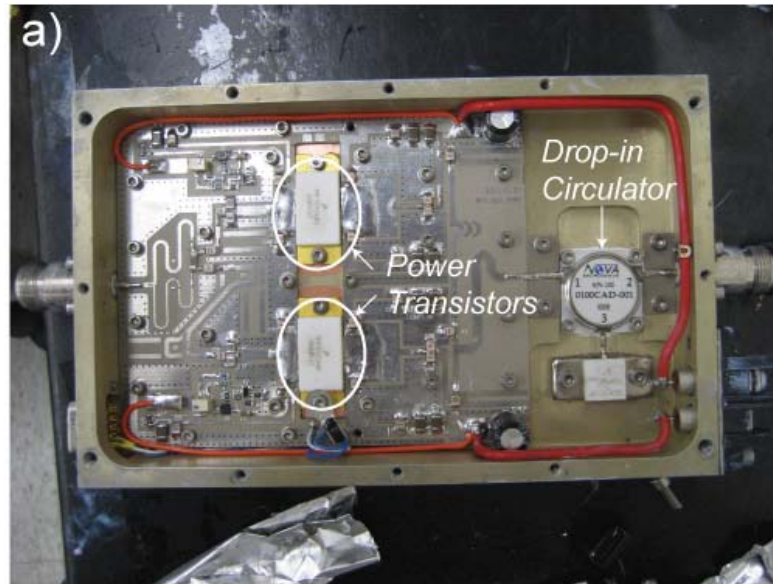


Figure 7.2: a) Circuit layout of a solid-state amplifier b) Typical gain measurement for the solid-state amplifier. Blue solid line is the fitting with 16.2 dB gain. Red dashed line shows the 1 dB compression power.

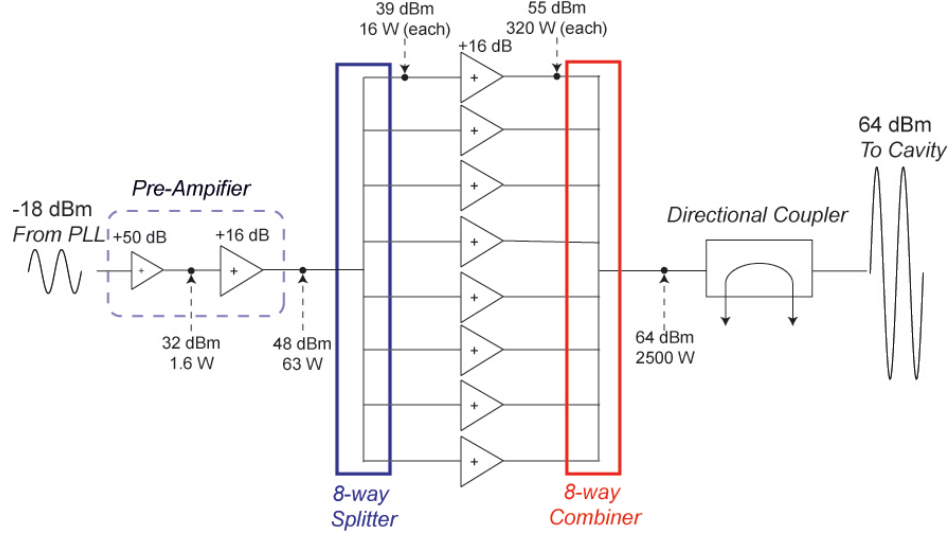


Figure 7.3: Schematic diagram of RF amplifier station. The power level is marked for the situation, when the amplifier station is running at its maximum power output.

by the single solid-state amplifier with +16 dB gain in each path. After amplification, ~ 320 W (55 dBm) power is supplied in each path before the 8-way power combiner. In the 8-way combiner, the high-power RF signal is combined and delivered to the RF cavity, through a directional coupler. The purpose of the directional coupler is to help monitoring the transmitted and the reflected RF power in the transmission line. We need to note that the power level shown in Fig. 7.3 is when the amplifier station is running at its maximum power output. To vary the field strength in the cavity, the power level of the input needs to be adjusted, while the RF-amplifier station only provides a constant gain. To achieve that, a voltage-controlled attenuator (VCA) needs to be installed before the amplifier station to attenuate the input power to a proper level.

Because the drain efficiency of the power transistors in the solid-state amplifier is $\sim 40\%$, each solid-state amplifier requires a dc power supply, which supplies more than double of its output RF power. At the same time, the low efficiency also means there is a large amount of power (as much as the output power) dissipated in the transistors as heat. With so much heat deposited in the amplifiers, the power transistors will be

burned, unless enough cooling is supplied. In practice, all the solid-state amplifiers are mounted on cold plates with thermal grease to improve the thermal conductivity. A chiller with a cooling capacity over 5 kW, and a flow rate ~ 5 GPM is used to carry away the heat generated by the RF amplifiers. In normal operation, the temperature of the power transistors is found to be ~ 50 °C.

7.2.4 Characterization of RF-amplifier Station

Fig. 7.4 shows a direct measurement of the output power from the amplifier station as a function of the input power. The output power is measured through the forward-coupling port of the directional coupler, which was well calibrated before the measurements. The first measurement (on 04Aug2012, black squares) was operated in a continuous mode. The amplified RF signal was terminated by a high-power 50-ohm terminator. The power test was operated up to ~ 1250 W (61 dBm), which was limited by the heating of the terminator. According to the linear fitting of data, the gain of the RF-amplifier station is 82.6 dB, which is consistent with the specifications of the amplifiers and our estimation. In the second measurement (on 03Jan2013, red triangles), the RF power was directly delivered to the RF cavity in a pulsed mode to prevent the cavity coupler from overheating. The test runs all the way to 3100 W (peak power) and the 1 dB compression point is ~ 1300 W (61.4 dBm). We note that the result indicates the 1 dB compression for each solid-state amplifier happens at ~ 162.5 W, which is much lower compared to the previous measured power-limit of the single amplifiers, the reason for which is still not clear. One possibility is that the 1 dB compression at 61.4 dBm is not caused by the 8 amplifiers in each path, but caused by the power limit of the pre-amplifier set. Another reason might be because of the power loss throughout the power splitting and combining processes, which can also limit the maximum power output from the station. In addition, in the pulsed mode, the gain of the amplifier station approaches more than 84 dB, which is higher than expected. We believe that

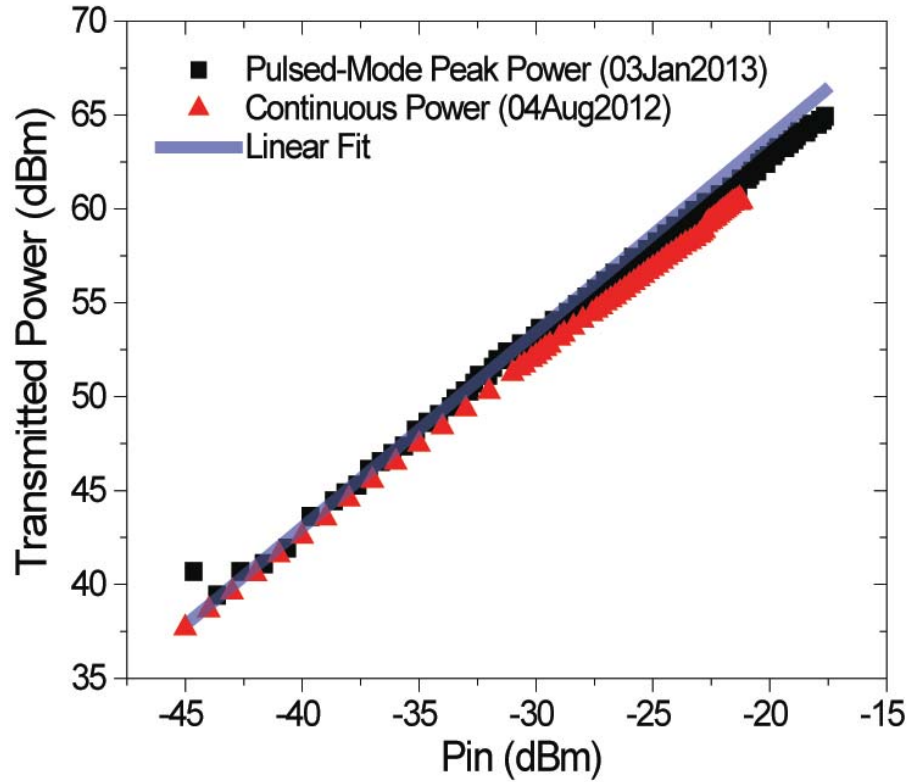


Figure 7.4: Output power measurement for RF amplifier station in 1) Pulsed mode (black square) and 2) Continuous mode (red triangle). The solid line is a linear fit on the low-power data.

it might be caused by the enhanced efficiency of the power transistors, when they are working at a lower temperature due to the pulsed-mode operation. From the results in Fig. 7.4, it is found that output power up to 3 kW (pulsed mode, peak power) can be available. However, special cares need to be taken when operating the amplifier station above its 1 dB compression point, because it means running one or more amplifiers to their design limits, which might cause damage in the long term.

Except for the directional coupler, the RF power inside the cavity can also be directly measured through the pick-up coil. A pick-up coil is usually a circuit loop, which senses the magnetic component of the RF field inside the cavity and converts it to a voltage signal. So, through the pick-up coil, not only the RF power, but also the phase of the RF field inside the cavity can be directly monitored,, which will be

very useful in the following sections. The pick-up coil was pre-calibrated when the RF cavity was assembled in the factory. The measured power on the pick-up coil is 0.1% of the absorbed power in the cavity.

7.3 Cavity Phase for the Resonant Condition

7.3.1 Equivalent Circuit Model

An RF cavity is a high-Q RF component, which means it has a very restrictive resonant condition. The cavity oscillation mode is usually very conveniently described in an equivalent RLC circuit [214], as shown in Fig. 7.5. The electric field of the TM_{010} mode is contained in between the two metal plates of the cavity, which is identical to the capacitor C in the equivalent circuit. The inductor L represents the magnetic field generated by the surface current that travels around outside the cavity. Finally, in a normal cavity (not superconducting), the surface current flows in a skin depth and causes ohmic heating, which is represented by the resistor R in the circuit. Note that the cavity voltage is applied across the resistor, so the definition of R is different from the surface resistor, R_{surf} . It plays a role as the shunt resistor. In the equivalent circuit model, the driving current $I(t)$, which is supplied by a current generator, produces a voltage $V(t)$ across the circuit. This voltage is identified as the axial voltage in the cavity. Both $I(t)$ and $V(t)$ have the same frequency, ω . The total current in the circuit is simply the sum of the current passing through each component

$$I(t) = C\dot{V} + \frac{\int V dt}{L} + \frac{V}{R}, \quad (7.3.1)$$

and by differentiation, we have

$$\frac{\dot{I}}{C} = \frac{\omega_0 R \dot{I}}{Q} = \ddot{V} + \frac{\omega_0 \dot{V}}{Q} + \omega_0^2 V, \quad (7.3.2)$$

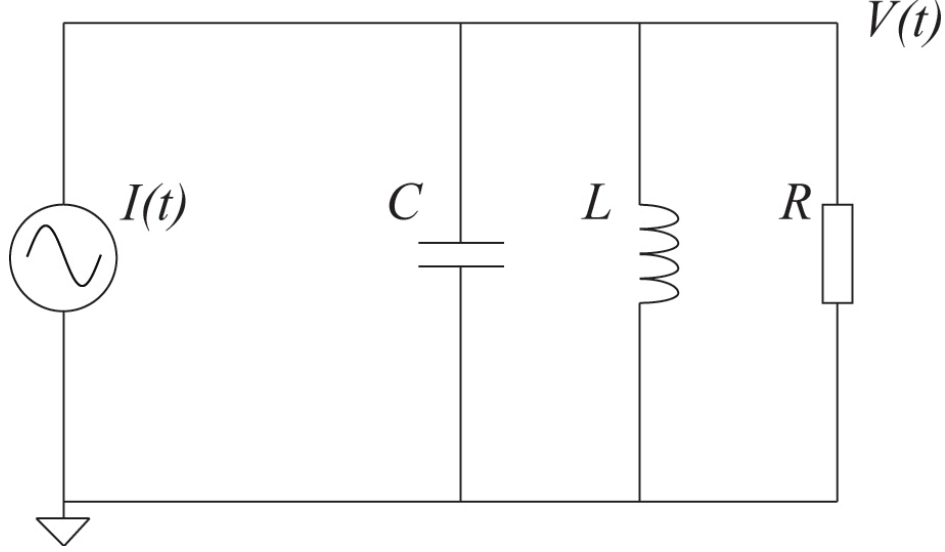


Figure 7.5: Equivalent circuit model of RF cavity

where $\omega_0 = \frac{1}{\sqrt{LC}}$ is the cavity resonant frequency (CRF) and $Q = \omega_0 RC$ is the Q-factor. The stored energy in the cavity is $U = \frac{CV_0^2}{2}$, while the average dissipated power is $P = \frac{V_0^2}{2R}$, where V_0 is the peak voltage across the circuit. So, the Q-factor can be expressed in another form, $Q = \omega_0 \frac{U}{P}$, which is exactly the same as the definition of the Q-factor in Section 7.1.

From Eqn. 7.3.2, a particular steady-state solution of $V(t)$ is $V(t) = V_0 e^{j(\omega t + \phi)}$, if we assume $I(t) = I_0 e^{j\omega t}$, where V_0 and I_0 are the amplitudes and ϕ is the phase of the resonator voltage relative to the driving current. By substituting this solution into Eqn. 7.3.2 and matching the real and imaginary parts on both sides of the equation, we can get the following steady-state solution

$$V(t) = \frac{RI_0 e^{j(\omega t + \phi)}}{\sqrt{1 + y^2}}, \quad (7.3.3)$$

where

$$\phi = -\tan^{-1} y. \quad (7.3.4)$$

y is called the detuning factor, which is defined by

$$y \equiv Q\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right). \quad (7.3.5)$$

7.3.2 Effects of Frequency Detuning

For the cavity in the experiment, the resonant frequency f_0 is ~ 1013 MHz ($f_0 = \omega_0/(2\pi)$) and the Q-factor is ~ 10130 . Substituting the parameters in Eqn. 7.3.3, we plot the power response and phase response as a function of the cavity-frequency detuning Δf , as shown in Fig. 7.6. The detuning frequency Δf is defined as the difference between the input frequency and the CRF, $\Delta f = f - f_0$. In the figure, the power is defined as $P = V^2/R$ and normalized to the power on resonance. With driving source frequency detuned from the CRF f_0 , two effects can be observed on the electric field inside the cavity. First, the power of the field rolls down quickly from the peak power on resonance, as shown in Fig. 7.6(a), which means a decrease of the available field strength for pulse compression. The change of the field strength can directly shift the focal point of the RF compressor, which degrades the quality of pulse compression. The FWHM (3 dB bandwidth) Δf_{BW} is ~ 100 kHz, which is simply given by f_0/Q . The second effect is a shift of the relative phase between cavity electric field and the driving source, as shown in Fig. 7.6(b). Because the input RF signal is tightly phase-locked to the laser pulse train through the low-level PLL, the change of phase $\Delta\phi$ indicates the relative shift of the cavity electric field phase to the arrivals of the electron pulses. This effect can cause energy gain or loss in the electron pulse, and hence an arriving time jitter of the electron pulses, which has been extensively discussed in Chapter 6.

The frequency detuning effects of the RF cavity can also be directly measured using the setup shown in Fig. 7.7(a). The frequency detuning is generated by changing the input frequency of the RF synthesizer (Agilent E8247C), f , relative to the CRF f_0 . The power response is measured using 1) a spectrum analyzer (ROHDE&SCHWARZ

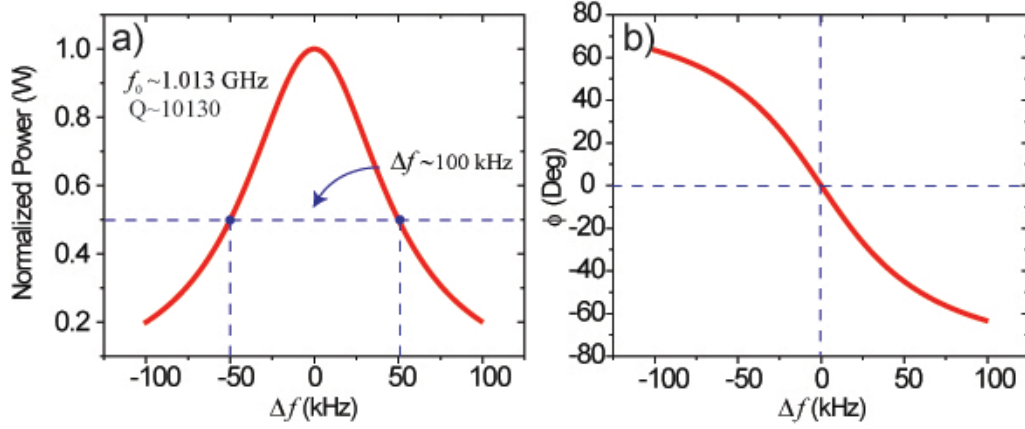


Figure 7.6: Power (a) and relative phase (b) response, when the frequency is detuned from CRF by Δf .

FSP13) through the pick-up coil; 2) an RF power detector measuring the reflected power through the reflection-coupling port of the directional coupler. The relative phase, on the other hand, is measured between the input signal phase (measured through the forward-coupling port of the directional coupler) and the cavity field (measured through the pick-up coil) using the phase detector (AD8302). Fig. 7.7(b) plots a typical result of the power and phase responses to Δf . Note that the vertical axis for power is in the unit of dB, which causes a slightly different line-shape compared to the theoretical calculation in Fig. 7.6(a).

In Fig. 7.7(b), the 3 dB bandwidth of the power measured on the picked-up coil is ~ 100 kHz, which agrees very well with the Q-factor of the cavity. The power inside the cavity (a picked-up coil measurement) increases to the maximum, when the frequency of the input signal is on resonance with the CRF, while the reflected power reduced to its minimum on resonance. The 3 dB bandwidth of reflection power is found to be ≈ 4 times narrower compared to power in the cavity, reason of which is still a puzzle. One possible explanation can be the nonlinear response of the reflection-coupling port of the directional coupler. In practice, both power measurements can be used as the reference for searching for the resonance, but the reflection-power measurement has better sensitivity. The phase shift induced by the input-frequency detuning is also

observed in the experiment, as shown by the blue symbols in Fig. 7.7(b). The phase's sensitivity to the offset frequency, $d\phi/df$ is ~ 1.25 Deg/kHz, when the input frequency is very close to the resonance. Under the framework of the equivalent circuit model, the sensitivity can be calculated by taking derivatives of Eqn. 7.3.4 at the frequency $f = f_0$ and the result is given by

$$\frac{d\phi}{df} = \frac{2Q}{f_0}. \quad (7.3.6)$$

The predicted phase sensitivity is ~ 1.15 Deg/kHz, which is in good agreement with the experimentally measured value.

In the RF-compression experiments, the frequency detuning can happen either due to the shift of the reference frequency (f_{Ref}) from the laser oscillator or due to the shift of CRF. The shift of f_{Ref} is usually on the order of several kHz (see Chapter 6), but the shift of the CRF can be as much as hundreds of kHz, if the cavity temperature is not well regulated. The change of the cavity temperature mechanically deforms the RF cavity, which increases or decreases the cavity chamber volume, and hence, changes the CRF. So, a chiller with sufficient cooling capacity is necessary to carry away the heat generated by the RF power from the cavity and keeps the cavity at a stable working temperature. However, residual temperature fluctuations can still persist and affect the relative phase ϕ .

On the other hand, the CRF of the cavity can also be tuned by a mechanical tuner, which deforms the RF cavity by applying a stretching force along the cavity axis. Inset of Fig. 7.8(a) shows the design of an external tuning plate. The results of the CRF measurement shown in Fig. 7.8(a) confirm the shift of the CRF with different tuner angles. The change of CRF as a function of the external tuner angle is plotted in the inset of Fig. 7.8(b). At the same time, the relative phase ϕ is measured using the setup shown in Fig. 7.7(b). Very interestingly, it is found that when the power detuning is plotted as a function of the phase difference ϕ , the results overlap with each other, regardless of the values of the CRF, which is shown in Fig. 7.8(b). The results in Fig.

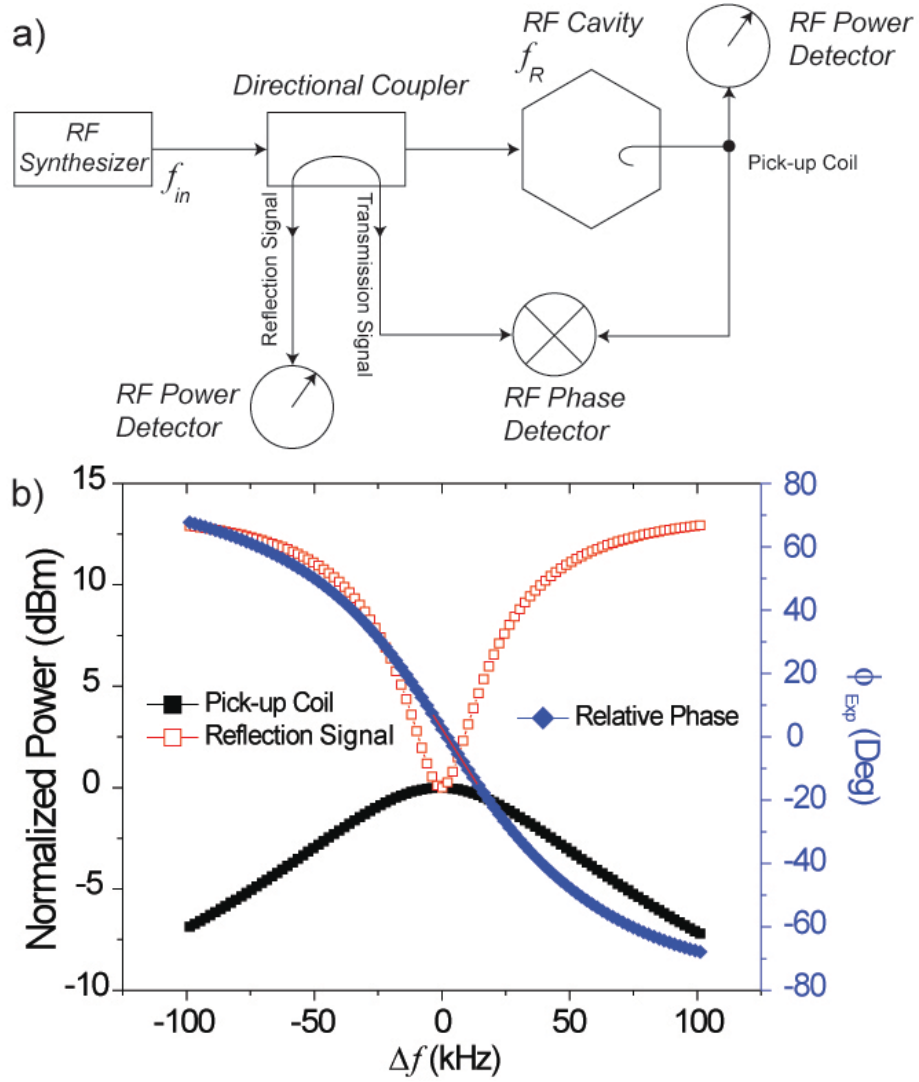


Figure 7.7: a) Experimental setup for measuring the cavity frequency detuning effects b) A typical power and phase response to the frequency detuning, Δf . Cavity power (black) reaches the maximum at the resonant frequency ($\Delta f = 0$), while the reflected power (red) drops to a minimum. The relative phase (blue) shifts with the frequency detuning Δf .

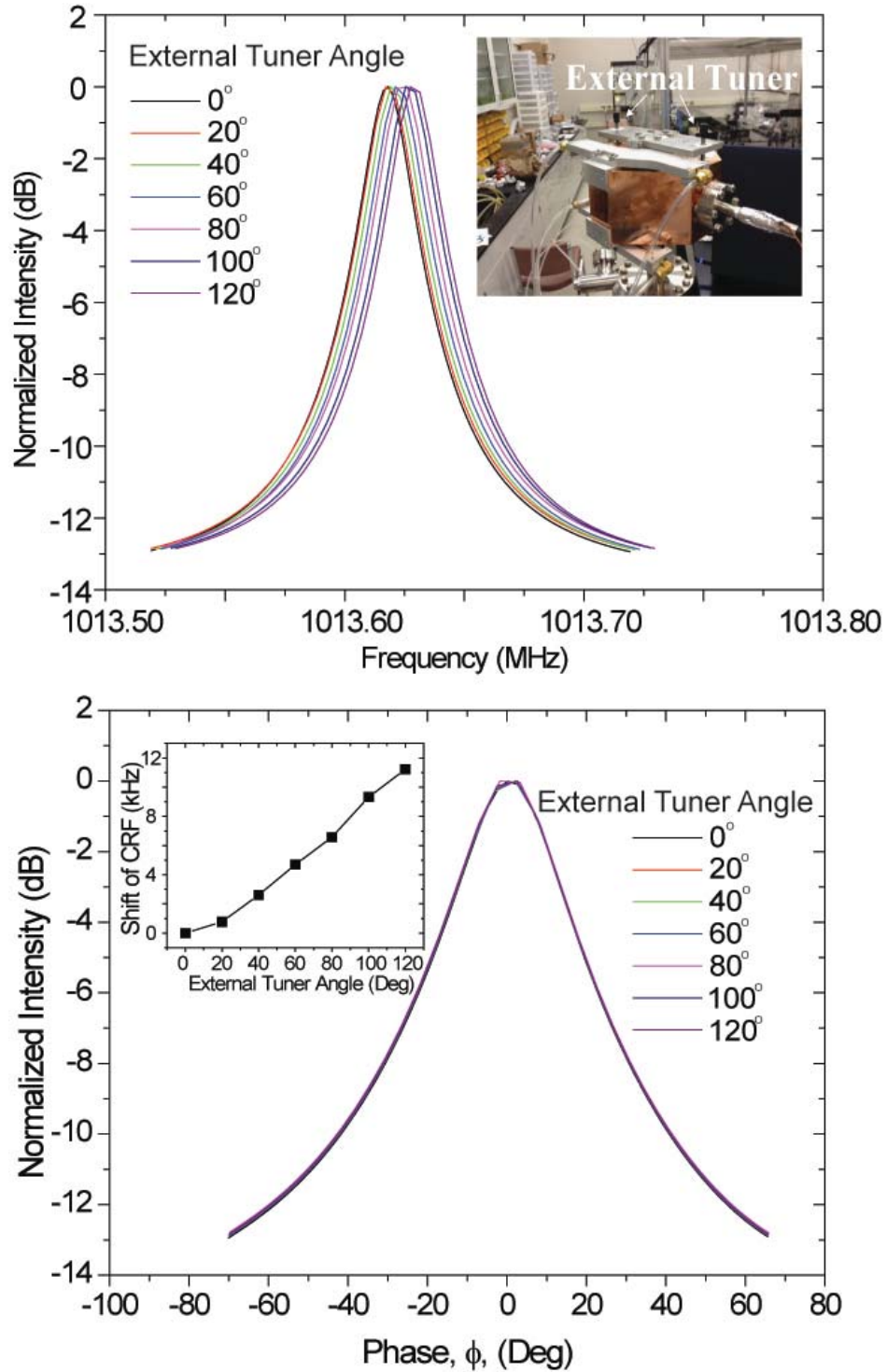


Figure 7.8: a) Cavity resonant frequency measurements with different angles of the external tuner; Inset: setup of the cavity external tuner b) Reflection power as a function of relative phase with different external tuner angles, ϕ ; Inset: the shift of cavity resonant frequency (CRF) as a function of external tuner angle. CRF at 0° tuner angle is taken as the reference.

7.8(b) carry important information, that is, when the input frequency is on resonance with the cavity CRF, the relative phase between the driving source and the RF field inside the cavity is a fixed value, which does not depend on the cavity CRF. In the framework of the equivalent circuit model (Eqn. 7.3.4) and in the resonant condition ($\omega = \omega_0$), the phase difference ϕ is always zero, regardless of the value of ω_0 , which is consistent with the results shown in Fig. 7.8(b). Eqn. 7.3.6 indicates that the sensitivity of phase to frequency offset can be affected by the value of the CRF (f_0). However, in practice, the shift of the CRF is so small compared to f_0 that the change of the sensitivity is also negligible. The consequence is that the resonance curves are almost identical with different f_0 values, as shown in Fig. 7.8(b).

7.4 RF Cavity Phase-locked Loop

In the previous section, we show that frequency detuning will cause 1) an amplitude change of the compressing electric field and 2) a relative phase shift between the input RF signal and the cavity electric field, both of which degrades the quality of the pulse compression. The frequency detuning can be caused by 1) the shift of the reference frequency from the PLL, which is directly phase-locked to the laser oscillator; 2) the shift of CRF, which is usually caused by the cavity-temperature instability. Due to the complexity of the problem, we will find that stabilizing either frequency, the shift of the other frequency will cause the cavity detuning. So, we have to take the strategy of actively adjusting one of the frequencies to track the other. Fortunately, the solution is already implied by Fig. 7.8(b), that if the relative phase ϕ is stabilized at the right position, the resonance condition of the cavity power can be satisfied at the same time, regardless of the shift of two frequencies. So, a cavity phase-locked loop (PLL) is designed to execute this job.

The block diagram of the cavity PLL is shown in Fig. 7.9. The input RF signal

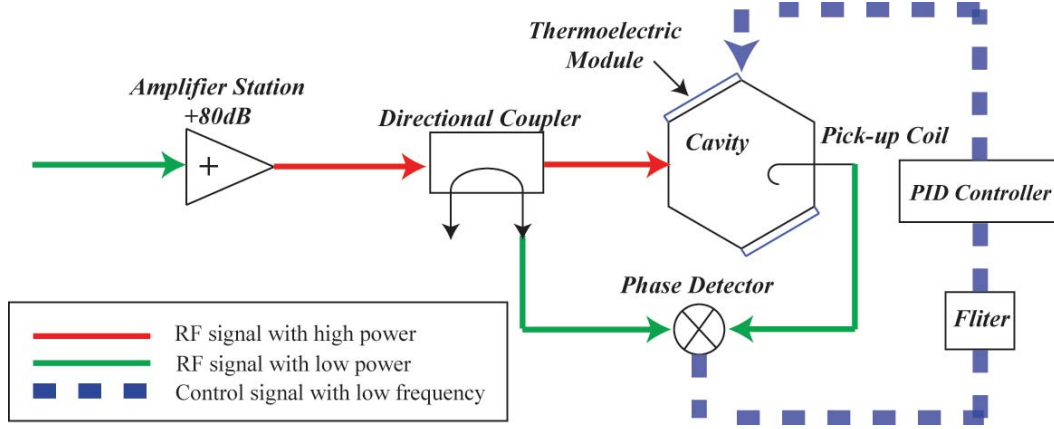


Figure 7.9: Block diagram of the setup for cavity phase-lock

is first picked up through the directional coupler. The RF field inside the cavity is extracted from the picked-up coil. Their phases are thus compared with a phase detector (AD8302) and the phase error is output as a voltage from the phase detector. To control the cavity resonance, four thermoelectric modules are mounted on the cavity serving as the temperature actuators. The voltage signal from phase detector is fed back to a PID controller (Thorlab ITC4005), which drives the thermoelectric modules to actively adjust the CRF by changing the cavity temperature. By doing so, the CRF can actively track the change of the input frequency from the low-level PLL and keep the cavity always on resonance.

To calibrate the performance of cavity PLL, a stable RF synthesizer (Agilent E8247C) is used to provide the low-level RF signal as the input, which provides more flexibility for quantifying the contributions from both the shifts of the CRF and of the reference frequency f_{Ref} . The phase error is measured by the voltage output from the phase detector. In Fig. 7.10, the calibration results with fixed f_{Ref} are first presented. As shown in Fig. 7.10(c), obvious phase fluctuations can be observed due to the cavity-temperature fluctuations, when the cavity PLL is not working. Throughout the period of measurement, the standard deviation of the phase fluctuations for a free-running cavity is ~ 340 fs and ~ 535 fs, for low RF power and high RF power, respectively. The results depend on the environmental conditions of the cavity. When the cavity PLL

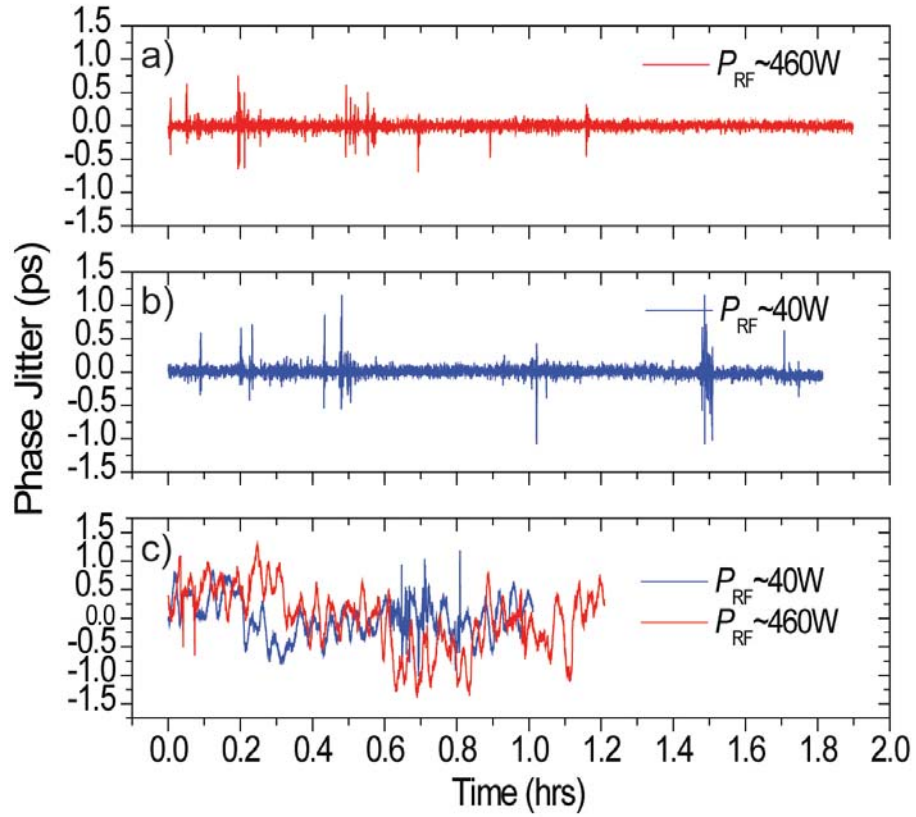


Figure 7.10: Cavity-level phase jitter measurement with a) input RF power ~ 460 W and cavity phase-lock working; b) input RF power ~ 40 W and cavity phase-lock working; c) input RF power ~ 40 W and ~ 460 W, but cavity phase-lock not working

is turned on, we can see that the low-frequency fluctuation is completely suppressed, which tells us that the cavity PLL successfully tracks the change of phase error and actively adjusts the cavity temperature to compensate it. The standard deviation of phase error for a phase-locked cavity is ~ 94 fs and ~ 74 fs for low RF power and high RF power, respectively, which is an obvious improvement compared to the free-running cavity.

Now let us take a look at the responses of the phase-locked cavity to the frequency detuning of f_{Ref} . Here, the frequency change of f_{Ref} is simply achieved by changing the frequency of Agilent E8247C. The calibration results are shown in Fig. 7.11. f_{Ref} is shifted right before each measurement with the amount specified in each figure. The experiment is continuous without any adjustments on the control parameters. As shown in Fig. 7.11, the CRF can track the change of f_{Ref} very well and the phase between the input RF signal to the cavity field is well locked, when f_{Ref} is changed by ± 4 kHz. This frequency detuning of f_{Ref} is already twice as much as the frequency shifts observed on the laser oscillator (see Chapter 6).

However, it is also observed that the overall phase fluctuations are worse in Fig. 7.11 compared to the results in Fig. 7.10. This might be caused by two processes: 1) the PID control parameters were not well optimized; 2) there were more environmental fluctuations in the second experiment. Especially, we need to take note of the second problem. The RF cavity is built with copper, which has very high thermal conductivity and large thermal mass. To completely compensate a temperature difference of 1 K between the cavity temperature and the environmental temperature, hundreds of watts of power might be needed, which is larger than the capacity of the thermoelectric modules we use here. Ultimately, the cavity phase-lock will be pulled out of lock, if the environmental temperature changes too much. At the edge of instability, additional phase noise appears in the signal. For example, the phase is shifted by ~ 250 fs at the moment ~ 0.3 hours after the measurement starts, as shown in Fig.

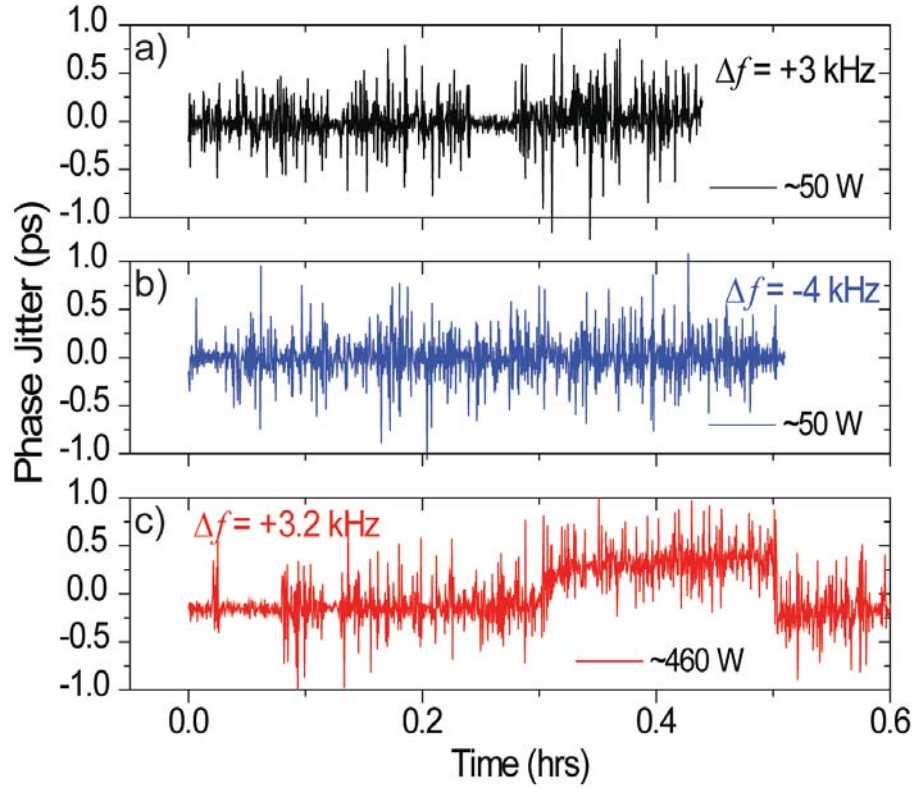


Figure 7.11: Cavity-level phase jitter measurement with a) input RF power ~ 50 W and detuning frequency, $\Delta f = +3$ kHz; b) input RF power ~ 50 W and $\Delta f = -4$ kHz; c) input RF power ~ 460 W and $\Delta f = +3.2$ kHz. (positive Δf means increase of input signal frequency; negative Δf means decrease)

7.11(c), which is caused by a change of room temperature. Further isolating the RF cavity to the environment using thermal insulating materials is expected to improve the performance.

7.5 Conclusion

In this chapter, the systems of the RF amplifiers and the RF cavity are introduced. An RF cavity is a hollow metal structure, which can resonantly store the oscillating electric field. In the RF-compression UEM, a strong oscillating electric field along the pulse longitudinal direction is used to execute the pulse compression, which is provided by the cavity's TM_{010} mode. A high-Q cavity is designed and built with

‘donut’ structure, which reduces the energy loss significantly. The power requirement for pulse compression is first estimated to be ~ 500 W. To supply enough RF power for compression, an amplifier-station, which can deliver up to 2500 W and +80 dB gain, is built using the solid-state amplifiers in parallel. Because of the very high Q-factor of the cavity, the phase and power of the cavity field is found to be sensitive to the frequency detuning. To compensate the detuning between the CRF and the low-level reference frequency, a cavity PLL is designed and characterized. With the cavity PLL, the CRF can be actively adjusted by changing the cavity temperature to track the change of the input RF frequency from the low-level PLL. In optimum, the timing jitter between the low-level reference signal and the cavity field can be as little as 100 fs.

Utilizing the results of Chapter 6 and Chapter 7, an overall setup of the RF system for the RF-compression UEM can be depicted, as shown in Fig. 7.12. The RF system is separated into two independent systems: the low-level RF system and the cavity PLL. In the low-level RF system, the combined output of the VCO and the RF synthesizer (Agilent E8247C) is synchronized with the laser oscillator pulse train through a phase-locked loop. Additional timing drift is corrected with a feedback temperature control by stabilizing the temperature of the RF components. So, the output of the low-level RF system is phase-locked to the laser oscillator pulse train, and hence the arrivals of the electron pulses. Its output power is ~ -11 dBm, which is then sent into the amplifier-station with +80 dB gain. The power of the signal is adjusted by the voltage-control attenuator before the amplifier-station. At the cavity, the timing jitter due to the fluctuations of the CRF and the laser oscillator reference frequency is further corrected by the cavity PLL, which actively tunes the cavity temperature and forces the CRF to follow the changes of f_{Ref} . So, the cavity PLL guarantees a phase-lock between the PLL output and the electric field inside the cavity. By combining these two feedback loops, the ultimate phase-lock between the cavity field and the electron pulse arrivals

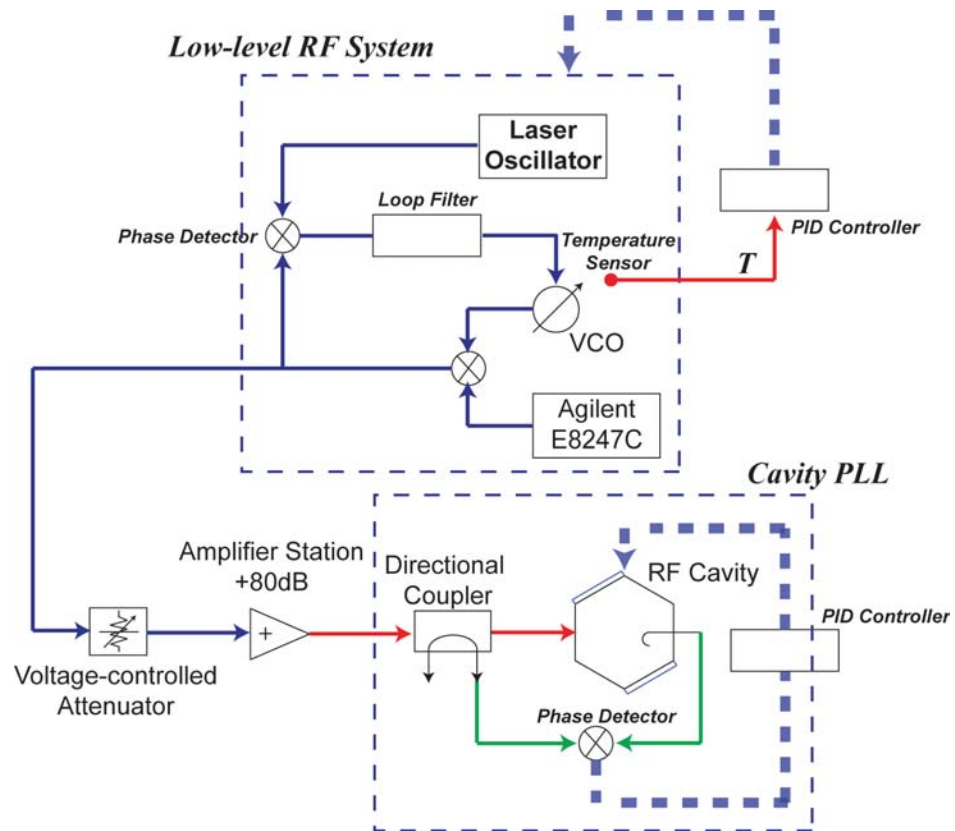


Figure 7.12: Schematic diagram for overall RF system setup.

is achieved.

Chapter 8

Preliminary Characterizations of the RF-compression UEM and Conclusions

In previous chapters, the principles of the RF-compression UEM have been introduced, and its RF-systems, including the low-level RF PLL and the cavity PLL, have been described in detail. The purpose of the RF-compression UEM is to tightly focus the electron pulse in both the transverse and the longitudinal directions, thus, to achieve high brightness and high temporal resolution at the same time. In this chapter, some of the preliminary characterizations for the RF-compression UEM will be presented in Section 8.1. The conclusion for the whole thesis will be given in Section 8.2.

8.1 Preliminary Characterizations for RF-compression UEM

The structure of the UEM is shown in Fig. 5.17. In the UEM, the photoelectron pulses are first focused by the 1st and the 2nd condenser lenses, before they arrive in

the cavity. The transverse focusing before the cavity is to make the electron beam transverse size small enough to pass through the RF cavity. In the cavity, the chirping of the electron pulses' longitudinal phase space is reversed by the oscillating electric field, as shown in Fig. 6.1, which leads to a self-compression along the longitudinal direction in the following free-drift region. Before the objective lens, an aperture is applied to select the electrons in a smaller phase space, hence to improve the beam quality in the transverse direction. A strong objective lens is used to focus the electron pulse to achieve small electron beam footprint on the sample plane. With sufficient 266 nm laser excitation, up to 5×10^6 electrons per pulse are available, when the aperture is not placed in the beam path. After being cut by a $170 \mu\text{m}$ (in diameter) aperture, more than 10^5 electrons per pulse are available for diffraction. The measurements of the number of electrons per pulse and the electron beam's transverse emittance can be found in Ref. [44]. The improvement of electron beam brightness is directly proven by the improvement of the diffraction patterns. In Fig. 8.1, the diffraction pattern from polycrystalline thin-film membrane sample of VO_2 is revisited using the RF-compression UEM. The sample geometry has been described in Chapter 3. The diffraction pattern shown in Fig. 8.1(b) is taken with 20 electron pulses accumulation, while ~ 200 pulses are required to obtain a pattern with similar SNR in the first-generation UED system. The quality of the diffraction pattern in Fig. 8.1(b) and its 1D reduction in Fig. 8.1(c) confirms that the brightness of the electron beam in the UEM system is approximately one order higher compared to the first-generation UED system [44].

8.1.1 Sensitivity of Temporal Resolution to the Timing Jitter

As has been discussed in previous chapters, the temporal resolution is related not only to the electron pulse duration, but also to the arriving-time jitter of the electron pulses to the sample plane, which can be caused by the timing jitter of the RF field.

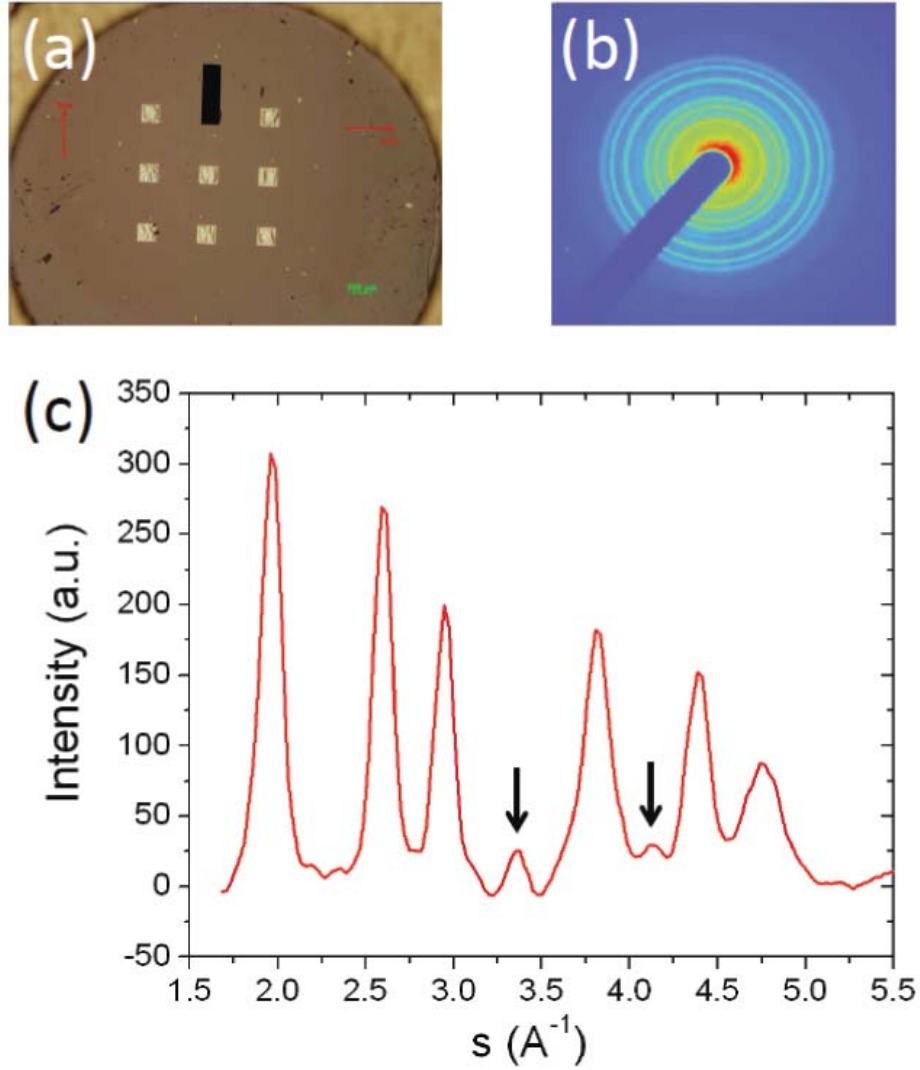


Figure 8.1: a) Optical microscope image of VO_2 thin-film membrane sample. The VO_2 is directly deposited on the silicon membrane in the square opening windows. The window size is $100 \times 100 \mu\text{m}^2$; b) Powder diffraction pattern from thin-film membrane VO_2 ; c) 1D curve reduced from the diffraction pattern shown in b), which shows the intensity distribution of Bragg peaks as a function of momentum transfer [44].

The mechanism is presented in Fig. 6.4. According to Bragg's law (Eqn. 1.2.1), the scattering angle is directly related to the electron wavelength, hence the electron energy. In other words, the diffraction patterns can be used as a probe to investigate the energy variation of the electron pulses and to characterize the arriving-time jitter induced by the timing jitter of the RF field. The schematic of the experimental setup is shown in Fig. 8.2. The relative phase is adjusted by moving the delay stage in the 266 nm laser path, which triggers the photoemission on the cathode. Because the RF field is phase-locked to the laser oscillator, by delaying the 266 nm pulse for different delay times Δt , the relative phase between the electron pulse and the RF field inside the cavity can be adjusted to a high precision. As shown in Fig. 6.4, at different phases, the electron pulse is accelerated or decelerated with different amount of energy, which serves as a probe for the strength of the cavity electric field. The delay for zero-crossing phase can be determined by matching the scattering angles with the measurements without the RF field.

The experimental results are plotted in Fig. 8.3 with 418 W RF power delivered to the cavity [215]. The energy of the electrons varies with the delay of the 266 nm laser pulse. The physical process is already described in Eqn. 7.2.1. Fitting the experimental data using Eqn. 7.2.1 yields the field strength $E_0 = 2.43$ MV/m, while the Superfish simulation gives us a field strength ~ 3.12 MV/m for the same power. The additional power loss can be attributed to the coupler, which is not considered in the Superfish calculation. The optimum condition of the pulse compression is at time delay zero, where the electrons' energy is not affected by the cavity field. What we care about is the effects of the timing jitter to the electrons' velocity, which presents a linear relationship at the vicinity of phase zero, as shown in the inset of Fig. 8.3. The slope is $dv_e/dt_{jitter} = 0.0021 \times 10^8$ m/(s·ps). With the knowledge that the cavity-to-sample distance l_{c-s} is 0.44 m, the sensitivity of the electron pulse arriving time t_e to

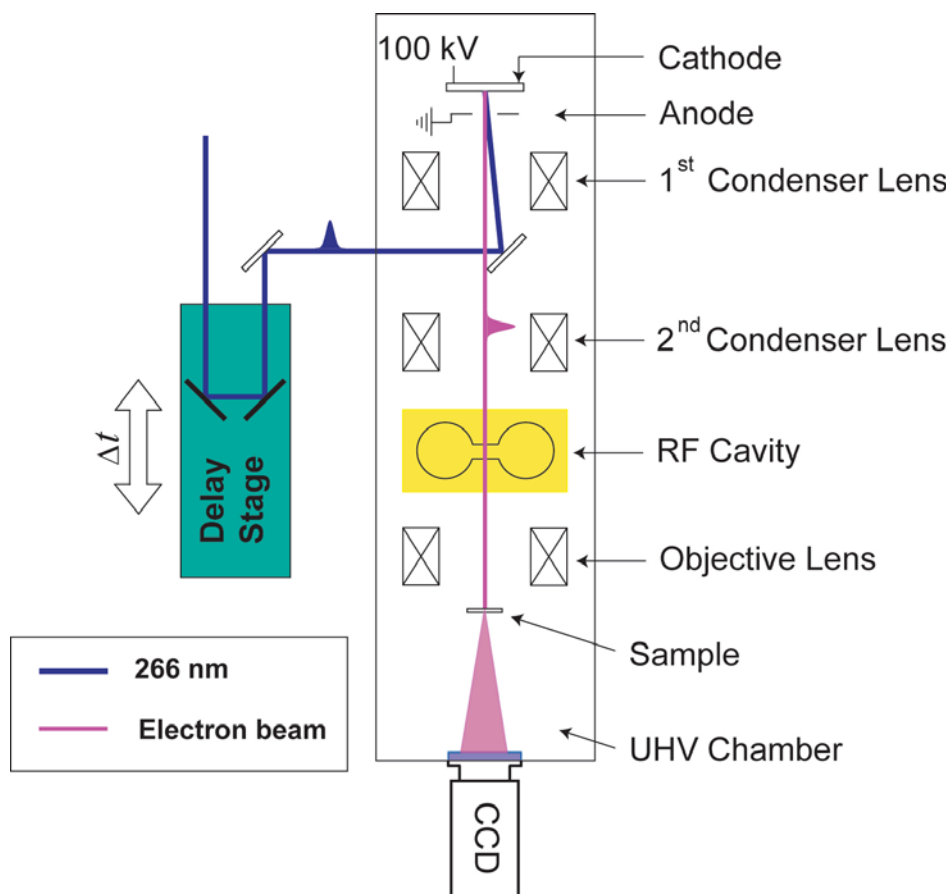


Figure 8.2: Experimental setup for measuring the cavity electric field strength using electron pulses. The optical delay line is installed in the path of 266 nm laser, which adjusts the relative timing difference between the electron pulses and the RF field. The electron energy is probed using the diffraction patterns from the sample collected with an intensified CCD camera.

the timing jitter t_{jitter} is given by

$$\begin{aligned}\eta &= \frac{dt_e}{dt_{jitter}} = \frac{dt_e}{dv_e} \cdot \frac{dv_e}{dt_{jitter}} \\ &= -\frac{l_{c-s}}{v_{e0}^2} \frac{dv_e}{dt_{jitter}},\end{aligned}\tag{8.1.1}$$

where v_{e0} is the velocity of 100 keV electrons. The sensitivity is ~ 3.43 , which means every 100 fs timing jitter (t_{jitter}) produces ~ 343 fs jitter on the arriving-time of the electron pulses. So, without considering the electron pulse duration, to have sub-ps overall temporal resolution, the accumulated timing jitter (low-level RF PLL + cavity PLL) must be less than 300 fs. More rigidly, the temporal resolution is determined by the convolution of both the pulse duration and the arriving-time jitter. At the same time, η can be changed by the cavity field strength, and it increases with the input RF power increment.

8.1.2 Characterizing the Pulse Compression

The temporal resolution is directly characterized on the VO₂ thin-film membrane sample using the cross-correlation method, which was described in Sec. 5.1. The diffraction pattern and reduced 1D curve are shown in Fig. 8.1. Here, the ultrafast dynamics of the Bragg reflection $[30\bar{2}]$ are taken as the reference for the VO₂ phase transitions, which have been discussed in detail in Chapter 3. According to the results presented in Chapter 3, the dynamics of the peak $[30\bar{2}]$ can be as fast as ~ 1 ps, which sets the limit on the best temporal resolution we can characterize here. As shown in Fig. 8.4(a), with more RF power input into cavity, the VO₂ dynamics appear to be faster. A minimum temporal resolution of ~ 5 ps is extracted through error function fitting, when a 300 W RF power is input into the cavity. In Fig. 8.4(b), we plot the temporal resolution as a function of the RF power in the cavity.

The preliminary characterizations demonstrated that the method of RF compres-

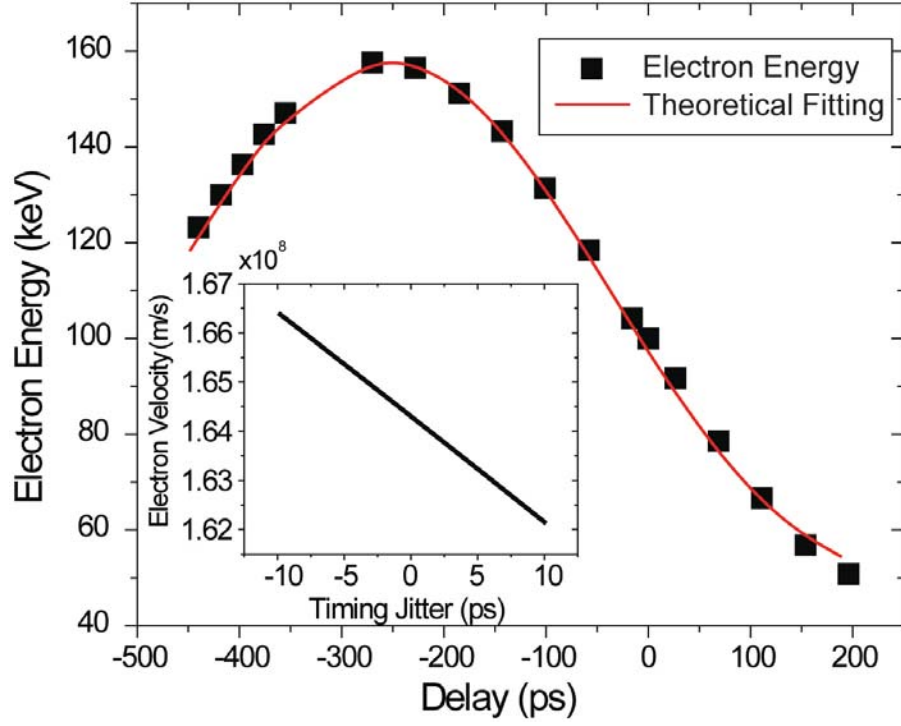


Figure 8.3: Experimental results for the electron energy as a function of the delay of the 266 nm laser pulses. The electron energy is extracted from the diffraction pattern using the Bragg's law (Eqn. 1.2.1). The red curve is the theoretical fitting of the experimental data using Eqn. 7.2.1, and the relativistic effects are considered. According to the fitting, the electric field strength E_0 is 2.43 MV/m and the RF field frequency is 1.0125 GHz. Inset: electron velocity as a function of timing jitter.

sion reduces the electron-pulse duration at the sample plane. However, the preliminary test suffers from the following issues:

1) Electron pulses' transverse expansion

The measurements are all started with $\sim 10^6$ e⁻/pulse, before the RF cavity. However, with RF compression in the longitudinal direction, the expansion of the electron pulses in the transverse direction can be observed, which significantly dilutes the electron density at the aperture level. So, after being cut by the aperture, $\sim 10^5$ e⁻/pulse is obtained for 0 W and 210 W RF power, but only $\sim 10^4$ e⁻/pulse is available for 300 W RF power. A direct consequence of this is the reduced SNR of the results with high-power compression, as shown in Fig. 8.4(a).

2) Distortion of diffraction pattern

Distortion of the diffraction patterns is found with the high RF power input (not shown here). The precise reason for this distortion is still not clearly known. There are two suspects. First, the distortion might be caused by the aberration of the objective lens, which increases significantly, when the RF compression leads to more energy spread in the electron pulses. Second, the non-linearity of the RF field in space and time may strongly disturb the phase-space distribution of the electrons. The appearance of these effects is likely amplified given the fact that the electron beam column is not yet fully aligned.

3) Temperature instability with the high-RF power

With the RF power increased to 300 - 400 W, more than 1 kW power is expected to be deposited from the RF amplifier-station and the chillers to the environment, which can cause the room temperature to increase by ~ 1 °C (depending on the specific conditions of the building air conditioners) and air flows due to the imbalance of the temperature in the lab. Many instruments, which have high sensitivity to the environmental temperature, can become less stable, including the laser oscillator, the RF cavity and the low-level RF PLL. A better water cooling system and a control of air

circulation could help reduce the temperature fluctuations at the experiment station.

8.2 Conclusions

The results presented in this thesis demonstrated the application of ultrafast electron diffraction (UED) to investigate the fundamental physical principles of many-body correlations in solids, in which VO_2 serves as an exemplary system. The observations highlight the important interplay between the strongly correlated electrons and the lattice degrees of freedom in the strongly correlated systems.

The photoinduced phase transitions in VO_2 polycrystalline thin film was first investigated using the UED method with different pump wavelengths. The phase transition process was identified to be a charge-driven non-equilibrium process, which requires less energy compared to its thermodynamic counterpart. The details of the atomic movements were extracted from the UED results, which illustrates the step-wise atomic motions after photo-excitation. The initial dilation of V-V bonding strongly suggests the excitation of d bonding electrons initiated by fs laser pulses, which have strong correlation features in VO_2 's electronic structure [142]. It is also consistent with recent theoretical investigations, which indicate the important contribution of the electron correlations on V-V dimers [69]. The phase transitions of VO_2 were then studied at the sub-micrometer scale using VO_2 single-crystal nanobeams. The structural phase transition was found to be decoupled from the metal-to-insulator transition investigated with the optical microscope and electron diffraction, which clearly distinguish the Mott-Hubbard nature of the insulating band gap in VO_2 [119]. The decoupling effect was attributed to the interface charge doping from different types of substrates. The results here suggest a new phase of VO_2 (monoclinic but metallic), which has not been covered by theoretical investigations. At the same time, it also highlights the sensitivity of the strongly correlated material to its environmental parameters, when

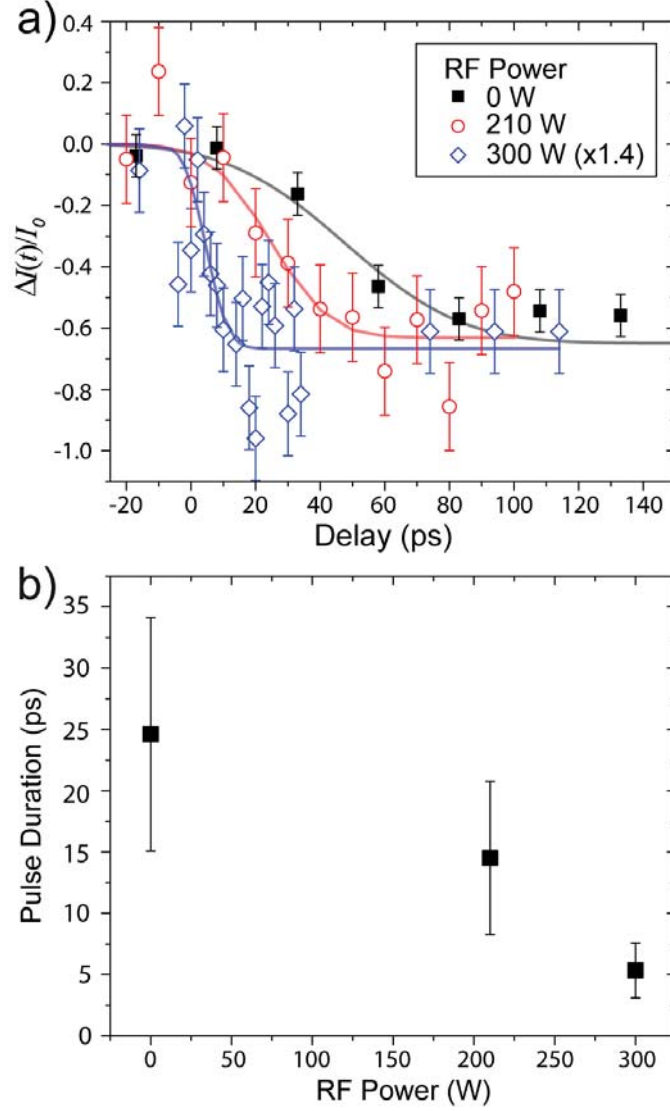


Figure 8.4: a) Intensity dynamics of the Bragg reflection $[30\bar{2}]$ with different RF powers in the cavity. The solid lines are the fitting of the dynamics with error functions, whose width yields the temporal resolution of the RF-compression UEM. The total number of electrons are all $\sim 10^6$ e^- /pulse before the RF cavity. Due to the aperture cutting, $\sim 10^5$ e^- /pulse for 0 W and 210 W RF power and $\sim 10^4$ e^- /pulse for 300 W power are used for diffraction. b) Temporal resolution, which is extracted from the error-function fitting in (a), as a function of RF power.

its dimension is reduced to microscopic scale. In the UED measurements, anisotropic phonon response and strong reduction of the optically accessible spectral weight in the new M_3 phase have been identified, indicating a spectral weight transfer throughout the metal-to-insulator transition, which has been confirmed by other experimental [58] and theoretical [69] studies.

The success of these studies is based on the direct access of short electron probes to the ultrafast temporal evolution of the lattice structures in solids after optical excitation. From the studies, we found we reached the limit of the first-generation UED system on both the beam brightness and the temporal resolution, which is limited by the space charge effects. Using the projection-shadow-imaging technique, the space charge effect is investigated at the near-cathode-surface region, where the strong Coulomb repulsion plays an important role, because of the high electron density in this region. The following space-charge-led features have been identified [159]:

1. Power-law scaling between pulse width to the number of electrons in the pulse;
2. Broadening of initial phase space of electron pulses during the photoemission process;
3. Virtual cathode effect, which limits the electron brightness.

To increase the electron pulse brightness and reduce its duration, an RF-compression scheme is proposed and its performance is discussed under the framework of the analytical Gaussian model (AGM) [177, 178, 188], serving as the fundamental model for the RF-compression ultrafast electron microscope (UEM), which is under development at Michigan State University.

In the RF-compression UEM, an RF cavity is implemented to compress the space-charge-broadened electron pulse in its longitudinal direction by applying an oscillating electric field to reverse the momentum-space correlation of the electron pulse in its phase space [159, 212]. The optimum pulse-compression condition relies on the

high timing accuracy between the electron pulse arrivals to the RF electric field inside the cavity, which can be achieved through phase-lock loops (PLLs). The phase-locking problems at the low-power level and cavity level have been introduced and discussed in detail. At the low-power level, ~ 50 fs phase jitter has been demonstrated by temperature-stabilizing RF components of the low-level RF PLL. The fundamental limit is found to be the amplitude and pointing fluctuations of the laser oscillator. At the cavity level, the temperature of the cavity is actively adjusted by the cavity PLL to track the frequency shift of the laser oscillator. An optimum timing jitter of ~ 100 fs is demonstrated at the high-power level. The performance of the RF-compression UEM is, then, characterized and the pulse compression is demonstrated on the VO_2 thin-film sample, which shows fast response of the lattice structure after optical excitation. The best temporal response determined from the experiments is ~ 5 ps, which is largely limited by instabilities in the setup at the current stage. At the time of writing this thesis, a new system parameters are set up and under calibration, which is able to achieve a better temporal resolution with both the beam and the temperature stability. Ongoing efforts to improve the performance of the UEM include:

- 1) Reduce the temperature fluctuations induced by the high-power amplifiers

Currently, the RF-amplifier station is overrated and generates huge amount of heat, which affects the room temperature, hence causes fluctuations to other temperature-sensitive instruments, including the laser oscillator, the low-level PLL and the RF cavity. With better knowledge of the compression power requirement, a simpler amplifier structure can be implemented to improve the stability of the overall system.

- 2) Actively stabilize the laser oscillator amplitude and pointing for low-level PLL

The timing jitter of the low-level PLL is still limited to several hundred fs for a long-term operation, which is caused by the amplitude and pointing fluctuation of the laser oscillator. Different approaches have been proposed to reduce its contribution to the timing jitter [204, 206, 207] and sub-10 fs timing jitter for long-term operation has

been demonstrated to be achievable [207].

3) Tune the electron optics for high-density electron pulses

The electron optics in conventional electron microscopes is designed for single (or low-density) electron operations, where the strong electron-electron repulsion is usually ignored. However, in the RF-compression UEM, the performance of the electron optics depends highly on the electron beam density. At different beam densities, the electron optical arrangement needs to be re-optimized. Experimental protocol must account for this effect. A better defined optical tuning procedure informed by the efficient multi-particle simulations [216], taking the space charge effect into account, needs to be put in place.

With these improvements, manifold perspectives arise for investigations in the strongly correlated materials with micrometer, even nanometer, dimensions, such as electron and lattice interplay in high- T_c superconductors, charge-density wave materials, etc. Especially, with the improvements of electron optics for high-density electron pulses, atomic resolution images can be taken with high temporal resolution, which will be useful to illustrate interesting science in many disciplines, such as protein folding in biological molecules, making and breaking bonds in chemical reactions, transient states in solid-state phase transitions, to name just a few exciting possibilities.

APPENDICES

Appendix A

VO₂ Structures and Diffraction Simulation

A.1 Rutile Structure

The rutile structure of metallic VO₂ is based on a tetragonal lattice with space group: $P4_2/mnm$. The metal atoms are located at Wyckoff position (2a): $(0, 0, 0)$ and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. The oxygen atoms are located at position (4f): $(u, u, 0)$, $(-u, -u, 0)$, $(\frac{1}{2} + u, \frac{1}{2} - u, \frac{1}{2})$, $(\frac{1}{2} - u, \frac{1}{2} + u, \frac{1}{2})$, where u is called the internal parameter. McWhan et al. reported the lattice constants and the value of u of the rutile structure [149]: $a_R = 4.5546 \text{ \AA}$, $c_R = 2.8514 \text{ \AA}$ and $u = 0.3001$. According to the symmetry of space group $P4_2/mnm$, the unit cell of the rutile structure contains two formula units and the atomic coordinates can be derived, as shown in Table A.1. In Table A.1, the unit cell vectors are defined in a way that the a_R axis is aligned with the Cartesian x axis, b_R with the y axis and c_R with the z axis.

	Orthogonal Coordinates			Fractional Coordinates		
Atom	x (Å)	y (Å)	z (Å)	x_{frac}	y_{frac}	z_{frac}
V ₁	0.0000	0.0000	0.0000	0.0	0.0	0.0
V ₂	2.2773	2.2773	1.4257	0.5	0.5	0.5
O ₁	1.3664	1.3664	0.0000	0.3	0.3	0.0
O ₂	-1.3664	-1.3664	0.0000	-0.3	-0.3	0.0
O ₃	3.6437	0.9109	1.4257	0.8	0.2	0.5
O ₄	0.9109	3.6437	1.4257	0.2	0.8	0.5

Table A.1: List of orthogonal atomic coordinates and internal fractional coordinates of VO₂ Rutile (*R*) structure.

A.2 M₁ Structure

The M₁ phase of VO₂ is characterized by the monoclinic structure with the space group $P2_1/c$. Anderson first reported the lattice constants of the M₁ structure [122]: $a_{M1} = 5.743$ Å, $b_{M1} = 4.517$ Å, $c_{M1} = 5.375$ Å and $\beta_{M1} = 122.61^\circ$. Later on, Longo and Kierkegaard refined lattice constants [124]: $a_{M1} = 5.7517$ Å, $b_{M1} = 4.5378$ Å, $c_{M1} = 5.3825$ Å and $\beta_{M1} = 122.646^\circ$. In the subsequent calculation, the lattice constants from Anderson will be used. The unit cell of the M₁ structure contains four formula units. The metal atoms, as well as the two different types of oxygen atoms, occupy the Wyckoff positions (4e): (u, v, w) , $(-u, -v, -w)$, $(u, \frac{1}{2} - v, \frac{1}{2} + w)$ and $(-u, \frac{1}{2} + v, \frac{1}{2} - w)$, where u , v and w are the internal structural parameters shown in Table A.2.

With the structure parameters in Table A.2 and lattice constants, the orthogonal atomic coordinates and internal fractional coordinates can be derived, which are listed in Table A.3.

Atom	u	v	w
V	2.424	0.975	0.025
O ₁	0.10	0.21	0.20
O ₂	0.39	0.69	0.29

Table A.2: Crystal structure parameters of M₁-VO₂ as given by Anderson [122]

	Orthogonal Coordinates			Fractional Coordinates		
Atom	x (Å)	y (Å)	z (Å)	x_{frac}	y_{frac}	z_{frac}
V ₁	-4.4041	0.1132	-1.3174	0.242	0.975	0.025
V ₂	4.4041	-0.1132	1.3174	-0.242	-0.975	-0.025
V ₃	2.1456	2.3770	0.1310	0.242	-0.475	0.525
V ₄	-6.6626	2.1507	2.7657	-0.242	1.475	0.475
O ₁	-0.9486	0.9055	0.0050	0.1	0.21	0.2
O ₂	-3.1167	1.3130	-1.3997	0.39	0.69	0.29
O ₃	0.9486	-0.9055	-0.0050	-0.1	-0.21	-0.2
O ₄	3.1167	-1.3130	1.3997	-0.39	-0.69	-0.29
O ₅	-1.3099	3.1694	1.4534	0.1	0.29	0.7
O ₆	0.8582	3.5769	0.0486	0.39	-0.19	0.79
O ₇	-3.2071	1.3583	1.4433	-0.1	0.71	0.3
O ₈	-5.3752	0.9508	2.8481	-0.39	1.19	0.21

Table A.3: List of the orthogonal atomic coordinates and the internal fractional coordinates of M₁-VO₂ structure.

There are several details worth noting in Table A.3:

1. In the M_1 structure, the V-V dimer forms along the a_{M1} axis, which is equivalent to the c_R axis in the rutile structure. In Table A.3, the c_R axis is defined to be along the Cartesian z-axis. So, to be consistent, in the M_1 -VO₂, a_{M1} is along the z-axis, b_{M1} along the x-axis, and c_{M1} is located in y-z plane with an angle $\beta_{M1} = 122.61^\circ$ to the x-axis (a_{M1}). For convenience, we can define the M_1 lattice axis in the Cartesian coordinates: $a_{M1} = (0, 0, -5.743)$, $b_{M1} = (-4.517, 0, 0)$ and $c_{M1} = (0, 4.5277, 2.8967)$.
2. In the definition of the internal fraction coordinates, x_{frac} is the fractional atomic coordinate along the a_{M1} axis, y_{frac} along the b_{M1} axis and z_{frac} along c_{M1} axis. So, z_{frac} is not orthogonal to x_{frac} .

The structure of M_1 -VO₂ is closely related to the high-temperature rutile structure. The most significant difference is the V-V atom dimerization along the c_R (a_{M1}) axis, which breaks the high-temperature tetragonal symmetry and doubles the unit cell size along the c_R (a_{M1}) axis. Comparing the lattice constants of the rutile and M_1 structures, we find that the M_1 's translational symmetry can be approximately written as:

$$a_{M1} \approx \begin{pmatrix} 0 \\ 0 \\ -2c_R \end{pmatrix}, b_{M1} \approx \begin{pmatrix} -a_R \\ 0 \\ 0 \end{pmatrix}, c_{M1} \approx \begin{pmatrix} 0 \\ a_R \\ c_R \end{pmatrix}. \quad (\text{A.2.1})$$

This representation emphasizes that the unit-cell size of the monoclinic structure is approximately double of the rutile unit-cell size along the c_R axis. There is a lattice strain, which causes deviations of the ratios: $\frac{c_{M1} \sin(\beta_{M1})}{c_R} = 1.0159$, $\frac{-c_{M1} \cos(\beta_{M1})}{a_R} = 0.9941$ and $\frac{a_{M1}}{2c_R} = 1.0070$, from unity. To simplify the problem, we are going to ignore the lattice strain and assume the ratios above are unity, then the M_1 structure can be directly derived by distort V-V bonding in the rutile structure within a supercell containing

	Orthogonal Coordinates			Unit cell shift
Atom	x (Å)	y (Å)	z (Å)	
V ₁	0.1129	0.1132	-1.3174	(0,b-1,0)
V ₂	-0.1129	-0.1132	1.3174	(0, b+1,0)
V ₃	2.1456	2.3770	0.1310	(0, 0, 0)
V ₄	2.3714	2.1507	2.7657	(0, b-2,0)
V ₅	0.1129	4.6409	1.5793	(0, b-1,c+1)
V ₆	-0.1129	4.4145	-1.5289	(a+1, b+1,c+1)
V ₇	2.1456	6.9047	3.0277	(a+1, b+1, c+1)
V ₈	2.3714	6.6784	-0.0806	(a+1, b-2,c+1)

Table A.4: List of the orthogonal atomic coordinates of the M₁-VO₂ with the supercell definition. The ‘Unit cell shift’ shows the movements of the atom from its original atomic coordinates in Table A.3 to the new coordinate. For example, ‘(a + 1, b − 1, 0)’ means the new atomic coordinates are derived by moving the atom along +a_{M1} axis by one unit cell and along the −b_{M1} axis by one unit cell from its original coordinates.

eight formula units: $a = (a_R, 0, 0)$, $b = (0, 2a_R, 0)$ and $c = (0, 0, 2c_R)$, as shown in Fig. A.1. To confine the atoms of the M₁ structure into a supercell, the V atoms’ orthogonal coordinates have the following confinements:

$$x \in [0, a_R], y \in [0, 2a_R], z \in [0, 2c_R]. \quad (\text{A.2.2})$$

To agree with the above confinements, the V atoms listed in Table A.3 can be moved by one unit cell along the a_{M1}, b_{M1} and c_{M1} axes or any combinations of these to find the coordinates of their equivalent atoms in the supercell. Table A.4 shows the new V atoms coordinates after the movements.

A.3 Atomic Movements from Rutile Structure to M_1 structure

In order to figure out how the V-V bonds are distorted from the rutile structure to the M_1 structure, we need to find out the correspondence of the atoms in Table A.4 to the atoms in Table A.5. The list of the V atoms in Table A.5 is sorted that a direct correspondence between Table A.5 and Table A.4 can be found. By comparing the atomic coordinates, it is not difficult to derive the distortions within the V-V bonds, as shown in Table A.6. We find that the supercell contains two long bonds (V_5 - V_6 and V_7 - V_8) and two short bonds (V_1 - V_2 and V_3 - V_4). None of the four V-V bonds are equivalent. In addition, the bond V_3 - V_4 is offset from the bond V_1 - V_2 by $(\frac{1}{2}a_R, \frac{1}{2}a_R, \frac{1}{2}c_R)$, the bond V_5 - V_6 is offset by $(0, a_R, 0)$ and the bond V_7 - V_8 is offset by $(\frac{1}{2}a_R, \frac{3}{2}a_R, \frac{1}{2}c_R)$. Fig. A.1 represents the V atom movement from the rutile (dashed red circles) to the monoclinic (solid blue circles) structures.

We can shift all the atoms' coordinates in Table A.4 by 1.4357 Å in the z direction and rearrange the oxygen atoms' coordinates. Thus, a direct correspondence between the atomic coordinates in the two structures can be seen. Table A.7 lists the orthogonal coordinates of V atoms and O atoms in the rutile structure and the monoclinic structure with the supercell definition.

	Orthogonal Coordinates			Unit cell shift
Atom	x (Å)	y (Å)	z (Å)	
V ₁	0.0000	0.0000	0.0000	(0,0,0)
V ₂	0.0000	0.0000	2.8514	(0,0,c+1)
V ₃	2.2773	2.2773	1.4257	(0,0,0)
V ₄	2.2773	2.2773	4.2771	(0,0,c+1)
V ₅	0.0000	4.5546	2.8514	(0,b+1,c+1)
V ₆	0.0000	4.5546	0.0000	(0,b+1,0)
V ₇	2.2773	6.8319	4.2771	(0,b+1,c+1)
V ₈	2.2773	6.8319	1.4257	(0,b+1,0)

Table A.5: List of the orthogonal atomic coordinates of $R\text{-VO}_2$ with the supercell definition. The ‘Unit cell shift’ shows the movements of the atom from its original atomic coordinates in Table A.1 to the new coordinates. The list is ordered in a one-to-one correspondence with the M_1 structure atomic coordinates in Table A.4

V-V bonds	Δx (Å)	Δy (Å)	Δz (Å)	Bond Distance (Å)
V ₁ -V ₂	-0.2258	-0.2264	2.6348	2.6541
V ₃ -V ₄	0.2258	-0.2264	2.6348	2.6541
V ₅ -V ₆	-0.2258	-0.2264	-3.1082	3.2146
V ₇ -V ₈	0.2258	-0.2264	-3.1082	3.2146

Table A.6: Four pairs of the V-V bonds in the supercell. $\Delta x_i = x_{i(m)} - x_{i(n)}$ for the bond V_n-V_m, where $i = \{x, y, z\}$

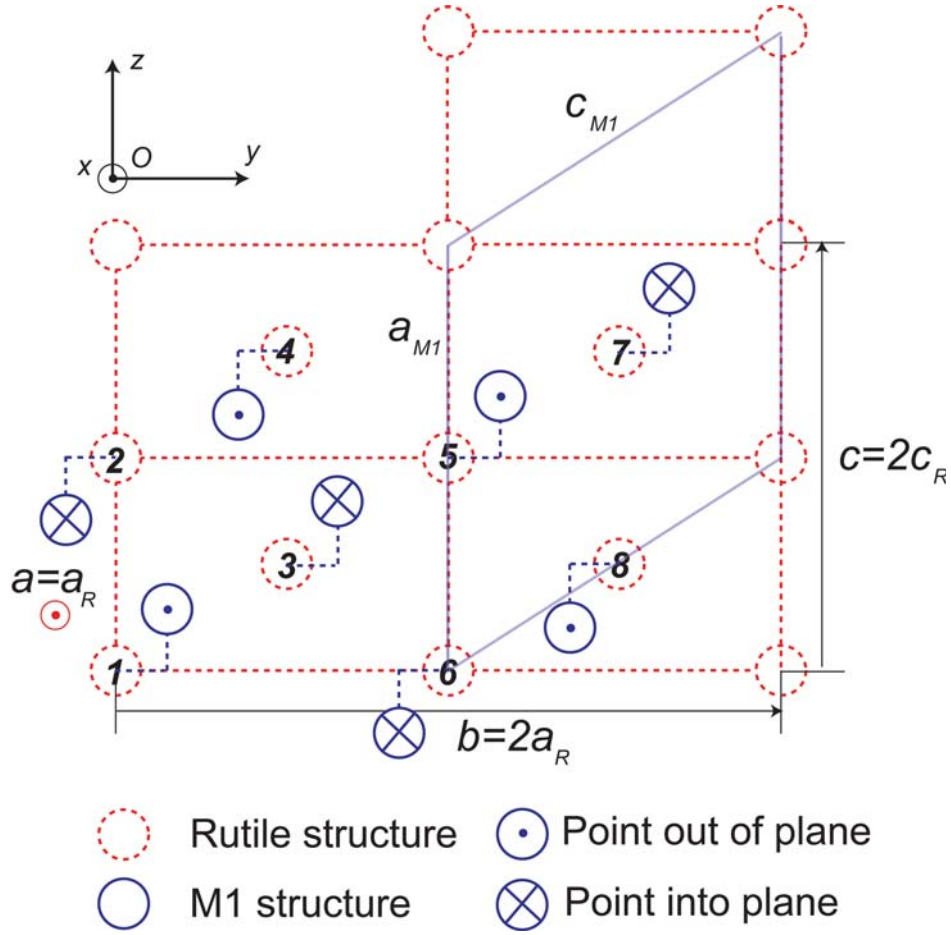


Figure A.1: V-V bond distortion illustrated under supercell definition. The dashed red circles represent the rutile structure, in which the V atoms are equally spacing along rutile c_R axis. The blue solid circles represent the M_1 structure. The blue dashed lines show the atomic distortion from rutile position to monoclinic position in y-z plane. The cross means atom moves along -x direction, while the dot +x direction. The displacements of V atoms are exaggerated for the purpose of illustration. Blue solid line encircles the unit cell of $M_1\text{-VO}_2$.

	Rutile			M₁		
Atom	x (Å)	y (Å)	z (Å)	x (Å)	y (Å)	z (Å)
V ₁	0.0000	0.0000	0.0000	0.1129	0.1132	0.1183
V ₂	0.0000	0.0000	2.8514	-0.1129	-0.1132	2.7531
V ₃	2.2773	2.2773	1.4257	2.1456	2.3770	1.5667
V ₄	2.2773	2.2773	4.2771	2.3714	2.1507	4.2014
V ₅	0.0000	4.5546	2.8514	0.1129	4.6409	3.0150
V ₆	0.0000	4.5546	0.0000	-0.1129	4.4145	-0.0932
V ₇	2.2773	6.8319	4.2771	2.1456	6.9047	4.4634
V ₈	2.2773	6.8319	1.4257	2.3714	6.6783	1.3551
O ₁	1.3664	1.3664	0.0000	1.4003	1.3130	0.0360
O ₂	-1.3664	-1.3664	0.0000	-1.3099	-1.3583	-0.0076
O ₃	3.6437	0.9109	1.4257	3.5684	0.9055	1.4407
O ₄	0.9109	3.6437	1.4257	0.8582	3.5769	1.4843
O ₅	1.3664	1.3664	2.8514	1.3099	1.3583	2.8790
O ₆	-1.3664	-1.3664	2.8514	-1.4003	-1.3130	2.8354
O ₇	3.6437	0.9109	4.2771	3.6588	0.9508	4.2838
O ₈	0.9109	3.6437	4.2771	0.9486	3.6222	4.3274
O ₉	1.3664	5.9210	0.0000	1.3099	5.8860	0.0327
O ₁₀	-1.3664	3.1882	0.0000	-1.4003	3.2147	-0.0109
O ₁₁	3.6437	5.4655	1.4257	3.6588	5.4785	1.4375
O ₁₂	0.9109	8.1983	1.4257	0.9486	8.1499	1.4811
O ₁₃	1.3664	5.9210	2.8514	1.4003	5.8407	2.9327
O ₁₄	-1.3664	3.1882	2.8514	-1.3099	3.1694	2.8891
O ₁₅	3.6437	5.4655	4.2771	3.5684	5.4332	4.3374
O ₁₆	0.9109	8.1983	4.2771	0.8582	8.1046	4.3810

Table A.7: List of the orthogonal atomic coordinates of the rutile and the M₁ structures in the supercell coordinates. The atoms are moved that one-to-one correspondence on the atomic coordinates is formed between two structures.

Appendix B

Optical Constants of VO₂ M₁ Phase

The real part ε_1 and imaginary part ε_2 of the dielectric constants can be extrapolated from Ref. [126]. The refractive index n is given by

$$n = \sqrt{\frac{\sqrt{\varepsilon_1^2 + \varepsilon_2^2} + \varepsilon_1}{2}}, \quad (\text{B.0.1})$$

and the extinction coefficient k is given by

$$k = \sqrt{\frac{\sqrt{\varepsilon_1^2 + \varepsilon_2^2} - \varepsilon_1}{2}}. \quad (\text{B.0.2})$$

Using the results from Ref. [126], n and k can be calculated and are plotted as a function of the photon energy in Fig. B.1. In Ref. [74], n and k for 790 nm photons (1.57 eV) are reported to be 2.9 and 0.5, respectively, which is in fair agreement with Fig. B.1.

The optical reflectivity R can be calculated using n and k by

$$R = \frac{(1 - n)^2 + k^2}{(1 + n)^2 + k^2}, \quad (\text{B.0.3})$$

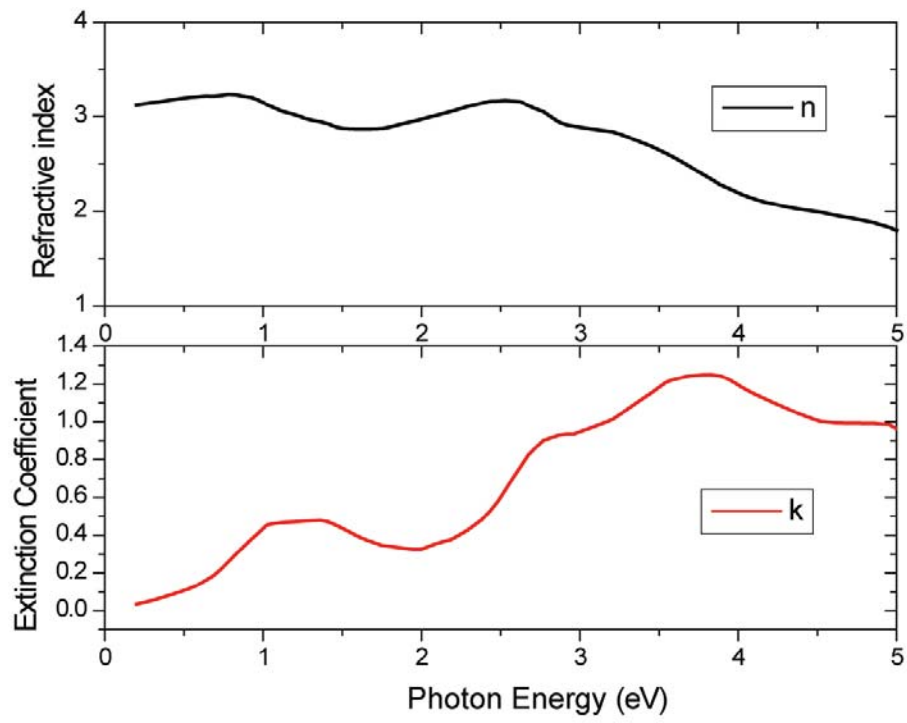


Figure B.1: Refractive index n and extinction coefficient k as a function of the photon energy, derived using the data from Ref. [126].

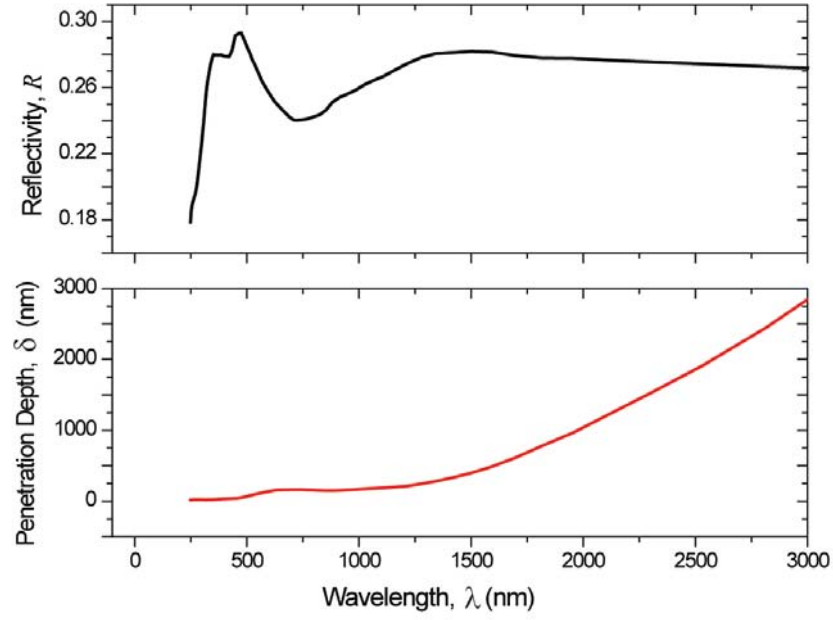


Figure B.2: Reflectivity R and optical penetration depth δ as the function of the wavelength λ .

and the penetration depth δ is given by

$$\delta = \frac{\lambda}{4\pi k}. \quad (\text{B.0.4})$$

R and δ as a function of the wavelength λ is plotted in Fig. B.2.

Appendix C

Supplementary Materials for VO₂ Nanobeam Experiments

C.1 Extended Steady-state Experiments of VO₂ Nano- beams Supported by Different Substrates

To investigate the reproducibility of the separation of MIT and SPT of VO₂ nanobeams on different substrates, a total of 11 different samples were prepared and measured (4 on the Au grids, 3 on the Si window, 2 on the Cu grid, and 2 on the Ni grids). The TEM windows and grids used are listed below:

1. Gold: Gilder Extra Fine Bar Grids, Ted Pella, Prod NO. G1500HSG
2. Silicon: Silicon Aperture Frame (no support film), Ted Pella, Prod NO. 21542-10
3. Copper: Gilder Extra Fine Bar Grids, Ted Pella, Prod NO. G1500HS
4. Nickel: Gilder Extra Fine Bar Grids, Ted Pella, Prod NO. G1500HSN

These TEM grids/windows were used without additional surface cleaning treatment.

The different results of the MIT and SPT measurements are tabulated in Table C.1. Generally, there are distinct differences in the SPT critical temperature during heating (T_c^H) on different substrates. The Cu grid produces the highest T_c^H (higher than 363 K - the temperature limit of the TEM stage), followed by the Au grid (361 K) and the Ni grid (342 K). In contrast, all the MIT experiments show similar critical temperature $T_c^H=339\pm2$ K, independent of the substrates. On the silicon substrate the T_c^H of MIT and SPT are indistinguishable within the precision of our temperature calibration at two different stages. The critical temperatures determined during cooling are less reproducible and highly depend on the history of heating and cooling, resulting in a broad range of ΔT from 4 to 30K.

Similar measurements were also carried out on as-grown VO_2 nanobeams on a SiO_2 surface. Under an optical microscope, periodic insulating and metallic domains can be clearly observed with sample temperature is raised across the critical temperature, which indicates the strain effects induced from the substrate. The structural phase transition is studied using our ultrafast electron diffraction (UED) system, where the sample temperature can be increased across the phase transition critical temperature. The diffraction pattern is shown in Fig. C.1(b), where the dimer peaks, $(5\ 1\ \bar{2})$, $(3\ 0\ \bar{1})$ and $(1\ \bar{1}\ 0)$, are the signatures of the V-V dimerization in the monoclinic structure.

C.2 Estimation of Temperature Rise after Pump Laser Excitation

To estimate the absorbed pump energy and understand the ultrafast pump-probe dynamics, it is very important to apply the correct optical constants in the calculation. We determine the optical absorbance and penetration depth of VO_2 at wavelength $\lambda = 800$ nm using the real (n) and imaginary (k) parts of the reflection index. For mono-

Sample	Substrate	MIT			SPT		
		T_c^H (K)	T_c^C (K)	ΔT (K)	T_c^H (K)	T_c^C (K)	ΔT (K)
Au-001	Gold	340.7	326.7	12.0	358.9	330.5	28.4
Au-002	Gold	338.0	327.3	12.0	362.8	336.5	26.3
Au-003	Gold	339.3	332.2	7.1	360.3	346.1	14.2
Au-004	Gold	340.4	335.2	5.2	>356	–	–
Si-001	Silicon	336.8	332.3	4.8	338.1	333.2	4.9
Si-002	Silicon	338.4	333.2	5.2	337.1	331.0	6.1
Si-003	Silicon	338.6	334.8	3.8	339.6	333.8	5.8
Cu-001	Copper	337.0	331.9	5.1	>363.0	–	–
Cu-002	Copper	339.3	333.2	6.1	>363.0	–	–
Ni-001	Nickel	338.9	324.7	14.2	343.6	329.8	13.8
Ni-002	Nickel	335.7	326.6	9.1	339.0	327.7	11.3

Table C.1: Various critical temperatures obtained during heating (T_c^H) and cooling periods (T_c^C) in the metal-insulator transition (MIT) and the structural phase transition (SPT) characterizations of VO₂ nano-beams supported on different substrates. The supercooling effect can be gauged by the downshift of the critical temperature obtained during cooling compared to that during heating period (ΔT).

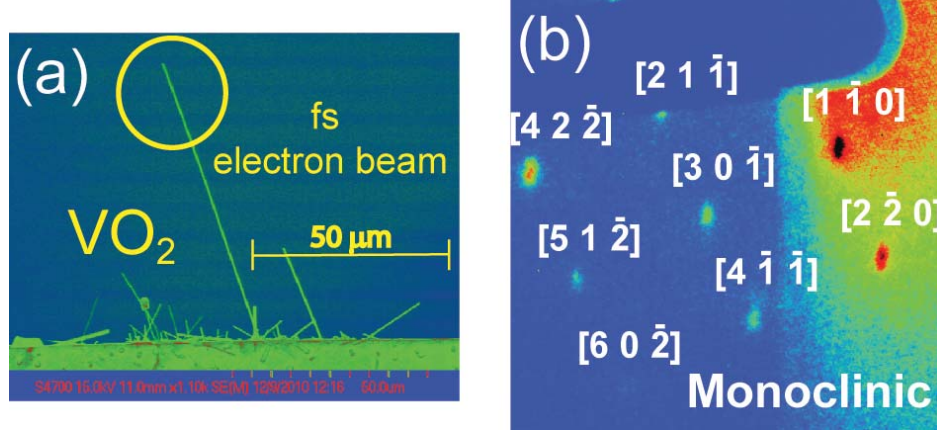


Figure C.1: a. The SEM image of VO₂ nano-beams grown on a SiO₂ substrate. The electron beam cross-section is presented in the circular area. b. Diffraction pattern of VO₂ nano-beam as-grown on SiO₂ substrate at room temperature

clinic M₁ phase, $n=2.9$, $k=0.5$; and for rutile phase $n=2.3$, $k=0.79$ [74]. The reflectivity R can be calculated by

$$R = \frac{(1 - n)^2 + k^2}{(1 + n)^2 + k^2}. \quad (\text{C.2.1})$$

We obtain, for monoclinic M₁ phase $R=0.25$, and for rutile phase $R=0.19$. The absorbed optical energy density E_{ph} can be calculated by

$$E_{ph} = F(1 - R)/d, \quad (\text{C.2.2})$$

where F is the pump laser fluence, and $d = \lambda/4\pi k$ is the penetration depth, which is 127.3 nm for the monoclinic M₁ phase and 88.4 nm for the rutile phase, respectively. The laser-induced temperature rise (ΔT) can be determined by

$$\Delta T = E_{ph}/C_v, \quad (\text{C.2.3})$$

where C_v is the heat capacity. With $F=7 \text{ mJ cm}^{-2}$ and $C_v=3.1 \text{ J K}^{-1}\text{cm}^{-3}$ near the room temperature [133], we can get a temperature rise of 133 K for M₁ phase (assuming no phase transition) and 207 K for rutile phase, which is shown as dashed lines in Fig.

4.8(c).

C.3 Anisotropic Debye-Waller Analysis of Pump-probe Measurements on Au-supported VO₂ Nanobeams

In addition to the temperature estimation using the optical and thermal constants, the Debye-Waller analysis also allows us to extract the lattice temperature directly from the UED results.

The analysis starts from fitting the diffraction patterns. We used the program ‘Crystal-maker’ to simulate the VO₂ diffraction pattern and compare it with the experimental data. The VO₂ nanobeam’s orientation can be resolved through the comparison, and the diffraction pattern can be correctly indexed, as shown in Fig. C.2.

According to the Debye-Waller (DW) theory, the intensity of the Bragg peaks are related to the amplitude of atomic displacement variances for a reciprocal lattice vector q^* , which is given by

$$\sigma_{q^*}^2 = -\frac{\ln\left(\frac{I_{q^*}(t)}{I_{q^*}(t < 0)}\right)}{q^{*2}}, \quad (\text{C.3.1})$$

where $I_{q^*}(t)$ is the intensity of the Bragg peak with the reciprocal lattice vector q^* , $I_{q^*}(t < 0)$ is the ground-state peak intensity before the pump pulse arrives. In the single-crystal diffraction, each Bragg peak carries the information of a crystal momentum q^* , which represents a set of crystal planes with a certain orientation in real space. The anisotropic structural dynamics can be distinguished from the difference in the intensity dynamics of different Bragg peaks. However, because the signal-to-noise ratio (SNR) of a single Bragg peak is not high enough to extract the dynamics of the corresponding atomic planes, averages among different Bragg peaks need to be taken

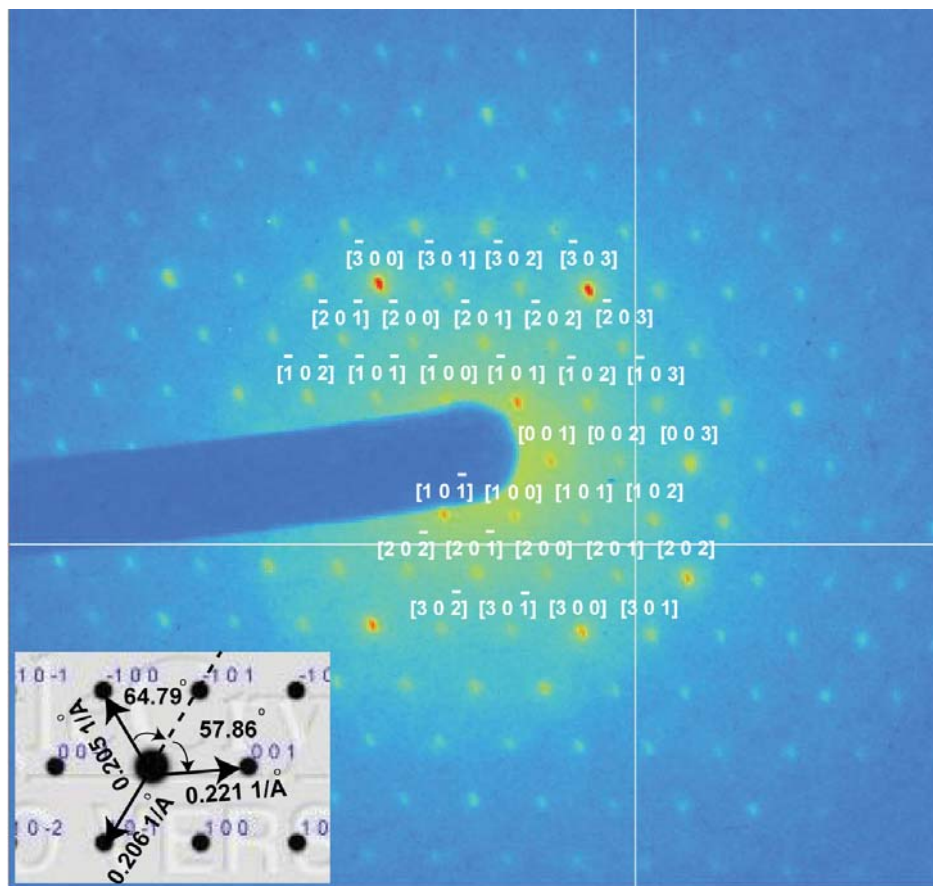


Figure C.2: a) Electron diffraction pattern with Miller indices of VO₂ single crystal nanobeam measured using UED. Inset: Crystal-maker simulation of the diffraction pattern.

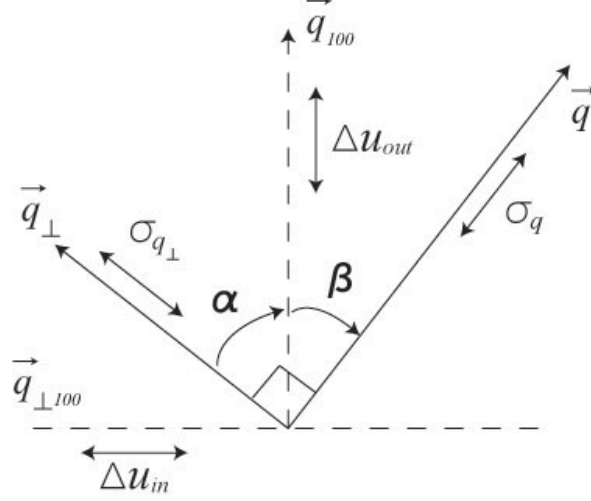


Figure C.3: Principle of projection of atomic variance with arbitrary vector q^* and $q_{\perp}^*$ onto $q_{[100]}^*$

to improve the (SNR). In VO_2 , the V-V dimerization is along the a_M (c_R) axis with Miller index $[100]$, and its electronic structure is also strongly anisotropic [58, 70]. To capture the anisotropic nature of the atomic variances, we can project the atomic variances of other directions onto the $[100]$ direction and the direction perpendicular to $[100]$ ($\perp[100]$).

In Eqn. C.3.1, the atomic displacement variances along the vector q can be quantified by the intensity change of the Bragg peak of q^* , where q^* is the reciprocal lattice vector of the real-space vector q . Similarly, the atomic fluctuations perpendicular to q is related to the Bragg peak intensity of $q_{\perp}^*$, where $q_{\perp}^*$ is also perpendicular to q^* in the reciprocal space. The two atomic variances σ_q and $\sigma_{q_{\perp}}$ can be projected onto the directions of $[100]$ and $\perp[100]$ by considering the angles α and β , as shown in Fig. C.3. In the analysis, we include all the available pairs of q^* and $q_{\perp}^*$ in the diffraction pattern and average their intensity variations to get the averaged atomic variances Δu_{out}^2 (out-of- bc -plane) along the $[100]$ direction and Δu_{in}^2 (in- bc -plane) along the $\perp[100]$ direction.

The following is the mathematical derivations of Δu_{out}^2 and Δu_{in}^2 from the intensity variations of the Bragg peaks. The averaged atomic variances along the $[100]$ direction

can be calculated by

$$\begin{aligned}\Delta u_{out}^2 &= \frac{\sum_i^{N_{pair}} (\sigma_{q_i}^2 \cos^2 \beta_i + \sigma_{q_{i,\perp}}^2 \cos^2 \alpha_i)}{N_{pair}} \\ &= \frac{\sum_i^{N_{pair}} \left[-\frac{\ln(I_{q_i}(t)/I_{q_i}(t < 0))}{q_i^2} \cos^2 \beta_i - \frac{\ln(I_{q_{i,\perp}}(t)/I_{q_{i,\perp}}(t < 0))}{q_{i,\perp}^2} \cos^2 \alpha_i \right]}{N_{pair}},\end{aligned}\tag{C.3.2}$$

where N_{pair} is the number of the pairs of the Bragg reflections. Similarly, in the direction perpendicular to $[1\ 0\ 0]$ direction, we have

$$\begin{aligned}\Delta u_{in}^2 &= \frac{\sum_i^{N_{pair}} (\sigma_{q_i}^2 \sin^2 \beta_i + \sigma_{q_{i,\perp}}^2 \sin^2 \alpha_i)}{N_{pair}} \\ &= \frac{\sum_i^{N_{pair}} \left[-\frac{\ln(I_{q_i}(t)/I_{q_i}(t < 0))}{q_i^2} \sin^2 \beta_i - \frac{\ln(I_{q_{i,\perp}}(t)/I_{q_{i,\perp}}(t < 0))}{q_{i,\perp}^2} \sin^2 \alpha_i \right]}{N_{pair}}.\end{aligned}\tag{C.3.3}$$

The anisotropic atomic displacement variances after pump-laser excitation is shown in Fig. 4.8(c). We found that there is a difference between Δu_{out}^2 and Δu_{in}^2 in the M_1^* phase, while this difference disappears when VO_2 is in the M_3 phase. This result indicates that there might be soft-phonon mode excited along the a_M axis during the M_1^* -to- M_3 MIT, which was also indicated by previous X-ray diffraction experiments [149].

The relationship between the mean square of the atomic variances and the crystal temperature has been characterized by early X-ray experiments [143, 149, 217], which showed an anomalously large DW factor in the metallic rutile structure suggesting a lattice softening. Fig. C.4(a) shows the data of the atomic variances projected along the rutile a_R and c_R axes, reconstructed using the results from earlier X-ray diffraction [143, 149] and X-ray near-edge absorption spectroscopy [217] experiments. A sharp jump is clearly visible (up to a factor of 4) following the insulator-metal transitions. A linear correlation between the temperature rise and the change of atomic variances can

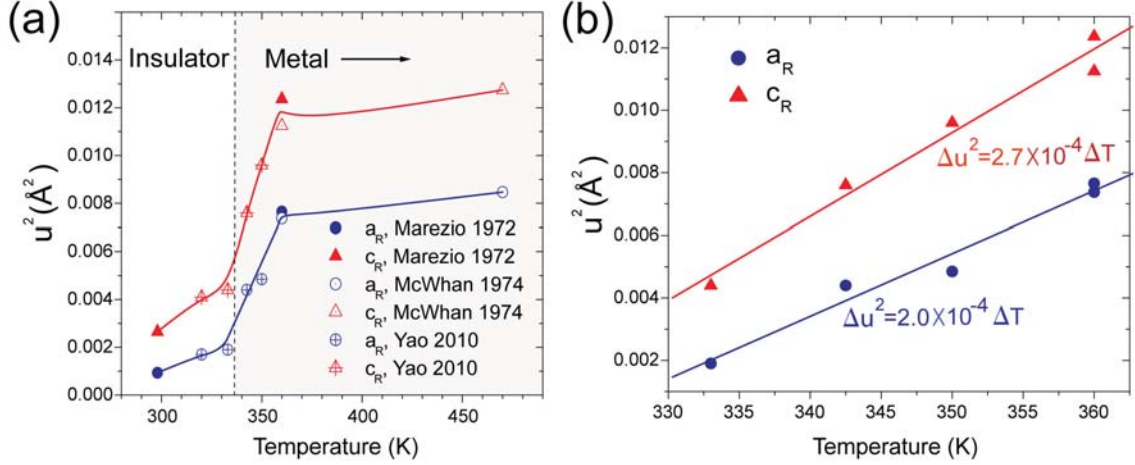


Figure C.4: (a) The projected atomic displacement variance u^2 , compiled based on data extracted from X-ray diffraction experiments (see Marezio et al. [143] and McWhan et al. [149]), and X-ray near-edge absorption spectroscopy experiment (see Yao et al. [217]). Generally, u^2 is larger along the rutile c-axis (c_R) than along the rutile a-axis (a_R) and b-axis (b_R , not shown, but similar to a_R values). A significant rise of u^2 following insulator-metal transition is clearly visible. (b) The correlations between the temperature rise (ΔT) and the increase in atomic displacement variance (Δu^2) in the transition region are shown.

be established in the transition region, as shown in Fig. C.4(b). Using this correlation, we can estimate the temperature increase following photo-excitation in the M_3 state. Note that, in Fig. C.4, u^2 stands for the absolute value of the atomic variances, while Δu^2 is the relative change of the atomic variances from the ground state. Because, in the UED measurements, the ground-state Bragg peak intensities are always taken as the reference, only the changes of the atomic variances induced by the pump laser are measured. Based on the maximum change in $\Delta u^2 \sim 0.0035 \text{\AA}^2$ along c_R , we determine ΔT to be no more than 20 K at $F=7 \text{ mJ cm}^{-2}$.

Appendix D

Supplementary Materials for Projection-shadow-imaging Technique

D.1 Normalization Factor for Analytical Description

The integrated intensity of the shadow image in Eqn. 5.2.5 should only depend on the electron sheet density, if we assume the electron-electron scattering cross-section λ does not change. However, if we plot the function in Eqn. 5.2.5 with a fixed electron sheet density, Σ , but with a varying point-projection distance (PPD), x_0 , we will find that the integrated intensity under the shadow image profile is not a constant, as shown in Fig. D.1(a). To normalize Eqn. 5.2.5, we can integrate the function over the range on

the screen

$$\begin{aligned}
\int_{-\infty}^{\infty} F(d) dd &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \frac{\lambda \xi N_e}{\pi \sqrt{2\pi} \sigma_x \sigma_y \sigma_z} \exp \left[-\frac{x^2}{2\sigma_x^2} \right] \exp \left[-\frac{(z - z_0)^2}{2\sigma_z^2} \right] dx \right) \frac{L}{x_0} dz \\
&= \frac{L}{x_0} \sqrt{\frac{2}{\pi}} \frac{\lambda \xi}{\sigma_y} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dz \frac{N_e}{2\pi \sigma_x \sigma_z} \exp \left[-\frac{x^2}{2\sigma_x^2} \right] \exp \left[-\frac{(z - z_0)^2}{2\sigma_z^2} \right] \\
&= \frac{L}{x_0} \sqrt{\frac{2}{\pi}} \frac{\lambda \xi}{\sigma_y} N_e \\
&= \frac{L \sigma_x}{x_0} \sqrt{2\pi} (\lambda \xi \Sigma),
\end{aligned} \tag{D.1.1}$$

where d is the position on the screen. The definitions of other parameters can be find in Chapter 5. To cancel the effects of the dimensional parameters (x_0 , L and σ_x) on the integrated intensity, a normalization factor $x_0/(L\sigma_x)$ needs to be applied. The normalized form of Eqn. 5.2.5 is

$$\begin{aligned}
M(d) &= \frac{x_0}{L\sigma_x} F(d) \\
&= \frac{x_0}{L\sigma_x} \frac{(\sqrt{2\pi} \lambda \xi) \Sigma}{\sigma_z} \frac{\exp \left[-\frac{(dx_0 - Lz_0)^2}{2(d^2 \sigma_x^2 + L^2 \sigma_z^2)} \right]}{\sqrt{\frac{1}{\sigma_x^2} + \frac{d^2}{L^2 \sigma_z^2}}}.
\end{aligned} \tag{D.1.2}$$

In Fig. D.1(b), we can see that the integrated intensity of the normalized analytical function, $M(d)$, is a constant for a fixed electron sheet density.

D.2 Calibration of Projection Shadow Imaging Technique

D.2.1 Image the Photoemitted Electron Pulses from HOPG with Different x_0

To understand and calibrate the projection shadow image technique, a series of experiments are carried out on the HOPG surface, following R. K. Raman's early work

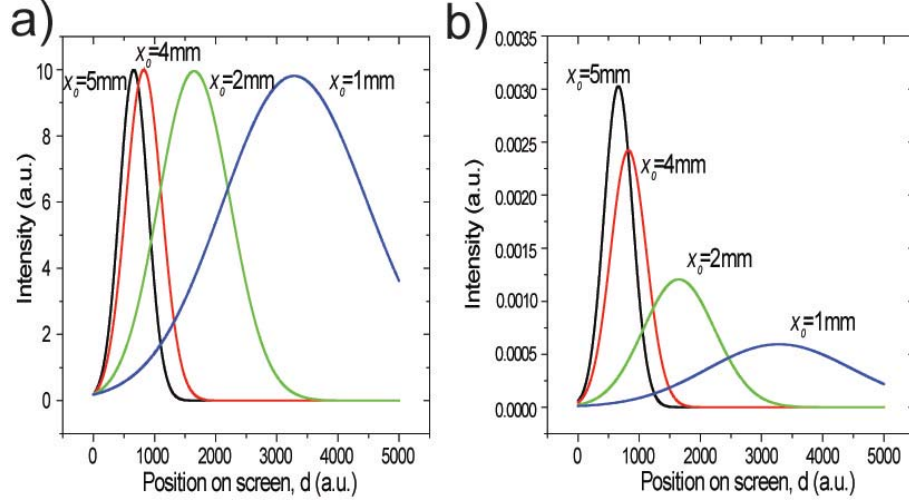


Figure D.1: a) Analytical description of the shadow images (Eqn. 5.2.5) on the camera, with $\lambda\xi\Sigma = 1.0$, $\sigma_z = 10 \mu\text{m}$, $L = 165 \text{ mm}$, $z_0 = 20 \mu\text{m}$ and varying PPD ($x_0 = 5$ (black), 4 (red), 2 (green), 1 (blue) mm). The curves have the same intensity, but different integrated intensities. (b) Analytical function after normalization, $M(d)$ (Eqn. D.1.2), with the same set of parameters used in (a). The integrated intensity is a constant with varying x_0 .

[158]. The experiments are operated with different point-projection distances (PPDs), x_0 , with the same pump laser configuration, and hence, the same photoemission conditions. x_0 is changed by bending the electron beam using a front-deflector, thus, adjusting the location of point P on the sample surface, as shown in Fig. D.2(a). The exact x_0 values are confirmed by the spatial cross-correlation measurements using laser pumps and electron probes. The obtained shadow images are analyzed using the analytical model in Eqn. 5.2.6. The center of mass (COM) locations and the pulse longitudinal widths as a function of delay time are plotted in Fig. D.2(b) and (c) for different values of x_0 . Ideally, because of the identical photoemission conditions, the extracted dynamics of the COMs and pulse widths should be the same with different values of x_0 , if Eqn. 5.2.6 be reliable to deconvolute the projection-imaging geometry. From the results shown in Fig. D.2(b) and (c), we can conclude that fitting the experimental image profile using Eqn. 5.2.6 can give us a good estimation on the pulse COM locations and longitudinal widths. We do find that when x_0 is too small, the

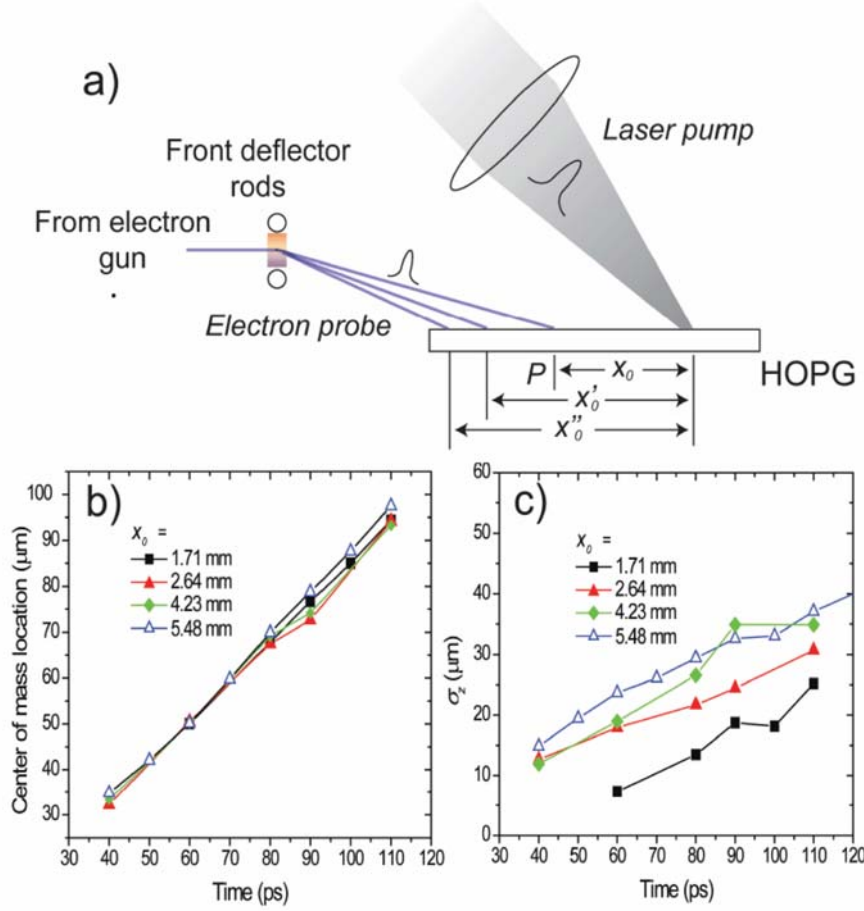


Figure D.2: a) Side-view of the experimental setup. Pump laser is aimed at a fixed position on HOPG surface with fluence 54 mJ/cm^2 , to ensure the reproducible photoemission conditions, when x_0 is changing. b) Center of mass (COM) location above HOPG surface as the function time, which is obtained from analytical-model analysis (Eqn. 5.2.6). c) Electron pulse longitudinal width, σ_z , evolves as the function of time.

results deviate from the others, like $x_0 = 1.71 \text{ mm}$ in Fig. D.2(c), so short x_0 should be avoided in experimental setups, in which the shallow angle of the incident electron beam elongates the footprint of probing electron pulse and P can no longer be treated as a point source.

On the other hand, the fitting using Eqn. 5.2.6 cannot yield a consistent result for the electron density. The fitting results of $\lambda\xi\Sigma$ are not a constant for different values of x_0 , even though the electron pulses should have the same electron density. This result indicates that Eqn. 5.2.6 is physically not correct, because it completely ignores the

Coulomb field of the photoelectron pulses.

D.2.2 Ray-Tracing Imaging Simulation

To further understand the projection shadow imaging technique, we need to measure the number of electrons and find the relationship between the electron sheet density and fitting parameter $\lambda\xi\Sigma$ in Eqn. 5.2.5 and 5.2.6. If we assume every pulse contains 10^6 electrons, with 1 kHz repetition rate, the total current we will have is only ~ 100 pA, which is very difficult to be measured accurately. In addition, the experimental setup also limits the space available for mounting the current sensor. So, we choose to use the ray-tracing simulations of the projection shadow imaging process to predict the density of electrons.

The ray-tracing program calculates the probing electrons' trajectory and numerically rebuilds the shadow images. To avoid the expensive multi-particle simulations, the program uses disk mean-field model of the target electron pulse to calculate the probing electrons' trajectory. The errors from the disk mean-field model is first corrected by comparing its electric field distribution with that of the real electron pulse with a three-dimensional Gaussian distribution. The ray-tracing simulation program is constituted by two parts:

1. *Field Program* is to correct the errors of the electric field distribution caused by the disk mean-field approximation.
2. *Imaging Program* is to calculate the probing electrons' trajectories and the normalized shadow images on the 'screen'

In the *Field Program*, the mean-field of the slices of the electron pulse is calculated and compared with a multi-particle calculation with high precision. Then, the 'near field' and the 'far field' corrections are derived and to be used in the *Imaging Program*.

In the *Imaging Program*, the imaging process is simulated. For convenience of the calculation, the probing electron-source is assumed to be a single point at the same height as the target electron pulse, as shown in Fig. D.3(a). This approximation is justified by the fact that the imaging projection angles in the experiments are usually very small ($z_0/x_0 < 0.02$ rad). To get enough signal-to-noise ratio (SNR), 10^7 electrons are ‘emitted’ from the source point, P . The trajectories of the electrons are calculated considering the electric field of the target electron pulse and the electron counts are recorded and pixelated by the ‘CCD screen’, which is 16.5 cm away from P . The pixel size is $44.95 \mu\text{m}$, which is the same as the pixel size of the real CCD camera in the experiments. Following the data analysis procedure in real experiments, the simulated electron distribution on the ‘screen’ is subtracted and normalized by the reference distribution, which is taken, when the target electron pulse is not in space. The dimensional parameters ($\sigma_x, \sigma_y, \sigma_z$) of the target electron pulse are directly set in the program, while the simulated shadow images are fitted using Eqn. 5.2.6. The results confirm that the analytical description (Eqn. 5.2.6) can be used to reliably extract the dimensions of photoelectron pulses, when the electron density is not very high ($\rho < 100 \text{ e}^-/\mu\text{m}^3$). The *Imaging Program*, on the other hand, can be used to fit the experimental shadow-image profile to extract the electron density, which cannot be accurately obtained by the analytical fittings. Fig. D.3(b) and (c) show the simulated shadow images compared with the experimental results with $x_0 = 5.48 \text{ mm}$ and 1.71 mm , respectively. The fitting gives electron sheet density to be $\Sigma = 8.7 \times 10^{12} \text{ m}^{-2}$ for $x_0 = 5.48 \text{ mm}$ and $\Sigma = 8.5 \times 10^{12} \text{ m}^{-2}$ for $x_0 = 1.71 \text{ mm}$, which is consistent with each other.

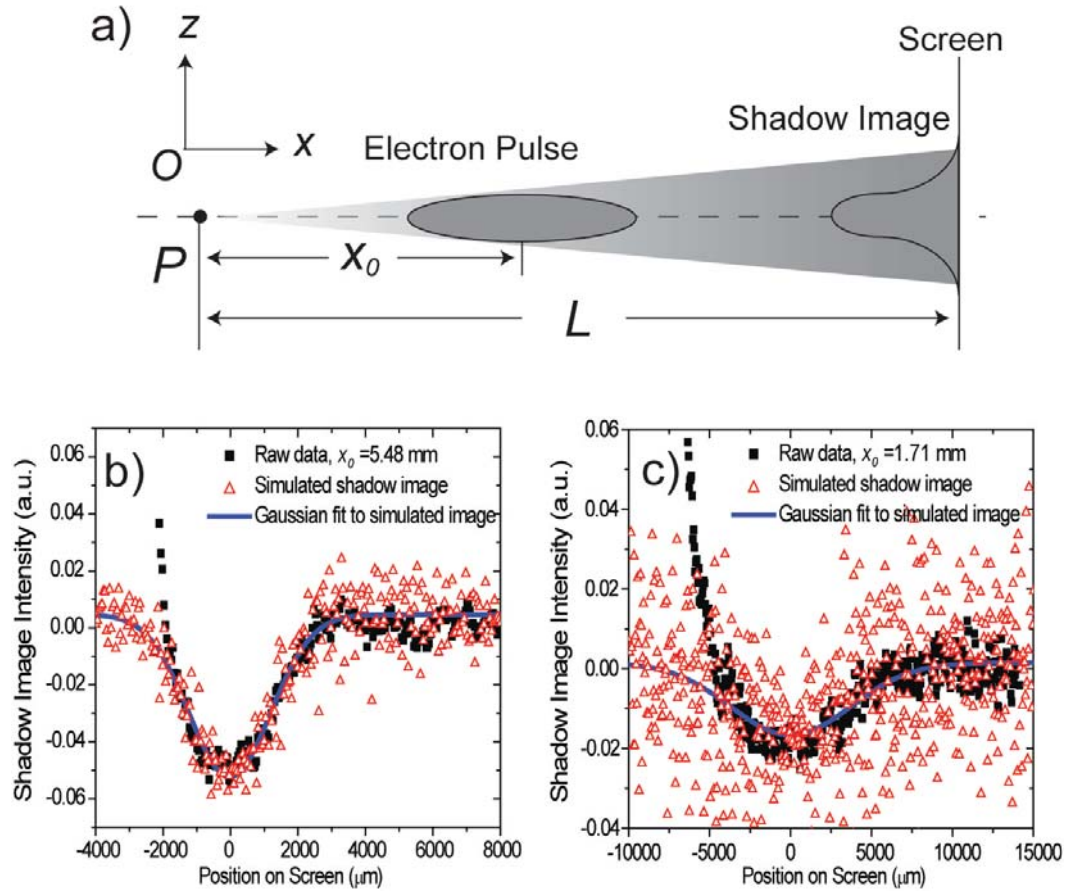


Figure D.3: a) Geometry of shadow image simulation. The source point P is at the same height as the target electron pulse. b) Comparison of simulated shadow image (red open triangle) with the experimental results (black solid square) with $x_0 = 5.48$ mm. The data is taken with delay time $\Delta t = 100$ ps after photoemission. The parameters in the shadow-imaging simulation are $\sigma_x = 370$ μm , $\sigma_y = 260$ μm , $\sigma_z = 40$ μm and $\Sigma = 8.7 \times 10^{12}$ m^{-2} . c) Comparison of simulated shadow image (red open triangle) with the experimental results (black solid square) with $x_0 = 1.71$ mm. The dimensional parameters are the same as b), while $\Sigma = 8.5 \times 10^{12}$ m^{-2} .

D.2.3 Experimentally Measuring Electron Sheet Density with Bragg-beam Deflecting Effects

The electron sheet density, Σ , can also be experimentally measured from the dynamics of the Bragg peaks' movements. In HOPG diffraction patterns, there are three Bragg peaks available at the center of the electron pulse image, which are indexed as (0,0,6), (0,0,8) and (0,0,10), as shown in Fig. D.4(a) [158]. Due to the Coulomb repulsion of the photoelectron pulses, the Bragg beams are deflected, which presents as movements of the Bragg peaks on the CCD camera. The dynamics of the Bragg-peak movements are plotted in Fig. D.4(b). The movements is initially upwards (away from the sample surface), since the photoelectron pulse is lower than Bragg peaks, while the movement direction is switched to downwards, after the electron pulse crosses the Bragg beams. The solid lines in Fig. D.4(b) are the simulation results for the Bragg-beam deflecting. The dimensional parameters of the photoelectron pulses are extracted from the experimental data using Eqn. 5.2.6. The electron sheet density Σ is $8.0 \times 10^{12} \text{ m}^{-2}$ in the Bragg-beam deflection simulation. Here, we can see that the Bragg-beam deflection simulation successfully captures the process of the Bragg-peak movements, and it also helps to confirm the charge density.

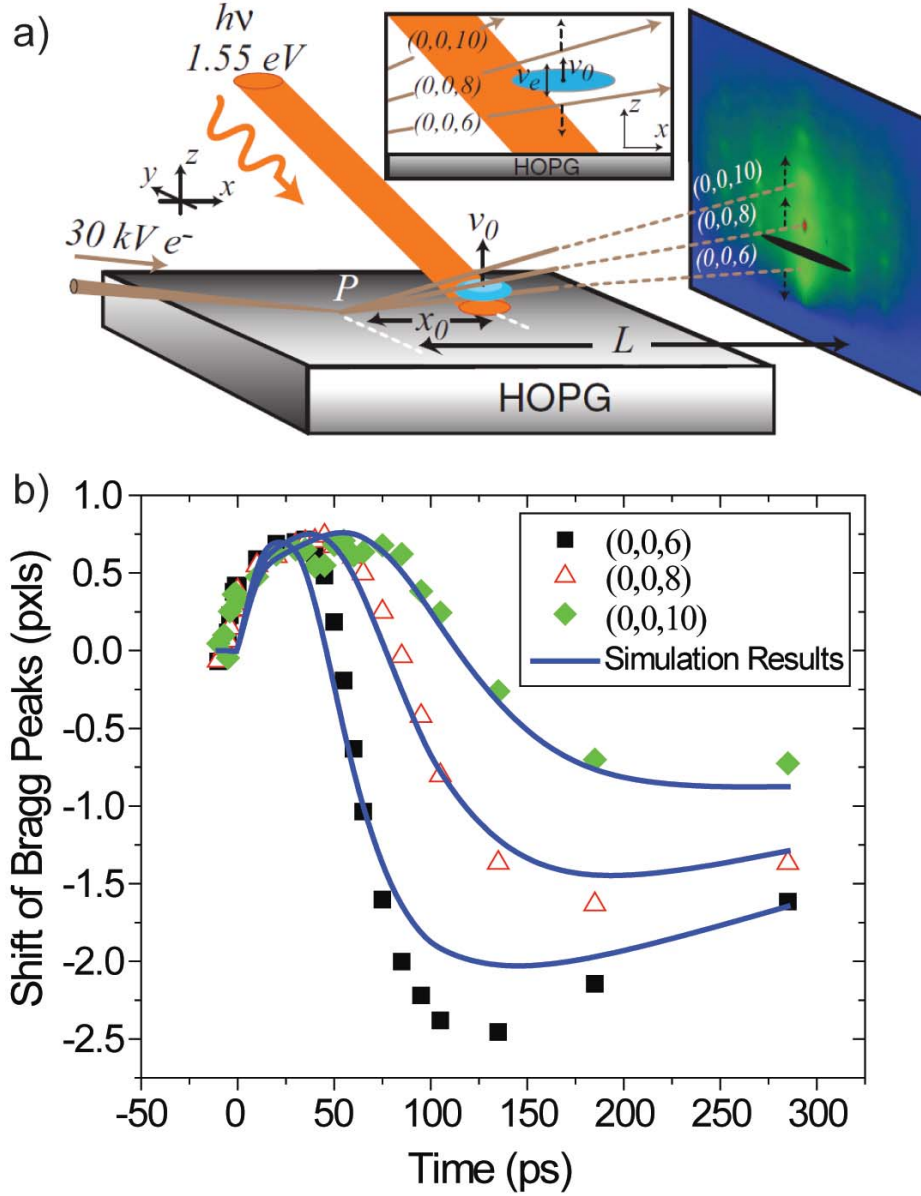


Figure D.4: a) Schematics of the PSImg technique [158]. Bragg peaks from HOPG sample cross the photoelectron pulse above the HOPG surface, as shown in the inset. The dashed arrows illustrate the effects of the photoelectron pulse applying on the Bragg beams. b) Bragg beams' movements as a function of time delays. The symbols are the experimental results. The solid lines are the results of the Bragg-beam deflection simulation with the electron sheet density $\Sigma = 8.0 \times 10^{12} \text{ m}^{-2}$.

REFERENCES

REFERENCES

- [1] A. H. Zewail. The new age of structural dynamics. *Acta Cryst. Sect. A*, A66:153–156, 2010.
- [2] M. Chergui and A. H. Zewail. Electron and X-Ray Methods of Ultrafast Structural Dynamics: Advances and Applications. *Chem. Phys. Chem.*, 10:28 43, 2009.
- [3] A. H. Zewail. Micrographia of the twenty-first century: from camera obscura to 4D microscopy. *Phil. Trans. R. Soc. A*, 368:1191–1204, 2010.
- [4] A. W. Dweydari and C. H. B Mee. Work Function Measurements on (100) and (110) Surfaces of Silver. *Phys. Stat. Sol. (a)*, 27:223–230, 1975.
- [5] R. K. Raman. *Electronically Induced Structure Transformations in Graphite & Silver, Studied Using Ultrafast Electron Crystallography*. PhD thesis, Michigan State university, 2010.
- [6] C. Y. Ruan, Y. Murooka, R. K. Raman, R. A. Murdick, R. J. Worhatch, and A. Pell. The Development and Applications of Ultrafast Electron Nanocrystallography. *Microscopy and Microanalysis*, 15:323–337, 2009.
- [7] D. B. Williams and C. B. Carter. *Transmission Electron Microscopy: A Textbook for Materials Science*. Springer, 2009.
- [8] P. J. Brown, A. G. Fox, E. N. Maslen, M. A. O’Keefe, and B. T. M. Willis. 6.1. Intensity of diffracted intensities. *International Tables for Crystallography*, C:554–590, 2006.
- [9] H. Bethe. Theorie der Beugung von Elektronen an Kristallen. *Ann. Phys.*, 392:55 – 129, 1928.
- [10] M. Blackman. On the Intensities of Electron Diffraction Rings. *Proc. R. Soc. London, Ser. A*, 173:68 – 82, 1939.
- [11] S. Speer, J. C. H. Spence, and E. Ihrig. On differentiation of the scattering matrix in dynamical transmission electron diffraction. *Acta. Cryst.*, A46:763–772, 1990.
- [12] J. M. Zuo and J. C. H. Spence. *Electron Microdiffraction* . Springer, 1992. 257
- [13] B. McNeil. Free electron lasers: First light from hard X-ray laser. *Nature Photonics*, 3:375 – 377, 2009.
- [14] B. J. Siwick, J. R. Dwyer, R. E. Jordan, and Miller R. J. D. Ultrafast electron

optics: Propagation dynamics of femtosecond electron packets. *J. Appl. Phys.* 92:1643, 2002.

- [15] M. Harb, W. Peng, G. Sciaini, C. T. Hebeisen, R. Ernstorfer, M. A. Eriksson, M. G. Lagally, S. G. Kruglik, and R. J. D. Miller. Excitation of longitudinal and transverse coherent acoustic phonons in nanometer free-standing films of (001) Si. *Phys. Rev. B*, 79:94301, 2009.
- [16] C. T. Hebeisen, G. Sciaini, M. Harb, R. Ernstorfer, T. Dartigalongue, S. G. Kruglik, and R. J. D. Miller. Grating enhanced ponderomotive scattering for visualization and full characterization of femtosecond electron pulses. *Opt. Express*, 16:3334 – 3341, 2008.
- [17] C. Gerbig, S. Morgenstern, C. Sarpe, M. Wollenhaupt, and T. Baumert. Femtosecond Transmission Electron Diffraction on Single Crystalline Graphite. *Research in Optical Sciences*, 2012.
- [18] J. D. Geiser and P. M. Weber. High-repetition-rate time-resolved gas phase electron diffraction. *SPIE Proceeding*, 2521:136 – 144, 1995.
- [19] V. A. Lobastov, R. Srinivasan, and A. H. Zewail. Four-dimensional ultrafast electron microscopy. *Proceedings of the National Academy of Sciences*, 102:70697073, 2005.
- [20] Michael S. Grinolds, Vladimir A. Lobastov, Jonas Weissenrieder, and Ahmed H. Zewail. Four-dimensional ultrafast electron microscopy of phase transitions. *Proceedings of the National Academy of Sciences*, 103:18427 – 18431, 2006.
- [21] M. Aidelsburger, F. O. Kirchner, F. Krausz, and P. Baum. Single-electron pulses for ultrafast diffraction. *Proceedings of the National Academy of Sciences*, 107:1971419719, 2010.
- [22] E. Fill, L. Veisz, A. Apolonski, and F. Krausz. Sub-fs electron pulses for ultrafast electron diffraction. *New J. Phys.*, 8:272, 2006.
- [23] H. S. Park, J. S. Baskin, O. H. Kwon, and A. H. Zewail. Atomic-Scale Imaging in Real and Energy Space Developed in Ultrafast Electron Microscopy. *Nano Lett.*, 7:2545 – 2551, 2007.
- [24] B. Barwick, D. J. Flannigan, and A. H. Zewail. Photon-induced near-field electron microscopy. *Nature*, 462:902 – 906, 2009.
- [25] A. Rose. Television Pickup Tubes and the Problem of Vision. *Adv. Electron.*, 1:131 – 166, 1948.
- [26] H. D'omer and O. Bostanjoglo. High-speed transmission electron microscope.

Rev. Sci. Instrum., 74:4369 – 4372, 2003.

- [27] T. LaGrange, K. Armstrong, M. R. and Boyden, C. G. Brown, G. H. Campbell, J. D. Colvin, W. J. DeHope, A. M. Frank, D. J. Gibson, F. V. Hartemann, J. S. Kim, W. E. King, B. J. Pyke, B. W. Reed, M. D. Shirk, R. M. Shuttlesworth, B. C. Stuart, B. R. Torralva, and N. D. Browning. Single-shot dynamic transmission electron microscopy. *Appl. Phys. Lett.*, 89:044105, 2006.
- [28] T. LaGrange, G. H. Campbell, B. W. Reed, M. Taheri, J. B. Pesavento, Kim J. S., and N. D Browning. Nanosecond time-resolved investigations using the in situ of dynamic transmission electron microscope (DTEM). *Ultramicroscopy*, 108:14411449, 2008.
- [29] B. W. Reed, M. R. Armstrong, N. D. Browning, G. H. Campbell, J. E. Evans, T. LaGrange, and D. J. Masiel. The Evolution of Ultrafast Electron Microscope Instrumentation. *Microscopy and Microanalysis*, 15:272 – 281, 2009.
- [30] T. LaGrange, B. Reed, W. DeHope, R. Shuttlesworth, and G. Huete. Movie Mode Dynamic Transmission Electron Microscopy (DTEM): Multiple Frame Movies of Transient States in Materials with Nanosecond Time Resolution. *Microscopy and Microanalysis*, 17:458 – 459, 2011.
- [31] T. T. Han, Z. Tao, S. D. Mahanti, K. Chang, C. D. Ruan, C. Y. and Malliakas, and M. G. Kanatzidis. Structural dynamics of two-dimensional charge-density waves in CeTe₃ investigated by ultrafast electron crystallography. *Phys. Rev. B*, 86:075145, 2012.
- [32] M. Eichberger, H. Schöfer, M. Krumova, M. Beyer, J. Demsar, H. Berger, G. Moriena, G. Sciaini, and R. J. D. Miller. Snapshots of cooperative atomic motions in the optical suppression of charge density waves. *Nature*, 468:799802, 2010.
- [33] X. J. Wang, Z. Wu, and H. Ihee. Femto-seconds Electron Beam Diffraction using Photocathode RF Gun. *Proc. PAC'03*, 2003.
- [34] J. B. Hastings, F. M. Rudakov, D. H. Dowell, J. F. Schmerge, J. D. Cardoza, J. M. Castro, S. M. Gierman, H. Loos, and P. M. Weber. Ultrafast time-resolved electron diffraction with megavolt electron beams. *Appl. Phys. Lett.*, 89:184109, 2006.
- [35] S. J. Park, J. H. Park, Y. W. Parc, C. B. Kim, C. D. Park, J. S. Oh, J. Y. Huang, I. S. Ko, X. Ding, and X. Wang. PHYSICS AND ENGINEERING ISSUES OF PPI (POHANG PHOTO-INJECTOR) FOR PAL XFEL. *Proc. 27th FEL Conf.*, 2005.
- [36] P. Musumeci, J. T. Moody, C. M. Scoby, M. S. Gutierrez, and T. Tran. Rf streak camera based ultrafast relativistic electron diffraction. *Rev. Sci. Instrum.*,

- 80:13302–1 – 13302–4, 2009.
- [37] P. Musumeci, J. T. Moody, C. M. Scoby, M. S. Gutierrez, H. A. Bender, and N. S. Wilcox. High quality single shot diffraction patterns using ultrashort megaelectron volt electron beams from a radio frequency photoinjector. *Rev. Sci. Instrum.*, 81:13306–1 – 13306–3, 2010.
- [38] X. Wang. MeV Ultrafast Electron Diffraction. *Research in Optical Sciences*, 2012.
- [39] R. Li, C. Tang, Y. Du, W. Huang, Q. Du, J. Shi, L. Yan, and X. Wang. Experimental demonstration of high quality MeV ultrafast electron diffraction. *Rev. Sci. Instrum.*, 80:083303, 2009.
- [40] J. Yang, K. Kan, T. Kondoh, Y. Yoshida, K. Tanimura, and J. Urakawa. Femtosecond pulse radiolysis and femtosecond electron diffraction. *Nuclear Instruments and Methods in Physics Research A*, 637:S24 – S29, 2011.
- [41] R. G. Carter. RF power generation. *CAS - CERN Accelerator School: RF for Accelerators*, 2011.
- [42] O. J. Luiten, S. B. van der Geer, M. J. de Loos, F. B. Kiewiet, and M. J. van der Wiel. How to Realize Uniform Three-Dimensional Ellipsoidal Electron Bunches. *Phys. Rev. Lett.*, 93:94802, 2004.
- [43] R. P. Chatelain, V. R. Morrison, C. Godbout, and B. J. Siwick. Ultrafast electron diffraction with radio-frequency compressed electron pulses. *Appl. Phys. Lett.*, 101:081901, 2012.
- [44] K. Chang. *Development and Application of RF-Enabled High-Flux Ultrafast Electron Microscope*. PhD thesis, Michigan State university, 2014.
- [45] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Cengage Learning, 1976.
- [46] C. Kittel. *Introduction to Solid State Physics (8th edition)*. Wiley, 2004.
- [47] W. A. Harrison. *Electronic Structure and the Properties of Solids*. Dover Publications, 1989.
- [48] P. Phillips. *Advanced Solid State Physics*. Perseus Books, 2008.
- [49] R. O. Jones and O. Gunnarsson. The density functional formalism, its applications and prospects. *Rev. Mod. Phys.*, 61:689 – 746, 1989.
- [50] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136:B864, 1964.

- [51] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140:A1133, 1965.
- [52] M. Imada, A. Fugimori, and Y. Tokura. Metal-insulator transitions. *Rev. Mod. Phys.*, 70(4):1039–1263, 1998.
- [53] N. F. Mott and R. Peierls. Discussion of the paper by de Boer and Verwey. *Proc. R. Phys. Soc.*, 49:72, 1937.
- [54] P. W. Anderson. New Approach to the Theory of Superexchange Interactions. *Phys. Rev.*, 115:2–13, 1959.
- [55] J. Hubbard. Electron Correlations in Narrow Energy Bands. *Proc. R. Soc. London, Ser. A*, 276:238–257, 1963.
- [56] Meinders, M. B. J. and Eskes, H. and Sawatzky, G. A. Spectral-weight transfer: Breakdown of low-energy-scale sum rules in correlated systems. *Phys. Rev. B*, 48:3916 – 3926, 1993.
- [57] J. B. Goodenough. The Two Components of the Crystallographic Transition in VO₂. *Solid State Chem.*, 3:490–500, 1971.
- [58] T. C. Koethe, Z. Hu, M.W. Haverkort, C. F. Schüßler-Langeheine, C. Venturini, N. B. Brookes, O. Tjernberg, W. Reichelt, H. H. Hsieh, H.-J. Lin, C. T. Chen, and L. H. Tjeng. Transfer of Spectral Weight and Symmetry across the Metal-Insulator Transition in VO₂. *Phys. Rev. Lett.*, 97:116402, 2006.
- [59] B. Himmetoglu, A. Floris, S. de Gironcoli, and M. Cococcioni. Hubbard-Corrected DFT Energy Functionals: The LDA+U Description of Correlated Systems. *International Journal of Quantum Chemistry*, 114:14–49, 2014.
- [60] V. Eyert. The metal-insulator transitions of VO₂: A band theoretical approach. *Ann. Phys. (Leipzig)*, 11:650–704, 2002.
- [61] A. Liebsch, H. Ishida, and G. Bihlmayer. Coulomb correlations and orbital polarization in the metal-insulator transition of VO₂. *Phys. Rev. B*, 71:85109, 2005.
- [62] Ed Caruthers and Leonard Kleinman. Energy Bands of Semiconducting VO₂. *Phys. Rev. B*, 7:3760 – 3766, 1973.
- [63] N. A. Skorikov M. A. Korotin and V. I. Anisimov. Variation of orbital symmetry of the localized 3d¹ electron of the V⁴⁺ ion upon the metal-insulator transition in VO₂. *Phys. Met. Metallogr.*, 94:17 – 23, 2002.
- [64] A. Georges and G. Kotliar. Hubbard model in infinite dimensions. *Phys. Rev. B*, 45:6479 – 6483, 1992.

- [65] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.*, 68:13 – 125, 1996.
- [66] G. Kotliar and D. Vollhardt. Strongly Correlated Materials: Insights from Dynamical Mean-field Theory. *Physics Today*, March 2004.
- [67] P. Weiss and P. Curie. L’hypothèse du champ moléculaire et la propriété ferromagnétique. *J. Phys. Theor. Appl.*, 6:661–690, 1907.
- [68] P. W. Anderson. Localized Magnetic States in Metals. *Phys. Rev.*, 124:41 – 53, 1961.
- [69] S. Biermann, A. Poteryaev, A. I. Lichtenstein, and A. Georges. Dynamical Singlets and Correlation-Assisted Peierls Transition in VO₂. *Phys. Rev. Lett.*, 94:026404, 2005.
- [70] M.W. Haverkort, Z. Hu, A. Tanaka, W. Reichelt, S.V. Streltsov, M. A. Korotin, V. I. Anisimov, H. H. Hsieh, H.-J. Lin, C. T. Chen, D. I. Khomskii, and L. H. Tjeng. Orbital-Assisted Metal-Insulator Transition in VO₂. *Phys. Rev. Lett.*, 95:196404, 2002.
- [71] M. M. Qazilbash, M. Brehm, B. G. Chae, P.-C. Ho, G. O. Andreev, B. J. Kim, S. J. Yun, A. V. Balatsky, Maple, F. Keilmann, H. T. Kim, and Basov D. N. Mott Transition in VO₂ Revealed by Infrared Spectroscopy and Nano-Imaging. *Science*, 318:1750–1753, 2007.
- [72] J. Wu, Q. Gu, B. S. Guiton, N. Leon, L. Ouyang, and H. Park. Strain-Induced Self Organization of Metal-Insulator Domains in Single-Crystalline VO₂ Nanobeams. *Nano Lett.*, 6(10):2313–2317, 2006.
- [73] M. F. Becker, A. B. Buckman, R. M. Walser, T. Lépine, P. Georges, and A. Brun. Femtosecond laser excitation of the semiconductor-metal phase transition in VO₂. *Appl. Phys. Lett.*, 65:1507–1509, 1994.
- [74] A. Cavalleri, Th. Dekorsy, H. H. W. Chong, J. C. Kieffer, and R. W. Schoenlein. Evidence for a structurally-driven insulator-to-metal transition in VO₂: A view from the ultrafast timescale. *Phys. Rev. B*, 70:161102, 2004.
- [75] G.H. Jonker and J.H. Van Santen. Ferromagnetic compounds of manganese with perovskite structure. *Physica*, 16:337 – 349, 1950.
- [76] R.M. Kusters, J. Singleton, D.A. Keen, R. McGreevy, and W. Hayes. Magnetoresistance measurements on the magnetic semiconductor Nd_{0.5}Pb_{0.5}MnO₃. *Physica B: Condensed Matter*, 155:362 – 365, 1989.

- [77] E. Dagotto, T. Hotta, and A. Moreo. Colossal Magnetoresistant materials: the Key Role of Phase Separation. *Phys. Rep.*, 155:1–153, 2001.
- [78] S. Jin, T. H. Tiefel, M. McCormack, R. A. Fastnacht, R. Ramesh, and L. H. Chen. Thousandfold Change in Resistivity in Magnetoresistive La-Ca-Mn-O Films. *Science*, 264:413, 1994.
- [79] C. G. Xiong, Q. Li, L. H. Ju, S. N. Mao, L. Senapati, X. X. Xi, R. L. Greene, and T. Venkatesan. Giant magnetoresistance in epitaxial $\text{Nd}_{0.7}\text{Sr}_{0.3}\text{MnO}_{3-\delta}$ thin films. *Appl. Phys. Lett.*, 66:1427, 1995.
- [80] M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas. Giant Magnetoresistance of (001)Fe/(001)Cr Magnetic Superlattices. *Phys. Rev. Lett.*, 61:2472 – 2475, 1988.
- [81] Clarence Zener. Interaction between the d -Shells in the Transition Metals. II. Ferromagnetic Compounds of Manganese with Perovskite Structure. *Phys. Rev.*, 82:403 – 405, 1951.
- [82] P. W. Anderson and H. Hasegawa. Considerations on Double Exchange. *Phys. Rev.*, 100:675 – 681, 1955.
- [83] A. J. Millis, P. B. Littlewood, and B. I. Shraiman. Double Exchange Alone Does Not Explain the Resistivity of $\text{La}_{1-x}\text{Sr}_x\text{MnO}_3$. *Phys. Rev. Lett.*, 74:5144 – 5147, 1995.
- [84] A. J. Millis, Boris I. Shraiman, and R. Mueller. Dynamic Jahn-Teller Effect and Colossal Magnetoresistance in $\text{La}_{1-x}\text{Sr}_x\text{MnO}_3$. *Phys. Rev. Lett.*, 77:175 – 178, 1996.
- [85] J. Bardeen. Theory of the Meissner Effect in Superconductors. *Phys. Rev.*, 97:1724 – 1725, 1955.
- [86] Leon N. Cooper. Bound Electron Pairs in a Degenerate Fermi Gas. *Phys. Rev.*, 104:1189 – 1190, 1956.
- [87] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Microscopic Theory of Superconductivity. *Phys. Rev.*, 106:162 – 164, 1957.
- [88] J.G. Bednorz and K. A. Müller. Possible high TC superconductivity in the Ba-La-Cu-O system. *Zeitschrift für Physik B*, 64:189193, 1986.
- [89] P. W. Anderson. The Resonating Valence Bond State in La_2CuO_4 and Superconductivity. *Science*, 235:1196–1198, 1987.

- [90] Gabriel Kotliar and Jialin Liu. Superexchange mechanism and d -wave superconductivity. *Phys. Rev. B*, 38:5142 – 5145, 1988.
- [91] P. W. Anderson, P. A. Lee, M. Randeria, T. M. Rice, N. Trivedi, and F. C. Zhang. The physics behind high-temperature superconducting cuprates: the 'plain vanilla' version of RVB. *J. Phys.: Condens. Matter*, 16:R755–R769, 2004.
- [92] P. A. Lee, N. Nagaosa, and X.-G. Wen. Doping a Mott insulator: Physics of high-temperature superconductivity. *Rev. Mod. Phys.*, 78:17 – 84, 2006.
- [93] Norman M. R. and Pepin C. The electronic nature of high temperature cuprate superconductors. *Rep. Prog. Phys.*, 66:15471610, 2003.
- [94] A. Schilling, M. Cantoni, J. D. Guo, and H. R. Ott. Superconductivity above 130 K in the Hg-Ba-Ca-Cu-O system. *Nature*, 363:56–58, 1993.
- [95] Y. Kamihara, H. Hiramatsu, M. Hirano, R. Kawamura, H. Yanagi, T. Kamiya, and H. Hosono. Iron-Based Layered Superconductor: LaOFeP. *J. Am. Chem. Soc.*, 128:10012 – 10013, 2006.
- [96] C. Chen and Z. Zhou. Optical phonons assisted infrared absorption in VO₂ based bolometer. *Appl. Phys. Lett.*, 91:011107, 2007.
- [97] T. Driscoll, H.-T. Kim, B.-G. Chae, M. D. Ventra, and D. N. Basov. Phase transition driven memristive system. *Appl. Phys. Lett.*, 95:043503, 2009.
- [98] T. Driscoll, S. Palit, M. M. Qazilbash, M. Brehm, F. Keilmann, B.-G. Chae, S.-J. Yun, H.-T. Kim, S. Y. Cho, N. M. Jokerst, D. R. Smith, and D. N. Basov. Dynamic tuning of an infrared hybrid-metamaterial resonance using vanadium dioxide. *Appl. Phys. Lett.*, 93:024101, 2008.
- [99] M.-J. Lee, Y. Park, D.-S. Suh, E.-H. Lee, S. Seo, D.-C. Kim, R. Jung, B.-S. Kang, S.-E. Ahn, C. B. Lee, D. H. Seo, Y.-K. Cha, I.-K. Yoo, J.-S. Kim, and B. H. Park. Two Series Oxide Resistors Applicable to High Speed and High Density Nonvolatile Memory. *Adv. Mater.*, 19:3919–3923, 2007.
- [100] T. D. Manning, I. Parkin, M. E. Pemble, D. Sheel, and D. Vernardou. Intelligent Window Coatings: Atmospheric Pressure Chemical Vapor Deposition of Tungsten-Doped Vanadium Oxide. *Chem. Mater.*, 16:744–749, 2004.
- [101] C. Tang, P. Georgopoulos, M. E. Fine, J. B. Cohen, M. Nygren, G. S. Knapp, and A. Aldred. Local Atomic and Electronic Arrangements in W_xV_{1-x}O₂. *Phys. Rev. B*, 31:1000–1011, 1985.
- [102] B. G. Chae, H. T. Kim, and S. J. Yun. Characteristics of W- and Ti-Doped VO₂ Thin Films Prepared by Sol-Gel Method. *Electrochem. Solid-State Lett.*,

- 11:D53–D55, 2008.
- [103] J. P. Pouget, H. Launois, J. P. D’Haenens, P. Merenda, and T. M. Rice. Electron Localization Induced by Uniaxial Stress in Pure VO₂. *Phys. Rev. Lett.*, 35:873–875, 1975.
- [104] J. Cao, E. Ertekin, V. Srinivasan, W. Fan, S. Huang, H. Zheng, J. W. L. Yim, D. R. Khanal, D. F. Ogletree, J. C. Grossman, and J. Wu. Strain engineering and one-dimensional organization of metalinsulator domains in single-crystal vanadium dioxide beams. *Nature Nanotechnology*, 4:732–737, 2009.
- [105] V. Eyert. VO₂: A Novel View from Band Theory. *Phys. Rev. Lett.*, 107:016407, 2011.
- [106] C. Weber, D. D. ORegan, N. D. M. Hine, M. C. Payne, G. Kotliar, and Littlewood P. B. Vanadium Dioxide: A Peierls-Mott Insulator Stable against Disorder. *Phys. Rev. Lett.*, 108:256402, 2012.
- [107] M. Rini, A. Cavalleri, R. W. Schoenlein, R. Lpez, L. C. Feldman, R. F. Haglund, Jr., L. A. Boatner, and T. E. Haynes. Photoinduced phase transition in VO₂ nanocrystals: ultrafast control of surface-plasmon resonance. *Opt. Lett.*, 30:558 – 560, 2005.
- [108] M. Hada, D. Zhang, A. Casandruc, R. J. D. Miller, Y. Hontani, J. Matsuo, R. E. Marvel, and R. F. Haglund. Hot electron injection driven phase transitions. *Phys. Rev. B*, 86:134101, 2012.
- [109] A. Cavalleri, H. H. W. Chong, S. Fourmaux, T. E. Glover, P. A. Heimann, J. C. Kieffer, B. S. Mun, H. A. Padmore, and R. W. Schoenlein. Picosecond soft xray absorption measurement of the photoinduced insulator-to-metal transition in VO₂. *Phys. Rev. B*, 69:153106, 2004.
- [110] A. Cavalleri, Cs. Tóth, C. W. Siders, J. A. Squier, F. Ráksi, P. Forget, and J. C. Kieffer. Femtosecond Structural Dynamics in VO₂ during an Ultrafast Solid-Solid Phase Transition. *Phys. Rev. Lett.*, 87:237401, 2001.
- [111] A. Cavalleri, M. Rini, H. H.W. Chong, S. Fourmaux, T. E. Glover, P. A. Heimann, J. C. Kieffer, and R.W. Schoenlein. Band-Selective Measurements of Electron Dynamics in VO₂ Using Femtosecond Near-Edge X-Ray Absorption. *Phys. Rev. Lett.*, 95:067405, 2005.
- [112] M. Hada, K. Okimura, and J. Matsuo. Characterization of structural dynamics of VO₂ thin film on c-Al₂O₃ using in-air time-resolved x-ray diffraction. *Phys. Rev. B*, 82:153401, 2010.
- [113] H. Wen, L. Guo, E. Barnes, J. H. Lee, D. A. Walko, R. D. Schaller, J. A. Moyer, R. Misra, Y. Li, E. M. Dufresne, D. G. Schlom, V. Gopalan, and J. W. Freeland.

Structural and electronic recovery pathways of a photoexcited ultrathin VO₂ film. *Phys. Rev. B*, 88:165424, 2013.

- [114] C. Kübler, H. Ehrke, R. Huber, R. Lopez, A. Halabica, R. F. Jr. Haglund, and A. Leitenstorfer. Coherent Structural Dynamics and Electronic Correlations during an Ultrafast Insulator-to-Metal Phase Transition in VO₂. *Phys. Rev. Lett.*, 99:116401, 2007.
- [115] A. Pashkin, C. Kübler, H. Ehrke, R. Lopez, A. Halabica, R. F. Haglund Jr., R. Huber, and A. Leitenstorfer. Ultrafast insulator-metal phase transition in VO₂ studied by multiterahertz spectroscopy. *Phys. Rev. B*, 83:195120, 2011.
- [116] M. Liu, H. Y. Hwang, H. Tao, A. C. Strikwerda, K. Fan, G. R. Keiser, A. J. Sternbach, K. G. West, S. Kittiwatanakul, J. Lu, S. A. Wolf, F. G. Omenetto, X. Zhang, K. A. Nelson, and R. D. Averitt. Terahertz-field-induced insulator-to-metal transition in vanadium dioxide metamaterial. *Nature*, 487:345 – 348, 2012.
- [117] E. Abreu, M. Liu, J. Lu, K. G. West, S. Kittiwatanakul, W. Yin, S. W. Wolf, and R. D. Averitt. THz spectroscopy of VO₂ epitaxial films: controlling the anisotropic properties through strain. *New J. Phys.*, 14:083026, 2012.
- [118] P. Baum, D. Sh. Yang, and A. H. Zewail. 4D Visualization of Transitional Structures in Phase Transformations by Electron Diffraction. *Science*, 318:788, 2007.
- [119] Z. Tao, T. R. T. Han, S. D. Mahanti, P. M. Duxbury, F. Yuan, C. Y. Ruan, K. Wang, and J. Wu. Decoupling of Structural and Electronic Phase Transitions in VO₂. *Phys. Rev. Lett.*, 109:166406, 2012.
- [120] F. J. Morin. Oxides Which Show a Metal-to-Insulator Transition at the Neel Temperature. *Phys. Rev. Lett.*, 3:3436, 1959.
- [121] L. A. Ladd and W. Paul. Optical and transport properties of high quality crystals of V₂O₄ near the metallic transition temperature. *Solid State Commu.*, 7:425–428, 1969.
- [122] G. Anderson. Crystal Structure of Vanadium Dioxide. *Acta Chem. Scand.*, 10:623–628, 1956.
- [123] S. Westman. Note on a phase transition in VO₂. *Acta Chem. Scand.*, 15:217, 1961.
- [124] J. M. Longo and P. Kierkegaard. Refinement of the structure of VO₂. *Acta Chem. Scand.*, 24(1):420–426, 1970.

- [125] A. Rúa, F. E. Fernández, and N. Sepálveda. Bending in VO₂-coated microcantilevers suitable for thermally activated actuators. *J. Appl. Phys.*, 107:074506–074506–4, 2010.
- [126] Hans W. Verleur, A. S. Barker, and C. N. Berglund. Optical Properties of VO₂ between 0.25 and 5 eV. *Phys. Rev.*, 172:788 – 798, 1968.
- [127] M. Mulato, I. Chambouleyron, E. G. Birgin, and J. M. Martínez. Determination of thickness and optical constants of amorphous silicon films for transmittance data. *Appl. Phys. Lett.*, 77:2133, 2000.
- [128] R. W. Wood. *Physical Optics*. Macmillin Co, 1936.
- [129] E. D. Palik. *Handbook of Optical Constants of Solids*. Academic Press, 1985.
- [130] T. L. Cocker, L. V. Titova, S. Fourmaux, G. Holloway, H.-C. Bandulet, D. Brassard, J. C. Kieffer, M. A. El Khakani, and F. A. Hegmann. Phase diagram of the ultrafast photoinduced insulator-metal transition in vanadium dioxide. *Phys. Rev. B*, 85:155120, 2012.
- [131] T. L. Cocker, L. V. Titova, S. Fourmaux, H. C. Bandulet, D. Brassard, J. C. Kieffer, M. A. El Khakani, and F. A. Hegmann. Terahertz conductivity of the metalinsulator transition in a nanogranular VO₂ film. *Appl. Phys. Lett.*, 97:221905, 2010.
- [132] G. V. Chandrashekhar, H. L. C Barros, and J. M. Honig. Heat capacity of VO₂ single crystals. *Mater. Res. Bull.*, 8:369–374, 1973.
- [133] C. N. Berglund and H. J. Guggenheim. Electronic Properties of VO₂, near the Semiconductor-Metal Transition. *Phys. Rev.*, 185:1022, 1969.
- [134] A. S. Barker, H. W. Verleur, and H. J. Guggenheim. Infrared Optical Properties of Vanadium Dioxide Above and Below the Transition Temperature. *Phys. Rev. Lett.*, 17:1286 – 1289, 1966.
- [135] H. Dachraoui, N. Müller, G. Obermeier, C. Oberer, S. Horn, and U. Heinzmann. Interplay between electronic correlations and coherent structural dynamics during the monoclinic insulator-to-rutile metal phase transition in VO₂. *J. Phys. Condens. Matter.*, 23:435402, 2011.
- [136] R. Lopez, L. C. Feldman, and Jr. Haglund, R. F. Size-Dependent Optical Properties of VO₂ Nanoparticle Arrays. *Phys. Rev. Lett.*, 93:177403, 2004.
- [137] L. Whittaker, C. Jaye, Z. Fu, D. A. Fischer, and S. Banerjee. Depressed Phase Transition in Solution-Grown VO₂ Nanostructures. *J. Am. Chem. Soc.*, 131(25):8884–8894, 2009.

- [138] J. Wei, Z. Wang, W. Chen, and D. H. Cobden. New Aspects of the Metal-Insulator Transition in Single-Domain Vanadium Dioxide Nanobeams. *Nature Nanotechnology*, 4:420–424, 2009.
- [139] L. Whittaker, C. J. Patridge, and S. Banerjee. Microscopic and Nanoscale Perspective of the Metal-Insulator Phase Transitions of VO₂: Some New Twists to an Old Tale. *J. Phys. Chem. Lett.*, 2:745–758, 2011.
- [140] R. M. Wentzcovitch, W. W. Schulz, and P. B. Allen. VO₂: Peierls or Mott-Hubbard? A View from Band Theory. *Phys. Rev. Lett.*, 72:3389–3392, 1994.
- [141] J. P. Pouget, H. Launois, T. M. Rice, P. Dernier, A. Gossard, G. Villeneuve, and P. Hagenmuller. Dimerization of a linear Heisenberg chain in the insulating phases of V_{1-x}Cr_xO₂. *Phys. Rev. B*, 10:1801–1815, 1974.
- [142] A. Zylbersztein and N. F. Mott. Metal-insulator transition in vanadium dioxide. *Phys. Rev. B*, 11:4383–4395, 1975.
- [143] M. Marezio, D.B. McWhan, J.P. Remeika, and P.D. Dernier. Structural Aspects of the Metal-Insulator Transitions in Cr-Doped VO₂. *Phys. Rev. B*, 5:2541–2551, 1972.
- [144] D. J. Hilton, R. P. Prasankumar, S. Fourmaux, A. Cavalleri, D. Brassard, M. A. El Khakani, J. C. Kieffer, A. J. Taylor, and R. D. Averitt. Enhanced Photosusceptibility near T_c for the Light-Induced Insulator-to-Metal Phase Transition in Vanadium Dioxide. *Phys. Rev. Lett.*, 99:2265401, 2007.
- [145] S.-K. Mo, Denlinger, H.-D. Kim, J.-H. Park, J.W. Allen, A. Sekiyama, A. Yamasaki, Kadono, S. Suga, Y. Saitoh, T. Muro, P. Metcalf, G. Keller, K. Held, V. Eyert, V. I. Anisimov, and D. Vollhardt. Prominent Quasiparticle Peak in the Photoemission Spectrum of the Metallic Phase of V₂O₃. *Phys. Rev. Lett.*, 90:186403, 2003.
- [146] T. Mueller, F. Xia, M. Freitag, T. Tsang, and Ph. Avouris. Role of contacts in graphene transistors: A scanning photocurrent study. *Phys. Rev. B*, 79:245430, 2009.
- [147] B. J. Kim, Y. W. Lee, S. Choi, J. W. Lim, S. J. Yun, H. T. Kim, T. J. Shin, and H. S. Yun. Micrometer x-ray diffraction study of VO₂ films: Separation between metal-insulator transition and structural phase transition. *Phys. Rev. B*, 77:235401, 2008.
- [148] M. K. Liu, M. Wagner, E. Abreu, S. Kittiwatanakul, A. McLeod, Z. Fei, M. Goldflam, S. Dai, M. M. Fogler, J. Lu, S. A. Wolf, R. D. Averitt, and D. N. Basov. Anisotropic Electronic State via Spontaneous Phase Separation in Strained Vanadium Dioxide Films. *Phys. Rev. Lett.*, 111:096602, 2013.

- [149] D. B. McWhan, M. Marezio, J.P. Remeika, and P.D. Dernier. X-ray diffraction study of metallic VO₂. *Phys. Rev. B*, 10:490–495, 1974.
- [150] J. Worster. The brightness of electron beam. *J. Phys. D: Appl. Phys.*, 2:457 – 462, 1968.
- [151] S. B. van der Geer, M. J. de Loos, E. J. D. Vredenburg, and O. J. Luiten. Ultracold Electron Source for Single-Shot, Ultrafast Electron Diffraction. *Microsc. Microanal.*, 15:282 – 289, 2009.
- [152] H. Hertz. Ueber einen Einfluss des ultravioletten Lichtes auf die electrische Entladung. *Annalen der Physik*, 267:9831000, 1887.
- [153] A. Einstein. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik*, 322:132148, 1905.
- [154] A. Damascelli, Z. Hussain, and Z. X. Shen. Angle-resolved photoemission studies of the cuprate superconductors. *Rev. Mod. Phys*, 75(4):473541, 2003.
- [155] F. J. M. Heringdorf, L. I. Chelaru, S. Möllenbeck, D. Thien, and M. H. Hoegen. Femtosecond photoemission microscopy. *Surf. Sci.* 604:4700-4705, 2007
- [156] H. N. Chapman, A. Barty, M. J. Bogan, S. Boutet, M. Frank, S. P. Hau-Riege, S. Marchesini, B. W. Woods, S. C. Bajt, W. H. Benner, R. A. London, E. Plönjes, M. Kuhlmann, R. Treusch, S. Düsterer, T. Tschentscher, E. Schneider, J. R. and Spiller, T. Möller, C. Bostedt, M. Hoener, D. A. Shapiro, K. O. Hodgson, D. van der Spoel, F. Burmeister, M. Bergh, C. Caleman, G. Hultdt, M. M. Seibert, F. R. N. C. Maia, R. W. Lee¹, A. Szöke, N. Timneanu, and J. Hajdu. Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nature Physics*, 2:839–843, 2006.
- [157] B. W. J. McNeil and N. R. Thompson. X-ray free-electron lasers. *Nature Photonics*, 4:814821, 2010.
- [158] R. K. Raman, Z. Tao, T. R. Han, and C. Y. Ruan. Ultrafast imaging of photoelectron packets generated from graphite surface. *Appl. Phys. Lett.*, 95:181108, 2009.
- [159] Z. Tao, H. Zhang, P. M. Duxbury, M. Berz, and C. Y. Ruan. Space charge effects in ultrafast electron diffraction and imaging. *J. Appl. Phys.*, 111:044316, 2012.
- [160] T. T. Han. Private communication, 2013.
- [161] J. R. Dwyer, C. T. Hebeisen, R. Ernstorfer, M. Harb, V. B. Deyirmenjian, R. E. Jordan, and R. J. D Miller. Femtosecond electron diffraction: ‘making the molecular movie’. *Phil. Trans. R. Soc. A*, 364:741778, 2006.

- [162] J. Cao, Z. Hao, H. Park, C. Tao, Kau D., and L. Blaszczyk. Femtosecond electron diffraction for direct measurement of ultrafast atomic motions. *Applied Physics Letters*, 83(5):1044–1046, 2003.
- [163] Ramesh Srinivasan, Vladimir A. Lobastov, Chong-Yu Ruan, and Ahmed H. Zewail. Ultrafast Electron Diffraction (UED) A New Development for the 4D Determination of Transient Molecular Structures. *HCA*, 86:1761 – 1799, 2003.
- [164] B. J. Siwick, J. R. Dwyer, R. E. Jordan, and R. J. D Miller. Femtosecond electron diffraction studies of strongly driven structural phase transitions. *Chemical Physics*, 299:285–305, 2004.
- [165] A. H. Zewail. Diffraction, crystallography and microscopy beyond three dimensions: structural dynamics in space and time. *Phil. Trans. R. Soc. A*, 363(1827):315–329, 2005.
- [166] T. van Oudheusden, P. L. E. M. Pasmans, S. B. van der Geer, M. J. de Loos, M. J. van der Wiel, and O. J. Luiten. Compression of Subrelativistic Space-Charge-Dominated Electron Bunches for Single-Shot Femtosecond Electron Diffraction. *Phys. Rev. Lett.*, 105:264801, 2010.
- [167] B. J. Siwick, A. A. Green, C. T. Hebeisen, and R. J. D Miller. Characterization of ultrashort electron pulses by electron-laser pulse cross correlation. *Opt. Lett.*, 30(9):1057–1059, 2005.
- [168] T. Srinivasan-Rao, J. Fischer, and T. Tsang. Photoemission studies on metals using picosecond ultraviolet laser pulses. *J. Appl. Phys.*, 69:3291 – 3296, 1991.
- [169] Xinrong Jiang, C. N. Berglund, Anthony E. Bell, and William A. Mackie. Photoemission from gold thin films for application in multi-photocathode arrays for electron beam lithography. *J. Vac. Sci. Technol. B*, 16:3374 – 3379, 1998.
- [170] L. M. Rangarajan and G. K. Bhide. Photoemission energy distribution studies of gold thin films under uv excitation by a photoelectron spectroscopic method. *Vacuum*, 30:515 – 522, 1980.
- [171] A. Janzen, B. Krenzer, O. Heinz, P. Zhou, D. Thien, A. Hanisch, F. J. Meyer zu Heringdorf, D. von der Linde, and M. Horn von Hoegen. A pulsed electron gun for ultrafast electron diffraction at surfaces. *Rev. Sci. Instrum.*, 78:13906–1 – 13906–7, 2007.
- [172] K. Chang, R. A. Murdick, Z. Tao, T-R T. Han, and C-Y Ruan. ULTRAFAST ELECTRON DIFFRACTIVE VOLTAMMETRY: GENERAL FORMALISM AND APPLICATIONS. *Mod. Phys. Lett. B*, 25:2099 – 2129, 2011.

- [173] Kevin L. Jensen, P. G. O'Shea, D. W. Feldman, and N. A. Moody. Theoretical model of the intrinsic emittance of a photocathode. *Appl. Phys. Lett.*, 89:224103–1 – 224103–3, 2006.
- [174] K. L. Jensen, P. G. O'Shea, D. W. Feldman, and J. L. Shaw. Emittance of a field emission electron source. *J. Appl. Phys.*, 107:14903–1 – 14903–14, 2010.
- [175] Bao-Liang Qian and Hani E. Elsayed-Ali. Electron pulse broadening due to space charge effects in a photoelectron gun for electron diffraction and streak camera systems. *J. Appl. Phys.*, 91:462 – 468, 2002.
- [176] Bryan W. Reed. Femtosecond electron pulse propagation for ultrafast electron diffraction. *J. Appl. Phys.*, 100:34916–1 – 34916–16, 2006.
- [177] A. M. Michalik and J. E. Sipe. Analytic model of electron pulse propagation in ultrafast electron diffraction experiments. *J. Appl. Phys.*, 99:54908–1 – 54908–9, 2006.
- [178] A. M. Michalik and J. E. Sipe. Evolution of non-Gaussian electron bunches in ultrafast electron diffraction experiments: Comparison to analytic model. *J. Appl. Phys.*, 105:84913–1 – 84913–10, 2009.
- [179] S. Hellmann, K. Rossnagel, M. Marczyński-Bühlow, and L. Kipp. Vacuum space charge effects in solid-state photoemission. *Phys. Rev. B*, 79:35402, 2009.
- [180] X.J. Zhou, B. Wannberg, W.L. Yang, V. Brouet, Z. Sun, J.F. Douglas, D. Dessau, Z. Hussain, and Z.-X. Shen. Space charge effect and mirror charge effect in photoemission spectroscopy. *Journal of Electron Spectroscopy and Related Phenomena*, 142:27 – 38, 2005.
- [181] S. Passlack, S. Mathias, O. Andreyev, D. Mitnacht, M. Aeschlimann, and M. Bauer. Space charge effects in photoemission with a low repetition, high intensity femtosecond laser source. *J. Appl. Phys.*, 100:24912–1 – 24912–8, 2006.
- [182] W. F. Krolikowski and W. E. Spicer. Photoemission Studies of the Noble Metals. II. Gold. *Phys. Rev. B*, 1:478 – 487, 1970.
- [183] Kevin L. Jensen, Donald W. Feldman, Nathan A. Moody, and Patrick G. O'Shea. A photoemission model for low work function coated metal surfaces and its experimental validation. *J. Appl. Phys.*, 99:124905–1 – 124905–19, 2006.
- [184] David H. Dowell and John F. Schmerge. Quantum efficiency and thermal emittance of metal photocathodes. *Phys. Rev. ST Accel. Beams*, 12:74201, 2009.
- [185] D. R. Lide. *CRC Handbook of Chemistry and Physics*, 74th ed. CRC, Boca Raton, FL, 1993.

- [186] Agust Valfells, D. W. Feldman, M. Virgo, P. G. O'Shea, and Y. Y. Lau. Effects of pulse-length and emitter area on virtual cathode formation in electron guns. *Phys. Plasmas*, 9:2377 – 2382, 2002.
- [187] P. Musumeci, J. T. Moody, R. J. England, J. B. Rosenzweig, and T. Tran. Experimental Generation and Characterization of Uniformly Filled Ellipsoidal Electron-Beam Distributions. *Phys. Rev. Lett.*, 100:244801, 2008.
- [188] J. A. Berger and W. A. Schroeder. Semianalytic model of electron pulse propagation: Magnetic lenses and rf pulse compression cavities. *J. Appl. Phys.*, 108:124905–1 – 124905–12, 2010.
- [189] Andreas Gahlmann, Sang Tae Park, and Ahmed H. Zewail. Ultrashort electron pulses for diffraction, crystallography and microscopy: theoretical and experimental resolutions. *Phys. Chem. Chem. Phys.*, 10:2894 – 2909, 2008.
- [190] German Sciaini, Maher Harb, Sergei G. Kruglik, Thomas Payer, Christoph T. Hebeisen, Frank-J. Meyer zu Heringdorf, Mariko Yamaguchi, Michael Horn-von Hoegen, Ralph Ernstorfer, and R. J. Dwayne Miller. Electronic acceleration of atomic motions and disordering in bismuth. *Nature*, 458:56 – 59, 2009.
- [191] A. H. Zewail and J. M. Thomas. *4D Electron Microscopy: Imaging in Space and Time*. Imperial College Press, London, 2010.
- [192] Judy S. Kim, Thomas LaGrange, Bryan W. Reed, Mitra L. Taheri, Michael R. Armstrong, Wayne E. King, Nigel D. Browning, and Geoffrey H. Campbell. Imaging of Transient Structures Using Nanosecond in Situ TEM. *Science*, 321:1472 – 1475, 2008.
- [193] J. M. Zuo, I. Vartanyants, M. Gao, R. Zhang, and L. A. Nagahara. Atomic Resolution Imaging of a Carbon Nanotube from Diffraction Intensities. *Science*, 300:1419 – 1421, 2003.
- [194] Knut W. Urban. Studying Atomic Structures by Aberration-Corrected Transmission Electron Microscopy. *Science*, 321:506 – 510, 2008.
- [195] Robert K. Shelton, Long-Sheng Ma, Henry C. Kapteyn, Margaret M. Murnane, John L. Hall, and Jun Ye. Phase-Coherent Optical Pulse Synthesis from Separate Femtosecond Lasers. *Science*, 293:1286 – 1289, 2001.
- [196] Steven T. Cundiff and Jun Ye. Colloquium: Femtosecond optical frequency combs. *Rev. Mod. Phys.*, 75:325 – 342, 2003.
- [197] Jun Ye, Long Sheng Ma, and John L. Hall. Molecular Iodine Clock. *Phys. Rev. Lett.*, 87:270801, 2001.

- [198] T. W. Hänsch. Application of high resolution laser spectroscopy. In A. Mooradain, T. Jaeger, and P. Stokseth, editors, *Application of high resolution laser spectroscopy*, page 326. Springer-Verlag, Berlin, 1976.
- [199] Y. V. Baklanov and V.P. Chebotayev. Narrow resonances of two-photon absorption of super-narrow pulses in a gas. *Applied Physics*, 12:97, 1977.
- [200] T. M. Fortier, M. S. Kirchner, F. Quinlan, J. Taylor, J. C. Bergquist, T. Rosenband, N. Lemke, A. Ludlow, Y. Jiang, C. W. Oates, and S. A. Diddams. Generation of ultrastable microwaves via optical frequency division. *Nature Photonics*, 5:425 – 429, 2011.
- [201] J. Portmann. Private communication, 2012.
- [202] W. J. Gruen. Theory of AFC Synchronization. *PROCEEDINGS OF THE I.R.E.*, 41:1043–1048, 1953.
- [203] D. H. Wolaver. *Phase-Locked Loop Circuit Design*. Prentice-Hall, Inc, 1991.
- [204] E.N. Ivanov, S.A. Diddams, and L. Hollberg. Analysis of noise mechanisms limiting the frequency stability of microwave signals generated with a femtosecond laser. *Selected Topics in Quantum Electronics, IEEE Journal of*, 9:1059 – 1065, 2003.
- [205] J. Kim, F.O. Ilday, A. Winter, and F.X. Kärtner. Timing distribution and RF synchronization with mode-locked lasers. *LEOS Summer Topical Meetings, 2005 Digest of the*, 2005.
- [206] J. Kim, F. X. Kärtner, and M. H. Perrott. Femtosecond synchronization of radio frequency signals with opticalpulse trains. *Opt. Lett.*, 29:2076 – 2078, 2004.
- [207] J. Kim, F. X. Kärtner, and F. Ludwig. Balanced optical-microwave phase detectors for optoelectronic phase-locked loops. *Opt. Lett.*, 31:3659 – 3661, 2006.
- [208] R. C. Dorf and R. H. Bishop. *Modern control systems*. Prentice-Hall, Inc, 2011. 272
- [209] D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls. *Characterization of Clocks and Oscillators*. NIST, 1990.
- [210] B. Drakhlis. Calculate Oscillator Jitter by using Phase-Noise Analysis, Part 1. *Microwaves and RF*, 2001.
- [211] B. Drakhlis. Calculate Oscillator Jitter by using Phase-Noise Analysis, Part 2. *Microwaves and RF*, 2001.

- [212] T. van Oudheusden, E. F. de Jong, S. B. van der Geer, W. P. E. M. Op't Root, O. J. Luiten, and B. J. Siwick. Electron source concept for single-shot sub-100 fs electron diffraction in the 100 keV range. *J. Appl. Phys.*, 102:093501, 2007.
- [213] J. H. Billen and L. M. Young. *Los Alamos National Laboratory Report NO. LA-UR-96-1834 (Posson Superfish)*. LANL, 1996.
- [214] T. P. Wangler. *RF Linear Accelerator*. Wiley-VCH, 2004.
- [215] N. Du and K. Chang. Private communication, 2014.
- [216] He Zhang and Martin Berz. The fast multipole method in the differential algebra framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 645:338 – 344, 2011.
- [217] T. Yao, X. D. Zhang, Z. H. Sun, S. J. Liu, Y. Y. Huang, Y. Xie, C. Z. Wu, X. Yuan, W. Q. Zhang, Z. Y. Wu, G. Q. Pan, F. C. Hu, L. H. Wu, Q. H. Liu, and S. Q. Wei. Understanding the Nature of the Kinetic Process in a VO₂ Metal-Insulator Transition. *Phys. Rev. Lett.*, 105:226405, 2010.