A SURVEY FOR HIGH-REDSHIFT GRAVITATIONALLY LENSED QUASARS

by

Minghao Yue

Copyright © Minghao Yue 2022

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF ASTRONOMY
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
WITH A MAJOR IN ASTRONOMY AND ASTROPHYSICS
In the Graduate College
THE UNIVERSITY OF ARIZONA

2022
As members of the Dissertation Committee, we certify that we have read the dissertation prepared by: Minghao Yue
titled: A Survey for High-redshift Gravitationally Lensed Quasars
and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Xiaohui Fan
Date: Jul 29, 2022
Ann Zabludoff
Date: Jul 28, 2022
Daniel Marrone
Date: Jul 28, 2022
George Rieke
Date: Jul 29, 2022
Richard Green
Date: Jul 28, 2022

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Xiaohui Fan
Date: Jul 29, 2022

Department of Astronomy
I would like to express my deepest appreciation to Prof. Xiaohui Fan, my PhD advisor, for his support and advises through the past years. Prof. Fan has been inspiring me with exciting ideas and providing me a wealth of research opportunities. He is so supportive and considerate in both my research and my life. I still remember when I first met him in 2015 when I visited Steward Observatory as an undergraduate student. After the first meeting with Prof. Fan, I knew that he was the advisor that I hoped to work with. My experience of working with Prof. Fan is so amazing that I will never forget in my life.

I would also like to sincerely appreciate Prof. Ann Zabludoff, Prof. George H. Rieke, Prof. Richard Green, Prof. Daniel P. Marrone, and Prof. Peter Behroozi. Prof. Zabludoff, Prof. Rieke, and Prof. Green have been on my mentoring committee, my prelim committee and my dissertation committee. They have provided me invaluable advices not only about this thesis, but also about my general research and my life as a graduate student. Prof. Marrone is on my dissertation committee, and Prof. Behroozi is on my prelim committee. I am very grateful for their help with my second-year project and my thesis project.

I would like to thank all my collaborators for their contributions in my research projects. Special thanks to Dr. Jinyi Yang and Dr. Feige Wang, who has helped me in many ways, from scientific suggestions to technical supports. I am also very grateful to the staff, technician, and astronomers at the Kitt Peak National Observatory, the Mount Graham International Observatory, the Large Binocular Telescope Observatory, the MMT Observatory, and the Las Campanas Observatory, for all the help I received when observing with their facilities.

Thanks should also go to Prof. Daniel P. Stark and Prof. Edward W. Olszewski. I was a teaching assistant in the courses they taught, from whom I learned a lot about how to work as a teacher. I would like to thank all the students in the classes as well for being nice and hard-working.

I am very grateful to all my colleagues at Steward for building a friendly and inclusive environment in the department. As an international student, I was unfamiliar with things in Tucson and at the University of Arizona when I arrived in 2016. People here at Steward are so warm and supportive, who really make a difference in my graduate life. Special thanks to Ms. Michelle Cournoyer, who is always there whenever I seek for help.

And I would like to thank Miss. Jiani Ding, for being my partner and for her support. Jiani is the person who understand my happiness and upsets; she knows well about the tough times I have experienced these years, and she is always there to encourage and help me. Jiani is also pursuing a PhD degree now, and I hope I
can help her in the same way she did.

Finally, let me express my appreciation to my family and friends, in particular my parents, for always being here in my life. My parents brought me to the world of natural science and astronomy in my childhood, which eventually lead to my decision of pursuing a career as an astronomer. Whenever I make little achievements, they are always more excited than me; whenever I face difficulties, they are always there to encourage me and listen to my complaints. I would not be able to achieve what I have done without my parents and other members in my family.
DEDICATION

I dedicate the thesis to my family.
# TABLE OF CONTENTS

LIST OF FIGURES ............................................. 10

LIST OF TABLES ............................................. 13

ABSTRACT ................................................... 14

CHAPTER 1 Introduction ..................................... 16
  1.1 Quasars: the Most Luminous Active Galactic Nuclei .... 16
    1.1.1 Quasar Luminosity Functions ..................... 18
    1.1.2 Triggering Mechanisms of Quasars ............... 19
  1.2 Gravitationally Lensed Quasars as Probes in Cosmology and Extra-galactic Astronomy .......... 20
  1.3 Previous Efforts in Searches of Gravitationally Lensed Quasars .... 25
    1.3.1 The Progress of High-redshift Lensed Quasar Searches .... 28
  1.4 Close Quasar Pairs: Another Population of Valuable Objects ... 31
  1.5 Scope of This Thesis ................................ 32

CHAPTER 2 A Mock Catalog of Gravitationally Lensed Quasars .... 34
  2.1 Introduction .......................................... 34
  2.2 Mock Catalog Generation ............................. 36
    2.2.1 Deflector Population .......................... 36
    2.2.2 Quasar Population ............................ 40
    2.2.3 Lensing Pipeline .............................. 44
    2.2.4 Broad-Band Photometry and Mock Images ...... 44
  2.3 Statistics of Lensed Quasars .......................... 47
    2.3.1 Comparison with the OM10 Mock Catalog ....... 52
    2.3.2 Quadruply-Lensed Quasars and Time Delays ... 54
    2.3.3 Statistics of Deflector Galaxies .............. 57
  2.4 Discussion .......................................... 58
    2.4.1 Choices of QLFs and VDFs ...................... 58
    2.4.2 Mock Catalog Completeness .................... 61
    2.4.3 Finding More Lensed Quasars Beyond the “Discoverable” Ones 61
  2.5 Summary and Conclusion .............................. 63
# TABLE OF CONTENTS – Continued

## CHAPTER 3 Analytical Model of the High-Redshift Gravitationally Lensed Quasar Population

3.1 Introduction ........................................ 65
3.2 Analytical Method .................................. 66
   3.2.1 Lensing Models .................................. 67
   3.2.2 Deflector Population ............................ 68
   3.2.3 Quasar Luminosity Functions .................... 69
3.3 Mock Catalog Method ............................... 70
3.4 Results ............................................. 71
   3.4.1 Lensing Optical Depth .......................... 72
   3.4.2 Magnification Bias and the Lensed Fraction ... 74
3.5 Discussion .......................................... 75
   3.5.1 The Lensed Fraction of High-redshift Quasars ... 75
   3.5.2 The Number of Detectable High-Redshift Lensed Quasars in Current and Future Sky Surveys .... 79
3.6 Conclusion .......................................... 81

## CHAPTER 4 Designing a Complete Survey for High-redshift Lensed Quasars

4.1 Introduction ........................................ 83
   4.1.1 An Overview of Previous Lensed Quasar Survey Methods ... 84
   4.1.2 What is a Good Survey Strategy for High-redshift Lensed Quasars? ......................... 87
   4.1.3 Sky Coverage and Depth of This Survey .......... 88
4.2 Data Sets ............................................ 89
   4.2.1 The DESI Legacy Imaging Surveys (the Legacy Survey) ... 90
   4.2.2 Other Imaging Surveys ............................ 91
   4.2.3 Other Data Sets .................................. 94
4.3 Observational Features of Gravitationally Lensed Quasars .............. 95
   4.3.1 Classification of Lensed Quasars Based on the Observed Features 95
   4.3.2 Discussions .................................... 103
4.4 Candidate Selection .................................. 103
   4.4.1 Morphology and Photometry Pre-Selection ........ 104
   4.4.2 Color Classification .............................. 110
   4.4.3 Supplementary Candidates ........................ 114
   4.4.4 Generalizing the Survey to Low-redshift Lensed Quasars ..... 117
4.5 Testing the Candidate Selection Method .......................... 118
4.6 Current Survey Progress ................................ 122
   4.6.1 Performance of the Prototype Candidate Selection Pipeline ... 122
   4.6.2 Visual Inspection and Follow-up Observations of Candidates . 123
TABLE OF CONTENTS – Continued

4.6.3 Expected Performance of the Current Candidate Pipeline  . . . 130
4.7 Future Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 130
4.7.1 Improving the Candidate Selection Pipeline . . . . . . . . . . . 130
4.7.2 Follow-up Studies of Lensed Quasars . . . . . . . . . . . . . . 132

CHAPTER 5 Discovery of a Close Quasar Pair at \( z = 5.66 \) . . . . . . . 133
5.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 133
5.2 Data and Analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . 134
5.2.1 Ground-Based Imaging and Spectroscopy . . . . . . . . . . . . 134
5.2.2 HST Imaging . . . . . . . . . . . . . . . . . . . . . . . . . . . . 136
5.2.3 Testing the Strong-Lensing Hypothesis . . . . . . . . . . . . . . 140
5.3 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 141
5.3.1 The Pair Fraction of High-Redshift Quasars . . . . . . . . . . . 141
5.3.2 The Triggering Mechanism of the Quasar Pair . . . . . . . . . . 144
5.4 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 145

CHAPTER 6 Follow-up Observations of a Lensed Quasar at \( z = 6.5 \) . . . 147
6.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 147
6.2 Data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 149
6.3 Lensing Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 152
6.3.1 Building the Lensing Model . . . . . . . . . . . . . . . . . . . . 152
6.3.2 Dust Continuum . . . . . . . . . . . . . . . . . . . . . . . . . . . 156
6.3.3 Integrated [C\( \text{II} \)] Flux . . . . . . . . . . . . . . . . . . . . . 157
6.3.4 [C\( \text{II} \)] Kinematics . . . . . . . . . . . . . . . . . . . . . . . . . 158
6.3.5 Systematic Uncertainties . . . . . . . . . . . . . . . . . . . . . . 162
6.4 Physical Properties of J0439+1634 . . . . . . . . . . . . . . . . . . . 162
6.4.1 Dust Continuum and [C\( \text{II} \)] Emission . . . . . . . . . . . . . 162
6.4.2 Host Galaxy Dynamics . . . . . . . . . . . . . . . . . . . . . . . . 167
6.5 A Maximum Starburst System With Oversized Black Hole at Cosmic
   Dawn . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 168
6.5.1 A Maximum Star Forming Rotating System . . . . . . . . . . . . 168
6.5.2 SMBH-Host Co-evolution . . . . . . . . . . . . . . . . . . . . . . . 169
6.6 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 170

CHAPTER 7 Other High-redshift Lensed Quasars and Close Quasar Pairs . 172
7.1 J0025–0145: a Lensed Quasar at \( z = 5.07 \) . . . . . . . . . . . . . . 172
7.1.1 Initial Identification . . . . . . . . . . . . . . . . . . . . . . . . . . 172
7.1.2 High-resolution HST Imaging . . . . . . . . . . . . . . . . . . . . 173
7.1.3 Discussion and Future Work . . . . . . . . . . . . . . . . . . . . . 176
7.2 J2329–0522: a Close Quasar Pair at \( z = 4.83 \) . . . . . . . . . . . . . 177
LIST OF FIGURES

1.1 An illustration of a doubly-imaged lensing system. .................. 21
1.2 Measuring $H_0$ using lensed quasars. .......................... 23
1.3 Quasar accretion disk sizes measured via microlensing events. 24
1.4 Predicted lensed fraction of $z \sim 6$ quasars from Pacucci & Loeb (2019) 30

2.1 The flowchart of lensed quasar mock catalog generation process. ... 36
2.2 The comparison between the derived VDFs of the CosmoDC2 catalog and the observed VDF at $z < 1.5$. .......................... 38
2.3 Quasar luminosity functions used for the mock catalog generation. 41
2.4 A gallery of mock lensed quasars. .......................... 46
2.5 Predicted number of discoverable lensed quasars in LSST and PS1 survey. .......................... 50
2.6 The distribution of the image separations of lensed quasars. 52
2.7 Comparison between the mock catalog of this work and Oguri & Marshall (2010). .......................... 53
2.8 The distribution of quadruply-lensed quasars in the mock catalog. 55
2.9 The magnitude difference and the time delays of the mock lensed quasars. .......................... 57
2.10 The distribution of deflector redshifts for mock lensed quasars 58
2.11 The deflector magnitudes of the mock lensed quasars. 59
2.12 Testing the mass-dependence term and the redshift-evolution term in the velocity dispersion of galaxies. .......................... 60

3.1 Comparison of galaxy velocity dispersion functions. .................. 69
3.2 The analytical model of the lensing optical depth for sources at $z_s = 6$. 73
3.3 Distribution of the lensing separation for sources at $z_s = 6$. .... 74
3.4 The magnification bias of $z_s \sim 6$ lensed quasars as a function of QLF parameters. .......................... 76
3.5 Lensed fraction of $z_s \sim 6$ quasars as a function of QLF parameters. 77
3.6 Numbers of detectable lensed quasars as a function of survey depth and quasar redshift. .......................... 81

4.1 The deblending power of Tractor. .......................... 92
4.2 The impact of foreground lensing galaxies on high-redshift quasar color selections. .......................... 96
4.3 An illustration of the subtypes of lensed quasars. .......................... 97
### LIST OF FIGURES – Continued

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4</td>
<td>Examples of lensed quasars for each subtype.</td>
<td>98</td>
</tr>
<tr>
<td>4.5</td>
<td>The flowchart of the candidate selection pipeline.</td>
<td>105</td>
</tr>
<tr>
<td>4.6</td>
<td>The impact of lensing magnification on the observed Eddington ratio.</td>
<td>116</td>
</tr>
<tr>
<td>4.7</td>
<td>Testing the morphological selection.</td>
<td>119</td>
</tr>
<tr>
<td>4.8</td>
<td>Testing the color classification.</td>
<td>121</td>
</tr>
<tr>
<td>4.9</td>
<td>Example of the PRF color selection in the prototype pipeline.</td>
<td>124</td>
</tr>
<tr>
<td>4.10</td>
<td>Example of the residual-based morphological selection in the prototype pipeline.</td>
<td>125</td>
</tr>
<tr>
<td>4.11</td>
<td>The information plot of J0918–0220.</td>
<td>128</td>
</tr>
<tr>
<td>4.12</td>
<td>The Magellan/Clay LDSS-3 spectra of J0918–0220.</td>
<td>129</td>
</tr>
<tr>
<td>5.1</td>
<td>Ground-based imaging and spectroscopy of J2037–4537.</td>
<td>135</td>
</tr>
<tr>
<td>5.2</td>
<td>The C IV emission lines of the two quasars in J2037–4537.</td>
<td>137</td>
</tr>
<tr>
<td>5.3</td>
<td>The HST ACS/WFC F850LP image of J2037–4537.</td>
<td>139</td>
</tr>
<tr>
<td>5.4</td>
<td>The pair fraction of quasars as a function of redshift.</td>
<td>144</td>
</tr>
<tr>
<td>6.1</td>
<td>The clean image and moments of J0439+1634 observed with ALMA.</td>
<td>151</td>
</tr>
<tr>
<td>6.2</td>
<td>The [C II] line profile of J0439+1634.</td>
<td>152</td>
</tr>
<tr>
<td>6.3</td>
<td>The default lensing model of J0439+1634 based on ALMA and HST observations.</td>
<td>155</td>
</tr>
<tr>
<td>6.4</td>
<td>The ALMA-only lensing model for J0439+1634.</td>
<td>156</td>
</tr>
<tr>
<td>6.5</td>
<td>The reconstructed source-plane moments for J0439+1634.</td>
<td>159</td>
</tr>
<tr>
<td>6.6</td>
<td>The velocity field of J0439+1634.</td>
<td>161</td>
</tr>
<tr>
<td>6.7</td>
<td>The [C II]-FIR ratio versus FIR surface brightness plot for J0439+1634.</td>
<td>166</td>
</tr>
<tr>
<td>7.1</td>
<td>Ground-based observations of J0025–0145.</td>
<td>174</td>
</tr>
<tr>
<td>7.2</td>
<td>HST imaging of J0025–0145.</td>
<td>175</td>
</tr>
<tr>
<td>7.3</td>
<td>The Legacy Survey images of J2329–0522.</td>
<td>178</td>
</tr>
<tr>
<td>7.4</td>
<td>The Magellan/Clay LDSS3 spectra of J2329–0522.</td>
<td>179</td>
</tr>
<tr>
<td>A.1</td>
<td>The redshift and luminosity distributions of the quasars in our master sample.</td>
<td>190</td>
</tr>
<tr>
<td>A.2</td>
<td>Number of quasars observed in each band.</td>
<td>191</td>
</tr>
<tr>
<td>A.3</td>
<td>The completeness of companion detection estimated by simulated companions.</td>
<td>197</td>
</tr>
<tr>
<td>A.4</td>
<td>The average companion numbers of sample A quasars.</td>
<td>198</td>
</tr>
<tr>
<td>A.5</td>
<td>An example of converting stellar mass limits into apparent magnitude limits.</td>
<td>201</td>
</tr>
<tr>
<td>A.6</td>
<td>The detected fraction of quasar companions with magnitudes of $m(10^9M_\odot)$, $m(10^{10}M_\odot)$ and $m(10^{11}M_\odot)$.</td>
<td>203</td>
</tr>
</tbody>
</table>
LIST OF FIGURES – Continued

A.7 The average number of companions of quasars and the control sample galaxies ................................................. 204
A.8 The redshift evolution of average number of companions around quasars 204
A.9 The evolution of the average number of companions around quasars with quasar luminosity. .......................... 205
A.10 Images of on-going merging systems in our master sample. .............. 217
10 (Continued) ............................................................................ 218
10 (Continued) ............................................................................ 219
## LIST OF TABLES

2.1 QLF parameters from the literature .................................. 43
2.2 PSF size and depth of the mock LSST images .......................... 47
2.3 Description of the columns in the mock catalog ...................... 48
2.4 Predicted number of discoverable lensed quasars .................... 49

4.1 Optical and near-IR photometry used for objects in different sky area 107
4.2 The PRFs for color classification ........................................ 112

5.1 Properties of J2037–4537 ..................................................... 138

6.1 Lens model parameters ..................................................... 153
6.2 De-lensed physical properties of J0439+1634 .......................... 164

A.1 The number of images excluded by the selection criteria. ............ 189
A.2 Summary of samples used in this study. ................................. 192
A.3 Information of on-going merging systems. .............................. 216
ABSTRACT

Gravitationally lensed quasars are valuable objects that enable various studies in extragalactic astronomy and cosmology. Despite intensive efforts, to date only one lensed quasar at $z > 5$ has been discovered, much fewer than the predictions of previous theoretical models. This discrepancy motivates us to conduct a new survey for high-redshift gravitationally lensed quasars at $z \gtrsim 5$. We revisit the models of high-redshift lensed quasar population using both mock catalogs and analytical methods, finding that there should be $\sim 10$ discoverable lensed quasars at $z > 5$ in present-day sky surveys. Previous theoretical models overestimated the number of lensed quasars at $z \sim 6$, and traditional survey strategies for high-redshift quasars are incomplete for lensed ones. We design a new candidate selection method for lensed quasars, utilizing the DESI Legacy Imaging Survey and other optical and infrared imaging surveys. The depth and deblending ability of the Legacy Survey allow discoveries of faint and small-separation lensed quasars. We select objects with both morphology and colors consistent with lensed quasars as candidates, and carry out follow-up observations to confirm the nature of these candidates. This process naturally yields close quasar pairs, a population of useful objects that exhibit observed features similar to doubly-imaged lensed quasars. In the initial stage of the survey, we have observed $\sim 200$ lens candidates, among which we identify one lensed quasar and two close quasar pairs at $z \gtrsim 5$. In particular, we report the discovery of J2037–4537 ($z = 5.66$), which is the most distant kpc-scale quasar pair known so far. J2037–4537 gives the first observational constraint on the pair fraction of $z > 5$ quasars, $f_{\text{pair}}(r < 30 \, \text{pkpc}) > 0.3\%$. Finally, we present the ALMA observation of J0439+1634, a lensed quasar at $z = 6.52$. With the help of lensing magnification, the observation reaches an effective spatial resolution of $\sim 0.8 \, \text{kpc}$. The host galaxy of J0439+1634 is a rotation-supported system with a regular Sérsic profile, yet detailed
analysis indicates complex structures in the quasar host galaxy. J0439+1634 is a gas-rich system that is undergoing extreme star formation processes.
1.1 Quasars: the Most Luminous Active Galactic Nuclei

It is believed that most (if not all) massive galaxies host supermassive black holes (SMBHs) in their centers (e.g., Kormendy & Gebhardt, 2001). Active galactic nuclei (AGN) are SMBHs that are actively accreting material. When falling into the SMBHs, the gravitational potential energy of the accreted material is converted into electromagnetic radiation and kinetic energy that powers outflows. SMBHs may have significant impacts on their host galaxies via AGN activities (e.g., Fabian, 2012), and understanding the evolution of AGN and their host galaxies is a fundamental task in extragalactic astronomy. The most powerful type of AGN, i.e., quasars, are among the most luminous non-transient objects in the universe (e.g., Fan et al., 2019a).

AGN emit from gamma-ray to radio wavelengths. Specifically, AGN exhibit power-law continua plus broad and/or narrow emission lines in the UV/optical wavelengths. Quasars were first recognized as point sources (i.e., quasi-stellar objects or QSOs) that exhibit large redshifts and non-stellar optical spectra (e.g., Schmidt, 1965). X-ray emission comes from the hot corona around the SMBHs and is the characteristic feature of accreting black holes. At infrared wavelengths, the emission from an AGN is dominated by the hot and warm dust components around the SMBH and in the host galaxy. The radio emission is mainly contributed by the jets driven by the outflowing material, which can be extended to tens to hundreds of kiloparsec.

AGN structure can be understood in the framework of the so-called “unified model” (Urry & Padovani, 1995), which explains most observed features of AGNs. In the unified model, the SMBH is surrounded by a compact accretion disk that
dominates the continuum emission from UV to optical. Outside the accretion disk there is the broad-line region (BLR), which consists of clouds of ionized gas orbiting around the SMBH and generates the broad emission lines in AGN spectra. The accretion disk and the BLR are surrounded by the dust torus, which absorbs the emission from the accretion disk and re-emits it in near-to-mid infrared wavelengths. The ionized gas at larger distances from the SMBH forms the narrow-line region (NLR) that generates narrow emission lines at UV and optical. At distances similar to the NLR, there are polar dusts emitting at mid-to-far infrared wavelengths. The far infrared to sub-mm fluxes of AGN are dominated by the dust in the host galaxy heated by star formation. In this framework, the difference between Type 1 (broad-line) and Type 2 (narrow-line) AGN is explained by the orientation effect, and in Type 2 objects, the AGN dust torus obscures direct emission from the accretion disk and the BLR.

Observations have shown that SMBHs co-evolve with their host galaxies. This coevolution is revealed by the $M_{\text{BH}} - \sigma$ relation, which states that the mass of the SMBH is tightly correlated with the velocity dispersion of the stellar bulge of its host galaxy (for a recent review, see Kormendy & Ho, 2013). Quasars (and the more general population of AGN) provide a promising link between the SMBHs and their host galaxies. On the one hand, the material that feeds the SMBH comes from the host galaxy, therefore the SMBH growth via accretion must be impacted by the properties of the host galaxy. On the other hand, the large amount of energy released by the AGN (especially luminous quasars) can have a significant impact on the host galaxy by heating the gas and driving galaxy-scale outflows, known as AGN feedback.

To understand the properties and evolution of AGN, intensive efforts have been carried out in the searches and follow-up observations of these objects. To date, quasars have been reported at redshifts $z \gtrsim 7$ (e.g., Mortlock et al., 2011; Bañados et al., 2018; Yang et al., 2020; Wang et al., 2021), i.e., when the universe is $\lesssim 700$ Myrs old. The discovery of billion-solar-mass SMBHs in the early universe challenges the theory of SMBH formation. At high redshifts, even luminous quasars appear to
be faint, and the most advanced facilities are needed to characterize distant quasars in detail.

In addition to the studies of SMBHs and their host galaxies, quasars also act as probes in the studies of intergalactic medium (IGM), circumgalactic medium (CGM), and cosmic structure formation. The Ly\(\alpha\) forest in quasar spectra has been used to characterize the evolution of the IGM (e.g., Palanque-Delabrouille et al., 2013). The metal line absorbers in quasar spectra probe the CGM of foreground galaxies, providing information about the enrichment of the CGM from the galaxy (e.g., Werk et al., 2014). Moreover, quasars trace overdensities, and mapping the environment of quasars sheds light on the evolution of the most massive structures in the universe (e.g., Onoue et al., 2018).

1.1.1 Quasar Luminosity Functions

The population of quasars can be described by the quasar luminosity function (QLF), which gives the comoving number density of quasars as a function of redshift and luminosity. To be specific, the comoving number density of quasars at redshift \(z\) that fall in the absolute magnitude bin \((M, M + dM)\) can be calculated as \(\Phi(M, z)dM\), where \(\Phi(M, z)\) is the QLF. QLF is usually parameterized as a double power-law (e.g., Richards et al., 2006a, and reference therein),

\[
\Phi(M, z) = \frac{\Phi^\ast(z)}{10^{0.4(\alpha+1)(M-M^\ast)} + 10^{0.4(\beta+1)(M-M^\ast)}},
\]

where \(\alpha\) and \(\beta\) are the faint- and bright-end slopes, \(M^\ast\) is the break magnitude and \(\Phi^\ast\) is the normalization.

Characterizing the redshift evolution of the QLF constrains models of the SMBH growth (e.g., Rosas-Guevara et al., 2016). Specifically, the luminosities of quasars are directly linked to the accretion rates of the SMBHs, and the redshift evolution of the QLF reflects the build-up history of SMBHs. QLF measurements is also important in understanding galaxy evolution given the coevolution of SMBHs and their host galaxies (e.g., Marulli et al., 2008). In addition, the QLF determines the contribution of quasars to the ionizing photons and is indispensable in the understanding of the
cosmic reionization (e.g., Onoue et al., 2017).

The QLF can be measured by conducting surveys of quasars (e.g., Yang et al., 2017; McGreer et al., 2018). So far, the QLF has been measured up to $z \sim 7$ (e.g., Kulkarni et al., 2019). The number density of quasars peaks at $z \sim 2$ and drops towards both high and low redshifts. At $z \gtrsim 6$, the number density of quasars evolves as $\rho(z) \propto 10^{-0.78z}$ (e.g., Wang et al., 2019b). The quick drop of the number density makes it challenging to find quasars at high redshifts.

1.1.2 Triggering Mechanisms of Quasars

To generate AGN activities in a galaxy, there must be sufficient material inflows towards the galaxy center to feed the SMBH. It has been proposed that gas-rich galaxy major mergers can trigger AGN activities (e.g., Di Matteo et al., 2005), where the merging event disturbs the gas content and generates strong inflows towards the SMBHs. Hopkins et al. (2008) provide a thorough description of the evolution of galaxy major mergers triggering quasars. A galaxy major merger starts from the first encounter and passing-by of the two progenitor galaxies, after which the two galaxies move towards and collide with each other once again. During the merging process, the gas in the progenitor galaxies loses angular momentum and flows into the galactic centers, fueling powerful quasar activities. The SMBH usually starts actively accreting material after the final coalesce of the galaxy merger. The quasar is usually obscured by dust at the beginning, with the host galaxy appears to be a luminous infrared galaxy (LIRG). The quasar then removes the surrounding dust via AGN feedback and becomes unobscured.

Some secular evolution processes can also generate instabilities in galaxies and drive gas inflows (e.g., Shlosman et al., 1989; Hopkins et al., 2014). These processes include minor mergers, perturbations from close companions, instabilities introduced by bars and spiral arms, etc. Hopkins et al. (2014) suggest that secular evolution can produce low-to-intermediate luminosity AGNs, while major mergers are needed to trigger the most luminous AGNs, i.e., quasars.

There have been efforts to determine whether major galaxy mergers or secular
evolution dominate the triggering of AGNs. The results are ambiguous, with some studies claim a low merger fraction in AGN (e.g., Cisternas et al., 2011; Villforth et al., 2017) and others suggest a high fraction of merger-triggered AGN (e.g., Fan et al., 2016; Goulding et al., 2018). The observed merger fraction depends on many properties of the AGN (e.g., redshifts, luminosities, obscurations). In Yue et al. (2019), we argue that previous observations are consistent with a picture where most quasars are triggered by mergers and evolve from obscured to unobscured phase.

1.2 Gravitationally Lensed Quasars as Probes in Cosmology and Extragalactic Astronomy

In this thesis, I focus on a special subset of quasars, i.e., quasars that are gravitationally lensed by foreground galaxies. Gravitational lensing is a phenomenon where the light from the background source is deflected by the gravity of a foreground massive object. The foreground massive object, often called the “deflector” or the “lens”, acts as a natural telescope that magnifies the background source. Lensing magnification enhances the sensitivity and spatial resolution of observations, making gravitational lensing a powerful tool in the studies of faint distant objects. Figure 1.1 presents a cartoon of a lensing system, which explains the basic concept of gravitational lensing.

Gravitationally lensed quasars enable many critical studies in cosmology and extragalactic astronomy. Lensed quasars have been used to measure cosmological parameters, to investigate the structure of quasar accretion disks, and to study quasar host galaxies with improved sensitivity and spatial resolution.

1) Constraining cosmological parameters

Even before the first discovery of gravitational lensing systems, it has been proposed that lensed variable objects can be used to constrain the Hubble constant ($H_0$) (e.g., Refsdal, 1964). In a multiply-imaged lensing system, each lensed image has a distinct light path, resulting in time lags between the lensed images. The difference
Figure 1.1: An illustration of a doubly-imaged lensing system. Adopted from Bisnovatyi-Kogan & Tsupko (2017). *Left:* the geometry of the source, the lensing galaxy (also called the deflector), and the observer. There are two possible paths through which the light can travel from the source to the observer, which forms two lensed images of the background source. Other lensing systems can generate more than two images of the background source. *Right:* what the observer sees on the sky.
in light path lengths, and thus the time lag, is directly linked to the underlying cosmology. Lensing time lags can be measured by monitoring the light curves of the lensed images of the variable source. Combined with lens modeling, such measurements provide independent constraints of cosmological parameters, especially $H_0$.

Quasars are intrinsically variable objects, which have stochastic light curves with typical amplitudes of $\sim 0.2$ mag (e.g., MacLeod et al., 2010). A few projects aim at constraining $H_0$ using lensed quasars, including H0LiCOW (Suyu et al., 2017) and STRIDES (Treu et al., 2018). To improve the accuracy of measurements and reduce possible systematics, these studies use carefully-selected lensed quasars with significant variability and large time delays, and carry out long-term photometric monitoring to accurately measure the time lags. In addition, follow-up spectroscopy is performed for the main deflector and the other galaxies close to the sightline of the lensing system, which provides the dynamical masses of these galaxies via stellar kinematics. These measurements significantly reduce the systematic uncertainties in lensing models.

To date, the time lag analysis has been performed for $\sim 10$ lensed quasars, resulting in $H_0$ measurements with errors of $\sim 2\%$. For example, Wong et al. (2020) gives $H_0 = 73.3^{+1.7}_{-1.8}$ km s$^{-1}$ Mpc$^{-1}$ based on six lensed quasars (Figure 1.2). As an independent test, the $H_0$ measured by lensed quasars provides critical information about the discrepancy between the supernovae measurements and the CMB experiments. Surveys of lensed quasars are thus critical, which can build large samples of lensed quasars for time lag analysis (up to several thousands; e.g., Oguri & Marshall, 2010; Yue et al., 2022a).

(2) Measuring the size of quasar accretion disks

One unique application of lensed quasars is to measure the sizes of quasar accretion disks via microlensing events. A microlensing event happens when a star in the foreground deflector galaxy passes through the sightline of a lensed quasar image. In this case, there will be an increase in the magnification of the lensed image, generating a “bump” in its light curve. Note that the microlensing effect depends on the
Figure 1.2: Measuring $H_0$ using lensed quasars in the H0LiCOW project. Adapted from Wong et al. (2020). Lensed quasars provide independent constraints on the Hubble constant in addition to the CMB experiment and the distance ladder measurements. By combining the analysis of six carefully-selected lensed quasars, the Hubble constant is measured to an accuracy of $\sim 2\%$.

position of the lensed quasar image, meaning that different lensed images exhibit distinct microlensing signals. It is thus possible to probe the microlensing signals by comparing the light curves of the lensed images. If we know the velocity of the star ($v_{\text{star}}$) and how long the additional magnification lasts ($\tau$), we can estimate the size of the accretion disk as $D \sim v_{\text{star}} \tau$.

Kochanek (2004) describes in detail the method of using microlensing to measure the quasar accretion disk size. First, models of lensed quasars are generated that predict the light curves of each lensed quasar image, which is a function of the intrinsic quasar light curve and the stellar population (mass function, spatial distributions, kinematics, etc.) of the deflector. Then Monte Carlo simulation is run to generate random realizations of the lensing system, and Bayesian analysis is used to derive the best-fit values and the errors of parameters. The observables in the fitting include the light curves of lensed quasar images, as well as the images and spectroscopy of the deflector.

The method proposed by Kochanek (2004) has been successfully applied on a
Figure 1.3: Quasar accretion disk sizes measured via microlensing events in gravitationally-lensed quasars. Adapted from Cornachione & Morgan (2020). The $y$-axis represents the half-light radius of the quasar accretion disk at rest-frame 2500Å. The points with error bars correspond to the microlensing measurements for 15 lensed quasars, and the solid line with blue shades shows the best-fit relation and errors. The crosses, the dashed line and the green shades illustrate the estimated disk size using a standard thin disk model, which is $\sim 3 - 4$ times smaller than the microlensing measurement.

number of lensed quasars (e.g., Cornachione et al., 2020). Cornachione & Morgan (2020) measure the accretion disk sizes for a sample of 15 gravitationally lensed quasars (Figure 1.3). Interestingly, the microlensing-measured sizes are $\sim 3 - 4$ times larger than the values estimated using their luminosity and assuming a standard thin disk model. This result suggests that quasar accretion disks are more complex than the standard thin disks, though this discrepancy is still not well-understood.

(3) Providing a magnified view of quasar host galaxies

Gravitationally lensed quasars offer opportunities to study SMBHs and their host galaxies with enhanced sensitivity and spatial resolution (e.g., Paraficz et al., 2018). In some cases, lensing magnification enables studies that are not accessible for un-lensed objects. This effect is especially important for high-redshift objects,
which are usually faint and spatially-unresolved in most observations.

One ideal example to illustrate the power of lensing magnification is J0439+1634 (Fan et al., 2019b), a gravitationally lensed quasar at $z = 6.52$. J0439+1634 has a large magnification of $\mu = 51$, allowing us to reach an effective spatial resolution of $\lesssim 20$ pc using ALMA or JWST (Yue et al., 2021b). This resolution enables critical experiments like measuring the SMBH mass via gas kinematics in the vicinity of the black hole. For normal high-redshift quasars, it is impossible to reach this spatial resolution in the foreseeable future. J0439+1634 thus offers the best opportunity to directly measure the mass of a SMBH beyond the local universe. J0439+1634 will be discussed in more detail in Chapter 6.

1.3 Previous Efforts in Searches of Gravitationally Lensed Quasars

The first gravitationally lensed quasar discovered is Q0957+561 at $z = 1.413$ (Walsh et al., 1979), which is also the first gravitational lensing system ever reported. Since the discovery of Q0957+561, there have been intensive efforts to search for lensed quasars. In the past two decades, large imaging sky surveys like the Sloan Digital Sky Survey (SDSS; York et al., 2000) have greatly enhanced our ability to identify lensed quasars. So far, about 220 lensed quasars have been discovered using various techniques\(^1\). In this subsection, I summarize some previous efforts of searching for gravitationally lensed quasars.

(1) The SDSS Quasar Lens Search

The SDSS Quasar Lens Search (SQLS; Oguri et al., 2006) aims at identifying gravitationally lensed quasars in the spectroscopically-confirmed SDSS quasar sample (e.g., Richards et al., 2006a). SQLS uses two methods to select lens candidates in the quasar sample: the “morphological selection” identifies quasars that cannot be well-fitted by a point spread function (PSF), which might be unresolved lensed quasars; the “color selection” looks for quasars that have a companion with colors similar to the quasar, which are likely resolved lensed quasar images. Follow-up

\(^1\)A compilation of lensed quasars can be found at https://research.ast.cam.ac.uk/lensedquasars/
imaging and spectroscopy with high spatial resolution were carried out to confirm the lensing nature of these candidates.

SQLS develops an effective way of searching for lensed quasars in large imaging surveys; Inada et al. (2012) report the final SQLS sample, which contains a well-defined statistical subsample of 26 lensed quasars brighter than $m_i = 19.1$, as well as 36 lenses discovered with various techniques. The search for lensed quasars continues in the subsequent SDSS observations, e.g., the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al., 2013) of SDSS-III. More et al. (2016) report the discovery of 13 lensed quasars in the BOSS quasar sample.

(2) The Cosmic Lens All-Sky Survey

Another early effort to search for gravitational lensed systems is the Cosmic Lens All-Sky Survey (CLASS; Myers et al., 2003). CLASS is a sky survey using the Very Large Array (VLA), aiming at finding gravitationally lensed radio sources. The VLA observations reach a beam size of $\sim 0.2^\prime$, which can resolve small-separation compact lensing systems. In their final sample, CLASS reports 18 lenses\(^2\). Optical follow-up observations found that CLASS lens sample contains both luminous quasars and radio galaxies that are faint in the optical wavelength.

(3) The STRong lensing Insights into the Dark Energy Survey

The STReong lensing Insights into the Dark Energy Survey (STRIDES; Treu et al., 2018) aims at searching for lensed quasars in the Dark Energy Survey (DES; DES Collaboration et al., 2021a) footprint and constraining cosmological parameters using these lensed quasars. In addition to the DES data, STRIDES also uses the data from the Wide-field Infrared Survey Explorer survey (WISE; Wright et al., 2010) and the Gaia survey (Gaia Collaboration et al., 2016). STRIDES adopted a number of different candidate selection strategies, mainly targeting at bright lenses that are suitable for cosmological studies. Some of these strategies will be discussed in Chapter 4.

To date, STRIDES has published 12 lensed quasars (e.g., Anguita et al., 2018; Lemon et al., 2020). A few of these lensed quasars have been analyzed to measure

\(^2\)http://www.aoc.nrao.edu/~smyers/class.html
the Hubble constant (e.g., Shajib et al., 2020). By the end of the project, STRIDES is expected to deliver time delay analysis of \( \sim 50 \) quadruply-lensed quasars and \( \sim 200 \) doubly-lensed quasars, expanding the current sample by a factor of \( \sim 10 \).

(4) Gaia-based lensed quasar searches

The *Gaia* mission (Gaia Collaboration et al., 2016) is an all-sky survey that aims at mapping the stars in the Milky Way. Using a space-based telescope, *Gaia* delivers a PSF size of \( \sim 0.04" \) in Data Release 2 (Gaia Collaboration et al., 2018) and is expected to further improve the spatial resolution in future data releases. *Gaia* offers both imaging and integrated-field spectroscopy in the optical wavelengths. The imaging observations use a “G” filter that covers from 3500Å to 10000Å, as well as a “blue” (\( G_{BP} \)) and a “red” (\( G_{RP} \)) filter to get the color of the objects. With high-precision astrometry, *Gaia* provides the parallax and proper motions of each object. The spectroscopic observations measure physical properties (temperature, gravity, metallicity, etc.) and radial velocities of stars.

The *Gaia* survey has revolutionized the searches for lensed quasars. The superb angular resolution of *Gaia* can resolve many compact lensed quasars that are unresolved in ground-based observations. In addition, the parallax and proper motions are efficient in rejecting galactic stars. Many recent surveys for lensed quasar use *Gaia* to identify promising lens candidates (e.g., Krone-Martins et al., 2018; Stern et al., 2021). In particular, Lemon et al. (2017) propose a number of catalog-level quantities in the *Gaia* survey that are useful in selecting lensed quasars. Objects that are resolved into multiple components, have large astrometric uncertainties, or have positional and photometric differences between *Gaia* and ground-based imaging surveys are ideal candidates of lensed quasars. Using these indicators, \( \sim 50 \) lensed quasars have been discovered in the subsequent studies (Lemon et al., 2018, 2019).

(5) Other lensed quasar surveys and investigations

Some other searches for lensed quasars have been carried out based on various sky surveys, for example, the KiDS-SQuaD (Spiniello et al., 2018) that uses the Kilo-Degree Survey (de Jong et al., 2013), the Survey of Gravitationally lensed Objects in
HSC Imaging (SuGOHI; e.g., Chan et al., 2020) that uses the Hyper Supreme-Cam Survey (HSC; Aihara et al., 2018a), and searches for lensed quasars (Schechter et al., 2017) in the VST-ATLAS survey (Shanks et al., 2015). Useful tools and techniques have been developed in these surveys. For example, Chao et al. (2021) use variability analysis to identify lensed quasar candidates in the HSC survey; Chan et al. (2015) develop an automatic image fitting and lens modeling tool that gives the probability for each object to be a lensed quasar in the HSC survey.

Deep learning has been found to be a powerful tool in surveys for lensed galaxies (e.g., Lanusse et al., 2018), which identifies the distinguishing features of a lensed galaxy (e.g., arcs) in images. Agnello et al. (2015) investigate the possibility of using deep learning to identify lensed quasar candidates. However, the subsequent studies (e.g., Agnello, 2017; Agnello et al., 2018; Agnello & Spiniello, 2019) still heavily rely on traditional catalog-based indicators instead of deep learning methods. One possible reason is that lensed quasars have fundamental differences from normal lensed galaxies. Specifically, lensed galaxies exhibit unique features in their images, i.e., the lensed arcs, which can be identified by image recognition algorithms. Lensed quasars, instead, appear as a galaxy plus several point sources, and are usually confused with projected groups of galaxies and stars. To date, it is still unclear how to efficiently use deep learning in lensed quasar searches.

1.3.1 The Progress of High-redshift Lensed Quasar Searches

Lensing magnification is especially important in the observations of high-redshift objects, for which it is usually challenging to obtain high-resolution and high signal-to-noise ratio (S/N) data. Probing high-redshift quasars that are magnified by gravitational lensing is crucial in at least three aspects: (1) Identifying bright lensed quasars that are intrinsically faint is necessary for an accurate measurement of the quasar luminosity function (e.g., Maoz & Rix, 1993). (2) Gravitational lensing enables studies of the intrinsically faint quasar population with high S/N data. (3) For lensed objects, observations can reach a physical resolution that is several times higher than un-lensed ones.
The previous and on-going projects discussed in Section 1.3 are for the general lensed quasar population. Despite the success in lensed quasar surveys at relatively low redshifts \((z \lesssim 4)\), little progress has been made in the searches of high-redshift \((z \gtrsim 5)\) lensed quasars. The most distant lensed quasars discovered prior to 2019 are at \(z = 4.8\), including SDSSJ0946+1835 (McGreer et al., 2010) and SDSSJ1452+4224 (More et al., 2016). These two lensed quasars were initially targeted by the SDSS spectroscopy as galaxies, and their spectra reveal the existence of the background quasars.

Shortly after the SDSS built the first sample of quasars at \(z > 5\), there have been efforts to search for lensed quasars at similar redshifts. Richards et al. (2004) and Richards et al. (2006b) use \(HST\) to obtain high-resolution images of \(\sim 160\) quasars at \(z > 4\), finding no lensed ones in this sample. The first known lensed quasar at \(z > 5\) is J0439+1634 \((z = 6.5;\) Fan et al., 2019b), which is the most distant lensed quasar known so far and is still the only lensed quasar reported at \(z > 5\). J0439+1634 was discovered serendipitously, indicating that a large number of high-redshift lensed quasars might still be missed in previous quasar surveys.

Meanwhile, theoretical models have predicted that the lensed fraction of quasars at \(z \gtrsim 6\) can be as high as \(\sim 20\%\) (Comerford et al., 2002; Wyithe & Loeb, 2002). The high lensed fraction is a result of magnification bias, i.e., lensing magnification boosts the observed flux of intrinsically faint objects, making them detectable in flux-limited surveys. Recently, Pacucci & Loeb (2019) re-calculate the lensed fraction of high-redshift quasars using updated models for the deflector galaxy population and the quasar luminosity functions. The prediction is shown in Figure 1.4, which suggests that the lensed fraction of \(z \sim 6\) quasars is \(\sim 4\% - 20\%\), depending on the slopes of the quasar luminosity function.

To date, more than 500 quasars at \(z > 5\) have been discovered (e.g., Jiang et al., 2016; Matsuoka et al., 2018a; Yang et al., 2019a; Wang et al., 2019b). The predicted lensed fraction of high-redshift quasars suggests that at least 20 lensed quasars should have been discovered at \(z > 5\). The discrepancy between the predicted number of high-redshift lensed quasars and the observations suggests that either
Figure 1.4: Predicted lensed fraction of $z \sim 6$ quasars from Pacucci & Loeb (2019), as a function of lensing magnification and the bright-end slope of the quasar luminosity function (QLF). The arrows mark the probabilities to have magnification $\mu > 2$, which corresponding to multiply-imaged cases. The observed lensed fraction is much lower than the predicted ones for both QLF slopes. See Chapter 3 for related discussions.
the theoretical models have critical deficiencies, or previous surveys for high-redshift lensed quasars are highly incomplete. In particular, none of the previous surveys focuses on lensed quasars at \( z \gtrsim 5 \), and it is still unclear how to find these objects efficiently.

### 1.4 Close Quasar Pairs: Another Population of Valuable Objects

Surveys of lensed quasars will naturally yield another type of valuable objects, i.e., close pairs of quasars. A close quasar pair consists of two quasars that are physically associated with each other, with separations of \( \lesssim 10 \) kpc. Given their small separations, the two quasar host galaxies in a quasar pair must be merging with each other. Close quasar pairs and doubly-imaged lensed quasars look similar to each other in imaging surveys. Indeed, sometimes it can be challenging to distinguish a close quasar pair from a doubly-imaged lensed quasar (e.g., Yue et al., 2021a). As a result, a complete search for lensed quasars will also be a search for close quasar pairs.

Quasar pairs reflect a unique phase in galaxy merger evolution where both SMBHs in the progenitor galaxies are ignited by the merging event. The evolution of galaxy mergers depends on the properties of progenitor galaxies (e.g., masses, morphology, gas contents; Hopkins et al., 2008), and only mergers with certain orbital parameters can trigger luminous quasar pairs. Capelo et al. (2015, 2017) run hydrodynamical simulations of galaxy mergers, showing that close quasar pairs usually have SMBH mass ratios close to one and separations of \( \sim 10 \) kpc.

Observations of quasar pairs constrains the evolution models of SMBHs and galaxy mergers in a few ways. First, the statistics of quasar pairs (e.g., the fraction of pairs among all quasars) can be directly compared to cosmological simulations (e.g., Volonteri et al., 2016; Steinborn et al., 2016), putting unique constraints on the models of SMBH growth, galaxy merger evolution and AGN feedback. Second, detailed observations of individual quasar pairs can be compared to hydrodynamical simulations of galaxy mergers (e.g., Green et al., 2010; Capelo et al., 2017), providing
information about how galaxy mergers trigger quasar activity.

In addition to studies of SMBHs and galaxy mergers, quasar pairs also enable various studies in extragalactic astronomy. Specifically, quasar pairs probe the small-scale structures in the IGM or the CGM of foreground galaxies via the absorption lines in the quasar spectra (e.g., Rorai et al., 2017). In addition, quasar pairs are good probes of massive structures, since quasars usually reside in overdensities (Onoue et al., 2018).

1.5 Scope of This Thesis

In Section 1.3.1, I describe the discrepancy between the predicted number of high-redshift lensed quasars and previous observations. This discrepancy motivates us to carry out this thesis project, i.e., conducting a new survey for high-redshift gravitationally lensed quasars. We aim at building the first sample of lensed quasars at $z > 5$. We will analyze the observed features of high-redshift lensed quasars using simulations, and design a candidate selection pipeline explicitly for these objects. We will then carry out follow-up observations to confirm the nature of the lens candidates. This survey is the first systematic survey for lensed quasars at $z \gtrsim 5$; the lensed quasars discovered in this survey will not only resolve the discrepancy between previous models and observations, but also enable various critical studies of the high-redshift SMBHs and their host galaxies.

In this thesis, I summarize my efforts in this new survey for high-redshift lensed quasars. Specifically, I will explore the following open questions:

- How accurate can we model the high-redshift lensed quasar population? One key step towards resolving the “missing high-redshift lensed quasars problem” is to fully understand the uncertainties in theoretical models of the lensed quasar population. Previous studies adopt various assumptions about deflectors and background quasars. These assumptions introduce complex systematic uncertainties that are not well-understood.

- What are the observed features of high-redshift lensed quasars, and how can we
identify these objects? High-redshift lensed quasars exhibit observational features that are different from un-lensed high-redshift quasars and lensed quasars at low redshifts. It is unclear what high-redshift lensed quasars look like in current and upcoming sky surveys, as well as how to identify them effectively.

- How can we use high-redshift lensed quasars to explore the evolution of distant SMBHs and their host galaxies? Lensing magnification is a powerful tool in probing faint and distant objects. It is critical to explore how to use lensed quasars to carry out studies of distant SMBHs that are not accessible in other objects. I will use J0439+1634 (a lensed quasar at $z = 6.5$) to investigate what we can learn from high-redshift quasars magnified by gravitational lensing.

- What can we learn from the close quasar pairs discovered in this survey? Close quasar pairs are natural yields of surveys for lensed quasars. Limited by the sensitivity and spatial resolution of most observations, close quasar pairs at $z \gtrsim 3$ remain largely unexplored. Our survey for high-redshift quasars will also build the first sample of quasar pairs at $z \gtrsim 5$, and I will use this sample to study the evolution of SMBHs in the early universe.

This thesis is organized as follows. In Chapter 2 and Chapter 3, I revisit the theoretical models of high-redshift lensed quasars. Specifically, Chapter 2 describes the method of simulating the population of lensed quasars, which are applied in Chapter 3 to calculate the number of lensed quasars we can find in various surveys. I then describe how to design a complete survey for high-redshift lensed quasars in Chapter 4. In Chapter 5 and Chapter 7, I present lensed quasars and quasar pairs at $z \gtrsim 5$ discovered in this survey. In Chapter 6, I investigate the host galaxy properties of J0439+1634, illustrating the usefulness of lensed quasars in studying high-redshift SMBHs and their hosts. Chapter 8 summarizes this thesis. In addition, I present a study of the close environments of quasars in Appendix A, providing supplementary information about the merger-triggering model of quasars.
2.1 Introduction

Gravitationally lensed quasars are valuable objects which enable a variety of unique studies in extragalactic astronomy. Examples include high-resolution observations of quasar host galaxies (e.g., Paraficz et al., 2018), studies of the dark matter profile and the circumgalactic medium of the foreground deflector galaxy (e.g., Gilman et al., 2020; Cashman et al., 2021), measuring the accretion disk size of background quasars (e.g., Cornachione & Morgan, 2020), and constraining key cosmological parameters such as the Hubble constant (e.g., Suyu et al., 2017). High-redshift lensed quasars are especially interesting, as they also probe the faint quasar population that is inaccessible without lensing magnification (e.g., McGreer et al., 2010; Yang et al., 2019b; Yue et al., 2021b).

Despite their usefulness, gravitationally lensed quasars are rare, with about 120 of them discovered prior to 2015 (e.g., Inada et al., 2012). In the last few years, many searches for lensed quasars have been carried out utilizing recent wide-area imaging surveys, and have nearly doubled the sample size of lensed quasars (e.g., More et al., 2016; Lemon et al., 2018, 2019, 2020; Agnello et al., 2018; Treu et al., 2018; Spiniello et al., 2018; Jaelani et al., 2021). These searches have utilized effective candidate selection strategies for lensed quasars, based on their observed features like colors, morphologies and variabilities. In the near future, the Rubin Observatory Legacy Survey of Space and Time (LSST, Ivezić et al., 2019) will deliver deep, multi-band and multi-epoch imaging with sub-arcsec spatial resolution in a wide sky area, making it promising to build a large sample of lensed quasars in the next decade.

†This Chapter was published as Yue et al. (2022a)
As the commissioning of LSST is approaching, there is a need to investigate the expected outputs of future surveys for lensed quasars, and to design a complete and efficient candidate selection strategy with LSST data. Mock catalogs and simulated observations are powerful tools to accomplish these tasks. In Oguri & Marshall (2010) (hereafter OM10), the authors present a mock catalog of lensed quasars, and discuss the number of lensed quasars that can be found in a number of surveys. Agnello et al. (2015) further generate simulated observations for the mock lensing systems using the properties of real galaxies and quasars from various of sky surveys. The simulated observations, as well as the candidate selection techniques developed based on it, have been proven to be effective in many lensed quasar searches (e.g., Agnello et al., 2018; Spiniello et al., 2018; Lemon et al., 2020).

However, the mock catalog in OM10 and the simulated observations have a few limitations. First, the OM10 mock catalog is based on early measurements of quasar luminosity functions (QLFs) and galaxy velocity dispersion functions (VDFs), while recent observations have provided more accurate measurements of the quasar and galaxy populations, especially at high redshifts. Second, the OM10 catalog only includes quasars at $z_{\text{qso}} < 5.5$, while the depth and wavelength coverage of LSST will allow finding lensed quasars at redshift $z \gtrsim 7$. In addition, the simulated observations in Agnello et al. (2015) were largely based on wide-area sky surveys like the Sloan Digital Sky Survey (SDSS, York et al., 2000) and the Wide-field Infrared Survey Explorer (WISE, Wright et al., 2010), which are usually not deep enough to characterize faint galaxies at relatively high redshift. As such, the observed features of lensed quasars with faint (and thus less massive) deflector galaxies are not well-represented.

In this work, we present a new mock catalog of lensed quasars at $z_{\text{qso}} < 7.5$ with simulated spectral energy distributions (SEDs) and LSST images, extending previous studies to higher quasar redshifts and smaller deflector galaxy masses. We use updated QLFs and VDFs to ensure accurate modeling of the lensed quasar population. This mock catalog predicts the number of discoverable lensed quasars in current and next-generation sky surveys, and can serve as training and testing sets.
in future searches for lensed quasars. This paper is organized as follows. We describe
the method of generating the mock catalog in Section 2.2. Section 2.3 describes the
predicted statistics of lensed quasars. Section 2.4 discusses some systematic uncer-
tainties of the mock catalog and investigates possible strategies for future lensed
quasar searches. We summarize in Section 2.5. Throughout this paper we use a
flat Lambda cold dark matter ($\Lambda$CDM) cosmology with $H_0 = 70$ km s$^{-1}$ kpc$^{-1}$ and
$\Omega_M = 0.3$.

2.2 Mock Catalog Generation

In this section, we describe the method we use to generate the mock catalog, in-
cluding simulations of the quasar population and the deflector galaxy population,
as well as the matching and lensing algorithm. Figure 2.1 shows the flow chart of
these processes.  

2.2.1 Deflector Population

We use the CosmoDC2 mock galaxy catalog (Korytov et al., 2019) to model the
deflector galaxies. CosmoDC2 is a synthetic galaxy catalog that covers an area of
440 deg$^2$ out to a redshift of $z = 3$. CosmoDC2 first identifies dark matter halos in
the $N$–body simulation Outer Rim (Heitmann et al., 2019), and assigns simulated
galaxies from the UniverseMachine (Behroozi et al., 2019) to the dark matter halos.
The algorithm is tuned to match the observed stellar mass versus star formation rate

1The mock catalog and the code are available at https://github.com/yuemh/lensQSOsim.git
\( M_\ast, \) SFR \) distributions of galaxies. CosmoDC2 then assigns various of properties to the mock galaxies, including stellar masses, optical-to-near infrared (NIR) SEDs, half-light radii and ellipticities, using empirical relations or simulated galaxy spectral libraries. Korytov et al. (2019) show that the mock catalog reproduces the observed colors and number counts of galaxies at a wide range of redshifts. The large area of the CosmoDC2 catalog ensures a small statistical error of the simulated lens sample, and the galaxy properties provide all the information needed to generate mock observations of lensing systems.

We use singular isothermal ellipsoids (SIE, e.g., Kormann et al., 1994) to model the lensing potential of the deflector galaxies. An SIE lens is parameterized by its ellipticity, \( e = 1 - q \) where \( q \) is the axis ratio\(^2\), and its Einstein radius,

\[
\theta_E = 4\pi \left( \frac{\sigma}{c} \right)^2 \frac{D_{ds}}{D_s}
\]

where \( \sigma \) is the line-of-sight velocity dispersion of the galaxy, \( D_s \) is the angular diameter distance from the observer to the source, and \( D_{ds} \) is the angular diameter distance from the deflector to the source.

We estimate the velocity dispersion of a galaxy using its stellar mass. Specifically, we convert stellar masses to dynamical masses using a stellar-to-dynamical mass ratio, and calculate the velocity dispersion by applying the virial theorem (e.g., Cappellari et al., 2006):

\[
\sigma = \sqrt{\frac{GM_{\text{dyn}}}{KR_e}}
\]

We adopt \( M_\ast/M_{\text{dyn}} = 0.557 \), which gives a good description for galaxies in a wide mass (100 km s\(^{-1}\) < \( \sigma < 300 \) km s\(^{-1}\)) and redshift (0 < \( z < 1 \)) range (e.g., Bezanson et al., 2011, 2012). Since early-type galaxies dominate the deflector population in galaxy-scale lenses (e.g., Oguri, 2006; Inada et al., 2012), we adopt the typical scaling factor \( K = 6 \) for early-type galaxies (e.g., Beifiori et al., 2014; Nigoche-Netro et al., 2016). We do not use the Sérsic-index-dependent scaling factor as suggested in, e.g.,\(^{2}\)

\(^{2}\)The ellipticity defined here follows the convention in extragalactic astronomy; it is different from that of the eccentricity of an elliptical orbit, which is \( \sqrt{1-q^2} \).
Figure 2.2: The comparison between the derived VDFs of the CosmoDC2 catalog and the observed VDF at z < 1.5. Left: The local (z < 0.1) VDF. The VDF of the CosmoDC2 catalog accurately matches the spectroscopically measured VDF in Sohn et al. (2017) and Hasan & Crocker (2019). Right: The VDF at 0.3 < z < 1.5. The gray area marks the region where Bezanson et al. (2012) suggest that their galaxy sample has a low completeness. The CosmoDC2 VDF is in good agreement with the observed VDF up to z = 1.5. At the less massive end, the CosmoDC2 catalog gives a higher number density than the observation, which can be explained by the incompleteness of the observed galaxy sample.

Bertin et al. (2002), as CosmoDC2 uses a bulge + disk composition to describe the galaxy morphology instead of a Sérsic profile.

Figure 2.2 compares the derived VDF of the CosmoDC2 mock galaxy catalog with the observed galaxy VDF. The local VDF has been accurately measured spectroscopically (e.g., Choi et al., 2007). We find that the VDF of the CosmoDC2 catalog at z < 0.1 is fully consistent with the recent spectroscopic measurements (Sohn et al., 2017; Hasan & Crocker, 2019). At higher redshifts, it is challenging to measure the velocity dispersion of a large galaxy sample and thus the VDF spectroscopically. The right panel of Figure 2.2 compares the CosmoDC2 catalog VDF with the observed VDF at z < 1.5 from Bezanson et al. (2012), who measure the dynamical mass of galaxies in the COSMOS and UDF fields using photometric data. The CosmoDC2 catalog is also in good agreement with the observed values except at the low-mass end at z > 0.6, which can be due to incompleteness of the galaxy sample in the observations. Figure 2.2 suggests that our method successfully reproduces the observed VDF at least at z < 1.5.
In this work, we only include galaxies with $\sigma > 50$ km s$^{-1}$, which corresponds to an Einstein radius limit of $\theta_{E,\text{lim}} = 4\pi (\sigma_{\text{lim}}/c)^2 = 0\arcsec 072$. Lensing systems with smaller $\theta_E$ are difficult to identify even with space telescopes like the Hubble Space Telescope (HST).

We notice that the observed $M_* - \sigma$ relation has significant scatter, which can be partly explained by the effect of galaxy sizes ($R_e$). The observed scatter of Equation 2.2 is $\sim 0.02$ dex for the velocity dispersion (e.g., Auger et al., 2010). This small scatter has little influence on the galaxy VDF, and is not considered in this work.

In addition to the deflector galaxy, we add an external shear and convergence to each lensing system. We use the CosmoDC2 catalog to estimate the redshift evolution of external shears and convergences. Specifically, we divide the mock galaxy catalog into redshift bins with $\Delta z = 0.01$, then fit the distribution of external convergences ($\kappa$) and the two components of the external shear ($\gamma_1, \gamma_2$) in each redshift bin as normal distributions centered at zero. For both $\kappa$ and $\gamma$, the best-fit standard deviation can be well-described by a linear function of log$(1 + z)$:

$$
\sigma_\kappa = 0.057 \times \log(1 + z) - 0.001 \\
\sigma_{\gamma_1 \text{ or } \gamma_2} = 0.040 \times \log(1 + z) - 0.001
$$

(2.3)

For each lensing system, we use the best-fit normal distribution to generate a random set of $(\kappa, \gamma_1, \gamma_2)$ values.

Note that the convergence and shears in the CosmoDC2 catalog have some limitations. Specifically, the convergence and shears from ray-tracing simulations depend on the resolution (i.e., the smoothing scale) of the density field, especially at small scales (e.g., Hilbert et al., 2009). To perform the ray-tracing simulation, CosmoDC2 divides the mass particles from the cosmological simulation into discrete shells and estimates the surface mass density of each shell on a HEALPix grid with $\text{Nside}=4096$, which corresponds to a resolution of 0.5 arcminutes on the

---

3The actual distributions of external convergence and shears can be skewed. We use normal distributions in the sake of simplicity, since the convergence and shears have effectively no impact on the predicted number of lensed quasars.
sky. Korytov et al. (2019) compares the cosmic shear power spectrum of the CosmoDC2 catalog and theoretical predictions, showing that the difference is negligible at $\ell < 1000$ and is $\lesssim 10\%$ at $\ell \sim 4000$. As the values of $\kappa$ and $\gamma$ are small (Equation 2.3), this systematic uncertainty has effectively no impact on the number of lensed quasars.

### 2.2.2 Quasar Population

We simulate quasars at $z < 7.5$ using python code SIMQSO\(^4\) (McGreer et al., 2013). SIMQSO generates mock catalogs of quasars by randomly sampling the absolute magnitude and the redshift of quasars according to the input QLF. For each simulated quasar, SIMQSO generates a mock spectrum which includes a broken power-law continuum and a number of emission line components, and sets a random sightline to model the intergalactic medium (IGM) absorption based on observed IGM models. We use the default continuum, emission line and IGM parameters in SIMQSO which has been shown to successfully reproduce the observed quasar color distribution (e.g., Ross et al., 2013) and has been widely used to estimate the completeness of quasar surveys (e.g., Jiang et al., 2016; Yang et al., 2016; McGreer et al., 2018; Wang et al., 2019b).

We use a double power-law to describe the QLF:

$$
\Phi(M, z) = \frac{\Phi_*}{10^{0.4(M - M^*)}(\alpha + 1) + 10^{0.4(M - M^*)(\beta + 1)}}
$$

(2.4)

where $\alpha$ and $\beta$ are faint and bright end slopes of the QLF, $M^*$ is the break absolute magnitude and $\Phi_*$ is the normalization. Specifically, the QLF we adopt at different redshifts are as follows:

1. At $0 < z < 3.5$, we adopt the pure luminosity evolution (PLE, for $z < 2.2$) and the luminosity evolution and density evolution (LEDE, for $2.2 < z < 3.5$) model from Ross et al. (2013). We use the relations in the Appendix B of Ross et al. (2013) to convert the absolute magnitudes to $M_{1450}$, i.e., the absolute magnitude at rest-frame 1450Å.

\(^4\)https://github.com/imcgreer/simqso
Figure 2.3: QLFs adopted by this work. We use the QLF from Ross et al. (2013) to model quasars at $z < 3.5$, and compile a number of literature to construct the QLF at $z > 3.5$ (see text for details). *Left:* The Ross et al. (2013) QLFs. *Right:* The high-redshift QLFs we use from the literature. At $z \gtrsim 4$, the values of $M^*$, $\alpha$ and $\beta$ do not show strong redshift evolution, suggesting that a linear interpolation between the redshifts is adequate.

2. At $3.5 < z < 7.5$, we compile a number of QLFs at different redshifts from the literature. These QLFs are summarized in Table 2.2.2. At $3.5 < z < 6.7$, we use linear interpolation to obtain the QLF parameters at a certain redshift. At $z > 6.7$, we keep $\alpha$, $\beta$ and $M^*$ fixed to the $z = 6.7$ value, and apply a density evolution of $\log \Phi_*(z) = \log \Phi_*(z = 6.7) - 0.78 \times (z - 6.7)$, following Wang et al. (2019b).

Figure 2.3 shows the QLF used in this work at different redshifts. Both Table 2.2.2 and the right panel of Figure 2.3 suggest a subtle redshift evolution of $M^*$, $\alpha$ and $\beta$ at high redshift. In addition, a linear interpolation of $\log \Phi_*$ is equivalent to an exponential redshift evolution of $\Phi_*$, which has been shown to well-describe the number density of high-redshift quasars (e.g., Jiang et al., 2016; McGreer et al., 2018; Wang et al., 2019b). Using linear interpolations between redshifts is thus sufficient to accurately capture the evolution of QLFs at high redshift.

Following the Active Galactic Nuclei (AGN) chapter in the LSST science book (LSST Science Collaboration et al., 2009), we only include quasars with $M_i < -20$ in our simulated quasar sample, which is equivalent to $M_{1450} < -19.1$ according to
the relation in Ross et al. (2013). For faint AGN below this flux limit, host galaxies
start to dominate their emissions. These objects require a survey strategy that is
different from classical quasars and is out of the scope of this paper.

The area of the CosmoDC2 catalog is 440 deg$^2$, which is 1.067% of the whole
sky. To make the simulated lens catalog equivalent to an all-sky catalog, we scale
the QLF by a factor of 2%/1.067% = 1.874, so that the number of expected lensed
quasars in the CosmoDC2 field equals to 1/50 of the sky. By running 50 random
realizations (see Section 2.2.3), we can obtain a mock catalog that is equivalent to
the all-sky area.
Table 2.1. QLF parameters from the literature

<table>
<thead>
<tr>
<th>Redshift</th>
<th>$\log \Phi_*$</th>
<th>$M^*$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Mpc$^{-3}$mag$^{-1}$)</td>
<td>(mag)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.1 &lt; z &lt; 2.2$</td>
<td>$-5.96^{+0.02}_{-0.06}$</td>
<td>$-21.364^{+0.05}<em>{-0.11} - 2.5 \times (1.241^{+0.10}</em>{-0.028} z - 0.249^{+0.006}_{-0.017} z^2)$</td>
<td>$-1.16^{+0.02}_{-0.04}$</td>
<td>$-3.37^{+0.03}_{-0.05}$</td>
</tr>
<tr>
<td>$2.2 &lt; z &lt; 3.5$</td>
<td>$-5.89^{+0.15}<em>{-0.25} - 0.689^{+0.021}</em>{-0.027} \times (z - 2.2)$</td>
<td>$-25.00^{+0.34}<em>{-0.46} - 0.809^{+0.033}</em>{-0.166} \times (z - 2.2)$</td>
<td>$-1.31^{+0.52}_{-0.19}$</td>
<td>$-3.45^{+0.35}_{-0.21}$</td>
</tr>
<tr>
<td>$z = 3.9$</td>
<td>$-6.58 \pm 0.01$</td>
<td>$-25.36 \pm 0.13$</td>
<td>$-1.30 \pm 0.05$</td>
<td>$-3.11 \pm 0.07$</td>
</tr>
<tr>
<td>$z = 5.0$</td>
<td>$-7.36^{+0.56}_{-0.81}$</td>
<td>$-25.78^{+1.35}_{-1.10}$</td>
<td>$-1.21^{+1.36}_{-0.64}$</td>
<td>$-3.44^{+0.66}_{-0.84}$</td>
</tr>
<tr>
<td>$z = 6.0$</td>
<td>$-7.96^{+0.28}_{-0.42}$</td>
<td>$-24.90^{+0.75}_{-0.90}$</td>
<td>$-1.23^{+0.44}_{-0.34}$</td>
<td>$-2.73^{+0.23}_{-0.31}$</td>
</tr>
<tr>
<td>$z = 6.7$</td>
<td>$-8.43 \pm 0.23$</td>
<td>$[-24.90]$</td>
<td>$[-1.23]$</td>
<td>$-2.51 \pm 0.39$</td>
</tr>
</tbody>
</table>

1The pure luminosity evolution (PLE) model in Ross et al. (2013).
2The luminosity evolution and density evolution (LEDE) model in Ross et al. (2013).
3The parameterized QLF at $z = 3.9$ from Akiyama et al. (2018).
4The parameterized QLF at $z = 5$ from Kim et al. (2020).
5The parameterized QLF at $z = 6$ from Matsuoka et al. (2018a) (the standard model in their Table 5).
6The QLF at $z \sim 6.7$ by fitting the binned QLF from Wang et al. (2019b), with the faint-end slope and $M^*$ fixed to the values in Matsuoka et al. (2018a).

Note. — The parameters are defined in Equation 2.4. Absolute magnitudes in this table are $M_{1450}$. The square brackets marks quantities that are fixed when fitting the QLFs. Specifically, Wang et al. (2019b) only measures the bright end of $z \sim 6.7$ QLF, and we fit the double power-law by fixing the $M^*$ and $\alpha$ values to that of Matsuoka et al. (2018a). For $3.5 < z < 6.7$, we perform linear interpolation to obtain QLF model parameters. For $z > 6.7$, we assume a pure density evolution of $\log \Phi_*(z) = \log \Phi_*(z = 6.7) - 0.78 \times (z - 6.7)$, following Wang et al. (2019b).
2.2.3 Lensing Pipeline

After producing the mock deflector galaxies and the simulated background quasars, we generate one realization of the mock lens catalog as follows. We first assign random positions in the CosmoDC2 field to the simulated quasars, and assign random position angles to the mock galaxies. Then, for each galaxy, we identify all quasars that have distances to the galaxy smaller than $5\theta_{E,max}$, where $\theta_{E,max} = 4\pi(\sigma/c)^2$ is the maximum possible Einstein radius for the deflector galaxy. For each matched galaxy-quasar pair, we assign a random external convergence and a random external shear according to the best-fit probabilistic distribution described in Section 2.2.1, and use GLAFIC (Oguri, 2010) to model the lensing configuration. GLAFIC is software that provides fast and flexible lensing analysis for a variety of deflector mass and source emission models. We then select multiply-imaged quasars into the mock lens catalog.

We run 50 random realizations to generate a mock lens catalog that is equivalent to the all-sky area. For quasars at $z > 5$, due to their low number density, we run 1,000 realizations (i.e., $20 \times$ all-sky area) to reduce the Poisson noise.

2.2.4 Broad-Band Photometry and Mock Images

We calculate the magnitudes of the simulated lensed quasars in the LSST $ugrizy$, UKIDSS $JHK$ (Lawrence et al., 2007), and WISE W1, W2 bands. We use the simulated spectra generated by SIMQSO to calculate the synthetic magnitudes of quasars. For deflector galaxies, CosmoDC2 provides the synthetic LSST $ugrizy$ magnitudes and 30 top-hat SED points in the wavelength range $1000\,\text{Å} < \lambda < 20000\,\text{Å}$. We directly adopt the synthetic $ugrizy$ magnitudes from the CosmoDC2 catalog, and estimate the $JHK$ and W1, W2 magnitudes using the SED. Specifically, the SED covers $J$ and $H$ at all redshifts, and we linearly interpolate the SED points to build spectra and calculate the magnitudes. At low redshifts, the SED does not cover the wavelengths of $K$, W1 and/or W2. In this case, we fit the SED using the galaxy templates from Brown et al. (2014), and use the best-fit template to estimate
the magnitude in that filter.

To further explore how survey strategies work for real data, we generate simulated LSST images in the ugrizy bands for the mock lensed quasars. We use point sources to describe the lensed images of the background quasar, and use a bulge (de Vaucouleurs) + disk (exponential) two-component model to describe the foreground galaxy. CosmoDC2 provides the radius and the SED of both components for each galaxy. We then generate mock images using \textit{galfit} (Peng et al., 2002) for each lensing system. Specifically, the images are created as follows:

1. \textit{The Point Spread Function (PSF)}: The PSF of each image is modeled as an 2-D Gaussian profile. The full-width half maximum (FWHM) is drawn from a normal distribution with a standard deviation $\sigma_{\text{FWHM}} = 0.02 \times \text{mean}$, where \text{mean} is the mean PSF size of the corresponding filter. The axis ratio is uniformly drawn in $[0, 0.95]$, and the position angle is uniformly drawn from $[0, 2\pi)$. This approach mimics the PSF variation in real observations.

2. \textit{Image Noises}: We follow the methods in the LSST review paper (Ivezić et al., 2019, also see https://smtn-002.lsst.io/) to calculate the errors of image pixels. The random noise of a pixel consists of the background noise and the Poisson noise of the source flux. These noises are calculated based on Equation 5 in Ivezić et al. (2019).

Table 2.2.4 summarizes the mean PSF sizes and depths we used in the simulation. Figure 2.4 shows some examples of the SED and simulated LSST images of a mock lensed quasar, which exhibit a large diversity in observational features.

The final data products are a mock catalog of lensed quasars and mock LSST images for each simulated lensed quasar. We provide a number of properties in the mock catalog, including the lensing configuration, the time delay of each lensed image, and the magnitudes of the quasar and the deflector. These properties are summarized in Table 2.2.4.
Figure 2.4: A gallery of mock lensed quasars. We show several typical cases for both the low-redshift and the high-redshift sample. The synthetic magnitudes are in LSST $ugrizy$, UKIDSS $JHK$, and WISE W1, W2 bands. The mock lensed quasars show a large diversity in observational features in magnitudes, colors, lensing separations, etc. It is worth noticing that (1) the deflector galaxy is always at the center of the image, and is usually blended with the fainter lensed image in doubly-lensed systems; (2) the low-redshift example of a “marginally discoverable” lens (third row, left) has a small lensing separation of $0''5$, and the high-redshift example (third row, right) has a marginally detectable fainter image (which is blended with the deflector galaxy). See section 2.3 for the definition of discoverable lenses.
Table 2.2. PSF size and depth of the mock LSST images

<table>
<thead>
<tr>
<th>Filter</th>
<th>5 - σ Depth</th>
<th>PSF FWHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>u</td>
<td>26.1</td>
<td>0.81</td>
</tr>
<tr>
<td>g</td>
<td>27.4</td>
<td>0.77</td>
</tr>
<tr>
<td>r</td>
<td>27.5</td>
<td>0.73</td>
</tr>
<tr>
<td>i</td>
<td>26.8</td>
<td>0.71</td>
</tr>
<tr>
<td>z</td>
<td>26.1</td>
<td>0.69</td>
</tr>
<tr>
<td>y</td>
<td>24.9</td>
<td>0.68</td>
</tr>
</tbody>
</table>

1For point sources.

2.3 Statistics of Lensed Quasars

In this section we present the numbers of lensed quasars that are discoverable in various imaging surveys. Following Oguri & Marshall (2010), we define a lensed quasar to be discoverable if: (1) it has a separation of $\Delta \theta > \frac{2}{3} \times \text{FWHM}$ and (2) has the second (for doubly-imaged lenses) or third (for quadruply-imaged lenses) brightest lensed image detectable at a 5σ level. Here, the separation of a lensed quasar is defined as the largest separation between any pairs of the lensed images. These criteria ensure that the lensing structure is at least marginally resolved.

Note that while lensed quasars that meet the criteria above are in principle discoverable, in practice, their recovery faces a number of technical challenges. In real observations, the flux contribution of the lensing galaxy, low signal to noise ratios, and the presence of large number of both astrophysical and instrumental contaminants could result in significant challenges and low success rate in lensed quasar selection, especially for the marginally discoverable cases. As a result, most of the
## Table 2.3. Description of the columns in the mock catalog

<table>
<thead>
<tr>
<th>Name</th>
<th>Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quasar Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qidx</td>
<td>int</td>
<td>The unique identifier of the mock quasar.</td>
</tr>
<tr>
<td>zq</td>
<td>float</td>
<td>The redshift of the mock quasar.</td>
</tr>
<tr>
<td>qra</td>
<td>float</td>
<td>The RA of the mock quasar, in degrees.</td>
</tr>
<tr>
<td>qdec</td>
<td>float</td>
<td>The Dec of the mock quasar, in degrees.</td>
</tr>
<tr>
<td>absMag</td>
<td>float</td>
<td>The absolute magnitude at rest-frame 1450Å ($M_{1450}$).</td>
</tr>
<tr>
<td>QSOMags$^1$</td>
<td>list of floats</td>
<td>The synthetic, unmagnified magnitudes of the mock quasar in LSST $ugrizy$, UKIDSS $JHK$ and WISE W1, W2.</td>
</tr>
<tr>
<td><strong>Deflector Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>galaxy_id</td>
<td>int</td>
<td>The “galaxy_id” field of the mock galaxy in the CosmoDC2 catalog.</td>
</tr>
<tr>
<td>zg</td>
<td>float</td>
<td>The redshift of the mock galaxy.</td>
</tr>
<tr>
<td>gra</td>
<td>float</td>
<td>The RA of the mock galaxy, in degrees.</td>
</tr>
<tr>
<td>gdec</td>
<td>float</td>
<td>The Dec of the mock galaxy, in degrees.</td>
</tr>
<tr>
<td>rmaj</td>
<td>float</td>
<td>The semi-major axis of the galaxy, in arcsec.</td>
</tr>
<tr>
<td>rmin</td>
<td>float</td>
<td>The semi-minor axis of the galaxy, in arcsec.</td>
</tr>
<tr>
<td>pa_lens</td>
<td>float</td>
<td>The position angle of the galaxy (following GLAFIC conversion).</td>
</tr>
<tr>
<td>stellar_mass</td>
<td>float</td>
<td>The stellar mass in $M_\odot$.</td>
</tr>
<tr>
<td>sigma</td>
<td>float</td>
<td>The velocity dispersion, in km/s.</td>
</tr>
<tr>
<td>GalMags</td>
<td>list of floats</td>
<td>The synthetic total magnitudes of the deflector galaxy in LSST $ugrizy$, UKIDSS $JHK$ and WISE W1, W2.</td>
</tr>
<tr>
<td>shear_1, shear_2</td>
<td>float</td>
<td>The $x$– and $y$– components of the external shear.</td>
</tr>
<tr>
<td>convergence</td>
<td>float</td>
<td>The external convergence.</td>
</tr>
<tr>
<td><strong>Lensing Configuration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nimg</td>
<td>int</td>
<td>The number of lensed images.</td>
</tr>
<tr>
<td>image_x$^2$</td>
<td>list of floats</td>
<td>the $x$-coordinates of the lensed images, in arcsec (following GLAFIC conversion).</td>
</tr>
<tr>
<td>image_y</td>
<td>list of floats</td>
<td>the $y$-coordinates of the lensed images, in arcsec.</td>
</tr>
<tr>
<td>image_mu</td>
<td>list of floats</td>
<td>the magnification of the lensed images.</td>
</tr>
<tr>
<td>image_dt</td>
<td>list of floats</td>
<td>the time delay of the lensed images.</td>
</tr>
</tbody>
</table>

$^1$All magnitudes are noiseless.

$^2$In our conversion, the deflector galaxy locates at $x = 0$ and $y = 0$. 
Table 2.4. Predicted number of discoverable lensed quasars

<table>
<thead>
<tr>
<th>Survey</th>
<th>Area</th>
<th>Filter</th>
<th>5σ Depth</th>
<th>FWHM</th>
<th>(N_{\text{qso}})</th>
<th>(N_{\text{lens}})</th>
<th>(N_{\text{quad}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS1(^2)</td>
<td>(3 \times 10^4)</td>
<td>(i \text{ or } z)</td>
<td>23.1 or 22.3</td>
<td>1.11 or 1.07</td>
<td>(4.8 \times 10^6)</td>
<td>928</td>
<td>132</td>
</tr>
<tr>
<td>Legacy(^3)</td>
<td>(1.4 \times 10^4)</td>
<td>(r \text{ or } z)</td>
<td>23.5 or 22.5</td>
<td>1.18 or 1.11</td>
<td>(2.6 \times 10^6)</td>
<td>474</td>
<td>70</td>
</tr>
<tr>
<td>LSST(^4)</td>
<td>(2 \times 10^4)</td>
<td>(i \text{ or } z)</td>
<td>26.8 or 26.1</td>
<td>0.71 or 0.69</td>
<td>(7.4 \times 10^6)</td>
<td>2377</td>
<td>193</td>
</tr>
<tr>
<td>DES(^5)</td>
<td>(5 \times 10^3)</td>
<td>(i \text{ or } z)</td>
<td>23.8 or 23.1</td>
<td>0.88 or 0.83</td>
<td>(1.0 \times 10^6)</td>
<td>241</td>
<td>33</td>
</tr>
<tr>
<td>HSC(^6)</td>
<td>(1.4 \times 10^3)</td>
<td>(i \text{ or } z)</td>
<td>26.4 or 25.5</td>
<td>0.56 or 0.63</td>
<td>(5.1 \times 10^5)</td>
<td>165</td>
<td>14</td>
</tr>
</tbody>
</table>

\(^1\)The columns: \(N_{\text{qso}}\) is the total number of quasars beyond the flux limit (estimated using the mock quasar sample), \(N_{\text{lens}}\) is the predicted number of all discoverable lensed quasars, and \(N_{\text{quad}}\) is the predicted number of discoverable quad lenses. Magnitude limits are for point sources.

\(^2\)The Pan-STARRS1 (PS1) survey, adopted from Chambers et al. (2016)

\(^3\)The DESI legacy imaging survey, adopted from Dey et al. (2019)

\(^4\)The LSST survey, adopted from Ivezić et al. (2019)

\(^5\)The Dark Energy Survey (DES), adopted from DES Collaboration et al. (2021b)

\(^6\)The Hyper Suprime-Cam Wide Survey (HSC), adopted from Aihara et al. (2018b) and Aihara et al. (2018a)

Note. — See Section 2.3 for the criteria of a discoverable lensed quasar. We consider two filters for each survey, and a lensed quasar will be counted if it is discoverable in either band.
Figure 2.5: Predicted number of discoverable lensed quasars in LSST and PS1 survey, as a function of redshift. See Section 2.3 for the definition of a discoverable lensed quasar. **Left:** the low-redshift \((z_{\text{qso}} < 5)\) sample. We run 50 realizations which is equivalent to an all-sky area, and scale the predicted number of discoverable lenses according to the survey area (listed in Table 2.3). There are \(\sim 900\) discoverable lensed quasars in the PS1 survey, and LSST will increase this number to about \(2.4 \times 10^3\). Note that there is a bump at \(z_{\text{qso}} \sim 4\) in lensed quasar numbers, where we switch from the low-redshift QLF to the high-redshift one. The faint-end values of the two QLFs are not fully consistent at this redshift. **Right:** the high-redshift \((z_{\text{qso}} > 5)\) sample. We run 1,000 realizations to reduce the statistical error since the number density of high-redshift quasars is low. The mock catalog suggests that only a handful of lensed quasars at \(z_{\text{qso}} > 5\) are discoverable in current imaging surveys like PS1, and there will be \(\sim 45\) discoverable in LSST. At \(z \gtrsim 6.5\), there are effectively no discoverable lensed quasars even for LSST.
lensed quasars reported in recent surveys are bright ones with large separation (e.g., Lemon et al., 2020; Chan et al., 2020). A main application of this catalog is to facilitate development of selection algorithm to enable effective recoveries of these discoverable systems.

In Table 2.3, we list the predictions for a number of current and next-generation imaging surveys. As specific examples, Figure 2.5 shows the redshift distribution of discoverable lensed quasars in the PS1 survey and the LSST survey. Our mock catalog suggests that current sky surveys such as PS1 can in principle find \( \sim 900 \) lensed quasars, and LSST will increase this number to \( \sim 2.4 \times 10^3 \). Most of the lensed quasars have redshift \( 1.5 < z_{\text{qso}} < 3.5 \), which is consistent with the statistics of the current sample of discovered lensed quasars. Due to the decline of the quasar number density, the numbers of discoverable lensed quasars drop quickly at \( z_{\text{qso}} \gtrsim 4 \). Figure 2.5 suggests that only a handful of lensed quasars at \( z_{\text{qso}} > 5 \) are discoverable in imaging surveys such as PS1, and there will be \( \sim 45 \) discoverable in LSST. There are effectively no discoverable lensed quasars at \( z > 7 \) even in the LSST survey. We will discuss in more detail the lensed fraction of \( z \gtrsim 6 \) quasars using both the mock catalog and analytical models in a subsequent paper (Yue et al., 2022b).

It is worth noticing that, even though LSST reaches more than three magnitudes deeper than PS1, the number of total quasars and discoverable lensed quasars are only 1.5 and 2.6 times higher. This is due to the fact that we only include quasars with \( M_i < -20 \) in the catalog so that the AGNs are not host-dominated. LSST reaches significantly fainter flux levels than this limit especially at low redshift.

Figure 2.6 shows the distribution of lensing separations of the mock lensed quasars. About \( \sim 80\% (50\%) \) of these lenses have separation larger than \( 0''5 (1''0) \). These numbers are consistent with the results in, e.g., Hilbert et al. (2008) and Collett (2015). A ground-based imaging survey with a PSF size of \( \sim 0''7 \) will be sufficient to make most of the lenses marginally resolved, which suggests that the bottleneck of lensed quasar surveys is the image depth rather than the PSF size.

We note that the predicted number of lensed quasars in this work is about two to three times lower than that of OM10, especially for deep surveys like HSC and
Figure 2.6: The distribution of the image separations of lensed quasars. The dashed line marks the limit above which the mock catalog is complete. About $\sim 50\%$ of the lensed quasars have separation $\Delta \theta > 1''$, and $\sim 80\%$ of them have $\Delta \theta > 0'.5$.

LSST. The main reason is the difference in the simulated quasar sample; OM10 adopt a steeper QLF faint-end slope, and more importantly, they do not apply a cut in the absolute magnitude, while we only include quasars with $M_i < -20$. In addition, OM10 use a deflector VDF without redshift evolution, in contrast to the CosmoDC2 VDF which declines with redshift. In Section 2.3.1 we discuss in more detail the difference between our mock catalog and that of OM10, where we show that the number of lensed quasars is consistent after taking the difference in QLFs and VDFs into account.

2.3.1 Comparison with the OM10 Mock Catalog

The mock catalog in OM10 and this work have a number of significant differences: (1) The two mock catalogs adopt different QLFs and VDFs. (2) We only include quasars with $M_i < -20$, while OM10 extend the quasar sample to as faint as $M_i \sim -18$. And (3) OM10 only include lensing systems with separation $\Delta \theta > 0'.5$, while our mock catalog has more compact lenses down to $\Delta \theta \sim 0'.1$.

To perform meaningful comparisons, we select subsamples of the two mock catalogs that have $M_{1450} < -24$, $z_{\text{qso}} < 3$ and $\Delta \theta > 0'.5$. In this redshift and magnitude range, the QLFs used in OM10 and this work are close to each other (Figure 2.7, left panel). We also compare the VDFs adopted in OM10 and this work in the middle
Figure 2.7: Comparison between this work and OM10. **Left:** the QLF adopted in this work and OM10. The two QLFs are nearly identical for quasars at $z \lesssim 2.5$ with $M_{1450} < -24$. **Middle:** The VDF adopted in this work and OM10. We show the CosmoDC2 VDF at $z < 1.5$ which covers the majority of deflector galaxies for $z_{\text{qso}} \lesssim 2.5$. OM10 uses the local VDF in Choi et al. (2007) and assumes no redshift evolution. The local VDF of the CosmoDC2 catalog is close to the one used by OM10, and the OM10 VDF is slightly higher than the CosmoDC2 VDF at higher redshift. **Right:** The predicted number of lensed quasars at $z < 3$. We only select lensed quasars with $\Delta \theta > 0.5$ and $M_{1450} < -24$, at which the QLFs used by OM10 and this work are close to each other. Our mock catalog predicts slightly smaller numbers of lensed quasars than OM10 when only restricted to these parameter ranges, and the different can be fully explained by the adopted QLFs and VDFs (see text for details).

In short, OM10 adopt the local VDF from Choi et al. (2007) and assume no redshift evolution. This VDF is close to the CosmoDC2 catalog VDF at $z < 0.1$, but it predicts higher number density than the CosmoDC2 VDF beyond the local universe. These differences can fully account for the predicted numbers of lensed quasars, shown in the right panel of Figure 2.7. After applying the luminosity, redshift and lensing separation cuts mentioned above, our mock catalog gives a slightly smaller number of lensed quasars compared to OM10. The main difference between the two mock catalogs appears at $z_{\text{qso}} \sim 2.2$, where the OM10 catalog has $\sim 30\%$ more lenses than this work. This can be explained by the difference in the VDFs at $z_d \gtrsim 0.3$, where the OM10 VDF is $\sim 0.3$ dex higher than the CosmoDC2 VDF.

The comparison suggests that our method of generating a mock catalog is reliable, and illustrates the importance of using accurate and updated QLFs and VDFs. Another major difference between OM10 and this work is that we apply an absolute
magnitude cut of $M_i < -20$ for quasars. This cut is used in the LSST science book, and the main motivation is that fainter AGNs have distinct observational features compared to normal type-I quasars. For these objects, host galaxy emissions start to dominate the SEDs and make their morphology extended. Modeling the host galaxy emission is out of the scope of this work, and we expect that these faint AGNs require a very different candidate selection strategy.

AGN variability offers a possible way of finding host-dominated lensed AGNs (e.g., Kochanek et al., 2006). OM10 presented the number of lensed quasars that are brighter than the $10\sigma$ limits of yearly-stacked images for each sky survey, which can be identified via their variability. Variability analysis will be especially useful in the LSST survey, providing a critical supplement to image-based candidate selection methods.

### 2.3.2 Quadruply-Lensed Quasars and Time Delays

One important application of lensed quasars is to measure key cosmological parameters, especially the Hubble constant $H_0$. This relies on measuring the time delays between the lensed images of the background quasar. Compared to doubly-imaged ones, quadruply lensed quasars are preferred since they provide more time delays and give stronger constraints.

Table 2.3 gives the number of discoverable quadruply-lensed quasars in various sky surveys. Our mock catalog suggests that there are $\sim 200$ quadruply-lensed quasars discoverable in LSST. The impact of survey depth on the number of discoverable lenses is less significant for quads compared to doubly-lensed quasars, which is a result of magnification bias and the criteria of discoverable lenses (Section 2.3). In general, the fainter image in doubly-lensed systems has a low magnification and is often de-magnified ($\mu = 0.6$), while the third brightest images in quad lenses have significantly higher magnifications ($\mu = 2.6$). Given the luminosity limit of our quasar sample ($M_i < -20$), most quadruply-lensed quasars are brighter than the LSST survey limit. Nonetheless, the image quality of LSST is still critical to the search of these bright quads, since the better depth and spatial resolution are
essential for a high completeness and efficiency in the candidate selection.

Figure 2.8 further illustrates the number of discoverable quadruply lensed quasars as a function of survey depth. We assume an LSST-like PSF size of 0.07 when counting the number of discoverable lenses, and normalize the numbers to the area of LSST. We also include the number of doubly-lensed quasars and unlensed quasars for comparison. Note that the magnitudes in Figure 2.8 correspond to the second brightest image for doubly-lensed quasars and the third brightest image for quadruply-lensed ones, in accordance with the definition of discoverable lensed quasars. At $m_i \lesssim 25$, about $\sim 10\% - 15\%$ of the discoverable lensed quasars are quads. The magnification bias makes the fraction of quads decreases with survey depth, in agreement with the result in OM10. The number of quad lenses drops quickly at $m_i \gtrsim 26$, which is a result of the absolute magnitude cutoff.

The probability for a deflector to generate a quadruple lens increases with its ellipticity. The right panel of Figure 2.8 illustrates the distribution of the deflector
ellipticity, $e_{\text{galaxy}}$, for the CosmoDC2 mock catalog and the one adopted by OM10. OM10 assumes a Gaussian distribution of ellipticity for all the galaxies, while CosmoDC2 use a more complicated approach, where the probabilistic distribution of $e_{\text{galaxy}}$ depends on the bulge-to-disk ratio and the galaxy luminosity. The overall ellipticity distribution of OM10 and CosmoDC2 are similar; correspondingly, the predicted quad fractions at $m_i \lesssim 26$ of this work and OM10 are consistent within statistical errors.

The external shears are another key factor in generating quadruply-lensed systems. Luhtaru et al. (2021) analyzed 39 quadruply-lensed quasars, 15 of which were found to be “shear-dominated”. This result illustrates the significant contribution of external shears in making quad lenses. OM10 assumes a log-normal distribution of the total external shear, $\gamma$, with a mean of 0.05 and a deviation of 0.2 dex. In comparison, we adopt a redshift-evolving distribution for the external shear. At most redshifts ($z \lesssim 6$), the shear distribution in this work has a lower mean value and a more prominent tail at large shears. We thus expect that the quad lenses in our mock catalog and OM10 have different properties (e.g., lensing configurations).

It might be useful to keep in mind that the external convergence and shear distributions used in this work are derived from ray-tracing simulations, which are subject to a number of systematic uncertainties (Section 2.2.1).

Figure 2.9 illustrates the distribution of time delays and the magnitude differences between lensed images. The typical time delays are $10^1 - 10^3$ days. In most of the doubly-imaged lenses, the brighter image arrives earlier than the fainter one, while for quad lenses, the second or third brightest images arrives first in most cases. These results are consistent with the findings in OM10. For a typical time delay of $\sim 10^2$ days, a cadence of $\sim 5$ days gives a good measurement of the light curves (Suyu et al., 2017). The cadence of the LSST survey is not sufficient to provide accurate light curves for most of the lensed quasars. As such, LSST will mainly serve as a deep imaging survey for discoveries, and follow-up monitoring is needed to measure the time delays of lensed quasars.
The magnitude difference and the time delays of the mock lensed quasars. The subscripts represent the order of arrival, e.g., \( m_1 \) is the magnitude of the image that arrives first. Most lensed quasars have time lags \( \sim 10^1 - 10^3 \) days. For doubly-imaged quasars, the brighter image usually arrives earlier than the fainter one. For quad lenses, the first-arriving image is the second or third brightest in most cases.

### 2.3.3 Statistics of Deflector Galaxies

Figure 2.10 shows the distribution of the deflector redshift, \( z_d \), of simulated lensed quasars. The distribution peaks at \( z_d \sim 0.7 \) and drops to a negligible level at \( z_d \gtrsim 2.5 \). This picture is consistent with OM10 and previous modeling of the deflector galaxy population (e.g., Hilbert et al., 2008; Wyithe et al., 2011). For comparison, we also include two samples of sources at fixed redshift, \( z_s = 2 \) and \( z_s = 5 \). Although CosmoDC2 only includes galaxies at \( z < 3 \), Figure 2.10 indicates that only a small fraction of lensed quasars has deflector redshift \( z_d > 3 \).

In Figure 2.11, we present the \( i \)-band magnitudes of the deflectors and the deflector-to-quasar magnitude difference of the simulated lensed quasars. In nearly all of the LSST-discoverable lensed quasars, the deflector galaxy is brighter than the LSST limit, which suggests that the impact of the foreground galaxy must be considered when selecting lensed quasar candidates. The deflector-to-quasar magnitude difference spreads over a wide range, and the number of quasar-flux-dominated lenses is comparable to that of galaxy-flux-dominated ones. Correspondingly, we expect a large diversity in observed features for lensed quasars. The galaxy-dominated lenses are likely generated by a massive, luminous deflector and have large image separations, and the quasar-dominated ones usually have compact lensing structures.
Figure 2.10: The distribution of deflector redshifts. The $y$–axis denotes the probability for a background source to be lensed by a foreground deflector at a certain redshift. The gray histogram represents all the simulated lensed quasars, which peaks at $\sim 0.7$ and has a median of 0.87. The distribution drops to a negligible level at $z_d > 2.5$. We also include two samples of background sources at $z_s = 2$ and $z_s = 5$ for comparison.

These objects may require very different candidate selection techniques.

2.4 Discussion

2.4.1 Choices of QLFs and VDFs

Section 2.3.1 shows the impact of QLFs and VDFs on the modeling of the lensed quasar population. In this section, we discuss the systematic uncertainties introduced by our choices of QLFs and VDFs.

At $z \lesssim 4$, the QLFs are well-measured down to at least $M_{1450} \sim -22$ (e.g., Richards et al., 2006a; Ross et al., 2013; Palanque-Delabrouille et al., 2016), and these measurements are in good agreement with each other. Since most LSST-discoverable lensed quasars have $z_{\text{qso}} < 4$, the QLFs adopted by this study provide a good description of the quasar population. Meanwhile, there are still debates about the QLFs at $z \gtrsim 5$. A few studies give a faint-end slope of $\alpha \sim -2$ (e.g., Jiang et al., 2016; Yang et al., 2016; McGreer et al., 2018), while others suggest $\alpha \sim -1.3$ (Niida et al., 2020; Kim et al., 2020). In this work, we use the HSC-based high-redshift QLFs which have $\alpha \sim -1.3$. If the faint-end slope turns out to be
Figure 2.11: Left: The $i$-band magnitudes of the deflector galaxies of LSST-discoverable lensed quasars. Nearly all of them are beyond the LSST magnitude limit. The impact of deflector galaxy flux must be considered when selecting lensed quasar candidates. Right: the difference between the magnitude of the deflector galaxy and the magnified magnitude of the quasar. The quasar magnitude corresponds to the total flux of all the lensed images of the background quasar. The magnitude difference spreads over a wide range, which suggests a large diversity in observed features for lensed quasars.

steeper, there might be more faint lensed quasars at high redshift.

Unlike the QLFs, the deflector VDFs are not well constrained by spectroscopic surveys beyond the local universe. Many previous studies assume no evolution of the VDF (e.g., OM10; Wyithe & Loeb, 2002). However, recent analysis of strong lensing systems indicate that the VDF drops towards high redshift (e.g., Geng et al., 2021), which is consistent with photometric measurements (e.g., Bezanson et al., 2011, 2012) and the CosmoDC2 VDFs used in this work. It is thus critical to use redshift-evolving VDF models when modeling the lens population.

In addition, we consider the uncertainties introduced by the scaling relation used to calculate the velocity dispersion (Equation 2.2). Many studies suggest that the $M_{\text{dyn}}/M_*$ ratio evolves with galaxy mass (e.g., Auger et al., 2010; Nigoche-Netro et al., 2016) and redshift (e.g., Beifiori et al., 2014). We thus consider the impact of an evolving $M_{\text{dyn}}/M_*$ ratio in the following form:

$$\log \frac{M_{\text{dyn}}}{M_*} = a \log(M_*) + b \log(1 + z) + c$$ (2.5)

This is equivalent to an evolution of $\sigma \propto M_*^{a/2}(1 + z)^{b/2}$. 
Figure 2.12: Left: Testing the mass-dependence term in the velocity dispersion. We use the $M_\ast - M_{\text{dyn}}$ relation in Auger et al. (2010) to calculate the velocity dispersion of CosmoDC2 mock galaxies. The resulting local VDF does not match the observations. Right: Testing the redshift-evolution term in the velocity dispersion. We add the redshift-evolution from Beifiori et al. (2014) to the velocity dispersion inferred by Equation 2.2 for mock galaxies. The redshift evolution is subtle, and the inferred VDF is consistent with observation at $z < 1$. However, at $z > 1$, the redshift evolution term starts to overpredict the number of galaxies at the massive end.

We first set $z = 0$ to check the mass-dependence. We adopt the relation in Auger et al. (2010), which suggests $\log\left(\frac{M_\ast}{M_\odot}\right) = 0.885 \log\left(\frac{M_{\text{dyn}}}{M_\odot}\right) + 0.905$. The left panel of Figure 2.12 shows the derived VDF at $z < 0.1$, which does not reproduce the observed local VDF. We thus conclude that adding the mass-dependent term does not give a good description of the galaxy velocity dispersion for the CosmoDC2 catalog.

We then keep the relation $M_\ast/M_{\text{dyn}} = 0.557$, and adopt the redshift evolution from Beifiori et al. (2014), which gives $\sigma \propto (1 + z)^{0.12}$. The right panel of Figure 2.12 compares the resulted CosmoDC2 VDF and the observed values in Bezanson et al. (2012). At $z < 1$, the redshift evolution makes little difference compared to the original one, and the resulting VDF is still consistent with observation. At $1.2 < z < 1.5$, the redshift-evolving estimation starts to over-predict the number of galaxies at the massive end, where the observed galaxy sample should be complete. This difference might be more significant at $z > 1.5$ where the redshift evolution term gets larger.
In addition to the redshift evolution of the velocity dispersion, Beifiori et al. (2014) also report a redshift evolution of the $M_{\text{dyn}}/M_*$ ratio, which follows $M_{\text{dyn}}/M_* \propto (1 + z)^{-0.30 \pm 0.12}$. By including both trends into Equation 2.2, we get a very subtle evolution of $\sigma \propto (1 + z)^{-0.03}$. This result might explain why a non-evolving $M_{\text{dyn}}/M_*$ ratio correctly reproduces the observed VDF. In any case, the redshift dependence is subtle. Given the median redshift of deflectors ($z_d \sim 0.7$, Figure 2.10), the evolution will not strongly influence the resulting lensing statistics.

### 2.4.2 Mock Catalog Completeness

The CosmoDC2 catalog models dark matter halos down to $10^{9.8} M_\odot$. According to the $M_{\text{galaxy}} - M_{\text{halo}}$ relation (e.g., Behroozi et al., 2013), a halo of $\sim 10^{10} M_\odot$ usually has $M_{\text{galaxy}}/M_{\text{halo}} \sim 10^{-3}$. The CosmoDC2 catalog should thus be complete down to $M_* \sim 10^7 M_\odot$. Our $\sigma$-selected deflector sample has a minimum stellar mass of $M_* = 10^7.8 M_\odot$, which means that the mock catalog is complete for deflectors with $\sigma > 50 \text{ km s}^{-1}$ and $z_d < 3$.

The main incompleteness comes from the fact that CosmoDC2 catalog only covers $0 < z < 3$. Figure 2.10 suggests that the majority of mock lenses have $z_d < 3$. Wyithe et al. (2011) studies the strong lensing statistics for $z \sim 8$ galaxies, and their results suggest that $> 80\%$ of lensing systems have $z_d < 3$. This fraction will be higher for lower source redshifts, and we conclude that our mock catalog is highly complete.

### 2.4.3 Finding More Lensed Quasars Beyond the “Discoverable” Ones

In the discussion above, we adopt the definition of a discoverable lensed quasar in OM10, which requires an image separation $\Delta \theta > \frac{2}{3} \times \text{FWHM}$ and a $5\sigma$ detection of the second or third brightest lensed image. The motivation of these criteria is that the lensed quasar images need to be at least marginally resolved and barely visible. Under this definition, about one third of the lensed quasars in the mock catalog are
not discoverable in LSST. In this subsection, we discuss possible strategies to find these undiscoverable lenses.

Identifying undiscoverable lensed quasars is important because the discoverable criteria can miss some unique and interesting objects. As an example, Fan et al. (2019b) report a compact lens with separation of 0\textquoteleft\textquoteleft\textquoteright\textquoteright 2 at $z = 6.51$, J0439+1634, which is the only known multiply-imaged quasar at $z > 5$ to date. J0439+1634 has a large magnification of $\mu = 51$, making it the brightest quasar at $z > 6$ in nearly all wavelengths, and provides so far the best case study of a high-redshift quasar. Although J0439+1634 is bright ($m_z = 19.49 \pm 0.02$), the small separation makes it undiscoverable in ground-based imaging surveys. Besides, undiscoverable lenses enable many studies that are impossible otherwise. For instance, small-separation lenses can probe the mass distribution of less massive deflector galaxies.

We first consider the image separation criteria. Figure 2.6 suggests that LSST-like seeing ($\sim 0\textquoteleft\textquoteleft\textquoteright\textquoteright 7$) will be sufficient to resolve the majority of lensed quasars. Lensing systems with smaller separation have less massive and thus fainter deflector galaxies. These objects are likely to have quasar-dominated flux with a marginally extended morphology. Since the deflector galaxy and the lensed quasar images are blended, we expect that neither a PSF nor a regular Sérsic profile can provide a good description of their morphology, according to which we can identify these compact lenses. This method is used in Fan et al. (2019b) to identify the $z = 6.51$ compact lensed quasar using ground-based imaging, which is later confirmed by HST imaging.

For lensed quasars that do not meet the second criterion, i.e., the fainter lensed image being brighter than the survey limit, some of them can be discovered by identifying the brighter image as a quasar. If there is a bright galaxy next to it, it is a hint that the quasar might be lensed. Figure 2.11 suggests that the deflector galaxy is detectable in the LSST survey in most lensing systems. Follow-up deep imaging and/or spectroscopy can confirm if there is a fainter lensed image. This technique requires decent image de-blending to disentangle the brighter lensed quasar image and the deflector galaxy.

As we show in Figure 2.11, the observational features of lensed quasars have a
large variety. A complete search of them requires a number of distinct candidate selection methods. The mock catalog can serve as training and testing sets for future surveys of lensed quasars with LSST data.

2.5 Summary and Conclusion

We present a mock catalog of gravitationally lensed quasars at $z_{\text{qso}} < 7.5$, which covers quasars with $M_i < -20$ and deflector galaxies with $\sigma > 50 \text{ km s}^{-1}$, using updated QLFs and deflector VDFs. We generate synthetic magnitudes and simulated LSST images for the mock lensed quasars. Using the mock catalog, we explore the expected outputs of lensed quasar surveys, and investigate possible strategies to find more lensed quasars. Our main conclusions are:

1. The number of discoverable lensed quasars in current sky surveys is $\sim 10^3$, and LSST will increase this number to $\sim 2.4 \times 10^3$. Most of the lensed quasars have image separation $\Delta \theta > 0.5\text{'}$, which will at least get marginally resolved in the LSST survey.

2. There will be $\sim 200$ discoverable quadruply lensed quasars in the LSST survey. The fraction of quad lenses among all discoverable lensed quasars is $\sim 10\% - 15\%$ at $m_i \lesssim 25$, and decreases with survey depth. The typical time delays between lensed images are $\sim 10^1 - 10^3$ days. Follow-up high-cadence observations are necessary to accurately measure the time lags.

3. The mock lensed quasars show a large diversity in their observational features, from deflector-dominated ones which are lensed by bright, massive galaxies with large lensing separations, to quasar-dominated ones which are likely more compact. The variety in observational features requires complex candidate selection techniques.

We estimate that the mock catalog has a high completeness for lensed quasars with Einstein radius $\theta_E > 0.07\text{'}$. This range covers nearly all the lensed quasars that can be identified in the near future, including compact ones that can be discovered
by the Euclid Telescope (Scaramella et al., 2021) and the Roman Space Telescope (Spergel et al., 2015). The mock catalog and the simulated images will be powerful tools in future lensed quasar surveys, especially for high-redshift lensed quasars which are not well-explored by previous studies. As an example, we have used the mock catalog to analyze the population of lensed quasars at $z > 5$ and design a new survey for these objects (Yue et al. in prep). Besides, although we focus on traditional (point-like) type-I quasars, the code can be easily adapted to study other background source populations, including fainter AGNs and normal galaxies.
CHAPTER 3

Analytical Model of the High-Redshift Gravitationally Lensed Quasar Population†

3.1 Introduction

In the past two decades, extensive searches have been carried out for high-redshift quasars (e.g., Jiang et al., 2016; Matsuoka et al., 2018a; Wang et al., 2019b, 2021; Yang et al., 2019a, 2020), which led to a sample of more than 1200 quasars at $z > 4.5$ (more than 500 at $z > 5$). In Fan et al. (2019b), we reported the discovery of the first gravitationally lensed quasar at $z > 5$ (J0439+1634 at $z = 6.52$). High-redshift lensed quasars such as J0439+1634 are valuable, as they enable in-depth investigations of distant quasars with enhanced sensitivity and resolution (e.g., Yang et al., 2019b; Yue et al., 2021b).

Since the commissioning of the Sloan Digital Sky Survey (SDSS, York et al., 2000), there have been efforts to search for lensed quasars at $z \gtrsim 4$ (e.g., Richards et al., 2004, 2006b). However, there are only three lensed quasars at $z > 4.5$ known so far, including a couple at $z = 4.8$ (McGreer et al., 2010; More et al., 2016) and J0439+1634 at $z = 6.52$. The fraction of lensed ones among high-redshift quasars is $\sim 0.2\%$ (without correcting for survey incompleteness). Meanwhile, a number of theoretical studies have predicted a high lensed fraction ($\gtrsim 4\%$) for high-redshift quasars due to strong magnification bias (e.g., Wyithe & Loeb, 2002; Comerford et al., 2002; Pacucci & Loeb, 2019). Fan et al. (2019b) and Pacucci & Loeb (2019) suggested that survey incompleteness is a plausible reason for the discrepancy.

We notice, however, that previous models were based on various assumptions on the quasar population and the deflector population, which might introduce systematic errors that have not been fully explored. In particular, many previous studies were based on early measurements of quasar luminosity functions (QLFs)

†This Chapter was published as Yue et al. (2022b)
and deflector velocity dispersion functions (VDFs). Since recent observations have provided more accurate measurements of the QLFs and VDFs, there is a need to re-investigate the models of high-redshift lensed quasars to fully understand the discrepancy between theoretical predictions and observations.

In this paper, we revisit the models of high-redshift lensed quasar population and investigate possible systematic uncertainties. We adopt recent measurements of VDFs and QLFs, and use both analytical methods and mock catalogs to model the properties of high-redshift lensed quasars. Our analysis addresses the following questions: (1) what are the uncertainties of the theoretical models, (2) how many lensed quasars can we detect in current and future sky surveys, and (3) how strong is the tension between the theoretical predictions and the observations? Answering these questions is critical to the design of a complete search for high-redshift lensed quasars with the upcoming sky surveys like the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST, Ivezić et al., 2019).

This paper is organized as follows. We describe the analytical method in Section 3.2 and the mock catalog method in Section 3.3. We present the predicted statistics of high-redshift lensed quasars in Section 3.4, discuss the implications in Section 3.5, and summarize our conclusions in Section 3.6. We use a flat $\Lambda$CDM cosmology with $H_0 = 70$ km s$^{-1}$Mpc$^{-1}$ and $\Omega_M = 0.3$ throughout this work. All magnitudes are z-band AB magnitudes unless further specified.

### 3.2 Analytical Method

In this Section, we describe the analytical model of the high-redshift lensed quasar population, including the lensing model, the deflector galaxies, and the source quasars. In this paper, we use “strongly-lensed” and “multiply-imaged” interchangeably.
3.2.1 Lensing Models

We use singular isothermal spheres (SIS) to describe the mass profile of deflector galaxies. SIS and singular isothermal ellipsoids (SIE, e.g., Kormann et al., 1994) are widely used in models of strong-lensing systems (e.g., Wyithe & Loeb, 2002; Oguri & Marshall, 2010). SIS is parameterized by its Einstein radius, \( \theta_E = 4\pi \left( \frac{\sigma}{c} \right)^2 \frac{D_{ds}}{D_s} \), where \( \sigma \) is the velocity dispersion of the deflector, \( D_{ds} \) and \( D_s \) are the angular diameter distances from the source to the deflector and the observer, respectively.

Following Wyithe et al. (2011), the lensing optical depth of a population of SIS deflectors, i.e., the probability for a source at redshift \( z_s \) to be multiply-imaged, is

\[
\tau_m = \int_0^{z_s} dz_d \int d\sigma \phi(\sigma, z_d) \frac{d^2V_c}{d\Omega dz_d} \pi \theta_E(\sigma, z_d, z_s)^2 D_d^2
\]  

(3.1)

where \( z_d \) is the deflector redshift, \( \phi(\sigma, z_d) \) is the deflector VDF, \( \frac{d^2V_c}{d\Omega dz_d} = (1 + z_d)^3 \frac{c}{dz_d} \frac{dt}{dz} \) is the differential comoving volume, and \( D_d \) is the angular diameter distance at \( z_d \).

The lensing optical depth \( \tau_m \) describes the fraction of lensed ones among all background sources. In flux-limited surveys, however, some lensed sources can get detected which are intrinsically fainter than the survey limit. Consequently, we expect that the observed lensed fraction in flux-limited surveys should be higher than \( \tau_m \). The magnification bias \( B \) quantifies this effect,

\[
B = \int_{\mu_{\text{min}}}^{+\infty} d\mu \frac{p(\mu)N(> L_{\text{lim}}/\mu)}{N(> L_{\text{lim}})}
\]  

(3.2)

where \( p(\mu) \) is the probabilistic distribution of the magnification of lensed sources, \( L_{\text{lim}} \) is the survey flux limit, and \( N(> L_{\text{lim}}) \) is the number of background sources brighter than \( L_{\text{lim}} \). According to Equation 3.2, if there are \( N(> L_{\text{lim}}) \) strongly-lensed objects that are intrinsically brighter than \( L_{\text{lim}} \), the number of strongly-lensed objects that have magnified flux brighter than \( L_{\text{lim}} \) should be \( B \times N(> L_{\text{lim}}) \).

In our analytical model, we use \( \mu_{\text{min}} = 2 \) which is the minimum magnification for multiply-imaged systems generated by SIS. If the survey has a high spatial resolution which can resolve the lensing structure, we shall use \( p(\mu) = 2/(\mu - 1)^3 \) which describes the magnification of the brighter lensed image. Similarly, for surveys
with poor resolution where the lensed images are blended, we use 
\[ p(\mu) = \frac{8}{\mu^3} \]
which corresponds to the total magnification.  

Although SIS deflectors cannot produce quad lenses, the fraction of quad lenses among all lenses is small (\(\sim 10\%\), e.g., Oguri & Marshall, 2010).

The fraction of lensed sources among all sources that have *apparent* luminosity brighter than \(L_{\text{lim}}\) is (e.g., Mason et al., 2015):

\[
F_{\text{multi}} = \frac{B\tau_m}{B\tau_m + B'(1 - \tau_m)}
\]

where \(B'\) is the magnification bias of sources that are not multiply imaged. The major source of \(B'\) is weak lensing. As shown in Mason et al. (2015), for sources at \(z_s > 6\), the magnification caused by weak lensing (\(\mu_{\text{weak}}\)) is distributed within a narrow range around \(\mu_{\text{weak}} \approx 1\), with almost all sources having \(0.7 < \mu_{\text{weak}} < 1.3\). According to Equation 3.2, we use \(B' \approx 1\) in our model. This approximation is also adopted in Wyithe et al. (2011).

### 3.2.2 Deflector Population

The observed galaxy VDFs are well described by Schechter functions (e.g., Choi et al., 2007; Chae, 2010). In this work, we use the following form of the Schechter function:

\[
\phi(\sigma, z_d)d\sigma = \phi_* \left( \frac{\sigma}{\sigma_*} \right)^a \exp \left[ - \left( \frac{\sigma}{\sigma_*} \right)^b \right] \frac{d\sigma}{\sigma}
\]

where \(\phi(\sigma, z_d)d\sigma\) is the number density of galaxies at redshift \(z_d\) which have velocity dispersions within \((\sigma, \sigma + d\sigma)\). In principle, the parameters \(\phi_*, \sigma_*, a\) and \(b\) may evolve with redshift. We adopt the parameterized local VDF from Hasan & Crocker (2019) and assume the redshift evolution of \(\phi_*\) and \(\sigma_*\) suggested by Geng et al. (2021). The resulting VDF has \(\phi_* = 6.92 \times 10^{-3}(1 + z_d)^{-1.18} \text{Mpc}^{-3}\), \(\sigma_* = 172.2 \times (1 + z_d)^{0.18} \text{km s}^{-1}\), \(a = -0.15\) and \(b = 2.35\). In Figure 3.1, we compare our parameterized VDF and the observed values in the local universe from Sohn.

---

1. We derive \(p(\mu)\) from Equation 8.34(c) – 8.35(b) from Schneider et al. (1992). Also see Wyithe et al. (2011).
Figure 3.1: The velocity dispersion functions. **Left:** The local VDF. The black solid line shows the analytical VDF adopted by this work, and the red solid line illustrates the VDF of the CosmoDC2 catalog. The open squares and circles are the observed VDFs from Sohn et al. (2017) and Hasan & Crocker (2019). All these VDFs are in good agreement with each other. We also include the VDFs used in previous studies of high-redshift lensed quasars for comparison (Section 3.5). **Right:** the VDF beyond the local universe. The black dots are the observed VDF from Bezanson et al. (2012), and the gray area marks the mass range where Bezanson et al. (2012) suggested that their galaxy sample has low completeness. Both the analytical VDF and the CosmoDC2 VDF are consistent with the observed ones except at low-mass end, which could be a result of the incompleteness of the observed galaxy sample. We also include the local VDF from Choi et al. (2007) (black dashed line) which are used in previous studies to model the deflectors.

et al. (2017) and Hasan & Crocker (2019), and at 0.3 < z_d < 1.5 from Bezanson et al. (2012). This parameterized, redshift-evolving VDF is in good agreement with the observations at z_d < 1.5, which covers the majority of the deflector population (Section 3.4.1).

### 3.2.3 Quasar Luminosity Functions

QLFs are usually parameterized by a broken power-law (Pei, 1995), which accurately describes the QLF at all redshifts (e.g., Kulkarni et al., 2019):

\[
\Phi(M, z_d) = \frac{\Phi^*(z_d)}{10^{0.4(\alpha+1)(M-M^*)} + 10^{0.4(\beta+1)(M-M^*)}}
\]  

(3.5)
where $\alpha$ and $\beta$ are the faint- and bright-end slopes, $M^*$ is the break magnitude and $\Phi^*$ is the normalization. For $M$ and $M^*$, we use the absolute magnitude at rest-frame 1450 Å ($M_{1450}$) throughout this paper. We use the median value of $m - M_{1450}$ of $z_s \sim 6$ mock quasars (Section 3.3) to convert apparent magnitudes $m$ to absolute magnitudes. Since the shape of QLF determines the magnification bias, we explore a wide range of $\alpha, \beta$ and $M^*$ values in Section 3.4.2. The normalization $\Phi^*$ has no impact on the magnification bias and thus the multiply-imaged fraction.

### 3.3 Mock Catalog Method

In addition to the analytical model, we also use mock catalogs to investigate the lensed quasar population. The mock catalogs not only serve as independent tests which validate the analytical model, but also illustrate the impact of some simplifications, e.g., using SIS instead of SIE to describe deflectors. The methods of mock catalog generation is described in Yue et al. (2022a) and is briefly summarized here.

We use the CosmoDC2 mock galaxy catalog (Korytov et al., 2019) to model the deflector galaxy population. CosmoDC2 is a synthetic mock catalog that covers a sky area of 440 deg$^2$ and out to redshift $z = 3$. We use SIEs to describe the mass profile of the deflector galaxies. CosmoDC2 provides the mass, half-light radius and ellipticity of each mock galaxy, and we estimate the velocity dispersion using a simple relation based on the virial theorem. The resulting VDF is shown in Figure 3.1, which is consistent with observations and is in good agreement with the analytical VDF out to $z_d \sim 1.5$.

We then generate two types of mock background sources to study the statistics related to $\tau_m$ and $B$, respectively:

1. Sources at fixed redshifts. They are used to model the redshift evolution of $\tau_m$. Since the CosmoDC2 catalog only contains mock galaxies at $z_d < 3$, it does not provide complete descriptions of the deflectors for sources at $z_s > 3$. As such, we generate three sets of mock sources at $z_s = 1, 2$ and 3.

2. Simulated quasars at $5.5 < z_s < 6.5$, which are used to investigate the im-
pact of QLF parameters on the magnification bias, $B$, for quasars at $z_s \sim 6$. We generate these simulated quasars using SIMQSO (McGreer et al., 2013). SIMQSO generates mock quasar catalogs with simulated spectra according to the input QLF. Specifically, we consider two sets of QLFs: the “steep” set which has $\alpha = -2, M^* = -27$ (e.g., McGreer et al., 2018), and the “shallow” set which has $\alpha = -1.3, M^* = -25$ (e.g., Matsuoka et al., 2018a). For each set, we consider a range of bright-end slopes at $-3.4 \leq \beta \leq -2.6$. We adopt the normalization $\Phi^*$ from Matsuoka et al. (2018a).

We assign random positions to the sources in the CosmoDC2 field, and use GLAFIC (Oguri, 2010) to perform lens modeling and identify multiply-imaged systems. For the second type of sources (i.e., mock quasars), we run multiple random realizations to suppress the Poisson noise, such that the number of lensed quasars is equivalent to $20 \times$ the sky area.

It is useful to keep in mind the differences between the mock catalog and the analytical model. The VDFs of the two approaches are close but have small differences, and we use SIE instead of SIS to model the deflectors for mock catalogs. Besides, the mock quasars generated by SIMQSO have a variety of continuum slopes and emission line properties to match the observed distribution, while in the analytical model, we perform a simple $k$–correction to convert absolute to apparent magnitudes. As a result, we do not expect that the two methods should give the same result. Nonetheless, as we will show in Section 3.4, the predictions of the two approaches are largely consistent, which suggests that both approaches are solid.

3.4 Results

We focus on the multiply-imaged fraction $F_{\text{multi}}$ of $z_s \sim 6$ quasars in this work. According to Equation 3.3, $F_{\text{multi}}$ is a function of the lensing optical depth $\tau_m$ and the magnification bias $B$. We describe the two factors respectively in the following subsections.
3.4.1 Lensing Optical Depth

We first present the lensing optical depth, $\tau_m$, predicted by the analytical model and the mock catalog. Since the CosmoDC2 catalog does not have galaxies at $z_d > 3$, we only use the mock catalog to calculate the lensing optical depth for sources at $z_s \leq 3$. In general, we validate our methods by comparing the mock catalog and the analytical model at $z_s \leq 3$, and make predictions using the analytical model at higher redshifts.

The left panel of Figure 3.2 illustrates the predicted lensing optical depth as a function of source redshift $z_s$. At $z_s \leq 3$, the outputs of the analytical model and the mock catalog are similar. This comparison suggests that both the analytical method and the mock catalog method are reliable. We use the analytical model to predict the lensing optical depth at $z_s > 3$, which gives $\tau_m = 1.8 \times 10^{-3}$ for $z_s = 6$.

To further validate our analysis, we include two additional experiments: the “CosmoDC2 Discrete Sum” where we change the integration in Equation 3.1 to discrete sums and calculate $\tau_m$ using the CosmoDC2 VDF, and the “Mock Catalog SIS” where we use SIS instead of SIE deflectors when making the mock catalogs. In other words, the two experiments calculate $\tau_m$ using the analytical method and the mock catalog method assuming the same VDF and deflector mass profile. Figure 3.2 shows that the two experiments provide the same result, further proving that both methods are reliable. The comparison between the default mock catalogs and the Mock Catalog SIS suggests that using SIS instead of SIE will overestimate the lensing optical depth by $\sim 10\%$. This is consistent with the result in Kormann et al. (1994), who shows that the cross section of an SIE decreases with ellipticity.

In the right panel of Figure 3.2, we show the redshift distribution of the deflector galaxies for sources at $z_s = 6$, predicted by the analytical model. We also include the predictions for $z_s = 3$ from both the analytical model and the mock catalog for comparison. Again, the two methods give very similar results for $z_s = 3$. For sources at $z_s = 6$, the distribution of deflectors peaks at $z_d \sim 1$, and the majority of deflector galaxies have $0.5 \lesssim z_d \lesssim 2.5$. It is thus important to correctly model the
Figure 3.2: Left: The lensing optical depth as a function of source redshift. Note that the CosmoDC2 catalog do not contain deflectors at $z_d > 3$, so that the mock catalog is incomplete at $z_s > 3$. The analytical model has $\tau_m = 1.8 \times 10^{-3}$ at $z_s = 6$. At $z_s < 3$, the outputs of the analytical model and the mock catalog are similar. We include two additional experiments to validate the results, the CosmoDC2 Discrete Sum and the Mock Catalog SIS, which further demonstrate that our analysis is reliable (see text for details). Right: The distribution of deflector redshifts. For $z_s = 6$, the analytical model suggests that most of the deflectors have $0.5 < z_d < 2.5$, and the distribution peaks at $z_d \sim 1$. We also include the predictions for $z_s = 3$ as a test of our results, for which the analytical method is consistent with the mock catalog.

redshift evolution of VDF out to $z_d \gtrsim 2$ in order to make precise predictions of $\tau_m$ for high-redshift sources.

It is worth noticing that the analytical model gives a higher lensing optical depth than the mock catalogs. The reason is that the analytical VDF is higher than the CosmoDC2 VDF at $z_d \sim 0.6$ (as shown in Figure 3.1) and $z_d \gtrsim 2$. Since VDFs beyond the local universe has yet been well-constrained by observations, the predicted lensing optical depth will inevitably has some systematic errors. We will discuss this point in more detail in Section 3.5.1.

Figure 3.3 shows the distribution of lensing separation for $z_s = 6$ sources. About 65% (85%) of the lenses have separation larger than $1''$ (0''.5). Most of these lenses will be resolved in surveys such as the LSST which has a seeing of $\sim 0''.7$. The only lensed quasar at $z > 5$ known to date, J0439+1634 at $z = 6.52$, is a compact system with $\Delta\theta = 0''.2$. Figure 3.3 suggests that such systems are rare but possible.
Figure 3.3: Distribution of the lensing separation for sources at $z_s = 6$, predicted by the analytical model. About 85% of the lenses have $\Delta \theta > 0\farcs5$, and the fraction is $\sim 65\%$ for $\Delta \theta > 1\farcs0$. The dashed line shows the separation of J0439+1634 at $z_s = 6.52$, which is the only known lensed quasar at $z_s > 5$. Systems like J0439+1634 are rare but are still possible.

3.4.2 Magnification Bias and the Lensed Fraction

The magnification bias $B$ is a function of the QLF and $p(\mu)$ (Equation 3.2). We use both the analytical model and the mock catalogs to calculate the magnification bias. Note that although the mock catalog does not have deflectors at $z_d > 3$, it still provides correct description for the magnification bias of high-redshift sources, since $z_d$ and the VDF do not go into the calculation of $B$.

Figure 3.4 shows the predicted magnification bias as a function of QLF parameters. We consider two specific cases, where the top panel shows the magnification bias for the brightest lensed image ($B_1$) and the lower panel shows the bias for the total magnification ($B_{\text{total}}$). According to Figure 3.4, the magnification biases of the two cases differ by $\sim 50\%$. Observationally, these two cases correspond to the high and low spatial resolution limits, respectively, and real surveys lie between the two special cases.

The predictions of the analytical model and the mock catalogs are consistent in all cases. This comparison suggests that using SIS in the analytical model gives a
good approximation of the real case where deflectors are elliptical. The magnification bias is higher for steeper bright-end slope and decrease with the survey depth, which is a result of the flatter faint-end slope compared to the bright-end. To calculate the lensed fraction $F_{\text{multi}}$, we adopt $\tau_m = 1.8 \times 10^{-3}$ for $z_s = 6$. High-redshift quasar searches in the SDSS have a depth of $m_{\text{lim}} = 20.2$ (e.g., Richards et al., 2002), for which the lensed fraction is $F_{\text{multi}} \sim 1\% - 3\%$. Present-day surveys for $z \sim 6$ quasars usually have $m_{\text{lim}} \gtrsim 22$ (e.g., Jiang et al., 2016; Matsuoka et al., 2018a), which corresponds to $F_{\text{multi}} \sim 0.4\% - 0.8\%$. The lensed fraction of the brightest $z_s \sim 6$ quasar population ($m_z \lesssim 19$) can be as high as $\sim 6\%$ for the steep bright-end slopes.

In Figure 3.5, we further explore the impact of QLF parameters on the lensed fraction using the analytical model. Note that an absolute magnitude of $-25$ corresponds to an apparent magnitude of 21.5 at $z_s = 6$ under our assumed $k$–correction. The multiply-imaged fraction can be as high as $\gtrsim 10\%$ at the corner of $M_{\text{lim}} - M^* \lesssim -2$ and $\beta \lesssim -3.5$. The lensed fraction drops quickly to other regions. In particular, we highlight the predictions for our fiducial QLF of $\alpha = -1.3, \beta = -2.75$ and $M^* = -25$, which are close to the measurements in Matsuoka et al. (2018a) and Wang et al. (2019b) for $z_s \gtrsim 6$ quasars. This fiducial QLF gives $F_{\text{multi}} \sim 1\% - 2\%$ for the brightest quasars ($m_z \lesssim 19$) and $F_{\text{multi}} \sim 0.4\% - 0.8\%$ for a survey depth of $m_z \approx 22$.

3.5 Discussion

3.5.1 The Lensed Fraction of High-redshift Quasars

We can now investigate the discrepancy between observations and theoretical models in the lensed fraction of high-redshift quasars. According to Equation 3.3, the multiply-imaged fraction $F_{\text{multi}}$ is a function of the lensing optical depth $\tau_m$ and the magnification bias $B$. We first consider the lensing optical depth $\tau_m$, which is determined by the deflector VDF. Figure 3.1 summarizes the VDFs adopted by previous models of high-redshift lensed quasars. In general, two approaches have
Figure 3.4: The magnification bias of $z_s \sim 6$ lensed quasars as a function of QLF parameters. We illustrate two cases, where the top panel shows the magnification bias for the brightest lensed image, and the bottom panel is for the total flux of all lensed images. In all cases, the analytical model (solid lines) agrees well with the mock catalog (dots with error bars), suggesting that the difference between SIS and SIE deflectors is small. We assume Poisson statistics for the mock catalog (note that the mock catalogs are equivalent to $20 \times$ sky area). The secondary y–axis marks the multiply-imaged fraction assuming $\tau_m = 1.8 \times 10^{-3}$. The magnification bias is higher for steeper faint-end slope and drops with survey depth. For the depth of current surveys of high-redshift quasars ($m_{\lim} \sim 22$), the lensed fraction is $F_{\text{multi}} \sim 0.4\% - 0.8\%$. 
Figure 3.5: Lensed fraction of $z_s \sim 6$ quasars as a function of QLF parameters, predicted by the analytical model. Similar to Figure 3.4, we assume $\tau_m = 1.8 \times 10^{-3}$, and illustrate the two cases that correspond to the brightest image (top) and the total flux (bottom). The contours mark levels of 10%, 3%, 1% and 0.3%, and these levels are also marked in the colorbars by yellow lines. For shallow surveys ($M_{\text{lim}} - M^* \lesssim -2$) and steep bright-end slopes ($\beta \lesssim -3.5$), the lensed fraction can be as high as $F_{\text{multi}} \gtrsim 10\%$. The fraction drops quickly towards other regions in the parameter space. The cyan dashed lines and asterisks highlight the fiducial QLF with $M^* = -25$, $\alpha = -1.3$ and $\beta = -2.7$, which has $F_{\text{multi}} \sim 1\% - 2\%$ for the brightest quasars and $F_{\text{multi}} \sim 0.4\% - 0.8\%$ for the depths of current surveys.
been adopted to model the deflector VDF:

(1) Convert galaxy luminosity functions to VDFs using empirical relations such as the Faber-Jackson relation (FJR, Faber & Jackson, 1976). This method was used in Wyithe & Loeb (2002) and Pacucci & Loeb (2019). Figure 3.1 shows that the VDFs in these two studies are close to each other, and both of them are significantly higher than the observed VDFs. Correspondingly, both studies predicted a lensed fraction that is several times higher than our model ($F_{\text{multi}} \gtrsim 5\%$ for $m_{\text{lim}} = 20.2$).

The reason why FJR leads to a high VDF is complicated. First, the FJR has a large scatter ($\gtrsim 10\, \text{km}\,\text{s}^{-1}$, e.g., Barone-Nugent et al., 2015). Applying a single FJR without scatters will not accurately reproduce the VDF. Second, the FJR is usually calibrated for massive galaxies with $\log \sigma \gtrsim 2.3$ (e.g., Focardi & Malavasi, 2012), which might not describe less massive galaxies correctly.

(2) Directly adopt the VDF from galaxy surveys. Most previous studies of this kind adopted the local VDF in Choi et al. (2007) and assumed no redshift evolution. For example, Wyithe et al. (2011) calculated the lensing optical depth out to $z_s \sim 11$ and gave $\tau_m = 3 \times 10^{-3}$ for $z_s = 6$. Recent observations suggested that the number density of massive galaxies decreases towards high redshifts (e.g., Chae, 2010; Bezanson et al., 2012), indicating that a non-evolving VDF overestimates the lensing optical depth. As shown in Figure 3.1, at $z_d \gtrsim 0.9$, the Choi et al. (2007) VDF is higher than the observations and the VDFs in this study, which explains the difference in the predicted $\tau_m$. Deflectors at this redshift range dominate the lensing optical depth (Figure 3.2). This comparison illustrates the importance of using a redshift-evolving VDF.

In addition to using VDFs to describe the deflector population, some studies also use cosmological simulations to calculate the lensing optical depth. For example, Hilbert et al. (2008) used the matter distribution from the Millennium simulation (Springel et al., 2005b) and performed ray-tracing to model the strong-lensing statistics. They predict a lensing optical depth that is comparable to our results ($\tau_m \sim 1.2 \times 10^{-3}$ at $z_s = 5.7$).

To summarize, it is critical to have an accurate description of the deflector popu-
lation when modeling the lensing statistics. The VDFs adopted by previous studies may have significant uncertainties and are usually overestimated, which could bias the predicted lensing optical depth by a factor of $\sim 2 - 4$. By applying recent measurements of VDFs, we reach a lower lensing optical depth and thus lower lensed fraction, which relieves the tension between the theoretical models and the observations. Since the VDF at $z_d \gtrsim 1.5$ is still largely unconstrained, the predicted $\tau_m$ in our model also has some systematic errors. We estimate this uncertainty to be $\sim 30\%$ by comparing the analytical VDF and the CosmoDC2 VDF, which are in good agreement at $z_d \lesssim 1.5$ but shows more differences at higher redshifts. After taking the uncertainties into account, the difference between the observed lensed fraction ($F_{\text{multi}} \approx 0.2\%$) and our model prediction ($F_{\text{multi}} \approx 0.4\% - 0.8\%$) is marginal.

We then consider the magnification bias $B$. Previous studies explored a wide range of QLF parameters, and the results are similar to this study. This is largely due to that SIS accurately describes the magnification distribution, $p(\mu)$, as suggested by Figure 3.4. However, it is worth noticing that most of the previous studies used the SDSS quasar survey depth ($m_{\text{lim}} = 20.2$) when predicting the lensed fraction. Present-day high-redshift quasar surveys have depths of $m_{\text{lim}} \gtrsim 22$, which lead to a lower expected lensed fraction than the SDSS depth.

3.5.2 The Number of Detectable High-Redshift Lensed Quasars in Current and Future Sky Surveys

Large and deep imaging surveys (e.g., the LSST survey and the Euclid survey (Scaramella et al., 2021)) will greatly enhance our ability of finding high-redshift quasars. In this Section, we estimate the number of high-redshift lensed quasars that can be detected in current and future sky surveys.

We use the analytical model to calculate the number of lensed quasars that can be detected given a certain survey depth.\footnote{The code for the analytical model is available at https://github.com/yuemh/lensQSOsim/tree/main/analytical} Specifically, we estimate the number of quasars that are intrinsically brighter than the survey limit using the QLF from
Matsuoka et al. (2018a), then calculate the number of detectable strongly-lensed quasars using Equation 3.3. To convert apparent magnitudes to absolute magnitudes for quasars at redshift \(z_s\), we generate mock quasars at \(z_s - 0.1 < z_{\text{mock}} < z_s + 0.1\) following the method in Section 3.3, and adopt the median value of \(m - M_{1450}\) of these mock quasars. Since the majority of the \(z_s \approx 6\) lenses will be resolved in upcoming sky surveys that have resolution of \(\lesssim 0\.
\text{'0}7\), we use the magnification bias for the brightest image, \(B = B_1\).

Figure 3.6 presents the predicted number density of lensed quasars that have the brightest lensed image detectable as a function of survey depth and redshift. We first consider optical imaging surveys, for which we use their \(z\)-band depths to compute the number of detectable lensed quasars. Current optical surveys have a typical depth of \(m_z = 22\) (e.g., the DESI Legacy Imaging Survey, Dey et al., 2019), yielding 14 detectable lensed quasars at \(z_s > 5.5\) over the whole sky. The LSST survey will reach a depth of \(m_z \approx 26\), with which we can detect 89 lensed quasars at \(z_s > 5.5\) over the whole sky (\(\sim 45\) over the 20,000 deg\(^2\) LSST footprint).

For quasars at \(z_s \gtrsim 6.5\), the \(\text{Ly}\alpha\) wavelength is redshifted to \(\lambda_{\text{obs}} \gtrsim 9120\AA\), and a substantial fraction of flux in \(z\)-band is absorbed by the highly neutral intergalactic medium. Consequently, the numbers of detectable lensed quasars in optical surveys drop quickly at this redshift range. We thus consider the expected outputs of near-infrared (NIR) imaging surveys. Present-day NIR surveys like the UKIRT Hemisphere Survey (UHS, Dye et al., 2018) and the VISTA Hemisphere Survey (VHS, McMahon et al., 2013) have depths of \(m_J \approx 21\) (AB magnitude), leading to \(\sim 1 - 2\) detectable lensed quasars at \(z_s > 6.5\) over the entire sky. In the upcoming \textit{Euclid} survey that will reach a depth of \(m_J = 24.5\), there will be 12 detectable lensed quasars at \(z_s > 6.5\) over the sky (\(\sim 4\) over the 14,000 deg\(^2\) \textit{Euclid} footprint).

To summarize, current sky surveys can only detect about 15 lensed quasars at \(z_s > 5.5\) over the entire sky. Next-generation sky surveys are needed to build the first large sample of high-redshift lensed quasars. Besides, our model predicts that the \textit{Euclid} survey can only detect \(\sim 0.6\) lensed quasars at \(z_s > 7.5\), meaning that
Figure 3.6: Numbers of detectable lensed quasars as a function of survey depth and quasar redshift. The solid lines correspond to the depths of the LSST survey ($m_z \approx 22$) and the Euclid survey ($m_{J,AB} \approx 24.5$), and the dashed lines are depths of current imaging surveys (see text for details). Current sky surveys (optical + NIR) can detect $\sim 15$ lensed quasars at $z_s > 5.5$ over the whole sky. The LSST survey can detect $\sim 45$ lensed quasars at $z_s > 5.5$ in its footprint, and the Euclid survey can detect $\sim 4$ at $z_s > 6.5$.

we are not likely to find a lensed quasar beyond this redshift in the near future.

Note that the numbers quoted above correspond to detectable lensed quasars; the outputs of real surveys heavily depend on the completeness of candidate selection methods and can be exceedingly complicated (see related discussion in Yue et al. 2021, submitted). If we only count the area with high galactic latitude (e.g., $|b| \gtrsim 30$ deg) and assume a high survey completeness (e.g., $\gtrsim 80\%$), the model suggests that we can only find a handful of lensed quasars at $z_s > 5.5$ in current sky surveys. The completeness of real surveys could be much lower (e.g., Fan et al., 2019b; Pacucci & Loeb, 2019). We expect that upcoming sky surveys like the LSST will find several tens of lensed quasars at $z_s > 5.5$.

3.6 Conclusion

We revisit the models of the high-redshift lensed quasar population, focusing on the lensed fraction $F_{\text{multi}}$ of $z_s \sim 6$ quasars. We adopt recent measurements of VDFs
and explore a wide range of QLF parameters, using both analytical methods and mock catalogs. Our models suggest a lensing optical depth of $\tau_m = 1.8 \times 10^{-3}$ for sources at $z_s = 6$, and a lensed fraction of $F_{\text{multi}} \sim 0.4\% - 0.8\%$ for ordinary QLF parameters and current survey depth ($m_{\text{lim}} \approx 22$). For the brightest $z_s \sim 6$ quasars ($m_z \lesssim 19$), the fraction is $\sim 2\%$ for ordinary QLF parameters and can be as high as $\sim 6\%$ for steep QLF bright-end slopes.

By comparing our models to previous studies, we illustrate that it is critical to use an accurate, redshift-evolving VDF. Inaccurate VDFs will bias the predicted $\tau_m$ by a factor of several. Adopting VDFs from recent measurements relieves the tension between the observed lensed fraction and the model predictions. As the VDF at $z_d \gtrsim 2$ is still poorly constrained, we estimate that our model still has a systematic uncertainty of $\sim 30\%$ for $\tau_m$ and thus $F_{\text{multi}}$.

Finally, we estimate the number of high-redshift lensed quasars that can get detected in present-day and future imaging surveys. Our model suggests that there are $\sim 15$ lensed quasars beyond the magnitude limit of current wide-area imaging surveys. Deeper surveys such as the LSST survey and the Euclid survey will find several tens of lensed quasars at $z_s > 5.5$ and are necessary for building the first large sample of high-redshift lensed quasars.
CHAPTER 4

Designing a Complete Survey for High-redshift Lensed Quasars

4.1 Introduction

Gravitationally lensed quasars are valuable but rare objects in extragalactic astronomy and cosmology. Intensive efforts have been taken in the past several decades to search for these objects. To date, there are ~ 220 known lensed quasars\(^1\), one of which is at \(z > 5\) (J0439+1634 at \(z = 6.5\); Fan et al., 2019b). In contrast, Chapter 2 shows that there are \(\gtrsim 1000\) lensed quasars discoverable in current sky surveys, among which \(\sim 10\) are at \(z > 5\). The fact that a substantial fraction of lensed quasars is missing is a result of the complex nature of lensed quasars. Most lensed quasars are unresolved or marginally-resolved in present-day ground-based imaging surveys, which have typical PSF sizes of \(\sim 1'' - 2''\) (Yue et al., 2022a). With different deflector SEDs, quasar SEDs and lensing configurations, the appearance of lensed quasars in imaging surveys exhibits a large diversity. A certain candidate selection strategy will be effective only for a subset of lensed quasars.

This project focuses on searching for lensed quasars at \(z \gtrsim 5\). Compared to low-redshift ones, high-redshift lensed quasars are fainter, featured by red optical colors due to IGM absorption. Before designing the candidate selection method for high-redshift lensed quasars, it is useful to overview the techniques used in previous studies, based on which I can find effective ways of identifying lensed quasars at high redshifts.

\(^1\)A compilation of discovered lensed quasars can be found in https://research.ast.cam.ac.uk/lensedquasars/
4.1.1 An Overview of Previous Lensed Quasar Survey Methods

The most common way of finding lensed quasars is to select objects with a quasar-like color and a lens-like shape in imaging surveys. Some other techniques have also been proven to be useful, including variability analysis, spectroscopic analysis and radio interferometry. In this subsection, I describe in detail the concepts of these methods, as well as their pros and cons.

(1) Image and photometry analysis

Most known lensed quasars are discovered based on their images in ground-based optical and near-infrared (NIR) sky surveys. Some of these surveys are described in Chapter 1. These studies adopted various candidate selection strategies. For example, the SDSS Quasar Lens Search (SQLS; e.g., Oguri et al., 2006) starts from the spectroscopically-confirmed SDSS quasar sample, and looks for quasars that are not well-fitted by a point spread function (PSF) or have a close companion with colors similar to the quasar. Agnello (2017) uses photometry from the WISE survey (Wright et al., 2010) to select objects that have quasar-like colors and are resolved into multiple components in the Gaia survey (Gaia Collaboration et al., 2016); Lemon et al. (2017) propose catalog-level quantities in the Gaia survey that can distinguish lensed quasars from other objects, which lead to the discovery of \( \sim 50 \) new lensed quasars (Lemon et al., 2018, 2019). Khramtsov et al. (2019) use random forests to classify objects from the Kilo Degree Survey (de Jong et al., 2013) as stars, galaxies and quasars, then select objects that are classified as a quasar with a non-stellar companion as lens candidates.

Despite the differences in the technical details, all these studies contain two fundamental pieces, i.e., color classification and morphological classification. The main idea is to find objects that have a mixed color between galaxies and quasars, and have shapes that indicate a possible lensing structure. Large-separation lenses that are resolved in imaging surveys are relatively easy to identify, which are featured by a quasar plus a close non-stellar companion. Unresolved lenses with small separation, instead, are more challenging.
To be more specific, unresolved lensed quasars can have a quasar-like color, a galaxy-like color, or something in between. Given the huge amount of galaxies and the small number of lensed quasars, it is practically impossible to find a simple color classification with both acceptable completeness and efficiency. Moreover, for unresolved systems, it is impossible to carry out lens modeling to reveal possible lensing structures. Some morphological indicators have been proposed to be useful to distinguish lensed quasars, including image fitting residuals (e.g., Oguri et al., 2006) and astrometric offsets between blue and red bands (e.g., Agnello & Spiniello, 2019). However, many other objects can also fulfill these criteria, like irregular galaxies, galaxy mergers, galaxies with strong star formation clumps, quasars blended with galactic stars, etc. Oguri et al. (2006) estimate that the selection completeness for lensed quasars with separation $\Delta \theta \lesssim 1''$ is lower than 50% for the SQLS. In the end, there is no practical way to identify unresolved lensed quasars efficiently based on their images and photometry.

It is thus critical to improve the angular resolution of observations. The recent commissioning of Gaia has revolutionized the searches for lensed quasars, aided by its all-sky coverage and superb angular resolution ($\sim 0.4''$ in DR2; e.g., Gaia Collaboration et al., 2018). Furthermore, Gaia provides the proper motion and parallax of objects, which are useful in excluding galactic stars in the candidate selection. Recent searches for lensed quasars thus heavily rely on the Gaia survey (e.g., Lemon et al., 2017; Spiniello et al., 2018; Lemon et al., 2020; Stern et al., 2021).

Another way to effectively reach a higher angular resolution is to improve the source detection and deblending algorithm of imaging surveys. Specifically, Oguri & Marshall (2010) propose the criteria of a discoverable lensed quasar, which should have a lensing separation larger than two thirds of the PSF FWHM. This definition corresponds to an optimal case where marginally resolved systems can be identified. In reality, most source detection and fitting algorithms (e.g., SExtractor; Bertin &
Arnouts, 1996) can deblend two point sources separated by $\gtrsim 1.5 \times \text{PSF FWHM}$². To maximize the output of current and upcoming imaging surveys, especially the Vera C. Rubin Observatory Legacy Survey for Space and Time (LSST; Ivezić et al., 2019), it is necessary to optimize the algorithms of source deblending.

To summarize, imaging and photometric analysis is an efficient way to identify large-separation lenses that are resolved in imaging surveys. Some unresolved ones can be selected by features like image fitting residuals and astrometric offsets between blue and red bands, yet the selection efficiency is usually low.

(2) Other techniques to select lensed quasars

- **Variability analysis.** Quasars are intrinsically variable objects, while normal galaxies do not show variabilities at observable timescales. It is thus possible to find lensed quasars by identifying extended sources that show variability, which could be unresolved lensed quasars. For example, Chao et al. (2021) explore the transients in the Hyper Supreme-Cam Survey (HSC; Aihara et al., 2018b) and find eight lens candidates, including one previously known lensed quasar. Bag et al. (2022) propose that seeking for possible time lags in light curves is a practical way of identifying unresolved lensed quasars. The limit of variability analysis is also obvious. The typical amplitude of quasar variability is $\sim 0.2 \text{ mag}$ (MacLeod et al., 2010), and variability analysis requires long-time monitoring with high S/N photometry. As such, this method is only useful for bright lensed quasars with significant variability.

In addition to lensed quasars, variability analysis can also select some other useful objects, including transients (e.g., supernovae) and low-luminosity AGNs. Practically, lensed quasars will be identified together with these objects in variability analysis.

- **Radio surveys.** Some efforts of finding gravitational lenses have been carried out in radio wavelengths. Since quasars are featured by radio emission, these surveys will naturally yield lensed quasars. For example, the Cosmic Lens All-Sky Survey

²The limit of $\sim 1.5 \times \text{PSF FWHM}$ is a rough estimate based on the performance of current imaging surveys (like DES and PS1; section 4.2). The exact deblending limit relies on the flux profile of the sources and the signal-to-noise ratio.
(CLASS; Myers et al., 2003) used the Very Large Array (VLA) to map the radio sky with a spatial resolution of \( \sim 0.2' \), leading to the discovery of \( \sim 20 \) lensed quasars. The radio searches only work for quasars with bright radio flux; CLASS only targeted objects brighter than 30 mJy at 4.85 GHz.

- Spectroscopic analysis. Some lensed quasars are initially identified as a normal galaxy, and follow-up spectroscopy serendipitously find features of a background quasar. A good example is SDSS J094604.90+183541.8 (McGreer et al., 2010), which was targeted by SDSS spectroscopy as a luminous red galaxy (LRG). Its spectrum exhibits quasar emission lines at \( z = 4.8 \), and follow-up observations confirmed this object as a lensed quasar.

### 4.1.2 What is a Good Survey Strategy for High-redshift Lensed Quasars?

The survey methods discussed above are for the general lensed quasar population. This project focuses on lensed quasars at high redshifts \( (z \gtrsim 5) \), which have some distinct features compared to low-redshift ones. In this subsection, I briefly discuss how the features of high-redshift quasars impact the strategy of our survey.

The most prominent feature of high-redshift quasars is the so-called “Ly\( \alpha \) break”. Due to the absorption of neutral hydrogen in the IGM, quasars at \( z > 5 \) have little or no flux at rest-frame wavelengths \( \lambda_{\text{rest}} < 1216\text{ Å} \). This effect generates a sharp flux break at the Ly\( \alpha \) wavelength, which is redshifted to \( \lambda_{\text{obs}} \gtrsim 7300\text{ Å} \) at \( z \gtrsim 5 \). Accordingly, previous surveys for quasars at \( z \gtrsim 5 \) select candidates that are not detected in blue bands (e.g., the g or r band) and show a sharp increase of flux at red bands (e.g., i or z band).

The distinct features of high-redshift quasars lead to the following considerations:

- High-redshift quasars appear to be fainter than low-redshift ones, and only have prominent flux in bands redder than the Ly\( \alpha \) break. I thus expect that a large fraction of high-redshift lensed quasars have fluxes dominated by the foreground deflector. Only selecting objects with quasar-like color will result
in low completeness. Near-infrared photometry offers SED points redder than the Ly$\alpha$ break and is critical in color classification. Variability analysis is less useful, both due to the reduced S/N and the time dilation effect at high redshifts.

- *Gaia* is useful only for the brightest high-redshift quasars. *Gaia* is designed for optical photometry and spectroscopy of galactic stars, with a wavelength coverage of $3500\,\text{Å} - 10000\,\text{Å}$. High-redshift quasars have little flux at $\lambda_{\text{obs}} \lesssim 7300\,\text{Å}$ and are faint due to their large luminosity distances. As a result, only the brightest high-redshift quasars can be detected by *Gaia*. Nonetheless, *Gaia* is still useful in identifying and rejecting galactic stars.

- High-redshift quasars are not detected in $g$ or $r$-bands, while the foreground lensing galaxy can be detected in these wavelengths. For lensed quasars at $z \gtrsim 5$, $g$-band imaging can be used to measure the properties of the deflector galaxy (positions, shapes, magnitudes, etc.). Specifically, if a high-redshift quasar shows some flux in the IGM absorption trough, the quasar is likely lensed by a foreground galaxy.

Based on these considerations, I adopt a combination of imaging + photometry selection and several supplementary candidate selections. The imaging + photometry selection uses optical and NIR sky surveys, where optimal image deblending technique should be performed to resolve objects with small angular separations. The supplementary candidate selections use various indicators of lensing, which enhance the completeness for unresolved lensed quasars. I do not rely on the *Gaia* survey in the candidate selection. Instead, *Gaia* and other available external datasets are used to prioritize the selected candidates.

### 4.1.3 Sky Coverage and Depth of This Survey

This survey focuses on the DESI Legacy Imaging Survey (Dey et al., 2019, hereafter the Legacy Survey) and the Dark Energy Survey (DES; DES Collaboration et al.,
footprint. I choose the combination of the Legacy Survey and the DES because these two surveys give more discoverable lensed quasars at \( z > 5 \) than other present-day imaging surveys, as discussed in Chapter 2. The Legacy Survey and the DES combined covers an area of \( \sim 20,000 \text{ deg}^2 \) in the high-galactic-latitude region \( (|b| \gtrsim 20^\circ) \), delivering a PSF size of \( \sim 1\text{"}1 \) and a depth of \( \approx 23 \) mag in \( z \)-band. In comparison, the Pan-STARRS1 survey (PS1; Chambers et al., 2016) is shallower, with a substantial fraction of its footprint falling in the low galactic latitude regions. Deep surveys like the HSC have footprints much smaller than the Legacy+DES combination. Although these deep surveys can detect fainter objects, most of these faint objects are not detected in present-day infrared sky surveys, which is a major drawback when selecting high-redshift quasar using their colors. Another crucial advantage of the Legacy Survey is its excellent performance in source deblending; I will discuss this point in Section 4.2.1.

Near-IR and mid-IR imaging surveys play critical roles in the color selection of high-redshift quasars. Previous high-redshift quasar searches have shown that the \textit{WISE} photometry is especially powerful in rejecting stellar contaminants, in particular cool dwarfs that also have red optical colors. In this survey, I only consider objects that are detected in \( W1 \)-band (\( \sim 3.4 \mu \text{m} \)) at \( > 5\sigma \) levels. The deepest \textit{WISE} all-sky data release available by the time of writing this thesis is the CatWISE2020 catalog (Marocco et al., 2021), which has a depth of \( W1_{\text{lim}} \approx 18 \) (Vega). Note that quasars at \( z \gtrsim 5 \) have \( m_z - W1 \sim 3 \) (e.g., Wang et al., 2016). As such, the primary targets of this survey are lensed quasars at \( z \gtrsim 5 \) that have at least one lensed image brighter than \( m_z = 21 \). I will show in Section 4.5 that \( m_z = 21 \) is also the limit where the lensing structure of a lensed quasar is still detectable in the Legacy Survey.

### 4.2 Data Sets

This section describes the data sets that I use in this project. In the following text, I use AB magnitudes for optical surveys and Vega magnitudes for IR surveys.
4.2.1 The DESI Legacy Imaging Surveys (the Legacy Survey)

The Legacy Survey includes three imaging sky surveys that are used for the target selection of the DESI spectroscopic survey (DESI Collaboration et al., 2016):

- the DECaLS survey (Dey et al., 2019), which scans $\sim 9,000$ deg$^2$ at $\delta \lesssim 32^\circ$ in the $g, r$ and $z-$bands using the Dark Energy Camera (DECam) on the Blanco 4m telescope;

- the BASS survey (Zou et al., 2017), which covers $\sim 5,500$ deg$^2$ at $\delta \gtrsim 32^\circ$ in the $g$ and $r-$bands using the 90Prime camera on the Bok 2.3-m telescope at Kitt Peak;

- the MzLS survey (Silva et al., 2016), which covers $\sim 5,500$ deg$^2$ at $\delta \gtrsim 32^\circ$ in the $z-$band using the MOSAIC-3 camera at the prime focus of the 4-meter Mayall telescope at Kitt Peak.

In addition, the Legacy Survey analyzes the $g, r$ and $z-$band observations of the DES, which also uses the DECam on the Blanco 4m telescope and covers $\sim 5,000$ deg$^2$ in the southern sky. The combination of these datasets provides a sky coverage of $\sim 20,000$ deg$^2$, reaching $5\sigma$ depths of $m_g = 24.7$, $m_r = 23.9$ and $m_z = 23.0$ for point sources. The $z-$band images have the sharpest PSF among all the three bands, which has an FWHM of $\sim 1''$.1.

The Legacy Survey uses Tractor (Lang et al., 2016) to measure the properties of objects. Tractor is a python-based source fitting and photometry tool. Briefly speaking, Tractor fits the images of an object as a point source and a Sérsic profile, then picks up the best-fit model based on the reduced $\chi^2$. For each object, the Legacy Survey performs joint fitting of all the survey images that contain the object, providing consistent photometry in different bands. Tractor also generates residual images for the best-fit models. The source detection and fitting pipeline can be executed iteratively, i.e., new sources can be detected in the residual image of the previous iteration. In this way, the Legacy Survey can detect faint objects close to a bright object.
One crucial strength of Tractor is its deblending ability. With sufficient S/N, the Legacy Survey can resolve heavily-blended objects with separations of \( \sim 0'7 \). For example, Figure 4.1 shows two sub-arcsecond separation lensed quasars that are resolved in the Legacy Survey. This deblending power of Tractor makes it possible to resolve the most compact discoverable lensed quasars in the Legacy Survey (which have lensing separations \( \Delta \theta \approx \frac{2}{3} \times \text{FWHM} \), Chapter 2).

4.2.2 Other Imaging Surveys

The Legacy Survey delivers imaging and photometry in the \( g, r, z, W1 \) and \( W2 \) bands. I use other public imaging surveys to gather the SED of objects from optical to mid-infrared, which are briefly described below.

1. The Dark Energy Survey (DES).

DES is a wide-area imaging survey that covers \( \sim 5,000 \text{ deg}^2 \) in the southern galactic cap. DES uses DECam on the Blanco 4m telescope to image the sky in \( g, r, i, z \) and \( Y \) bands. The 5\( \sigma \) magnitude limits in these bands are 24.7, 24.4, 23.8, 23.1, 21.7 for point sources. Note that the DES \( g, r \) and \( z \) band images are also processed by the Legacy Survey and are included in data releases after DR8. This project uses DES DR2 catalogs and images to gather the \( i \) and \( Y \)-band photometry of objects.

2. The Pan-STARRS1 Survey.

The Pan-STARRS1 Survey is a wide-area imaging survey that covers sky regions with \( \delta > -30^\circ \). The PS1 survey scans the sky in \( g, r, i, z \) and \( Y \) bands using a 1.8 meter telescope located at Haleakala, Hawaii. The 5\( \sigma \) magnitude limits in these bands are 23.3, 23.2, 23.1, 22.3, 21.3 for point sources. The typical PSF size of PS1 is \( \sim 1''07 \) (\( z \)-band). For objects that are not covered by the DES, I use PS1 DR2 catalogs and images to obtain the \( i \) and \( Y \)-band photometry of objects.

3. VISTA infrared imaging surveys.

Several near-IR imaging surveys have been carried out with the VIRCAM on the 4.1-m Visible and Infrared Survey Telescope for Astronomy (VISTA). These surveys cover different regions in the southern sky with various depth. In this project, I use
Figure 4.1: Two lensed quasars with sub-arcsecond separation that are resolved in the Legacy Survey Tractor catalog. These two examples illustrate the deblending power of Tractor. Note that the PSF size of the legacy survey is \( \sim 1'' \). Top: HE0512–3329 (Gregg et al., 2000), which has a redshift of 1.57 and a lensing separation of 0''65. HE0512–3329 is resolved into two point sources in the Tractor catalog, which are marked by the black crosses. The high-resolution HST image shows that the Tractor model correctly describe this lensing system. Bottom: SDSSJ1128+2402 (Inada et al., 2014), which has a redshift of 1.61 and a separation of 0''78. Similar to the top panel, SDSSJ1128+2402 is resolved into two point sources in the Tractor catalog, and the model is consistent with the Gaia catalog.
the VISTA Hemisphere Survey (VHS; McMahon et al., 2013) and the VISTA Kilo-Degree Infrared Galaxy Survey (VIKINGS; Edge et al., 2013) which cover the whole southern hemisphere. VIKINGS provides imaging in $z, Y, J, H$ and $K_s$ bands in 1,500 deg$^2$ of high-galactic-latitude area, reaching depths of 23.1, 22.3, 22.1, 21.5, 21.2, respectively. VHS images the rest of the southern hemisphere in $Y, J, H$ and $K_s$ bands with a depth that is $\sim$ 1 magnitude shallower than VIKINGS. By the time of writing this thesis, VHS has finished the imaging in the $J$ and $K_s$ bands, while the $H$-band observations are still on-going. The typical PSF sizes of VIKINGS and VHS are $\sim 1''$ in $J$-band.

4. The UKIRT infrared imaging surveys.

The northern hemisphere counterpart of the VISTA infrared imaging surveys is the UKIRT infrared imaging surveys. These surveys are carried out using the Wide Field Camera on the United Kingdom Infrared Telescope (UKIRT), including the UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al., 2007) and the UKIRT Hemisphere Survey (UHS; Dye et al., 2018) . UKIDSS covers 7,000 deg$^2$ in the northern hemisphere with the $Z, Y, J, H$ and $K$-bands, reaching depths of 20.3, 19.5, 18.6, and 18.2, respectively. UHS will deliver the $J$ and $K$-band photometry for the area at $0^\circ < \delta < 60^\circ$ that is not observed by UKIDSS, aiming at reaching the same depth of UKIDSS. By the time of writing this thesis, the UHS has finished the $J$-band observations, and the $K$-band data will be released in the future. The typical PSF sizes of UKIDSS and UHS are $\sim 1''$ in $J$-band.

5. The WISE survey.

The Wide-Field Infrared Survey Explorer (WISE; Wright et al., 2010) is a space-based telescope that provides mid-infrared imaging over the whole sky. In this project, we use the CatWISE2020 catalog (Marocco et al., 2021), which co-adds all the observations from 2008 to 2020 to improve the survey depth. The depths are $W1 = 17.7$ and $W2 = 17.3$ (Vega magnitudes). Note that WISE has a PSF size of $\sim 5''$, which means that galaxy-scale lensing systems will be unresolved in WISE. Although the Legacy Survey also offers $W1$ and $W2$ magnitudes of objects, CatWISE2020 has a better depth and is thus used in this project.

The Gaia survey uses a space telescope to scan the entire sky in optical wavelengths. The primary goal of Gaia is to map the galactic stars in the six-dimensional space (position + velocity), probing the structure and evolution of the Milky Way. Gaia delivers imaging in three bands: the $G$—band that covers 3500Å to 10000Å, the $G_{BP}$—band that covers 3500Å to 6500Å, and the $G_{RP}$—band that covers 6500Å to 10000Å. The detection limit is $G \sim 21$ in Gaia DR2. Gaia also provides integrated-field unit (IFU) spectroscopy at 8450Å–8720Å for objects with $G < 17$. Gaia measures the parallax and proper motion of each object, which are useful in removing galactic contaminants in this project. The spatial resolution of Gaia is $\sim 0.4$ for data release 2 (DR2; Gaia Collaboration et al., 2018).

4.2.3 Other Data Sets

In addition to imaging surveys, I use the following data sets to obtain additional information about the candidates:

1. X-ray and radio source catalogs.

Since quasars produce emission in X-ray and radio wavelengths, I include X-ray and radio surveys and archival catalogs. These data sets include the Chandra Source Catalog (Evans et al., 2010), the XMM-Newton Source Catalog (Zolotukhin et al., 2017), the ROSAT All-Sky Survey (Voges et al., 1999), and the VLA Sky Survey (VLASS; Condon et al., 1998). These data sets are used to prioritize the lensed quasar candidates (Section 4.6.2).

2. The DESI spectroscopic survey.

The Dark Energy Spectroscopy Instrument (DESI; DESI Collaboration et al., 2016) is a spectroscopic survey using the 4-meter Mayall telescope at Kitt Peak. DESI will deliver spectra of galaxies in the Legacy Survey footprint down to $m_z = 21$. These spectra will be used to identify unresolved lensing systems, which have spectra that show features of both the deflector and the background source.

\(^3\text{http://xmm-catalog.irap.omp.eu/}\)
4.3 Observational Features of Gravitationally Lensed Quasars

Chapter 2 suggests that there are about 10 lensed quasars at \( z > 5 \) that are discoverable in current imaging surveys. In contrast, so far only one has been reported (J0439+1634 at \( z = 6.52 \)). One reason for this discrepancy is that previous high-redshift quasar surveys have low completeness for lensed objects. Traditionally, searches for quasars rely heavily on identifying objects with quasar-like colors in public imaging surveys. For lensed quasars, however, the foreground lensing galaxy contributes to the fluxes and the colors of the whole system, which are usually unresolved or marginally resolved in ground-based imaging. As a result, traditional high-redshift quasar surveys have low completeness for lensed ones. To illustrate this, I test the performance of the candidate selection method in Jiang et al. (2016) on the simulated lensed quasars described in Chapter 2. Figure 4.2 shows the result, which clearly suggests that the foreground lensing galaxy can significantly reduce the completeness of traditional color selection. A new survey strategy is required for a more complete survey of lensed quasars. Specifically, different methods might be needed depending on the observed features of lensed quasars.

4.3.1 Classification of Lensed Quasars Based on the Observed Features

Due to the differences in lensing configurations, foreground galaxy SEDs and background quasar SEDs, lensed quasars exhibit a large diversity in their observed features. To design a complete survey, I classify lensed quasars into several groups based on how they look in ground-based images. To start with, I divide lensed quasars into three types according to their shapes in an imaging survey:

1. *Type I*: the lensed quasar has a small separation and appears to be a point source.

2. *Type II*: the lensed quasar has an intermediate separation and appears to be an extended source, but is not resolved into multiple components.

3. *Type III*: the lensed quasar is resolved into multiple components.
Figure 4.2: The impact of foreground lensing galaxies on high-redshift quasar color selections. We use simulated lensed quasars described in Chapter 2 and test them against the color selection criteria for \( z \sim 6 \) quasars in Jiang et al. (2016). The gray histogram shows the case where we only count the flux from the quasar, and the dashed histogram takes the lensing galaxy into consideration. Although the completeness of the color selection is high for quasar-only colors, it drops significantly with the contamination of the lensing galaxy.
I further define subtypes for each class that reflect the overall color of the lensing system. These subtypes are marked by suffixes $a$, $b$ and $c$, where the quasar-to-deflector flux ratio increase from $a$ (quasar-dominated) to $c$ (deflector-dominated). Figure 4.3 summarizes the definition of the subtypes and how we can identify them in imaging surveys. Note that the classifications are discussed qualitatively in this Section; the quantitative criteria of candidate selection will be discussed in Section 4.4. Figure 4.4 shows some examples of the subtypes.

It is useful to keep in mind a few considerations about lensed quasars. First, the lensing separation increases with the mass and thus the flux of the deflector, meaning that a bright deflector galaxy often leads to a large lensing separation. Second, in most lensing systems where the deflector is close to an singular isothermal ellipsoid (SIE; Kormann et al., 1994), the brighter lensed image is located farther away from
Figure 4.4: Examples of lensed quasars for each subtype. 

**Top:** J0439+1634 at $z = 6.52$ (Fan et al., 2019b), which is a typical subtype Ia lensed quasar. J0439+1634 appears to be a point source in the PS1 survey, and HST images suggest a small lensing separation of 0".2. There are no examples of subtype Ib and Ic in the known lensed quasar sample. 

**Middle:** examples of type II lensed quasars. From left to right: HE0512-3329 ($z = 1.57$; Gregg et al., 2000), ULAS0743+2457 ($z = 2.17$; Inada et al., 2014), SDSSJ1452+4224 ($z = 4.82$; More et al., 2016). The images are Legacy grz images with sizes of 6" × 6".

**Bottom:** examples of type III lensed quasars. From left to right: HE1104–1805 ($z = 2.32$; Wisotzki et al., 1993); J0941+0518 ($z = 1.54$; Lemon et al., 2018); J0124-0033 ($z = 2.84$; Lemon et al., 2019). The image for HE1104–1805 is the PS1 giy image, and the rest two images are the Legacy grz images. These three images have sizes of 9" × 9".
the deflector. Although this might not always be the case in the real universe, I notice that most known lensed quasars do follow this behavior.

Now I can discuss the observed features of each subtype in imaging surveys and how to identify these objects. Contaminants that might be confused with lensed quasars are discussed in Section 4.6.2. I describe an object as “point-like” or “extended” in an imaging survey according to its morphological type assigned by the survey. For example, a point-like object in the Legacy Survey has type=PSF in the Tractor catalog. I describe a lensing system as “resolved” if the lensing system is resolved into multiple components in the imaging survey. Furthermore, I always assume that quasars are point sources, i.e., the flux from the quasar host galaxy is negligible.

- **Subtype Ia.**

  Subtype Ia includes the most compact lensing systems that are dominated by quasar flux. These objects appear to be point sources and have a quasar-like color. In other words, these objects can be selected by traditional quasar survey methods. A representative example is J0439+1634, which has a lensing separation of $\sim 0.002$ and is discovered in a survey for normal quasars at $z \gtrsim 6.5$ (e.g., Wang et al., 2019b).

  Although these lensed quasars look similar to unlensed ones, careful analysis can find hints of strong lensing. First, the foreground lensing galaxy might still be detected in the survey images, which can be identified via the residuals of PSF subtraction. In particular, high-redshift quasars have no flux in blue bands, where the foreground deflector can be detected. Second, lensing magnification will change the apparent luminosity of the quasar without changing the spectral shape. Some quantities of quasars are related both to the spectral shape (e.g., broad emission lines) and the luminosity, which can be used to identify quasars with high lensing magnification (Section 4.4.3).

- **Subtype Ib.**

  Subtype Ib lenses appear to be point sources, where the quasar flux and the deflector flux are comparable. There is only one reasonable case for this subtype, i.e., the deflector galaxy is compact and appears to be a point source, and the lensing
separation is so small that the lensing system is not extended in the survey image. This subtype is expected to be rare, as a small lensing separation usually means a faint deflector galaxy and thus a large quasar-to-deflector flux ratio. There are no known lensed quasars that belong to this subtype.

Given their mixed color, subtype Ib lensed quasars cannot be selected in traditional lensed quasar surveys. Instead, we can find subtype Ib lenses by selecting point sources that exhibit a mixed (quasar+galaxy) color and have large PSF-subtraction residuals. The other practical ways to identify these objects is to search for objects that show both features of a low-redshift galaxy and a high-redshift quasar in large-area spectroscopic surveys like DESI.

– Subtype Ic.

Subtype Ic lenses appear to be a point sources and are dominated by deflector flux. These objects correspond to a very rare case, where the lensing separation is small, the deflector galaxy is compact (and is thus likely faint given the size-luminosity relation of galaxies), while the flux of the system is still dominated by the deflector galaxy. I consider subtype Ic lenses highly unlikely and do not consider this subtype further in this project. Specifically, since the primary targets of this project are lensed quasars brighter than $m_z = 21$ (Section 4.1.3), a deflector-flux-dominated lensed quasar must have a deflector galaxy with $m_z \lesssim 20$ to become a target of this survey. It is implausible that such a bright galaxy appear to be a point source in imaging surveys like the Legacy Survey.

– Subtype IIa.

Subtype IIa lenses are extended objects that are dominated by quasar flux. The lensing separation is sufficient to make these objects deviate from point sources, yet these objects are not resolved into multiple components. For example, a quasar-flux-dominated, doubly-lensed quasar that has a separation close to the PSF FWHM will appear as a subtype IIa lens. An example for this subtype is HE0512-3329 ($z = 1.57$, Gregg et al., 2000), which is unresolved in most ground-based sky surveys (excluding the Legacy Survey, where the deblending ability of Tractor makes it resolved; see Figure 4.1).
Unlike other extended objects in sky surveys (i.e., galaxies), subtype IIa lensed quasars cannot be well-fitted by regular Sérsic profiles. We can use this feature to distinguish these objects from normal galaxies. In addition, subtype IIa lenses have quasar-like color, allowing us to identify them using traditional color selection methods in quasar surveys.

**Subtype IIb.**

Subtype IIb lensed quasars are extended but unresolved objects, where the quasar and the deflector flux are comparable. These objects share similar features with subtype IIa, except that subtype IIb objects do not pass traditional color selection methods for quasars. An example is ULASJ0743+2457 ($z = 2.17$, Inada et al., 2014).

To identify subtype IIb lenses, we can select extended objects that exhibit a mixed (quasar+galaxy) color and cannot be well-fitted by a Sérsic profile. These objects can also be recognized via positional offsets between blue and red bands, given that the background quasar and the foreground galaxy have different colors (e.g., Agnello & Spiniello, 2019). For instance, the southeast and the northwest components of ULASJ0743+2457 show different colors in Figure 4.4.

**Subtype IIc.**

Subtype IIc lensed quasars are extended but unresolved objects, where the deflector flux dominates. In other words, these lensing systems have relatively small separations and faint background quasars, such that the lensed images of the quasar are overwhelmed by the deflector emission. A representative example is SDSSJ1452+4224 ($z = 4.8$; More et al., 2016), which is well-fitted by a Sérsic profile in the Legacy Survey.

The best way to find subtype IIc lenses is via spectroscopic surveys like SDSS and DESI. In some cases where the lensed quasar images are bright enough, we can identify them via significant residual of image fitting. However, large image fitting residuals can also be results of complex structures of galaxies (e.g., star formation regions with bright emission lines, minor mergers, bars and spiral arms, etc.), which can introduce substantial contamination in the candidate selection. Note that color
selection does not work for this subtype given the overwhelmingly large number of normal galaxies.

- **Subtype IIIa.**

Subtype IIIa lensed quasars are resolved into multiple components and are dominated by quasar flux. In other words, subtype IIIa objects appear to be several point sources, all of which have quasar-like colors. Many of the lensed quasars discovered in early ages belong to this subtype, for example, HE1104–1805 \((z = 2.32; \text{Wisotzki et al., 1993})\)

To identify subtype IIIa lenses, we can use traditional quasar survey methods to select quasar candidates, then search for close pairs among the quasar candidates. This process naturally yields physical pairs of quasars, which are further discussed in Chapter 5.

- **Subtype IIIb.**

Subtype IIIb lensed quasars are resolved into multiple components, where the quasar and the deflector flux are comparable. In other words, subtype IIIb objects appear to be (at least) a point source and an extended source, where the point source is a lensed image of the background quasar. Many of the known lenses fall in this subtype, and a representative example is J0941+0518 \((z = 1.54; \text{Lemon et al., 2018})\). To identify subtype IIIb lenses, we can select point sources with quasar-like colors that have an extended source next to them.

- **Subtype IIIc.**

Subtype IIIc lensed quasars are resolved into multiple components and are dominated by deflector flux. These objects share similar features with subtype IIIb, except for that the deflector galaxy in subtype IIIc lenses is brighter. An example is J0124–0033 \((z = 2.84; \text{Lemon et al., 2019})\). In principle, subtype IIIc lensed quasars can be identified in the same way as subtype IIIb ones, i.e., searching for point sources with quasar-like colors that have an extended source next to it. However, this task might be practically challenging, as the background quasar is likely overwhelmed by the flux from the foreground galaxy.
4.3.2 Discussions

The classifications described above are based on fluxes and shapes of lensed quasars. In principle, this classifications can be applied to quasars at all redshifts. Note that the type of a lensed quasar depends on the spatial resolution and the wavelength of observations. Since this project focuses on high-redshift lensed quasars that have little flux in blue bands, the subtypes mentioned in the following sections all correspond to the Legacy Survey $z-$band.

In addition, I briefly mention some limitations of the classification introduced above. The discussion in this Section only covers galaxy-scale lensing systems and does not include quasars that are lensed by galaxy groups and clusters. Quasars lensed by groups and clusters are much rarer than galaxy-scale lensed quasars and are not considered in this survey. These objects have large lensing separations ($\gtrsim 10''$), and the lensed images of the background quasar can be identified in traditional quasar surveys. In addition, it is possible that the quasar host galaxy gets strongly magnified and makes the lensed quasar image deviate from a point source. An example is RXJ1131-1231 ($z = 0.658$, Sluse et al., 2003). This case is also rare especially for high-redshift lensed quasars, which have smaller angular sizes for their host galaxies.

4.4 Candidate Selection

In this Section, I describe the candidate selection pipeline of our survey. Following the discussion in Section 4.3, I consider two samples of candidates, i.e., the companion-selected sample and the residual-selected sample. Generally speaking, candidates should either be resolved and have at least one quasar-like component (subtype IIIa, IIIb, IIIc), or be unresolved, have a mixed color between normal galaxies and quasars, and show large image fitting residuals when modeled by regular flux profiles (subtype Ia, IIa, IIb). I run morphological selection (Section 4.4.1) and color selection (Section 4.4.2) to identify these candidates. Finally, I build several supplementary candidate samples to cover lensed quasars that are missed by
the main selection pipeline (i.e., subtype Ia, Ib, IIb, IIc). The candidate selection procedure is summarized as a flowchart in Figure 4.5.

4.4.1 Morphology and Photometry Pre-Selection

The first step of the candidate selection pipeline is to identify objects that have shapes indicating possible lensing structures. The exact methods for companion-selected candidates and residual-selected candidates are different, which are described separately below.

(1) Residual-selected Candidates

This sample contains unresolved lensed quasars (type I and II), which cannot be well-fitted by regular flux profiles (i.e., a PSF or a Sérsic profile). To start with, I select objects in the Legacy Survey DR9 Tractor catalog using the following criteria:

- Morphological selection criteria for residual-selected candidates:

1. $z$–band magnitude $\text{mag}_z < 21$ and photometric error $\sigma_{\text{mag}_z} < 0.1$,
2. $\text{MASKBITS}=0$, i.e., no defects in survey images like bad pixels, saturation, etc.
3. Has a $z$–band image fitting reduced $\chi^2$ ($\text{rchisq}_z$) that is larger than 97.5% of the objects with similar shapes and magnitudes.

The third criterion above aims at selecting objects that cannot be well-fitted by the image model in the Legacy Survey. The exact way to apply this criterion is complicated, as the reduced $\chi^2$ of image fitting is influenced by many factors, including how bright the object is and the exact flux profile used for fitting. In this project, I set a threshold of $\text{rchisq}_z$ that depends on the shape and the magnitude of objects. I first classify all objects in the Legacy Survey as point sources (type=PSF) and extended objects (type!=PSF). I then bin each object class based on the $z$–band magnitude using a bin width of 0.1 magnitude, and select the top 2.5% objects in each bin in terms of $\text{rchisq}_z$. Here I use $z$–band as it has a smaller

\footnote{Context written in the typewriter font are columns in the Legacy Survey catalog.}
Figure 4.5: The flowchart of the candidate selection pipeline.
PSF size compared to $g$– and $r$–band in the Legacy Survey, and the fraction of 2.5% corresponds to $2\sigma$ outliers in Normal Distributions.

I then crossmatch the Legacy Survey catalog with other optical and NIR imaging surveys to obtain the $i, y(Y), J, H, K(K_s), W1$ and $W2$–magnitudes for the objects. The exact survey to crossmatch depends on the position of the candidate, and the detailed information is summarized in Table 4.1. When crossmatching the survey catalogs, I always use a matching radius of $2''$. I further select objects with photometric errors smaller than 0.2 mag in all bands except for $g, r$ and $W2$–bands. I do not set requirements in the errors of the $g$ and $r$–band magnitude because high-redshift quasars have little flux in these two bands, and I require that the photometric error in $W2$ should be smaller than 0.3 mag as WISE $W2$ photometry is shallower than $W1$. The selected objects are called residual-selected candidates and will enter the color classification described in Section 4.4.2.
Table 4.1. Optical and near-IR photometry used for objects in different sky area

<table>
<thead>
<tr>
<th>Sky Region Name</th>
<th>Legacy Survey</th>
<th>Optical Survey</th>
<th>Near-IR Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECaLS+DES+VIKING&lt;sup&gt;3&lt;/sup&gt;</td>
<td>DECaLS (&lt;i,g,r,z&lt;/i&gt;)</td>
<td>DES (&lt;i,Y&lt;/i&gt;)</td>
<td>VIKINGS (&lt;i,J,H,K&lt;/i&gt;)</td>
</tr>
<tr>
<td>DECaLS+DES+VHS&lt;sup&gt;4&lt;/sup&gt;</td>
<td>DECaLS (&lt;i,g,r,z&lt;/i&gt;)</td>
<td>DES (&lt;i,Y&lt;/i&gt;)</td>
<td>VHS (&lt;i,J,K&lt;/i&gt;)</td>
</tr>
<tr>
<td>DECaLS+PS1+VIKING&lt;sup&gt;5&lt;/sup&gt;</td>
<td>DECaLS (&lt;i,g,r,z&lt;/i&gt;)</td>
<td>PS1 (&lt;i,i,y&lt;/i&gt;)</td>
<td>VIKINGS (&lt;i,J,K&lt;/i&gt;)</td>
</tr>
<tr>
<td>DECaLS+PS1+VHS&lt;sup&gt;6&lt;/sup&gt;</td>
<td>DECaLS (&lt;i,g,r,z&lt;/i&gt;)</td>
<td>PS1 (&lt;i,i,y&lt;/i&gt;)</td>
<td>VHS (&lt;i,J,K&lt;/i&gt;)</td>
</tr>
<tr>
<td>DECaLS+PS1+UKIDSS&lt;sup&gt;7&lt;/sup&gt;</td>
<td>DECaLS (&lt;i,g,r,z&lt;/i&gt;)</td>
<td>PS1 (&lt;i,i,y&lt;/i&gt;)</td>
<td>UKIDSS (&lt;i,J,H,K&lt;/i&gt;)</td>
</tr>
<tr>
<td>DECaLS+PS1+UHS&lt;sup&gt;8&lt;/sup&gt;</td>
<td>DECaLS (&lt;i,g,r,z&lt;/i&gt;)</td>
<td>PS1 (&lt;i,i,y&lt;/i&gt;)</td>
<td>UHS (&lt;i&gt;J&lt;/i&gt;)</td>
</tr>
<tr>
<td>BASS+PS1+UKIDSS&lt;sup&gt;9&lt;/sup&gt;</td>
<td>BASS (&lt;i,g,r&lt;/i&gt;)+MaLS (&lt;i&gt;z&lt;/i&gt;)</td>
<td>PS1 (&lt;i,i,y&lt;/i&gt;)</td>
<td>UKIDSS (&lt;i&gt;J,H,K&lt;/i&gt;)</td>
</tr>
<tr>
<td>BASS+PS1+UHS&lt;sup&gt;10&lt;/sup&gt;</td>
<td>BASS (&lt;i,g,r&lt;/i&gt;)+MaLS (&lt;i&gt;z&lt;/i&gt;)</td>
<td>PS1 (&lt;i,i,y&lt;/i&gt;)</td>
<td>UHS (&lt;i&gt;J&lt;/i&gt;)</td>
</tr>
</tbody>
</table>

<sup>1</sup>Sky regions are defined as the overlapping regions of the corresponding sky surveys.

<sup>2</sup>In case that one object falls in multiple sky regions, the one listed on the top in this Table is adopted.

<sup>3</sup>Roughly covers RA from 22<sup>h</sup> to 3.5<sup>h</sup> and Dec from −35° to −26°.

<sup>4</sup>Roughly covers the DES footprint that is not included in DECaLS+DES+VIKING.

<sup>5</sup>Roughly covers RA from 10<sup>h</sup> to 16<sup>h</sup> and Dec from −5° to 5°.

<sup>6</sup>Covers all Legacy Survey footprint at −30° < Dec < 0°.

<sup>7</sup>Roughly covers the UKIDSS footprint at 0° < Dec < 32°.

<sup>8</sup>Roughly covers the Legacy Survey footprint at 0° < Dec < 32° that is not included in DECaLS+PS1+UKIDSS.

<sup>9</sup>Roughly covers the UKIDSS footprint at 32° < Dec < 60°.

<sup>10</sup>Roughly covers the Legacy Survey footprint at 32° < Dec < 60° that is not included in BASS+PS1+UKIDSS.

Note. — *WISE* photometry is available for all the sky regions. VHS and VIKINGS use the same filter set, though VIKINGS is deeper than VHS. The H-band observations of VHS has not been completed. Similarly, UKIDSS and UHS use the same filter set. UHS has only finished the J-band observations, reaching the same depth as UKIDSS.
(2) Companion-selected Candidates

This sample corresponds to type III lensed quasars. These lensing systems are resolved into multiple components, where at least one component should be the lensed image of the background quasar. I thus start from finding point sources with quasar-like color that have a close companion. Specifically, I select objects from the Legacy Survey DR9 Tractor catalog using the following criteria:

- **Morphological selection criteria for companion-selected candidates:**

  1. **type=PSF**, i.e., the object is a point source,
  2. $z$-band magnitude $\text{mag}_{z} < 21$ and photometric error $\sigma_{\text{mag}_{z}} < 0.1$,
  3. **MASKBITS=0**, i.e., no deficits in survey images like bad pixels, saturation, etc.,
  4. Has a companion within $3''$ (the companion does not need to satisfy the above criteria.)

Here, the selected object corresponds to one of the lensed images of the background quasar, and the companion corresponds to the lensing galaxy and/or the other lensed images of the quasar.

I then gather the $i, g(Y), J, H, K(K_s), W1$ and $W2$—magnitudes for the objects from other optical and NIR imaging surveys. This task is a bit trickier than the residual-selected candidates, given the deblending ability of Tractor. Specifically, two objects with a separation of $\lesssim 1''$ can get resolved in the Legacy Survey catalog. However, such objects might be unresolved in other imaging surveys and appear to be a single object. To obtain the magnitudes of a resolved candidate in all bands consistently, I run forced photometry on images from DES, PS1, VIKING, VHS, UKIDSS, and UHS for the resolved candidates selected above. Again, the exact surveys used for an object depend on its position on the sky, as summarized in Table 4.1. I do not run forced photometry for WISE images because the PSF size of WISE is too large ($\gtrsim 4''$). Correspondingly, WISE magnitudes are not used in the color classification of companion-selected candidates.
I use galfit (Peng et al., 2002) to perform forced photometry. For a resolved candidate, I fit the images from other sky surveys with all source parameters fixed to the values in the Legacy Survey catalog, except for the fluxes of the objects. I use bright stars as PSF models. To select PSF stars, I first identify point sources (type=PSF) in the Legacy Survey Catalog that have photometric errors less than 0.05 mags in all bands, then select objects that satisfy the following: (1) have no companions within 4", (2) have no bad flags (MASKBITS=0), and (3) have a rchisq.z value that is smaller than 16% of the objects in the corresponding magnitude bin (see the description of Criteria 4.1.3 above for more details; this fraction corresponds to 1σ boundary in Normal Distributions). For each object, I identify the closest PSF star and use its images as the PSF model for the object. I use the background area around the object to estimate the noise of pixels and assume that the noise is background-limited.

Galfit returns the best-fit magnitudes and errors for each object. I then select objects with photometric errors smaller than 0.2 mag in all bands except for g and r−bands. Furthermore, I require that the candidate should be detected in W1−band at a 5σ level, and in W2−band in a 3σ level. Note that I am not able to get the W1 and W2 magnitudes for individual resolved components given their small separation. However, requiring detections in WISE bands reduces the contamination from galactic stars, which usually have a bluer z − W1 colors compared to those of high-redshift quasars. Objects selected above are named companion-selected candidates and will be processed by the color classification (Section 4.4.2).

Lastly, it is worth noticing that some companion-selected candidates may have incorrect image models, especially objects that are only marginally resolved. Incorrect image models may result in unreliable photometry. However, this issue has little impact on the whole candidate selection, since candidates that have incorrect image models will fall into the residual-selected candidate sample as well. As I will show in Section 4.4.2, residual-selected candidates will pass the color selection criteria if the candidate has a quasar-like color or a mixed color between galaxies and quasars. Even if a marginally-resolved lensed quasar has an incorrect image model
in the Legacy Survey, the measured colors of the object will still lie between quasars and galaxies. The relatively loose color selection criteria ensure that lensed quasars with incorrect image models can still be selected.

4.4.2 Color Classification

The second step is to identify objects that have SEDs consistent with lensed quasars. For companion-selected candidates (type III), I require that the point source component in the candidate should have a quasar-like color. For residual-selected candidates, the color should be either quasar-like (subtype Ia, IIa) or mixed between high-redshift quasars and low-redshift galaxies (subtype Ib, IIb). In the latter case, I further require a significant astrometric offset between blue and red bands to minimize the contamination from normal galaxies.

I use Probabilistic Random Forests (PRFs; e.g., Reis et al., 2019) to estimate whether an object has a “quasar-like color” or not. PRFs are variations of Random Forests, a class of supervised machine learning algorithm that is widely used in classification and regression. Unlike regular RFs where the features of data points have no errors, PRFs take the error of features into account by considering each data point as a probabilistic distribution. PRFs are thus the ideal choice for classifying and fitting noisy data.

The first step of using PRFs is to construct training sets. I aim at classifying all the objects into three classes: stars, galaxies and (high-redshift) quasars. I select/simulate objects for training sets as follows:

- **Stars.** I identify isolated point sources with small image fitting residuals in the Legacy Survey as the training set for stars. I require that these objects have photometric errors in all bands smaller than 0.1 mag and have \( \text{rchi} \text{sq.Z} \) smaller than 16% of the objects in the corresponding magnitude bin, following the method in Section 4.4.1. The requirement in \( \text{rchi} \text{sq.Z} \) ensures that the object is correctly classified as a point source. I further require that the object has no companions within 4'' to avoid flux contamination. The sample of
stars selected this way will contain some real quasars and compact galaxies. Nonetheless, these point-like extragalactic objects have negligible impact on the PRF training, as galactic stars outnumbers non-stellar point sources by orders of magnitudes.

- **Galaxies.** Similar to stars, I identify isolated extended sources with small image fitting residuals in the Legacy Survey. Specifically, I select objects that have photometric errors smaller than 0.1 mag in all bands, have $\text{rchisq}_{z}$ smaller than 16% of the objects in the corresponding magnitude bin, and have no companions within $4''$.

- **High-Redshift Quasars.** I use SIMQSO (McGreer et al., 2013) to generate mock quasars at $4.5 < z < 7$, which covers the redshift range of interest in this project. More details about SIMQSO can be found in Chapter 2 and Chapter 3. To homogeneously sample quasars at all redshifts, I adopt a “flat” QLF normalization such that $dN_{\text{QSO}}/dz$ is a constant. All the other settings of SIMQSO are the same as described in Chapter 2. I use the simulated spectra to calculate the magnitudes, and assume a photometric error of 0.05 mag for all the mock quasars. This magnitude error is used arbitrarily to mimic photometry with decent S/N.

When training a PRF, I randomly select 100,000 objects from each sample described above. Since the exact magnitudes used for a candidate depend on its position (see Table 4.1), it is necessary to train multiple PRFs to cover all the possible combination of filters. Also note that VIKINGS and VHS use the same instrument and filter set, while VIKINGS is deeper than VHS. I thus select objects in the VIKINGS survey to build star and galaxy training sets for both the VIKING and VHS footprint. Similarly, UKIDSS and UHS have the same depth in $J$–band, and I use objects from UKIDSS to build star and galaxy training sets for both the UKIDSS and UHS footprint. Table 4.2 summarizes the training set used to build each PRF, and the PRF I use for color classification for different sky regions.
<table>
<thead>
<tr>
<th>PRF Name</th>
<th>Training Set Sky Region</th>
<th>Candidate Sky Region</th>
<th>Magnitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECaLS,DES,VIKINGS,(WISE)</td>
<td>DECaLS+DES+VIKINGS</td>
<td>DECaLS+DES+VIKINGS</td>
<td>$g,r,i,z,Y,J,H,K_s,(W1,W2)$</td>
</tr>
<tr>
<td>DECaLS,DES,VHS,(WISE)</td>
<td>DECaLS+DES+VIKINGS</td>
<td>DECaLS+DES+VHS</td>
<td>$g,r,i,z,Y,J,H,K_s,(W1,W2)$</td>
</tr>
<tr>
<td>DECaLS,PS1,VIKINGS,(WISE)</td>
<td>DECaLS+PS1+VIKINGS</td>
<td>DECaLS+PS1+VIKINGS</td>
<td>$g,r,i,z,Y,J,H,K_s,(W1,W2)$</td>
</tr>
<tr>
<td>DECaLS,PS1,VHS,(WISE)</td>
<td>DECaLS+PS1+VIKINGS</td>
<td>DECaLS+PS1+VHS</td>
<td>$g,r,i,z,Y,J,K_s,(W1,W2)$</td>
</tr>
<tr>
<td>DECaLS,PS1,UKIDSS,(WISE)</td>
<td>DECaLS+PS1+UKIDSS</td>
<td>DECaLS+PS1+UKIDSS</td>
<td>$g,r,i,z,Y,J,H,K_s,(W1,W2)$</td>
</tr>
<tr>
<td>DECaLS,PS1,UHS,(WISE)</td>
<td>DECaLS+PS1+UKIDSS</td>
<td>DECaLS+PS1+UHS</td>
<td>$g,r,i,z,Y,J,(W1,W2)$</td>
</tr>
<tr>
<td>BASS,PS1,UKIDSS,(WISE)</td>
<td>BASS+PS1+UKIDSS</td>
<td>BASS+PS1+UKIDSS</td>
<td>$g,r,i,z,Y,J,H,K_s,(W1,W2)$</td>
</tr>
<tr>
<td>BASS,PS1,UHS,(WISE)</td>
<td>BASS+PS1+UHS</td>
<td>BASS+PS1+UHS</td>
<td>$g,r,i,z,Y,J,(W1,W2)$</td>
</tr>
</tbody>
</table>

1The sky region from which the training set objects are gathered.

2The sky region where the PRF is used to perform color classification.

3The magnitudes used to generate PRF features. Note that VIKINGS and VHS use the same filter set, while VIKINGS is deeper than VHS; UKIDSS and UHS use the same filter set and have similar depths in $J$–band.

Note. — The sky regions are defined in Table 4.1. Due to the large PSF size of WISE, WISE photometry are used only for unresolved candidates. In other words, I use PRFs with WISE photometry to classify residual-selected candidates, and use PRFs without WISE photometry to classify companion-selected candidates. See Section 4.4.1 for details.
It is worth mentioning that, although the training sets of stars and galaxies directly come from the datasets I use for candidate selection, there are no overlaps between the candidates and the training sets. Specifically, the candidates have either companions or large image fitting residuals, while the training set objects have no companion and have small image fitting residuals. Unlike some other quasar surveys that use colors as input features of RFs, I use flux ratios between other bands and $z-$band as features. This choice makes it easier to deal with large flux errors (which lead to non-Gaussian errors in colors) and non-detections.

Using the trained PRFs, I predict the probabilities for the residual- and companion-selected candidates to be a star, a galaxy, or a quasar. I then select high-redshift lensed quasar candidates based on these probabilities. For companion-selected candidates (corresponding to subtype IIIa, IIIb, IIIc), the color selection criteria are the following:

- **Color selection criteria for companion-selected candidates:**

1. Has a quasar-like ($p_{QSO} > 0.8$), non-stellar color ($p_{\text{star}} < 0.01$),

2. Has a companion with a non-stellar color ($p_{\text{star}} < 0.01$) or being extended (type! = PSF).

The above criteria select Type III lensed quasars. The criteria above aim at selecting quasars with a non-stellar projected companion, which could be either a foreground galaxy (the deflector) or the other lensed image of the quasar.

For residual-selected candidates, the criteria are:

- **Color selection criteria for residual-selected candidates:**

1. Has a non-stellar color ($p_{\text{star}} < 0.01$),

2. Satisfies one of the following:

   - Has a quasar-like color ($p_{QSO} > 0.8$);
Has a mixed color between high-redshift quasars and galaxies ($p_{\text{QSO}} > 0.3$), and has a significant ($> 3\sigma$) astrometric offset between $g-$ and $z-$band.

Here, I identify residual-selected candidates that have quasar-like colors (subtype Ia, IIa) or mixed colors between a quasar and a galaxy (subtype Ib, IIb). The latter case suffers heavy contamination from normal galaxies. To minimize the contamination, I add the astrometry offset criterion, i.e., that there should be a significant astrometric offset between blue and red bands. Since high-redshift quasars have little or no flux in $g-$band, the $g-$band center reflects the position of the lensing galaxy, while the $z-$band flux center has a significant contribution from the background quasar. As a result, there should be an offset between $g-$ and $z-$band flux centers. This method has been proven to be useful in searches for low-redshift lensed quasars (Agnello & Spiniello, 2019). In this survey, I use the $3'' \times 3''$ image cutouts from the Legacy Survey to calculate the flux centers (and their errors) of objects. The noise map is estimated following the method for forced photometry, described in Section 4.4.1.

Candidates selected via color classification are called the main sample candidates, in opposite to the supplementary candidate samples described in the next subsection.

### 4.4.3 Supplementary Candidates

The morphological + color candidate selection above can select lensed quasars of subtype Ia, Ib, IIa, IIb, IIIa, IIIb, and IIIc. However, this selection method will miss the majority of subtype Ic, IIc lenses, and some subtype Ia, Ib, and IIb lenses. I thus include two supplementary candidate samples to cover these subtypes.

1. **Quasars with features suggesting lensing magnification**

Lensed quasars have their observed flux magnified, while their spectral shapes remain unchanged. I can thus search for potential indicators of lensing magnification from quantities that are related to both the spectral shape and the luminosity of quasars. These indicators include Eddington ratios and the ages of quasars, which allows me to identify lensed quasars with high magnification. In particular, some
subtype Ia lensed quasars are too compact to be distinguishable from normal quasars in imaging surveys. These objects can be selected via their apparent Eddington ratios and small proximity zone sizes.

- **High Eddington ratio quasars.** The Eddington ratio reflects the strength of the quasar emission relative to the Eddington accretion limit. The Eddington ratio is defined as $R_{\text{Edd}} = \frac{L_{\text{bol}}}{L_{\text{Edd}}}$, where $L_{\text{bol}}$ is the bolometric luminosity of the quasar, and $L_{\text{Edd}} = 1.26 \times 10^{38} \frac{M_{\text{BH}}}{M_\odot}$ erg s$^{-1}$ is the Eddington luminosity of the SMBH. Practically, the SMBH mass of a quasar is estimated by applying the Virial theorem on the broad line region. Observations suggest an empirical relation of $\log M_{\text{BH}} = a \log L + b \log \text{FWHM} + c$, where $L$ is the luminosity of the quasar, FWHM is the width of the broad emission line, and $a \approx 0.5$, $b \approx 2$, and $c$ are constant coefficients (e.g., Vestergaard & Peterson, 2006). The exact values of $a$, $b$ and $c$ depend on the luminosity and the broad emission line used in the empirical relation. Combining the equations above gives $R_{\text{Edd}} \propto L^{1-a} \text{FWHM}^{-b}$. Strong lensing will change the observed luminosity but not the line FWHM, and a lensing magnification of $\mu$ will lead to an overestimation of $R_{\text{Edd}}$ by a factor of $\mu^{1-a}$. This effect is shown in Figure 4.6, with examples of known lensed quasars at $z > 5$. In particular, quasars with $R_{\text{Edd}} > 1$ (i.e., super-Eddington accretion) are more likely magnified. In this survey, I select quasars at $z > 5$ from the literature that have high Eddington ratios ($R_{\text{Edd}} > 3$) as lensed quasar candidates.

- **Very young quasars.** The age of a quasar, defined as the time since the quasar starts to get unobscured and ionize the surrounding IGM, can be estimated via the size of its proximity zone and its luminosity. Specifically, Eilers et al. (2017) suggest $\tau \propto r^3/L$, where $\tau$ is the age of the quasar, $r$ is the size of the proximity zone and $L$ is the quasar luminosity. The proximity zone size is measured via the Ly$\alpha$ emission line profile of the quasar, which is not influenced by strong lensing. Lensed quasars thus appear to be younger than the general quasar population. This effect is confirmed by Davies et al. (2020), who analyze the spectra of J0439+1634 and show that this lensed quasar has an apparent proximity zone size that is $\sim 2 - 3$ times smaller than the general quasar population. In this survey, I select young
Figure 4.6: The $M_{\text{BH}}$ vs luminosity distribution of quasars. Adapted from Fan et al. (2019b). Lensing magnification will lead to overestimated Eddington ratios. The open symbols illustrate two examples of lensed quasars at $z > 5$, J0439+1634 and J0025–0145 (also see Chapter 6). The filled purple asterisk shows the intrinsic $M_{\text{BH}}$ and luminosity of J0439+1634 after correcting the lensing effect. Similar analysis is still pending for J0025–0145. I include $z \sim 6$ quasars from the literature (black dots) and SDSS quasars (gray scatters and contours) for comparison.

quasars compiled in Eilers et al. (2020) as lens candidates. Note that the proximity zone analysis only works for quasars at $z \gtrsim 5.5$, where the cosmic UV background is low and the neutral fraction of the IGM is sufficiently high (see Eilers et al. (2017) for details).

2. DESI-observed galaxies

It could be extremely challenging to distinguish deflector-dominated lensing systems from normal galaxies. One possible way of finding these lenses is to use large spectroscopy surveys like SDSS and DESI. These surveys target a wide range of extragalactic objects, from which I can search for candidates that show features of both a galaxy and a quasar in their spectra. A good example is SDSSJ1452+4224 ($z = 4.8$; More et al., 2016), which is deflector-dominated and appears to be a regular Sérsic profile in Legacy Survey images. This lensed quasar is first identified via the broad emission line in its SDSS spectrum, then confirmed as a lensed quasar by follow-up high-resolution imaging and spectroscopy.
DESI has four extragalactic target samples: bright galaxies (BGS), emission-line galaxies (ELG), luminous red galaxies (LRG), and quasars (QSO). Specifically, the BGS sample will provide a flux-limited sample down to $m_r = 20$, and the LRG sample will reach a depth of $m_z = 21$. Note that LRGs contribute to about 80% of the lensing optical depth (e.g., Oguri & Marshall, 2010). Since this project focuses on lensed quasars with $m_z < 21$, the DESI target samples will cover nearly all the deflector-dominated lensed quasars. Checking DESI spectra that are classified as extragalactic objects (i.e., not a star) will identify candidates of lensed quasars, which show features of both a low-redshift galaxy and a high-redshift quasar.

By the time of writing this thesis, the DESI-based candidate selection is only designed conceptually, and no real candidates have been selected yet. One reason is that the DESI survey is in its early stage, with a small fraction of its targets observed. In the future, I will develop utilities for spectra fitting, which fits the DESI spectra by a low-redshift galaxy plus a high-redshift quasar template. If the composite (quasar + galaxy) spectra model describes the object better than the DESI model, the object will be selected as a lens candidate.

### 4.4.4 Generalizing the Survey to Low-redshift Lensed Quasars

The major improvement of this project compared to previous lensed quasar searches is that I minimize the dependence on the Gaia survey and make full use of the de-blending ability of the Legacy Survey. The candidate selection pipeline is designed to select high-redshift lensed quasars that are faint and red in optical imaging surveys, most of which have a small quasar-to-deflector flux ratio and are not detected in Gaia. Nevertheless, I notice that the candidate selection pipeline can be straightforwardly generalized to low-redshift lenses. Specifically, I use mock quasars at $z > 4.5$ as training sets for the PRF in the current version. By including low-redshift quasars in the training set, the candidate selection pipeline will work for lensed quasars at all redshifts.

In the future, LSST will provide optical images that are several magnitudes deeper than current ground-based imaging surveys and the Gaia survey. The meth-
ods developed in this project can be used to search for faint lensed quasars in the LSST survey that are not detected in Gaia and other space-based surveys.

4.5 Testing the Candidate Selection Method

A key task in quasar surveys is to estimate the completeness of candidate selection. Since the candidate selection in this project relies on the shapes and colors of objects, the survey completeness depends on lensing configuration, deflector morphology, and the SEDs of the deflector and the quasar. At this point, I perform sanity checks of the candidate selection pipeline using known lensed quasars and un-lensed high-redshift quasars. In the future, new data releases of public imaging surveys (e.g., the Legacy Survey and the Gaia survey) and the updates of my survey strategy (Section 4.7) will improve the performance of the candidate selection pipeline. A full analysis of the completeness function will be carried out at the end this survey.

I first test the morphological selection described in Section 4.4.1 using known lensed quasars. There are 220 known lensed quasars by the time of writing this thesis, 177 of which are detected by the Legacy Survey. I remove the requirements in magnitudes and photometric errors, allowing to investigate the magnitude limit of the morphological selection. The morphological selection successfully recovers 162 out of 177 lensed quasars. This project focuses on lensed quasars with at least one component brighter than $m_z = 21$ and lensing separation smaller than $\Delta \theta = 3\arcsec$. For the subsample in this magnitude and separation range, the selection completeness is $146/152 \approx 96\%$.

To further investigate the factors that influence the morphological selection completeness, I plot the distribution of $z-$band magnitude versus lensing separation for lensed quasars in Figure 4.7. At $m_z \lesssim 20$, almost all known lensed quasars are resolved in the Legacy survey. The smallest lensing separation that is resolvable is $\sim 0\!.07$, confirming the argument in Section 4.1.3. The completeness of the morphological selection is close to one for known lensed quasars with $m_z \lesssim 20$. The completeness drops quickly at $m_z > 20$ and drops to $\lesssim 20\%$ at $m_z > 21$. This result
Figure 4.7: The distribution of $z$-band magnitude versus lensing separation for lensed quasars that are detected in the Legacy Survey. The blue symbols are lensed quasars that have more than one matches in the Legacy Survey using a radius of 3" (i.e., resolved); the orange symbols are lensed quasars that have only one match in the Legacy Survey (i.e., unresolved). The open circles and the crosses mark lensed quasars that are selected and missed by the morphological selection criteria, respectively. The dashed line shows the designed depth of this survey, $m_z < 21$. The morphological selection has a high completeness ($\approx 96\%$) for the known lensed quasars that fall in the targeted magnitude and separation range (see text for details).

explains why I set $m_z = 21$ to be the targeted survey depth.

The lensed quasars missed by the morphological selection are all objects with fluxes dominated by the deflector (i.e., subype IIc and IIIc). Many of these lensed quasars appear to be a regular elliptical galaxy in Legacy Survey images and require high-resolution imaging to reveal the lensed images of the background quasar. The DESI-based supplementary sample will cover these lensed quasars, further improving the survey completeness of this project.

I then test the PRF color classification. Previous quasar surveys have suggested a completeness of $\geq 90\%$ for RF-based classifiers (e.g., Schindler et al., 2017). Here I use a test PRF and known quasars at $z > 4.5$ from the literature to evaluate the performance of PRFs. I use objects in the PS1 + UKIDSS footprint and simulated
quasars as training sets and use PS1 $grizy$ + UKIDSS $J + WISE W1, W2$ to generate features for the test PRF, following the method in Section 4.4.2. I then obtain the photometry of known quasars at $z > 4.5$ that fall in the UHS footprint and at $\delta > 30^\circ$, and select the subsample that satisfies the magnitude error requirements described in Section 4.4.1\textsuperscript{5}.

Figure 4.8 illustrates the performance of the PRF on the subsample of high-redshift quasars. It is clear that the predicted $p_{\text{QSO}}$ is high except for faint quasars at $z \sim 4.5$. One possible reason is that the training set for quasars does not contain simulated quasars at $z < 4.5$. Nonetheless, most $z > 5$ quasars have high predicted $p_{\text{QSO}}$. For quasars at $z > 4.8$, requiring $p_{\text{QSO}} > 0.8$ selects 98\% of the known quasars. As such, I use $p_{\text{QSO}} > 0.8$ as the criteria to determine if an object has a quasar-like color in Section 4.4.2.

To summarize, the candidate selection pipeline successfully recover the majority of known lensed quasars (for morphological selection) and un-lensed high-redshift quasars (for color selection). In the future, I will use simulated lensed quasars (e.g., the mock catalog presented in Chapter 2) to estimate the survey completeness as a function of deflector properties and background quasar properties.

The completeness of the supplementary samples depends on the strategy of previous quasar surveys and the target selection of DESI. The complex nature of the supplementary sample makes it difficult to estimate its completeness function. Also note that the supplementary candidate sample contains lensed quasars that are not “discoverable” in the Legacy Survey as defined in Oguri & Marshall (2010) (also see Section 2). As such, I will not consider the supplementary sample when estimating the completeness of this survey for the discoverable lensed quasars.

\textsuperscript{5}This PRF and the test sample are constructed when I develop the color classification method. The PRFs in Table 4.2 will be tested in the future.
Figure 4.8: The performance of the test PRF on a sample of known quasars at $z > 4.5$ (see text for details). *Top:* the distribution of redshift versus $z$–band magnitude of the quasars. The color of the scatter points represent the predicted $p_{\text{QSO}}$. Most known quasars have $p_{\text{QSO}} \gtrsim 0.9$ except for faint ones at $z \sim 4.5$. *Bottom:* the distribution of $p_{\text{QSO}}$ for $z > 4.8$ quasars, which are the primary targets of this survey. Requiring $p_{\text{QSO}} > 0.8$ (Section 4.4.2) leads to a completeness of 98%. 


4.6 Current Survey Progress

So far I have described the candidate selection pipeline of this survey. The design of this pipeline was finished recently, and the candidate samples are not yet available. In the past years, I have developed several prototypes of the current candidate selection pipeline, using which I have identified and observed some candidate lensed quasars. In this Section, I describe the performance of the most recent prototype pipeline and the follow-up observations of the lens candidates.

4.6.1 Performance of the Prototype Candidate Selection Pipeline

The prototype candidate selection pipeline has a few essential differences from the current version. Briefly speaking,

1. The prototype pipeline takes Legacy Survey DR8 as the input catalog, whereas the current version uses DR9. Compared to DR8, the Legacy Survey DR9 has significant improvement in source deblending. See Chapter 7 for some examples. As a result, many objects that get resolved in the current pipeline are unresolved in the prototype.

2. The prototype pipeline collects the magnitudes of companion-selected and residual-selected candidates in the same way. All magnitudes are aperture magnitudes with a diameter of 5″, except for WISE bands that have PSF sizes of \( \sim 4″ \). The large aperture was chosen to include the flux from possible companions.

3. When selecting the residual-selected candidates, the prototype uses a parameterized description for the limit of \( r\text{chisq}_z \) as a function of object magnitude, whereas the current version uses a non-parameterized description.

4. In the color classification step, the prototype considers two types of objects, i.e., point sources and extended objects. For point sources, the prototype selects objects with \( p_{\text{QSO}} > 0.8 \) and \( p_{\text{star}} < 0.01 \), which corresponds to subtype-Ia lenses. For extended sources, the criteria are \( p_{\text{QSO}} > 0.3 \) and \( p_{\text{star}} < 0.01 \),
which selects type II and III lenses. The prototype pipeline does not require
astrometric offsets between blue and red images for candidates.

Here I use the PS1+UHS sky region at $\delta > 30^\circ$ to illustrate how this prototype
pipeline works. I first gather the photometry from the PS1 survey for objects at
$\delta > 30^\circ$, and match the PS1 catalog to the UHS, the \textit{WISE} survey, and the Legacy
Survey catalogs. I only select objects with $m_z < 20$, since the large aperture used
in photometry increases the magnitude errors. I then apply color selection for the
objects, which is illustrated in Figure 4.9. Note that the PRF used here is the same
as the one used to test the color selection completeness in Section 4.5. There are
6,338,034 extended objects in the field, out of which 12,483 pass the color selection
criteria. For point sources, these numbers are 6,702,097 and 563, respectively.

The objects that pass the color selection enter the morphological selection.
The companion-selected sample includes PS1 objects with multiple matches in the
Legacy Survey, where the matched Legacy Survey objects are separated by less than
3''. The residual-selected sample contains objects with $rchiq.z > 1.55 + 0.05 \times
|m_z - 20.5|^{4.5}$. This criterion in $rchiq.z$ is tuned to include $> 90\%$ of the known
lensed quasars. Among the 12,483 color-selected extended sources, 2,588 pass the
morphological selection. Similarly, 198 out of 563 point sources pass the morpho-
logical selection. Figure 4.10 shows the residual-based morphological selection for
the extended objects.

4.6.2 Visual Inspection and Follow-up Observations of Candidates

The candidate sample selected by the prototype pipeline is still dominated by con-
taminants. To enhance the efficiency of follow-up observations, I visually inspect all
the candidates and assign priorities to them. For companion-selected candidates,
I check the fluxes and positions of the resolved components and the image fitting
residual. In particular, I look for point sources in the survey images, which are likely
lensed images of the background quasar. If the positions and flux ratios of the point
sources are consistent with a typical lensing model, the candidate is assigned a high
Figure 4.9: Example of the PRF color selection in the prototype pipeline, for the sky area covered by PS1 and UHS at $\delta > 30^\circ$. The PRF used here is the same one as the PRF used to test the color selection completeness in Section 4.5. The histograms show the distribution of predicted probabilities of objects to be a star, galaxy, or a quasar. The black dashed lines marks the selection criteria for $p_{\text{QSO}}$. Top: extended sources; Bottom: point sources.
Figure 4.10: Example of the residual-based morphological selection in the prototype pipeline. The scatters are the extended sources that pass the color selection in Figure 4.9. The $x$– and $y$–axis are the Legacy Survey $z$–band magnitude and the $z$–band image fitting reduced $\chi^2$. The black dashed line marks the selection criteria, $rchi$ $z > 1.55 + 0.05 \times |mag$ $z - 20.5|^{4.5}$. This function is tuned so that most of the known lenses can get selected. Note that this Figure contains scatters at $mag$ $z > 20$, which are faint companions in the Legacy Survey matched to the PS1 objects.
priority. For residual-selected candidates, if the candidate exhibits the typical sign of a high-redshift quasar, i.e., a sharp increase of flux from blue to red wavelengths, it is assigned a high priority. Based on these considerations, I mark candidates as priority 1 (likely), 2 (uncertain), or 3 (unlikely). In addition, a candidate will be moved one priority higher if it is detected in X-ray or radio (see Section 4.1.3), and will be moved one priority lower if it has a non-zero parallax or proper motion (in 3σ level) in Gaia.

I then carry out follow-up spectroscopy and high-resolution imaging for the priority 1 candidates to verify whether the candidate is a lensed quasar. Specifically, I first take the spectrum of the candidate and seek for features of a high-redshift quasar. If the spectrum confirms the existence of a quasar, I will take images and/or spectra of the system with sufficient spatial resolution. In the case of quadruply-lensed quasars, I consider the candidate to be confirmed if the lensing structure is clearly resolved in the image. Confirming doubly-imaged lensing systems is less straightforward. For example, many candidates that are resolved into two point sources turn out to be projected pairs of quasars and stars. It is thus necessary to take the spectra of both components and show that the two spectra exhibit the same features (redshift, emission line profile, etc.). I further require that the deflector galaxy should be detected in the image, which ensures that the object is a doubly-imaged lensed quasar instead of a physical quasar pair.

So far, I have observed ~ 200 priority-1 candidates, among which I have confirmed one lensed quasar and two close quasar pairs at $z \gtrsim 5$, plus a handful at lower redshifts. The $z \gtrsim 5$ lensed quasars and quasar pairs will be described in detail in Chapter 5 and Chapter 7. Here I use a low-redshift lensed quasar discovered in this survey, J0918–0220 ($z = 0.8$), to illustrate the visual inspection and follow-up observation process.

For each pipeline-selected candidate, I generate an information plot for quick visual inspection. Figure 4.11 presents the information plot of J0918–0220. This plot includes the color-color plot, the SED and the survey images of the candidate, as well as its magnitudes, Gaia information (proper motion and parallax), X-ray and radio
detection, and the result of crossmatching to previous quasar catalogs. J0918–0220 is selected based on its mixed color between galaxies and quasars ($p_{\text{GALAXY}} = 0.43$ and $p_{\text{QSO}} = 0.57$). I assign a priority-1 to this object because the Legacy Survey images show two point sources with similar colors and a red galaxy in the middle. Also note that J0918–0220 is not detected in the Gaia survey DR2; this object is a good example of how the Legacy Survey unveils lensed quasars that are not detected in Gaia.

J0918–0220 was observed by Magellan/Clay LDSS-3 in April 2021. Figure 4.12 illustrates the spectra of J0918–0220. I put the slit along the two blue point sources in the Legacy Survey image as shown in Figure 4.11, and the 2D spectrum is clearly resolved into two traces. The extracted 1D spectrum suggests that both traces exhibit broad emission lines at $z = 0.8$, which confirms that J0918–0220 is a lensed quasar. Specifically, trace 1 is the fainter lensed image of the quasar blended with the deflector galaxy, and trace 2 is the brighter lensed image of the quasar.

Even after visual inspection, the high-priority candidate sample is still dominated by contaminants. Typical contaminants include emission line galaxies, projected pairs of galaxy+star and quasar+star (where the star has similar colors to high-redshift quasars). Specifically, the narrow emission lines in galaxies can mimic the broad emission line of quasars in broad-band photometry, and possible off-center star forming regions will also generate astrometric offsets between blue and red bands (Criteria 4.4). Late-type stars, especially M, L and T dwarfs, have been long recognized as the main contaminant in high-redshift quasar surveys (e.g., Yang et al., 2019a). These stars have red color in optical wavelengths and can be confused with high-redshift quasars when the photometry is noisy. M, L and T dwarfs are numerous in the Milky Way, and many candidates turn out to be a galaxy blended with one or two cool dwarfs.

J0918–0220 has a large lensing separation that can be well-resolved with seeing-limited observations. Given the distribution of lensing separations (e.g., Figure 3.3), about 50% of the high-redshift lensed quasars have lensing separations of $\lesssim 1''$. and require observation with adaptive optics (AO) or space telescopes like HST.
Figure 4.11: The information plot of J0918–0220. This information plot consists of the color-color plot (top left), the SED (middle left), the survey images (bottom left), and some information about the object (right). The information includes magnitudes, $Gaia$ proper motions and parallax, X-ray and radio detections, PRF-predicted probabilities, matches to previous quasar catalogs, as well as Legacy Survey shape and companion information. Such information plots are used for quick visual inspection.
Figure 4.12: The Magellan/Clay LDSS-3 spectra of J0918–0220, taken in April 2021 using the VPH-All grism and the 1" slit. Top: the 2D spectra. With a seeing of $\sim 0.7$, J0918–0220 is clearly resolved into two traces. The upper trace is the brighter quasar image, and the lower trace is the fainter quasar image blended with the deflector galaxy. The two traces exhibits some emission lines at the same wavelengths. Bottom: the extracted 1D spectra for the two traces. Both traces show broad emission lines at $z = 0.8$, confirming that J0918–0220 is a lensed quasar.
and *JWST* to clearly resolve the lensing structure. For candidates that cannot be well-resolved in seeing-limited observations, I will first carry out ground-based spectroscopy and identify objects that exhibit features of both high-redshift quasars and low-redshift galaxies. I will then take high-resolution images to resolve the lensing structures using either space telescopes or AO if possible. In addition, sub-mm interferometers like ALMA are useful in distinguishing close quasar pairs and doubly-imaged lensed quasars. Since the far-infrared emission of quasars is dominated by their host galaxies, lensed quasars can be unambiguously confirmed by detecting the lensed arcs of the quasar hosts.

### 4.6.3 Expected Performance of the Current Candidate Pipeline

The updates in the new candidate selection pipeline will improve the selection efficiencies in a number of ways. First, the improved source deblending in the Legacy Survey DR9 allows resolving objects with small separation ($\sim 0.007$), which are not resolved in Legacy Survey DR8. It is thus possible to reject galaxy+star or quasar+star pairs using forced-photometry and color classification of the resolved components. Second, the magnitudes used in the current pipeline are based on image fitting, which have smaller errors compared to the aperture magnitudes used in the prototype. Third, by introducing the astrometric offset criteria, the majority of normal galaxies will get excluded in the current pipeline, which are a major source of contaminants in the prototype pipeline. I thus expect that the current pipeline will reach a better depth and a higher efficiency compared to the prototype.

### 4.7 Future Work

#### 4.7.1 Improving the Candidate Selection Pipeline

This survey for high-redshift quasars is still ongoing. The main framework of the candidate selection pipeline has been recently finalized, with some expected updates in the future. In this Section, I describe some improvements of the candidate selection that I am working on or will work on soon.
(1) **Morphological Selection.** The most important part of the morphological selection is to get the correct image models for candidates, especially for blended ones. Specifically, this project uses the best-fit model in the Legacy Survey *Tractor* catalog. In the upcoming Legacy Survey DR10, there will be major updates in the source detection and photometry (e.g., Meisner et al., 2022). It is worth noticing that the Legacy Survey DR9 shows a significant improvement in source deblending compared to DR8 (see Chapter 7 for some examples). I thus expect that the Legacy Survey DR10 will result in a better performance of the candidate selection pipeline.

(2) **Color Classification.** The PRF classification can be improved by upgrading the training sets. One obvious update is to include low-redshift \( (z < 4.5) \) simulated quasars in the quasar training set, with which the candidate selection pipeline will also work for low-redshift lensed quasars. In addition, more features can be added to SIMQSO to make the simulated quasars more realistic, including dust attenuation and broad absorption line (BAL) features.

(3) **Supplementary Sample.** The main pending task about the supplementary sample is to finish the DESI-based candidate selection, as described in Section 4.4.3. I will also investigate the variability-based candidate selection, which will further improve the survey completeness. Although current sky surveys like PS1 do not have sufficient depth to probe the variability of faint high-redshift quasars, the upcoming LSST survey will offer sufficient depth and time resolution, opening a new window of searching for lensed quasars based on their variability.

(4) **Prioritizing the Candidates.** The current pipeline relies on visual inspection to identify high-priority candidates. In the future, I will develop several automatic tools to replace the visual inspection step, taking the images and the SEDs of the candidates as inputs. For resolved candidates (especially subtype IIIb and IIIc where the deflector galaxy flux is significant), a fast, automatic lens modeling tool can pick up candidates that can be well-described by lensing models (e.g., Chan et al., 2015). Recent studies also show that deep learning are useful in selecting quadruply-lensed quasars (e.g., Akhazhanov et al., 2022). For unresolved candidates, I will develop SED fitting tools to select objects that can be well-described by the combination
of a galaxy spectrum and a quasar spectrum. These automated tools will reduce the amount of contaminants and make the candidate selection less subjective. In addition, these tools will allow us to get ready for the upcoming large sky surveys like LSST, where visual inspection becomes impractical given the huge data volumes.

4.7.2 Follow-up Studies of Lensed Quasars

The lensed quasars will enable important studies about the high-redshift lensed quasar population. In Section 1.3.1, I mention that surveys for high-redshift lensed quasars are critical to the measurements of quasar luminosity functions (QLFs) and the understanding of SMBH-host coevolution in the early universe. Here I discuss more specifically how this survey will contribute to these two topics.

(1) Quasar luminosity functions. The lensed fraction of high-redshift quasars have been used to constrain the bright-end slope of QLFs (e.g., Richards et al., 2004, 2006b). In Chapter 3, I show that the lensed fraction of quasars at \( z \sim 6 \) is \( \sim 0.4\% - 0.8\% \) for current quasar surveys. This fraction is much smaller than the Poisson noises of the current high-redshift quasar samples (e.g., Wang et al., 2016; Matsuoka et al., 2018a; Yang et al., 2019a). As such, the lensed quasar sample itself will not put stronger constraints on the QLFs compared to surveys for unlensed quasars. Instead, by conducting a systematic survey for lensed quasars, I will confirm the predictions of the theoretical models, illustrating that current QLF measurements are accurate even without correcting the magnification bias (or vice versa).

(2) SMBH-host coevolution in the early universe. This survey will offer the first sample of lensed quasars at \( z \gtrsim 5 \). These lensed quasars provide magnified views of SMBHs and their host galaxies in the early universe. In particular, high-redshift quasar host galaxies exhibit complex structures at sub-kpc scales (e.g., Venemans et al., 2017; Yue et al., 2021b). Lensing magnification helps us to probe and understand the small-scale structures in the host galaxies lensed quasars.
Discovery of a Close Quasar Pair at $z = 5.66^\dagger$

5.1 Introduction

Galaxy mergers are natural consequences of hierarchical structure formation of the universe (e.g., Cole et al., 2000). It has been proposed that galaxy mergers can trigger powerful quasar activity which regulates the star formation of their host galaxies by significant feedback (e.g., Di Matteo et al., 2005). In some rare cases, the supermassive black holes (SMBHs) in both progenitor galaxies are ignited by the merging event, forming a close pair of quasars (for a recent review, see De Rosa et al., 2019). High-resolution hydrodynamical simulations suggest that close quasar pairs correspond to a special, short-lived phase of galaxy mergers, which most frequently appear when the SMBHs have a mass ratio close to one and a separation smaller than 10 kpc (e.g., Capelo et al., 2015, 2017). Quasar pairs are also predicted in cosmological simulations (e.g., Steinborn et al., 2016; Volonteri et al., 2016; Rosas-Guevara et al., 2019), and their statistics (e.g., the fraction of pairs among all quasars) constrain the evolution of SMBHs and their host galaxies. In addition, close quasar pairs trace overdensities (e.g., Onoue et al., 2018), and are unique probes of the small-scale structure of intergalactic medium (e.g., Rorai et al., 2017).

Searches of quasar pairs have been carried out using wide-area optical and infrared (IR) sky surveys (e.g., Hennawi et al., 2010; Silverman et al., 2020). These studies have discovered several tens of close quasar pairs (with projected separation $\Delta d \lesssim 10$ kpc) out to $z \sim 3$ (e.g., Shen et al., 2021; Tang et al., 2021). Due to the rapid decline of quasar number densities (e.g., Kulkarni et al., 2019) and the relatively poor physical resolution for most observations, it becomes difficult to identify

\textsuperscript{\dagger}This Chapter was published as Yue et al. (2021a)
close quasar pairs at $z \gtrsim 3$. Finding high-redshift quasar pairs is critical to the understanding of the evolution and environment of high-redshift quasars and galaxy mergers.

In this letter, we report a close quasar pair candidate at redshift $z = 5.66$ with a projected proper distance of $\Delta d = 7.3$ kpc, J2037–4537. Following Hennawi et al. (2010), in this Letter, we define a quasar pair as two quasars with $\Delta d < 1$ Mpc and a radial velocity difference of $\Delta v < 2000$ km s$^{-1}$. We further define close quasar pairs to be those with $\Delta d < 10$ kpc. The most distant quasar pair previously known is at $z = 5.1$ with $\Delta d = 135$ kpc (McGreer et al., 2016), and for close quasar pairs, the highest-redshift one is at $z = 3.1$ with $\Delta d = 7.4$ kpc (Tang et al., 2021). J2037–4537 significantly extends the frontier of studies of quasar pairs.

This letter is organized as follows. We describe the ground-based observations and the Hubble Space Telescope (HST) imaging in Section 5.2. We discuss the interpretation of the data in Section 5.3 and summarize in Section 5.4. We use a flat $\Lambda$CDM cosmology with $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$ and $\Omega_M = 0.3$.

5.2 Data and Analysis

5.2.1 Ground-Based Imaging and Spectroscopy

J2037–4537 was discovered in our on-going survey for high-redshift gravitationally lensed quasars (Yue et al. in prep) in the Dark Energy Survey (DES, e.g., DES Collaboration et al., 2021a) field. It is not detected in the DES $g$–band, and has the DES $r = 22.8$, $i = 20.4$, $z = 19.6$ and $Y = 19.5$, exhibiting typical colors of a quasar at redshift $z \gtrsim 5.5$ (e.g., Yang et al., 2019a). In the DES images (Figure 5.1, upper left), J2037–4537 appears to be two point sources separated by 1″24, where the fainter one (object B) is redder than the brighter one (object A). Table 5.2.1 lists the photometric properties of the two objects. The magnitudes are obtained by fitting the images as two point sources using galfit (Peng et al., 2002), where the point spread function (PSF) models are built by the IRAF task psf.

The optical spectrum of J2037–4537, taken by the Low Dispersion Survey
Figure 5.1: Ground-based imaging and spectroscopy of J2037–4537. **Upper Left:** the DES $izY$ images shown in RGB format. The PSF FWHM of the image is \( \sim 0''9 \). The fainter component (object B) is redder than the brighter one (object A). The two objects are separated by 1''24. The cyan lines mark the slit position in the spectroscopy. **Upper Right:** the 2-D spectrum taken by Magellan/Clay LDSS-3. The two objects are well-resolved under a seeing of \( \sim 0''7 \), allowing us to extract the two traces accurately. The white and yellow lines mark the apertures used to extract the 1-D spectrum. The Magellan/Baade FIRE spectrum has a better seeing of \( \sim 0''6 \). **Middle Panel:** The extracted 1-D spectrum of the two objects, which is a combination of the Magellan/Clay LDSS-3 spectra (\( \lambda < 1\mu m \)) and the Magellan/Baade FIRE spectra (\( \lambda > 1\mu m \)). Both objects exhibit features of a $z = 5.66$ quasar, and object B is redder than object A. The dashed line shows the F850LP filter used in HST ACS/WFC imaging. **Lower Panel:** The flux ratio between the two objects. The shaded area marks wavelengths bluer than the Ly$\alpha$ break, where the quasars have little flux due to the IGM absorption. There are complex features around the emission lines, which suggest that the two objects have different line profiles, and disfavor the strong lensing hypothesis.
Spectrograph (LDSS-3) on Magellan/Clay telescope, reveals that both objects are quasars at redshift $z = 5.66$ (Figure 5.1). We use the VPH-Red grism and the 1\textquoteleft 0 slit which deliver a resolution of $R = 1350$. The spectrum is reduced with the standard IRAF pipeline. The two objects are clearly resolved in the spectrum and have the same Lyα-break wavelength, suggesting a redshift difference $\Delta z \lesssim 0.01$. Object B has a redder spectral energy distribution (SED) shape than object A, which agrees with the broad-band photometries. We also obtained near-IR spectra of the two objects using the Folded-port InfraRed Echelle (FIRE) on Magellan/Baade telescope. The spectra are reduced with PypeIt (Prochaska et al., 2020).

J2037–4537 could either be a physical quasar pair, or two images of a strongly lensed quasar. Both the image and the spectra do not show signs of a third object (e.g., a foreground lensing galaxy). This is further confirmed by the HST image (Section 5.2.2). In addition, the two objects have different emission line profiles. The difference is the most obvious for the C\textsc{iii}$\lambda$ line and can also be seen in the A/B flux ratio (Figure 5.1, lower panel). These features disfavor the strong-lensing scenario and indicate that J2037–4537 is a physical quasar pair.

Figure 5.2 illustrates the C\textsc{iv} broad emission lines of the two objects, which are fitted as a power-law continuum plus a Gaussian profile. Using the C\textsc{iv} line width and the empirical relation in Vestergaard & Peterson (2006), we estimate the SMBH masses to be $\log M_{\text{BH}}/M_\odot = 8.60$ for quasar A and $\log M_{\text{BH}}/M_\odot = 8.45$ for quasar B. Although the Mg\textsc{ii} line is a better SMBH mass indicator for high-redshift quasars, it falls in the wavelengths with strong sky emission lines at $z = 5.66$. The results are listed in Table 5.2.1.

### 5.2.2 HST Imaging

Gravitational lensing can also generate two close quasar images at the same redshift. To fully explore the lensing hypothesis, we observed J2037–4537 using HST ACS/WFC in the F850LP filter, aiming at detecting or ruling out the foreground lensing galaxy (Program GO-16507). The F850LP filter has a wavelength coverage of $0.8\mu m < \lambda < 1.1\mu m$, where the two objects show significantly different SEDs in
Figure 5.2: The $\text{CIV}$ emission lines of the two quasars (black) and the best-fit model (red). We use a power-law continuum and a Gaussian profile to describe the emission lines. The absorption features (gray dashed area) are masked during the fitting.
Table 5.1. Properties of J2037–4537

<table>
<thead>
<tr>
<th>Object</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>20:37:40.29</td>
<td>-45:37:57.7</td>
</tr>
<tr>
<td>Dec</td>
<td>20:37:40.28</td>
<td>-45:37:56.5</td>
</tr>
</tbody>
</table>

Magnitudes\(^1\)

<table>
<thead>
<tr>
<th></th>
<th>DES g</th>
<th>DES r</th>
<th>DES i</th>
<th>DES z</th>
<th>DES Y</th>
<th>F850LP</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>&gt; 25.1</td>
<td>23.09</td>
<td>20.84</td>
<td>20.17</td>
<td>20.07</td>
<td>20.10</td>
</tr>
<tr>
<td>Dec</td>
<td>&gt; 25.1</td>
<td>24.8</td>
<td>21.80</td>
<td>20.67</td>
<td>20.39</td>
<td>20.67</td>
</tr>
</tbody>
</table>

Luminosities

<table>
<thead>
<tr>
<th></th>
<th>(M_{1450}^2)</th>
<th>(\log L_{\text{bol}}(\text{erg s}^{-1})^3)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>-26.42</td>
<td>47.1</td>
<td>-25.99</td>
</tr>
<tr>
<td>Dec</td>
<td>-25.99</td>
<td>46.9</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)All magnitudes are AB magnitudes. Detection limits are 5\(\sigma\).

\(^2\)The absolute magnitude at rest-frame 1450Å.

\(^3\)Calculated based on \(M_{1450}\) and the bolometric correction in Runnoe et al. (2012).

\(^4\)The SMBH mass based on the relation in Vestergaard & Peterson (2006). The intrinsic scatter of the relation is 0.36 dex which dominates the uncertainties.
Figure 5.3: The HST ACS/WFC F850LP image of J2037–4537 (left) and the residual of the two-PSF model (right). There is no sign of a third object besides the two point sources. We measure the flux in the area C (the red circle, with a radius of 0.35”) to estimate the flux limit of the possible foreground lensing galaxy. We also measure the fluxes in two additional areas, C1 and C2 (yellow circles), which correct the contributions from the PSF-subtraction residual (see text for details). Our estimation gives a 3σ flux limit of $m_{F850LP} > 26.7$ for region C.

The ground-based spectroscopy. Figure 5.3 presents the HST image and the two-PSF fitting residual. J2037–4537 is well-described by the two-PSF model and there is no evidence for a third object in this field.

To further set a flux limit for the possible lensing galaxy, we measure the flux in the region between the two point sources (region C, marked by the red circle in the right panel of Figure 5.3). Since PSF models of ACS/WFC usually have non-negligible errors in PSF wings (Jee et al., 2007), we measure two additional regions (C1 and C2, marked by the yellow circles) to correct this systematic uncertainty. The yellow regions have the same size as the red one, and the distance from region C1 (C2) to object A (B) equals to the distance from region C to object A (B). The flux in region C1 (C2) thus estimates the contribution of the PSF wing residuals from object A (B) to the flux in region C. We estimate the flux of the possible foreground galaxy as

$$F_{\text{foreground}} = F_C - F_{C1} - F_{C2}$$

(5.1)
This gives a non-detection and rules out the existence of a foreground lensing galaxy with $m_{F850LP} < 26.7$ at $3\sigma$ level. Assuming a typical redshift of $z_l \sim 1$ for the lensing galaxy (e.g., Hilbert et al., 2008; Wyithe et al., 2011), this limit gives an absolute magnitude of $M \gtrsim -17.6$, which is about four magnitudes fainter than the breaking magnitude of the galaxy luminosity function at $z \sim 1$ ($M_* \sim -21.5$; e.g., Faber et al., 2007).

5.2.3 Testing the Strong-Lensing Hypothesis

The distinct SEDs of the two quasar images and the non-detection of the lensing galaxy strongly disfavor the lensing hypothesis. Although microlensing and/or differential reddening can lead to different spectral features in lensed quasar images (e.g., Sluse et al., 2012), it is difficult for these effects to explain the observational features of J2037–4537. If object A and B are lensed images of the same quasar, the extinction of image B must be much stronger than A given their colors. At long wavelengths ($\lambda_{\text{obs}} \gtrsim 1.5\mu$m), image B is brighter than image A, meaning that image A is intrinsically fainter. Meanwhile, in most galaxy-scale lensing systems, the fainter image is closer to the deflector (e.g., Mason et al., 2015) and should have a stronger extinction. This conflicts with the fact that image B is redder than A, suggesting that the lensing scenario is unlikely.

In principle, a deflector galaxy with highly irregular mass profiles and/or dust distributions might be able to generate a lensing system like J2037–4537. We thus report J2037–4537 as a high-confidence close quasar pair candidate. The definitive test of the lensing scenario can be achieved with the Atacama Large Millimeter/submillimeter Array (ALMA). ALMA will measure the $[\text{C}\text{II}]$ redshifts of the two quasar images to an accuracy of $< 0.001$ (e.g., Decarli et al., 2018), and will reveal the lensed arcs of the quasar host galaxy. In the following discussions, however, we will treat J2037–4537 as a physical quasar pair.
5.3 Discussion

5.3.1 The Pair Fraction of High-Redshift Quasars

J2037−4537 has a projected separation of 1′′24 (7.3 kpc at $z = 5.66$). Kiloparsec-scale quasar pairs are rare especially at high redshift, due to the quick decline of the number density of high-redshift quasars. De Rosa et al. (2019) summarize a number of predictions of the active galactic nuclei (AGNs) pair fractions (e.g., Steinborn et al., 2016; Volonteri et al., 2016; Rosas-Guevara et al., 2019). These studies use large-volume cosmological simulations like Magneticum$^1$, EAGLES (Schaye et al., 2015), and Horizon-AGN (Dubois et al., 2014), which model the SMBH growth and AGN activities using analytical relations. These studies then count close SMBH pairs that have bolometric luminosities beyond a certain threshold. Specifically, De Rosa et al. (2019) consider AGNs with bolometric luminosity $L_{\text{bol}} > 10^{43}$ erg s$^{-1}$, and count AGN pairs with a separation of $r < 30$ proper kpc (pkpc). The predicted pair fraction, $f_{\text{pair}}(r < 30 \text{ pkpc})$, is about 1% with little redshift evolution.

J2037−4537 allows us to constrain the quasar pair fraction at high redshift for the first time. This is done by estimating the expected number of quasars we can find in our survey. J2037−4537 is discovered serendipitously in our survey for high-redshift lensed quasars in the DES field, which covers $\sim 5000 \text{ deg}^2$. We identify J2037−4537 as a lens candidate with quasar-dominated flux, for which we request a non-detection in DES−$g$, a DES−$z$ magnitude of $m_z < 21$, and a projected separation of $\Delta d < 3''0$. The $g$− and $z$−band flux cuts restrict our targets to $5 \lesssim z \lesssim 6$ quasars (e.g., Wang et al., 2016; Yang et al., 2017, 2019a). As the survey is not completed yet, we simply assume a selection function of one for $5 < z < 6$ quasars, and provide the lower limit of the quasar pair fraction.

We use SIMQSO (McGreer et al., 2013) to generate a mock catalog of quasars in the DES field at $5 < z < 6$ that are brighter than $m_z = 21$. SIMQSO is a Python-based package which generates mock catalogs of quasars with simulated spectra, and has been shown to accurately reproduce the observed colors and magnitudes of

$^1$http://www.magneticum.org/index.html
quasars (e.g., Ross et al., 2013). We use a broken power-law to parameterize the quasar luminosity function (QLF):

\[
\phi = \frac{\Phi^*}{10^{0.4(M-M^*)(\alpha+1)} + 10^{0.4(M-M^*)(\beta+1)}}
\]

(5.2)

We adopt the \(z=5\) QLF from Kim et al. (2020) and the \(z=6\) QLF from Matsuoka et al. (2018a), and linearly interpolate the QLF parameters (\(\log \Phi^*, M^*, \alpha, \beta\)) for other redshifts at \(5 < z < 6\). SIMQSO suggests that there are \(N_{\text{all}} = 485\) quasars brighter than \(m_z = 21\) at \(5 < z < 6\) in the DES field.

J2037–4537 directly constrains the number of quasar pairs at \(z > 5\) with \(0'9 < \Delta d < 3'0\). The lower boundary of the projected separation corresponds to the spatial resolution of the DES survey. To avoid confusion, we always use \(r\) to denote the physical separation of a quasar pair, and use \(\Delta d\) to denote the projected separation. By converting the angular separations to proper distances, we find \(f_{\text{pair}}[5.3 < \Delta d (\text{pkpc}) < 17.7] \geq 1/485 = 0.2\%\). As a comparison, the expected number of quasar pairs is \(7 \times 10^{-6}\) if quasars are randomly distributed. We estimate this number by assigning random positions to the quasars in the mock catalog, and counting pairs of mock quasars with \(0'9 < \Delta d < 3'0\) and \(\Delta \nu < 2000\) km s\(^{-1}\), following Hennawi et al. (2006). The clustering signal for \(z > 5\) quasars is thus \(W_p(0'9 < \Delta d < 3'0) \sim 10^5\). Hennawi et al. (2006) measure the quasar correlation function at \(1 \lesssim z \lesssim 3\) and \(\Delta d \sim 0.1\) pMpc. Extrapolating the correlation function in Hennawi et al. (2006) to small scales gives a clustering signal of \(W_p \sim 10^3\). This comparison indicates that close quasar pairs are regulated by processes (likely galaxy mergers) that are different from quasar clustering at larger scale.

In order to make direct comparisons to the simulated pair fractions, we convert \(f_{\text{pair}}(0'9 < \Delta d < 3'0)\) to \(f_{\text{pair}}(r < 30\) pkpc) as follows. We assume that the small-scale quasar correlation function can be described by a power-law, i.e., \(\xi(r) = (r/R_0)^{-\gamma}\). Following the analysis in Hennawi et al. (2006) and McGreer et al. (2016), the expected number of quasar pairs with separation \(r < R\) can be
calculated as

\[
N_{\text{pair}}(r < R) = \frac{1}{2} \sum_j N_{\text{all}} \int_0^R n(z_j) \times 4\pi r^2 [1 + \xi(r)] dr
\]  

(5.3)

where the sum goes through all the mock quasars and \(n(z_j)\) is the number density of quasars at redshift \(z_j\) that have \(m_z < 21\). We use the quasar spectrum template from Vanden Berk et al. (2001) to convert the absolute magnitude \(M_{1450}\) to the \(z\)-band apparent magnitude. Similarly, the expected number of quasar pairs with projected separation \(d_{\text{min}} < \Delta d < d_{\text{max}}\) is

\[
N_{\text{pair}}(d_{\text{min}} < \Delta d < d_{\text{max}}) = \frac{1}{2} \sum_j N_{\text{all}} \int_{\frac{v_{\text{max}}}{H(z_j)}}^{\frac{v_{\text{max}}}{aH(z_j)}} dy \int_{d_{\text{min}}}^{d_{\text{max}}} dx n(z_j) \times 2\pi x [1 + \xi(\sqrt{y^2 + x^2})]
\]  

(5.4)

where \(a = (1 + z)^{-1}\) is the scale factor, \(H(z_j)\) is the Hubble constant at redshift \(z_j\), and \(v_{\text{max}} = 2000\ \text{km s}^{-1}\) is the maximum velocity difference of the quasar pair. Note that all the quantities in Equation 5.3 and 5.4 are in comoving units. We adopt a fiducial power-law index of \(\gamma = 2\), which gives a good description of quasar pairs down to a separation of \(\lesssim 100\ \text{pkpc}\) (e.g., Shen et al., 2010). At small separations, we have \(\xi(r) \gg 1\), and \(\gamma = 2\) gives a flat distribution of \(\frac{dN_{\text{pair}}(r < R)}{dR}\), in agreement with the simulated AGN pairs in Rosas-Guevara et al. (2019).

By applying \(N_{\text{pair}}(0.9' < \Delta d < 3.0') > 1\) to the above equations, we get \(R_0 > 254h^{-1}\ \text{comoving Mpc (cMpc)}\), \(N_{\text{pair}}(r < 30\ \text{pkpc}) > 1.5\) and \(f_{\text{pair}}(r < 30\ \text{pkpc}) > 0.3\%\). Figure 5.4 shows the comparison between our results and the predictions from the cosmological simulations. The lower limit is several times lower than the predicted pair fraction. Note that the simulations count both obscured and unobscured AGNs. Since our survey only targets unobscured type-I quasars, we expect that the observed quasar pair fraction in our survey should be lower than the simulated AGN pair fractions. In addition, the AGN pair fraction depends on the luminosity and the SMBH mass, as discussed in De Rosa et al. (2019). Silverman et al. (2020) give an observed pair fraction of \(\sim 0.26\%\) for \(z \lesssim 4\) quasars, and their sample has a luminosity cut (\(\log L_{\text{bol}} > 45.3\)) that is comparable to our
Figure 5.4: The pair fraction of quasars, $f_{\text{pair}}(r < 30\text{ pkpc})$. The red line marks the lower limit set by J2037–4537. We also include the predictions of cosmological simulations, which are adapted from the Figure 6 in De Rosa et al. (2019). The simulated AGN samples have $L_{\text{bol}} > 10^{43}\text{erg s}^{-1}$, and include both obscured and unobscured AGNs. The observed lower limit is several times lower than the simulated sample. Note that the two quasars in J2037–4537 have $L_{\text{bol}} \sim 10^{47}\text{erg s}^{-1}$ (Table 5.2.1). Observations at lower redshift indicate that the pair fraction decreases with AGN luminosity, which could explain the difference between the simulations and the lower limit set by J2037–4537 (see text for details).

Silverman et al. (2020) show that the observed fraction is consistent with a luminosity-matched sample in the cosmological simulations, which indicates that the pair fraction might be lower for luminous AGNs. Current cosmological simulations do not have sufficient volumes to produce a quasar pair like J2037–4537 at $z > 5$. Future simulations with larger volumes will enable direct comparisons with observations and provide unique constraints on the SMBH evolution models at high redshift.

5.3.2 The Triggering Mechanism of the Quasar Pair

The tangential separation between the two quasars in J2037–4537 is 7.3 kpc, and the spectra suggest a redshift difference of $\Delta z \lesssim 0.01$. As such, this quasar pair must
reside in a galaxy merger. Although current data cannot rule out the possibility that the two quasars are triggered independently and coincidentally before the merging event, we argue that J2037–4537 is likely triggered by the galaxy merger. This is because the correlation length indicated by J2037–4537 ($R_0 > 254h^{-1}\text{cMpc}$) is much larger than the correlation length of $z \sim 5$ quasars ($R_0 \sim 20h^{-1}\text{cMpc}$, e.g., McGreer et al., 2016). The significant difference in $R_0$ at large and small scales indicates that close quasar pairs and quasar pairs with separation $\gtrsim 100\text{pkpc}$ might have different origins. The former are more likely results of galaxy mergers, while the latter are related to structure formation at sub-Mpc scale.

The observational features of J2037–4537 are consistent with merger-triggered quasar pairs in hydrodynamical simulations. Capelo et al. (2017) explore the factors that control the emergence of AGN pairs in galaxy mergers, and conclude that close quasar pairs usually appear in galaxy mergers with SMBH mass ratios close to one and separations less than 10 pkpc. It is thus plausible that the two quasars are triggered by the merging process. Future observations (likely with ALMA or the James Webb Space Telescope) will reveal more details of the gas kinematics and will allow investigating how the gas feeds the SMBHs in this system. Besides, we notice that the two quasars show similar profiles for a few broad emission lines (e.g., the Ly$\alpha$, N$\upsilon$, O$\upsilon$ and Si$\upsilon$ lines). If J2037–4537 is a physical quasar pair, these similarities might be causally connected to the physical association of the two quasars and encode critical information about how the quasar activities are triggered.

5.4 Summary

We report the discovery of a close quasar pair candidate at $z = 5.66$, J2037–4537, which has a projected separation of 1''24 (7.3 kpc). The ground-based spectroscopy shows that the resolved two objects are two quasars at the same redshift, and the high-resolution HST imaging does not detect the foreground lensing galaxy. Given the features in the spectra and the high-resolution images, it is highly unlikely that
J2037–4537 is a lensed quasar. J2037–4537 sets the first constraint on the pair fraction of quasars at $5 < z < 6$, $f_{\text{pair}}(r < 30 \text{ pkpc}) > 0.3\%$. Future observations of the gas kinematics in J2037–4537 will reveal more information about the triggering mechanism of the quasar pair.
6.1 Introduction

In the past two decades, more than 200 quasars at $z > 6$ have been discovered (e.g., Matsuoka et al., 2016, 2018b,c, 2019; Venemans et al., 2013, 2015; Yang et al., 2019c, 2020; Wang et al., 2017, 2019b; Jiang et al., 2016; Bañados et al., 2016, 2018). Studies of the quasar host galaxies provide crucial knowledge about the co-evolution of supermassive black holes (SMBHs) with their host galaxies and environment in the early universe. Detecting the host galaxies of high-redshift quasars is challenging at rest-frame ultraviolet to near-infrared wavelengths, where the emission from the central quasar overwhelms the host galaxy (e.g., Mechtley et al., 2012; Marshall et al., 2020). As such, information about quasar host galaxies is mostly from the far-infrared (FIR) and sub-millimeter (sub-mm) regime (e.g., Riechers et al., 2009; Wang et al., 2008, 2010; Venemans et al., 2012). The dust continuum and atomic and molecular emission lines (for example, the [C\(\text{II}\)] fine structure line and CO rotational lines) contain a wealth of information about the interstellar medium (ISM), including the dust mass and temperature (e.g., Beelen et al., 2006; Schreiber et al., 2018), the atomic and molecular gas mass (e.g., Weiß et al., 2005; Bolatto et al., 2013), and the gas-phase metallicity (e.g., Rigopoulou et al., 2018). Spatially resolved line emission also directly probes the gas-phase kinematics of quasar host galaxies and provides the only current way to measure their dynamical masses (e.g., Walter et al., 2009). The total-infrared (TIR) luminosity is widely used to estimate the star formation rate (SFR) (e.g., Murphy et al., 2011), assuming that the cool dust in the quasar host is dominantly heated by star formation (e.g., Beelen et al., 2006; Leipski et al., 2014; Lyu et al., 2016).

†This Chapter was published as Yue et al. (2021b)
With unprecedented sensitivity and resolving power, the Atacama Large Millimeter/Submillimeter Array (ALMA) has greatly improved our understanding of high-redshift quasars. To date, several tens of quasars at $z > 6$ have been observed by ALMA, including about 15 at $z > 6.5$. These observations led to an overall picture of the high-redshift quasar population: most of the high-redshift quasars are hosted by infrared-luminous, gas rich galaxies (e.g., Decarli et al., 2018; Venemans et al., 2018; Wang et al., 2019a; Shao et al., 2019), indicating active star formation ($\text{SFR} \gtrsim 10^2 \text{M}_\odot \text{yr}^{-1}$). The kinematics of the bright [C\text{II}] emission line constrains the dynamical masses of quasar host galaxies. Compared to the local $M_{\text{BH}} - M_{\text{host}}$ relation (e.g., Kormendy & Ho, 2013), SMBHs in quasars at $z \gtrsim 6$ are oversized (e.g., Venemans et al., 2016; Decarli et al., 2018; Wang et al., 2019c). While SMBHs might grow earlier than their hosts at high redshift, this difference may be a result of selection effects, i.e., current quasar surveys are biased toward luminous quasars, which have massive SMBHs (e.g., Lauer et al., 2007; Willott et al., 2015; Izumi et al., 2019).

At $z > 6$, quasar host galaxies usually have sizes of $\sim 2 - 4$ kpc (measured via the far-infrared dust continuum, e.g., Decarli et al., 2018), although they can be as compact as $\sim 1$ kpc (e.g., Venemans et al., 2017). Most ALMA observations of high-redshift quasars use beam sizes of $\gtrsim 0.3$, which marginally resolve these quasar host galaxies. These hosts have a variety of morphologies, ranging from a regular Gaussian profile (e.g., Shao et al., 2017; Venemans et al., 2018) to highly irregular, indicating an on-going merging system (e.g., Bañados et al., 2019; Neeleman et al., 2019). In a recent study, Venemans et al. (2019) reported 400-pc resolution imaging of a quasar host galaxy at redshift $6.6$, which shows complex structures of dust continuum and [C\text{II}] emission, including cavities with sizes of $\sim 0.5$ kpc. The authors propose that these cavities might be relevant to the energy output of the central active galactic nucleus (AGN). Sub-kpc resolution is thus necessary to the investigation of the structures in high-redshift quasars and to the understanding of SMBH-host coevolution.

Gravitational lensing acts as a natural telescope, significantly enhancing the
angular resolution and the sensitivity of observations (e.g., Hezaveh et al., 2016; Litke et al., 2019; Inoue et al., 2020; Cheng et al., 2020; Spilker et al., 2020). In Fan et al. (2019b), we reported the discovery of a gravitationally lensed quasar at $z = 6.51$, J043947.08+163415.7 (hereafter J0439+1634). High-resolution images taken by Hubble Space Telescope (HST) reveal the multiple images of the quasar generated by gravitational lensing. The lensing model based on HST data suggests that J0439+1634 is a naked-cusp lens with three images, with a total magnification of $51.3 \pm 1.4$. J0439+1634 is the only known lensed quasar at $z > 5$ to date and provides an excellent chance to study a high-redshift quasar in enhanced spatial resolution due to its large lensing magnification.

Here we report the sub-mm continuum and $[\text{C}\text{\textsc{ii}}]$ 158 $\mu$m emission line of J0439+1634 observed by ALMA at a resolution of $\sim 0'.3$. With the help of lensing, we reach a physical resolution of $\sim 0.8$ kpc. We describe our data in Section 6.2. In Section 6.3, we describe the measurement of the dust continuum and $[\text{C}\text{\textsc{ii}}]$ emission line, including the lens modeling and the reconstruction of the velocity field. We present the physical properties of J0439+1634 in Section 6.4 and discuss their implications for the evolutionary state of the quasar host galaxy in Section 6.5. We summarize this paper in Section 6.6. Throughout this paper, we use a $\Lambda$CDM universe with $H_0 = 70$ km s$^{-1}$Mpc$^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$.

### 6.2 Data

J0439+1634 was observed in ALMA Band 6 under configuration C43-5 in October 2018. The configuration contains 48 12-m antennas, which has a maximum baseline of 1.24 km. We tuned the four 1.875 GHz-wide spectral windows (SPWs) at 238.593 GHz, 236.718 GHz, 252.206 GHz, and 253.894 GHz with channel widths of 15.625 MHz, 15.625 MHz, 7.8125 MHz, and 7.8125 MHz, respectively. The $[\text{C}\text{\textsc{ii}}]$ emission line falls in the third SPW. The on-source exposure time is 99 minutes. We use J0510+1800 as the bandpass calibrator and J0440+1437 as the phase calibrator. The C43-5 observations are a part of Program 2018.1.00566.S, which aims at map-
ping the dust continuum and [C II] emission line to a spatial resolution of 0.03. The high-resolution observation with configuration C43-8 is not completed at the time of this paper’s writing.

We reduce the ALMA data using the Common Astronomy Software Applications (CASA) version 5.6.1 (McMullin et al., 2007). We use the task UVCONTSUB\footnote{This step is completed prior to the release of CASA version 5.6.1, and we use CASA version 5.4.0 when running UVCONTSUB.} to fit a linear function to the line-free channels, which models the continuum, and subtract the continuum model to obtain the line-only visibility. We then use the continuum data to perform phase self-calibration and apply the self-calibration model to the line-only data. We clean the continuum and line data with the CASA task TCLEAN using Briggs weighting, setting robust = 0.5. The synthesized beam has a size of 0.31 × 0.27 and a position angle of 39.4 degrees. Figure 6.1 shows the cleaned image of the dust continuum and the zeroth, first, and second moments of the [C II] emission. J0439+1634 is clearly resolved as an arc-like shape, which is typical for lensed galaxies. The zeroth moment (integrated flux) of the [C II] line is more extended than the continuum flux. The first moment (mean velocity map) shows ordered motion.

We extract the continuum and [C II] fluxes of J0439+1634 with a 2″ diameter aperture, which gives $S_{245\,\text{GHz}} = 16.0 \pm 0.1$ mJy for the continuum and $F_{[\text{C II}]} = 14.5 \pm 0.2$ Jy km s$^{-1}$ for the integrated [C II] flux. We also fit a Gaussian profile to the [C II] line using the CASA task SPECFIT. The [C II] line is centered at 252.7744 ± 0.0011 GHz with an FWHM of 270.0 ± 2.8 km s$^{-1}$, which gives a redshift $z_{[\text{C II}]} = 6.51871 \pm 0.00003$. In the rest of the paper, we set 252.7744 GHz as the rest frequency for [C II]. Figure 6.2 shows the extracted [C II] line profile, which is well-fitted by a Gaussian function and shows no evidence for an excess redshifted or blueshifted component.
Figure 6.1: The clean image and moments of J0439+1634 observed with ALMA. Upper Left: the continuum; Upper Right: the zeroth moment of [C\textsc{ii}] emission; Lower Left and Lower Right: the first and second moments of the [C\textsc{ii}] emission. The data are cleaned using Briggs weighting with robust = 0.5. In the first and second moment maps, we only show pixels that have integrated flux signal-to-noise ratio larger than 10. Contours in the continuum map are at 20\(\sigma\), 40\(\sigma\), 80\(\sigma\) and 160\(\sigma\) levels, where \(\sigma\) is the flux error per beam estimated using an annulus with 1" < \(r\) < 2". Similarly, Contours in the [C\textsc{ii}] moment 0 map are at 10\(\sigma\), 20\(\sigma\) and 40\(\sigma\) levels. The moment 1 map shows rotation-like ordered motion.
Figure 6.2: The [C\text{II}] line profile of J0439+1634, extracted from the cleaned image using a 2\arcsec diameter aperture. The error bar in the upper left corner shows the uncertainty of the flux density. The red line presents the best-fit Gaussian profile, which has a FWHM of 270.0\pm 2.8 km s\(^{-1}\) and a central frequency of 252.7744\pm 0.0011 GHz.

6.3 Lensing Model

We use VISILENS (Spilker et al., 2016) to model the visibility of J0439+1634. VISILENS is a parameterized lens modeling tool for interferometry data. In short, VISILENS models the \(uv\)-plane response of a lens system and obtains the posterior distributions of model parameters using Markov Chain Monte Carlo (MCMC).

6.3.1 Building the Lensing Model

Fan et al. (2019b) built the lensing model of J0439+1634 based on the \textit{HST} image, where they used a singular isothermal ellipsoid (SIE; e.g., Kormann et al., 1994) to describe the mass distribution of the lens galaxy. In their fiducial model, the lensing galaxy has a high ellipticity \((e = 0.65)\), lies at the west side of the quasar, and generates three quasar images. The position, the ellipticity, and the position angle of the modeled lens galaxy are consistent with the observed \textit{HST} optical image.

The \textit{HST} images have spatial resolution of \(\sim 0\farcs075\), which is several times better
Table 6.1. Lens model parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Default $^1$</th>
<th>ALMA only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lens</td>
<td>[C\text{,\textsc{ii}}]</td>
</tr>
<tr>
<td>Redshift</td>
<td>0.67</td>
<td>[6.5187]</td>
</tr>
<tr>
<td>$\Delta$RA (&quot;)</td>
<td>0.459 ± 0.002</td>
<td>0.236 ± 0.003</td>
</tr>
<tr>
<td>$\Delta$Dec (&quot;)</td>
<td>−0.050 ± 0.001</td>
<td>−0.019 ± 0.001</td>
</tr>
<tr>
<td>Mass (M$_\odot$)</td>
<td>[2.06 × 10$^{10}$]</td>
<td>-</td>
</tr>
<tr>
<td>$e$</td>
<td>0.65</td>
<td>0.409 ± 0.011</td>
</tr>
<tr>
<td>PA (deg)</td>
<td>94.58</td>
<td>-9.6 ± 1.2</td>
</tr>
<tr>
<td>Flux</td>
<td>-</td>
<td>3.46 ± 0.04</td>
</tr>
<tr>
<td>$R_{\text{eff}}$</td>
<td>-</td>
<td>0.136 ± 0.002</td>
</tr>
<tr>
<td>$n$</td>
<td>-</td>
<td>1.71 ± 0.06</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-</td>
<td>4.53 ± 0.05</td>
</tr>
</tbody>
</table>

$^1$In the default model, the parameters of the lens galaxy, except its position, are fixed to the fiducial model in Fan et al. (2019b).

$^2$ΔRA and ΔDec are relative to the phase center.

$^3$Mass of the lens galaxy.

$^4$Ellipticity of the lens or source.

$^5$Position angle (from north to east) of the lens or source. PA = 0 means that the major axis lies east-to-west.

$^6$The source flux, in mJy for the continuum and in Jy km s$^{-1}$ for the [C\text{\,\textsc{ii}}] line.

$^7$The half-light radius.

$^8$The S´ersic index.

$^9$The flux magnification.

Note. — The quantities in the square brackets are fixed. For the lens galaxy, Fan et al. (2019b) use Einstein radius instead of mass. Here we follow the convention in VISILENS. The redshift and the mass of the lens galaxy are degenerate, and the lens parameters in this table gives the same lensing model as Fan et al. (2019b). Besides, note that the uncertainties only include statistical errors, and do not take into account the systematic errors introduced by the model choice (see Section 6.3.5 for more discussion).
than the current ALMA data. We thus adopt the lens mass distribution from the fiducial model in Fan et al. (2019b). Specifically, we use an SIE to describe the lens galaxy. The Einstein radius, ellipticity, and position angle of the SIE are fixed to the values in the fiducial HST model, while the position of the lens is left free, which accounts for any pointing offsets between HST and ALMA. Because the continuum has a higher signal-to-noise ratio (SNR), we first fit the continuum to obtain the best-fit lens position, then apply the lens position when fitting the [C\text{II}] emission. We use a Sérsic profile to describe the source, both for the continuum and the [C\text{II}] line emission. This model is referred to as the default model in this paper.

For comparison, we also build an alternative model, hereafter referred as the “ALMA-only” model, in which we leave all parameters free when fitting the ALMA data and do not use any information from the HST observations. Again, we use an SIE to describe the lens galaxy and a Sérsic profile to describe the source emission, both for the continuum and the [C\text{II}] line. We first fit the continuum to obtain the best-fit values of the lens parameters, then apply these values when fitting the [C\text{II}] line.

Figure 6.3 shows the fitting result of the default model, and Figure 6.4 illustrates the ALMA-only model. Table 6.3 summarizes the best-fit parameters for both models. Despite tiny differences in details, the two models give the same overall lensing structure. Because the HST images have better resolution, we use the default model to derive the properties of J0439+1634 from this point on. We will discuss the systematic errors introduced by the choice of the model in Section 6.3.5.

In addition to the fiducial model, Fan et al. (2019b) raise two alternative models, in which the lens mass distribution differs significantly from the fiducial model and produces either double or quadruple quasar images. See Figure 4 in Fan et al. (2019b) for more information. The alternative models do not provide suitable fits to the ALMA data. We conclude that the fiducial HST model has the correct lensing configuration.
Figure 6.3: The default lensing model of J0439+1634 based on ALMA and HST observations. **Upper Panel:** the continuum model. From left to right: the observed dirty image with natural weighting, the modeled dirty image with natural weighting, the residual image, and the source model. Contours in the observed and model images are $-10\sigma$, $-5\sigma$, $5\sigma$, $10\sigma$, $20\sigma$, $50\sigma$, $100\sigma$, $200\sigma$ and $400\sigma$ levels, where $1\sigma$ equals to the “error per beam” in the observed image. Contours in the residual image are $-5\sigma$, $5\sigma$, $10\sigma$, $15\sigma$, $20\sigma$ and $25\sigma$ levels. In all images, dashed black lines are negative contours and solid black lines are positive ones. The red line marks the caustics of the lens. When fitting the continuum, we fix the lens parameters, except the position, to the fiducial model in Fan et al. (2019b). **Lower Panel:** Same as the upper panel, but for [C\textsc{ii}] emission. The lens parameters in the [C\textsc{ii}] fitting are fixed to the best-fit values in the continuum model.
Figure 6.4: Same as Figure 6.3, but for the ALMA-only lensing model. We first fit the continuum visibility with all parameters left free, then fit the $[\text{C} \text{II}]$ emission with lens galaxy parameters fixed to the best-fit values in the continuum model. This ALMA-only model and the default model in Figure 6.3 are nearly identical.

### 6.3.2 Dust Continuum

The upper panel of Figure 6.3 shows the best-fit dust continuum in the default model. As described in Section 6.3.1, in the default model, we fix the deflector galaxy mass distribution to the fiducial model in Fan et al. (2019b) which is based on $HST$ imaging, and fit the quasar host galaxy emission in ALMA data as a Sérsic profile. The dirty images are generated with natural weighting to enhance the SNR. The dust continuum of J0439+1634 can be well-fitted by a single Sérsic profile, with a reduced $\chi^2 = 1.035$. The best-fit Sérsic index is $1.71 \pm 0.06$ and the half-light radius is $0.136 \pm 0.002$ (0.74 \pm 0.01 kpc), suggesting a compact, exponential-disk-like profile. (See Section 6.4.1 for further discussion.) The overall magnification is $4.53 \pm 0.05$ when averaged over the entire galaxy. Compared to the fiducial $HST$ model, the position of the optical quasar deviates from the continuum center by $0.014$. The typical astrometric error for ALMA is about 5\% of the resolution, which translates to $\sim 0.015$ given a beam size of $\sim 0.3$ (the ALMA technical handbook,
e.g., Cortes et al., 2020). The positions of the optical quasar and the host galaxy are thus consistent.

The residual map shows some statistically significant structures. The peak of these structures is 4.9% of the peak in the observed dirty image. We expect such features given that we use a simple SIE + Sérsic model and the SNR of the data is high (with natural weighting, the peak SNR in the dirty image is \( \sim 600 \)). When we add a Gaussian profile to the source model, where we allow the Gaussian profile to have negative flux, the flux of the Gaussian profile converges to zero within the error. We thus argue that the structures in the residual image cannot be explained by a single bump or void in the source galaxy. The structures might result from an over-simplification of the lens and source model.

### 6.3.3 Integrated [C\textsc{ii}] Flux

The lower panel of Figure 6.3 shows the best-fit result of the [C\textsc{ii}] emission in the default model, and Table 6.3 shows the parameters of [C\textsc{ii}] line observations. The reduced \( \chi^2 \) of the best-fit model is 1.014. The best-fit [C\textsc{ii}] emission has a Sérsic index of \( 0.82 \pm 0.05 \), consistent with an exponential \((n = 1)\) profile, and a half-light radius of \( 0''.233 \pm 0''.006 \) \((1.27 \pm 0.03 \text{ kpc})\). The position and the ellipticity of the integrated [C\textsc{ii}] emission are consistent with those of the dust continuum within 2\( \sigma \), while the [C\textsc{ii}] line has a smaller Sérsic index and larger half-light radius. This difference suggests that the [C\textsc{ii}] line is more diffuse than the dust, as shown in the clean images. The overall magnification of the [C\textsc{ii}] emission is \( 3.44 \pm 0.05 \), which is smaller than that of the dust continuum, mainly because the [C\textsc{ii}] profile is more diffused.

Similar to the continuum, a Sérsic profile captures the major features of the [C\textsc{ii}] emission. The residual image of the [C\textsc{ii}] emission is similar to the continuum residual. The peak in the residual is 8.1% of the peak in the dirty image.
6.3.4 \[\text{[C}\ II]\ Kinematics\]

Figure 6.1 suggests that the host galaxy of J0439+1634 has an ordered, rotation-like velocity field. We thus fit the \([\text{C}\ II]\) emission using an axisymmetric rotating thin disk, following the method described in Neeleman et al. (2019). In short, we set up parameterized models for the flux distribution, the mean velocity field, and the velocity dispersion field. We then use VISILENS to calculate the lensed \([\text{C}\ II]\) emission and the \(uv\)-plane response in each velocity channel. We obtain the best-fit model parameters by minimizing the residual of the visibility in all channels. To keep maximum flexibility, we do not constrain the parameters using the Sérsic model for the integrated \([\text{C}\ II]\) flux. We assume a Sérsic profile for the flux distribution and apply various forms for the rotation curve and the velocity dispersion profile. However, all of these models return large residuals and unphysical best-fit parameters. We thus conclude that J0439+1634 cannot be described by an axisymmetric rotating thin disk.

The main reason for the poor fit is the apparent misalignment between the major axis of the flux distribution and the velocity gradient. For an axisymmetric rotation disk, the major axis and the velocity gradient should be in the same direction. In contrast, the major axis of the flux distribution of J0439+1634 is roughly aligned east-to-west (Figure 6.3, right panel), while the velocity gradient is roughly north-to-south (Figure 6.1). To further investigate this problem, we estimate the source-plane flux distribution using a simple inverse ray-tracing method. Specifically, we reconstruct the source (i.e., un-lensed) data cube on a grid with a pixel size of 0\('\)04. Using the overall \([\text{C}\ II]\) magnification \(\mu_{\text{[C}\ II]} = 3.44\), we estimate the average source-plane resolution to be \(\sim \sqrt{0.31 \times 0.27}/3.44 = 0\('\)156\) for the \([\text{C}\ II]\) emission. A pixel size of 0\('\)04 gives a super-Nyquist sampling, which helps to resolve the regions with higher magnification than the average value. We then trace all the pixels in the image-plane data cube (i.e., the clean image) to the source plane according to the default lensing model. If more than one image-plane pixels are traced to the same source-plane pixel, these image pixels are averaged. This simple method captures...
Figure 6.5: The reconstructed source-plane moments. From left to right: the reconstructed zeroth, first, and second moments. The coordinates are relative to the phase center. The images have pixel sizes of 0\"04 (0.22 kpc). The black cross marks the center of the \[\text{[CII]}\] emission from the default lensing model. The red line illustrates the lensing caustics. The moment 2 map has some pixels missing due to low SNR which leads to mathematical errors (i.e., getting a squared root of a negative value). The lower-right corner of the left panel shows the output when performing the inverse ray-tracing analysis to a beam on the image plane which locates at the flux peak. The “reconstructed” beam has a size of 0\"26 × 0\"09 and is a rough estimate of the source-plane beam shape.

The main features of the quasar host galaxy without expensive pixelized lensing reconstruction.

We generate source-plane moment maps using the reconstructed data cube. When calculating the first and the second moments, we only include pixels that have integrated flux SNR larger than 10. Figure 6.5 shows the reconstructed moments, which confirm the overall picture of J0439+1634: a regular profile for the integrated emission (moment 0) and a rotation-like mean velocity field (moment 1). The major axis of moment 0 is significantly offset from the velocity gradient in the moment 1 map, confirming the argument we made with the lensed image. Another hint is the structures in the moment 2 map. For a rotating thin disk, we expect a peak at the center of the moment 2 map due to the beam smearing effect where the line-of-sight velocity gradient is large. This peak is not seen in Figure 6.5; instead, the moment 2 map has complex structures, which indicates complicated velocity field in the host galaxy.

The lower-right corner of the reconstructed moment 0 map illustrates the output
when we perform the inverse ray-tracing analysis to an image-plane beam located at the image-plane flux peak. This “reconstructed” beam is a rough estimate of the beam shape on the source plane. The source-plane beam has a size of $0.26 \times 0.09$, which further illustrates that a source-plane pixel size of $0.04$ is appropriate.

Note that the blue and red wings in the moment 1 map are located outside of the caustics, which means that they are not multiply-imaged. For these areas, the effect of gravitational lensing on the observations is equivalent to shrinking the beam size and applying some image distortions. As such, the inverse ray-tracing reconstruction can capture the structure of the velocity field, especially in the blue and red wings. In Figure 6.6, we illustrate the position-velocity plot of J0439+1634, generated using the reconstructed moment 0 and moment 1. We extract the velocities along the black line, which connects the pixels with the maximum and minimum moment 1 value (i.e., the red and blue peaks). The position-velocity plot clearly shows a rotation-like feature. The velocity rises at $r \lesssim 1$ kpc and flattens beyond this radius. The maximum rotation velocity is roughly $v_{\text{max}} \sin(i) = 130$ km s$^{-1}$ where $i$ is the inclination angle of the rotation axis, and the velocity is measured out to $r_{\text{max}} = 2$ kpc.

In addition to the rotation-like feature, the position-velocity plot contains another component with low velocity ($|v| \lesssim 50$ km s$^{-1}$, marked by the yellow box). We mark pixels that contribute to this feature with black dots in the moment 1 map in Figure 6.6. These pixels either lie inside the caustics, which means that they are multiply imaged, and have a complicated lens-mapping function, or have low signal and large errors. The simple inverse ray-tracing method will fail for these areas, and this component is likely an artifact. If it is physical, it might reflect some complex structures in the host galaxy, which can be resolved in the upcoming high-resolution ALMA observations. Possible scenarios include minor mergers and clumpy star-formation regions. In any case, this structure only contributes $\sim 10\%$ of the total flux, and the rotation-like feature dominates the velocity field.

J0439+1634 has a rotation-like velocity field, but cannot be described by an axisymmetric rotating thin disk. It is likely that the [C II] emission is not axisym-
Figure 6.6: **Left**: The moment 1 map. The black line illustrates the axis used to extract the position-velocity diagram. The axis connects the pixels with the maximum and minimum moment 1 value (i.e., the blue peak and the red peak). The black dots mark the pixels that contributes to the minor feature in the position-velocity diagram (marked by the yellow box in the right panel). The red curve marks the caustics. **Right**: The position-velocity diagram. The diagram contains a rotation-like feature and a minor feature marked by the yellow box. The two cyan asterisks correspond to the value we adopt for the maximum velocity and the corresponding radius: $v_{\text{max}} \sin(i) = 130$ km s$^{-1}$ and $r_{\text{max}} = 2$ kpc.

metric and/or that the host galaxy has a thick geometry. Both cases are common for high-redshift galaxies (e.g., Pensabene et al., 2020; Förster Schreiber & Wuyts, 2020). Specifically, many star-forming galaxies at $z \gtrsim 2$ have velocity dispersion $\sigma \gtrsim 45$ km s$^{-1}$ and show thick geometry (for a recent review, see Förster Schreiber & Wuyts, 2020). For J0439+1634, we can estimate its velocity dispersion using the regions where the rotation curve has flattened. The moment 2 map suggests that these regions have $\sigma \sim 70$ km s$^{-1}$, which means that a thick geometry is likely.

Fitting a non-axisymmetric model and/or a thick disk model requires high spatial resolution, and our data cannot put a strong constraint on these models. In future work, we will perform pixelized lensing reconstruction and detailed dynamical modeling once the high-resolution ALMA observations are carried out (Project 2018.1.00566.S, PI: Fan).
6.3.5 Systematic Uncertainties

We first consider the systematic errors in the fluxes and sizes of the continuum and the \([\text{C}\ II]\) emission. The main source of systematic errors is that the SIE + Sérsic lensing model is over-simplified. As a result, there are positive clumps in the residual images in Figure 6.3. The flux of the clumps is much smaller than the flux calibration error and is negligible in the error analysis. Properly modeling the structures in the residual images requires expensive pixelized modeling of the source flux and the lens mass (e.g., Hezaveh et al., 2016), which is beyond the scope of this paper.

We then consider the systematic uncertainties in the source-plane reconstruction in Section 6.3.4. The major uncertainty is the beam-smearing effect. The source-plane resolution is \(\sim 0\farcs15\) (\(\sim 0.8\) kpc). In this study, we focus on the maximum rotation velocity rather than detailed velocity field structure. Beam-smearing effects have little influence on our main result, because (1) the blue and red peaks are only singly imaged, and (2) the rotation velocity flattens at \(r \gtrsim 1\) kpc, so the central, low-velocity area does not influence the edge at \(r_{\text{max}} = 2\) kpc where we measure the maximum velocity. Similar methods have been adopted by recent studies to measure the rotation velocity of lensed galaxies (e.g., Cheng et al., 2020).

6.4 Physical Properties of J0439+1634

6.4.1 Dust Continuum and \([\text{C}\ II]\) Emission

The host galaxy of J0439+1634 has a regular Sérsic profile, both for the continuum and \([\text{C}\ II]\) emission. The Sérsic index is close to one, which suggests that J0439+1634 is more similar to an exponential disk than a de-Vaucouleurs bulge with \(n = 4\). The \([\text{C}\ II]\) line profile is also well-described by a single-peaked Gaussian profile, with no excess of blueshifted or redshifted components. The smooth structures in the moment maps and the position-velocity plot disfavor the scenario of a close, ongoing major merger. In addition, no other objects are detected within the ALMA field of view (\(\sim 12''\)). These results suggest that J0439+1634 is not an on-going
major merger and does not exhibit significant outflow features in the [C\text{II}] velocity field. However, it is possible that J0439+1634 is a minor merger or a remnant of a recent major merger.

Yang et al. (2019b) measure the far-infrared (FIR) to centimeter-wavelength spectral energy distribution (SED), as well as the CO, [C\text{I}], [C\text{II}], O\text{I}, and H$_2$O emission lines of J0439+1634. Their NOrthern Extended Millimeter Array (NOEMA) observation gives $z_{[C\text{II}]} = 6.5188 \pm 0.0002$, $F_{[C\text{II}]} = 11.7 \pm 0.8$ Jy km s$^{-1}$, FWHM$_{[C\text{II}]} = 328.1 \pm 18.0$ km s$^{-1}$, and $S_{239\text{GHz}} = 14.0 \pm 0.1$ mJy. The typical flux calibration uncertainty is $\sim 15\%$ for NOEMA and $\sim 10\%$ for ALMA. The continuum and [C\text{II}] line fluxes reported in Section 6.2 are consistent with those in Yang et al. (2019b). The difference in the FWHM of the [C\text{II}] line is about $3\sigma$.

Based on these measurements, Yang et al. (2019b) calculate the infrared luminosities, emission line luminosities, star formation rate, dust mass, and gas mass without correcting for the lensing magnification. We refer the reader to Yang et al. (2019b) for the details of how these properties are calculated. In this work, with spatially resolved ALMA images, we calculate the de-lensed values of these quantities. For [C\text{II}]-based quantities, we apply the magnification of [C\text{II}] emission, $\mu_{[C\text{II}]} = 3.44$; for FIR-based quantities, we apply the continuum magnification, $\mu_{\text{cont}} = 4.53$. In addition, we apply the continuum magnification to the CO-based molecular gas mass, assuming that the dust continuum traces the molecular gas. The results are listed in Table 6.4.1. Similar to other high-redshift quasars (e.g., Walter et al., 2009; Decarli et al., 2018; Wang et al., 2019a), J0439+1634 is hosted by a gas-rich ultra luminous infrared galaxy (ULIRG) with intense star formation activity, with a TIR luminosity of $\sim 10^{13} L_\odot$.

One interesting aspect of the [C\text{II}] emission is the so-called “[C\text{II}] deficit.” In systems with high FIR surface brightness ($\Sigma_{\text{FIR}} \gtrsim 10^{11} L_\odot$ kpc$^{-2}$; e.g., Díaz-Santos et al., 2017; Herrera-Camus et al., 2018), the [C\text{II}]-to-FIR ratio decreases with $\Sigma_{\text{FIR}}$. This relation has been observed in many different systems, including nebulae in the Milky Way (for example, the Orion Nebula; e.g., Goicoechea et al., 2015), local luminous infrared galaxies (LIRGs) and ULIRGs (e.g., Díaz-Santos et al., 2013),
Table 6.2. De-lensed physical properties of J0439+1634

<table>
<thead>
<tr>
<th>[C\textsc{II}]-Based Properties</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(L_{[\text{C\textsc{II}}]}(L_\odot))</td>
<td>((3.5 \pm 0.3) \times 10^9)</td>
</tr>
<tr>
<td>(L'_{[\text{C\textsc{II}}]}(\text{K km s}^{-1}\text{pc}^2))</td>
<td>((1.7 \pm 0.1) \times 10^{10})</td>
</tr>
<tr>
<td>(M_{\text{C}^+}(M_\odot))</td>
<td>((1.1 \pm 0.1) \times 10^7)</td>
</tr>
<tr>
<td>(\text{SFR}<em>{[\text{C\textsc{II}}]}(M</em>\odot \text{ year}^{-1}))</td>
<td>(\sim 233 \sim 1398)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other properties</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(L_{\text{FIR}}(L_\odot))</td>
<td>((7.5 \pm 0.4) \times 10^{12})</td>
</tr>
<tr>
<td>(L_{\text{TIR}}(L_\odot))</td>
<td>((1.06 \pm 0.04) \times 10^{13})</td>
</tr>
<tr>
<td>(\text{SFR}<em>{\text{TIR}}(M</em>\odot \text{ year}^{-1}))</td>
<td>(1.56 \times 10^3)</td>
</tr>
<tr>
<td>(M_{\text{dust}}(M_\odot))</td>
<td>(4.9 \pm 0.2 \times 10^8)</td>
</tr>
<tr>
<td>(M_{\text{H}<em>2,\text{CO}}(M</em>\odot))</td>
<td>(1.19 \times 10^{10})</td>
</tr>
</tbody>
</table>

Note. — These values are calculated according to Table 1 in Yang et al. (2019b). We apply \(\mu = \mu([\text{C\textsc{II}}]) = 3.44\) for [C\textsc{II}]-based quantities and \(\mu = \mu(\text{continuum}) = 4.53\) for the other quantities. FIR luminosity includes flux in rest-frame 42.5 – 122.5 \(\mu\)m, and TIR luminosity includes flux in rest-frame 8 – 1000 \(\mu\)m. The uncertainties only reflect statistical errors.
high-redshift submillimeter galaxies (SMGs) (e.g., Oteo et al., 2016; Spilker et al., 2016; Litke et al., 2019) and quasars (e.g., Decarli et al., 2018; Neeleman et al., 2019). The underlying mechanism of the $[\text{C}\,\text{ii}]$ deficit might be complex. Some plausible scenarios include: (1) In regions with higher surface density, most carbon atoms are in the form of CO molecules rather than $\text{C}^+$ ions (Narayanan & Krumholz, 2017); (2) $[\text{C}\,\text{ii}]$ might become optically thick at high surface density (Luhman et al., 1998); (3) Large FIR surface brightness indicates strong dust absorption of the UV radiation, which is the main heating source of $[\text{C}\,\text{ii}]$ emission (Herrera-Camus et al., 2018); (4) $[\text{C}\,\text{ii}]$ emission might be saturated in warm gas (Muñoz & Oh, 2016).

With spatially resolved data, we can investigate the $[\text{C}\,\text{ii}]$ deficit in different regions of J0439+1634. Using the continuum and $[\text{C}\,\text{ii}]$ model in Figure 6.3, we measure the $[\text{C}\,\text{ii}]$ and FIR flux of J0439+1634 in four regions. The $n^{th}$ region is defined as $n \times R_{\text{eff}} < R < (n + 1) \times R_{\text{eff}}$, where $n = 0, 1, 2, 3$, and $R_{\text{eff}} = 0'136$ is the half-light radius of the continuum emission. All regions have the same center, ellipticity, and position angle as the continuum emission. The widths of the rings are close to the average source-plane resolution of $0'15$. The flux calibration error ($\sim 15\%$) is the dominant source of uncertainty, and we ignore other uncertainties.

Figure 6.7 illustrates the position of J0439+1634 on the $[\text{C}\,\text{ii}]/\text{FIR} - \Sigma_{\text{FIR}}$ plot. We include $z \gtrsim 6$ quasars from the literature for comparison. Specifically, three quasars from Shao et al. (2017) and Wang et al. (2019c) have spatially resolved measurements. We also include LIRGs from the Great Observatories All-sky LIRG Survey (GOALS) sample (Díaz-Santos et al., 2013), SMGs at $4.5 < z < 6.5$ (Riechers et al., 2013; Neri et al., 2014; Gullberg et al., 2018), and resolved regions of a gravitationally lensed SMG at $z = 5.7$, SPT0346 (Litke et al., 2019). Similar to other objects with high-resolution data, different regions in J0439+1634 tightly follow the $[\text{C}\,\text{ii}]$ deficit. Our result thus suggests that the $[\text{C}\,\text{ii}]$ deficit is related to physical processes on $\lesssim 1$ kpc scales, i.e., it reflects the properties of the local ISM rather than the entire galaxy.
Figure 6.7: The [C\text{II}]-FIR ratio versus FIR surface brightness plot for resolved regions in J0439+1634, local LIRGs in the GOALS sample (Díaz-Santos et al., 2013), quasars at $z \gtrsim 6$ (Wang et al., 2013; Venemans et al., 2016, 2017; Shao et al., 2017; Decarli et al., 2018; Izumi et al., 2018; Wang et al., 2019a,c; Venemans et al., 2020), and $4.5 < z < 6.5$ SMGs (Riechers et al., 2013; Neri et al., 2014; Gullberg et al., 2018; Litke et al., 2019). Besides J0439+1634, three quasars and an SMG in this plot have resolved measurements, including J1319 (Shao et al., 2017), J0129, and J1044 (Wang et al., 2019c), and SPT0346 (Litke et al., 2019). The points representing different regions in one object are connected by a line.
6.4.2 Host Galaxy Dynamics

Using the maximum line-of-sight velocity $v_{\text{max}} \sin(i) = 130 \, \text{km} \, \text{s}^{-1}$ and the corresponding radius $r_{\text{max}} = 2 \, \text{kpc}$, we estimate the dynamical mass within $r_{\text{max}}$ for J0439+1634, $M_{\text{dyn}} \sin^2(i) = 7.9 \times 10^9 M_{\odot}$. We report in Table 6.4.1 the H$_2$ mass of J0439+1634, $M_{\text{H}_2, \text{CO}}(M_{\odot}) = 1.19 \times 10^{10} M_{\odot}$, which gives a gas-mass fraction of $1.5 \times \sin^2(i)$. The gas-mass fraction is high except for very low inclinations.

Although the host galaxy of J0439+1634 has a regular shape and a rotation-like velocity field, we show in Section 6.3.4 that this galaxy is not an axisymmetric thin disk. Plausible scenarios include (1) the emission in the galaxy is not axisymmetric, which can happen when the star-forming regions are not evenly distributed on the disk, or (2) the host galaxy has a “thick” geometry (e.g., a spheroid) and is not a thin disk. The upcoming high-resolution ALMA data will reveal small-scale structures that might distinguish these models. Here we briefly discuss the implications of the thick disk model, since the moment 2 map indicates that a thick geometry is likely (Section 6.3.4).

A thick disk with a large velocity dispersion has a non-negligible turbulent pressure gradient which needs to be considered when estimating the dynamical mass. Following the discussion in Förster Schreiber & Wuyts (2020), we estimate the circular velocity of the host galaxy, $v_c$:

$$v_c^2 = v_{\text{max}}^2 + 2\sigma^2 \times (r/R_d) \quad (6.1)$$

which is related to the dynamical mass by $M_{\text{dyn}}(r) = v_c^2 r/G$, where $R_d$ is the scaling radius of the exponential disk (i.e., $I(r) \propto e^{-r/R_d}$). Applying $v_{\text{max}} = 130/\sin(i) \, \text{km} \, \text{s}^{-1}$, $r = 2 \, \text{kpc}$, $\sigma = 70 \, \text{km} \, \text{s}^{-1}$ and $R_d = 0.8 \, \text{kpc}$ according to the best-fit [CII] model in Section 6.3.3 yields

$$v_c^2 = \left[ \frac{130 \, \text{km} / \text{s}}{\sin(i)} \right]^2 + (160 \, \text{km} / \text{s})^2 \quad (6.2)$$

which suggests that, in the thick-disk model, the contribution of velocity dispersion is significant. However, the velocity dispersion should be taken as an upper limit given the beam smearing effect. It is hard to correct the beam smearing effect under
the current resolution; as such, we still use $M_{\text{dyn}} \sin^2(i) = 7.9 \times 10^9 M_\odot$ in the rest of this paper. This result illustrates the need of careful modeling with high-resolution data when measuring the dynamical mass of high-redshift quasar host galaxies.

Pensabene et al. (2020) analyze the archival ALMA data of 32 quasars to model their kinematics, where ten quasars at $z > 5.7$ are found to have rotation-like velocity fields. Among these ten quasars, three have a significantly misaligned flux major axis and velocity gradient. This result suggests that complicated kinematics are common in high-redshift quasars and that high-resolution observations are crucial to understanding high-redshift quasar host galaxies.

6.5 A Maximum Starburst System With Oversized Black Hole at Cosmic Dawn

6.5.1 A Maximum Star Forming Rotating System

Our analysis shows that the host galaxy of J0439+1634 is a compact ULIRG with vigorous star formation. Assuming that the dust continuum traces the SFR surface density (SFRD), we estimate the SFRD within the continuum half-light radius to be $\Sigma_{\text{SFR}} \approx 800 M_\odot \text{year}^{-1} \text{kpc}^{-2}$. Such a high $\Sigma_{\text{SFR}}$ is close to the highest SFRD values seen in the universe ($\sim 10^3 M_\odot \text{year}^{-1} \text{kpc}^{-2}$; e.g., Walter et al., 2009) and approaches the Eddington-limit of star formation (Thompson et al., 2005). In addition, we estimate the maximum SFR proposed by Elmegreen (1999), where the gas is assumed to collapse on a free-fall timescale, $t_{\text{ff}} = \sqrt{2R^3/GM} = 1.06 \sin(i) \times 10^7 \text{year}$. The maximum possible SFR is $\epsilon M_{\text{gas}}/t_{\text{ff}}$, where $\epsilon$ is the efficiency of gas turning into stars. This argument suggests that J0439+1634 has a maximum possible SFR of $\epsilon \sin^{-1}(i) \times 1.12 \times 10^3 M_\odot \text{year}^{-1}$ within a radius of $R < 2 \text{kpc}$. For any reasonable inclination angle ($i \gtrsim 5^\circ$), a high star formation efficiency is required ($\epsilon \gtrsim 0.1$). Our analysis suggests that J0439+1634 is forming stars at the maximum possible rate.

The rich gas reservoir and the vigorous star formation of J0439+1634 could be a remnant of a recent major merger or strong cold gas inflow (e.g., Dekel et al., 2009). The major merger remnant scenario is promising because it provides a natural
explanation to the misalignment between the major axis and the velocity gradient, i.e., the star-formation regions are not yet evenly distributed in the rotating disk. Under the current resolution, small-scale structures will get smoothed out, and the flux distribution mimics a Sérsic profile. Upcoming high-resolution ALMA data could reveal these possible structures.

6.5.2 SMBH-Host Co-evolution

Fan et al. (2019b) measures the SMBH mass of J0439+1634 to be $M_{\text{BH}} = (4.29 \pm 0.60) \times 10^8 M_\odot$, which gives $M_{\text{BH}}/M_{\text{dyn}} = 0.055 \sin^2(i)$. Assuming J0439+1634 follows the local relation in Kormendy & Ho (2013) yields $M_{\text{BH}}/M_{\text{host}} = 0.005$. A face-on rotation model with inclination $i = 17^\circ$ moves J0439+1634 onto the local $M_{\text{BH}} - M_{\text{host}}$ relation, and a fiducial inclination angle of $i = 60^\circ$ yields $M_{\text{BH}}/M_{\text{host}} = 0.04$. This result is similar to that in many high-redshift quasars (e.g., Venemans et al., 2017; Decarli et al., 2018; Wang et al., 2019a), which have $M_{\text{BH}}/M_{\text{host}}$ several times higher than the local relation.

With the SMBH mass and the observed central velocity dispersion, we estimate the size of the SMBH’s sphere of influence:

$$r_h = \frac{GM_{\text{BH}}}{\sigma^2}. \quad (6.3)$$

Both the image-plane moment 2 (Figure 6.1) and the reconstructed source-plane moment 2 (Figure 6.5) show roughly constant velocity dispersion across the galaxy. We thus adopt the median value of the source-plane moment 2 map, $\sigma_{\text{med}} = 94 \text{ km s}^{-1}$, which gives $r_h = 0.18 \text{ kpc}$. The most extended configuration of ALMA delivers a resolution of $\sim 0.02$. For the region near the SMBH, we apply the magnification of the optical quasar from Fan et al. (2019b), $\mu_{\text{quasar}} = 51.3$. An image-plane resolution of $0.02$ thus corresponds to a source-plane resolution of $\sim 2.8 \text{ mas (} \sim 16 \text{ pc})$. Thus, high-resolution ALMA observations will allow us to sample the SMBH’s sphere of influence well and to measure the mass of the SMBH directly via gas kinematics.

Direct measurement of SMBH mass has been possible only at low-redshift (for a review, see Kormendy & Ho, 2013). For most high-redshift quasars, SMBH masses
are measured based on the empirical relation between the continuum luminosity and the broad line region size (e.g., Vestergaard & Peterson, 2006), which has only been calibrated at $z \lesssim 2$. J0439+1634 thus provides a unique opportunity to calibrate the SMBH mass measurement at high redshift.

### 6.6 Conclusions

We present ALMA observations of a gravitationally lensed quasar at $z = 6.52$, J0439+1634. We model the dust-continuum, the $\text{[C II]}$ emission, and the velocity field of the host galaxy. Our main conclusions are:

1. The ALMA observations demonstrate that the three-image fiducial model in Fan et al. (2019b) based on HST observations of the quasar is correct, ruling out the alternative models considered in Fan et al. (2019b). The default lensing model gives an overall magnification of $4.53 \pm 0.05$ and $3.44 \pm 0.05$ for the continuum and $\text{[C II]}$ emission of the host galaxy, respectively. The average source-plane resolution is $\sim 0''.15$ ($\sim 0.8$ kpc).

2. J0439+1634 is a compact ULIRG well-described by a compact Sérsic profile. The Sérsic index is close to one for both the continuum and $\text{[C II]}$ emission. The resolved regions in J0439+1634 follow the “[C II] deficit,” suggesting that the deficit is related to the sub-kpc properties of the ISM.

3. J0439+1634 has a rotation-like velocity field, but it cannot be well described as an axisymmetric rotating thin disk. The maximum line-of-sight rotation velocity is $v_{\text{max}} \sin(i) = 130$ km s$^{-1}$, with the inclination angle $i$ unconstrained. The dynamical mass within 2 kpc is $7.9 \times 10^9 \sin^2(i) M_\odot$. J0439+1634 is likely a gas-rich galaxy with a high gas-mass fraction.

4. J0439+1634 is forming stars at the maximum possible rate. The star-formation rate surface density of J0439+1634 approaches the largest value seen in the universe and the Eddington limit.
5. The SMBH-to-dynamical mass ratio of J0439+1634 is \(0.055 \times \sin^2(i)\), which suggests that J0439+1634 is likely to host an oversized SMBH compared to local relations. The size of the sphere of influence is 0.18 kpc. The most extended configuration of ALMA will resolve the sphere of influence and allow us to measure the SMBH mass directly using the gas kinematics.

Our lensing model incorporates the major features of J0439+1634 detected under low resolution ALMA data. Future high-resolution ALMA observations with higher-resolution, combined with pixelized lens modeling, will reveal more detailed structures in the foreground lens and quasar host galaxy. Specifically, as discussed in Section 6.5.2, the resolution around the SMBH will reach \(\sim 16\) pc, likely within the SMBH sphere of influence. The power of gravitational lensing makes J0439+1634 a valuable object for a case study, which will provide crucial and previously inaccessible information about the coevolution of SMBHs and their hosts at \(z > 6\).
CHAPTER 7

Other High-redshift Lensed Quasars and Close Quasar Pairs

In this Chapter, I describe the other high-redshift ($z \gtrsim 5$) lensed quasars and quasar pairs discovered in this project, including J0025–0145 (a lensed quasar at $z = 5.07$) and J2329–0522 (a close quasar pair at $z = 4.83$). These objects are good examples of how the candidate selection method in Chapter 4 works and how follow-up observations confirm candidates as lensed quasars or quasar pairs.

7.1 J0025–0145: a Lensed Quasar at $z = 5.07$

7.1.1 Initial Identification

J0025–0145 was initially discovered by Wang et al. (2016) as a luminous quasar with $M_{1450} = -28.5$. In this project, J0025–0145 was identified as a subtype Ia lens candidate (i.e., a point source with quasar-like color) via its large image fitting residual in Legacy Survey DR8. In Legacy Survey DR9, the iterative source detection process identifies a faint, extended object next to the point source. Figure 7.1 shows the Legacy Survey DR9 images of J0025–0145. The quasar corresponds to the point source, which has $m_g = 22.63$, $m_r = 19.79$ and $m_z = 18.05$. The image fitting residuals indicate a complicated structure.

Detailed analysis of the DECaLS $g$–band image supports the existence of a foreground deflector galaxy. Due to the absorption of the neutral hydrogen in the intergalactic medium (IGM), a quasar at $z = 5.07$ should have little flux at wavelengths bluer than the Lyman limit, and zero flux at rest wavelengths $\lambda_{\text{rest}} \lesssim 820\,\AA$ (e.g., Worseck et al., 2014). In contrast, there is clearly an extended source detected in the DECaLS $g$–band, which falls at the blue side of the Lyman limit at $z = 5.07$. The extended shape of the object in $g$–band rules out the possibility for it to be a
point source, e.g., a galactic star. In addition, the centers of the DECaLS \( r \)- and \( z \)-band flux, where the quasar flux dominates, have significant offsets from that of the \( g \)-band flux. These features suggest that the \( g \)-band flux comes from a foreground galaxy rather than the quasar. Similarly, the two-D spectrum of J0025–0145 (taken by Magellan/Clay LDSS-3; Figure 7.1, lower panel) shows a faint trace that extends to \( \sim 4200 \AA \) (rest-frame \( \sim 700 \AA \)), also suggesting the existence of a lensing galaxy.

J0025–0145 is also selected into the supplement candidate sample based on its high Eddington ratio. Figure 4.6 shows that J0025–0145 has one of the highest Eddington ratios \( R_{\text{Edd}} = 6.06 \pm 0.59 \) among the quasars at \( z \gtrsim 6 \). The extreme Eddington ratio indicates a large lensing magnification. Assuming that this quasar is accreting at the Eddington limit, I estimate the magnification of J0025–0145 to be \( \mu \sim 90.4 \).

### 7.1.2 High-resolution HST Imaging

J0025–0145 is a typical subtype Ia candidate that appears to be a point source in the Legacy Survey. Although the foreground deflector galaxy is detected in the survey images, the lensing structure is not resolved. To confirm its lensing nature, I obtained the high-resolution images of J0025–0145 using HST (Project ID: 16507). The upper panel of Figure 7.2 shows the 1D spectrum of J0025–0145 and the filters used for HST imaging. The F435W filter covers the rest-frame wavelengths \( \lambda_{\text{rest}} \lesssim 820 \AA \), which only detects the flux from the foreground lensing galaxy; the F606W filter detects both the foreground galaxy and the background quasar.

The lower panel of Figure 7.2 presents the high-resolution HST images. Both the F435W and the F475W image have an exposure time of 2072s, corresponding to one orbit. The F435W image clearly reveals the existence of a low-redshift deflector galaxy, which has a high ellipticity \( (e \approx 0.5) \) with some clumpy structures. Interestingly, there is only one point source in the F606W image, suggesting that the background quasar is not multiply-imaged. The quasar image and the foreground galaxy are separated by \( \sim 0''5 \). With such a small angular separation, the
Figure 7.1: Ground-based observations of J0025–0145. All images have sizes of $3'' \times 3''$. Top: the Legacy Survey $grz$ image shown in RGB format. The middle panel shows the Legacy Survey DR9 Tractor model, where J0025–0145 is modeled as a point source (the blue cross) plus a de Vaucouleurs profile (the blue ellipse). The residual image (top right) still has some noticeable residual. Middle: the DECaLS $g$–band image. I re-fit this image using galfit (middle), which is well-described by an exponential profile. The yellow and red contours mark the flux in the DECaLS $r$– and $z$–bands where the quasar dominates the flux. Bottom: the 2D spectrum taken by Magellan/Clay LDSS3. A trace is clearly visible at rest frame $\lambda < 820\AA$, where a quasar at $z > 5$ should have no flux. The $g$–band image and the faint trace in the two-D spectrum suggests the existence of a foreground galaxy.
Figure 7.2: The HST ACS/WFC images of J0025–0145. Top: the choice of filters. The F435W filter covers rest frame $\lambda < 820\ang$, where the background quasar has no flux. The F435W filter will thus only detect the foreground lensing galaxy. The F606W filter will detect both the quasar and the galaxy. Bottom: the HST ACS/WFC images. The foreground galaxy has a high ellipticity ($e \approx 0.5$) and lies $\sim 0\farcs5$ away from the quasar image. Note that there is only one lensed image of the background quasar.

background quasar must be strongly magnified by the deflector galaxy. However, without multiple lensed images of the background quasar, I am not able to construct a lensing model for J0025–0145. Note that it is possible to have a high magnification without generating multiple images of the background source. Specifically, in lensing systems that have highly eccentric deflectors and/or large external shears, the inner cusp of the lensing caustics may reach and exceed the boundary of the region where sources can get multiply-imaged. J0439+1634 provides a good example of this case, which turns out to be a triply-imaged naked cusp lens (Fan et al., 2019b). If the background source is
close to but does not fall inside the cusp, the source may have a high magnification (up to infinity) without being multiply imaged. See Kormann et al. (1994) for more discussions about the configurations of an elliptical deflector. As shown in Figure 7.2, J0025–0145 has a highly eccentric deflector, and a large magnification without multiple images is possible.

7.1.3 Discussion and Future Work

The most critical pending task about J0025–0145 is lens modeling. The only feasible way is to image the host galaxy, which will appear to be an extended arc and will reveal the lensing structure. Furthermore, given the typical sizes of quasar host galaxies at $z > 5$ ($\sim 2 - 4$ kpc, e.g., Decarli et al., 2018), part of the quasar host galaxy will fall into the caustics of the foreground lensing galaxy and generate multiple images. Given the strong flux from the central quasar in other wavelengths, this task can only be accomplished by sub-mm interferometers. In the upcoming semesters, I will use ALMA to image the host galaxy of J0025–0145 and build an accurate lensing model for J0025–0145.

The lensing model will provide the exact magnification of the optical quasar. If the high magnification indicated by the extreme Eddington ratio is confirmed, J0025–0145 will offer a unique chance to calibrate the $M_{\text{BH}} - \sigma$ (or $M_{\text{BH}} - M_{\text{host}}$) relation at high redshifts. Similar to J0439+1634, high-resolution ALMA observations with beam sizes $\sim 0''.02$ may be able to resolve the sphere of influence of the SMBH, enabling direct measurements of the SMBH mass via [C II] gas kinematics (see discussions in Section 6). J0025–0145 and J0439+1634 will calibrate the SMBH mass measurements at $z > 5$ that still suffer significant systematic errors (e.g., Grier et al., 2019), shedding light on the assembly history of high-redshift SMBHs and their host galaxies. Meanwhile, if the magnification turns out to be low, J0025–0145 will be confirmed as one of the most extreme super-Eddington accretion systems at high redshifts.

J0025–0145 and J0439+1634 are the only two lensed quasars at $z > 5$ known to date. Interestingly, both lensing systems have highly eccentric deflectors ($e \gtrsim 0.5$),
while the ellipticity of the deflectors in low-redshift lensing systems are usually low 
\((e < 0.5; \text{e.g., Koopmans et al., 2006})\). Moreover, J0439+1634 is a naked-cusp 

lens with three images of the background quasar, and Oguri & Marshall (2010) 

predict that only \(\sim 0.5\%\) of the lensed quasars are naked-cusp lenses. As such, 

J0025–0145 and J0439+1634 indicate that our knowledge about high-redshift lensed 

objects might be inaccurate. In particular, the properties of deflector galaxies might 

have significant redshift evolution. For example, Bruce et al. (2014) find that the 

bulge-to-disk ratio in galaxies increase significantly from \(z = 3\) to \(z = 1\), and 

Joachimi et al. (2013) suggest that disk-dominated galaxies have higher ellipticities 

than bulge-dominated ones. These trends might explain why both J0025–0145 and 

J0439+1634 have highly eccentric deflectors.

### 7.2 J2329–0522: a Close Quasar Pair at \(z = 4.83\)

#### 7.2.1 Initial Identification and Ground-Based Spectroscopy

J2329–0522 was initially identified as a subtype IIa lens candidate (i.e., unresolved, 

extended objects with quasar-like color). Similar to J0025–0145, J2329–0522 was 

fitted as a single object in Legacy Survey DR8, and the iterative source detection 

in Legacy Survey DR9 finds another faint source in the field. Figure 7.3 illustrates 

the Legacy Survey DR9 image of J2329–0522, where the system is modeled as a 

point source plus a de Vaucouleurs profile. The image fitting residual suggests that 

the Legacy Survey model might be inaccurate. I re-fit the \(z\)-band image using 

\textit{galfit}, which suggests that J2329–0522 can be well-described by two point sources 

separated by \(1.58\). The image fitting process closely follows the method in Chapter 

5.

J2329–0522 was observed using the Magellan/Clay LDSS-3 spectrograph in Oc-

tober 2021. I used the VPH-Red grating that delivers a wavelength coverage from 

5000Å to 10000 Å and a resolution of \(R \sim 1000\). Figure 7.4 shows the spectrum 

of the two components, where both traces exhibit features of a quasar at \(z = 4.83\). 

The seeing of the observation is \(\sim 0.6\), sufficient to well-resolve the two components.
Figure 7.3: The Legacy Survey images of J2329–0522. Top: the Legacy Survey DR9 grz–image shown in RGB format. The middle and right panels show the Tractor model and residual. J2329–0522 is modeled by a point source (the blue cross) plus a de Vaucouleurs profile (the blue ellipse) in the Legacy Survey DR9, where the point source is detected via iterative source detection. Bottom: the Legacy Survey (DECaLS) z–band image. I re-fit the image using galfit, finding that J2329–0522 can be well-described by two point sources. The two point sources are separated by 1.58.
Figure 7.4: The Magellan/Clay LDSS-3 2-D Spectrum. Taken with VPH-Red grism and 1'' slit. Top: the two-D spectra around the Lyα emission line. With a seeing of \( \sim 0.6'' \), the two objects are clearly resolved. Bottom: the extracted one-D spectrum. Both components show features of a \( z = 4.83 \) quasar. Note that the two quasars exhibit different emission line profiles, especially for the C IV line.

The two traces exhibit different C IV line profiles, suggesting that J2329–0522 is a physical pair of quasars instead of a doubly-imaged lensed quasar. The projected distance between the two quasars is 1''58 (10.1 kpc at \( z = 4.83 \)), and the two quasar host galaxies must be merging.

J2329–0522 shares many features with J2037–4537 (Chapter 5). In particular, the fainter component in the optical has a redder continuum in both quasar pairs. If the similarity is physical rather than coincidental, this result indicates some casual connections between the observed quasar SEDs and the galaxy merging event. This process is still poorly understood.

### 7.2.2 Future Work

The follow-up observations of J2329–0522 are still in the early stage. Specifically, the properties of the SMBHs (e.g., masses and Eddington ratios) remain unknown, which can be measured using the MgII broad emission line via near-IR spectroscopy.
High-resolution imaging will be useful to fully rule out the existence of a third object in this field, making sure that the differences in the quasar SEDs shown in Figure 7.4 is intrinsic.

J2329–0522 and J2037–4537 offer unique opportunities to investigate how high-redshift galaxy mergers trigger quasar activities. In the future, I will carry out follow-up observations of these two systems to characterize how the SMBHs, their host galaxies and the circumgalactic medium interact with each other. Specifically, JWST imaging and integrated field units (IFUs) in rest-frame optical will probe the emission from stars and ionized gas. Ground-based IFUs like MUSE will probe the circumgalactic medium around J2037–4537 via the Ly$\alpha$ emission, which will reveal gas inflows and/or outflows driven by accretion and/or quasar feedback. ALMA and VLA observations will measure molecular and atomic emission lines, characterizing the kinematics and the interstellar medium diagnostics in the galaxy merger. The final product will be multi-wavelength data set of the two quasar pairs (and probably newly-discovered ones), providing crucial clues about the evolution of SMBHs and galaxy mergers in the early universe.
CHAPTER 8

Summary

In my thesis, I present our new survey for high-redshift (\(z \gtrsim 5\)) gravitationally lensed quasars. We start from building mock catalogs and theoretical models of lensed quasars (Chapter 2 and Chapter 3), showing that there are \(\sim 10\) discoverable lensed quasars at \(z > 5\) in current sky surveys. We then describe the design of the survey, aiming at a complete and efficient search for all types of lensed quasars (Chapter 4). In the past few years, we have observed \(\sim 200\) of lens candidates, among which we have confirmed a few lensed quasars and close quasar pairs at \(z \gtrsim 5\) (Chapter 5 and Chapter 7). In addition, we present the ALMA observation of J0439+1634, a lensed quasar at \(z = 6.52\), illustrating the power of lensing magnification in observations of high-redshift SMBHs and their host galaxies (Chapter 6).

The main results are summarized as follows:

(1) We present a mock catalog of gravitationally lensed quasars at \(z_{\text{qso}} < 7.5\) with simulated images for the Rubin Observatory Legacy Survey of Space and Time (LSST). We adopt recent measurements of quasar luminosity functions to model the quasar population, and use the CosmoDC2 mock galaxy catalog to model the deflector galaxies, which successfully reproduces the observed galaxy velocity dispersion functions up to \(z_d \sim 1.5\). The mock catalog is highly complete for lensed quasars with Einstein radius \(\theta_E > 0.007\) and quasar absolute magnitude \(M_i < -20\). We estimate that there are \(\sim 10^3\) lensed quasars discoverable in current imaging surveys, and LSST will increase this number to \(\sim 2.4 \times 10^3\). Most of the lensed quasars have image separation \(\Delta \theta > 0.5\), which will at least be marginally resolved in LSST images with seeing of \(\sim 0.7\). There will be \(\sim 200\) quadruply-lensed quasars discoverable in the LSST. The fraction of quad lenses among all discoverable lensed quasars is about \(\sim 10\% - 15\%\), and this fraction decreases with survey depth. This
mock catalog shows a large diversity in the observational features of lensed quasars, in terms of lensing separation and quasar-to-deflector flux ratio. We discuss possible strategies for a complete search of lensed quasars in the LSST era.

(2) We revisit the lensed fraction of high-redshift quasars predicted by theoretical models, where we adopt recent measurements of galaxy velocity dispersion functions (VDFs) and explore a wide range of quasar luminosity function (QLF) parameters. We use both analytical methods and mock catalogs which give consistent results. For ordinary QLF parameters and the depth of current high-redshift quasar surveys ($m_z \lesssim 22$), our model suggests a multiply-imaged fraction of $F_{\text{multi}} \sim 0.4\% - 0.8\%$. The predicted lensed fraction is $\sim 1\% - 6\%$ for the brightest $z_s \sim 6$ quasars ($m_z \lesssim 19$), depending on the QLF. The systematic uncertainties of the predicted lensed fraction in previous models can be as large as factors of $2 - 4$ times and are dominated by the VDF. Applying VDFs from recent measurements decreases the predicted lensed fraction and relieves the tension between observations and theoretical models. Given the depth of current imaging surveys, there are $\sim 15$ lensed quasars at $z_s > 5.5$ detectable over the sky. Upcoming sky surveys like the LSST survey and the Euclid survey will find several tens of lensed quasars at this redshift range.

(3) Simulated lensed quasars suggest that previous surveys for high-redshift quasars have low completeness for lensed ones. Specifically, the foreground lensing galaxy has non-negligible contribution to the SED of the whole lensing system, such that color selection of high-redshift quasars fails to select lensed ones. Furthermore, most quasar surveys only consider point sources as candidates, while lensed quasars usually appear to be extended sources in imaging surveys.

(4) We developed a candidate selection pipeline for lensed quasars, utilizing the DESI Legacy Imaging Survey and other optical and infrared surveys. The depth and deblending ability of the Legacy Survey allows discovery of faint and small-separation lensed quasars. The main candidate selection pipeline identifies two types of lens candidates: “companion selected candidates” that are resolved into multiple components, where at least one component has quasar-like colors; “residual-selected candidates” that cannot be well-fitted by ordinary flux profiles and have a mixed
color between galaxies and quasars. We also include supplementary candidate samples to enhance the survey completeness. We estimate that the candidate selection should have a high completeness for lensed quasars.

(5) We report the discovery of a close quasar pair candidate at $z = 5.66$, J2037−4537. J2037−4537 is resolved into two quasar images at the same redshift in ground-based observations. Followup spectroscopy shows significant differences in both the continuum slopes and emission line properties of the two images. The two quasar images have a projected separation of $1''24$ (7.3 kpc at $z = 5.66$) and a redshift difference of $\Delta z \lesssim 0.01$. High-resolution images taken by Hubble Space Telescope do not detect the foreground lensing galaxy. The observational features of J2037−4537 strongly disfavor the lensing hypothesis. If J2037−4537 is a physical quasar pair, it indicates a quasar clustering signal of $\sim 10^5$ at a separation of $\sim 10$ proper kpc (pkpc), and gives the first observational constraint on the pair fraction of $z > 5$ quasars, $f_{\text{pair}}(r < 30 \text{ pkpc}) > 0.3\%$. The properties of J2037−4537 are consistent with those of merger-triggered quasar pairs in hydrodynamical simulations of galaxy mergers.

(6) We have observed $\sim 200$ lens candidates in the initial stage of our survey. In addition to J2037−4537, we have discovered one lensed quasar and one close quasar pair at $z \gtrsim 5$. These objects include J0025−0145 (a lensed quasar at $z = 5.07$) and J2329−0522 (a close quasar pair at $z = 4.83$). We have also discovered a few low-redshift lensed quasars and close quasar pairs.

(7) We present ALMA observation of J0439+1634, a gravitationally lensed quasar at $z = 6.5$. With the help of lensing magnification, the observation reaches an effective spatial resolution of $\sim 0.8$ kpc. The host galaxy of J0439+1634 can be well-fitted by a Sérsic profile consistent with an exponential disk, both in the far-infrared (FIR) continuum and the [C II] emission. The host galaxy of J0439+1634 is a compact ultra-luminous infrared galaxy, with a total star formation rate (SFR) of $1.56 \times 10^3 M_\odot$/year after correcting for lensing and an effective radius of 0.74 kpc. The reconstructed velocity field of J0439+1634 appears to be rotation-like. The maximum line-of-sight rotation velocity of 130 km/s at a radius of 2 kpc. However,
our data cannot be fit by an axisymmetric thin rotating disk, and the inclination of the rotation axis, $i$, remains unconstrained. We estimate the dynamical mass of the host galaxy to be $7.9 \sin^{-2}(i) \times 10^9 M_\odot$. J0439+1634 is likely to have a high gas-mass fraction and an oversized SMBH compared to local relations. The SFR of J0439+1634 reaches the maximum possible values, and the SFR surface density is close to the highest value seen in any star-forming galaxy currently known in the universe.

This survey for high-redshift lensed quasars is still on-going. In the future, the completeness and efficiency of the survey will be improved by new data releases of public sky surveys and updates in the candidate selection pipeline. When finished, this survey will yield the first statistical samples of lensed quasars and close quasar pairs at $z \gtrsim 5$, enabling various follow-up studies about the evolution of SMBHs and their hosts in the early universe.
APPENDIX A

Quasars Have Fewer Close Companions than Normal Galaxies†

A.1 Introduction

Active Galactic Nuclei (AGNs) play important roles in not only the growth of supermassive black holes (SMBHs) in the centers of galaxies, but also in the evolution of their host galaxies. SMBHs gain most of their mass through the AGN phase, and AGN activities can significantly influence the evolution of their host galaxies (for a recent review, see Kormendy & Ho, 2013). AGN can heat up and expel gas from their host galaxies and thus quench star formation (e.g., Croton et al., 2006; Cicone et al., 2014; Spacek et al., 2016). To draw the whole picture of SMBH and galaxy evolution through cosmic time, it is crucial to understand the triggering mechanism of AGNs.

Two scenarios have been proposed that can trigger AGN activities: major mergers of galaxies (especially gas-rich ones) and secular evolution. Major mergers can disturb gas in galaxies and generate gas inflows that are needed to feed the central SMBH (e.g., Barnes & Hernquist, 1991; Hopkins et al., 2006). Secular evolution happens when instabilities in galaxies, including those induced by galaxy bars or resulting from the gas inflows from the environment, drive gas to gradually move inward and fuel the SMBH (e.g., Shlosman et al., 1989; Hopkins & Quataert, 2010).

While simulations have shown that both scenarios can lead to rapid SMBH growth in the AGN phase, observational evidence regarding which one is the dominant mechanism remains ambiguous. A crucial test is to measure the merging fraction of their host galaxies (e.g., Grogin et al., 2005; Karouzos et al., 2014). The *Hubble Space Telescope (HST)* is powerful in identifying galaxy mergers in AGNs because of its small Point Spread Function (PSF). Cisternas et al. (2011) analyzed

†This Chapter was published as Yue et al. (2019)
the morphology of X-ray selected AGNs; they found that their merging fraction is the same as inactive galaxies within measurement errors. Villforth et al. (2017) worked on a sample of more luminous X-ray selected AGNs and reached a conclusion that was similar to Cisternas et al. (2011). However, Fan et al. (2016) reported an enhanced merging fraction of infrared-selected AGNs. Using imaging data from the Hyper-Supreme Camera (HSC) Survey (Aihara et al., 2018a) which used the Subaru Telescope, Goulding et al. (2018) also showed that the merging fraction of infrared-selected AGNs is larger than that of inactive galaxies. Treister et al. (2012) argued that major mergers are only responsible to the most luminous AGNs. These observational results, together with simulations (e.g., Hopkins et al., 2006; Hopkins & Quataert, 2010), suggest that the two triggering mechanisms may dominate for AGNs of different properties (e.g., luminosity, obscuration, or redshift).

The majority of previous studies used disturbed galaxy morphology to identify recent merging events. Galaxy pairs, or close companions of galaxies, is another frequently used indicator of galaxy mergers (e.g., Man et al., 2012). Ellison et al. (2011) measured the “pair fraction” (i.e., the average number of close companions) of 11060 galaxies from Sloan Digital Sky Survey (SDSS), claiming that the merging fraction of emission-line-selected low-luminosity AGNs is larger than that of the galaxy control sample.

Although there have been a number of previous studies to constrain the merging fraction of AGNs, few of them focused on the most luminous population of AGNs, i.e., type-1 quasars. Unlike the low-luminosity AGNs, quasars are usually much brighter than their host galaxies, which makes it very difficult to detect the disturbed features in quasar host galaxies using ground-based imaging. HST imaging can resolve some quasar host galaxies, mostly at low redshift (z \(\lesssim 1\)). However, the sample sizes of studies based on HST were small, usually containing several tens of AGNs, resulting in large statistical errors. On the other hand, using close companions as indicators of mergers is more accessible than disturbed host galaxy morphology for bright quasars, and can be expanded to larger samples. However, systematic studies on luminous quasar companions using HST are still lacking.
In this work, we measure the statistics of quasar companions and use the result to constrain the merging fraction of quasars. We use HST archival imaging for companion detection. Our master sample contains 532 quasars, which is much larger than previous studies based on HST imaging. The paper is organized as follows. §A.2 describes the selection of the quasar sample and archival images. §A.3 describes the detection of close companions, including the PSF subtraction method, as well as measurements of quasar companion fractions. The selection of the galaxy control sample and the comparison between companion fractions in quasars and normal galaxies are discussed in §A.4. §A.5 discusses the implication of the companion fraction in quasars in the context of merger-driver model of quasar triggering. §A.6 summarizes the paper. We use AB magnitude through this work, as well as a $\Lambda$CDM cosmology with $\Omega_M = 0.3$, $\Omega_A = 0.7$ and $H_0 = 70$ km s$^{-1}$.

A.2 The Quasar Sample

Our input parent quasar sample is based on the Véron Catalog of Quasars and AGN, 13th edition (Véron-Cetty & Véron, 2010, hereafter the Véron Catalog), the SDSS Data Release (DR) 7 (Schneider et al., 2010), DR12 (Pâris et al., 2017) and DR14 (Pâris et al., 2018) quasar catalogs. SDSS quasar catalogs provide $i$-band absolute magnitudes that is $K$-corrected to $z = 2$ ($M_i(z = 2)$, see Richards et al., 2006a) of quasars, and the Véron Catalog provide $B$ band absolute magnitude ($M_B$). We convert $M_B$ of quasars from the Véron Catalog to $M_i(z = 2)$ according to the relation in Richards et al. (2006a). Quasars which have $M_i(z = 2) < -23$ are selected as our parent sample. We further exclude $z < 0.3$ objects to avoid AGNs with extended emission which are confusing in PSF subtraction, and $z > 3$ objects because our control sample becomes incomplete at $z > 3$ (see §A.4.1 for details).

We use archival broad-band images of the Advanced Camera for Surveys Wide Field Camera (ACS/WFC) for quasar companion detection. The reason to use ACS/WFC images is the small PSF size and large field of view (FOV). A small PSF is crucial for detecting companions that are very close to bright quasars, and
a large FOV ensures a valid estimation for the number density of foreground and background objects. The ACS/WFC images that contain the selected quasars are fetched from the Hubble Legacy Archive (HLA)\footnote{https://hla.stsci.edu/}. These images are generated by the HLA using DrizzlePac tools (Gonzaga et al., 2012), which combine raw exposures with the same filter, same camera, and within the same visit. For each quasar, we choose the deepest image in each band to form our master image sample. In total, 595 quasars are found to appear in 806 images at this step. We run SExtractor (Bertin & Arnouts, 1996) on each image to generate a source catalog.

Though image coaddition can enhance the depth of images, we do not perform image coaddition because it will introduce difficulty to the background object number density estimation. In most cases, the overlapping area of different images containing the same quasar is a small part of the original images. The number density of background/foreground objects is estimated based on the number of objects in the whole image (see §A.3 for details). Most co-added images do not have enough area to perform a reliable background object density estimation.

We measure the statistics of quasar companions by counting all the projected companions and subtracting a background object density. We thus exclude all quasars that are strongly lensed, because the lensed images can be confusing when counting companions. We further exclude all images that satisfy any of the following:

1. Images with NCOMBINE = 1, where NCOMBINE is the “NCOMBINE” parameter in the header of the HST image, representing the number of images used for cosmic ray rejection when combining raw exposures. Images with NCOMBINE = 1 are severely polluted by cosmic rays and are not suitable for companion counting.

2. Images where the target quasar is located close (less than 2") to the edge of the CCD.

3. Images that are very crowded. The statistical errors on background object
Table A.1. The number of images excluded by the selection criteria.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Number of Images</th>
<th>Number of Quasars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>806</td>
<td>595</td>
</tr>
<tr>
<td>Lensed</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>NCOMBINE = 1</td>
<td>32</td>
<td>30</td>
</tr>
<tr>
<td>Close the Edge</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>Crowded</td>
<td>63</td>
<td>45</td>
</tr>
<tr>
<td>All Bad Images</td>
<td>126</td>
<td>98</td>
</tr>
<tr>
<td>Good Images</td>
<td>687</td>
<td>532</td>
</tr>
</tbody>
</table>

Note. — There are overlaps between different subsets of images / quasars. For example, one quasar may appear in both bad images and good images. As a result, the total number of quasars does not equal to the number of quasars in “good” images plus those in “bad” images.

density are high. Images which have more than 5000 objects that were brighter than 25 mag in the observed band are excluded in the further analysis.

We summarize the number of images excluded by each criterion in Table A.1. All the images that remain after the selection are referred to as “good” images in rest of the paper. The final master sample contains 532 quasars in 687 good images. Among the 532 quasars, 402 of them were observed in programs that were not related to AGN studies. The fact that most quasars were observed by chance ensures a small selection effect (see §A.5.1 for further discussion). In the master sample, one quasar might show up in multiple bands, but there will only be one image of the quasar given a certain filter. Figure A.1 shows the redshift and luminosity distribution of the quasars, and Figure A.2 shows the number of quasars observed in each band.
Figure A.1: The redshift and luminosity distributions of the quasars in our master sample. This sample contains 532 quasars which show up in 687 HST ACS/WFC archival images.

Images in the F814W band dominate the sample.

When studying quasar companions of certain magnitude, we will only analyze images that are deep enough to detect the faintest companions of interest at a 5σ level. Previous studies (e.g., Matsuoka et al., 2014) showed that quasar host galaxies have a typical stellar mass range $M_* \gtrsim 10^{10} M_\odot$ and a typical luminosity $M_g \lesssim -21$ in SDSS $g$-band. We are mainly interested in companions that have stellar masses close to the quasar host galaxy, which correspond to major mergers. We use two sets of magnitude limits:

1. A “simple” cut on absolute magnitude in each band. Companions with $M_{\text{abs}} < -19$ will be analyzed, regardless of in which band the quasar was observed. This sample is constructed as a “pure” observational result that can be directly compared with simulations without any further assumption. We convert the absolute magnitude $M_{\text{abs}} = -19$ to an apparent magnitude
Figure A.2: Number of quasars observed in each band. Note that one quasar can be observed by multiple bands.

$m(M_{\text{abs}} = -19)$ assuming that the companions have the same redshift as the quasar, and require that the image is deep enough to detect an object as faint as $m(M_{\text{abs}} = -19)$. We do not perform a $K$-correction when converting $M_{\text{abs}}$ to $m(M_{\text{abs}})$, because most quasars were observed in only one band. If a quasar is observed in multiple bands, the deepest image relative to $M_{\text{abs}} = -19$ is used.

Specifically, the depth of an image relative to $M_{\text{abs}} = -19$ (regardless of the filter) is quantified by $m_{\text{lim}} - m(M_{\text{abs}})$, where $m_{\text{lim}}$ stands for the 5σ depth of the image. The quasar sample corresponding to this magnitude cut is referred to as “sample A.” The analysis of this sample will be described in §A.3.

2. A cut based on stellar masses of companions. We convert stellar masses ($M_*$) to observed magnitudes (denoted by $m(M_*)$) using abundance matching (see §A.4.1 for details). This sample is constructed to enable physical interpretations of our result. Companions brighter than $m(10^9M_\odot)$ will be analyzed. Similar to Sample A, if a quasar is observed in multiple bands, the deepest image relative to $m(10^9M_\odot)$ will be used. The quasar sample corresponding to this magnitude cut is referred to as “sample B.” The analysis of this sample will be described in §A.4.
Table A.2. Summary of samples used in this study.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Description</th>
<th>Number of Quasars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master</td>
<td>The master sample</td>
<td>532</td>
</tr>
<tr>
<td>A</td>
<td>The image is deep enough to detect an object of absolute magnitude $M_{\text{abs}} = -19$ at the quasar’s redshift.</td>
<td>230</td>
</tr>
<tr>
<td>B</td>
<td>The image is deep enough to detect an object of apparent magnitude $m = m(10^9M_\odot)$.(^2)</td>
<td>354</td>
</tr>
</tbody>
</table>

\(^1\)All magnitude limits are for point sources at $5\sigma$ level.

\(^2\)The expression $m(M_*)$ refers to the apparent magnitude corresponding to stellar mass $M_*$. See §A.4.1 for the detailed definition.

The properties of the two quasar samples, together with the master sample, are summarized in Table A.2. Unlike in the master sample where a quasar may show up in multiple images, in sample A and B, one quasar only shows up in one image.

A.3 Detecting Close Companions of Quasars by PSF Subtraction

In this section, we will describe our PSF subtraction and companion detection method. We will also discuss the statistics of quasar companions with absolute magnitude $M_{\text{abs}} < -19$, regardless of the band in which the quasar image was taken. Accordingly, we use Sample A throughout this section.
A.3.1 Method

We use close companions of quasars to trace the merging history of quasar host galaxies. If the AGN activity appears in a certain stage of a galaxy merger, the number of companions around quasars will be different from that of inactive galaxies. Specifically, if two merging galaxies are still distinguishable when the quasar appears, we will see a pair of galaxies in the field; if the quasar emerges when the two progenitor galaxies have already merged into one galaxy, we will see one single quasar host galaxy rather than a pair. We will discuss this point in §A.5.3 in more details.

We first perform PSF subtraction to suppress the influence of quasar light on detecting their companions. The PSF model is generated by TinyTim (Krist et al., 2011). TinyTim takes the observation time, the position on the CCD chips and the spectrum shape of the source as input parameters. We use a power-law spectrum with a power-law index of $-2 (F_\lambda \propto \lambda^{-2})$ as the input spectrum to mimic the typical SED of a quasar. We use a modeled PSF rather than an empirical PSF because most quasars in our sample come from programs in which the quasars are not the primary science targets, and no separate PSF star observations are available. In addition, the wings of bright quasars can make it difficult to detect projected companions that are several arcseconds away from these quasars, thus a large PSF image is preferred. A modeled PSF can be as large as 20", which is difficult for an empirical PSF. The size of the PSF models used in this study is $20'' \times 20''$.

We use GALFIT (Peng et al., 2002) to perform PSF subtraction. Quasar images are fitted by a PSF component plus a Sérsic profile for the host galaxy. Examples of PSF subtraction can be found in Appendix A.6. We run SExtractor on the PSF-subtracted images and select all objects with a projected distance to the quasar of $10 \text{kpc} < d < 100 \text{kpc}$ as projected quasar companions. We divide the projected distance range $10 \text{kpc} < d < 100 \text{kpc}$ into 9 bins, with the $i$th bin at $10 \times i \text{kpc} < d < 10 \times (i+1) \text{kpc}$. The region where $d < 10 \text{kpc}$ is severely influenced by the PSF-subtraction residuals for many quasars and is not analyzed.
Companions selected in this way will be inevitably contaminated by foreground and background objects. In the following text, we use “projected companions” to represent all the companions detected (both physical and unphysical). To estimate the numbers of physical companions, a “background density” representing the background/foreground object surface density is calculated for each quasar. We use the entire image to estimate the background object density. Specifically, the surface number density of physical companions of the $i$th bin is estimated by

$$\sigma_{\text{phys}}(d_i, d_{i+1}) = \frac{N(d_i, d_{i+1})}{A(d_i, d_{i+1})C(d_i, d_{i+1})} - \frac{N_{\text{bkg}}}{A_{\text{bkg}}}$$  \hspace{1cm} (A.1)$$

where $N(d_i, d_{i+1})$ stands for the number of projected companions with $10 \times i \ \text{kpc} < d < 10 \times (i + 1) \ \text{kpc}$, $A(d_i, d_{i+1})$ is the corresponding area, and $C(d_i, d_{i+1})$ is the completeness of companion detection (see §A.3.2 for details). $N_{\text{bkg}}$ and $A_{\text{bkg}}$ describes the numbers of objects and the area of the whole image. The number of physical companions with $10 \times i \ \text{kpc} < d < 10 \times j \ \text{kpc} \ (1 \leq i < j \leq 9)$ is calculated by summing up the number of physical companions in the corresponding bins:

$$N_{\text{phys}}(d_i, d_j) = \sum_{i \leq k < j} \sigma_{\text{phys}}(d_k, d_{k+1})A(d_k, d_{k+1})$$  \hspace{1cm} (A.2)$$

and the error of $N_{\text{phys}}(d_i, d_j)$ is estimated assuming a Poisson distribution for $N(d_i, d_{i+1})$ and $N_{\text{bkg}}$ in Eq. A.1.

Unless specified, in the rest of the paper, “number of companions” refers for the estimated number of physical companions and “distance” means projected distance.

### A.3.2 Companion Detection Completeness

Some companions may be missed or misidentified as a result of imperfect PSF subtraction because it can be difficult to distinguish companions from the PSF-subtraction residuals. The probability of a companion to be detected is mainly influenced by the flux contrast between the companion and the PSF-subtraction residual. We find four factors that have a major influence on the completeness: (1) the flux of the quasar, (2) the accuracy of the PSF models, (3) the angular distance
from the quasar to the companion, and (4) the flux of the companion. Factors (1) and (2) vary from quasar to quasar, while factors (3) and (4) are determined by the companion itself. Accordingly, we add simulated companions with different flux and distance for each quasar image to estimate the fraction of missed companions. The simulated companions are generated to be point sources. For each quasar image, we simulate three sets of companions, with absolute magnitude $M_{\text{abs}} = -19, -20, -21$ in the corresponding band, assuming the companions to have the same redshift as the quasar. The completeness is estimated as a function of projected distance to the quasar. We generate 10 simulated companions at randomized positions in each distance bin. The simulated images are analyzed by the companion detection process described above, according to which we calculate the completeness of the companion detection for each quasar. A simulated companion is regarded as “detected” if an object is detected within 0.002 from the position of the simulated companion.

In the simulation, we do not consider false positives resulting from the PSF-subtraction residual because the probability for a PSF-subtraction residual to appear right at the position of a simulated companion is negligible. The simulation ensures that most companions of interest can be detected. To exclude false positives in the real images, we visually inspect all the images and remove suspicious detections that are likely PSF-subtraction residuals.

Figure A.3 shows the average completeness as a function of companion flux and the distance to the quasar from our simulation. The completeness is larger than 90% even for the faintest companion in the smallest distance bin (10 kpc < $d$ < 20 kpc). For companions that are more than 50 kpc away from the quasar (which corresponds to $\sim 6''$ at $z = 1.6$), the completeness does not change with the distance, which indicates that the influence of the quasar light becomes negligible at larger distances.

We test the potential influence of using point sources as simulated companions. We run another simulation where the shape of the companions are exponential disks with an effective radius of $r_e = 2$ kpc, assuming that the companion has a same redshift as the quasar. The difference between the completeness given by the two simulations is less than 1%.
A.3.3 Quasar Companion Fraction

Here we examine the statistics of quasar companions with $-20 < M_{\text{abs}} < -19$, $-21 < M_{\text{abs}} < -20$ and $M_{\text{abs}} < -21$, regardless of the band in which the quasar was observed\(^2\). We do not perform $K$-corrections on the magnitudes. Figure A.4 shows the average number of companions around quasars in Sample A. The number is negative in some distance ranges due to the subtraction of background object number density. We define companions with a projected distance of $10 \text{kpc} < d < 30 \text{kpc}$ to quasars as “close companions” and calculate the average number of close companions ($N_{\text{comp}}$) around quasars (galaxies). The distance range is chosen because companions at $10 \text{kpc} < d < 30 \text{kpc}$ are likely involved in a merging event, while companions at larger distances are not necessarily associated with mergers. A similar distance range has been adopted in previous studies on the galaxy merging rates (e.g., Lambas et al., 2003; Ellison et al., 2008; de Ravel et al., 2009; Man et al., 2012). The average numbers of close physical companions are $0.26 \pm 0.05$, $0.06 \pm 0.03$, and $0.05 \pm 0.03$ for $-20 < M_{\text{abs}} < -19$, $-21 < M_{\text{abs}} < -20$ and $M_{\text{abs}} < -21$ companions, and $0.38 \pm 0.07$ for all physical companions which have $M_{\text{abs}} < -19$.

One issue of this analysis is that the “absolute magnitude” we use here is difficult to be translated to physical properties of companions, given that we do not perform $K$-corrections on these magnitudes. $K$-corrections require knowledge of the companion SED, while we usually have only one-band measurement. This issue will be addressed in §A.4, by introducing a magnitude cut which is associated with galaxy stellar masses.

On-going Merging Systems

In addition to statistical studies of large samples, detailed modeling and observations of individual cases are also crucial to the understanding of the quasar triggering

\(^2\)The information of all the quasars and quasar companions is available at (https://github.com/yuemh/qso_companion)
Figure A.3: The completeness of companion detection estimated by simulated companions. The simulated objects are generated as point sources. We estimated that the fraction of missed companions is less than 10% for all the companions of interest.

mechanism. Here we report some on-going mergers with quasar activity. Follow-up observations on these objects, such as host galaxy morphology, gas kinetics and AGN obscuration can be compared directly with simulations. We visually inspect images of quasars in the master sample, and select objects that show disrupted features like tidal tails and asymmetric host galaxies. We find 22 quasars with features of recent mergers. The images of these objects can be found in Appendix A.6. Note that these objects are included in both sample A and B.

A.4 Comparing the Companion Fraction of Quasars with a Galaxy Control Sample

To compare the companion distribution of quasar host galaxies and normal galaxies, we use the 3D-HST galaxy catalog (Brammer et al., 2012; Skelton et al., 2014; Momcheva et al., 2016) to construct a control sample of galaxies. The 3D-HST is an HST Treasury program to provide ACS and WFC3 images and grism spectroscopy over five fields: COSMOS, GOODS-north, GOODS-south, AEGIS, and UDS with a combined usable area of $\sim 0.25$ deg$^2$. In this work, we use the 3D-HST photometry catalog v4.1.5, which is the latest version. It provides ACS and WFC3/IR broad-
Figure A.4: The average companion numbers of sample A quasars. Each data point corresponds to a distance bin of $\Delta d = 10$ kpc. All the error bars represent $1\sigma$ error assuming a Poisson distribution for the number of companions (same for all the other figures).
band fluxes of galaxies, including F160W, F140W and F125W for WFC3/IR images, as well as F814W and F606W for ACS images. The catalog also contains stellar mass and photometric redshifts of galaxies, derived with methods described in Skelton et al. (2014).

Quasar host galaxies are believed to be massive ($M_\star \gtrsim 10^{10} M_\odot$, e.g. Dunlop et al., 2003; Matsuoka et al., 2014). We thus construct a “massive galaxy sample” by selecting all the galaxies that have $M_\star > 10^{10} M_\odot$ and photometry flag USE\_PHOT=1 (which means good photometry) in the 3D-HST photometry catalog. The control sample galaxies are randomly drawn from the massive galaxy sample and share the same redshift distribution as our quasars. This is done by evenly dividing the redshift range ($0.3 < z < 3$) into 9 bins and randomly selecting galaxies in each redshift bin, so that the fraction of galaxies in a specific redshift bin among all the control sample galaxies equals to the fraction of quasars in that redshift bin among all the quasars.

All projected companions of the control sample galaxies with a projected distance $10 \text{ kpc} < d < 100 \text{ kpc}$ are selected in the same way as quasars. Since the 3D-HST contains five fields with different depths, the “background object density” of a control sample galaxy is approximated by the object surface number density of the 3D-HST field where the galaxy is located. We use F814W magnitudes to make magnitude cuts when counting companions, since F814W images dominate our quasar sample (Figure A.2).

### A.4.1 Making Absolute Magnitude Cuts Based on Stellar Mass

In §A.3.3, we show the number of companions around quasars as a function of companion flux, where the flux of companions are described by absolute magnitudes without $K$-corrections. This result is not suitable for studying the redshift evolution of quasar companions or comparing quasars with control sample galaxies, since different bands are used in the analysis of the quasars (all the 8 broad bands of ACS/WFC) and the galaxies (F814W only). It is difficult to perform $K$-corrections in this work, because most of the quasars have only been observed in one band.
Therefore, we use an abundance matching technique to convert stellar mass cuts to magnitude cuts in each band, which allows us to set magnitude cuts consistently at any redshift in all the eight bands. In short, for a given stellar mass $M_\star$, we find a magnitude (denoted by $m(M_\star)$) such that the number of galaxies that are more massive than $M_\star$ (denoted by $N(M_\star)$) is the same as the number of galaxies that are brighter than $m(M_\star)$. We describe the detailed procedures below.

The analysis is based on the photometric catalog of the UDS field in the 3D-HST project. The UDS photometric catalog is used because it contains the Johnson $B,V,R$ and SDSS $i', z'$ photometry, which can be converted to $HST$ broad-band magnitudes and be compared with the quasars. The conversion is done as follows. The SDSS $i', z'$ magnitudes are converted to Johnson $I$ magnitude according to Jordi et al. (2006), then the $BVRI$ magnitudes are converted to ACS/WFC broad band magnitudes according to Sirianni et al. (2005). The conversion from $BVRI$ to F850LP was not provided in Sirianni et al. (2005), so we used SDSS $z'$ as an approximation of the F850LP magnitudes. We estimate the difference between SDSS $z'$ and ACS/WFC F850LP magnitudes using the Exposure Time Calculator for $HST$ ACS/WFC using typical galaxy templates. The difference is smaller than 0.02 magnitude. We select galaxies in the UDS photometric catalog with $\text{USE\_PHOT}=1$ to construct the galaxy sample for the abundance matching technique. Note that this galaxy sample is different from the control sample; we only use the UDS field in the five 3D-HST fields and do not set any stellar mass cut on this sample.

For a quasar at redshift $z_q$, the magnitude cut (in an arbitrary band) corresponding to a stellar mass $M_\star$ is estimated as follows. First, all objects with photometric redshift ($z_{\text{phot}}$) that satisfy $z_q - 0.1 < z_{\text{phot}} < z_q + 0.1$ in the UDS catalog are selected, constructing a “magnitude-cut-setting galaxy sample”. Typically the redshift-matched galaxy sample contains 1000 ~ 3000 galaxies. The magnitudes of the objects in the magnitude-cut-setting galaxy sample are corrected to $z_q$ according to the photometric redshifts, i.e., given a galaxy with a photometric redshift $z_{\text{phot}}$
Figure A.5: An example of converting stellar mass limits into apparent magnitude limits. In this example, we convert a stellar mass of $10^{10} M_\odot$ to F814W magnitude for $z \sim 1.5$ objects. We first select objects in the UDS catalog with photometric redshift $1.4 < z_{\text{phot}} < 1.6$, and correct their F814W magnitude according to Eq. A.3. There are 291 objects with $M_* > 10^{10} M_\odot$ (objects in the green shade), which is also the number of objects with corrected F184W magnitude brighter than $m(10^{10} M_\odot) = 24.36$ (objects in the blue shade).

and an apparent magnitude $m_{\text{raw}}$, we calculate the corrected magnitude by

$$m = m_{\text{raw}} + 5 \log \left[ \frac{D_L(z_q)}{D_L(z_{\text{phot}})} \right]$$

(A.3)

where $D_L(z)$ is the luminosity distance. We then count the number of objects with stellar masses larger than the given value $M_*$ (referred to as $N(M_*)$) in the “magnitude-cut-setting galaxy sample”. The value of $N(M_*)$ varies from several hundred (for $M_* = 10^9 M_\odot$) to about twenty (for $M_* = 10^{11} M_\odot$). Finally, we find the magnitude limit $m(M_*)$ so that there are $N(M_*)$ objects that are brighter than $m(M_*)$.

As an example, Figure A.5 illustrates the process of estimating $m(10^{10} M_\odot)$ in F814W band at $z = 1.5$. The number of objects in the green and the blue shade is equal. By definition, we can estimate the number of quasar companions with stellar masses larger than $M_*$ by by counting objects that are brighter than $m(M_*)$.

For each quasar image, we set up three magnitude limits, $m(10^9 M_\odot)$, $m(10^{10} M_\odot)$ and $m(10^{11} M_\odot)$. When comparing quasars with the galaxy control samples, we will
discuss companions with $m(10^9 M_\odot) > m > m(10^{10} M_\odot)$ (faint), $m(10^{10} M_\odot) > m > m(10^{11} M_\odot)$ (intermediate) and $m < m(10^{11} M_\odot)$ (bright). Correspondingly, we use Sample B throughout §A.4. Since most quasar host galaxies have stellar mass $M_\star > 10^{10} M_\odot$, “intermediate” companions are mainly associated with major mergers (mass ratio is close to 1), while “faint” companions are mainly related to minor mergers (mass ratio is larger than 3). The case of “bright” companions is more complicated; these systems might be associated with minor mergers where the quasar host galaxy is the less massive progenitor, or major mergers if the quasar host galaxy is as massive as $10^{11} M_\odot$. Studies on quasar host galaxies (e.g., Matsuoka et al., 2014; Yue et al., 2018) show that only a small fraction of quasar host galaxies have stellar masses larger than $10^{11} M_\odot$, thus most systems with bright companions should be related to minor mergers. We estimate the completeness of companion detection for companions with $m(10^9 M_\odot), m(10^{10} M_\odot)$ and $m(10^{11} M_\odot)$, using the same method as described in §A.3.2. The result is presented in Figure A.6, which shows that the completeness of companion detection is higher than 95% in the simulated images.

Similar magnitude limits are calculated for galaxies in the control sample using the F814W magnitude. At $z = 3$, the faintest companions that we consider in this study have magnitude $m_{\text{F814W}}(10^9 M_\odot) = 26.3$. At this magnitude, about 2% objects in the 3D-HST catalog have signal-to-noise ratios smaller than 3, and the completeness of companion counting will drop toward fainter magnitudes (and thus higher redshift). This is the reason why our sample only contains $z < 3$ objects.

### A.4.2 Comparison Between Companions around Quasars and Galaxies

We calculate the number of faint, intermediate and bright companions of quasars in Sample B. Figure A.7 shows the average number of physical companions, both for quasars and the control sample galaxies. Quasars have fewer intermediate companions at $d < 60$ kpc. At larger distances, the number of companions around quasars and galaxies are roughly the same. No significant difference can be seen for faint and bright companions between quasars and galaxies.
Figure A.6: The detected fraction of quasar companions which have a magnitude of $m(10^9 M_\odot)$, $m(10^{10} M_\odot)$ and $m(10^{11} M_\odot)$, estimated by simulated images. The detected fraction is higher than 95% for all the companions of interest.

For Sample B quasars, the average numbers of “close companions” (companions with a project distance to quasars of $10 \text{ kpc} < d < 30 \text{ kpc}$) are

$$
\overline{N}_{\text{comp,Q}}(\text{faint}) = 0.233 \pm 0.043 \\
\overline{N}_{\text{comp,Q}}(\text{intermediate}) = 0.004 \pm 0.021 \\
\overline{N}_{\text{comp,Q}}(\text{bright}) = 0.012 \pm 0.018
$$

and the comparison galaxy sample has

$$
\overline{N}_{\text{comp,G}}(\text{faint}) = 0.178 \pm 0.012 \\
\overline{N}_{\text{comp,G}}(\text{intermediate}) = 0.087 \pm 0.008 \\
\overline{N}_{\text{comp,G}}(\text{bright}) = 0.016 \pm 0.005
$$

Quasars show a $3.7\sigma$ deficit of close intermediate companions, and have numbers of faint and bright companions similar to those of normal galaxies.

Figure A.8 shows the average number of close companions of quasars and control sample galaxies as a function of redshift. The average number of companions around
Figure A.7: The average number of companions of quasars and the control sample galaxies. Each data point corresponds to a distance bin of $\Delta d = 100\text{pc}$. Small x-axis offsets are added to the error bars to make them distinguishable.

Figure A.8: The redshift evolution of average number of companions around quasars. The average number of companions of galaxies in the control sample is also included.
control sample galaxies increases toward high redshift, both for faint and bright companions. This result is expected since the universe is more crowded at high redshift. Meanwhile, the average number of companions around quasars does not significantly evolve with redshift.

Figure A.9 shows the average number of close companions as a function of quasar luminosity. More luminous quasars have more faint companions, while the number of intermediate companions decreases with quasar luminosity. Given the statistical error, both trends are not significant. The number of bright companions does not show a luminosity dependence.

We also examine the relationship between average companion number of quasars and other quasar properties, including:

1. Broad absorption line (BAL) features. The SDSS quasar catalogs contain BAL flags, while the Veron catalog does not have such information, thus our BAL and non-BAL quasar sample only contain SDSS quasars. We find that BAL quasars and non-BAL quasars have consistent numbers of companions (both faint and intermediate) at a $1\sigma$ level.

2. Radio loudness. We match our quasar catalog with the *Faint Images of the
Radio Sky at Twenty-Centimeters (FIRST; e.g., Becker et al., 1995; White et al., 1997) survey catalog. We divide the quasars that were covered by the FIRST survey into two subsamples, namely “radio loud (RL) quasars” and “radio quiet (RQ) quasars”, according to whether they were detected by the FIRST survey. We find that RL quasars tend to have more “faint” companions than RQ quasars. The average number of “faint” companions is $0.523 \pm 0.144$ around RL quasars and $0.153 \pm 0.047$ around RQ quasars, which is a $2.4\sigma$ difference. The two samples have similar numbers of intermediate companions (consistent at a $1\sigma$ level).

3. Infrared (IR) brightness. We match our quasar catalog with the Wide-field Infrared Survey Explorer (WISE; e.g., Wright et al., 2010) ALLWISE source catalog. Since WISE W3 and W4 are usually not deep enough for faint quasars, we use the WISE W1 and W2 fluxes to estimate the rest-frame 2$\mu$m magnitude ($M_{2\mu}$) assuming that the near infrared SED of quasars can be represented by a power law. We then evenly divide the quasars into two subsamples according to their $M_{2\mu} - M_i(z = 2)$ color, referred to as “IR-bright quasars” and “IR-faint quasars”, respectively. One potential problem is that the near-infrared SED of some quasars cannot be well-fitted by a single power law (e.g., Glikman et al., 2006; Hernán-Caballero et al., 2016). As a sanity check, we use the quasar spectrum template in Hernán-Caballero et al. (2016) to fit the WISE W1 and W2 fluxes of the quasars and define the two subsamples based on the template-estimated rest-frame 2$\mu$m magnitude. Among all the IR-bright (IR-faint) quasars defined using the power-law fit, 11.6% are classified as IR-faint (IR-bright) in the template-based classification. In both cases, IR-bright and IR-faint quasars have similar numbers of faint and intermediate companions (consistent at a $1.5\sigma$ level).
A.5 Discussion

A.5.1 Selection Effects

Our sample consists of quasars from the SDSS quasar catalogs and the Véron quasar catalog that have been observed by *HST* ACS/WFC broad-band imaging. The parent sample (SDSS + Véron) includes almost all quasars known to date. Selection effects may be introduced by selecting quasars that have ACS/WFC imaging. Among the 532 quasars in our sample, 402 of them (76%) were observed by ACS/WFC for purposes that were irrelevant to AGN science (e.g., photometric surveys, studies on supernovae or local galaxies). Using the subsample of quasars which were observed in AGN-unrelated programs, we estimate the average companion numbers around quasars to be $0.194 \pm 0.052$, $0.003 \pm 0.028$ and $0.005 \pm 0.023$ for faint, intermediate and bright companions. These numbers are close to the result we get using the whole sample. For the rest of the sample, the goal of the original *HST* program was related to AGN science, which could introduce complicated selection effect, e.g., for the programs that targeted a specific class of AGN that could have higher merger fraction.

Our results described in §A.4.2 are based on several assumptions, which may also introduce systematic errors. The main assumption we make is that the companion fluxes can be converted to stellar masses as described in §A.4.1. This can not be applied to individual objects. However, if quasar companions follow the same stellar mass and flux distribution as normal galaxies, our method will provide the correct numbers of companions that fall in a certain stellar mass range. This method ignores the possible influence of quasars on their companions, which is a complicated effect and is difficult to correct. When applying the results, this potential systematic error must be kept in mind. On the other hand, the “primary results” in §A.3.3 are free from this systematic error.
A.5.2 The Fraction of Merger-Triggered Quasars

Our results suggest that there is a deficit of companions around quasars compared with inactive galaxies. We interpret the difference as a result of the difference between the merging history of quasars (especially merger-triggered ones) and normal galaxies. Here we first review the merger-triggering model of quasars briefly and discuss how many companions would we expect around a merger-triggered quasar.

According to simulations (e.g., Lotz et al., 2010), the evolution of a galaxy merger will experience five stages: the first encounter, the largest separation, the second encounter, the final coalesce and the post-merger remnant. Each stage last for $\lesssim 0.5\text{Gyr}$. The typical lifetime of a quasar is $10 \sim 100\text{Myr}$ (e.g., Hopkins et al., 2008), which means that the morphological properties of quasar host galaxies will not change significantly during the life of the quasar.

Previous simulations on galaxy major mergers triggering quasars (e.g., Springel et al., 2005a; Di Matteo et al., 2005; Hopkins et al., 2006; Newton & Kay, 2013) have suggested that quasars emerge at the final coalesce stage. This can be understood in the following picture. In the merger-triggering model, the quasar activity emerges when strong gas inflows feed the SMBH. This process requires the gas content to be highly disturbed. Simulations of galaxy major-mergers show that the disturbed features are the most prominent at the second encounter stage. It takes some time for the gas to reach and feed the SMBH, thus strong quasar activities are expected to emerge at the final coalesce. As an observational evidence, Ellison et al. (2013) found that post-mergers in SDSS have a high AGN-fraction.

With some simple calculations, we derive the fraction of merger-triggered quasars using the number of close companions. Assuming that quasars are either triggered by secular evolution or mergers, and the fraction of merger-triggered quasar is $\alpha$. We use $N_{\text{comp, QS}}$ and $N_{\text{comp, QM}}$ to denote their average number of companions, where “Q” means “quasar”, “S” stands for “secular evolution” and “M” means “merger”. The observed average companion number of quasars should be $N_{\text{comp, Q}} =$
$(1 - \alpha)N_{\text{comp, QS}} + \alpha N_{\text{comp, QM}}$, from which we have

$$\alpha = \frac{N_{\text{comp, QS}} - N_{\text{comp, Q}}}{N_{\text{comp, QS}} - N_{\text{comp, QM}}} \quad (A.6)$$

We further assume that (1) secular-evolution-triggered quasars have the same number of companions as normal galaxies ($N_{\text{comp, QS}} \approx N_{\text{comp, G}}$, where “G” means inactive galaxies), since the secular evolution will not influence the merging process; and (2) merger-triggered quasars have no close companions ($N_{\text{comp, QM}} \approx 0$), since the two progenitor galaxies of the merger-triggered quasar have already merged into a single galaxy. We then have

$$\alpha = \frac{N_{\text{comp, G}} - N_{\text{comp, Q}}}{N_{\text{comp, G}}} \quad (A.7)$$

Figure A.7 indicates that the difference between the average number of companions of quasars and galaxies varies with the companion stellar mass. As a result, $\alpha$ depends on the companion mass. This result is not surprising since we expect that major mergers and minor mergers have different probabilities to trigger quasars. We can interpret $\alpha(M_{\text{companion}})$ as the fraction of quasars that are triggered by a merger where one progenitor galaxy had a stellar mass of $M_{\text{companion}}$. Equation A.4 and A.5 describe the average number of companions of our sample. According to Equation A.7, we have $\alpha(\text{faint}) = -0.31 \pm 0.26$, $\alpha(\text{intermediate}) = 0.95 \pm 0.25$ and $\alpha(\text{bright}) = 0.23 \pm 1.14$. Since intermediate and faint companions are associated with major and minor mergers respectively, our toy model indicates that there should be a significant fraction of quasars that are triggered by major mergers, and that minor mergers have a small contribution in quasar triggering. The large error for the bright companions does not allow for a strong conclusion.

We now discuss the impact of our assumptions in the analysis above. The first assumption is that secular-evolution-triggered quasars have the same number of companions as normal galaxies. Close neighbors of galaxies can disturb their gas kinematics and lead to gas inflows. However, gas inflows generated in this way are usually not strong enough to feed a quasar, thus the environmental influence should be minor. Previous studies have also suggested that secular evolution is
only responsible for faint AGNs (e.g., Treister et al., 2012). Moreover, according to Equation A.6, a larger number of close companions around secular-evolution-triggered quasars will lead to a higher fraction of merger-triggered quasars. The second assumption is that there are no companions around merger-triggered quasars. This assumption is clearly over-simplified. Our main idea is that, in the picture of a merger-triggered quasar discussed above, the quasar host galaxy is a galaxy merger that has entered the final-coalesce stage, and we can only see one galaxy rather than a pair. Two possible errors come into this assumption. On the one hand, a merger-triggered quasar may not show up exactly in the final-coalesce stage. It is possible that the quasar is triggered earlier when the two merging galaxies are still distinguishable. We suggest that the influence of this possible error should be small, however, because we only count companions with distance larger than 10 kpc, and most observed merging galaxy pairs are closer. On the other hand, if there are nearby galaxies that are not involved in this merging event, we will have a non-zero companion number for merger-triggered quasars. Given the distance cut we applied (10 kpc < d < 30 kpc), such cases should be rare. In both cases, Equation A.6 suggests that, if $N_{\text{comp}, QM} > 0$, the fraction of merger-triggered quasars will become even larger, which will further strengthen our conclusion.

A.5.3 A Unified Picture of AGN Triggering

Conclusions from previous studies on AGN merger fractions have been ambiguous. There are both results indicating enhanced merging fractions of AGN (e.g., Ellison et al., 2011; Silverman et al., 2011; Satyapal et al., 2014; Fan et al., 2016; Weston et al., 2017; Goulding et al., 2018) and indicating no difference between AGN and normal galaxies (e.g., Cisternas et al., 2011; Schawinski et al., 2011; Kocevski et al., 2012; Villforth et al., 2017). Our result suggests that there should be a significant fraction of major-merger-triggered quasars. We consider several possible reasons for this ambiguity.

Firstly, most of these studies used distinct AGN samples. The AGN samples vary from emission-line-ratio selected (e.g., Ellison et al., 2011), near-IR selected (e.g.,
Satyapal et al., 2014; Fan et al., 2016; Weston et al., 2017; Goulding et al., 2018), X-ray selected (e.g., Cisternas et al., 2011; Silverman et al., 2011; Kocevski et al., 2012; Villforth et al., 2017), to optical selected (this work). We notice that all the studies using near-IR selected AGN samples reach a conclusion that AGN have an enhanced merging fraction, while most of the X-ray selected AGN samples do not show a significant difference from inactive galaxies. This result can be explained if different populations of AGN emerge in different stages of galaxy mergers. The luminosities of the AGN samples are also different, which is believed to have a crucial influence on AGN merging fraction. Previous observations using optical data focused mainly on low-luminosity AGNs to avoid the strong emission from the central nuclei. Most studies on high-luminosity AGNs are based on either X-ray or infrared data (for obscured ones). Comparing to these studies, our sample consists of optical-selected AGNs and spans a wide range of luminosity ($-31 < M_i(z = 2) < -23$).

Secondly, the method of these studies might introduce some biases. Most of the previous studies used disturbed features in quasar host galaxies as indicators of recent merger events. There have been arguments that these features should be able to survive until the quasar activity emerges (e.g., Cisternas et al., 2011; Villforth et al., 2017). However, this depends on the model of galaxy mergers and AGN triggering, which is highly uncertain and varies from object to object. Even if the average timescale of disturbed features may be long enough, it is still possible to miss some mergers and thus underestimate the merger fraction of AGN. The uncertainty in identifying disturbed features in AGN host galaxies might be more severe for bright unobscured AGNs that need PSF subtraction, given the difficulty of PSF modeling.

It is also difficult to distinguish major mergers and minor mergers based on the host galaxy morphology. If minor mergers do not contribute to the triggering of quasars (as indicated by our result), including them as “recent merging systems” will increase the number of identified merging systems and introduce extra uncertainties when testing the major-merger-triggering mechanism.

In comparison, we identify mergers by counting companions. This will not intro-
duce significant biases against the late-stage mergers, where the disturbed features in the galaxies might have already faded away. It is also more straightforward to distinguish major and minor mergers since we can estimate the companion mass. Counting companions is an accessible way for nearly all populations of AGN, and does not require superb angular resolution, which makes it easier to build a large, unbiased sample.

Keeping the possible biases in mind, we find that most of the results are consistent with the picture where:

(1) Mergers are the main triggering mechanism for high-luminosity AGNs, and secular evolution is mainly responsible for low-luminosity AGNs. There have been simulations claiming that secular evolution is not powerful enough to trigger the most luminous quasars (e.g., Treister et al., 2012). Previous observations have also reported that the merger fraction increases with AGN luminosity (e.g., Fan et al., 2016). We find that the average number of intermediate companions decreases with quasar luminosity. According to Figure A.9 and Equation A.7, our result indicates that luminous quasars are more likely to be triggered by major mergers.

(2) Merger-triggered AGNs evolve from an obscured to an unobscured phase. The transition happens when the radiation and material outflows blow away the dust around the active nucleus. Consequently, IR-selected AGNs (which are dustier) might represent an earlier stage of AGN evolution, and it is easier to detect the disturbed features in their host galaxies than other populations of AGN. This picture is supported by simulations (e.g., Di Matteo et al., 2005; Hopkins et al., 2008), and can explain the discrepancy with previous observations, which reported that IR-selected AGNs have a larger merging fraction than X-ray-selected AGNs.

**Merger Rate of AGN vs AGN Rate in Mergers**

In comparison to works such as ours, which measures the fraction of mergers in quasars, some studies instead compared the fraction of AGN in merging and non-
merging systems, quantified by the “AGN fraction ratio”:

\[
R = \frac{P(\text{AGN|merger})}{P(\text{AGN|non-merger})}
\]  

(A.8)

where \(P(\text{AGN|merger})\) \((P(\text{AGN|non-merger}))\) is the probability of finding an AGN in a (non-)merging system. Previous studies found that merging systems are more likely to host AGN (e.g., \(R \approx 2-7\) in Goulding et al. (2018), \(R \approx 5-17\) in Weston et al. (2017)), and concluded that galaxy mergers are the dominant triggering mechanism of the AGN in their sample. Using Bayesian analysis, we can calculate the AGN fraction ratio using the merger ratio in AGN, and thus compare our result directly with previous studies. According to Bayes’ Theorem,

\[
P(\text{AGN|merger}) = \frac{P(\text{merger|AGN})P(\text{AGN})}{P(\text{merger})}
\]

\[
P(\text{AGN|non-merger}) = \frac{P(\text{non-merger|AGN})P(\text{AGN})}{P(\text{non-merger})}
\]  

(A.9)

thus

\[
R = \frac{P(\text{non-merger})}{P(\text{merger})} \times \frac{P(\text{merger|AGN})}{P(\text{non-merger|AGN})}
\]

(A.10)

where we apply \(P(\text{merger}) + P(\text{non-merger}) = 1\).

In our sample, the estimated major-merger-triggered quasar fraction is \(\alpha = 0.95 \pm 0.25\), with a 3\(\sigma\) lower limit of 0.22. We assume that the secular-evolution-triggered quasars have the same merger fraction as inactive galaxies (i.e., \(P_S(\text{merger|AGN}) = P(\text{merger})\)). We also have \(P_M(\text{merger|AGN}) = 1\) by definition. We adopt the merger fraction of inactive galaxies from Villforth et al. (2017), which gave \(P(\text{merger}) \approx 0.2\). This gives

\[
P(\text{merger|AGN}) = \alpha P_M(\text{merger|AGN}) + (1 - \alpha) P_S(\text{merger|AGN}) \geq 0.37,
\]  

(A.11)

and \(R \geq 2.4\) according to Equation A.10, which is consistent with previous studies.
A.6 Summary

We investigate the numbers of companions around quasars which have $M_i(z = 2) < -23$ at $0.3 < z < 3$. Based on the SDSS quasar catalogs and the Véron quasar catalog, we construct a sample of 532 quasars which have been observed by HST ACS/WFC, and use the archival images to find all the companions around these quasars with projected distance of $10 \text{kpc} < d < 100 \text{kpc}$. PSF subtraction is done for all the quasars to enhance the detectability of close companions, and the fraction of missed companions was estimated to be less than 10% even for the faintest companion at the smallest projected distance of interest ($\S$A.3.2 and $\S$A.4.1). We use galaxies in the 3D-HST photometric catalog to construct a redshift-matched sample of massive inactive galaxies as our control sample. We define “faint”, “intermediate” and “bright” companions such that faint and bright companions are associated with minor mergers, and intermediate companions correspond to major mergers. We calculate the average number of companions of quasars and inactive galaxies, and raise an explanation to the difference between quasars and normal galaxies. Our main conclusions are:

1. Both quasars and inactive galaxies show excesses of companion surface densities in their neighborhoods. Quasars show a deficit of intermediate companions at projected distance $d \lesssim 60 \text{kpc}$, and have numbers of faint/bright companions similar to normal galaxies. The average number of companions around quasars show little evolution with redshift, and do not show significant dependence on absorption features, radio loudness and IR luminosity. More luminous quasars have more faint companions and fewer intermediate companions, though both trends are not significant. The number of bright companions does not evolve with quasar luminosity.

2. By assuming that merger-triggered quasars have no close companions and secular-evolution-triggered quasars have the same number of companions as inactive galaxies, the deficit of close companions around quasars indicates that a significant fraction of quasars are triggered by major mergers.
3. Most of the previous studies are consistent with the picture where the merger-triggered fraction increases with AGN luminosity, and merger-triggered AGN evolves from an obscured to an unobscured phase. The ambiguity of previous results may be a result of biases introduced by the samples and the methods.

Using close companions as identifiers of merging systems does not require superb angular resolution, which makes it possible to constrain the AGN merger fraction using ground-based imaging. Ground-based surveys like the Hyper Suprime-Cam Survey (Aihara et al., 2018a) and the Large Synoptic Survey Telescope (LSST Dark Energy Science Collaboration, 2012) have angular resolution that is good enough for companion counting, and can provide a large sample to decrease the statistical error. Future studies utilizing these surveys will provide a more accurate estimate on the fraction of merger-triggered AGN, and put better constraints on AGN triggering models.

Appendix: On-Going Merging Systems

In this appendix, we present the images and information of candidates of on-going merging systems with quasar activity mentioned in §A.3.3. We visually inspected all the quasars in the master sample and select systems that show either a pair of interacting galaxies or some disturbed features like tidal tails. Figure A.10 shows the images of the on-going merging systems. All the images are 100 kpc × 100 kpc in size. The information of these objects can be found in Table A.3.
Table A.3. Information of on-going merging systems.

<table>
<thead>
<tr>
<th>Quasar Name</th>
<th>RA</th>
<th>DEC</th>
<th>Redshift</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDSS J005009.81-003900.6</td>
<td>00:50:09.81</td>
<td>-00:39:00.6</td>
<td>0.728</td>
<td>Interacting Galaxies</td>
</tr>
<tr>
<td>SDSS J005916.10+153816.1</td>
<td>00:59:16.10</td>
<td>+15:38:16.1</td>
<td>0.354</td>
<td>Interacting Galaxies</td>
</tr>
<tr>
<td>SDSS J020258.94-002807.5</td>
<td>02:02:58.94</td>
<td>-00:28:07.5</td>
<td>0.339</td>
<td>Tidal Tail</td>
</tr>
<tr>
<td>SDSS J080908.13+461925.6</td>
<td>08:09:08.13</td>
<td>+46:19:25.6</td>
<td>0.657</td>
<td>Tidal Tail</td>
</tr>
<tr>
<td>SDSS J110556.18+031243.1</td>
<td>11:05:56.18</td>
<td>+03:12:43.1</td>
<td>0.353</td>
<td>Interacting Galaxies</td>
</tr>
</tbody>
</table>

Note. — Table 2 is published in its entirety in the electronic edition of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content. The full table is also available in FITS format at https://github.com/yuemh/qso_companion.
Figure A.10: Images of on-going merging systems in our master sample. The image sizes are $100\text{kpc} \times 100\text{kpc}$ at the redshift of the object. We present both the original image (left) and the PSF-subtracted image (right).
Figure 10 (Continued)
Figure 10 (Continued)
REFERENCES


Edge, A., Sutherland, W., Kuijken, K., et al. 2013, The Messenger, 154, 32


Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, Nature, 279, 381, doi: 10.1038/279381a0


—. 2022a, AJ, 163, 139, doi: 10.3847/1538-3881/ac4cb0


