Quasar Variability as seen by Large Optical Sky Surveys

Krzysztof Suberlak

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Željko Ivezić, Chair

Scott Anderson

Andy Connolly

Mario Jurić

Program Authorized to Offer Degree:
Astronomy

University of Washington

Abstract

Quasar Variability
as seen by
Large Optical Sky Surveys

Krzysztof Suberlak

Chair of the Supervisory Committee:
Professor Željko Ivezić
Astronomy Department

Quasars are powered by accretion of hot gas onto a growing, supermassive black hole, at cosmological distances. Such an active galactic nucleus (AGN) in the heart of distant galaxies is a source of intense radiation spanning the electromagnetic spectrum. Light passing through intervening intergalactic medium allows studies of He II reionization history and Damped Lyman α Absorbers. Counting quasars across cosmic time and luminosity (Quasar Luminosity Function) help relate their evolution to galactic build-up of stellar mass. AGN are also laboratories for high-energy astrophysics studying accretion phenomena, and general relativity - only this year we witnessed the first direct observation of the relativistic shadow cast by the supermassive black hole of M87.

Characteristics of quasar stochastic variability, such as observed timescales, and variability amplitude, are relevant to the physical properties of quasars: the black hole mass, bolometric luminosity, or Eddington ratio. In the era of large optical sky surveys (such as SDSS, CRTS, PTF, ZTF, PS1, LSST), an increasing availability of AGN time series makes it possible to combine these diverse sources of measurement.

SDSS survey data broadly supported variability being consistent with the Damped Random Walk model. Newer CRTS data seemed to indicate an enhanced variability on monthly

timescales, but in this dissertation I show that this was due to underestimated photometric uncertainties of CRTS.

Having shown that DRW is an appropriate description of quasar variability for timescales ranging from months to years, with the currently available data I proceed to combine the SDSS and PS1 data. I provide better constraints on the underlying DRW parameters and update the coefficients correlating the variability parameters to quasar physical properties. I also make prediction about the improvement that will be afforded with the arrival of the LSST.

Compared to its predecessors, LSST will have a larger etendue (300 m^2 deg^2) and will survey the sky to a new depth. For example, the SDSS etendue was only 5.9 m^2 deg^2. At the single-visit depth of 24.5 mag (r-band), crowdedness will become an issue even at moderate galactic latitudes. In preparation for LSST, I address the source of systematics in photometry resulting from the use of the LSST science pipelines, by reprocessing the DECAPS data. I show that with moderate software improvement it will be possible to fulfill the design requirements, and process the 18 000 deg^2 of the night sky.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I thank God, maker of Heaven and Earth, creator of the Universe, for making all things: visible and invisible. In His love He brought me to the fullness of faith in the Holy Catholic Church, which sustains me through the Holy Sacraments. He made it possible for me to come to the United States, and join the Blessed Sacrament parish, where I met my wonderful wife. I thank my parents, my family, my friends, my mentors, and all people of good will, for their continued support. I dedicate this work to my dear wife Michael Shelby, our daughter Stellamaris Danuta, and our little child that is still growing, snug in the womb.

Creator alme siderum, | Æterna lux credentium | Jesu Redemptor omnium, | Intende votis supplicum. Qui daemonis ne fraudibus | Periret orbis, impetu | Amoris actus, languidi | Mundi medela factus es. Commune qui mundi nefas | Ut expiares; ad crucem | E Virginis sacrario | Intacta prodis victima. Cujus potestas gloriæ,| Nomenque cum primum sonat, | Et cœlites, et inferi | Tremente curvantur genu. Te deprecamur ultimæ| Magnum diei Judicem; | Armis supemægratiæ| Defende nos ab hostibus. Virtus, honor, laus, gloria,| Deo Patri cum Filio, | Sancto simul Paracleto, | In sæculorum sæcula.

Ad maiorem Dei gloriam. Amen.

Chapter 1

INTRODUCTION

Quasars are some of the most intense sources of ionizing radiation in the Universe. Each quasar is a distant galaxy, harboring an actively accreting supermassive black hole - an active galactic nucleus (AGN) - at its center (Padovani et al., 2017). The majority of galaxies have gone through the quasar stage while growing a supermassive black hole at its center (Harrison, 2017; Heckman & Best, 2014; Kormendy & Ho, 2013). The peak of quasar activity is approximately at the redshift 2 (see eg. the distribution of SDSS eBOSS DR14 quasars in a catalog of Pâris et al. 2018, and work on the evolution of the quasar luminosity function by Kulkarni et al. 2019). Although the exact nature of the co-evolution of the supermassive black hole and the host galaxy is not yet fully understood (see recent observational work on AGN life cycle by Bianchini et al. 2019; Eilers et al. 2018; Sartori et al. 2018b; Schawinski et al. 2015; Schmidt et al. 2018; Yang et al. 2019; Yang et al. 2018), the majority of studies agree that the supermassive black hole grows via episodes of vigorous accretion, responding to the changes in the fuel supply (see theoretical work by Sartori et al. 2018a; Weigel et al. 2017, simulations of Jones et al. 2019; Ricarte et al. 2019, and recent summaries by Harrison 2017; Husemann & Harrison 2018; Padovani et al. 2017; Schawinski 2012). Apart from being interesting in their own right, quasars are also important for diverse branches of astrophysics. Distribution of quasars across the cosmic time - quasar luminosity function - is used to study quasar activity, and the evolution of the galactic initial mass function (AlSayyad, 2016; McGreer et al., 2013). Quasar light passing through structures at a range of redshifts allows studies of the intervening intergalactic medium (Hennawi & Prochaska, 2007; Prochaska et al., 2014), the He II reionization history (Khrykin et al., 2017), or Damped Lyman α Absorbers (Murphy & Bernet, 2016; Parks et al., 2018; Wolfe et al., 2005). Therefore, it

is crucial to increase the currently known sample of quasars, and characterize better their physical properties.

In this introduction we provide motivation for studying quasars in the time domain. We start with an overview of quasar properties as variable sources. We then proceed to briefly describe the content of each chapter that forms this dissertation.

1.1 Quasars as Variable Sources

According to the standard picture, each quasar contains a supermassive black hole (10^{6-10} $M_{\odot}$) surrounded by a geometrically thin, optically thick accretion disk (with many possible departures in shape and form of the disk and flow, eg. Advection-Dominated Accretion Flow, see Czerny et al. 2019; Ruan et al. 2019). Apart from winds of energetic particles (King, 2014; King & Pounds, 2015), energy is released via radiation, and jets perpendicular to the accretion disk, with the latter probably powered by a mechanism relying on magnetic fields (Begelman et al., 1984; Begelman & Silk, 2017; Blandford & Znajek, 1977).

Quasars are variable in the optical wavelengths at the root-mean-squared level of 0.2 mag (20%) (Kozłowski et al., 2010; Schmidt et al., 2010). A mathematical description of their variability by the Ornstein-Uhlenbeck process (or Damped Random Walk, DRW), has proven to be a robust approximation enabling to harness the inherent stochasticity into a space of two parameters - the characteristic timescale τ, and the asymptotic variability amplitude σ (Kozłowski et al., 2010; MacLeod et al., 2010). DRW is a stochastic process with the power exponential covariance matrix $k(\Delta t) = \sigma^2 \exp{(\Delta t/\tau)^{\beta}}$, with $\beta \approx 1$. Given the available quality of data, studies hitherto conducted (eg. Zu et al. 2013 with OGLE, and MacLeod et al. 2012 with SDSS) did not find a significant departure from $\beta = 1$ on timescales from months to years. Short, but well-sampled Kepler (Borucki et al., 2010) light curves were used to probe the high-frequency part of the light curve power spectral distribution (PSD), and work by Aranzana et al. (2018); Edelson et al. (2014); Mushotzky et al. (2011); Smith et al. (2018) seems to indicate that perhaps at short timescales (below ∼few days) there is a contribution from a stochastic process with β steeper than that of the pure DRW. In

this thesis however, we focus on ground-based optical surveys (Sloan Digital Sky Survey - SDSS, York et al. 2000, Catalina Real-time Transient Survey - CRTS, Djorgovski et al. 2011, PanSTARRS1 - PS1, Chambers et al. 2016, Palomar Transient Factory - PTF, Rau et al. 2009, Zwicky Transient Facility - ZTF, Bellm et al. 2019) that explore timescales longer than several days, and we assume that for these data the DRW is the best available description of quasar variability.

The process of accretion results in dynamical changes happening on a diverse hierarchy of timescales. A juxtaposition of theoretical and observed timescales is often employed to link the observed variability with the underlying physical processes. Timescales most often considered are light-crossing, orbital, thermal, viscous, and front propagation timescale (Stern et al., 2018) - they provide the first-order insight into the physics involved in more complex hydrodynamical modeling (Kokubo, 2015; Kubota & Done, 2018).

The shortest is the light-travel timescale, which is of the order of tens to several-hundreds of days for a typical supermassive black hole. This timescale plays a prominent role in a simple lamppost model, which involves a strong source of radiation close to the black hole, emanating light that within light-travel time is reverberated by structures of the broad line region (Cackett et al., 2007). These assumptions are part of the reverberation mapping (RM) - one of the most direct techniques (apart from microlensing Mosquera et al. 2013, and megamaser measurements Läsker et al. 2016) that allow us to infer a small size of the accretion disk (Cackett et al., 2018; Homayouni et al., 2019; Jiang et al., 2017; Mudd et al., 2018), and in turn - black hole masses (Bentz & Katz, 2015). While the complex shape of the broad line region (recently directly observed with interferometry for a 3C 273 by Gravity Collaboration et al. 2018) may cause complications for the details of the anomalies observed when performing the RM inference (Gaskell et al., 2019), the variations on light-crossing timescale are not a likely origin of the observed quasar optical variability.

The dynamical timescale is the inverse of the angular frequency for a gas on an approximately Keplerian orbit around the supermassive black hole (eg. for a 4×10^8 $M_\odot$ black hole, at 150 gravitational radii, the orbital timescale is approximately 10 days, see Stern

et al. 2018). However, besides this timescale being too short, quasar variability is generally aperiodic (stochastic). Indeed, any sign of periodicity convolved with the stochastic signal could indicate a presence of a binary supermassive black hole, as shown by recent searches (Bartos et al., 2017; Charisi et al., 2016; Charisi et al., 2015; Foord et al., 2017; Graham et al., 2015a,b; Liu et al., 2015, 2016; Zheng et al., 2016).

The thermal timescale is related to heating and cooling processes, and can be understood as the time needed for a small part of the disk to re-adjust to the thermal equilibrium (see derivation in Frank et al. 2002). Such 'hot spots' may be common in the disk - Ruan et al. (2014) showed that to explain the observed variability, and the bluer-when-brighter trend, the accretion disk needs to be strongly inhomogeneous, with large temperature fluctuations. Furthermore, Ruan et al. (2014) found that variations in accretion rate alone could not fully explain the variability on the months-to-year timescale. Although the details of propagation of thermal inhomogeneities are not fully understood (Dexter & Agol, 2011; Kokubo, 2015), the stochastic Ornstein-Uhlenbeck process describes well the response of a system to the input noise process, and fits well the morphology and the PSD of the observed quasar light curves (Kelly et al., 2011). On the order of 1 year for 4×10^8 $M_\odot$ black hole, thermal timescale is thus believed to be directly linked to the optical quasar variability (Kelly et al., 2009; Ruan et al., 2014; Zuo et al., 2012).

Two longest timescales are the front propagation timescale, and the viscous timescale. The thermal fronts propagate through the disk at the velocity proportional to the local sound speed and viscosity (Hameury et al., 2009). For the 4×10^8 $M_\odot$ black hole, the characteristic timescale of front propagation is $\sim$ 20 years (Stern et al., 2018). The viscous timescale is related to the mass flow across the disk, and in the same scenario is $\sim$ 400 years (Stern et al., 2018). It could be contributing variability at centuries-long timescales, and cannot be excluded with the available data. However, since both viscous and front propagation timescales are much longer than the observed timescales, these processes do not explain the currently observed variability with month-to-year timescales.

Some quasars also exhibit long-term trends appearing to be superimposed on the DRW

variability, reflecting the change in the strength of the continuum, and the broad emission line flux (MacLeod et al., 2019). These objects, termed 'Changing-look Quasars' (CLQs), could be undergoing variations in global accretion rate, but other explanations, such as obscurations, tidal disruption events, or microlensing have also been proposed (see observational work by Gezari et al. 2017; LaMassa et al. 2015; Ruan et al. 2016; Runnoe et al. 2016; Stern et al. 2018; Yang et al. 2018, and forward-modeling of AGN activity by Caplar et al. 2015, 2018; Sartori et al. 2018a, 2019; Schawinski et al. 2015). It appears that the most plausible scenario is that CLQs are the tail-end members of a probability distribution function describing quasar variability (MacLeod et al., 2012; Rumbaugh et al., 2018).

The structure of any light curve can be analyzed in terms of its self-similarity, or autocorrelation - the distribution of differences in brightness Δm as a function of time separation Δt. The rms scatter of Δm is the Structure Function (SF). SF can be characterized by the e-folding timescale τ, and the asymptotic amplitude SF_∞. In cases where there is insufficient temporal coverage for individual objects, ensemble SF is preferred over individual light curve modeling (since for 2-5 epochs the model would be unconstrained). Indeed, MacLeod et al. (2012) successfully used SDSS coupled with Palomar Observatory Sky Survey data to study quasar variability with SF. However, in this dissertation we focus on quasars in a small well-observed 290 deg^2 region of the celestial equator ($-50^\circ < \alpha < 60^\circ$, $-1^\circ.25 < \delta < +1^\circ.25$) known as the SDSS Stripe 82, with on average over 60 epochs per object. Thus Stripe 82 quasars can be individually fit with the DRW model, which allows correlating their variability with quasar physical properties, such as black hole mass, Eddington ratio, or luminosity (MacLeod et al., 2010). Although not explicitly employed in this study, characteristic variability of quasars can also be used for their selection and classification (see Butler & Bloom 2011; Clarke et al. 2019; Jin et al. 2019; MacLeod et al. 2012; Palanque-Delabrouille et al. 2016; Peters et al. 2015; Sánchez-Sáez et al. 2019; Tie et al. 2017; Yang et al. 2017).

1.2 Addressing the issue of enhanced variability on monthly timescales

MacLeod et al. (2010) (hereafter M10) modeled $\sim$ 9000 spectroscopically confirmed SDSS

Stripe 82 quasars as a damped random walk. They found that the DRW model could explain the data at the 1% level. They sought correlations between the fitted DRW model parameters - characteristic timescale τ and asymptotic variability at long timescales SF_∞ - and physical quasar properties, such as black hole mass, luminosity, or Eddington ratio. Their findings, especially fitted τ, were affected by the limited 8-year baseline of the SDSS dataset. In fact, a later study by Kozłowski (2016) (hereafter K16) claimed that to achieve an unbiased measurement of τ, the light curve baseline must be at least 10 times longer. Therefore, we initially set out to extend the SDSS quasar light curves with the available datasets (PS1, CRTS, PTF), and revisit correlations studied by M10.

The distribution of rest-frame τ in M10 SDSS dataset peaked at 200 days, similar to the results of Kelly et al. (2009), and Kozłowski et al. (2010). However, Graham et al. (2014), using a Slepian Wavelet Variance analysis of the Catalina Real-Time Transient Survey (CRTS) data, reported a discovery of the characteristic rest-frame timescale of approximately 54 days. The short characteristic timescale was at odds with previous studies and would jeopardize the validity of using the DRW model description. Therefore in Chapter 2 we set out to solve the puzzling 54 day timescale. Through our detailed error analysis of non-variable SDSS standard stars, we found that the photometric uncertainties reported by the CRTS image processing pipeline were underestimated by about 20 - 30 %. We thus reconciled the CRTS data with the SDSS, further supporting our conclusions with the PTF data.

1.3 Improving the DRW model parameters with PS1 data

Having addressed the issue of excess short-timescale variability, in Chapter 3 we extended the SDSS S82 dataset with other available survey data. We explored the viability of combining the SDSS data with PTF, PS1, CRTS, and ZTF data, and derived appropriate photometric offsets. With the results of fitting simulated quasar light curves with a DRW model, and light curve error analysis, we elected to use only PS1 r-band and SDSS r-band data, which carry sufficient similarity that photometric transformation is not necessary. We fitted the combined SDSS-PS1 light curves with the DRW model using the Gaussian process methodol-

ogy (Rasmussen & Williams, 2006). We found new correlation coefficients between the DRW model parameters and quasar luminosity, black hole mass, and Eddington ratio. With PS1 results, compared to SDSS-only results, the variability amplitude is more strongly dependent on virial black hole mass, and depends less on quasar luminosity. Also, the characteristic timescale is less strongly dependent on quasar luminosity. The correlation of variability amplitude and the Eddington ratio is similar to that found by M10 (a power-law slope of -0.208). We provide a framework for improving the analysis as more data becomes available with other sky surveys.

1.4 Preparing for the LSST - characterizing photometry in crowded stellar fields

With the beginning of the commissioning phase of the Large Synoptic Survey Telescope (LSST) in 2021, and the commencement of operations in 2022, each upcoming data release will bring a vast stream of deep imaging of the sky (Ivezić et al., 2019). Even the first few years of the wide-fast-deep survey will double the existing baselines of Stripe 82 quasars (from 8-12 years to 20-25 years in the observed frame), and millions more in the footprint of SDSS, PS1, DES, and other surveys. Therefore, a thorough understanding of systematics in the LSST science pipelines will play a crucial role in utilizing the LSST data and combining them with existing datasets. Since LSST will be much deeper than PS1 (5σ limiting magnitude of 24.5), confusion and crowding will become more important. Indeed, mapping the Milky Way is one of the four main LSST science themes, and there is a pressing need to establish the LSST performance in crowded fields (Ivezić et al., 2019). Furthermore, since quasars like stars are point sources, an analysis focused on stars informs the challenges that would affect preparation of quasar light curves. Thus, in preparation for LSST, in Chapter 4 we investigated the performance of the LSST science pipelines in crowded stellar fields.

We analyzed the DECam Plane Survey (DECaPS) data, performing a reprocessing of the calibrated observations with the LSST science pipelines, and comparing them to the results of the best currently available pipeline by Schlafly et al. (2018). We selected regions at varying stellar crowdedness based on the Galfast simulation of the night sky (Jurić et al., 2008).

According to Galfast, to process the full 18 000 deg^2 of the night sky, we need to be able to handle densities up to 1 million sources per deg^2 (Ivezić & the LSST Science Collaboration, 2018). We assessed source detection completeness of LSST to DECaPS, finding that there is a gradual degradation, mostly in completeness (85% at the edges of the Galactic plane region, single-visit depth of r-band 24.5 mag). For instance, for a region of moderate density of 200 000 sources per deg^2, the mean completeness between 18th and 20th mag is at 80% level, dropping to 50% at 21.5 mag. We also evaluated photometric repeatability, and found that both LSST and DECaPS are in systematics-dominated regime above 19th mag. The astrometric scatter for LSST is in the range of 25-30 miliarcsec, and is not strongly dependent on stellar densities. With custom-made metrics we made recommendations for improvements needed to bring the LSST science pipelines to their required performance level.

Chapter 2

SOLVING THE PUZZLE OF DISCREPANT QUASAR VARIABILITY ON MONTHLY TIME SCALES IMPLIED BY SDSS AND CRTS DATASETS

This Chapter was first published in the Monthly Notices of the Royal Astronomical Society as Suberlak et al. (2017). In this Chapter we present an improved photometric error analysis for the 7,100 CRTS (Catalina Real-Time Transient Survey) optical light curves for quasars from the SDSS (Sloan Digital Sky Survey) Stripe 82 catalogue. The SDSS imaging survey has provided a time-resolved photometric dataset which greatly improved our understanding of the quasar optical continuum variability: data for monthly and longer timescales are consistent with a damped random walk (DRW). Recently, newer data obtained by CRTS provided puzzling evidence for enhanced variability, compared to SDSS results, on monthly timescales. Quantitatively, SDSS results predict about 0.06 mag root-mean-square variability for monthly timescales, while CRTS data show about a factor of two larger rms, for spectroscopically confirmed SDSS quasars. Our analysis has successfully resolved this discrepancy as due to slightly underestimated photometric uncertainties from the CRTS image processing pipelines. As a result, the correction for observational noise is too small and the implied quasar variability is too large. The CRTS photometric error correction factors, derived from detailed analysis of non-variable SDSS standard stars that were re-observed by CRTS, are about 20-30%, and result in reconciling quasar variability behaviour implied by the CRTS data with earlier SDSS results. An additional analysis based on independent light curve data for the same objects obtained by the Palomar Transient Factory provides further support for this conclusion. In summary, the quasar variability constraints on weekly and monthly timescales from SDSS, CRTS and PTF surveys are mutually compatible, as well as

consistent with DRW model.

2.1 Introduction

Variability can be used to both select and characterise quasars in sky surveys (for a recent overview see Lawrence 2016). Although various time scales of variability can be linked to physical parameters, such as accretion disk viscosity, or corona geometry (Kelly et al. 2011; Graham et al. 2014), the physical mechanism remains elusive. Most viable explanations for observed variability include accretion disk instabilities (Kawaguchi et al., 1998), surface thermal fluctuations from magnetic field turbulence (Kelly et al., 2009), and coronal x-ray heating (Kelly et al., 2011, see Kozłowski 2016 for a review).

The diversity of physical scenarios available to explain the origin of quasar variability results in a variety of ways to characterise it. The two most widely used approaches to describing the variability of quasars include a structure function (SF) analysis and light curve fitting based on damped random walk (DRW, also known as the Ornstein–Uhlenbeck process) model (Kelly et al. 2007; MacLeod et al. 2011). An SF analysis essentially measures the width of the magnitude difference distribution as a function of the time separation, Δt. The DRW model approach is better suited for well-sampled light curves with a typical cadence of days (Kozłowski, 2016; Zu et al., 2013), whereas an ensemble SF analysis is better for sparsely sampled light curves (de Vries et al., 2005; Hawkins, 2002; Vanden Berk et al., 2004a); for a review and discussion see Kozłowski (2016). Although the sampling for CRTS (the Catalina Real-time Transient Survey) light curves analysed here (see §2.2.2) might be adequate for light curve fitting, we nevertheless opt for the SF approach because it allows for more straightforward analysis when data quality is suspect.

The observed SF is often characterised by a simple power law (Schmidt et al., 2010). If the probed time scales are long enough ($\sim$ years), the power law flattens above a characteristic timescale, τ (Ivezić et al., 2004; Kelly et al., 2007; MacLeod et al., 2010). This timescale may correspond to a transition from the stochastic thermal process that drives the variability to the physical response of the disk that successfully dampens the amplitude on longer

timescales (Collier & Peterson, 2001; Kelly et al., 2009; Kelly et al., 2007, 2011; Lawrence, 2016). In the context of a DRW model, the expected SF is described by

$$\mathrm{SF}(\Delta t) = \mathrm{SF}_\infty \left[1 - \exp(-\Delta t/\tau)\right]^{1/2}, \tag{2.1}$$

where SF_∞ is the asymptotic value of the structure function (for $\Delta t \ll \tau$, $\mathrm{SF}(\Delta t) \propto \Delta t^{1/2}$).

Most studies found that $\tau > 100$ days (MacLeod et al. 2010; Kozłowski 2016). It is a relatively short timescale compared to the dominant timescale of variation for quasars, that exceeds 10 years (Hawkins, 2007). Recently, Graham et al. (2014) found a characteristic time scale in quasar's rest frame of about 54 days, using the Slepian Wavelet Variance (SWV) analysis of CRTS light curves (the SWV time scale denotes the point at which the ensemble SWV for quasars deviates from the ensemble SWV for a DRW realization of the same data set, and is thus different from τ obtained in DRW analysis). This short timescale implies much stronger variability on monthly time scales than observed in SDSS data: SDSS results from MacLeod et al. (2010) predict about 0.06 mag root-mean-square (rms) variability for timescales below 50 days, while this CRTS-based analysis implies about a factor of two larger rms. These discrepancies have serious implications for physical interpretations of quasar variability: observed time scales are directly related to physical processes and increased variability levels call in question DRW as a viable model for describing quasar light curves (MacLeod et al. 2010; Kozłowski 2016).

It is not obvious whether these discrepancies are due to various problems with the CRTS and/or SDSS datasets (inadequate sampling, incorrect estimates of photometric errors, etc), or perhaps are due to different analysis methods (SWV vs. SF analysis). Here we reanalyse these CRTS data using the same SF method as used by MacLeod et al. (2010) to analyse SDSS data, and investigate the origin of these discrepant timescales and variability levels. We argue that the most likely explanation of these discrepancies are slightly under-estimated photometric errors for CRTS light curve data.

2.2 Data Sets

We study stars and quasars selected from the sky region known as SDSS Stripe 82 (S82; a $\sim$300 deg^2 large region along the Celestial Equator: $22^h24^m < \mathrm{RA} < 04^h08^m$ and $|\mathrm{Dec}| < 1.27°$). We utilise both SDSS and CRTS photometric data.

2.2.1 Sloan Digital Sky Survey (SDSS)

We use two SDSS catalogues, with five-band near-simultaneous photometry for 9,258 quasars, and 1,006,849 standard stars (non-variable stars, as implied by the repeated SDSS photometry, see Ivezić et al. 2007). The quasar catalogue[1] includes spectroscopically confirmed quasars from the SDSS Data Release 7 (Abazajian et al., 2009), based on the SDSS Quasar Catalogue V (Schneider et al., 2010), and was compiled by MacLeod et al. (2012). The SDSS standard stars catalogue[2] was constructed as described in Ivezić et al. (2007).

2.2.2 Catalina Real-time Transient Survey (CRTS)

The main goal of CRTS was to find near-Earth objects. Its short intra-night cadence (4 exposures per night) was designed to allow a rapid follow-up (Graham et al., 2015b), and white light (without filter) light curves maximise the sensitivity for faint objects. Three survey telescopes (0.7m Catalina Sky Survey Schmidt in Arizona, 1.5m Mount Lemmon Survey telescope in Arizona, and the 0.5m Siding Spring Survey Schmidt in Australia) were equipped with identical, 4kx4k CCDs (see Djorgovski et al. (2011) for technical details). Although in principle white light magnitudes can be calibrated to Johnson's V band zero point (Drake et al., 2013), this step was unnecessary in our analysis.

In this study we used a sample of 7,932 spectroscopically confirmed S82 quasars from the CRTS Data Release 2, based on the list by MacLeod et al. (2012). The majority (96%) of CRTS quasar light curves span the time of 7-9 years, with typical sampling of 1 to 4

[1] http://www.astro.washington.edu/users/ivezic/cmacleod/qso_dr7/Southern.html

[2] http://www.astro.washington.edu/users/ivezic/sdss/catalogs/stripe82.html

observations per night, 70 observing nights on average, and the median interval between two successive observing nights is 17.52 days (see Fig. 2.1). We also use CRTS light curves for 52,133 randomly chosen 10% subsample of the S82 standard stars from Ivezić et al. (2007).

2.2.3 Preprocessing

It is common to bin the data to reduce noise, by averaging over timescales shorter than what is required by the science goals. In this study, the hourly timescale of intra-night variability of CRTS light curves, with ~ 4 epochs each night, is much shorter than the timescales of interest (of the order of tens of days). We day-averaged all CRTS light curves following a procedure similar to Charisi et al. (2016). We adopt a convention that an index i runs over intra-night observations, and an index j separates distinct observing nights. Thus the day-averaged time stamp is :

$$t_j = \langle t_{ij} \rangle = N^{-1} \sum_{i=1}^{N} t_{ij} \tag{2.2}$$

where N is the number of observations per night. We similarly replace each set of N brightness measurements from the j-th night by their mean weighted by the inverse square of error:

$$m_j = \langle m_{ij} \rangle = \frac{\sum_{i=1}^{N} w_{i,j} m_{i,j}}{\sum_{i=1}^{N} w_{i,j}} \tag{2.3}$$

with weights $w_{i,j} = err_{i,j}^{-2}$, where $err_{i,j}$ are photometric uncertainty (colloquially, "error") estimates for individual photometric data points computed by the CRTS photometric pipeline. Averaging in flux space, instead of magnitude space, would not qualitatively change the results (because photometric uncertainties are sufficiently small).

Finally, we estimate the error on the weighted mean m_j by the inverse square of the sum of weights:

$$err_j = \left(\sum_{i=1}^{N} w_{i,j} \right)^{-1/2}, \tag{2.4}$$

and to avoid implausibly small error estimates, we add in quadrature 0.01^m to err_j if $err_j < 0.02^m$ (note that for homoscedastic errors, $err_{i,j} = \overline{err}$, $err_j = \overline{err}/\sqrt{N}$).

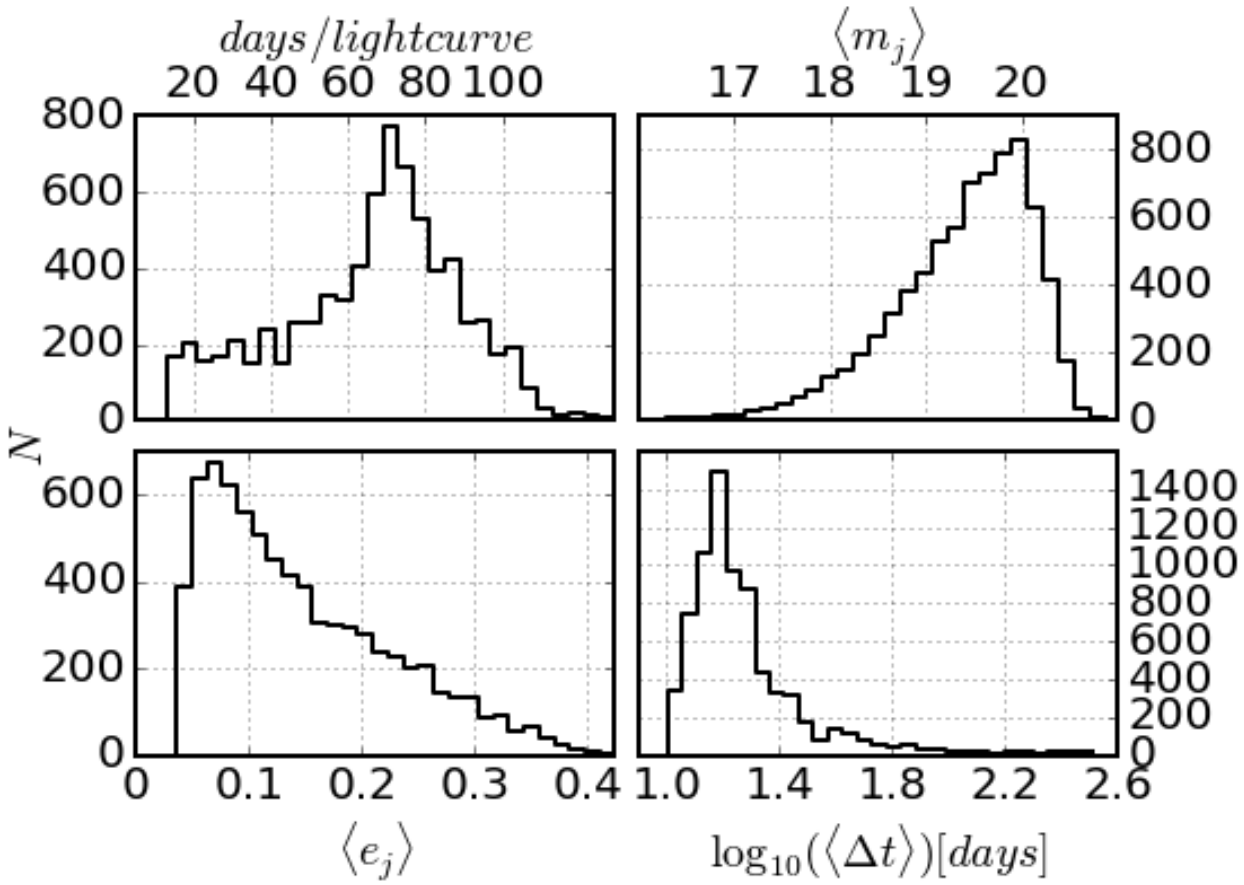

Figure 2.1: The distribution of properties of 7,601 CRTS quasar light curves for objects that were observed on at least 10 distinct nights (epochs). The distribution of the number of distinct nights is shown in the upper-left panel. Within that sample, 96% of light curves are longer than 7 years. The upper-right panel shows the mean day-averaged CRTS magnitude, $\langle m_j \rangle$ (see eq. 2.3). The bottom-left panel shows the mean day-averaged error, $\langle \sigma_j \rangle$ (see eq. 2.4). We use only quasars with light curve averaged error smaller than 0.3, leaving 7,108 quasars in the sample. The bottom-right panel shows the mean time difference $\langle \Delta t \rangle$ between day-averaged epochs. All means here are calculated per light curve.

Table 2.1: Count of stars and quasars, selected by their SDSS r magnitudes and $g-i$ colours.

r mag.	red stars	blue stars	quasars
17-18	2993	2795	185
18-18.5	2087	1400	333
18.5-19	2327	1496	747
total	7407	5691	1265

2.2.4 Final Sample Selection

We have selected both quasars and stars using a combination of information from SDSS and CRTS. To find magnitude difference between different observing nights, we first require that the raw light curves must have more than 10 photometric points (raw epochs). This step reduces the sample size from the initial 52,131 stars and 7,932 quasars to 49,385 stars and 7,707 quasars. After day-averaging, we also remove light curves with less than 10 observing nights (day-averaged epochs), leaving 48,250 stars and 7,601. In addition, we require that the light curve-average of nightly errors $\langle err_j \rangle < 0.3^m$ (see Fig. 2.1); this step removes fewer than 10% of light curves. Our final samples include 42,864 stars and 7,108 quasars.

A crucial part of our analysis below is a test of photometric uncertainties computed by the CRTS photometric pipeline using repeated CRTS observations of non-variable stars. In order to test for possible systematic effects with respect to magnitude (most notably the increase of photometric noise towards the faint end) and colour, we first select subsamples from three magnitude bins, using the SDSS r magnitudes: *bright:* 17-18, *medium:* 18-18.5, and *faint:* 18.5-19. We note that the faint completeness limit of the SDSS spectroscopic quasar sample is $r \sim 19$, and that the CRTS white light magnitudes are strongly correlated with the SDSS r magnitudes. Furthermore, we split the stellar sample using SDSS colour measurements into the "blue" ($-1 < g-i < 1$) and "red" ($1 < g-i < 3$) subsamples. Table 2.1 shows the number of objects in each type-magnitude bin.

2.3 Analysis

The structure function (SF) is a well-studied approach to characterizing light curves (de Vries et al., 2005; Graham et al., 2013; Ivezić et al., 2004; Kozłowski, 2016; MacLeod et al., 2010; Vanden Berk et al., 2004a). SF is closely related to the auto-correlation function (ACF), which in turn is the Fourier Transform of the frequency power spectrum (PS) (for a detailed discussion, see Ivezić et al. 2014; Kozłowski 2016). We choose to analyse light curves with SF over PS because the main motivation for our paper is to resolve the discrepancy between quasar timescales found with SDSS data using the SF method (MacLeod et al., 2010, 2011, 2012), and those based on CRTS data using the Slepian Wavelet Variance method (Graham et al., 2014). Given that we suspect the CRTS data quality to be the issue, we decided to also use the SF method with the CRTS dataset to ensure mathematical framework consistent with previous studies. Power spectrum analysis would introduce a third method, and thus would be less adequate to use in our study.

The SF for a light curve is a measure of the width of the magnitude difference distribution, as a function of the time separation, Δt (see below for a discussion of how to account for observational errors). For two (day-averaged) epochs j and k, with $j > k$, the magnitude difference is computed as $\Delta m_{j,k} = m_j - m_k$, the time difference is $\Delta t_{j,k} = t_j - t_k$, and the combined magnitude measurement error (measurement uncertainty for $\Delta m_{j,k}$) is $e_{j,k} = (err_j^2 + err_k^2)^{1/2}$ (where err_j is defined by eq. 2.4).

We compute SF as a function of time difference $\Delta t_{j,k}$ (hereafter, Δt for brevity and similarly, Δm for $\Delta m_{j,k}$ and e for $e_{j,k}$) by binning (Δt, Δm, e) data along Δt axis. With a mean number of data points per light curve of 70, on average we generate $\sum_{j=2}^{70}(j-1) = 2,415$ (Δt, Δm, e) data points. This large number allows us to simply use 200 linearly spaced bins of Δt, which provide adequate time resolution while ensuring sufficiently large number of Δm values per bin.

Given that we suspect data and data processing problems as a plausible explanation for discrepant results between SDSS-based and CRTS-based studies, we choose to study

variability in the observed frame (the available SDSS redshifts for all objects enable analysis in the rest frame, too - see Fig. 2.5).

The top two panels in Fig. 2.2 show the standard deviation for Δm, and the robust standard deviation ($\sigma_G = 0.741(q_{75} - q_{25})$, where q_{25} and q_{75} are 25% and 75% quartiles) estimate computed from the interquartile range, as a function of Δt for quasars, and separately for blue and red stars. σ_G is somewhat smaller than the standard deviation, which indicates mild non-Gaussianity of Δm distributions. For Δt below about 100 days, all three subsamples show similar behaviour, while for longer time scales quasars show appreciably larger scatter of observed Δm due to intrinsic variability. In order to estimate the intrinsic variability, these "raw" measurements need to be corrected for the effects of observational (measurement) errors, as described next.

2.3.1 Effects of Observational Errors on Structure Function

Given a bin with M values of (Δt_i, Δm_i, e_i), $i = 1...M$, SF will correspond to the rms width of the Δm_i distribution, σ_{tot}, *only if* all e_i are negligibly small compared to the true SF value. When measurement uncertainties are homoscedastic, $e_i = \bar{e}$, then simply SF $= (\sigma_{tot}^2 - \bar{e}^2)^{1/2}$. In a general case of heteroscedastic uncertainties, the correction for the effects of observational errors is more involved because each value Δm_i is drawn from a *different* Gaussian distribution whose width is given by $\sigma_i = (\mathrm{SF}^2 + \mathrm{e_i^2})^{1/2}$. Indeed, in this general case the distribution of all Δm_i in a given bin *need not be a Gaussian at all!*

We refer the reader for a detailed discussion of how to estimate SF in a general case to Ivezić et al. (2014), and here briefly summarise the gist of their maximum likelihood method. The likelihood of a set of M measurements Δm_i is given by

$$p(\{\Delta m_i\}|\mathrm{SF}, \mu, \{\mathrm{e_i}\}) = \prod_{\mathrm{i=1}}^{\mathrm{M}} \frac{1}{\sqrt{2\pi}\sigma_\mathrm{i}} \exp\left(\frac{-(\Delta \mathrm{m_i} - \mu)^2}{2\sigma_\mathrm{i}^2}\right), \tag{2.5}$$

where $\{.\}$ denotes a set of values and μ is introduced to account for possible systematic photometric errors between observing epochs that define the bin's Δt_i values. We note that

this expression is only an approximation to the true likelihood because it assumes that measurement errors for Δm_i are uncorrelated. This assumption is, strictly speaking, not true because different Δm_i values can be based on the same individual magnitude measurements. In practice, the covariance between errors can introduce a bias in maximum likelihood solutions, but only for M much larger than used here these errors become not negligible compared to the SF. Indeed, we used the same maximum likelihood method as Schmidt et al. (2010), Eq.2, that assumes no correlation between errors.

There is no closed form solution for maximizing the likelihood given by eq. 2.5 and we estimate SF numerically, using code[3] from *astroML* python module (Vanderplas et al., 2012). With the aid of Bayes Theorem and using uniform priors for SF and μ, the logarithm of the posterior pdf for SF and μ becomes

$$L_p(SF, \mu) = \text{const.} - \frac{1}{2}\sum_{i=1}^{M}\left(\ln(\mathrm{SF}^2 + \mathrm{e_i^2}) + \frac{(\Delta \mathrm{m_i} - \mu)^2}{\mathrm{SF}^2 + \mathrm{e_i^2}}\right). \tag{2.6}$$

We evaluate L_p on a grid[4] of μ and SF first, find its maximum which yields the maximum a posteriori (MAP) estimates for SF and μ, and then marginalise over μ to find the posterior pdf for SF:

$$p(\mathrm{SF}) = \int_0^\infty \mathrm{p(SF}, \mu | \{\Delta \mathrm{m_i}\}, \{\mathrm{e_i}\})\mathrm{d}\mu, \tag{2.7}$$

which is used to estimate the uncertainty (the credible region) of MAP estimate for SF. When there is no strong evidence for intrinsic variability, SF tends to zero.

The bottom two panels in Fig. 2.2 show SF and μ as a function of Δt for quasars, blue and red stars. For Δt below about 1000 days, μ for all three subsamples is within 0.01^m from zero, as expected. On the other hand, SF below about 100 days is in the range 0.05-0.10^m for all three subsamples. In case of quasars, the observed SF$\sim$0.1^m for 10 d $< \Delta t <$100 d demonstrates that the difference between SDSS results from MacLeod et al. (2010) (see the yellow dashed line in the third panel) and CRTS results from Graham et al. (2014) is *not*

[3]See http://www.astroml.org/book_figures/chapter5/index.html

[4]The grid size is set using approximate solutions described by Ivezić et al. (2014).

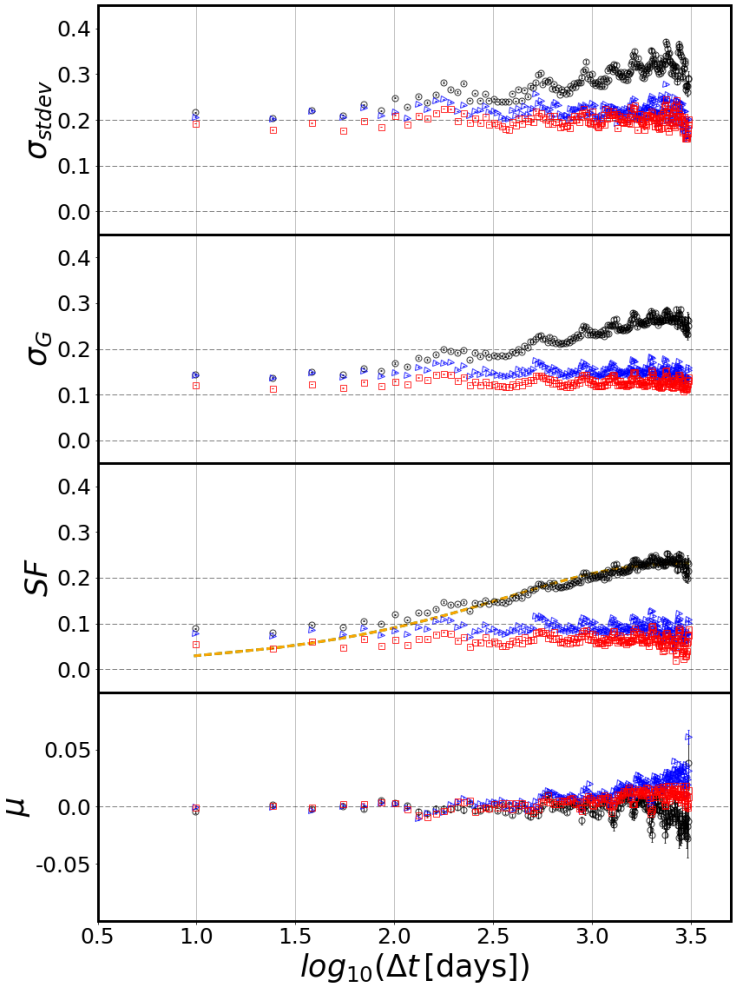

Figure 2.2: The four panels show various statistics computed for subsamples of 747 CRTS quasars (black circles), 1496 "blue" stars (blue triangles), and 2327 "red" stars (red squares), with SDSS r magnitudes in the range 18.5-19. Red and blue stars have SDSS colours $1 < \mathrm{g-i} < 3$ and $-1 < \mathrm{g-i} < 1$, respectively. All pairwise CRTS brightness differences are binned in 200 linearly spaced bins of time difference Δt. For each bin, we compute, from top to bottom: the standard deviation σ_{stdev}, the robust standard deviation estimate σ_G based on the interquartile range, the structure function SF, and the mean value of Δm per bin μ. The statistical (random) errors are often smaller than the symbol size due to large number of data points; systematic errors for all displayed quantities are probably of the order 0.01 mag (not shown). Both μ and SF are found from the two-dimensional maximum of the log-likelihood L_p on the $[\mu, \mathrm{SF}]$ grid (see eq. 2.6). The yellow dashed line in the third panel traces the fiducial Damped Random Walk model (see eq. 2.1). We address the peculiar wiggle behaviour in the Appendix 2.B, but it does not have any influence on our overall conclusions.

due to different analysis methods (SF vs. SWV, respectively): *here we fully reproduce this discrepancy using the SF method and CRTS data.*

Fig. 2.2 also indicates a plausible solution to this puzzle: the observed SF for both blue and red stars in the range $10^d < \Delta t < 100^d$ is *unexpectedly* large: the values are in the range 0.05-0.10^m rather than negligible (say, $\lesssim 0.01 - 0.02^m$). In other words, more variation is observed in light curves of non-variable stars than can be explained with reported photometric errors. The same result is obtained for all three chosen magnitude bins. Such a behaviour could be observed if photometric error estimates computed by the CRTS photometric pipeline are mis-estimated, resulting in an incorrect correction for observational errors. We proceed to perform an independent test of photometric errors using repeated observations of non-variable standard stars.

2.3.2 Tests of Observational Errors Using Non-variable Stars

Assuming that standard stars from SDSS are truly non-variable, if (Gaussian) photometric error estimates computed by the CRTS photometric pipeline are correct, then the distribution of $\chi_i = \Delta m_i / e_i$ for stars should be distributed as a unit Gaussian, N(0,1). Deviations of the distribution width for stars from unity indicate incorrect photometric error estimates. For quasars, we expect that the width should exceed unity because of their intrinsic variability, and that the width should increase with Δt. We perform this test in Fig. 2.3, where we show χ distributions for both blue stars and quasars, and for a grid of Δt and magnitude bins.

For the shortest Δt bin (<50 days), the distributions for stars and quasars appear indistinguishable for all three magnitude bins. This similarity immediately argues that there is no detected intrinsic variability for quasars. Furthermore, the width of χ distributions for stars appears to be a function of magnitude, with very little dependence on Δt. The distribution widths for stars in each magnitude bin (all Δt values), obtained using robust width estimator σ_G, are listed in Table 2.2. For example, the bin with $18.5 < r < 19$, which contains the majority of quasars, appears to have under-estimated photometric errors by a factor of 1.3 on average. The same conclusion is derived using red stars. For small Δt, where quasar SF

Table 2.2: The robust distribution widths for χ for blue stars.

mag	σ_G
17 - 18	0.870
18 - 18.5	1.107
18.5 - 19	1.288

is intrinsically small, the quasar SF will be thus significantly over-estimated, while for large Δt, where the quasar SF is intrinsically large, the effect on SF will be small. We extend this qualitative conclusion to a more quantitative analysis in the next section.

We note that problems with CRTS photometric uncertainty estimates have been reported before (e.g., Vaughan et al. 2016). Additional analysis of CRTS photometric uncertainty estimates, beyond magnitude limits of direct interest to quasar variability analysis, is presented in Appendix 2.A.

2.3.3 Structure Function with Corrected Observational Errors

Informed by the analysis from preceding section, we assume that correction factors for photometric error estimates are independent of colour and are only a function of magnitude. Depending on the magnitude of stars and quasars, we multiply their reported photometric errors by σ_G values listed in Table 2.2, and repeat SF analysis. By construction, we expect that the width of χ distributions for blue stars will be unity, and that their SF will tend to 0. For quasars, compared to SF values shown in the third panel in Fig. 2.2, we expect somewhat smaller SF at large Δt and much smaller SF at small Δt.

Fig. 2.4 shows SF for blue stars and quasars for subsamples from the three selected magnitude bins. As evident, both expectations are born out: for all three magnitude bins, SF for blue stars is essentially vanishing within noise ($\sim 0.05^m$), while SF for quasars at small Δt is about twice smaller than in Fig. 2.2 and thus consistent with the values based on SDSS data. In Fig. 2.5 we demonstrate that this agreement with SDSS results extends to rest

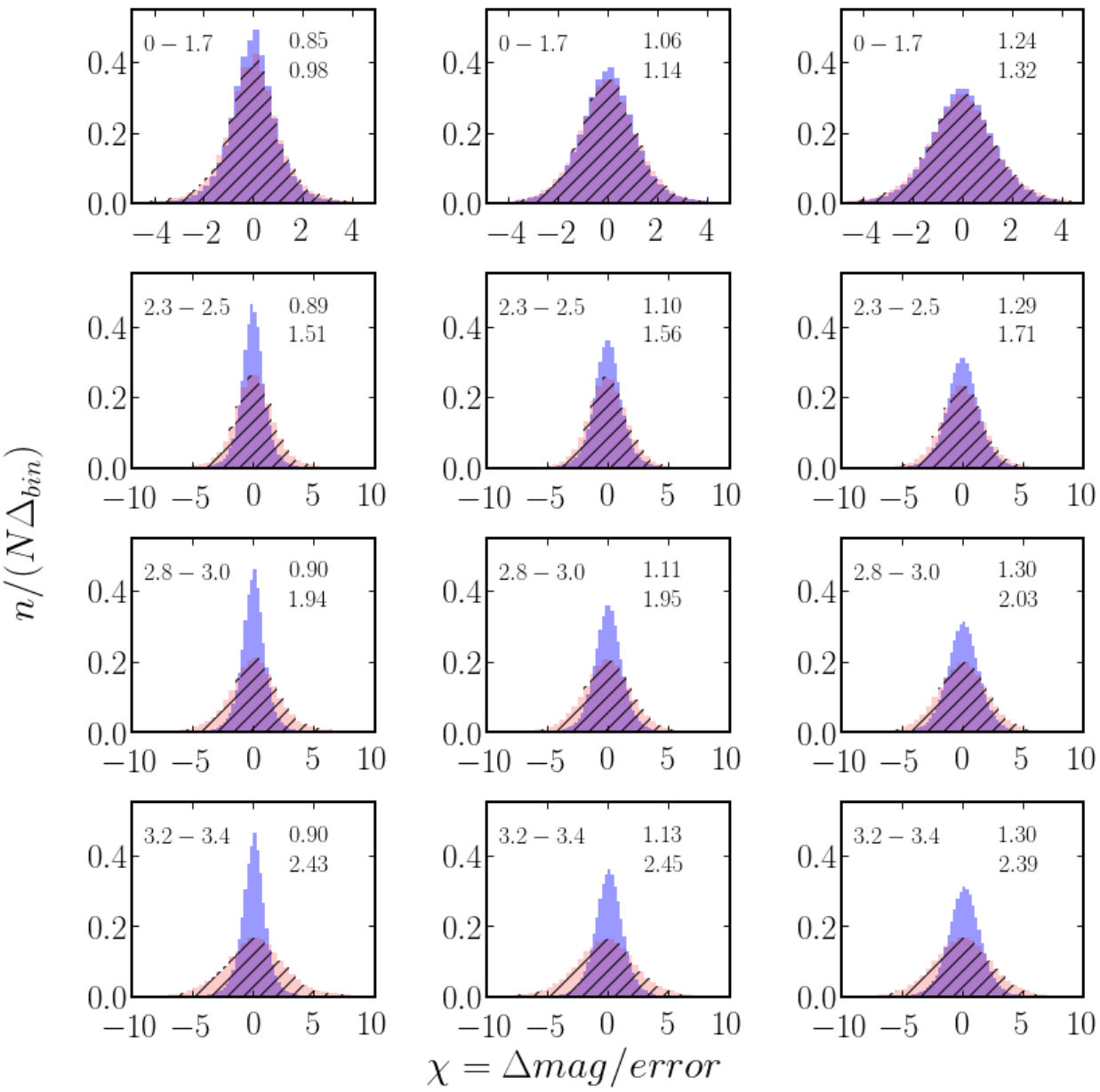

Figure 2.3: Histograms show CRTS-based $\chi = \Delta m/\text{error}$ for blue stars (blue shading) and quasars (red hatched shading), split into bins of $\log \Delta t$ (rows) and SDSS r magnitude (columns). Vertically, from top to bottom, $\log \Delta t$: $0 < \log \Delta t < 1.7$ ($t < 50$ days), $2.3 < \log \Delta t < 2.5$, $2.8 < \log \Delta t < 3.0$, and $3.2 < \log \Delta t < 3.4$ (indicated by numbers in the upper left corner of each subplot). Horizontally, from left to right, the SDSS r magnitude bins are: $17 - 18$, $18 - 18.5$, and $18.5 - 19$. The numbers in the upper-right corner of each subplot are the robust width of χ distributions determined using interquartile range (σ_G); upper value for blue stars and lower value for quasars.

Table 2.3: Count of stars and quasars, selected by their SDSS r magnitudes and $g-i$ colours. Analoguous to 2.1, except that here the counts of stars and quasars with PTF adequate data are listed.

r mag.	red stars	blue stars	quasars
17-18	1243	1077	90
18-18.5	825	497	160
18.5-19	913	548	377
total	2981	2122	627

frame analysis, too.

2.3.4 SF Estimated from PTF Data

Recent PTF (Palomar Transient Factory) Data Release 3 light curves[5] can be used for an independent test of our conclusions derived above. We queried the NASA/IPAC Infrared Science Archive[6] 'PTF Objects' catalogue using coordinates for 7,601 spectroscopically confirmed Stripe 82 quasars, and 48,250 standard stars (same as the final samples used for CRTS-based analysis). A positional multi-object search with a matching radius of 2 arcsec, with a flag 'ngoodobs' > 10, resulted in 6,471 quasars and 38,776 stars. For these objects we obtained time series data from the 'PTF Light Curve Table' catalogue (we grouped by SDSS coordinates).

We processed these PTF light curves in exactly the same way as the CRTS light curves. We first performed day-averaging, using the weighted error as the measure of uncertainty on day-averaged brightness measurement. We further selected only those objects that have been observed on at least 10 different nights, resulting in samples of 2,753 quasars and 15,714 stars. The counts of magnitude-limited subsamples are listed in Table 2.3.

[5] http://www.ptf.caltech.edu/page/lcdb

[6] https://irsa.ipac.caltech.edu

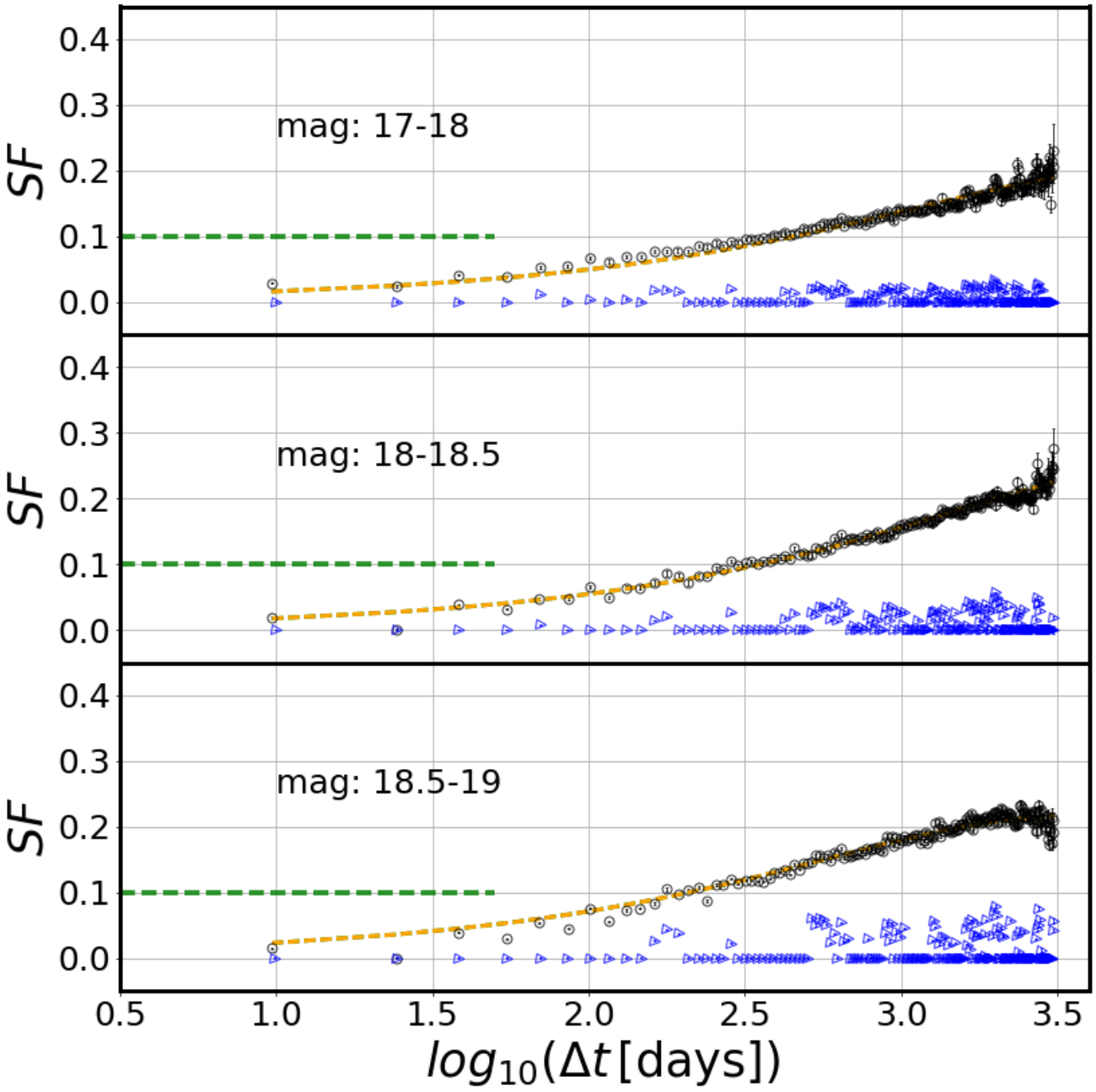

Figure 2.4: Analogous to the third panel in Fig. 2.2, except that here SF for blue stars and quasars in all three magnitude bins are shown, and photometric errors are modified by multiplicative correction factors listed in Table 2.2. Note that SF for stars in vanishing, while SF for quasars at $\log_{10}(\Delta t) < 1.7$ is about twice as small as in Fig. 2.2.

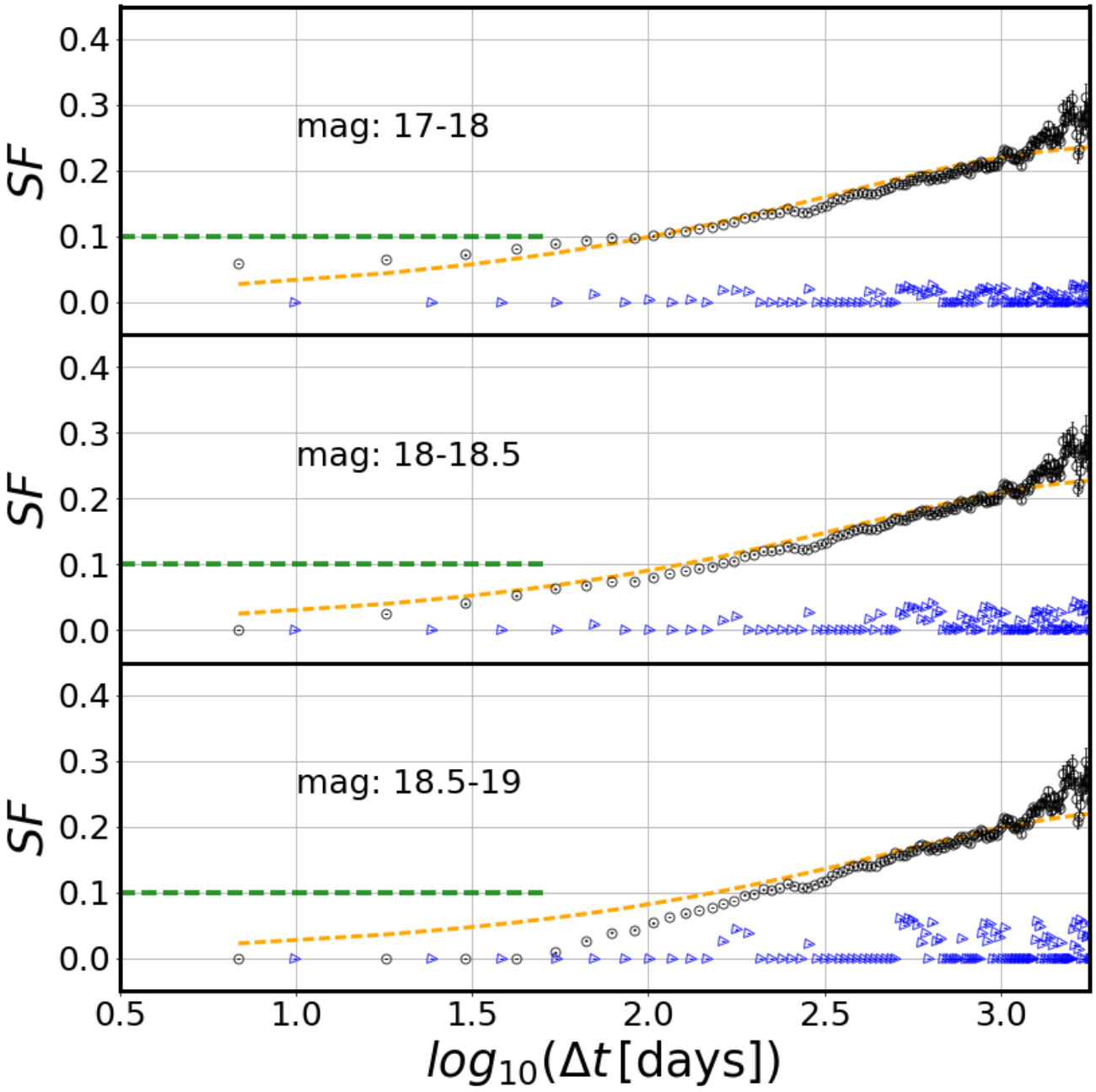

Figure 2.5: Analogous to Fig. 2.4, but here for Δt in the quasar rest frame : $t_{rest} = t_{obs} / (1+z)$, using known quasar redshifts from SDSS (MacLeod et al., 2010). The rest frame correction shifts time lags to shorter timescales and produces SF for quasars in agreement with corresponding results obtained by (MacLeod et al., 2010).

The SF results based on PTF light curve data are shown in Fig. 2.6. For these uncorrected PTF data, it is evident that there is no sign of variability for quasars on short timescales ($\Delta t < 100$ days) above the SDSS-level of $\sim 0.05^m$ (unlike for CRTS data, see Fig. 2.2). Note also that standard stars show no appreciable variability at any time scale (SF≈ 0). Therefore, this PTF-based analysis further supports our conclusion that extraneous quasar variability at short time scales was due to slightly underestimated photometric uncertainties.

2.4 Conclusions

We analysed the error properties of the CRTS sample of quasars and standard stars. Using repeated CRTS observations of non-variable stars, we found that the photometric error estimates computed by the CRTS photometric pipeline are slightly under-estimated for the majority of quasars. When quasar light curves are corrected for the impact of observational errors, the resulting corrections to the SF are thus too small. For small Δt, where quasar SF is intrinsically small, quasar SF is significantly over-estimated (akin to the subtraction of two large numbers to get a small number, when the second large number is under-estimated). In particular, at time scales of about 50 days, SF is over-estimated by about a factor of two. This behaviour provides a plausible explanation for the increased quasar variability level in CRTS light curves reported by Graham et al. (2014), compared to earlier SDSS-based results obtained by MacLeod et al. (2010). An additional analysis based on independent light curve data for the same objects obtained by the Palomar Transient Factory provides further support for this conclusion. We conclude that the quasar variability constraints on weekly and monthly time scales from SDSS, CRTS and PTF surveys are mutually compatible, as well as consistent with DRW model.

2.A Variation of the CRTS photometric uncertainty with magnitude

We found in Section 3.5 (see Table 2) that reported CRTS photometric uncertainty estimates are too large by $\sim$15% in the magnitude range 17–18, and too small by $\sim$10-25% in the magnitude range 18–19. Such problems have been reported before; for example, Vaughan

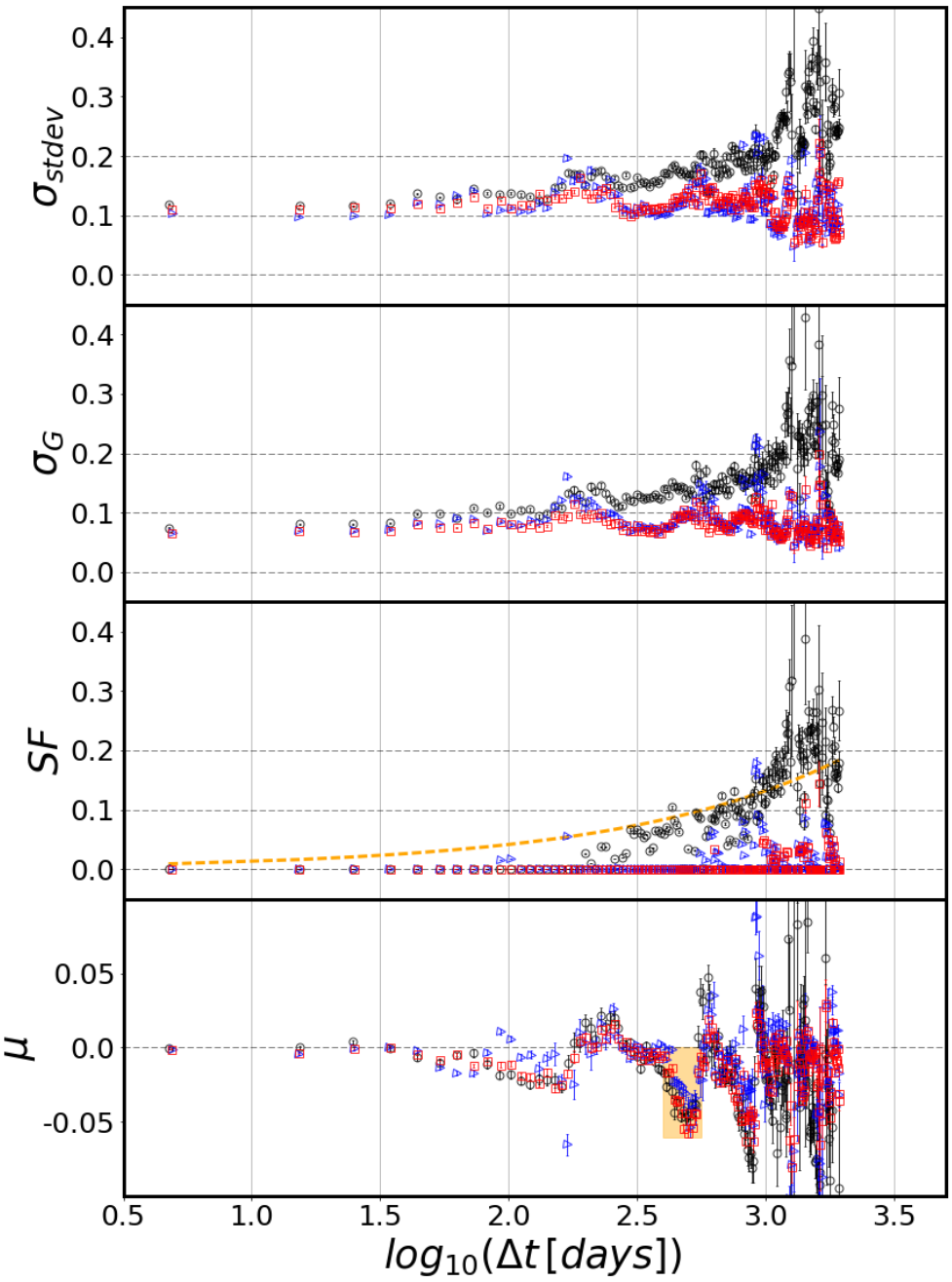

Figure 2.6: Analogous to Fig. 2.2, but here the statistics for subsamples of 377 quasars (black circles), 548 "blue" stars (blue triangles), and 913 "red" stars (red squares), with adequate PTF light curve data are shown. Note that the mean magnitude difference (μ, the bottom panel) does not stay as close to 0 as for CRTS data – a deviation around $\log_{10}\Delta t \approx 2.7$ might indicate some issues with photometric zero point calibration (at the level of 0.02-0.03^m).

et al. (2016) reported that for bright objects (magnitude ∼15) the error bars provided by the CRTS pipeline processing are overestimated by a factor of 4-5. Since this factor is much larger than we obtained for fainter magnitude bins, we extend our standard star analysis to the full CRTS magnitude range.

The top panel in Fig. 2.7 shows the variation with magnitude of the robust distribution width for the quantity

$$z_{ij} = \frac{m_{ij} - m_j}{err_{ij}}, \tag{2.8}$$

where m_j is the weighted mean magnitude for star indexed j, and index i runs over all observations of a given star. The quantity $\sigma_G(z)_j$ is the robust quartile-based distribution width of z_{ij} for a given star j. If the reported CRTS photometric uncertainties (err_{ij}) were correctly estimated, the $\sigma_G(z)$ distribution for standard (non-variable) stars would be centred on unity and independent of magnitude. As the top panel in Fig. 2.7 clearly demonstrates this is not the case: $\sigma_G(z)$ is ∼0.25 at the bright end, and increases to ∼1.5 at the faint end. In the magnitude range 17–19, the $\sigma_G(z)$ behaviour is consistent with the results listed in Table 2.

The middle and bottom panels show that the observed intrinsic scatter per light curve at the bright end is ∼ 0.01 mag, while reported photometric uncertainty is never smaller than 0.05 mag. In other words, we confirm the result reported by Vaughan et al. (2016) for the bright end and demonstrate that problems with reported CRTS photometric uncertainties are a strong function of magnitude.

2.B CSS calibration wiggles

We saw an oscillatory pattern on plots of structure function and standard deviation using CRTS data on Figs. 2.2, 2.4, 2.5. We ruled any astrophysical origin since the effect also persisted when using only standard stars. Despite an anti-correlation of the pattern with the number of points per bin, we ruled out the statistical origin by fixing the number of points per bin. Fig. 2.8 shows that wiggles persists even if we set the number of points per Δt bin to 20000. We see the effect when points are separated by (2k+1)/2 years, with $k = 0, 1, 2....$

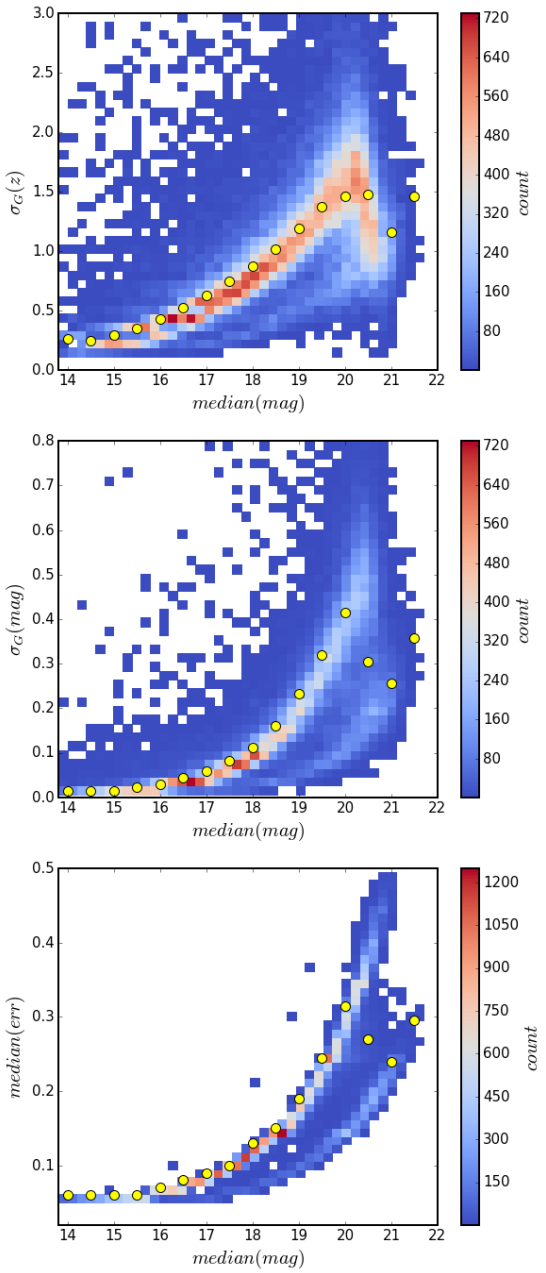

Figure 2.7: The top panel shows the variation with magnitude of the photometric scatter per light curve, normalised by reported CRTS photometric uncertainties (see eq. 2.8 for definition), using CRTS light curves for ∼48,000 standard (non-variable) stars from the SDSS catalogue. If the reported CRTS photometric uncertainties were correctly estimated, the $\sigma_G(z)$ distribution would be centred on unity and independent of magnitude. The middle panel shows the observed intrinsic scatter per light curve, and the bottom panel shows the distribution of reported photometric uncertainty, both as function of median magnitude (per light curve).

We conclude that this variation is related to the airmass which fluctuates seasonally, which was not properly accounted for in the CSS calibration process. This is because the primary aim of CSS was to detect moving objects which only requires intra-night consistency, and not long-term accuracy (Drake et al., 2013).

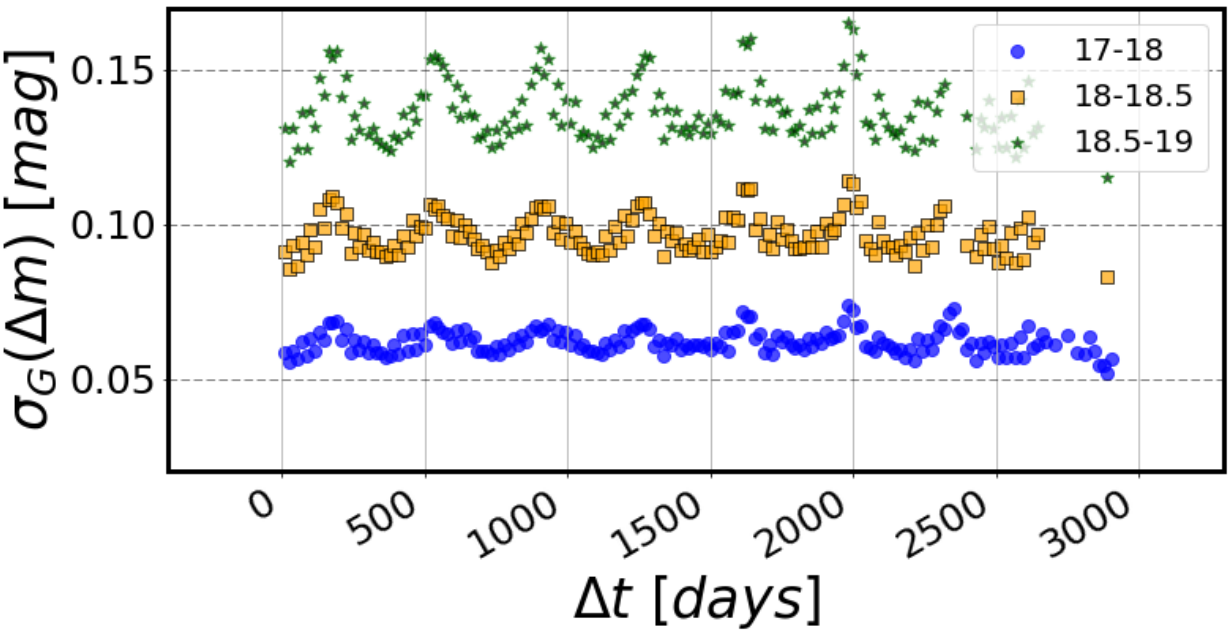

Figure 2.8: Robust standard deviation for CRTS standard stars, showing that the oscillatory pattern persists even with fixed number of points per bin. We combine the "blue" and "red" subsamples (-1< $g-i$ <3), yielding 5788, 3487 and 3823 stars in SDSS r-magnitude bins *bright* (green stars) *medium* (orange squares) and *faint* (blue circles), respectively (see Table 2.1 for counts in individual subsamples). For each Δt bin we randomly select 20000 Δm points. If there are less than 20000 points in a bin, we do not plot anything (this affects less than 35 bins per magnitude bin, mostly towards longer timescales). It illustrates that the wiggles are purely due to seasonal differences, and possibly hidden zero point errors, unaccounted for in the CSS pipeline. This pattern does not change our overall conclusions.

Chapter 3

IMPROVING DAMPED RANDOM WALK PARAMETERS FOR SDSS STRIPE 82 QUASARS WITH PAN-STARRS1

This Chapter has been submitted for publication to the Astrophysical Journal. We use the Panoramic Survey Telescope and Rapid Response System 1 Survey (Pan-STARRS1, PS1) data to extend the Sloan Digital Sky Survey (SDSS) Stripe 82 quasar light curves. Combining PS1 and SDSS light curves provides 15 years baseline for 9248 quasars - 5 years longer than prior studies that used only SDSS. We fit the light curves with Damped Random Walk (DRW) model, and correlate the DRW model parameters - asymptotic variability amplitude SF_∞, and characteristic timescale τ, with quasar physical properties - black hole mass, bolometric luminosity, and redshift. Using simulated light curves, we find that longer baseline allows us to better constrain DRW parameters. After adding PS1 data, the variability amplitude is a stronger function of the black hole mass, and has a weaker dependence on quasar luminosity. In addition, the characteristic timescale τ dependence on quasar luminosity is marginally weaker. We also make predictions for the fidelity of DRW model parameter retrieval when light curves will be further extended with Zwicky Transient Facility (ZTF) and Large Synoptic Survey Telescope (LSST) data. Finally, we show how updated DRW parameters lend an independent method of discovering Changing-Look Quasar Candidates (CLQSO). The candidates are outliers in terms of differences in magnitude and scatter between SDSS and PS1 segments. We identify 40 objects (35 newly reported) with tenfold increase in variability timescale between SDSS and SDSS-PS1 data, which is due to a large change in brightness (over 0.5 mag) - a characteristic for CLQSO.

3.1 Introduction

Quasars are variable at rest-frame optical wavelengths at the asymptotic root-mean-square (rms) level of about 0.2 mag. These distant galaxies harbor an actively accreting supermassive black hole - an active galactic nucleus. Although it is agreed upon that the majority of optical light originates from the thermal emission of the accretion disk, the detailed origin of variability has been debated for the past 50 years (see Sun et al. 2018 and references therein). Some favor a thermal origin of variability (Kelly et al., 2013), related to the propagation of inhomogeneities ('hot spots') in the disk (Cai et al., 2016; Dexter & Agol, 2011), others suggested magnetically elevated disks (Dexter & Begelman, 2019), or X-ray reprocessing (Kubota & Done, 2018). Indeed, it may well be that the answer involves combination of these - as Sánchez-Sáez et al. (2018) suggests, perhaps short-term variability (hours-days) is linked to the changes in X-ray flux, while long-term variability (months-years) is more intrinsic to the disk (Edelson et al., 2015; Lira et al., 2015). Nevertheless, quasar optical light curves have been successfully described using the Damped Random Walk (DRW) model (Kasliwal et al., 2015; Kelly et al., 2009; Kozłowski et al., 2010; MacLeod et al., 2010; Zu et al., 2011), and the DRW parameters have been linked to the physical quasar properties (MacLeod et al. 2010, hereafter M10).

Variability is also a classification tool, allowing to distinguish quasars from other variable sources that do not exhibit a stochastic variability pattern (MacLeod et al., 2011). This property is especially useful for selecting quasars in the intermediate redshift range, which overlaps the stellar locus in color-color diagrams (Sesar et al., 2007; Yang et al., 2017)). Variability has also been used to increase the completeness in measurements of Quasar Luminosity Function (see AlSayyad 2016; McGreer et al. 2018, 2013; Palanque-Delabrouille et al. 2013; Ross et al. 2013).

Power spectral density (PSD) informs us about the distribution of variability across frequency range: from short timescales (high frequencies) to long timescales (low frequencies). Quasar, or more broadly speaking, AGN variability, exhibits a broken power law PSD, of

the form $\log P(f) \propto \alpha \log(f)$, with α_l at low frequencies and α_h at high frequencies. For a pure DRW process, $\alpha_h = -2$ and $\alpha_l = 0$, so that:

$$P(f) = \frac{4\sigma^2\tau}{1 + (2\pi\tau f)^2} \tag{3.1}$$

(with $\sigma = \mathrm{SF}_\infty/\sqrt{2}$, τ the characteristic timescale, f the frequency), where $P(f) \propto f^{-2}$ at high frequencies $f > (2\pi\tau)^{-1}$, and levels to a constant value at lower frequencies (Kelly et al., 2014).

There is a debate in the literature about the exact shape of the quasar PSD, and of any possible departures from the pure DRW model. Studies using quasar data from wide-field photometric surveys (OGLE, SDSS, PS1) benefit from relatively long baselines (several years), which constrain the low frequency part of the PSD. Overall, there is no evidence of a significant departure from DRW at these long timescales, i.e. $\alpha_l \approx 0$ (Caplar et al., 2017; Guo et al., 2017; Kozłowski, 2016; Simm et al., 2016; Sun et al., 2018; Zu et al., 2013). However, these ground-based surveys suffer from a sparse sampling, which can be remedied by using a space-based telescope that can carry out near-continuous observations, like the Kepler mission (Borucki et al., 2010). Studies using Kepler data, that focused on a smaller number of well-sampled AGN with short baselines (<hundred days), found a range of power-law slopes at high frequencies - from -1 to -3.2, which includes the DRW $\alpha_h \approx -2$, but further study is needed (Aranzana et al., 2018; Edelson et al., 2014; Mushotzky et al., 2011; Smith et al., 2018). However, in this paper the timescales probed are larger than several days, and thus we can assume that DRW is the best working description of quasar variability for available optical light curves. Furthermore, in this work we directly compare the results of SDSS light curves extended with PS1 to M10 who used pure DRW description (see discussion therein on a possible departure from DRW). Therefore, to allow a better comparison of our results with M10, we use the DRW description of quasar PSD.

Due to its stochastic nature, for an unbiased parameter retrieval of DRW process the light curve is required to be several times longer than the characteristic timescale (Kozłowski et al.

2010; Kozłowski, Szymon 2017, hereafter K17). This means that DRW parameters recovered for short light curves (compared to the recovered timescale) may be biased, which in turn affects the correlations with physical parameters (black hole mass, Eddington ratio, absolute luminosity).

For this reason, while some studies have restricted the probed redshift range, limiting the quasar sample to where one would expect only shorter timescales based on previous studies (Guo et al., 2017; Kelly et al., 2013; Simm et al., 2016; Sun et al., 2018), some have elected not to study timescales at all (Sánchez-Sáez et al., 2018; Sun et al., 2018), or use the timescales recovered from short light curves primarily for classification (Hernitschek et al., 2016).

By extending available quasar light curves, we are able to better recover DRW timescales. Since almost a decade ago, when M10 published their study based on SDSS Stripe 82 data, new datasets (PS1,PTF,CRTS) have become available. They can extend the quasar light curves by almost 50%. Indeed, Li et al. (2018) combined SDSS and Dark Energy Camera Legacy Survey (DECaLS) data, to provide a 15 year baseline, but by using the entire SDSS footprint suffered from poor sampling and had to use the ensemble structure function approach.

Unlike previous studies, in this work, by combining SDSS and PS1 data for the well-observed Stripe 82, we afford both an extended baseline (15 years), a large number (9000) of quasars, and a good cadence (N $>$ 60 epochs) to which we fit the DRW model. The layout of this paper is as follows: first we confirm in Section 3.2 that extending the quasar baseline is an important improvement in providing unbiased estimates of the DRW model parameters (K17); in Section 3.3 we describe the datasets employed, and their combination onto a common photometric system; in Section 3.4 we simulate the improvement in the recovery of DRW parameters with baseline extension and realistic cadence; finally in Section 3.5 we describe the main results analyzing correlations between physical parameters and variability, in Section 3.6 we discuss the physical meaning of relevant timescales, and in Section 4.7 we summarize the main conclusions. In this work we adopt a ΛCDM cosmology with $h_0 = 0.7$

and $\Omega_m = 0.3$.

3.2 Methods

3.2.1 DRW as a Gaussian Process

Damped Random Walk (Ornstein-Uhlenbeck process, Rasmussen & Williams 2006) can be modeled as a member of a class of Gaussian Processes (GP). Each GP is described by a mean and a kernel - a covariance function that contains a measure of correlation between two points: x_n and x_m, separated by Δt_{nm} (autocorrelation). For the DRW process, the covariance between two observations spaced by Δt_{nm} is:

$$\begin{aligned} k(\Delta t_{nm}) &= \sigma^2 \exp\left(-\Delta t_{nm}/\tau\right) \\ &= \sigma^2 \mathrm{ACF}(\Delta t_{nm}) \end{aligned} \tag{3.2}$$

Here σ^2 is an amplitude of correlation decay as a function of Δt_{nm}, while τ is the characteristic timescale over which correlation drops by $1/e$. For a DRW, the correlation function $k(\Delta t_{nm})$ is also related to the autocorrelation function ACF.

Not explicitly used in this paper, but of direct relevance to the DRW modeling, is the structure function (SF). SF can be found from the data as the root-mean-square scatter of magnitude differences Δm calculated as a function of temporal separation Δt (we drop subscripts n,m for brevity). SF is directly related to a DRW kernel $k(\Delta t)$:

$$\mathrm{SF}(\Delta t) = \mathrm{SF}_\infty (1 - \exp\left(-|\Delta t|/\tau\right))^{1/2} \tag{3.3}$$

For quasars SF follows approximately a power law: SF $\propto \Delta t^\beta$, and for large time separation Δt, as epochs in the light curve cease to be correlated, it levels out to a constant value SF_∞ - the asymptotic SF. Note that σ in Eq. 3.2 is related to the asymptotic SF as $\mathrm{SF}_\infty = \sqrt{2}\sigma$ (also see Bauer et al. (2009); Graham et al. (2015a); MacLeod et al. (2012) for an overview).

3.2.2 Fitting

We evaluate the likelihood of the DRW model with a particular set of τ, σ given the data with celerite - a fast GP solver (Foreman-Mackey et al., 2017). The underlying matrix algebra is similar to that used by Rybicki & Press (1992), Kozłowski et al. (2010), and M10. Also, as in previous work, we use a prior on DRW parameters that is uniform in log space : $1/(\sigma\tau)$. The main difference in our approach is that rather than adopting the Maximum A-Posteriori (MAP) as the 'best-fit' value for the DRW parameters (as in Kozłowski et al. 2010, K17, Kozłowski 2016, M10, MacLeod et al. 2011), we find the expectation value of the marginalized posterior. This is advantageous because of a non-Gaussian shape of the posterior - otherwise, if the posterior was a 2D normal distribution, the expectation value would coincide with the maximum of the posterior (MAP solution).

3.2.3 The impact of light curve baseline

K17 reports that one cannot trust any results of DRW fitting unless the light curve length is at least ten times longer than the characteristic timescale. In this section we revisit the relationship between recovered and input timescales as a function of light curve baseline by following K17 setup. We confirm that the bias in retrieved DRW timescale depends on how many times the light curve is longer than the timescale. However, we find that the baseline does not have to be as many as ten times longer to provide meaningful, rather than unconstrained, results. Assuming fixed baseline of $\Delta T = 8$ years, we simulate 10 000 light curves, exploring 100 values of input timescales, but identical $\mathrm{SF}_\infty = 0.2$ mag, with either SDSS ($N = 60$), or OGLE-like ($N = 445$) cadence. Defining ρ as the ratio of input timescale to baseline, we probe a range of ρ between 0.01 and 15, uniform in a logarithmic grid.

For each light curve we simulate the underlying DRW signal $s(t)$ by iterating over the array of time steps t. At each step, we draw a point from a Gaussian distribution, for which the mean and standard deviation are re-calculated at each timestep (see eqs. A4 and A5 in Kelly et al. 2009, Sec. 2.2 in M10, and K17). Initially, at t_0 the signal is equal to the

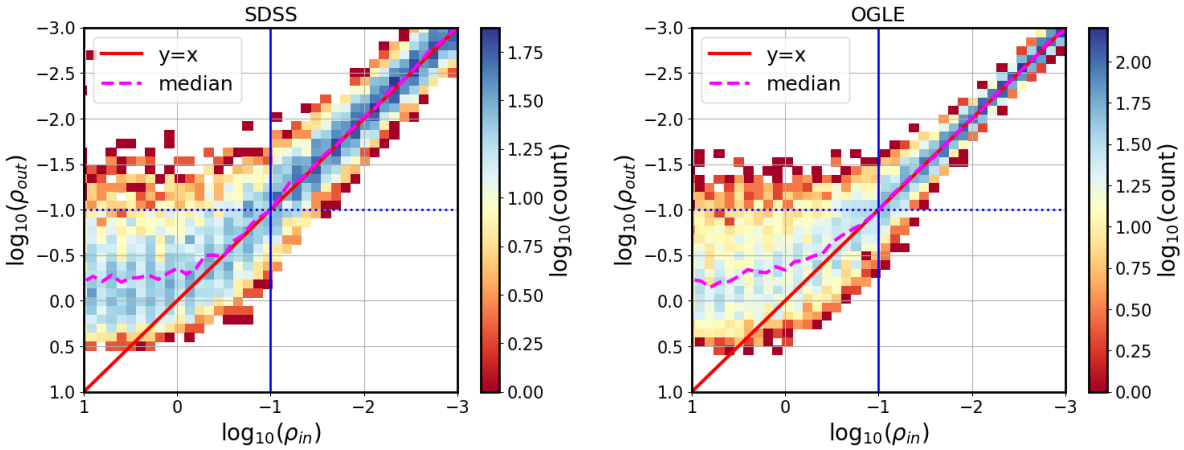

Figure 3.1: Recovery of the input DRW timescale, with baseline fixed to $\Delta T = 8$ years. We explore 100 logarithmically-spaced values of $\rho \equiv \tau/\Delta T$, simulating 100 realizations of DRW process at each ρ. The impact of photometric uncertainties and cadence is small in this case: the left panel (SDSS, N=60 epochs) does not significantly differ from the right panel (OGLE, N=445 epochs). The dotted horizontal and solid vertical lines mark $\rho = 0.1$, i.e. the baseline being ten times longer than the timescale. The solid diagonal line corresponds to a perfect recovery of DRW parameters (where $\rho_{in} = \rho_{out}$). For any quasar, extending its light curve moves it towards the upper-right (well-constrained) portion of the diagram, since for a fixed τ_{in}, increasing ΔT decreases ρ. For baselines shorter than τ, best-fit τ is underestimated and becomes biased to $\Delta T/2$.

mean magnitude, $s_0 = \langle m \rangle$. After a timestep $\Delta t_i = t_{i+1} - t_i$, the signal s_{i+1} is drawn from a normal distribution $\mathcal{N}(loc, stdev)$, with :

$$loc = s_i e^{-r} + \langle m \rangle \left(1 - e^{-r}\right) \tag{3.4}$$

and

$$stdev^2 = 0.5\,\mathrm{SF}_\infty^2 \left(1 - e^{-2r}\right) \tag{3.5}$$

where $r = \Delta t_i / \tau$, and τ is the damping timescale.

Like K17, we add to the true underlying signal with zero mean $s(t)$, a mean magnitude ($\mathrm{r_{SDSS}} = 17$ mag and $\mathrm{I_{OGLE}} = 18$ mag), and calculate magnitude-dependent estimate of photometric uncertainty:

$$\sigma^2_{\mathrm{SDSS}} = 0.013^2 + \exp\left[2(\mathrm{r_{SDSS}} - 23.36)\right] \tag{3.6}$$

$$\sigma^2_{\mathrm{OGLE}} = 0.004^2 + \exp\left[1.63(\mathrm{I_{OGLE}} - 22.55)\right] \tag{3.7}$$

To simulate observational conditions we add a Gaussian noise $n(t) = \mathcal{N}(0, \sigma(t))$:

$$y(t) = s(t) + n(t) \tag{3.8}$$

The resulting distribution of fitted timescales as a function of input timescales, scaled by the 8 year baseline: ρ_{out} vs ρ_{in}, is shown in Fig. 3.1. We confirm the findings of K17: for short light curves, the best-fit τ becomes $\sim 1/5$ of light curve length (where $\log_{10}(\rho_{out}) \approx -0.7$, the 'unconstrained' region, lower-left of each panel). However, as long as the light curve is several times longer than the timescale ($1/\rho \gtrapprox 3, i.e. \log_{10}(\rho) \lessapprox 0.5$), we can recover the timescale without substantial bias (the dashed line approaches the solid diagonal line on both panels). Therefore, by extending the baseline we can move from the biased region (bottom left) to the unbiased regime (top right). This is the basis for this study, in which we extend the baselines of quasar light curves from SDSS-only (10 years) to combined SDSS-PS1 (15 years).

3.3 Data

We focus on the data pertaining to a 290 deg^2 region of the southern sky known as Stripe 82 (S82), repeatedly observed by the Sloan Digital Sky Survey (SDSS) between 1998 and 2008. Originally aimed at supernova discovery, objects in this area were re-observed on average 60 times (see MacLeod et al. 2012, Sec. 2.2 for overview, and Annis et al. 2014 for details). Availability of well-calibrated (Ivezić et al., 2007), long-baseline light curves spurred variability research (Sesar et al., 2007). The DR9 catalog (Schneider et al., 2008) contains 9258 spectroscopically confirmed quasars within S82. Within $0.5''$ we find matching data for 9248 quasars from Pan-STARRS1 (PS1) DR2 (Chambers et al., 2016; Flewelling, 2018; Flewelling et al., 2016), 7737 from Catalina Real-Time Transient Survey (CRTS, Drake et al. 2009), 6455 from Palomar Transient Factory (PTF, Rau et al. 2009), and 8001 from Zwicky Transient Facility DR1 (ZTF, Bellm et al. 2019; Masci et al. 2019). Fig. 3.2 illustrates the improvement in baseline coverage when combining various surveys. The width of each rectangle corresponds to the duration of each survey (survey baseline), and the height to the area covered by each survey. The lower edge of each rectangle marks the 5σ depth in the r-band (or equivalent). LSST stands out in that it will provide the best extension of SDSS baseline and depth.

Combining data from different photometric standards requires applying color transformation, or photometric offsets. We first seek to combine PS1 gri, PTF gR, and CRTS V, into a common SDSS r-band (best photometry). To this end we calculate color terms using the SDSS standard stars catalog (Ivezić et al., 2007). Focusing on a 100 000 randomly chosen stars, we find their CRTS, PS1, and PTF matches [1].

The difference between the target (SDSS) and source (eg. PS1) photometry can be written as a function of the mean SDSS $g - i$ color:

---

[1] CRTS from B.Sesar, priv.comm., PS1 from MAST (`http://panstarrs.stsci.edu`), and PTF from IRSA PTF Object Catalog (`https://irsa.ipac.caltech.edu/`)

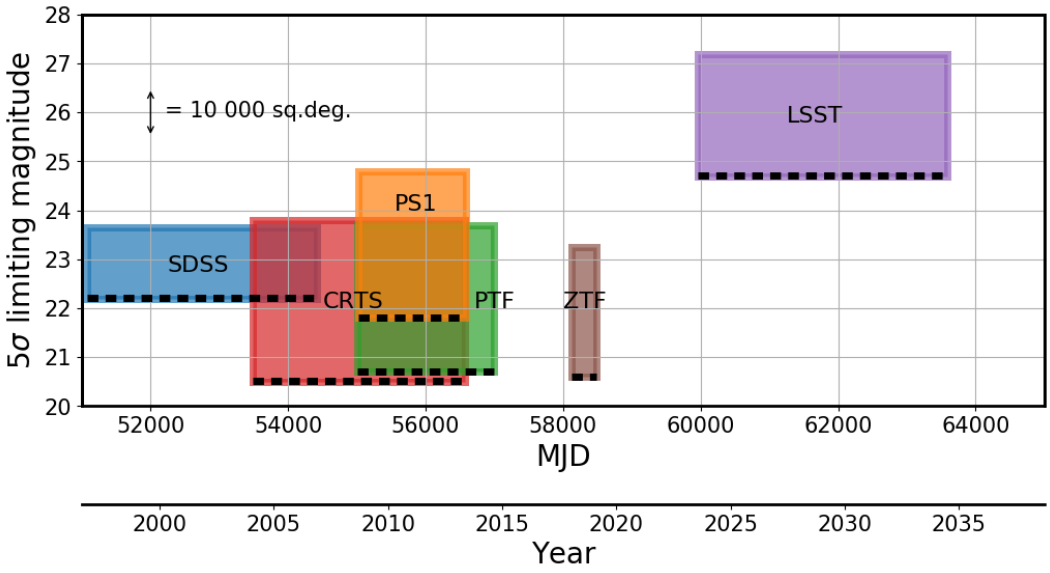

Figure 3.2: An illustration of survey baseline, sky area covered, and depth. The width of each rectangle corresponds to the extent of real or simulated light curves for Stripe 82 quasars for each survey. This includes SDSS DR7, CRTS DR2, PS1 DR2, PTF DR2, ZTF DR1, and for LSST the full 10-year survey. The lower edge of each rectangle (marked by a thick dashed line) corresponds to the 5σ limiting magnitude (SDSS r, PS1 r, PTF R, ZTF r, LSST r, CRTS V). The vertical extent of each rectangle corresponds to the total survey area (for SDSS, up to DR15). Note how PS1 and PTF extend the baseline of SDSS by approximately 50%, and how inclusion of LSST roughly triples the SDSS baseline. For reference, the area covered by LSST is 20000 sq.deg.

$$m_{\rm PS1} - m_{\rm SDSS} = f(g-i) \tag{3.9}$$

Some authors (eg. Li et al. 2018) allow the transformation to be a higher-order polynomial, but as Fig. 3.3 shows, quasars occupy a relatively narrow region of $g-i$ color space, and we find that the linear fit is sufficient. The derived linear coefficients for photometric transformations between SDSS r and PS1 gri, PTF gR, CRTS V, as a function of SDSS $g-i$ color, are listed in Table 3.1. We illustrate the process showing in Fig. 3.4 the SDSS-PS1 standard stars data used to calculate the offsets. Note that the PS1 r (middle panel) is very close to the SDSS r - within 0.01 mag (1% level) across the $g-i$ color range. We focus on SDSS r - PS1 r offset as a function of magnitude in Fig. 3.5 - the near-equivalence of bandpass coverage is valid at 1% level up to $r < 20.5$.

In selecting the most beneficial datasets to complement SDSS r we also consider the associated photometric uncertainties (aka 'errors'). As shown in Fig. 3.2, PTF and CRTS are shallower than SDSS or PS1. Therefore for faint objects, like quasars (for S82 sample the population median is SDSS $r \sim 20$ mag) PTF and CRTS have larger photometric uncertainties than SDSS or PS1. Indeed, as Fig. 3.6 shows, the distribution of median errors for PTF, CRTS, and ZTF quasar data is wider than the corresponding SDSS and PS1 data. As simulations show (Sec. 3.4), although PTF and CRTS data do extend the SDSS baseline, their error properties decreases their utility in complementing the SDSS dataset. After all, the SDSS baseline extension afforded with PTF and CRTS is comparable to that achieved with PS1 data alone (Fig. 3.2).

Furthermore, to mitigate problems that could arise when applying photometric transformations (such as spurious variability due to incorrect offsets, or color-dependent variability), we choose to combine SDSS r with only PS1 r, since as Figs. 3.4, 3.5 show, SDSS r and PS1 r are sufficiently similar (at 1% level up to 20.5 mag) that no photometric transformation is required.

Finally, we clean the combined SDSS r - PS1 r quasar light curves using standard procedures of σ-clipping in magnitude and error space, and error-weighted day-averaging, to

Table 3.1: Color terms (photometric offsets) between CRTS, PTF, PS1 pass-bands and SDSS, using the SDSS mean $g-i$ color to spread the stellar locus. Thus the SDSS r synthetic magnitude, r_s, can be found as $r_s = x - B_0 - B_1(g-i)$. This linear trend is illustrated in Fig.3.4, where we plot $(x - r_{\rm SDSS})$ as a function of $(g-i)_{\rm SDSS}$ for $x = g_{\rm P1}, r_{\rm P1}, i_{\rm P1}$.

Band (x)	B_0	B_1
CRTS V	-0.0464	-0.0128
PTF g	-0.0294	0.6404
PTF R	0.0058	-0.1019
PS1 g	0.0174	0.6194
PS1 r	0.0065	-0.0044
PS1 i	0.0260	-0.2830

NOTE – To derive the color terms we used a subset of 100 000 stars randomly chosen from the SDSS standard stars catalog (Ivezić et al., 2007). To minimize scatter we selected bright stars with $r < 19$ mag.

mitigate the impact of bad photometry, and average out the intra-night variability (as in Charisi et al. 2016; Suberlak et al. 2017). Of 9248 SDSS-PS1 quasars, 8516 have the PS1 r data with 662 092 epochs. We remove points that have error departing from median SDSS(PS1) light curve segment by more than 7σ, and we visually inspect all photometry with magnitudes departing by more than 7σ from median magnitude. Of 585 flagged light curves, 253 required removal of individual epochs containing bad photometry. To avoid unphysically small errors, we add in quadrature 0.01 mag if combined nightly error is < 0.02 mag. In the final sample there are 580 321 epochs.

3.4 Simulations : lessons learned

We simulate the theoretical improvement of the DRW parameter retrieval in extended light curves. We generate long and well-sampled 'master' light curves, all with input $\tau = 575$ days,

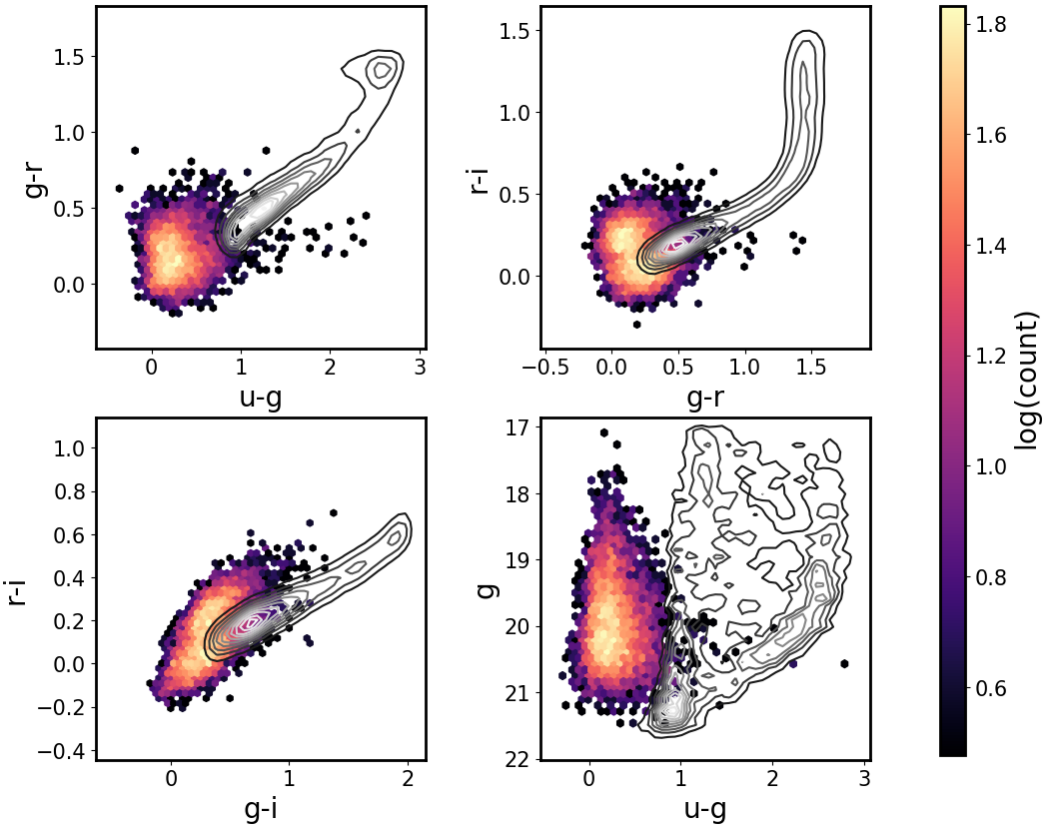

Figure 3.3: Regions of color-color (upper left, upper right, bottom left), and color-magnitude (bottom right) space occupied by SDSS S82 quasars (color) and stars (contours). We use quasar median photometry from Schneider et al. (2010), and standard stars catalog of Ivezić et al. (2007), showing a random subset of 10 000 stars. As seen in the bottom-left panel, quasars occupy a particular range of SDSS $g - i$ color. Therefore in fitting the linear color transformations we limit the color range to $-0.35 < (g - i) < 0.75$ (vertical dashed lines in Fig.3.4). Quasars also overlap other variable sources (eg. RR Lyrae), not shown here (Sesar et al., 2007).

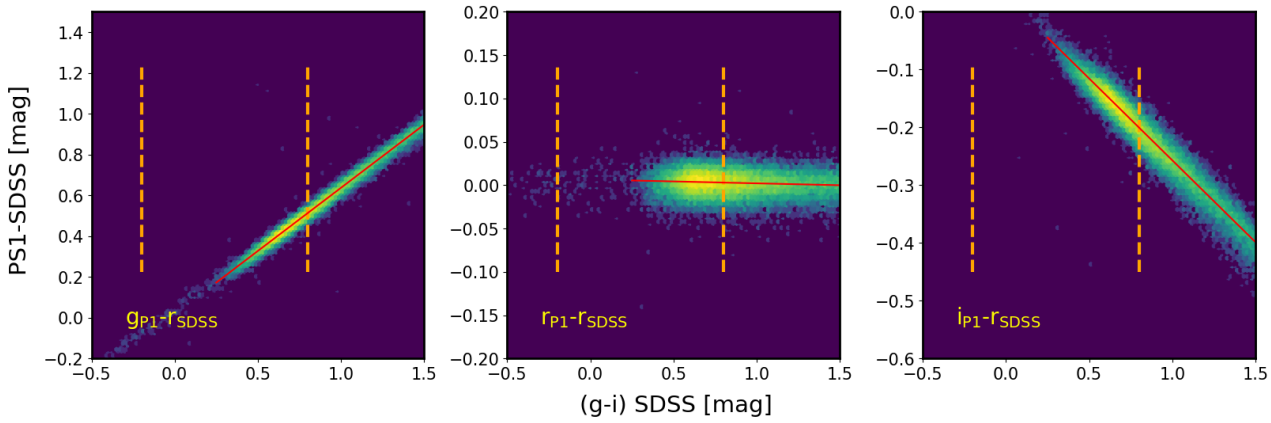

Figure 3.4: The SDSS-PS1 offsets, derived with the SDSS standard stars (Ivezić et al., 2007). From randomly chosen subset of 100 000 SDSS stars, 95 000 have PS1 DR2 data. To minimize scatter due to larger errors we select a subset of 40 000 stars with $r < 19$ mag. Vertical dashed lines mark the region in the SDSS color space occupied by quasars (see Fig. 3.3), used to fit the stellar locus with a first order polynomial, marked by the solid red line. The best-fit slopes: 0.619, -0.04, -0.283 for PS1 g, r, i, respectively, are listed as B_1 in Table 3.1.

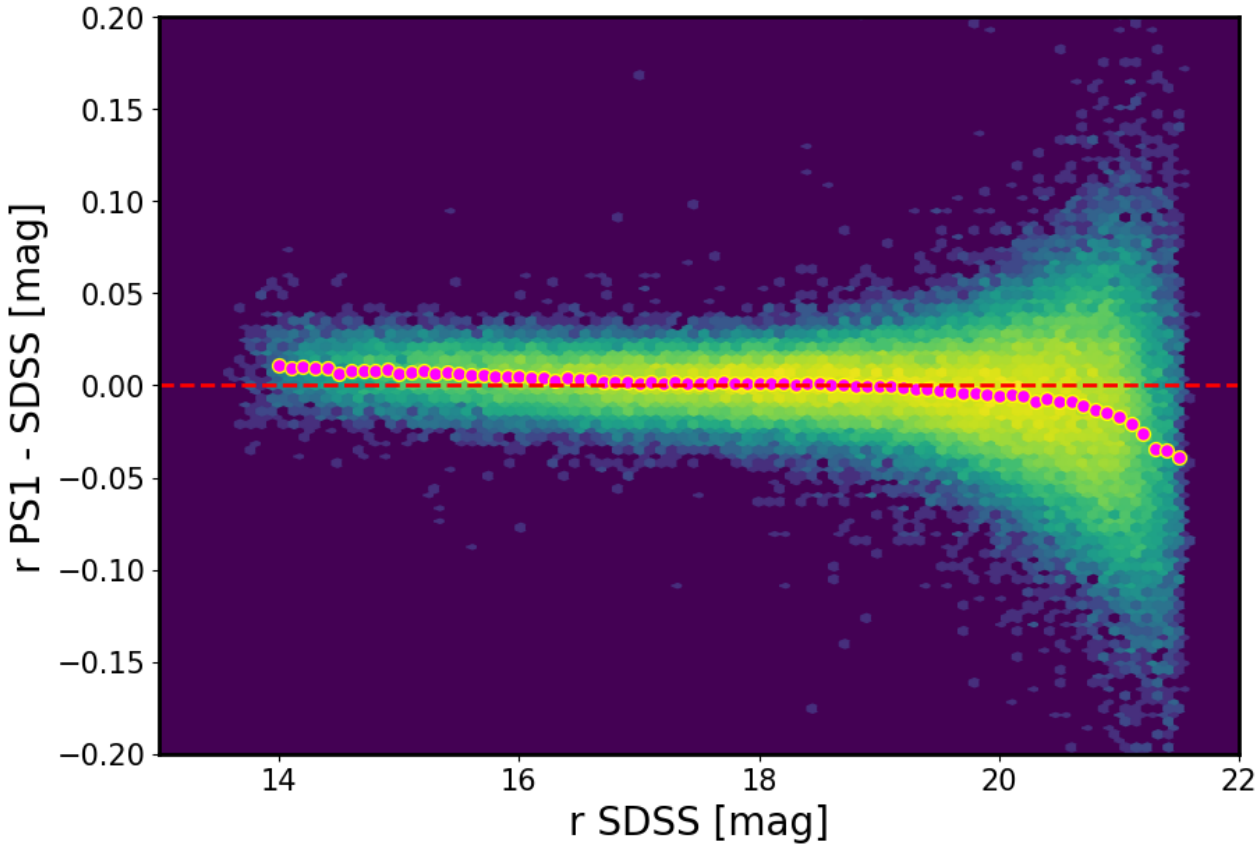

Figure 3.5: PS1 r vs SDSS r as a function of SDSS r for 100 000 randomly selected standard stars from Ivezić et al. (2007) catalog. Almost 95% of SDSS stars have PS1 DR2 photometry. The filled magenta circles represent the median offset - a slight slope at 1% (0.01 mag) level up to $r < 20.5$ mag.

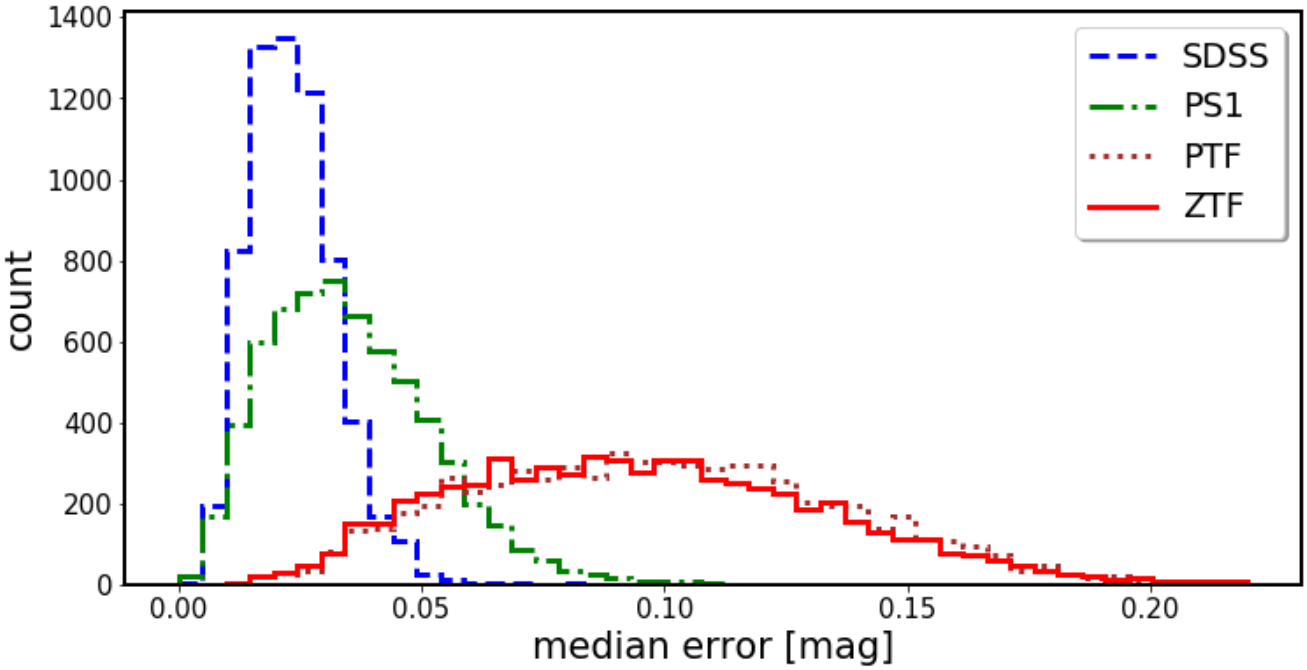

Figure 3.6: Distribution of median photometric uncertainties ('errors') in r-band real light curves. The PTF and ZTF segments have much larger errors than SDSS and PS1 due to shallower depth data. CRTS errors (not shown) are on average 50% larger than PTF.

$SF_\infty = 0.2$ mag (the median of S82 quasar distribution in M10), with 0 mean. We subsample at real observed epochs for SDSS and PS1, and at predicted cadences for ZTF and LSST (see Fig. 3.8). To each simulated light curve we add a magnitude offset corresponding to the mean of the combined SDSS-PS1 light curve. That way the magnitude distribution of simulated light curves is similar to that of the observed SDSS-PS1 data. For the LSST 10-year segment (finishing in 2031) we assumed 50 randomly distributed epochs per year, with the following error model:

$$\begin{aligned} \sigma_{LSST}(m)^2 &= \sigma_{sys}^2 + \sigma_{rand}^2 \ (\mathrm{mag})^2 \\ \sigma_{rand}^2 &= (0.04-\gamma)x + \gamma x^2 \\ x &= 10^{0.4(m-m_5)} \end{aligned} \tag{3.10}$$

with $\sigma_{sys} = 0.005$, $\gamma = 0.039$, $m_5 = 24.7$ (see Ivezić et al. 2019, Sec. 3.2). For the ZTF 1-year segment (Spring 2019 ZTF DR1 including the data from 2018) we assumed 120 observations (every three nights) in $g_{\rm ZTF}$ and $r_{\rm ZTF}$, deriving the magnitude-dependent error model by plotting best mag rms as a function of best median magnitude for ZTF matches to S82 standard stars in Fig. 3.7. We find that the LSST error model (Eq. 3.10) with $\gamma = 0.05$, $\sigma_{sys} = 0.005$, and $m_5 = 20.8$ adequately describes the ZTF photometric uncertainty.

To mirror observational conditions we add to the true underlying DRW signal a Gaussian noise, with variance defined by photometric uncertainties for corresponding surveys. Fig. 3.8 illustrates the simulated 'master' light curve (black dots, 4 per day), subsampled at SDSS (red), PS1 (green), and LSST (blue) cadence. While PS1 provides a 50% improvement of the SDSS baseline, LSST will nearly triple it. Fig. 3.9 shows how the simulated distribution of DRW parameters σ, τ, changes as the SDSS quasar light curves are extended with PS1, ZTF, and LSST data. In the future (after more data has been assembled and re-calibrated), ZTF will help, but not as dramatically as LSST. Note that ZTF, due to larger errors (Fig.3.6), causes a widening of the recovered τ distribution. Using PS1 data with its excellent deep photometry (as compared to ZTF or PTF) is the best improvement over existing SDSS

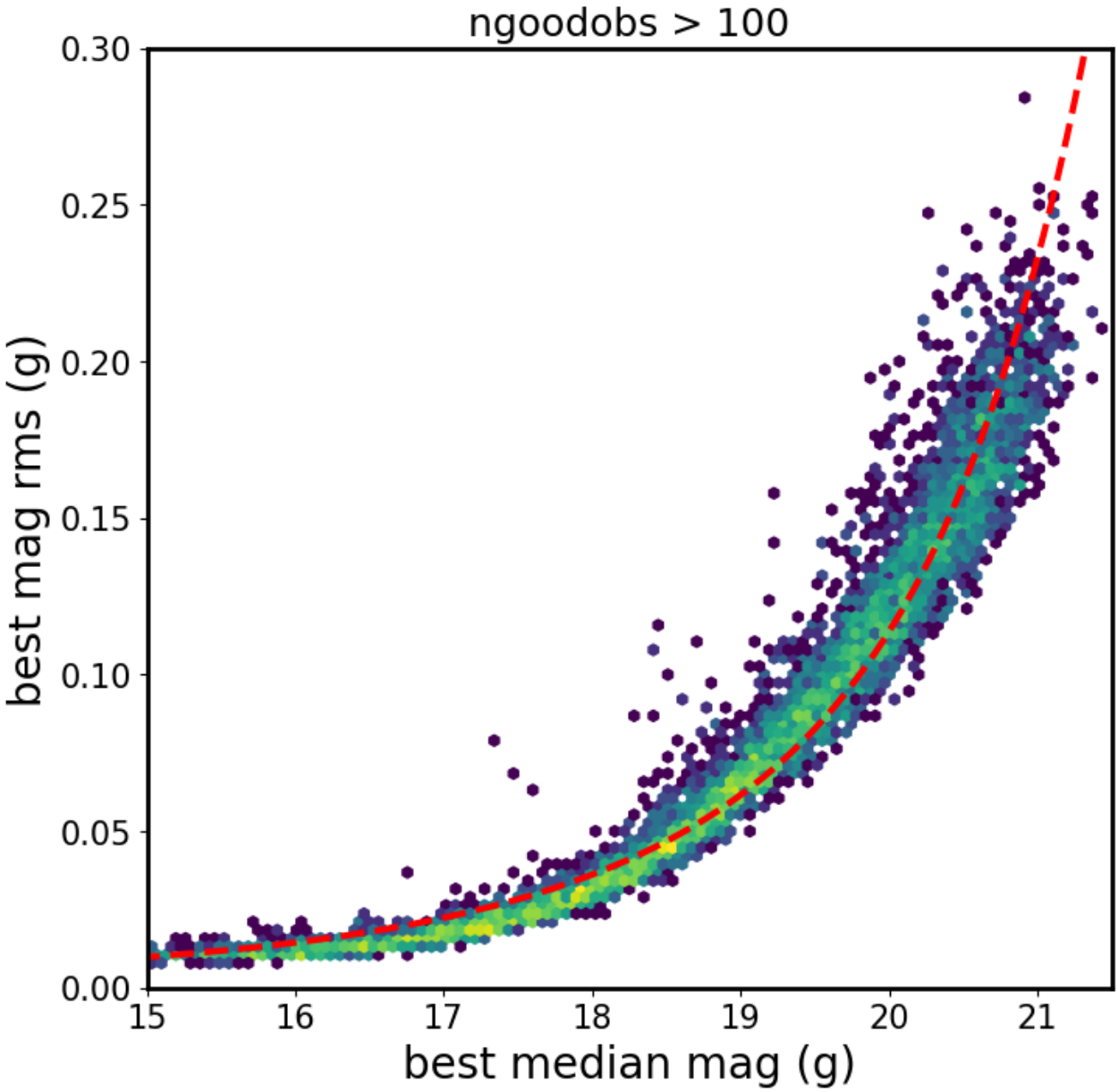

Figure 3.7: The best mag rms plotted as a function of magnitude for ZTF non-variable stars with over 100 observations. We overplot the adopted error model, with $\gamma = 0.05$, $\sigma_{sys} = 0.005$, and $m_5 = 20.8$ (see Eq. 3.10). Properties of ZTF photometric uncertainties are largely similar to the PTF uncertainties.

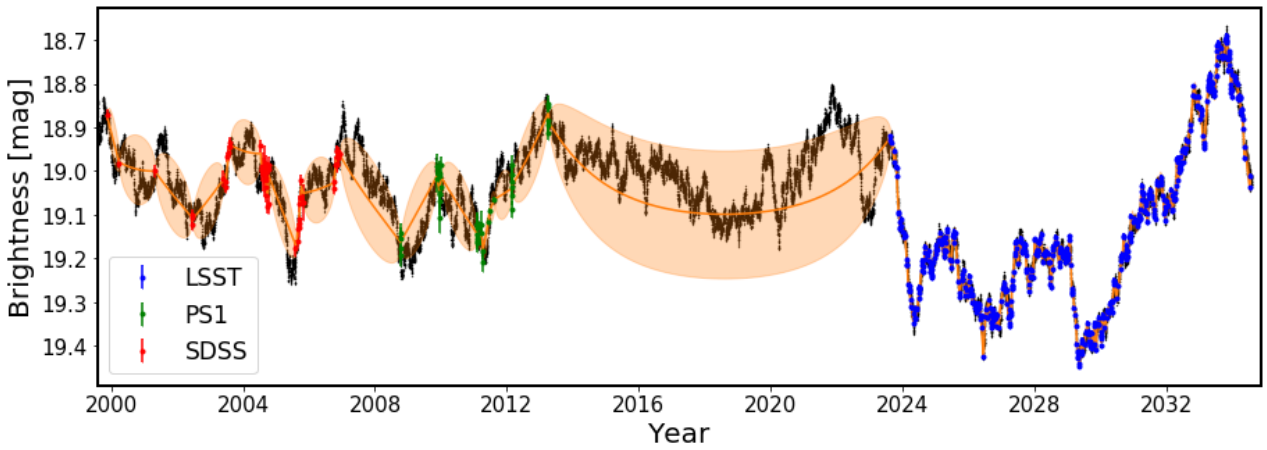

Figure 3.8: Simulated well-sampled underlying DRW process - one of 'master' light curves (τ = 575d, SF_∞ = 0.2 mag, 4 points per day) shown with small black dots. To simulate observations, the cadence is degraded (subsampled) to match the ground-based cadence corresponding to real quasar data from SDSS (red), PS1 (green) segments, and simulated LSST (blue) epochs (here we use SDSS-PS1 epochs for quasar dbID=3537034). The orange 'error snake' is an envelope marking the standard deviation of the fit to the data using a Gaussian process with DRW kernel (Sec. 3.4).

results. For this reason we use only SDSS-PS1 portion of quasar light curves, as the best trade-off between adding more baseline vs introducing more uncertainty with noisy data.

3.5 Results: variability parameters for S82 Quasars

We extend Stripe 82 quasar light curves by combining the SDSS r-band data with the PS1 r-band data, without any photometric offsets. For each quasar we fit the SDSS and SDSS-PS1 segments with the DRW model. This yields two sets of DRW parameters per quasar: $(\tau_{SDSS}, \sigma_{SDSS})$, and $(\tau_{SDSS-PS1}, \sigma_{SDSS-PS1})$. Because variability is inherent to the quasar, for the remaining analysis we shift all fitted timescales to quasar rest frame, and implicitly assume that the DRW timescales are considered in rest frame: $\tau_{\rm RF} = \tau_{\rm OBS}/(1+z)$.

In this section we first correct fitted τ, σ for wavelength dependence. Then we show consistency with M10 results, and consider the trends between DRW parameters and physical quasar properties: black hole mass $M_{\rm BH}$, absolute i-band magnitude M_i, or redshift z.

3.5.1 Comparison to M10

The DRW parameters recovered with celerite are broadly consistent with M10 - Fig. 3.10 shows the rest-frame τ, and SF_∞ distributions for our results for the SDSS segment (blue dashed contours), SDSS-PS1 combined light curves (green dot-dashed contours), and M10 SDSS for r-band only (red solid contours). When using exactly the same data as M10 (SDSS), our results agree. The offset of 0.05 dex between our and M10 results for SDSS, seen on the left panel of Fig. 3.11, can be attributed to data cleaning and software differences. The right panel of Fig. 3.11 shows the same distribution in terms of $K - \hat{\sigma}$ space, orthogonal to $\tau - \sigma$, where $\hat{\sigma} = \sigma\sqrt{2/\tau}$, and $K = \tau\sqrt{\sigma}2^{1/4}$.

3.5.2 Outliers: possible CLQSO candidates

Fig. 3.12 shows the change in recovered DRW parameters between SDSS and combined SDSS-PS1 light curves. The distribution of $f_\sigma \equiv \log_{10}(\sigma_{\rm SDSS-PS1}/\sigma_{\rm SDSS})$ and $f_\tau \equiv \log_{10}(\tau_{\rm SDSS-PS1}/\tau_{\rm SDSS})$

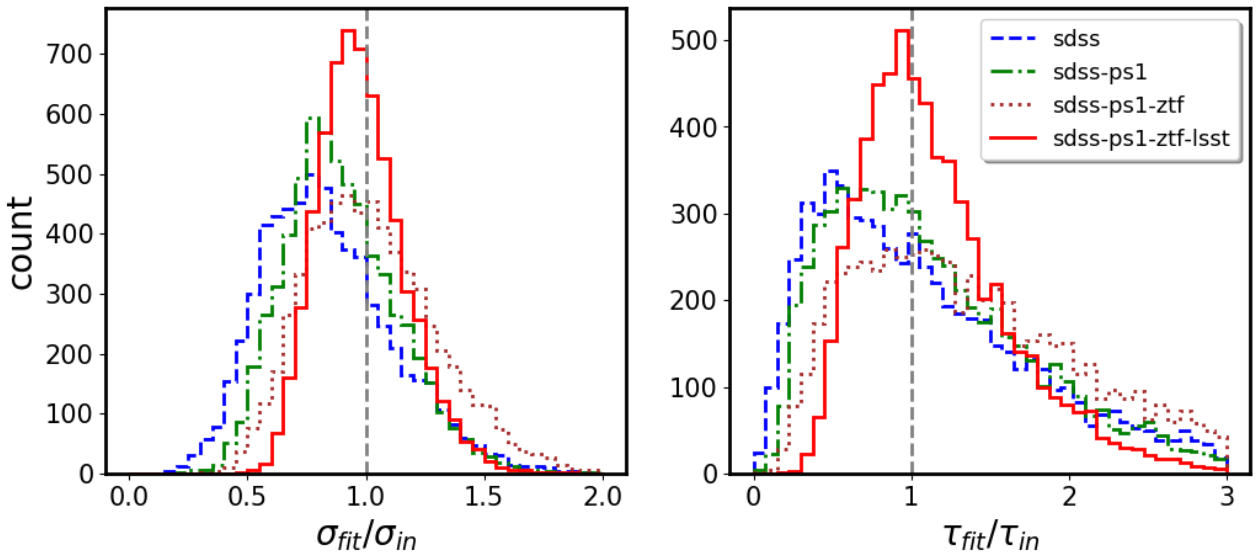

Figure 3.9: The ratio of DRW parameters fitted with celerite: τ and σ, to the input τ_{in} = 575 d, $\sigma_{in} = 0.2/\sqrt{2}$~0.14 ($SF_\infty$ = 0.2 mag). We simulated 9258 'master' light curves, and subsampled at real SDSS r-band or PS1 r-band cadence and photometric uncertainties, and simulated ZTF and LSST cadence. To simulate observing conditions, the underlying DRW signal was convolved with a Gaussian noise corresponding to epochal errors. For each light curve we start with SDSS segment only, and as we add more segments (PS1, ZTF, LSST), we refit for DRW model parameters with celerite. Thus each distribution corresponds to a different segment of simulated combined SDSS-PS1-ZTF-LSST light curves. Extending the baseline shifts the distribution of recovered DRW parameters towards unbiased regime - vertical dashed line marks input matching the output. This corresponds to the upper-right (well-constrained) portion of Fig. 3.1.

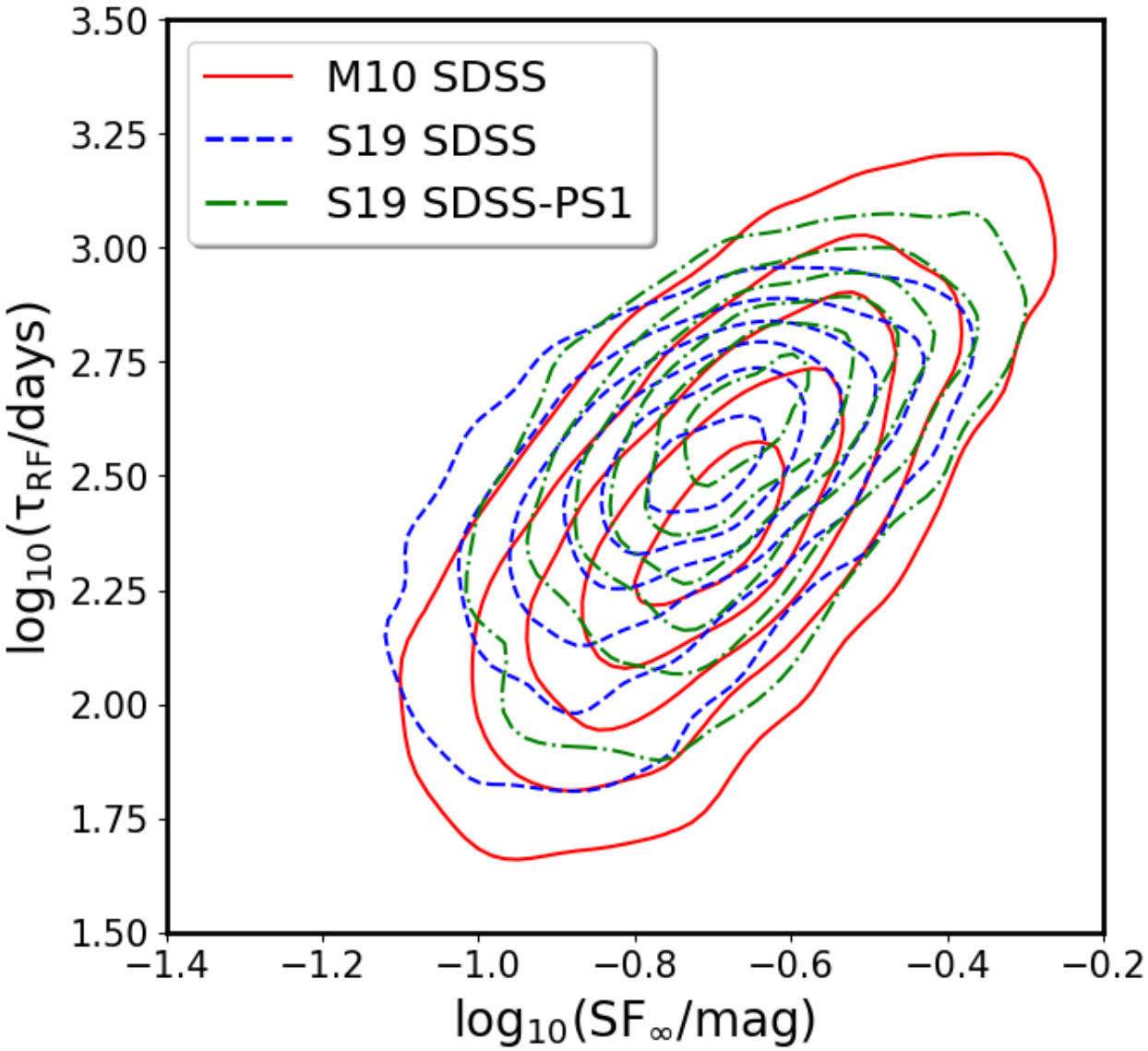

Figure 3.10: Comparison of distributions of the rest-frame variability timescale τ_{RF} against the asymptotic variability amplitude SF_∞, for M10 SDSS r-band (solid red) and celerite fits using SDSS (dashed blue) or SDSS-PS1 (dot-dashed green) segments of combined S82 quasar light curves. The timescales and SF_∞ from M10 and this work overlap, as we recover the same underlying distributions.

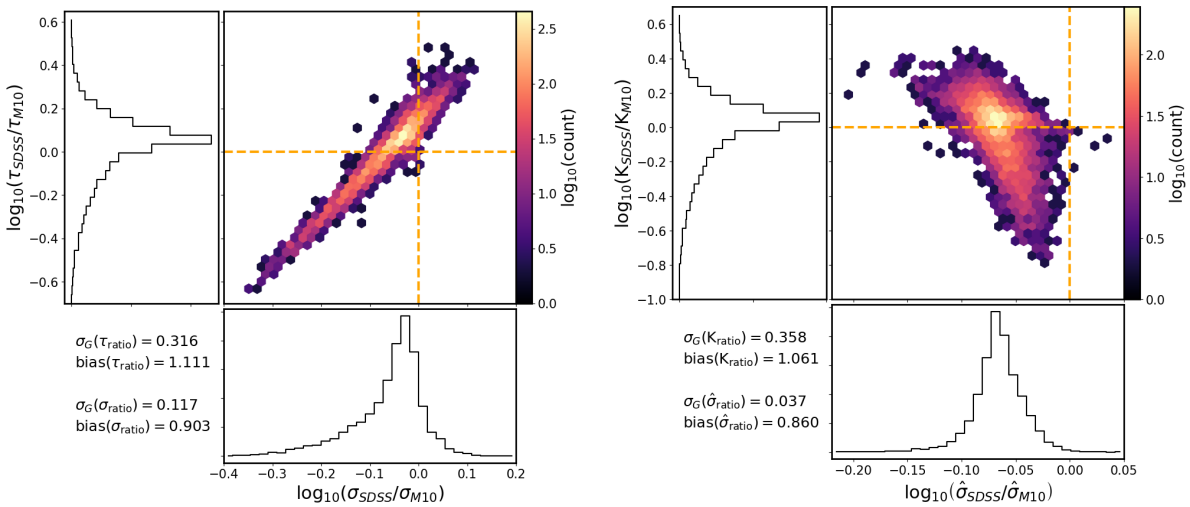

Figure 3.11: Comparison of celerite fits using only the SDSS r-band segments of S82 quasars ($\sigma_{\rm SDSS}, \tau_{\rm SDSS}$), against M10 results for SDSS r-band (σ_{M10}, τ_{M10}), object-by-object. The small offset (< 0.05 dex) can be attributed to data cleaning and software differences. See Fig. 3.10 for a comparison of rest-frame τ and SF_∞ distributions. This is similar to Fig.3 in M10, except we plot only the r-band SDSS results. The right-hand panel shows the comparison in an orthogonal $K - \hat{\sigma}$ space, where K is the direction along the diagonal on the left panel, and $\hat{\sigma}$-perpendicular to the diagonal. For this reason the right panel has a ten times smaller scatter along $\hat{\sigma}$ (0.037) than K (0.358).

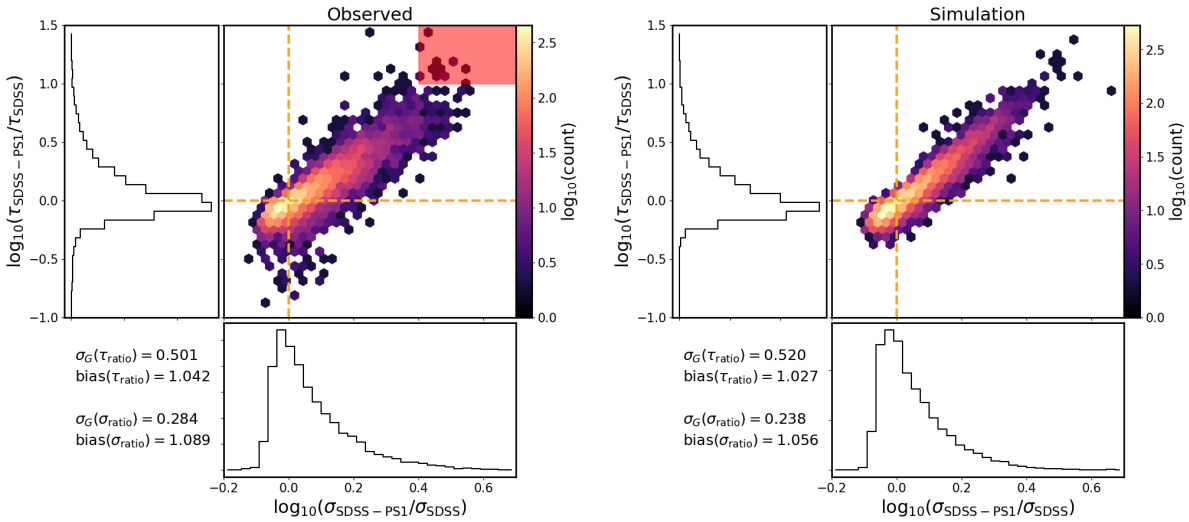

Figure 3.12: Ratios of fitted DRW parameters (τ, σ), comparing the value of parameter recovered using the combined light curve length (SDSS-PS1) to the shorter, SDSS-only light curve. The left panel shows the results for S82 quasars using real data, whereas the right panel shows the simulated quasars with realistic cadence, with $\tau_{in} = 575$ days and $\mathrm{SF}_\infty = 0.2$ (right). The general trend when using the real data (despite having a range of underlying timescales and amplitudes) is similar to that when using simulated data: the diagonal scatter is along the lines of constant $\hat{\sigma}$, and there is much less scatter in the perpendicular direction of K (see Fig. 3.13). There is no major change of shape of distribution as a function of mean quasar magnitude. The red rectangle marks the outliers with $\log(\tau_{\mathrm{SDSS-PS1}}/\tau_{\mathrm{SDSS}}) > 1$ and $\log(\sigma_{\mathrm{SDSS-PS1}}/\sigma_{\mathrm{SDSS}}) > 0.4$), discussed in Sec. 3.5.2.

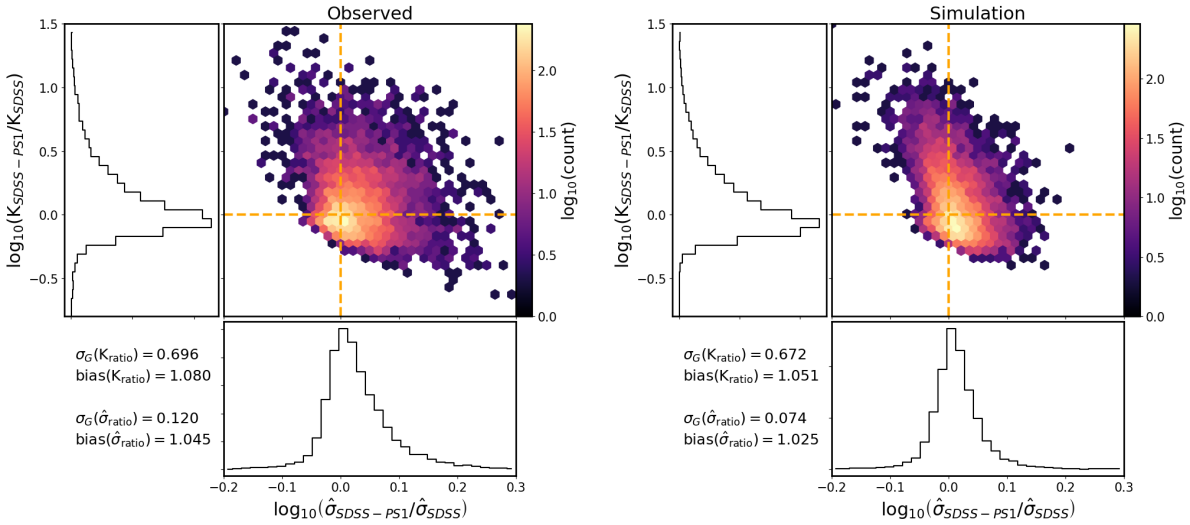

Figure 3.13: As Fig. 3.12, but in $K - \hat{\sigma}$ space, which is orthogonal to the $\tau - \sigma$ space, since $K = \tau\sqrt{\mathrm{SF}_\infty} = \tau\sqrt{\sigma}2^{1/4}$ and $\hat{\sigma} = \mathrm{SF}_\infty/\sqrt{\tau} = \sigma\sqrt{2/\tau}$.

for real light curves (left), matches the predicted distribution for simulated light curves (right). Studies show that about 0.1% quasars will exhibit large variability (in excess of 0.5 mag rms - see Fig.18 in MacLeod et al. 2012). Visual inspection of light curves in the upper-right region of the left panel of Fig. 3.12 (marked by the red rectangle) reveals large changes in brightness, similar to those seen in Changing-Look Quasars (Blanchard et al., 2017; Elitzur et al., 2014; Frederick et al., 2019; Gezari et al., 2017; Guo et al., 2016; LaMassa et al., 2015; Lawrence, 2018; MacLeod et al., 2019; Ross et al., 2018; Ruan et al., 2019, 2016; Runnoe et al., 2016; Schawinski et al., 2015; Shen et al., 2019; Sheng et al., 2017; Sheng et al., 2019; Stern et al., 2017, 2018; Trakhtenbrot et al., 2019; Yang et al., 2018). The light curves and properties of 40 CLQSO candidates for which $f_\tau > 1$, $f_\sigma > 0.4$, and $\langle r \rangle < 20.5$ are discussed in Appendix 3.B.

Such large differences in timescales and amplitude of variability can also be inferred directly from the light curves. Consider the difference in magnitude and scatter between the SDSS portion of the light curve (spanning approximately 10 years between 1998 and 2008),

and the PS1 portion (spanning ∼5 years between, 2009-2014 - see Fig. 3.2). We measure the median magnitudes offset as $\Delta(\text{median}) = \text{median}(SDSS) - \text{median}(PS1)$, and the scatter difference as $\Delta(\sigma_G) = \sigma_G(SDSS) - \sigma_G(PS1)$. The resulting distributions of Δ(median) and $\Delta(\sigma_G)$ for S82 quasars are shown in Fig. 3.14. Indeed, when plotting Δ(median) as a function of f_τ and f_σ there is a gradient indicating that the CLQSO candidates - outliers in (f_τ, f_σ) space, are also outliers in Δ(median)- $\Delta(\sigma_G)$ space. Thus the by-product of extending light curves to recalculate the DRW parameters with increased fidelity is an independent method to discover the CLQSO.

3.5.3 Rest-frame Wavelength Correction

Objects at cosmological distances are embedded in Hubble flow due to the expansion of the Universe (Riess et al., 2019). Therefore light observed from a distant quasar would have been emitted at shorter wavelength in quasar's rest-frame: $\lambda_{RF} = \lambda_{obs}/(1+z)$, where z is the cosmological redshift. Quasars at different redshifts probe different regions of rest-frame spectra (see Fig.7 in Shen et al. 2019). Thus before correlating the DRW parameters with quasar properties we correct σ, τ for the λ_{RF} dependence, studied by M10 with SDSS *ugriz* light curves. We plot in Fig. 3.15 the DRW parameters: SF_∞ and τ, as a function of λ_{RF}. A solid line marks the M10 best-fit power law trend:

$$f \propto \left(\frac{\lambda_{RF}}{4000\text{Å}}\right)^B \tag{3.11}$$

with $B = -0.479$ and 0.17 for SF_∞ and τ, respectively.

3.5.4 Trends with Black Hole Mass, Absolute Luminosity

In the era of large synoptic surveys such as ZTF or LSST, the large increase in the number of discovered quasars means that due to limited observational resources we will afford a spectroscopic follow-up for only a few percent of AGN with optical time-series (Ivezić et al., 2019). Therefore, a relationship between quasar variability parameters (τ, σ), and physical properties M_{BH}, M_i could provide an estimate of the latter for millions of quasars. We

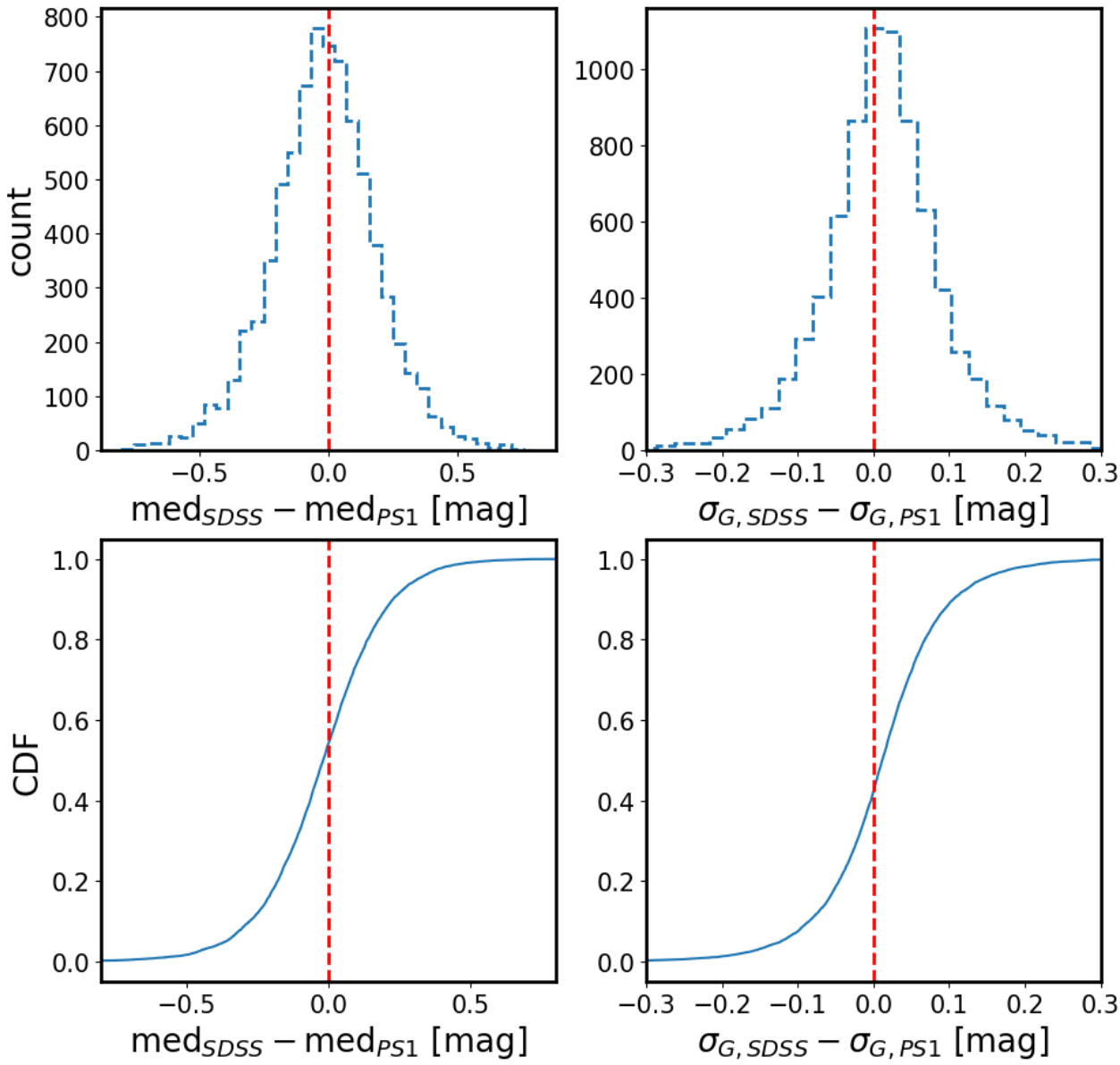

Figure 3.14: The differences between SDSS and PS1 segments of combined quasar r-band light curves. First, the difference between median SDSS and median PS1 portion, plotted as a histogram (upper-left panel), and cumulative distribution function (lower-left panel). Then, the difference between σ_G calculated for each portion of the light curve (σ_G is a robust estimate of the standard deviation, and is related to the difference between 75th and 25th percentile : $\sigma_G = 0.7413(Q_{75} - Q_{25})$). The outliers in the median offset space are also outliers in the DRW parameter space (eg. objects with $\log(\tau_{\mathrm{SDSS-PS1}}/\tau_{\mathrm{SDSS}}) > 1$ and $\log(\sigma_{\mathrm{SDSS-PS1}}/\sigma_{\mathrm{SDSS}}) > 0.4$) and $r > 20.5$ have $\Delta(\mathrm{median}) > 0.1$).

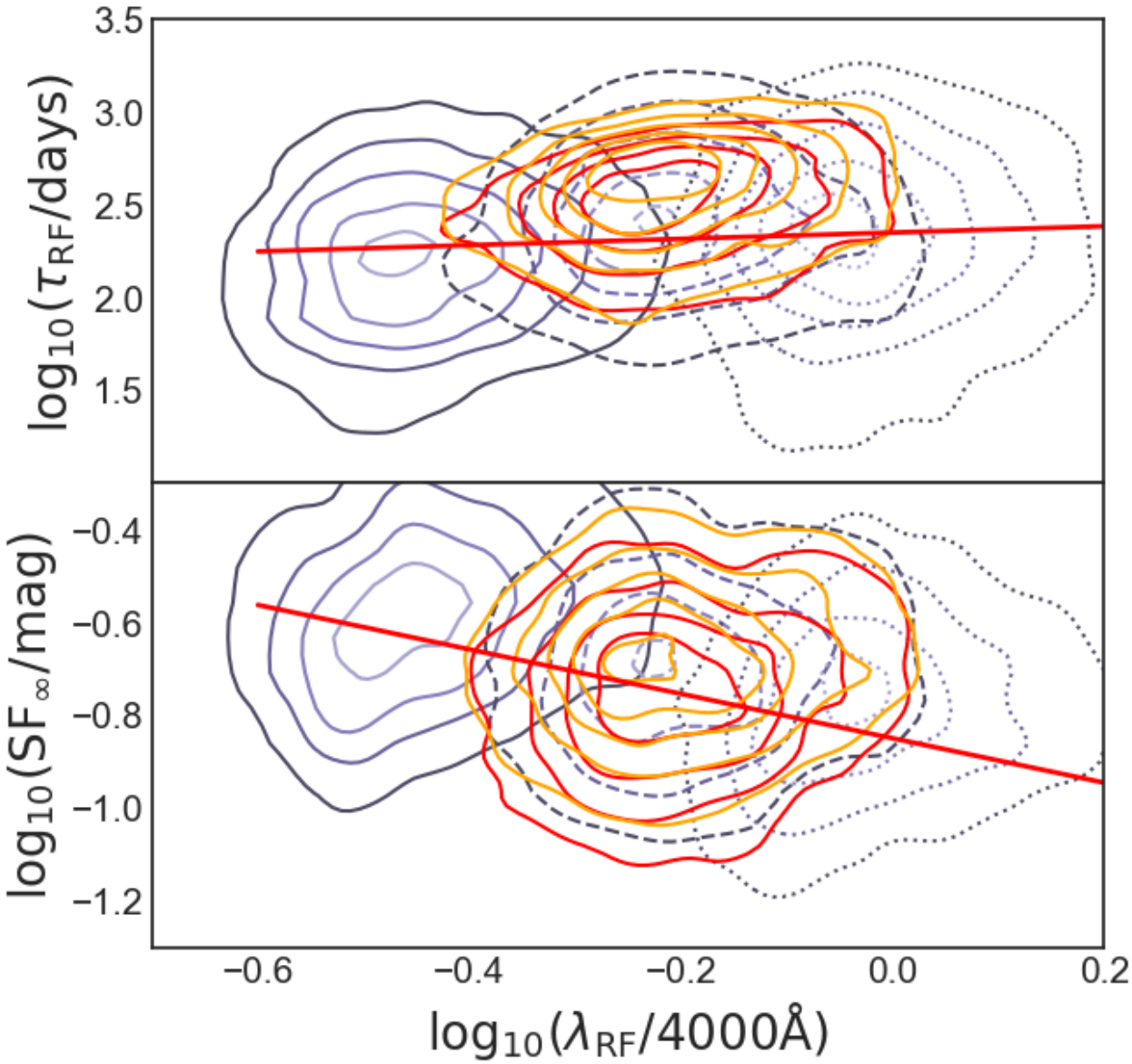

Figure 3.15: Rest-frame timescale τ (top panel), and asymptotic structure function SF_∞ (bottom panel), as a function of rest-frame wavelength λ_{RF}. The background contours show M10 SDSS urz data, and the foreground contours denote our results using SDSS (red) and SDSS-PS1 (orange) segments. The red line indicates the best-fit power law to M10 data, with $B = 0.17$ an -0.479 for τ_{RF}, and SF_∞, respectively. We take the center of each bandpass to approximate the observed wavelength: that is, for SDSS urz bandpasses, $\lambda_{obs} = 3520$, 6250, 9110 Å, respectively, and given the redshift of each quasar, find $\lambda_{RF} = \lambda_{obs}/(1+z)$.

inspect correlations between τ, σ and M_{BH}, M_i, using Shen et al. (2011) catalog, based on single-epoch SDSS spectra. M_i is K-corrected to $z = 2$, corresponding to the peak of quasar activity (Richards et al., 2006). For details, see Appendix 3.A.

On Fig. 3.16 we examine the distribution of M_{BH}, M_i, as a function of z for S82 quasars. The upward gradient in the top two panels reflects the selection effect that higher redshift quasars have to be brighter to be included in the magnitude-limited sample (luminosity-redshift degeneracy: see Sec.5, Fig.12 in M10, and Dong et al. 2018). Higher redshift quasars are also more active and have higher black hole masses due to cosmological downsizing (see Babić et al. 2007; Labita et al. 2009; McLure & Dunlop 2004). The distribution in the lower-left panel of Fig. 3.16 is peaked at $z = 2$ which corresponds to the peak of quasar activity.

Fig. 3.17 shows the DRW parameters for S82 quasars: τ and SF_∞, plotted as a function of quasar physical properties M_{BH}, M_i, and z. Upper-left and lower-left panels in Fig. 3.17 contain a gradient of SF_∞ with M_i,z - brighter quasars have lower variability amplitude, largely independent of black hole mass.

We investigate these relations in more detail by fitting f (τ or SF_∞) as a power-law function of M_{BH}, M_i, z :

$$\begin{aligned} \log_{10} f = \quad & A \quad + B \log_{10}\left(\lambda_{RF}/4000\text{Å}\right) + C(M_i + 23) \\ & + \quad D \log_{10}\left(M_{BH}/10^9 M_\odot\right) \end{aligned} \tag{3.12}$$

using a Bayesian linear regression method that incorporates measurement uncertainties in all latent variables (Kelly, 2007). This ansatz is identical to that used by M10, and very similar to the relation used by Scaringi et al. (2015), since for black holes their Eq.1 becomes $\log t_b = A' \log M_{BH} + B' \log L_{Bol} + C'$ where t_b is the PSD break timescale.

We compare the change in retrieved fit coefficients caused by adding PS1 data to SDSS against M10 SDSS-only study. Note that M10 fitted DRW model treating each of the 5 SDSS bands as a separate light curve, resulting in over thirty thousand values of τ, SF_∞ for nine

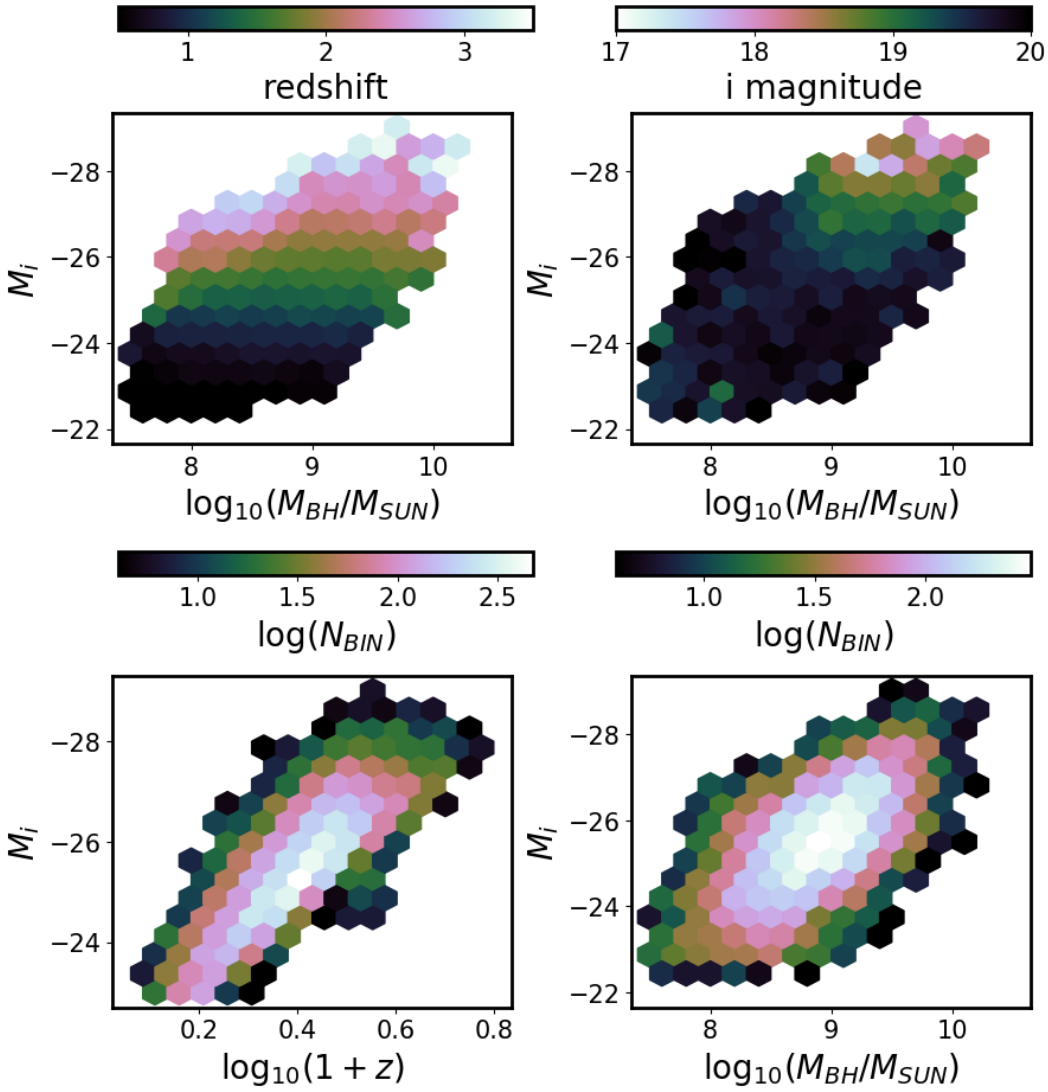

Figure 3.16: Distribution of quasars as a function of redshift, observed i-band magnitude, absolute i-band magnitude (K-corrected to z=2), and virial black hole mass. All data from Shen et al. (2011).

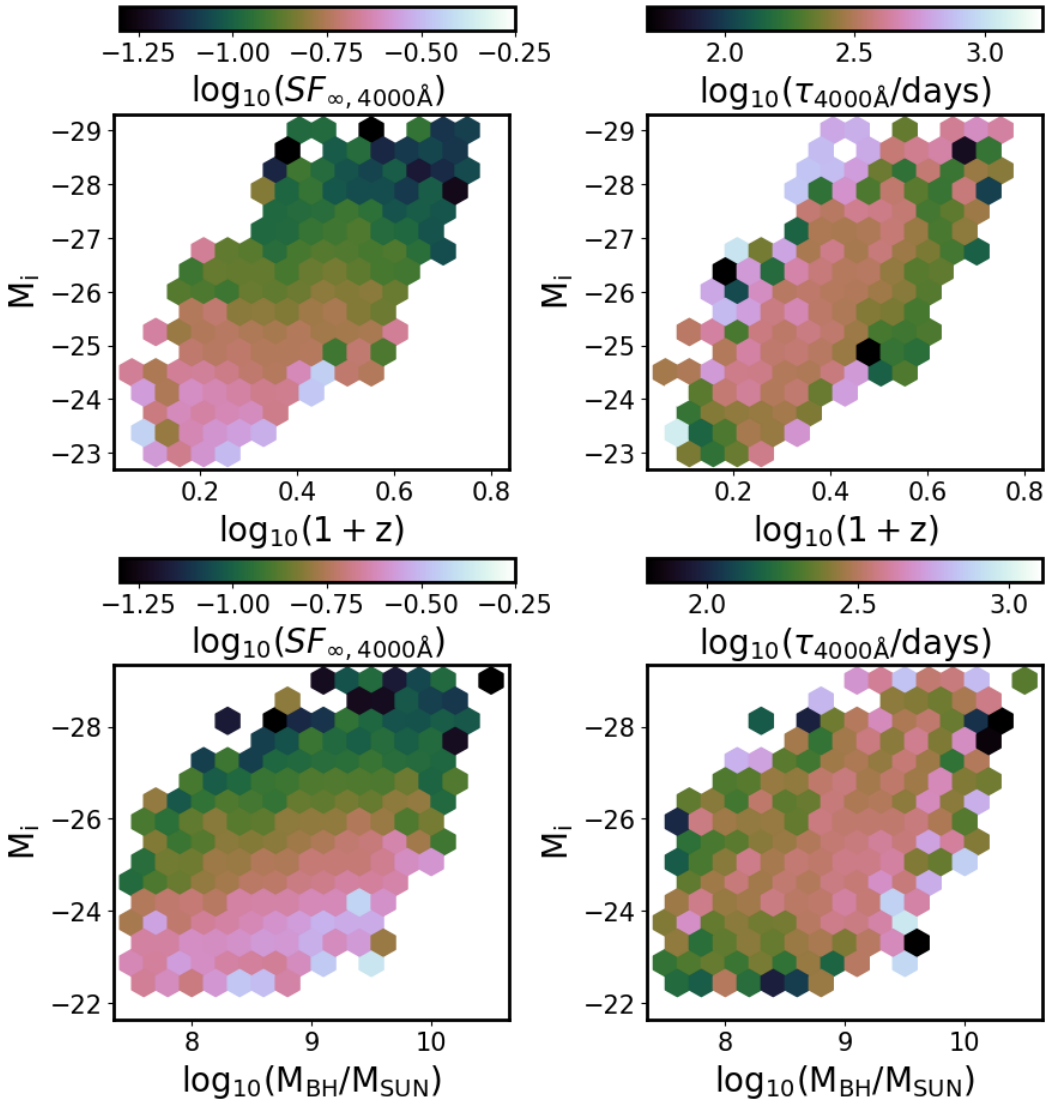

Figure 3.17: Long-term variability SF_∞, and characteristic timescale τ for SDSS-PS1 r band data, as a function of the absolute i-band magnitude ($M_i(z = 2)$, a proxy for bolometric luminosity), virial black hole mass M_{BH}, and redshift z, from Shen et al. (2011).

thousand S82 quasars. Grouping fitted quasar parameters by band, they were correlated to quasar physical parameters with Eq. 3.12. Fig. 3.18 shows the posterior samples for fitting Eq. 3.12 to $f = \mathrm{SF}_\infty$ for quasar data separately for each SDSS bandpass. Each band yields a slightly different fit coefficient. M10 reported as the fit result the band-mean (red vertical dashed line). Since this study uses only r band data, we compare the fit coefficients to M10 SDSS r data (green solid vertical line in Fig. 3.18). We show the results of fitting Eq. 3.12 to new SDSS and SDSS-PS1 parameters in Figs. 3.19, 3.20. First, with $f = \tau$ in Eq. 3.12 (Fig. 3.19), the SDSS-PS1 data confirms M10 for luminosity dependence (the posterior MCMC samples overlap), but the dependence of τ on M_{BH} is marginally weaker (by 0.007 dex). Second, in Fig. 3.20 with $f = \mathrm{SF}_\infty$, SF_∞ has a slightly weaker dependence on M_{BH} (by 0.04 dex compared to M10). The difference between celerite SDSS-only results and M10 can be attributed to data cleaning that was not performed by M10, and software differences. Each distribution from Figs. 3.19 and 3.20 is summarized in Table 3.2, with the uncertainty in A,C,D fit coefficients estimated from the standard deviation of the posterior samples.

We also searched for signal of Mg II variability (see Cackett et al. 2015 for a review). Mg II, like Hα and Hβ is a permitted low-ionization line (Yang et al., 2019), but on average is being emitted by gas further away from the ionizing source than the Hβ, possibly at the edge of the BLR (Guo et al., 2019). Ivezić et al. (2004) and MacLeod et al. (2012) studied the SDSS-POSS sample of quasars, and report the detection of a decrement in the data-model residuals around 2800 Å when plotting the residuals as a function of Δt and λ_{RF}. We investigated the data-model residuals for $f = \mathrm{SF}_\infty$ in Eq. 3.12. We find that using the SDSS data the decrement in the median(τ_{RF}) around 2800 Å is visible at $\sim 3\sigma$ relative to the smooth model, but adding the PS1 data the significance rises to $\sim 5\sigma$. We do not see much difference with regards to whether using the subset of 6371 quasars for which M10 had reliable results (listed in Table 3.2), versus the full set of 8516 quasars fitted with Celerite for DRW parameters. The effect is interesting, but does not produce very significant signal - see Appendix 3.C for more details.

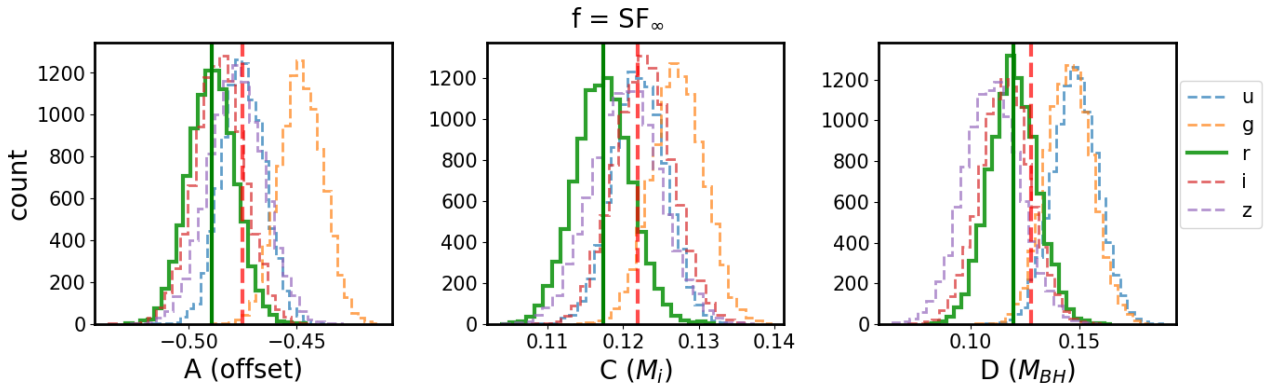

Figure 3.18: Posterior MCMC draws for fitting Eq. 3.12 with M10 variability amplitude SF_∞ against M_{BH}, M_i, z (Shen et al., 2011). Since M10 treated the near-simultaneous SDSS *ugriz* data for 9258 quasars, independently for each band, this resulted in DRW fit parameters for 7014 u, 7408 g, 6871 r, 6814 i, and 5111 z-band SDSS quasar light curves that fulfilled M10 quality of DRW fit selection criteria. M10 values for SF_∞ are corrected to 4000 Å using Eq. 3.11, with the power-law coefficient $B = -0.479$. Each distribution corresponds to a different SDSS band. We compare the results of fitting SDSS-PS1 r-band directly against M10 results for SDSS r-band (solid green). Note that Table 1 in M10 reported band-averaged values for A,C,D coefficients (vertical dashed red line), while we cite in Table 3.2 the mean for r-band (vertical solid green line).

Table 3.2: Comparison of best-fit coefficients for Eq. 3.12 using M10 results, and this work (S19). B is fixed to 0.17 or 0.479 from fitting a power law between λ_{RF} and τ, SF_∞ (see Fig. 3.15). Of 8516 quasars with SDSS-PS1 data, for consistency we use here an unbiased subset of 6371 quasars for which M10 had reliable results. For $f = \tau$, C is almost the same between M10 and this work for SDSS-PS1 (rows 1 and 3). However, D based on SDSS-PS1 data is larger than M10 by 0.01 dex (row 3). For $f = SF_\infty$, SDSS-PS1 based C is within 0.01 dex from M10 (rows 4,6), and D based on SDSS-PS1 data is almost identical to M10. When using celerite τ (σ) results for all 8516 quasars, the luminosity dependence is unchanged to within 0.01 dex, and the dependence on the black hole mass is stronger by 0.05 dex (0.02 dex), respectively.

f	Source	A(offset)	$B(\lambda_{RF})$	$C(M_i)$	$D(M_{\rm BH})$
τ	M10, SDSS	2.5 ± 0.027	0.17 ± 0.02	0.03 ± 0.009	0.178 ± 0.027
	S19, SDSS	2.515 ± 0.019	0.17 ± 0.02	0.042 ± 0.007	0.127 ± 0.019
	S19, SDSS-PS1	2.597 ± 0.02	0.17 ± 0.02	0.035 ± 0.007	0.141 ± 0.02
SF_∞	M10, SDSS	-0.486 ± 0.012	-0.479 ± 0.005	0.119 ± 0.004	0.121 ± 0.012
	S19, SDSS	-0.543 ± 0.009	-0.479 ± 0.005	0.125 ± 0.003	0.104 ± 0.008
	S19, SDSS-PS1	-0.476 ± 0.008	-0.479 ± 0.005	0.118 ± 0.003	0.118 ± 0.008

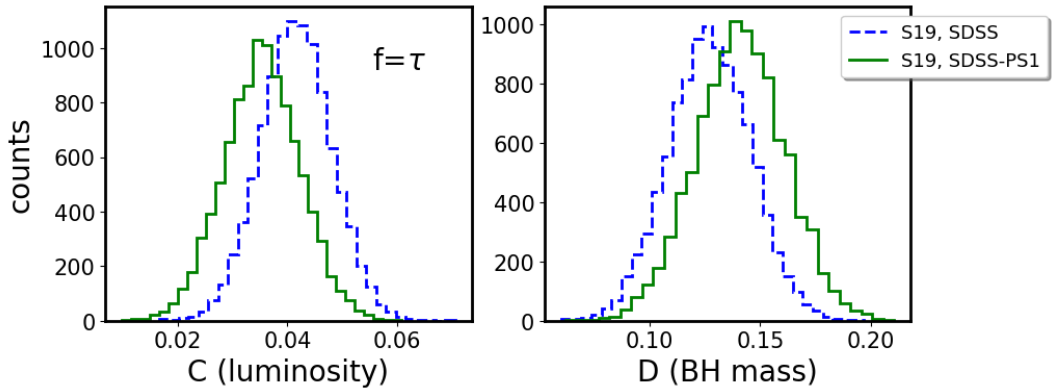

Figure 3.19: Distribution of MCMC posterior draws fitting Eq. 3.12 for characteristic timescale ($f = \tau$), based SDSS r-band results (dashed blue line), and new SDSS-PS1 combined r-band results (solid blue line). These are considered simultaneously as a function of quasar absolute magnitude M_i (left hand side panel), and black hole mass M_{BH} (right hand side panel). Of 9258 spectroscopically confirmed quasars in S82, we employed 8516 that had PS1 matches, of which 6371 fulfill M10 selection criteria (see M10, Sec 2.2). The results from SDSS-PS1 light curves are consistent with M10 for the SDSS r band.

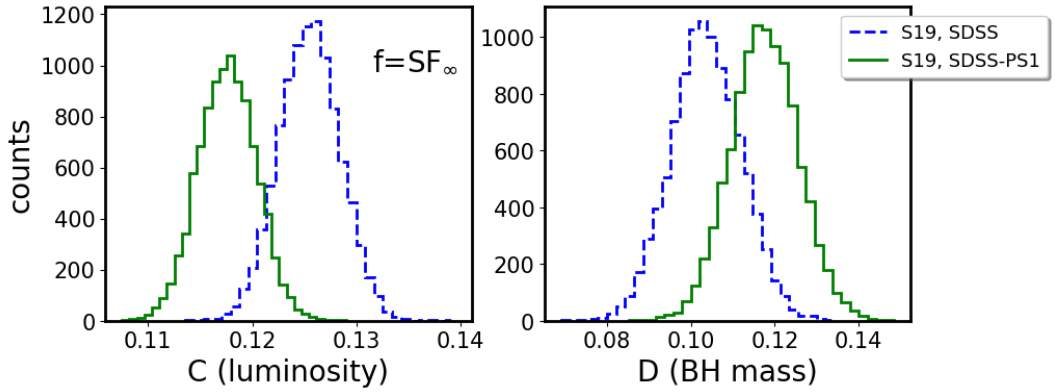

Figure 3.20: Same as Fig. 3.19, but fitting the DRW asymptotic amplitude ($f = \mathrm{SF}_\infty$ in Eq. 3.12) as a function of absolute magnitude M_i, and black hole mass M_{BH}. New data from PS1 is consistent with earlier results of M10 on luminosity dependence, but supports slightly weaker dependence of SF_∞ on M_{BH} (by 0.06 dex).

3.5.5 Comparison to other studies: Eddington ratio

Eddington ratio (f_{Edd}=$L_{Bol}/L_{\rm Edd}$) encodes accretion strength: the proximity of quasar bolometric luminosity to the theoretical Eddington limit, where $L_{\rm Edd}$=$1.26\times10^{38}(M_{\rm BH}/M_{\odot})$ erg/s (Shen et al., 2011). Since τ and SF_{∞} depend on M_i and M_{BH}, we investigate the possibility of the Eddington ratio being the driver of these observed trends. On Fig. 3.21 we show f_{Edd} as a function of M_i, M_{BH}, and SF_{∞}. The left two panels depict f_{Edd} and SF_{∞} binned as a function of M_i and M_{BH}. Third panel shows the quasar counts, and the fourth panel the bin means (black dots). The means are further binned along f_{Edd} (as in M10). Combined SDSS-PS1 data supports SF_{∞} being inversely related to f_{Edd}, with power-law slope of -0.208 ± 0.03, consistent with -0.23 ± 0.03 reported by M10. Observations are generally consistent with basic predictions from Table 3.3: A increases with increasing $M_{\rm BH}$, and f_{Edd} decreases as L_{Bol} increases, while τ increases with L_{Bol}. No model from Table 3.3 is rejected.

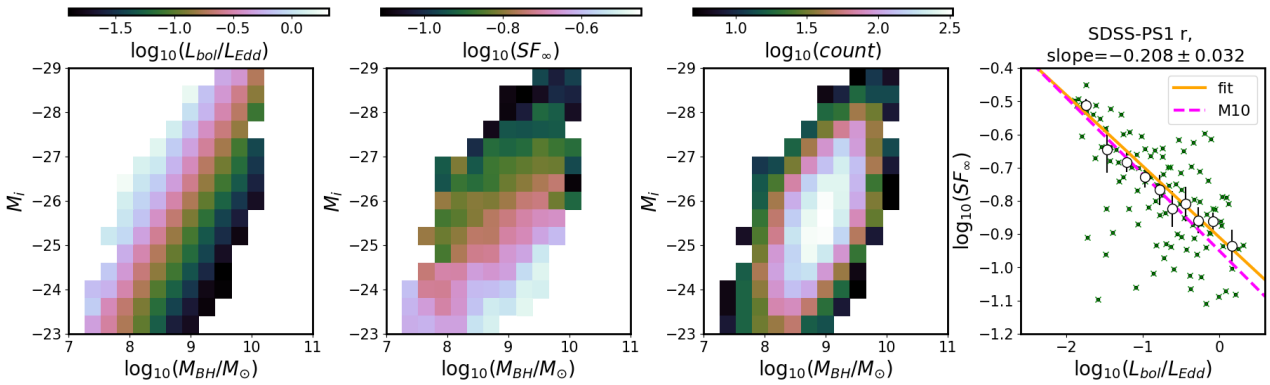

Figure 3.21: First panel: Eddington ratio $f_{Edd} = L/L_{Edd}$ plotted as a function of M_{BH} vs M_i(from Shen et al. 2011). Second panel: SF_∞ binned on the same grid of M_{BH}, M_i. Third panel: counts of quasars in each bin. We only plot bins with $N > 5$ quasars. Fourth panel : the median f_{Edd} and SF_∞ per bin are plotted as green crosses. These are aggregated along f_{Edd}, in 10 bins of such width that each has the same number of points. Open circles mark the median SF_∞ per f_{Edd} bin, with errors defined as $\sigma_y = 1.25\sigma_G(bin)/N$, where N is number of points per bin, and σ_G is the robust estimate of the standard deviation ($\sigma_G = 0.7413(Q_{75} - Q_{25})$). We assume the uncertainty along f_{Edd} as $\sigma_x = w/\sqrt{12}$, with w denoting the bin width (Ivezić et al., 2014). The solid orange line is the best-fit slope: -0.208 ± 0.032, with the slope uncertainty estimated from the standard deviation of the posterior samples. The best-fit slope agrees with M10 results (-0.23 ± 0.03), plotted as dashed magenta line.

Table 3.3: Theoretical predictions of various models concerning amplitude A and/or timescale τ of variability, Eddington ratio f_{Edd}, bolometric luminosity L_{Bol}, and accretion rate $\dot{m}$. Observations cannot reject any model.

Theory	Prediction
Standard thin disk (Netzer, 2013; Shakura & Sunyaev, 1973)	$A \nearrow$ as $f_{Edd} \searrow$, $\tau \propto L_{Bol}^{1/2}$ (inCaplar et al. 2017)
Strongly inhomogeneous disk (Dexter & Agol, 2011) / local fluctuations (Cai et al., 2018)	$A \nearrow$ as $f_{Edd} \searrow$
Variations in global accretion rate (Hawkins, 2007; Li & Cao, 2008; Zuo et al., 2012)	$A \nearrow$ as $M_{BH} \nearrow$, $A \searrow$ as $L_{Bol} \nearrow$
Eddington ratio reflecting AGN age (Hopkins et al., 2005; Martini & Schneider, 2003)	$A \nearrow$ as $f_{Edd} \searrow$
X-ray reprocessing (Kubota & Done, 2018)	$A \nearrow$ as $f_{Edd} \searrow$, and $A \nearrow$ as $\dot{m} \searrow$

Table 3.4: Comparison of published results on correlating observed quasar light curve properties: a measure of variability amplitude A, and timescale of variability τ, against the physical quasar properties: Black Hole mass $M_{\rm BH}$, Eddington ratio $f_{\rm Edd}$, bolometric luminosity L_{Bol}. We list the correlations as positive (+), negative (–, i.e. anti-correlation), not found (0), or not studied (x). We further note if the correlation is strong (!), or weak ($\sim$).

Publication	Measure of amplitude / timescale	A vs			τ vs		
		$M_{\rm BH}$	$f_{\rm Edd}$	$L_{\rm Bol}$	$M_{\rm BH}$	$f_{\rm Edd}$	$L_{\rm Bol}$
Wilhite et al. (2008)	SF (ensemble study)	+	-	-	x	x	x
Kelly et al. (2009)	DRW: τ-decorrelation timescale, σ	–	0	!–	+	x	$\sim$+
MacLeod et al. (2010)	DRW: τ,σ	+	!–	!–	+	x	$\sim$+
Morganson et al. (2014)[a]	SF: A, γ	x	x	-	x	x	+
Kozłowski et al. (2016)	DRW: τ,σ	x	x	!–	+	x	$\sim$+
Simm et al. (2016)[b]	EV , and PSD (break timescale)	0	–	–	0	0	0
Caplar et al. (2017)	SF: τ,σ	$\sim$	x	!–	x	x	+
Rakshit & Stalin (2017)[c]	DRW: τ,σ	+	!-	$\sim$+	x	x	x
Sun et al. (2018)[d]	SF: τ,σ	x	x	!–	x	+	!+
Li et al. (2018)[e]	SF: A, γ	$\sim$+	–	–	+	x	+
Sánchez-Sáez et al. (2018)[f]	SF: A	0	–	x	x	x	x
This work	DRW: τ-decorrelation timescale, σ	+	!–	+	+	x	$\sim$+

a: 105 783 quasars with SDSS-PS1 sparse data, SF parametrized as $V(\Delta t|A,\gamma) = A(\Delta t/1\text{year})^{\gamma}$: γ is the increase of SF with time lag Δt

b: 90 X-ray selected AGN, PS1 optical data. Variability characterized by normalized excess variance (EV). PSD characterized by break timescale.

c: Narrow- and Broad-Line Seyfert1 AGN, $z < 0.8$, CRTS data (> 50 epochs, $5 - 9$ years baseline), DRW fitted with JAVELIN (Zu et al., 2011), no timescale correlations considered due to short baseline.

d: 1004 SDSS quasars with $0.5 \leq z \leq 0.89$

e: 119 305 sparse quasar light curves from DECaLS and SDSS. SF as in Morganson et al. (2014) above.

f: 1348 QUEST-La Silla quasars, only amplitude of variability and excess variance.

3.6 Discussion

3.6.1 Trends with Eddington ratio

Anticorrelation of variability amplitude with Eddington ratio has a variety of possible theoretical explanations. In the thin disk theory (Frank et al., 2002; Netzer, 2013; Shakura & Sunyaev, 1973), radius of the emission region at given wavelength increases with Eddington ratio, and is inversely proportional to temperature (Rakshit & Stalin, 2017). Thus a hotter disk means that the emission observed in a given bandpass is emitted from a larger radius. From causality, a smaller region can be more variable than a larger one. Therefore, a hotter disk would be less variable at a given wavelength than a colder one, and the variability amplitude as studied in a particular bandpass (here, SDSS r-band) would be anticorrelated with Eddington ratio (Edelson et al., 2015; Fausnaugh et al., 2016).

On the other hand, in the strongly inhomogeneous disk model independent temperature fluctuations in N zones drive the variability (Dexter & Agol, 2011). In that framework the inverse trend of variability amplitude against L/L_{Edd} and L_{Bol} can be understood qualitatively if more luminous quasars also have higher mass accretion rate, and thus greater number of disk inhomogeneities, resulting in smaller flux variability (Simm et al., 2016). The inhomogeneous disk model was consistent with mean SDSS spectral analysis in Ruan et al. (2014), but was not a preferred explanation for Kokubo (2015).

Both Rumbaugh et al. (2018) (with Dark Energy Survey structure function study) and Sun et al. (2018) (with a low-z subsample of S82 SDSS quasars) confirm the anti-correlation between quasar variability and luminosity. However, Graham et al. (2019) do not find support for this trend with the sample of extremely variable quasars (EVQs) in the CRTS dataset, but when selecting for lower luminosity sources ($M_V < -23$), the anti-correlation is recovered. This agrees with an interpretation that a dwindling fuel supply may correspond to higher variability. Furthermore, Sánchez-Sáez et al. (2018) combined the SDSS spectra with 5 year light curves of 2345 quasars obtained with Quasar Equatorial Survey Team (QUEST)-La Silla AGN Variability Survey, and using the Bayesian parametrization of Structure Function

(Schmidt et al., 2010) they also found that the amplitude of variability A is anti-correlated with rest-frame emission wavelength, and Eddington ratio (also see Simm et al. 2016, Rakshit & Stalin 2017).

Indeed, f_{Edd} is a proxy for the strength of accretion, which together with orientation may be the key to explaining quasar main sequence (QMS: Marziani et al. 2018; Shen & Ho 2014). The QMS, defined by so-called Eigenvector-1, is the anti-correlation between the broad line FeII emission, and the strength of the narrow OIII (5007 Å) line (Wang et al., 1996). An analysis of quasar clustering by Shen & Ho (2014), later confirmed by Sun & Shen (2015) with measurements of black hole mass from the quasar host galaxy stellar dispersion (Ferrarese & Merritt, 2000; Kormendy & Ho, 2013), showed that the entire diversity of quasars in QMS can be explained by the variation in accretion (affecting $R_{\rm FeII}$ - the ratio of the FeII EQ between $4435 - 4685$ Å and Hβ), or orientation effects (affecting the FWHM of the Hβ). However, Panda et al. (2019a,b) found that these are insufficient, and variations in metallicity, as well as a range of cloud densities, and turbulences are required. Jiang et al. (2016) also found that metallicity, and in particular the iron opacity bump, may have a strong influence on the stability of an accretion disk, and thus linking metallicity to AGN variability. This is also consistent with findings of Sun et al. (2018): quasars with high FeII strength have higher metallicity, and have more stable disks.

3.6.2 Variability Timescales

In the era of changing-look active galaxies (including initially distinct classes of Changing-Look Quasars (LaMassa et al., 2015; MacLeod et al., 2019), Changing-Look AGNs (Bianchi et al., 2009; Marchese et al., 2012; Risaliti et al., 2009), Changing-Look LINERS (Frederick et al., 2019) to name a few) there is a revived interest in possibly linking the behavior of stellar-sized accreting systems (eg. Black Hole Binaries), to that of galactic scale (eg. AGN, QSO, LINERS - Noda & Done 2018; Ruan et al. 2019).

Several relevant timescales are involved, and there are various interlinked mechanisms that could drive the variability. A standard optically thick, geometrically thin, α-disk model

has a hierarchy of timescales: dynamical, thermal, front, viscous, with $t_{dyn} < t_{th} < t_{front} < t_{visc}$ (Frank et al., 2002; Netzer, 2013). We proceed to describe briefly each timescale, concluding with our interpretation of the mechanism that could drive the variability observed from the data.

The dynamical, or gas orbital, timescale is simply an inverse of the Keplerian orbital angular frequency Ω at radius R:

$$t_{dyn} \sim 1/\Omega = \left(\frac{GM}{R^3}\right)^{-1/2} \tag{3.13}$$

The main parameter describing the accretion disk is α - the ratio of the (vertically averaged) total stress to thermal (vertically averaged) pressure:

$$\alpha = \frac{\langle \tau_{r\varphi} \rangle_z}{\langle P \rangle_z} \tag{3.14}$$

After Lasota (2016), the hydrodynamical stress tensor (corresponding to kinematic viscosity ν) is:

$$\tau_{r\varphi} = \rho\nu \frac{\partial v_\varphi}{\partial R} = \rho\nu \frac{d\Omega}{d \ln R} = \frac{3\rho\nu\Omega}{2} \tag{3.15}$$

so with c_s - local sound speed at radius R (isothermal sound speed is $c_s = \sqrt{P/\rho}$),

$$\alpha = \frac{3\rho\nu\Omega}{2P} = \frac{3\Omega\nu}{2c_s^2} \tag{3.16}$$

This means that smaller α corresponds to less viscous disks.

The thermal timescale, related to the time needed for re-adjustment to the thermal equilibrium (derived in detail in Frank et al. 2002), is the ratio of heat content per unit disk area to dissipation rate per unit disk area: $(dE/A)/(dE/dt/A) = dt$. The heat content per unit volume is $\sim \rho kT/\mu m_p \sim \rho c_s^2$, and heat content per unit area is $\sim \rho c_s^2/h \sim \Sigma c_s^2$. Meanwhile, the dissipation rate per unit area, $D(R)$, is

$$D(R) = \frac{9}{8}\nu\Sigma R^{-3} GM \tag{3.17}$$

(eq. 4.30 in Frank et al. 2002), so :

$$t_{th} \sim \frac{c_s^2 R^3}{GM\nu} = \frac{c_s^2}{\nu\Omega} = \frac{t_{dyn}}{\alpha} \tag{3.18}$$

Thus if the disk were inviscid ($\nu \to 0$), then $t_{th} \to \infty$ i.e. there is no contact with adjacent disk elements.

The cooling and heating fronts propagate through the disk at αc_s (Hameury et al., 2009) - in that description with no viscosity there is no communication between neighboring disk annuli, and thus no front propagation (Balbus, 2003; Balbus & Hawley, 1998). Following Stern et al. (2018), if we define as h/R the the disk aspect ratio, with the disk height $h = c_s/\Omega$, the characteristic time for front propagation is:

$$t_{front} \sim (h/R)^{-1} t_{th} \tag{3.19}$$

The viscous timescale is the characteristic time it would take for a parcel of material to undergo a radial transport due to the viscous torques from the radius R to the black hole (Czerny, 2006). Note that while viscosity has probably magnetic origin (Eardley & Lightman, 1975; Grzędzielski et al., 2017), in this simplistic order of magnitude estimate we use a hydrodynamical description of accretion flow. With $\nu = \eta/\rho$ (kinematic viscosity being the ratio of dynamical viscosity to density), Frank et al. (2002) shows (Chap.5.2) that

$$t_{visc} \sim R^2/\nu \sim R/v_R = (h/R)^{-2} t_{th} \tag{3.20}$$

We can parametrize each timescale for a black hole mass $M_{BH} = 10^8 M_\odot$, at $R \sim 150 r_g$, with the gravitational radius $r_g = GM_{BH}/c^2 \sim 4$au, using Eqs.5-8 in Stern et al. (2018) :

$$t_{dyn} \sim 10\text{days} \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right) \left(\frac{R}{150 r_g}\right)^{3/2} \tag{3.21}$$

$$t_{th} \sim 1\,\text{year} \left(\frac{\alpha}{0.03}\right)^{-1} \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right) \left(\frac{R}{150 r_g}\right)^{3/2} \tag{3.22}$$

$$t_{front} \sim 20\,\text{years} \left(\frac{h/R}{0.05}\right)^{-1} \left(\frac{\alpha}{0.03}\right)^{-1} \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right) \left(\frac{R}{150 r_g}\right)^{3/2} \tag{3.23}$$

$$t_{visc} \sim 400\,\text{years} \left(\frac{h/R}{0.05}\right)^{-2} \left(\frac{\alpha}{0.03}\right)^{-1} \left(\frac{M_{\rm BH}}{10^8 M_\odot}\right) \left(\frac{R}{150 r_g}\right)^{3/2} \tag{3.24}$$

In summary, of considered timescales only thermal and dynamical are short enough to be related to the observed short-term stochastic variability. It may be that the variability on the scale of days is driven by local changes, and on the longer scale (perhaps hundreds of days) by a different mechanism (Kokubo, 2015). The other time scales may be more related to the dramatic changes in brightness of the continuum as observed in changing-look AGN. Indeed, Noda & Done (2018) favor a change in mass accretion rate, followed by a propagation of the cooling front (Lawrence, 2018; Simm et al., 2016). Noda & Done (2018) also suggest that perhaps some short-term variability could be related to the amount of the disk swept by the thermal front propagation due to Hydrogen ionization instability, similarly to white dwarf systems (also, see Ross et al. 2018; Ruan et al. 2019; Śniegowska & Czerny 2019).

The variability on several years timescale could also be explained by the X-ray reprocessing model (Kokubo, 2015; Kubota & Done, 2018), assuming that the AGN UV-optical variability is a result of reprocessing of X-ray or far-UV emission (Krolik et al., 1991). The idea of X-ray reprocessing over time has gained more and more support, with evidence from simultaneous X-ray-UV-optical AGN time series (Edelson et al., 2014; McHardy et al., 2018; Zhu et al., 2018). In particular, the accretion disk blackbody emission is insufficient to explain the broadband AGN spectrum. The total SED with a soft X-ray excess, and a hard X-ray tail, requires additional sources of emission. A recent model by Kubota & Done (2018) divides the flow into blackbody emission, warm Comptonization region (the disk), and hard

X-ray hot Comptonization component (corona, or a hot material filling the region close to the black hole below the truncation radius). Since the soft X-rays are correlated with the hard X-rays, at least part of the picture consists of reflection or reprocessing of hard X-rays by the disk (Lawrence, 2018). This model predicts an increase of variability amplitude (SF_∞) with M_{BH}, and adds an insight that the observed slope is due to changes in accretion rate $\dot{m}$, explaining that smaller $\dot{m}$ corresponds to highest variability. This qualitatively agrees with the picture that dwindling fuel supply makes the flow more variable. Previous worries about X-ray reprocessing concerned the seemingly insufficient solid angle subtended by the source of the hard X-rays to cause the observed soft X-ray and optical response. This is addressed by realizing that reprocessing could be taking place in the extended region (Gardner & Done, 2017), such as an inflated inner disk (corresponding to warm Comptonizing region in Kubota & Done 2018, or even the BLR region serving as an additional 'complex reprocessor'(McHardy et al., 2018). Also, for Panda et al. (2019a) warm corona helps decrease the dependence of $R_{\rm FeII}$ on f_{Edd}.

Thus while CLAGN may be related to the state-change to ADAF flow (Śniegowska & Czerny, 2019), similar to that of XRBs (Noda & Done, 2018; Ruan et al., 2019), with cooling and heating fronts (Ross et al., 2018), the short-timescale variability requires approximately three distinct emission regions (Kubota & Done, 2018), with extended reprocessor (such as diffuse, hot, puffed-up inner disk, and BLR - McHardy et al. 2018), that reverberates the rapid hard X-ray variability in soft X-rays to optical via UV (Fausnaugh et al., 2018). Some emission (especially soft X-rays) seems to require the warm Comptonizing corona (Kubota & Done, 2018). The warm corona, coupled with metallicity changes, and variation in turbulence level and cloud density, also helps explain the Quasar Main Sequence in the optical (Panda et al., 2019a,b). Finally, the Kubota & Done (2018) model, apart from being consistent with other mechanisms (Lawrence, 2018; McHardy et al., 2018; Panda et al., 2019a; Ross et al., 2018; Ruan et al., 2019; Śniegowska & Czerny, 2019), explains the observed correlation of variability amplitude with black hole mass as corresponding to variations in mass accretion rate.

3.7 Summary and Conclusions

We model the optical variability of ∼9000 Stripe 82 quasars as the Damped Random Walk (DRW, Kelly et al. 2009). DRW is a Gaussian Process (GP), described by two parameters - characteristic timescale τ (representing decorrelation timescale, or light curve smoothness), and the asymptotic amplitude SF_∞ (which relates to the amplitude of variability). We fit observed and simulated light curves with celerite - a fast GP solver (Foreman-Mackey et al., 2017). By simulating and fitting DRW light curves, we explore the impact of the ratio of input timescale and the light curve baseline. We find that the light curve length needs to be several larger than input timescale to allow unbiased timescale retrieval, confirming K17. Motivated by this result we consider extending SDSS with PS1, PTF, CRTS, and ZTF data. We calculated appropriate photometric offsets (color terms) to relate PS1 gri, PTF gR, CRTS V, and ZTF r to SDSS r-band. However, due to larger photometric uncertainties of PTF, ZTF ,CRTS at faint magnitudes of SDSS quasars, we decided to use only PS1 r-band data. Furthermore, SDSS r-band and PS1 r-band are sufficiently similar that no photometric transformation is required. Thus by extending the SDSS r-band light curves with PS1 DR2 r-band data we improve on the fidelity of recovered DRW parameters.

We identify 40 objects which exhibit tenfold increase in variability timescale when using the SDSS-PS1 dataset as compared to timescale inferred from SDSS alone. Their light curves show characteristics of changing-look quasars (magnitude difference larger than 0.5 mag, MacLeod et al. 2016). Of these, 5 are confirmed in the literature (LaMassa et al., 2015; MacLeod et al., 2019). We recommend spectroscopic follow-up and further monitoring of the brightest targets (see Appendix 3.B).

We investigate the correlation of quasar physical properties, such as black hole mass M_{BH} and absolute i-band magnitude M_i, with DRW model parameters. The SDSS-PS1 data, coupled with Shen et al. (2011) quasar catalog, imply that the damping timescale τ is correlated with M_{BH} with a power-law index of 0.141 ± 0.019, and almost independent of quasar bolometric luminosity, as in M10, Wilhite et al. (2008), and Vanden Berk et al.

(2004b). The asymptotic variability amplitude SF_∞ is anti-correlated with luminosity M_i with power law index 0.118 ± 0.003, and correlated with M_{BH} with slope 0.118 ± 0.008. This can be explained if the driving variable was Eddington ratio f_{Edd} (Wilhite et al., 2008). Indeed, there is an anti-correlation of SF_∞ and f_{Edd}, with power law slope of -0.208 ± 0.03 (similar to M10). As suggested by Kubota & Done (2018), this gradient of SF_∞ in the plane of M_{BH} vs M_i could be explained if the lower mass accretion rate corresponds to higher variability, so that when the supply of fuel decreases, the flow becomes less stable, more clumpy, and more inhomogeneous (Dexter & Agol, 2011; Kokubo, 2015; Rakshit & Stalin, 2017). This is also consistent with the X-ray reprocessing model, whereby the hard X-ray variability of the inner disk is reflected/reprocessed by the extended warm Comptonization region (inflated disk), and perhaps a complex reprocessor, including the clouds of the broad line region (Kubota & Done, 2018; Panda et al., 2019b). Changes on recovered timescales are too fast to be driven by changes in disk viscosity, or thermal front propagation alone - thermal or dynamical timescale of response to the changes in X-ray emission seems most consistent with our results (Stern et al., 2018).

More data extending the light curves would help improve the DRW fit coefficients, potentially decreasing the scatter in observed correlations. Moreover, given that the uncertainty in black hole mass is one of the biggest sources of error, better measurements of quasar properties would be of high utility (Shen et al., 2011). This will be possible by the upcoming AGN reverberation mapping campaigns (eg. SDSS-V Black Hole Mapper), providing better calibration for line width-based methods of estimating black hole masses (Kollmeier et al., 2017). All quasars in this study were spectroscopically confirmed, but some spectra had low signal-to-noise, resulting in higher likelihood of incorrect redshift measurement. Better spectroscopy and follow up of S82 quasars, afforded by SDSS-V panoptic spectroscopy, would not only help improve on the spectrum-based properties (redshift, absolute magnitude, black hole masses), but also allow to study spectral changes, and further new CLAGN discoveries (MacLeod et al., 2019).

If this study were to be expanded onto a sample of quasars with good photometry over

sufficiently long baselines, but lacking spectral information, the required physical information on quasars could be obtained by indirect methods of estimating the coarse spectral information from broad-band photometry (Kozłowski, 2015). This would benefit from better catalogs of existing spectroscopically-confirmed quasars (SDSS DR14) to improve the calibration, as well as better methods of estimating the redshift based on photometry alone (eg. photo-z: Curran & Moss 2019; Jin et al. 2019; Richards et al. 2015; Yang et al. 2017). This will be possible in short term with the ZTF (Bellm et al., 2019), and in the long term with LSST (Ivezić et al., 2019). Occasional coverage adding few epochs to some quasars may be possible with other surveys (eg. TESS, Ricker et al. 2014), but to improve the statistics of an entire sample of S82 quasars would require longer baselines. Combining SDSS and PS1 with LSST would provide unprecedented 35-year baseline, which assuming timescales below 1000 days is over 10 times longer, allowing unbiased DRW parameter retrieval, which coupled with correlations with quasar properties, would provide an estimate of black hole masses and bolometric luminosities for millions of quasars (Ivezić et al., 2019).

3.A Measuring Quasar properties

In this work we employ black hole masses, bolometric luminosities, and K-corrections from Shen et al. (2011) catalog, based on single-epoch SDSS spectra. Here we explain the choices made in the difficult art of estimating each of these quasar physical properties.

It is non-trivial to measure the mass of black holes living in the centers of active galaxies, even provided a detailed spectrum. The most common approach to estimate black hole masses in AGN is to assume that the broad-line region (BLR) is virialized:

$$M_{BH} = f\frac{R\Delta V^2}{G} = fM_{vir} \tag{3.25}$$

where f is a constant of order unity, R is the size of the BLR (estimated from emission-line lag Δt as $R = c\Delta t$), ΔV is virial velocity, G gravitational constant (Shen et al., 2008). From reverberation mapping studies (eg. Shen et al. 2019) we know that continuum luminosity L is related to the size of the BLR region as $R \propto L^{\gamma}$ (Vestergaard & Peterson, 2006), with γ

very close to 1/2 (eg. Bentz et al. 2009 finds from RM studies $\gamma = 0.519 \pm 0.06$). Thus we find that $R\Delta V^2 \propto L^{\gamma}\Delta V^2 \equiv \mu$. The virial velocity ΔV is usually estimated from the width of the broad emission lines (or line dispersion). In the absence of a quasar spectrum, there are alternative methods using a conversion of the broad-band photometry into monochromatic fluxes in the vicinity of reverberating lines (eg. Kozłowski 2015, used in Kozłowski 2017b to estimate black hole mass for 280 000 AGN). Depending on the redshift, different rest-frame calibrated emission lines shift into the observed passband: broad Hα at 6562Å, Hβ at 5100Å, Mg II at 2800Å, and C IV at 1350Å (see Fig.7 in Shen et al. 2019, and Vestergaard 2002). Some authors even consider separately C IV-based and Mg II-based black hole mass estimates. We refer the reader to Shen et al. (2008) who in detail describes various biases and inherent assumptions of virial black hole mass measurements.

Another important quasar property - bolometric luminosity, is most often estimated from the absolute i-band magnitude, M_i (see Shen et al. 2008, Fig.2). M_i is derived from the observed i-band magnitude, by correcting for Galactic extinction, and correcting for the fact that at different redshifts different portions of the spectral energy distribution are observed by the telescope filter bandpass. The latter, known as K-correction $K(z)$ (Oke & Sandage, 1968), is defined as $m_{intrinsic} = m_{observed} - K(z)$. In the early 2000's the common approach was to K-correct to redshift 0, but as (Richards et al., 2006) pointed out, since the distribution of quasars peaks at redshift 2, for most quasars correcting to the redshift of 0 required shifting the observed spectrum into the far infrared. Moreover, the procedure was to correct separately for the continuum and emission line contributions, assuming a particular spectral shape (eg. power law $f_{\nu} \propto \nu^{\alpha}$, with $\alpha = -0.5$ - see Richards et al. 2006; Schneider et al. 2010; Vanden Berk et al. 2001). This introduces a larger error for K(z=0) than for K(z=2) if the assumed spectral shape $\alpha = -0.5$ is far from the real spectral index. In early 2010's, after Blanton et al. 2003; Richards et al. 2006; Wisotzki 2000, the practice started shifting towards K-correcting to redshift 2, and including custom quasar spectral shapes, as reflected by the content of Shen et al. (2011) quasar catalog. Thus in this study we use the absolute i-band magnitude K-corrected to z=2: $M_i(z = 2)$.

These methods were used to create catalogs of quasar properties derived from spectra. Since quasars are variable at ∼0.2 mag level, the ideal is to use a single-epoch calibrated spectrum to estimate the continuum luminosity, and find virial black hole masses using relationships based on the monochromatic fluxes and broad line widths described above. A glance at the available quasar catalogs reveals that, given any SDSS data release, there is indeed first a catalog of basic quasar properties (redshift and photometry - eg. Schneider et al. 2007, 2010), and more detailed catalogs containing black hole masses and bolometric luminosities follow (eg. Shen et al. 2008, 2011). More recently, once SDSS DR12 Quasar Catalog (Pâris et al., 2017) was released, K17 followed using SDSS photometry as a proxy for monochromatic luminosities. Chen et al. (2018) added a detailed analysis of continuum luminosities in the Hα, Hβ regions for low-redshift quasars. Using the spectra from Chinese LAMOST survey Dong et al. (2018) also sought to estimate virial black hole masses, and the results, while consistent with Shen et al. (2011), suffered from the necessity to peg the non-calibrated spectra to the SDSS photometry which was taken a different epoch. Thus even though the SDSS DR12 Quasar Catalog of Pâris et al. (2018) is the most recent, like Pâris et al. (2017) it lacks black hole masses and bolometric luminosities, and there is no recent work that re-analyzed the spectral data. Therefore we use black hole mass estimates and monochromatic luminosities from Shen et al. (2011), based directly on single-epoch spectra.

3.B CLQSO candidates

Based on the DRW model parameters τ, σ fitted with celerite using the SDSS and PS1 data, we find that there are quasars for which there is a pronounced difference between τ, σ inferred from combined SDSS-PS1 data vs just SDSS. Specifically, Fig. 3.12 shows that there are objects where $f_\sigma = \log_{10}(\sigma_{\rm SDSS-PS1}/\sigma_{\rm SDSS}) > 0.4$ and $f_\tau = \log_{10}(\tau_{\rm SDSS-PS1} \,/\, \tau_{\rm SDSS}) > 1$ (a tenfold increase in τ and over twofold increase in σ). Visual inspection of objects simultaneously satisfying $f_\sigma > 0.4$ and $f_\tau > 1$ shows that these underwent a significant (> 0.5 mag) change in brightness between the SDSS (baseline 1998 - 2008) and PS1 DR2 observations (2009-2014 - see Fig. 3.2). Thus DRW fitting could also be a way of finding changing-look

quasar (and AGN) candidates. Figs. 3.22– 3.25 show the SDSS-PS1 r-band light curves of 40 CLQSO candidates, with median PS1 brightness larger than 20.5 mag. The open circles indicate day-averaged epochs (see Sec. 3.3). Table 3.5 contains the basic physical parameters for these quasars. Some quasars show a downward trend, like turn-off CLQSO (eg. 123909, 1412379, 1644710), while some are seen in a brightening stage, like turn-on CLQSO (eg. 1976348, 221006, 4069419, 4205621). Quasar 612585 has the largest amount of auxiliary multi-wavelength coverage: x-ray from XMM Newton, UV from Galex, IR from UKIDSS, VHS, WISE, and has been analyzed as part of the x-ray targeted sample of S82 quasars - S82X by LaMassa et al. (2016). Quasar 751557 was previously identified by MacLeod et al. (2019) as a CLQSO candidate, with detailed Magellan spectroscopy described therein. Two quasars: 1003694 and 1299803 have WISE data in the S82X catalog (LaMassa et al., 2016). Quasars 612585 and 3633437 have X-ray detections in the 3XMM-DR5 catalog (Rosen et al., 2016), but they have no matches in Chandra point source catalog (second release, Evans & Civano 2018; Evans et al. 2010). There are no matches against the unified radio catalog of Kimball & Ivezić (2008), which includes FIRST and NVSS data. We especially recommend spectroscopic follow-up of the brightest targets: 1976348 (mean 17.8 mag, turn-on, upper-left panel in Fig. 3.23), 2104791 (mean 18.4 mag, turn-off, middle-left panel in Fig. 3.23).

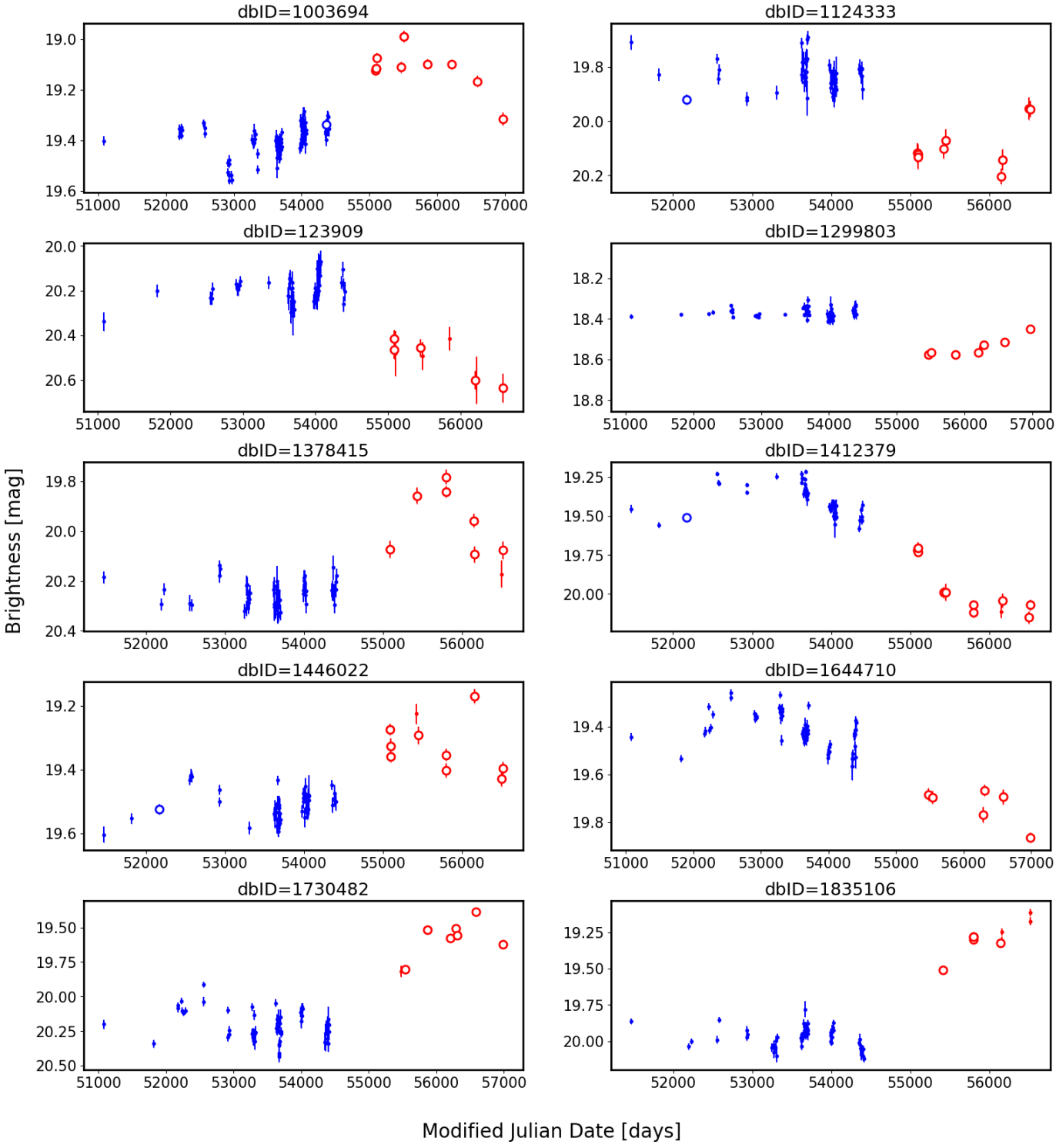

Figure 3.22: Outliers in the space of recovered DRW parameters between SDSS and SDSS-PS1, as well as median offsets. Page 1.

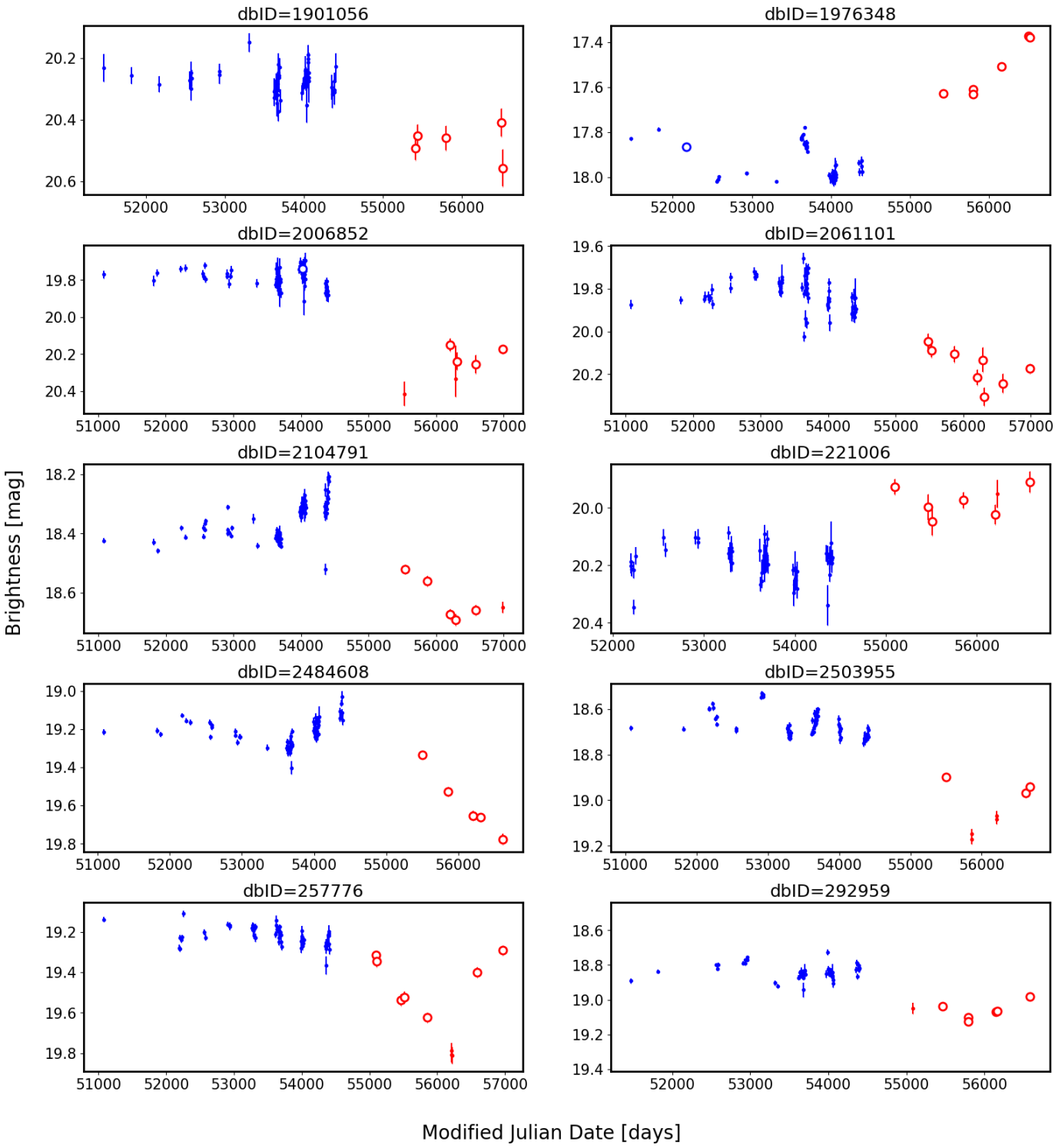

Figure 3.23: As Fig. 3.22, page 2.

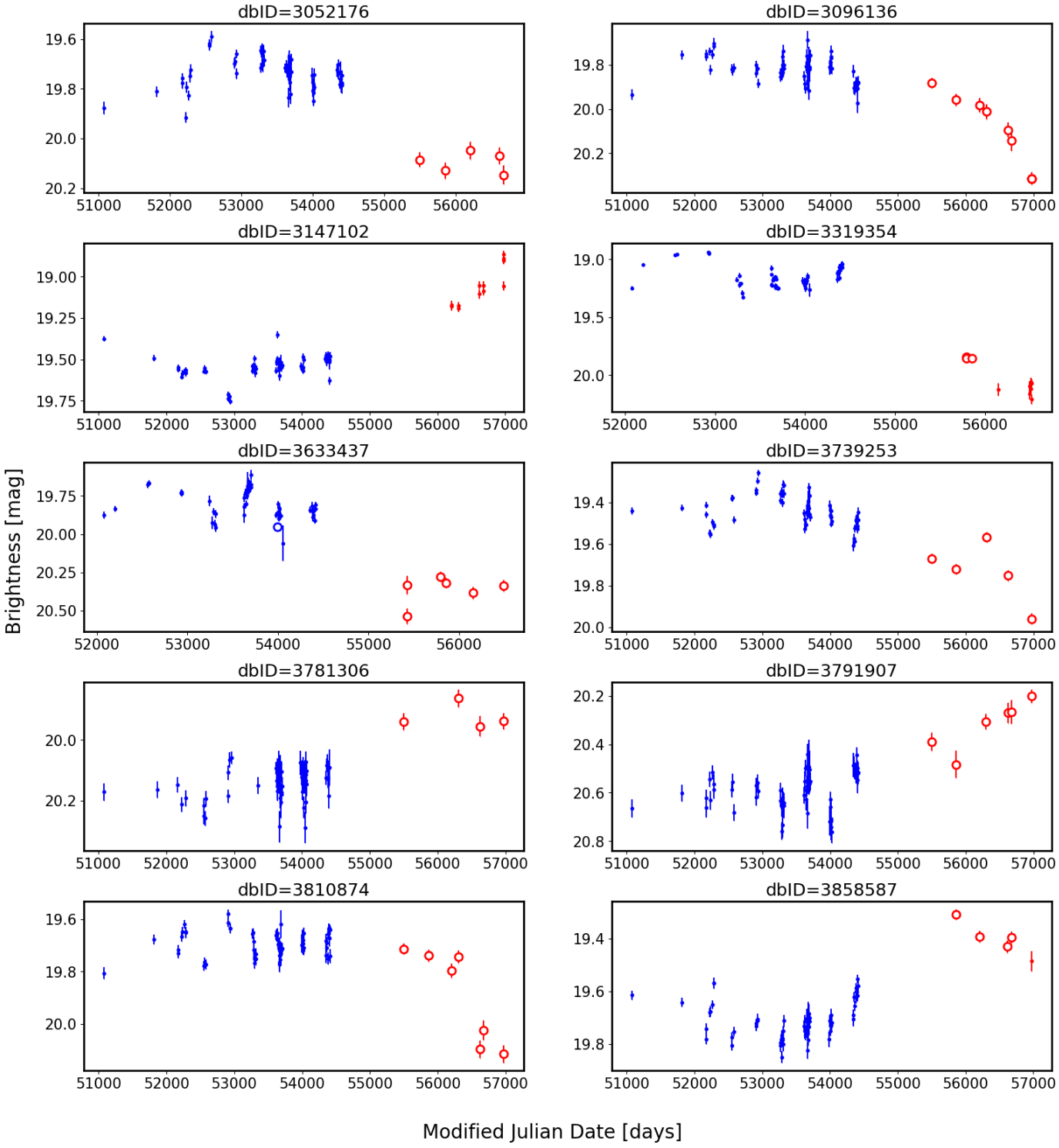

Figure 3.24: As Fig. 3.22, page 3.

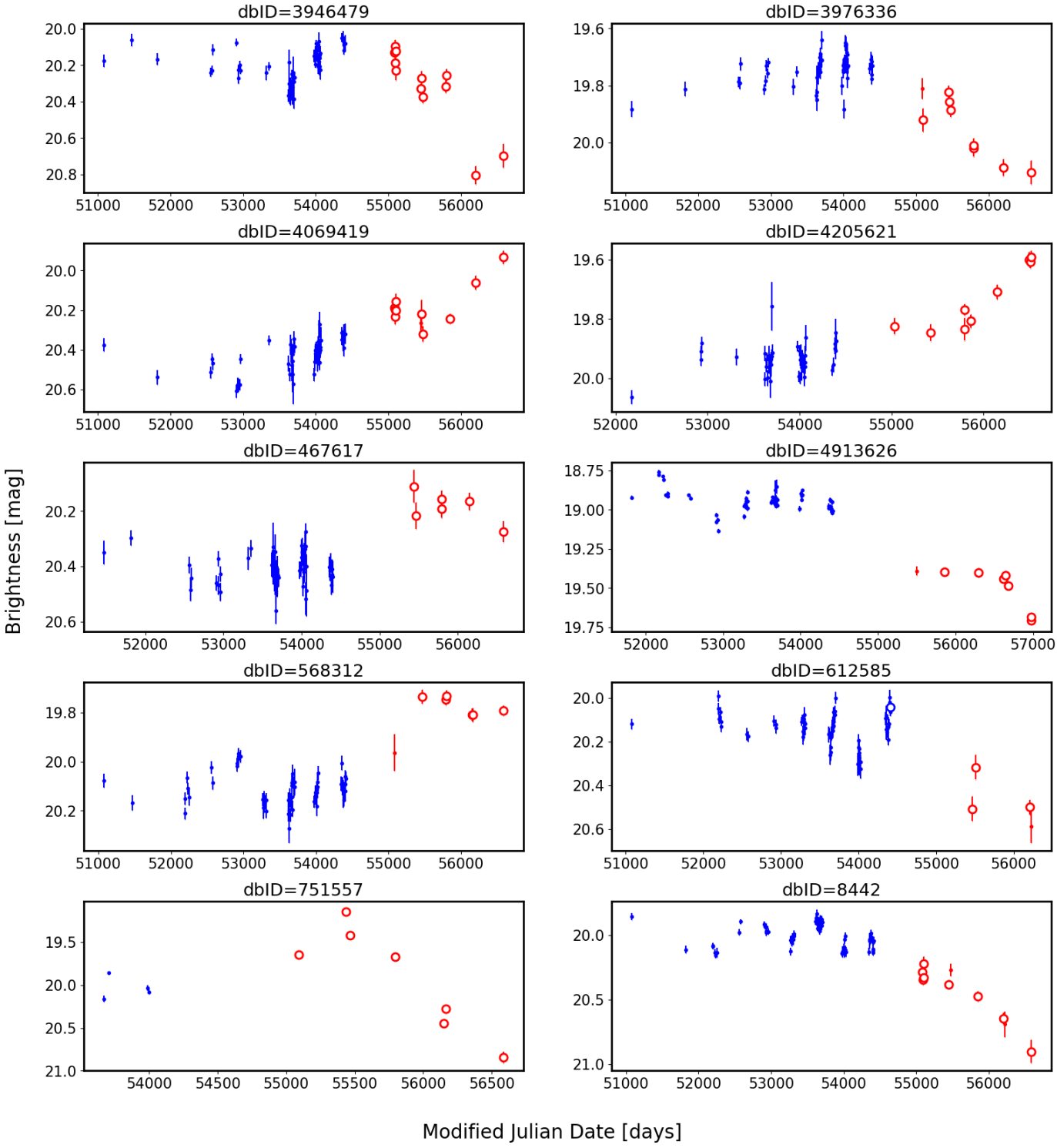

Figure 3.25: As Fig. 3.22, page 4.

Table 3.5: CLQSO candidates : catalog information from Shen et al. (2011), concerning DR7 name (dbID) and SDSSJID location (α, δ in degrees, J2000), distance (spectrum-based redshift), physical parameters (bolometric luminosity L_{Bol} erg s^{-1}, black hole mass ($M_{\rm BH}[M_\odot]$), Eddington ratio $f_{Edd} = L_{Bol}/L_{Edd}$), and light curve properties (the difference between median SDSS and PS1 magnitudes Δ(mag), the difference in scatter between SDSS and PS1 segments $\Delta(\sigma_G)$, and the median PS1 magnitude - see Fig. 3.14).

dbID	SDSSJID	α	δ	redshift	$\log_{10}(L_{Bol})$	$\log_{10}(M_{\rm BH})$	f_{Edd}	Δ(mag)	$\Delta(\sigma_G)$	medianPS1
123909	001626.54+003632.4	4.111	0.609	3.24	46.57	9.47	-1.0	-0.29	-0.054	20.48
8442	001731.70+004910.1	4.382	0.819	2.43	46.61	9.09	-0.58	-0.36	-0.107	20.36
4069419	003359.39+000230.0	8.497	0.042	1.64	45.95	9.05	-1.21	0.22	0.017	20.2
221006	005142.20+002129.0	12.926	0.358	1.55	45.95	8.24	-0.39	0.21	-0.01	19.97
257776	005513.15-005621.2	13.805	-0.939	3.61	47.13	9.58	-0.54	-0.32	-0.242	19.53
612585[a]	010812.00-000516.5	17.05	-0.088	1.0	45.52	9.06	-1.64	-0.37	0.006	20.5
1003694[b]	012114.19-010310.8	20.309	-1.053	1.89	46.59	8.83	-0.34	0.28	0.032	19.11
1299803[b]	014303.23-004354.0	25.763	-0.732	0.53	45.78	8.68	-1.0	-0.18	-0.008	18.56
1644710	021259.00-000550.1	33.246	-0.097	0.81	45.67	8.38	-0.81	-0.28	0.022	19.7
1730482	021529.02-005314.9	33.871	-0.887	1.37	45.98	8.8	-0.92	0.66	-0.009	19.57
2104791*	022239.83+000022.5	35.666	0.006	0.99	46.28	9.33	-1.16	-0.31	0.013	18.66
2061101	022505.06+001733.2	36.271	0.293	2.42	46.38	8.09	0.2	-0.34	-0.016	20.15
2006852	023917.86-001916.8	39.824	-0.321	1.41	46.07	8.73	-0.76	-0.45	-0.013	20.24
2503955	025316.46+010759.7	43.319	1.133	1.03	46.31	8.94	-0.73	-0.39	-0.071	19.07

2484608	025654.42-011455.4	44.227	-1.249	0.54	45.57	8.48	-1.01	-0.43	-0.023	19.65
3052176	030504.07+011324.5	46.267	1.223	0.61	45.29	9.2	-2.01	-0.35	0.016	20.09
3096136	031401.11+011131.6	48.505	1.192	1.31	46.06	9.21	-1.25	-0.27	-0.181	20.09
3147102	031846.13-005622.8	49.692	-0.94	2.12	46.57	9.07	-0.6	0.47	-0.078	19.07
3781306	032745.74+005217.2	51.941	0.871	1.16	45.77	N/A	N/A	0.19	0.032	19.94
3858587	032825.19-003252.3	52.105	-0.548	0.77	45.61	8.68	-1.17	0.33	0.031	19.39
3810874	033047.73+004859.4	52.699	0.816	0.86	45.8	8.35	-0.65	-0.1	-0.177	19.8
3739253	033059.05+010952.0	52.746	1.164	0.56	45.39	8.16	-0.86	-0.27	0.01	19.72
3791907	033431.17-000904.0	53.63	-0.151	1.64	46.01	9.12	-1.21	0.28	0.02	20.29
4913626	034512.62+002245.7	56.303	0.379	0.42	45.5	8.81	-1.41	-0.48	-0.055	19.43
4205621	203932.41-001818.3	309.885	-0.305	1.58	46.21	8.66	-0.55	0.17	-0.126	19.77
3319354	204952.62+011306.6	312.469	1.219	1.09	46.14	9.52	-1.48	-0.91	-0.048	20.08
3633437[c]	205105.02-005847.5	312.771	-0.98	0.54	45.34	8.47	-1.23	-0.51	0.071	20.34
1835106	215015.05-005331.4	327.563	-0.892	1.9	46.33	9.21	-0.98	0.69	0.035	19.28
1901056	215055.51-001739.4	327.731	-0.294	1.54	46.26	8.6	-0.44	-0.19	0.009	20.46
1976348*	215841.40-001507.7	329.673	-0.252	1.46	46.92	9.39	-0.57	0.39	-0.051	17.56
1446022	220535.23+000756.3	331.397	0.132	1.69	46.45	9.25	-0.9	0.17	-0.039	19.34
1378415	221347.32+001928.4	333.447	0.325	2.31	46.41	8.59	-0.29	0.23	-0.113	20.02
1412379	221831.58-004548.9	334.632	-0.764	1.23	46.15	9.48	-1.43	-0.61	-0.057	20.04
1124333	222918.25-004003.6	337.326	-0.668	1.16	45.81	8.35	-0.64	-0.28	-0.003	20.12
751557[d]	225240.37+010958.7	343.168	1.166	0.53	45.32	8.88	-1.66	0.39	-0.536	19.67

467617	231032.17-011449.5	347.634	-1.247	1.82	46.03	7.86	0.06	0.23	0.024	20.18
568312	231953.07-010139.0	349.971	-1.028	1.15	45.6	8.29	-0.79	0.33	0.005	19.79
292959	232030.97-004039.2	350.129	-0.678	1.72	46.69	9.39	-0.8	-0.22	-0.009	19.07
3976336	235213.27-004326.3	358.055	-0.724	0.9	45.64	8.85	-1.3	-0.18	-0.08	19.92
3946479	235248.71-001518.4	358.203	-0.255	1.34	45.8	9.01	-1.3	-0.07	-0.025	20.26

*: Recommended for follow-up

a: S82X (LaMassa et al., 2016), XMM Newton, Galex UV, UKIDSS, VHS, WISE

b: S82X (LaMassa et al., 2019), WISE

c: In 3XMM-DR5 x-ray catalog (Rosen et al., 2016)

d: M19, CLQSO candidate, Magellan follow-up

3.C Mg II variability

We searched for dependency of variability parameters on other physical properties, beyond the black hole mass and quasar luminosity employed in Eq. 3.12. The quasar optical spectrum has certain strong emission lines, depending on the redshift. In particular, the Mg II emission line lags the continuum and is less variable (Reichert et al., 1994). Mg II line is also an important virial black hole mass estimator for quasars - McLure & Jarvis (2002) noticed that the FWHM of the Mg II doublet trails that of the Hβ line (also see Shen 2013). Following Ivezić et al. (2004) and MacLeod et al. (2012), we investigated the residuals in data-model(SF_∞) as a function of wavelength. Left hand side panel of Fig. 3.26 shows the SF_∞model residuals in the λ_{RF} vs τ_{RF} space. The right hand side panel shows the residuals marginalized along τ_{RF}. The decrement around 2800 Å on the right hand side panel is more pronounced when using combined SDSS-PS1 light curves. We see that both aggregates based on raw data (blue dots) or on medians (black dots) agree - the two statistical methods show a small ($\sim 5\%$), but statistically significant detection of a depression in the 2800 Å region, as expected from theory.

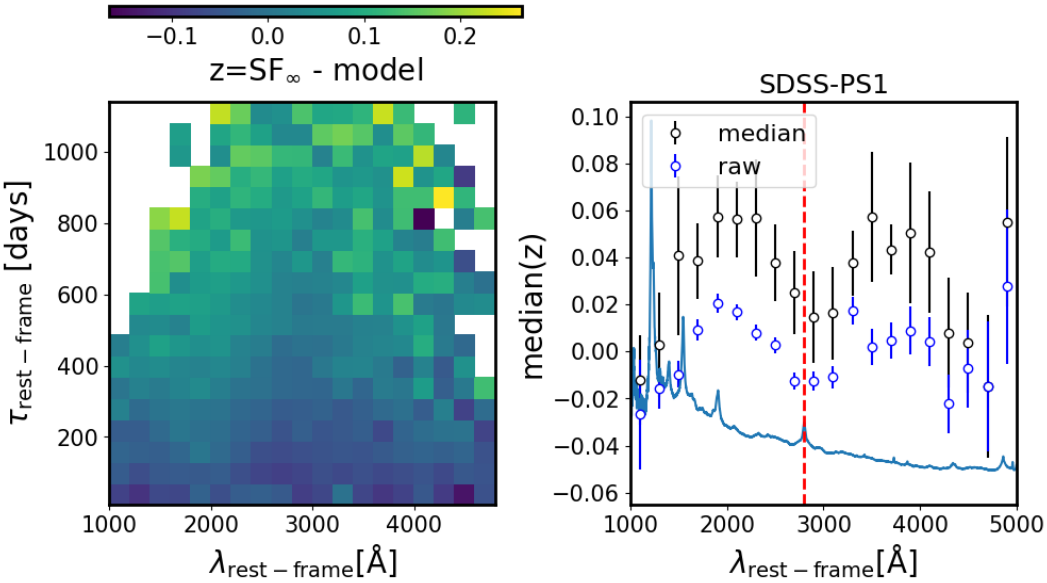

Figure 3.26: The data-model residuals after fitting Eq. 3.12 to the quasar SF_∞, using SDSS-PS1 quasar light curves. The left panel shows the residuals as a function of $\lambda_{\rm RF}$ and $\tau_{\rm RF}$. The right panel shows the median of residuals marginalized along $\tau_{\rm RF}$, with error on the median based on the scatter in $\lambda_{\rm RF}$ bins : $\sigma_{median} = 1.25{\rm RMS}_{\rm bin}/\sqrt{N_{\rm bin}}$, where ${\rm RMS}_{\rm bin}$ is estimated from the robust Gaussian interquartile-based σ_G (Ivezić et al., 2014). Black circles are medians based on the averages plotted on the left ($N_{\rm bin}$ is between 7 and 20, depending on the number of non-empty bins on the left hand side panel), and blue dots are medians based on the raw data ($N_{\rm bin}$ varies as a function of $\lambda_{\rm RF}$, from a few to 800, peaking at about 2500 Å, reflecting the fact that the quasar distribution peaks at redshift $z = 2$). We overplot the composite quasar spectrum from Vanden Berk et al. (2001), and mark with vertical dashed red line the location of the Mg II 2798 Å line.

Chapter 4

LSST FALL 2017 CROWDED FIELDS TESTING

We quantify the performance of the LSST pipeline for processing crowded fields, using images obtained from DECam and comparing to a specialized crowded field analysis performed as part of the DECAPS survey. Considering single-visit LSST depth, in an example field of roughly the highest density seen in the LSST Wide-Fast-Deep area, DECAPS detects 200 000 sources per sq. deg. to a limiting depth of 23rd magnitude. At this source density the mean LSST-DECAPS completeness between 18th and 20th mag is 80%, and it drops to 50% at 21.5 mag. For fields inside the Galactic plane cadence zone, source density rapidly increases. For instance, in a field in which DECAPS detects 500 000 sources per sq. deg. (5σ depth of 23.2), the mean completeness between 18th and 20th mag is 78%, and it drops to 50% at 20.2 mag. In terms of photometric repeatability, above 19th mag LSST and DECAPS are in a systematics-dominated regime, and there is only a slow dependence on source density. At fainter magnitudes, the scatter between LSST and DECAPS is less than the uncertainty from photon noise for source densities up to 100 000 per sq. deg, but the scatter grows to twice the photon noise at densities of 300 000 per sq. deg. and above. For repeat measurements of the same field with LSST, the astrometric scatter per source is at the level of 10-30 milliarcseconds for bright stars ($g < 19$), and is not strongly dependent on stellar crowdedness.

4.1 Introduction

We report on the performance of the Large Scale Synoptic Telescope (LSST) science pipelines[1], also known as 'the LSST stack', in stellar fields of varying levels of source crowdedness.

[1] `https://pipelines.lsst.io`

The LSST will sample on average over 500 regions in the sky every night, delivering terabytes of raw data in need of processing, including photometric and astrometric calibration, to deliver a calibrated image, as well as a source catalog, and other image products (Jurić et al., 2017).

Assuming a single-visit depth of 24.5 mag, the stellar density ranges from high density low-galactic latitude regions that have tens of millions of sources per square degree, to low-density regions towards the Galactic poles with less than a thousand source per square degree.

Deblending and accurate photometry is an inherent part of any astronomical data processing pipeline. There exists a body of research answering questions that are specific to crowded stellar fields, eg. how many beams (resolution elements) do we need per source (Hogg, 2001), or how the crowded fields photometry can be approached in the era of large telescopes (Olsen et al., 2003). Other studies involving Hyper Suprime-Cam pipeline (developed in parallel with the LSST Stack) recognized that the deeper the survey, the higher the stellar densities encountered, and therefore, the more challenging the process of deblending photometry (Bosch et al., 2018).

In this report we compare the 'out-of-the-box' LSST Stack processing pipeline, to the DECam [Galactic] Plane Survey (DECAPS) pipeline developed by Schlafly et al. (2018). To test the performance of these pipelines at different levels of stellar crowdedness, we choose regions of the sky at various densities based on the Galfast simulation of the night sky (Sec. 4.2). Given the expected stellar density as a function of position on the sky, we selected DECAPS fields, and processed them with LSST pipelines (Sec. 4.3). We compare the results of the LSST and DECAPS processing of the same visits by cross-matching the catalogs and comparing source counts, photometry (Sec. 4.4), and astrometry (Sec. 4.6). We summarize the key results and suggest future work in Sec. 4.7. There is an accompanying document in GitHub with longer data tables[2].

[2] https://github.com/lsst-dm/dmtn-077/blob/master/main_tables.pdf

4.2 Identifying density regions

To identify regions representing different stellar densities we use the Galfast simulated stellar density map prepared as part of Metrics Analysis Framework[3] by P. Yoachim and L. Jones[4].

The resulting dataset describes the simulated sky, divided into 49152 healpixels. Each healpixel contains 64 magnitude bins between 15 and 28 mag, each bin storing the cumulative count of sources per square degree[5]. In Table 4.1 we summarize the stellar count as a function of the limiting magnitude, and the area of the sky. In Table 4.2 we show what area of the survey would include regions of a particular stellar density. In this report we select the LSST single-visit depth limiting magnitude of r=24.5 . For each healpixel we find the fraction of pixels that have a higher stellar count (see Fig. 4.1).

Since by definition each healpixel has an equal area, the fraction of pixels corresponds to the fraction of the sky area. We choose to describe the level of stellar crowdedness by the percentage of the sky that has a higher density. Thus eg. '5%' density means that only 1 in 100 pixels has a higher density (see Fig. 4.1, 4.2).

4.3 DECam Plane Survey

To test the performance of the LSST Stack with real data, we used the Dark Energy Camera (DECam) images, taken as part of the DECam Plane Survey (DECAPS) (Schlafly et al., 2018), at the 4-m Cerro Tololo Inter-American Observatory telescope (CTIO)[6]. Each DECAPS image plane is tiled by a mosaic of 62 CCDs, each 2046x4094 px, 0.27 ″/px[7]. The FOV of the full mosaic is 2.2°wide — several times bigger than the full moon — which makes it comparable to the LSST 3.5°field-of-view. All DECAPS single-epoch images were

[3] https://www.lsst.org/scientists/simulations/maf

[4] https://github.com/lsst/sims_maf/blob/master/python/lsst/sims/maf/maps/createStarDensitymap.py

[5] Healpix stands for Hierarchical Equal Area isoLatitude Pixelization http://healpix.sourceforge.net (Górski et al., 2005)

[6] see http://www.ctio.noao.edu/noao/node/1033

[7] See Fig.4-3 in the NOAO Data Handbook (Shaw, 2015)

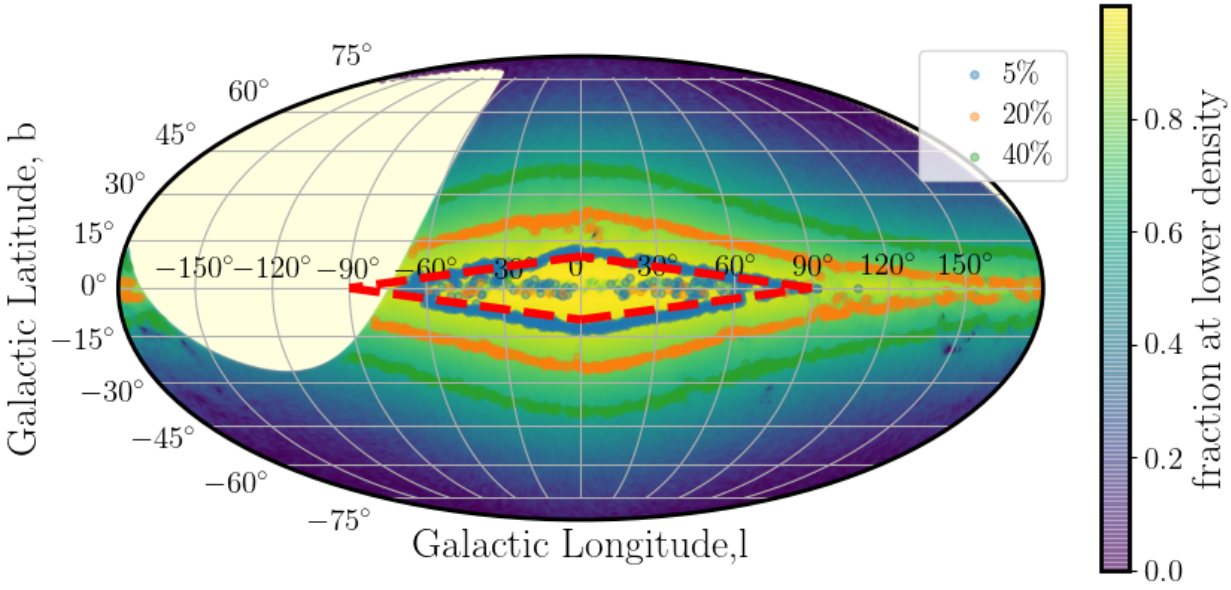

Figure 4.1: Galfast healpixels plotted in galactic coordinates in Mollweide projection. The brightest regions correspond to highest stellar densities. The blue, orange, and green points correspond to healpixels in given density brackets (top 5%, 20%, 40%). Note that the galactic exclusion zone, marked by red dashed lines, almost exactly traces the top 5% density contour. As expected, the highest density regions are located close to the galactic bulge, and regions of approximately constant stellar density trace isophotes of the Milky Way. The few 20% regions close to the galactic equator correspond to high extinction regions that appear to have less counts due to interstellar dust. The missing part at $\delta > 40^{\circ}$ is not observable from the southern location of Cerro Pachón, hence excluded from the simulation.

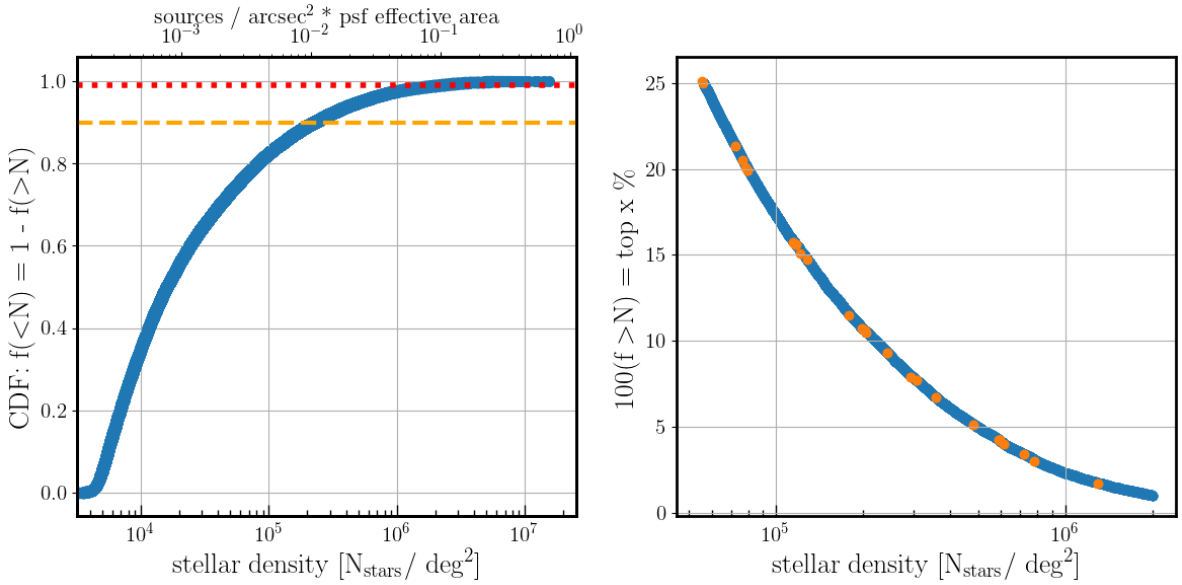

Figure 4.2: Using the Galfast sky simulation to choose DECAPS fields sampling different density regions. The left panel shows the fraction of the sky at a smaller density as a function of the stellar density. It is equivalent to the cumulative area of the sky up to given density. Given the stellar density per simulated healpixel, we count the number of healpixels at greater density. Normalized to the number of pixels, given their equal area, it corresponds to the fraction of the sky at greater stellar density. Horizontal dashed lines illustrate selecting pixels at top 1% or 10% density. The right panel focuses on the top 25% of density. It implies that according to the simulation, the density of 200,000 stars per sq.deg. corresponds to 5% of the sky, and only 1% of the sky has more than 10^6 stars per sq.deg. The upper axis represents the dimensionless density parameter $N_{beam} = N_{stars}/arcsec^2 * A_{PSF}$, with the PSF effective area $A_{PSF} = 0.64\ ''$.

Table 4.1: Dependence of the stellar count in Galfast simulation on the limiting magnitude. The Wide-Fast-Deep (WFD) survey is defined as $-65 < \delta < 5$, excluding the confusion zone (see `https://www.lsst.org/sites/default/files/skymap-2016.jpg`, and Awan et al. (2016)). All counts are in billions: 10^9. N_{all} is the count in all healpixels, $N_{top1\%(10\%)}$ in the pixels in top 1% (10%) density, N_{WFD} is the count in the WFD survey area, $N_{WFD+conf.}$ is count within WFD and confusion zone.

mag	N_{all}	$N_{top1\%}$	$N_{top10\%}$	N_{WFD}	$N_{WFD+conf.}$
24.5	3.76	1.09	2.84	3.16	1.1
27.5	8.50	2.65	7.76	7.28	2.19

processed with the DECAPS pipeline, resulting in single-epoch catalogs[8]. The details of the DECAPS pipeline can be found in Schlafly et al. (2018), but it was specifically designed for crowded field photometry, performing DAOPhot-like procedure (Stetson, 1987), without employing DAOPhot. The DECAPS algorithm performs repeated source detection, subtraction, and re-detection, which is different from the LSST pipeline. The DECAPS pipeline simultaneously solves for the positions and fluxes for all stars for a small fragment of the CCD (see Sec.4 in Schlafly et al. 2018). The headers of all DECAPS catalogs, assembled into the image database with information about single-visit exposure time, filter, time of observation, position, were used to select fields in the g filter, with 96 second exposures (to match the LSST 30 sec single-visit depth). Of these, we chose visits representative of given stellar densities based on the Galfast simulation (see Fig. 4.3). Postage stamp miniatures (Fig. 4.4) show that we indeed sample vastly different densities. Comparing DECAPS to Galfast counts (Fig. 4.5) we find that although the simulation may be not more accurate than up to a factor of a few, it is nevertheless useful for defining density regions.

[8]All available via `http://decaps.skymaps.info/catalogs.html`

Table 4.2: The first four columns contain the area and number of stars in Galfast healpixels with density $< \rho$ [sources per sq.deg] in the LSST Wide-Fast-Deep survey area. The WFD is defined as the area within, $-65 < \delta < 5$, excluding the confusion zone (CZ). The final four columns inform what percentage of WFD or CZ is at a density $< \rho$. Thus for instance if one decides to avoid regions of 200 k stars per sq.deg. and higher, it still includes 92.4 % of WFD(<24.5). Note that at r<24.5, all healpixels at densities at or above ≈1 mln stars per sq.deg. are in the confusion zone, so increasing the density cutoff does not include more healpixels. However, for r<27.5 there are regions outside the confusion zone even at $\rho > 1$ mln sources per sq.deg.

ρ	A(r<24.5)	A(r<27.5)	N(r<24.5)	N(r<27.5)	WFD(<24.5)	WFD(<27.5)	CZ(<24.5)	CZ(<27.5)
35 k	13002	10686	0.16 bln	0.18 bln	68.1 %	56.0 %	0.9 %	0.1 %
100 k	16226	14492	0.35 bln	0.41 bln	85.0 %	75.9 %	5.7 %	0.9 %
200 k	17631	16300	0.55 bln	0.66 bln	92.4 %	85.4 %	13.2 %	3.6 %
300 k	18235	17112	0.70 bln	0.86 bln	95.5 %	89.7 %	18.5 %	5.6 %
400 k	18581	17616	0.82 bln	1.04 bln	97.4 %	92.3 %	23.8 %	8.5 %
600 k	18920	18195	0.99 bln	1.32 bln	99.1 %	95.3 %	33.9 %	12.0 %
700 k	19003	18378	1.04 bln	1.44 bln	99.6 %	96.3 %	40.1 %	13.8 %
800 k	19053	18519	1.08 bln	1.55 bln	99.8 %	97.0 %	46.5 %	15.7 %
1 mln	19084	18742	1.10 bln	1.74 bln	100.0 %	98.2 %	57.5 %	20.5 %
2 mln	19084	19084	1.10 bln	2.19 bln	100.0 %	100.0 %	81.3 %	47.9 %

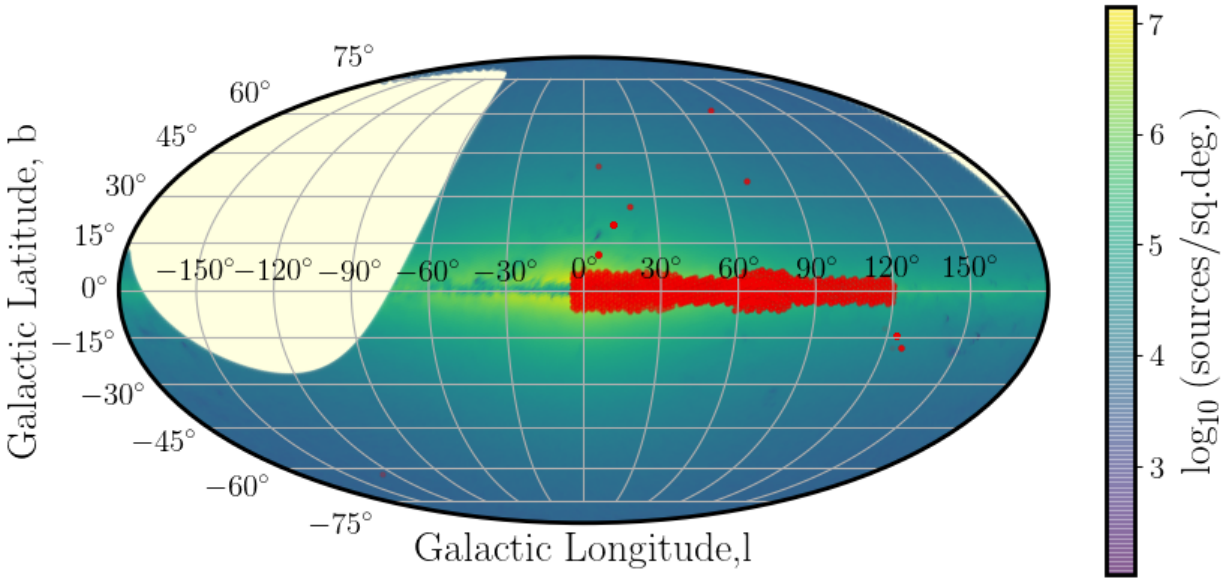

Figure 4.3: DECAPS fields (red) plotted on top of the Galfast simulated stellar density map (counts up to r<24.5). Cross-matching DECAPS catalog to Galfast simulation we selected visits representative of diverse range of stellar densities.

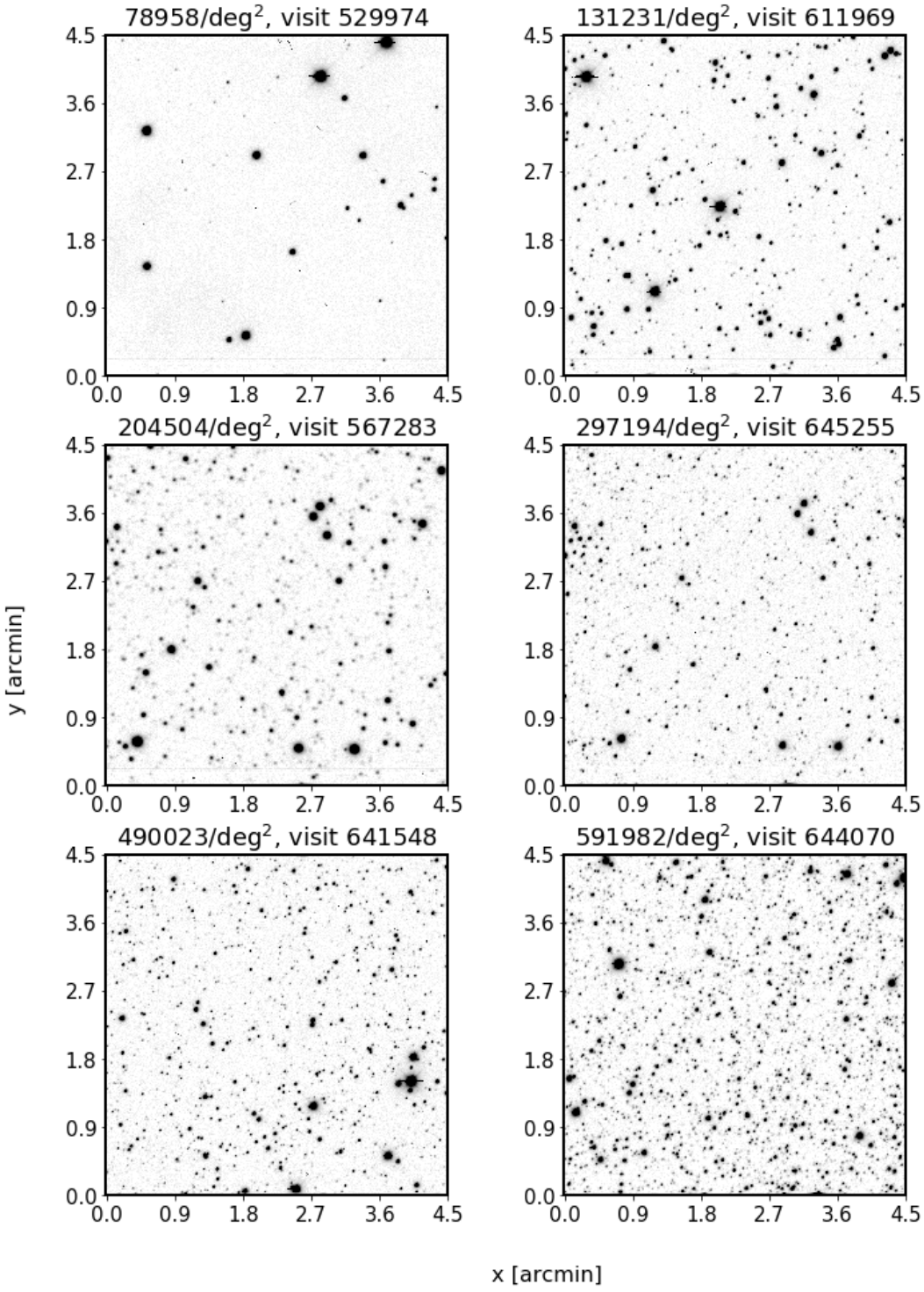

Figure 4.4: Illustrations of regions of different stellar count in the cleaned DECAPS single-epoch catalogs. As shown on Fig. 4.5, the Galfast count does not always correspond 1:1 to the DECAPS stellar count. For this reason we ordered DECAPS fields in terms of DECAPS source count rather than Galfast densities.

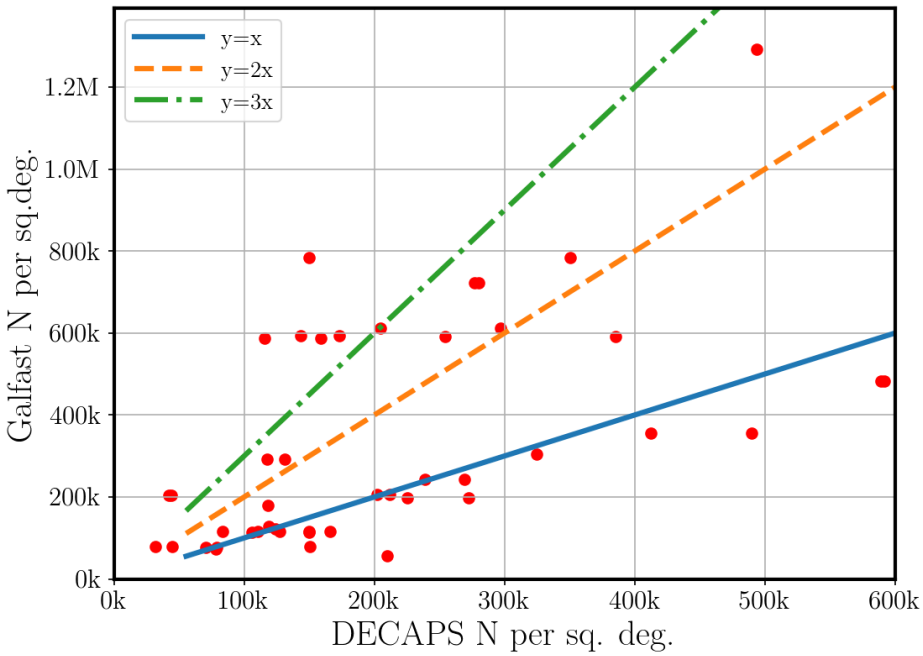

Figure 4.5: Comparison of DECAPS counts to Galfast simulated stellar counts. Overplotted are the line of equivalence y=x, and its multiplicities (2x,3x). Part of the reason for discrepancy could be the order-of-magnitude nature of the experiment - Galfast counts here assume the single-visit depth of 24.5 mag in r-band. The DECAPS exposure time (≈100 sec) and filter (g, or r) were chosen to mimic that depth as closely as possible, but the regions targeted include much extinction, which means that in some cases DECAPS counts may be less than what is implied by the simulation. However, there is a number of fields that lie along the blue line, implying that in some cases the Galfast counts were very close to the measured DECAPS counts.

4.4 LSST Processing of DECAPS data

Calibrated DECAPS imaging was processed with the LSST Science Pipelines installed on the LSST-dev machine at the NCSA[9], using the Stack version `d_2017_10_27`, and running `processCcd.py` with a mostly-default configuration. The only configuration changes made were to enable `PSFEx` for PSF determination and the new `MatchPessimisticB` astrometric matching code, which performs better over a wide ranges of source densities.

Transferring the resulting source catalogs and image files to a local machine we analyzed the output of LSST processing with jupyter notebooks and custom python tools[10]

Initially both DECAPS and LSST source catalogs contain good detections, as well as sources that are spurious, have a low S/N, or are flagged due to some other detection/processing issue. To clean both catalogs we used the DECAPS source flags, LSST source flags, and LSST image mask information.

First we compared whether the DECAPS source flags are consistent with the LSST image mask (Table 4.3). Confirming that they are, we decided to clean the DECAPS catalog with the DECAPS source level flags, removing edge detections, cosmic rays, or saturation spikes (see Table 4.4).

We followed a similar procedure with LSST source catalogs, removing sources flagged as 'edge' or 'interpolatedCenter'[11]. Moreover, only in the case of LSST catalog we are provided with the deblender-level information with 'parentId' and 'nchild' information for each source. Since the LSST pipeline deblends sources in a similar fashion to the SDSS Imaging Pipeline[12],

---

[9] `lsst-dev01.ncsa.illinois.edu` (141.142.237.49) OS: CentOS 7.4.1708 HW: Dell Inc. CPU: 48x 2.60GHz RAM: 252 GB

[10] Remote jupyter notebook access, which will be part of the Data Access Center, is not supported yet, as of early 2018.

[11] This is similar to the example in Sec.4 of SDSS Image Processing I: The Deblender (Lupton, 2005). Other flags would remove too many sources that have only small defects, eg. a bright source with a cosmic ray across its footprint can be flagged as 'interpolated', while any source which has even one bad pixel in the footprint would be flagged as 'bad' (Table 4.5).

[12] SDSS Image Processing I: The Deblender (Lupton, 2005), SDSS Image Processing II: The Photo Pipelines (Lupton et al., 2001), (Lupton et al., 2002), and (Lupton et al., 2005)

Table 4.3: LSST pixel mask. The decision is with reference to comparing specific LSST mask information to the DECAPS source flags.

Bit position	Description	Decision
0	bad	Remove
1	saturated	Remove
2	interpolated	Remove
3	cosmic ray	Remove
4	edge	Remove
5	detected	Keep
6	detected negative	Remove
7	suspect	Remove
8	no data	Remove

based on 'parentId' and 'nchild' we retain only successfully deblended children, or isolated parents (see Table 4.6, and Fig.4.6).

Finally, for both LSST and DECAPS catalogs we made a quality cut on S/N, keeping only sources where S/N > 5.

4.5 Source detection and photometry

We compare the LSST and DECAPS source catalogs in terms of source detection completeness and photometric accuracy. Considering the clean catalog source counts up to a given magnitude, we find that DECAPS catalogs contain more sources, which is especially noticeable at higher densities, with largest contribution from sources above 22 mag (Fig. 4.8).

4.5.1 Completeness

For any image analysis pipeline the ability to successfully detect and deblend sources should decrease as a function of increasing stellar crowdedness.

Table 4.4: DECAPS source flags.

Bit position	Description	Decision
1	Bad pixel	Remove
3	Saturated	Remove
4	Bleed trail	Remove
5	Cosmic ray	Remove
6	Low weight	Remove
8	Long streak	Remove
20	Additional bad pixel	Remove
21	Nebulosity	Keep
22	S7 amplifier B	Remove

Table 4.5: LSST source flags.

name	explanation
flag	general failure flag, set if anything went wrong
offimage	Source center is off image
edge	Source is outside usable exposure region
interpolated	Interpolated pixel in the Source footprint
saturated	Saturated pixel in the Source footprint
cr	Cosmic ray in the Source footprint
bad	Bad pixel in the Source footprint
suspect	Source's footprint includes suspect pixels
interpolatedCenter	Interpolated pixel in the Source center
saturatedCenter	Saturated pixel in the Source center
crCenter	Cosmic ray in the Source center
suspectCenter	Source's center is close to suspect pixels
flag	General Failure Flag

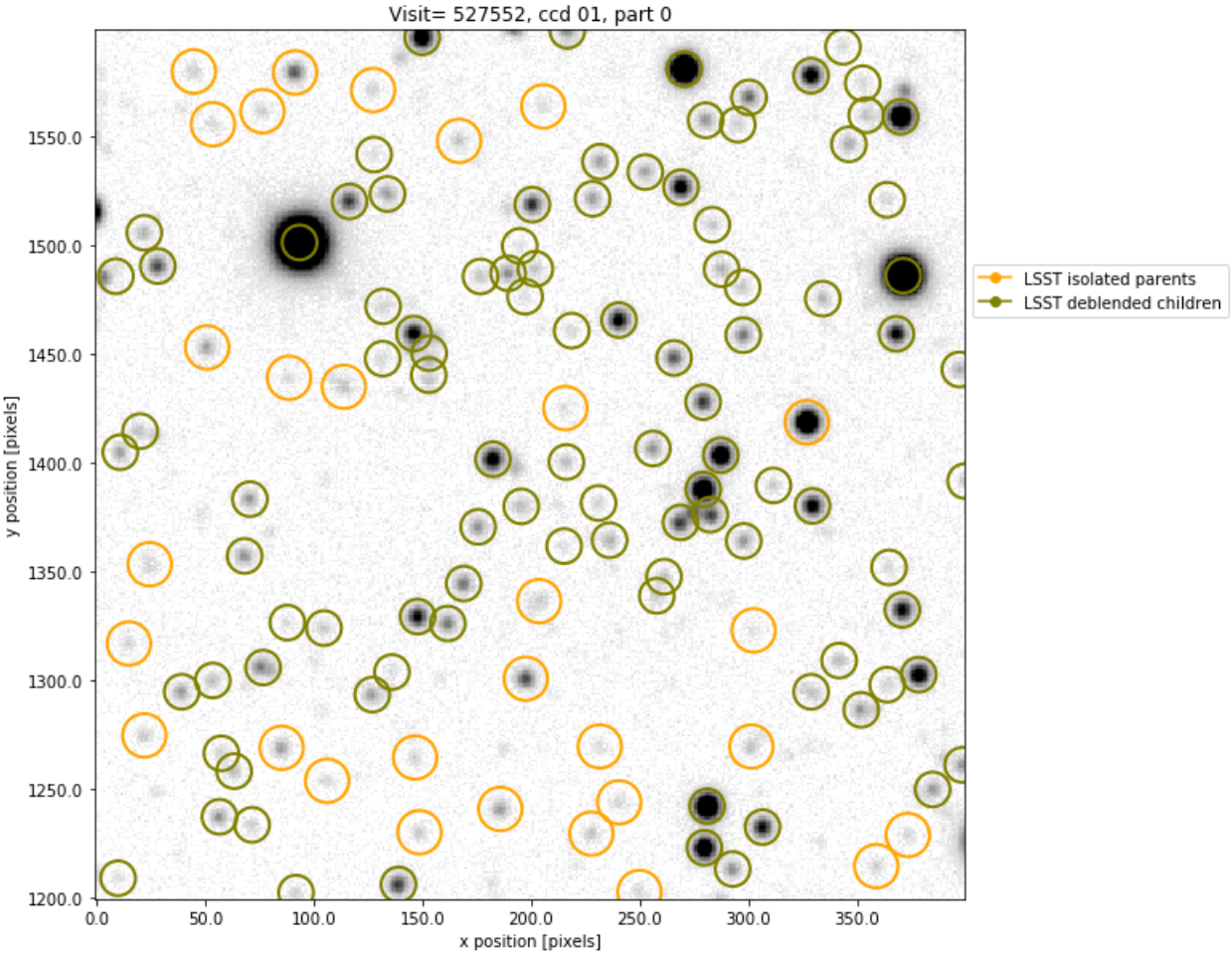

Figure 4.6: We illustrate the sources as reported by the LSST pipeline for a small region of CCD01 of visit 527552. A source may be reported as an isolated source (yellow), or a successfully deblended child (green). In this analysis we only keep isolated parents or deblended children.

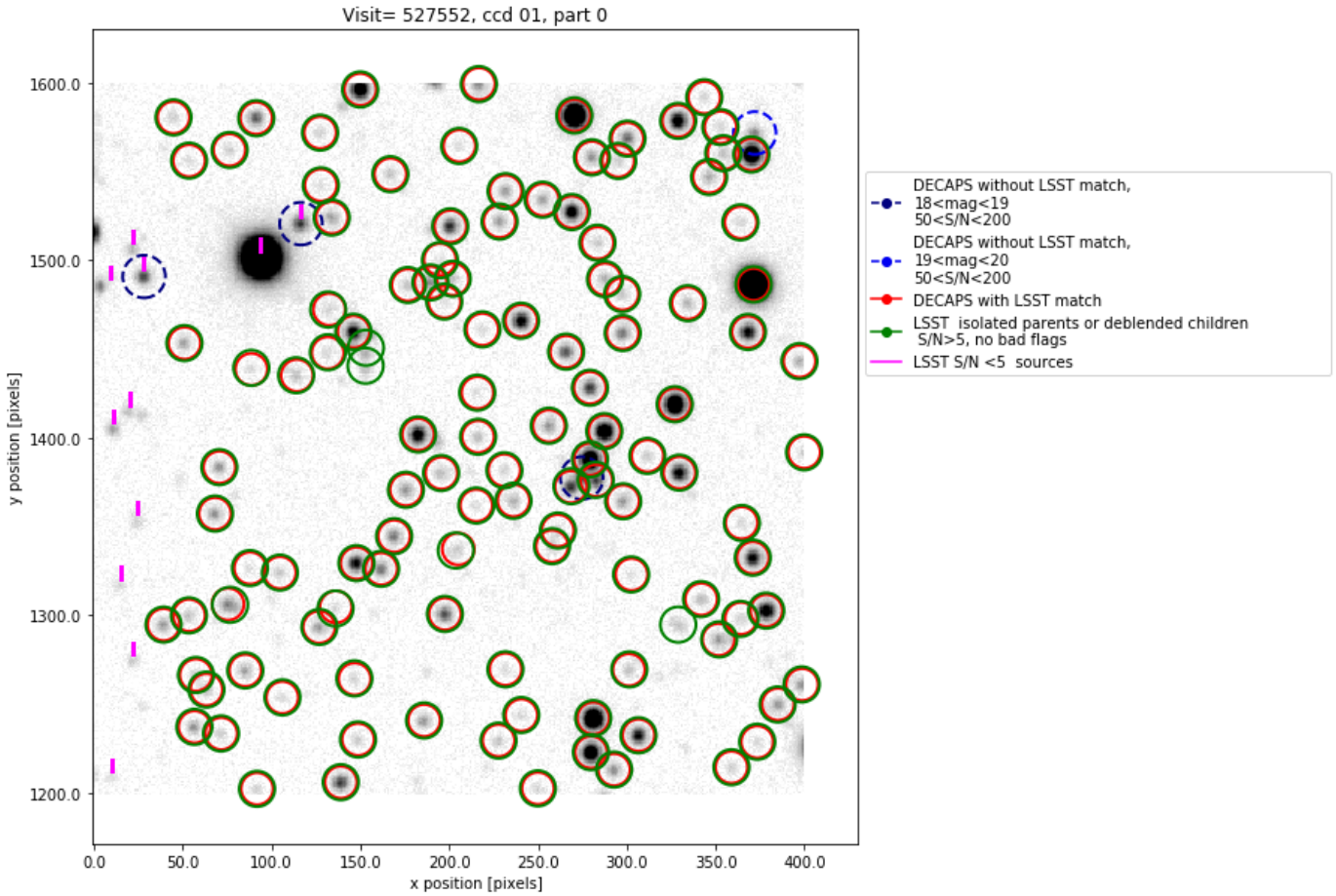

Figure 4.7: The same region as on Fig. 4.6. Green circles mark the position of retained LSST sources: isolated parents, or deblended children, with S/N $>$ 5, and no bad flags. Red circles mark the position of DECAPS detections with an LSST match. Vertical magenta dashes indicate LSST sources with S/N $<$ 5. Blue dashed circles mark location of DECAPS source without an LSST match. Note that eg. at (x,y) = 50,1490 an LSST source was detected, but since its S/N $<$ 5 it was not kept in the clean LSST catalog.

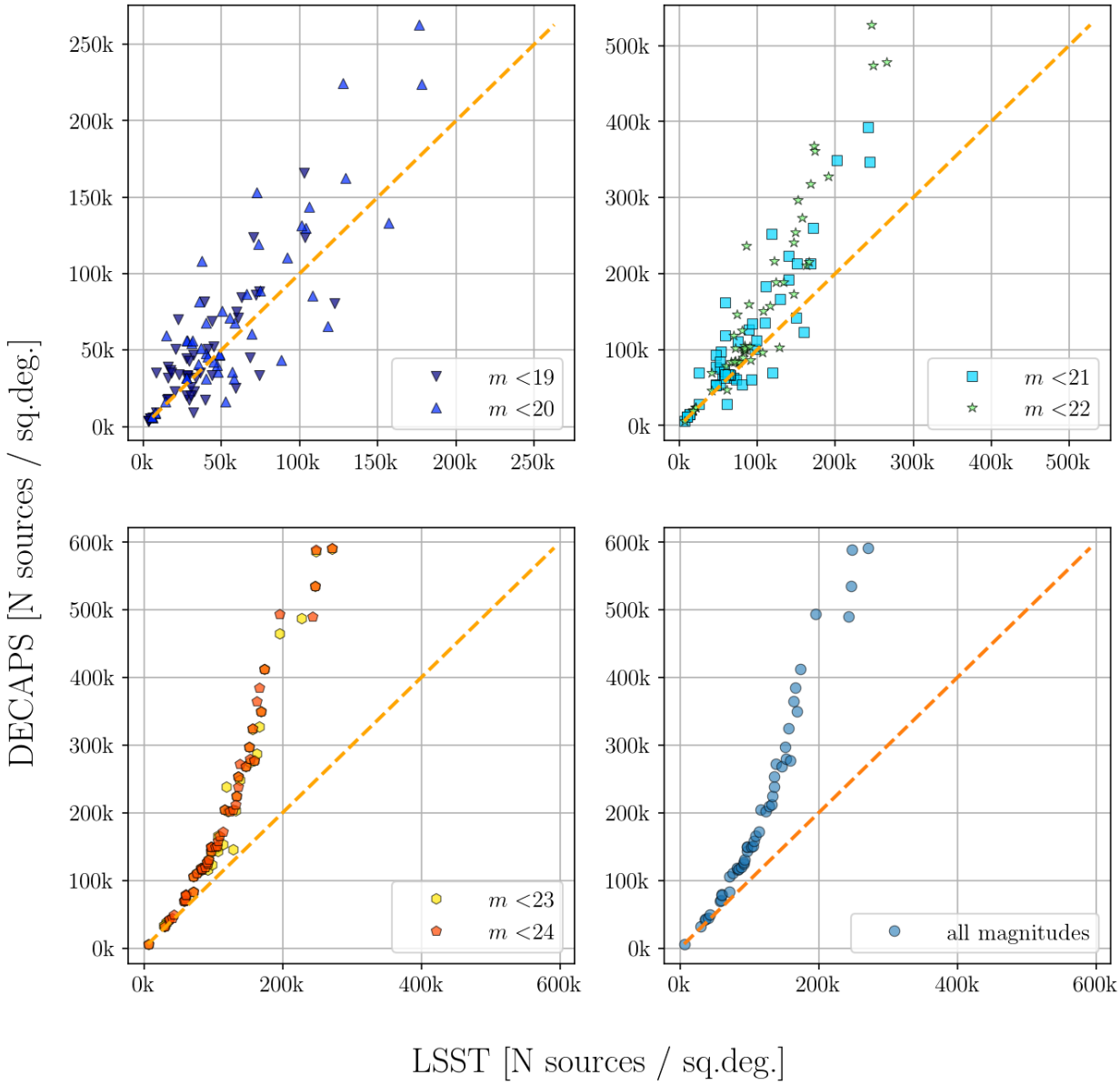

Figure 4.8: A plot of source count comparing LSST to DECAPS source catalogs of the same fields (visits) - each point per panel corresponds to a different visit. Clockwise from the upper-left panel we add progressively fainter sources, plotting the cumulative count up to a given magnitude, N(mag<cutoff).

Table 4.6: Summary of possible parentId and nchild combinations for blended sources in the LSST Science Pipeline. An example count in the final column is provided for visit 525814, a top 20% density region, which has the raw source count 235307. For that visit 16811 sources had bad flags, 49901 had S/N < 5, and in total 163093 were kept in the clean catalog.

parentID	nchild	type	decision	count
0	0	parent: isolated source	keep	104406
0	>0	blended source	remove	26981
!=0	0	deblended child	keep	103920
!=0	>0	failure case	remove	0

We compare LSST and DECAPS pipelines by cross-matching the source catalogs. For each DECAPS source we look for an LSST counterpart, and binning DECAPS sources along magnitudes we ask what is the completeness of LSST detections to DECAPS - what percentage of DECAPS sources has an LSST match within 0.5 ″(Fig. 4.9, top panel). We repeat this exercise the other way: looking for a DECAPS counterpart for each LSST source (Fig. 4.9, second panel). As expected, completeness decreases as a function of magnitude and increasing stellar crowdedness. We further characterize completeness by $\langle C_{18-20} \rangle$ - the mean completeness between 18-20 mag , and m_{50} - the magnitude at which completeness falls to a 50% level (see Fig. 4.10 - both exhibit a slight trend with stellar crowdedness. Given that the DECAPS pipeline detections are considered the gold standard, we also show that there is a very high degree of repeatability (>95%) for source detection between two epochs (Fig.4.11).

We illustrate on a CCD level relatively bright (18<m<20), high S/N DECAPS sources, that do not have an LSST match. In most cases we found that these sources have S/N<5 in the LSST catalog, which led to their exclusion from the analysis (see Fig. 4.7).

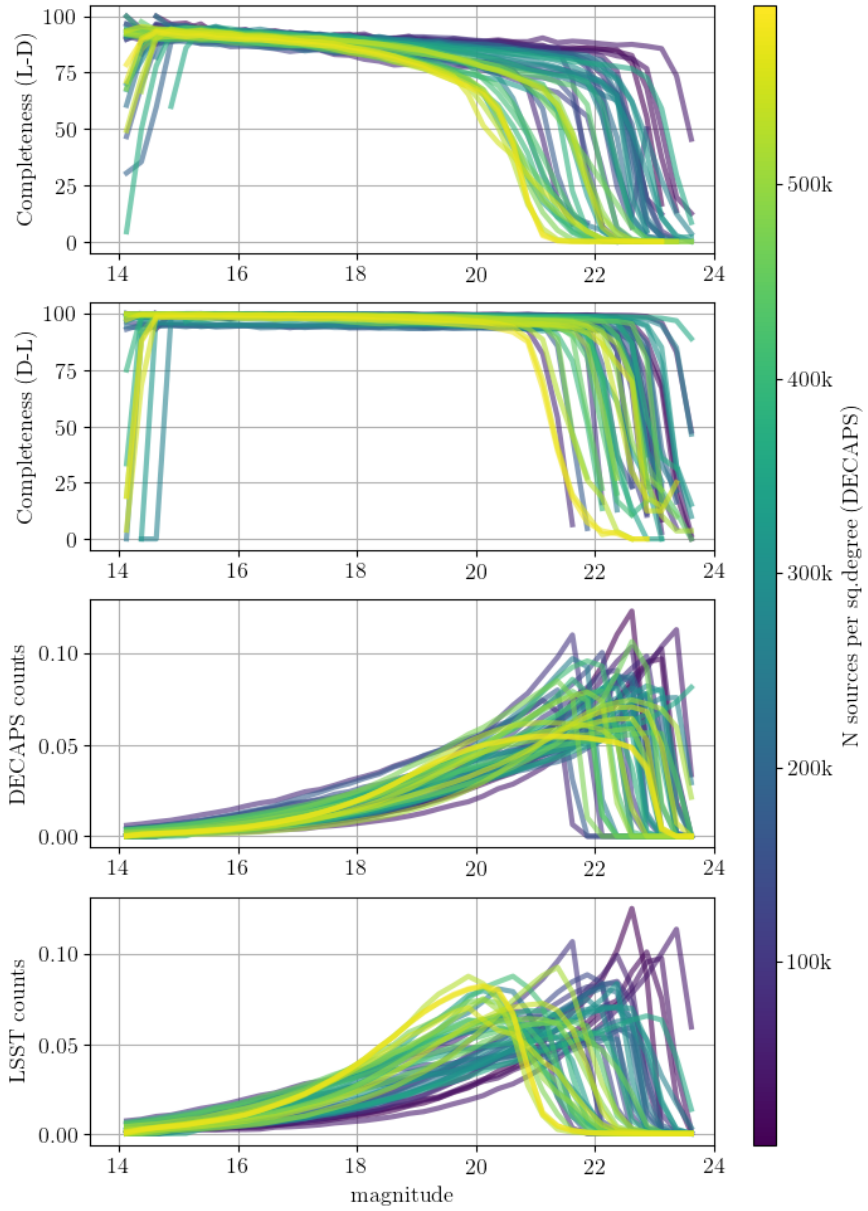

Figure 4.9: Top two panels show source-to-source completeness. The first panel is a measure of how complete is LSST catalog to DECAPS catalog (L-D), i.e. the fraction of DECAPS sources per magnitude bin that have an LSST match. The second panel shows an equivalent plot for the completeness of DECAPS to LSST (D-L), plotting the fraction of LSST sources that have a DECAPS match. The bottom two panels show the normalized source counts in the input catalogs. The LSST-DECAPS completeness falls off quicker than DECAPS-LSST, since DECAPS catalog contains more sources at fainter magnitudes (see Fig. 4.8). Different colors correspond to different level of stellar crowdedness, expressed in terms of the number of sources per square degree in DECAPS clean catalogs.

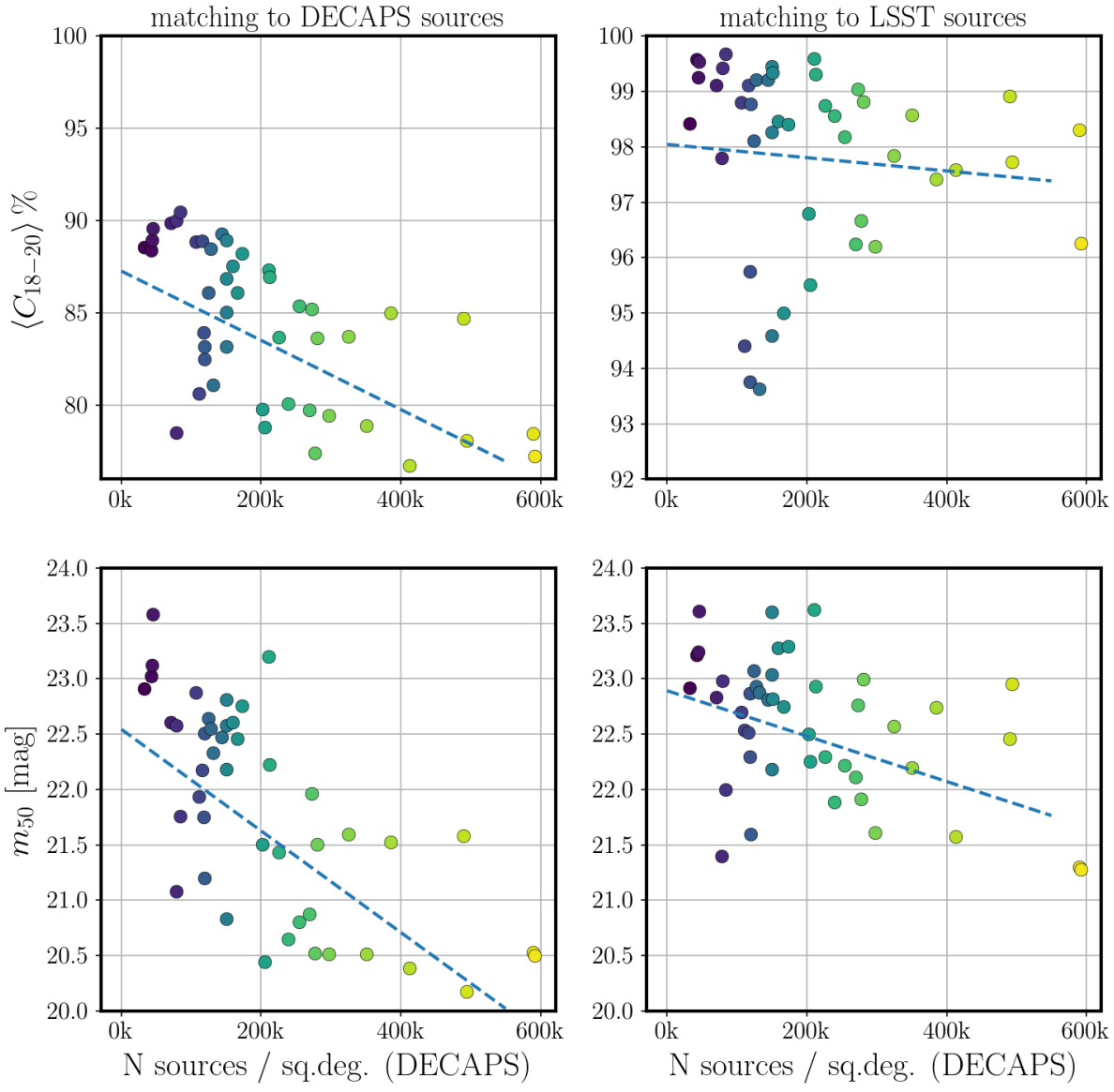

Figure 4.10: The mean completeness between 18 and 20 magnitudes (top two panels), and the magnitude at which completeness falls to 50% (bottom two panels). The panels on the left hand side correspond to the uppermost panel in Fig. 4.9, while the right hand side panels correspond to the second panel in Fig. 4.9. The color of all points corresponds to the stellar density. In each panel we overplot the linear best-fit to indicate the expected overall trend of decreasing $\langle C_{18-20} \rangle$ and m_{50} with source density.

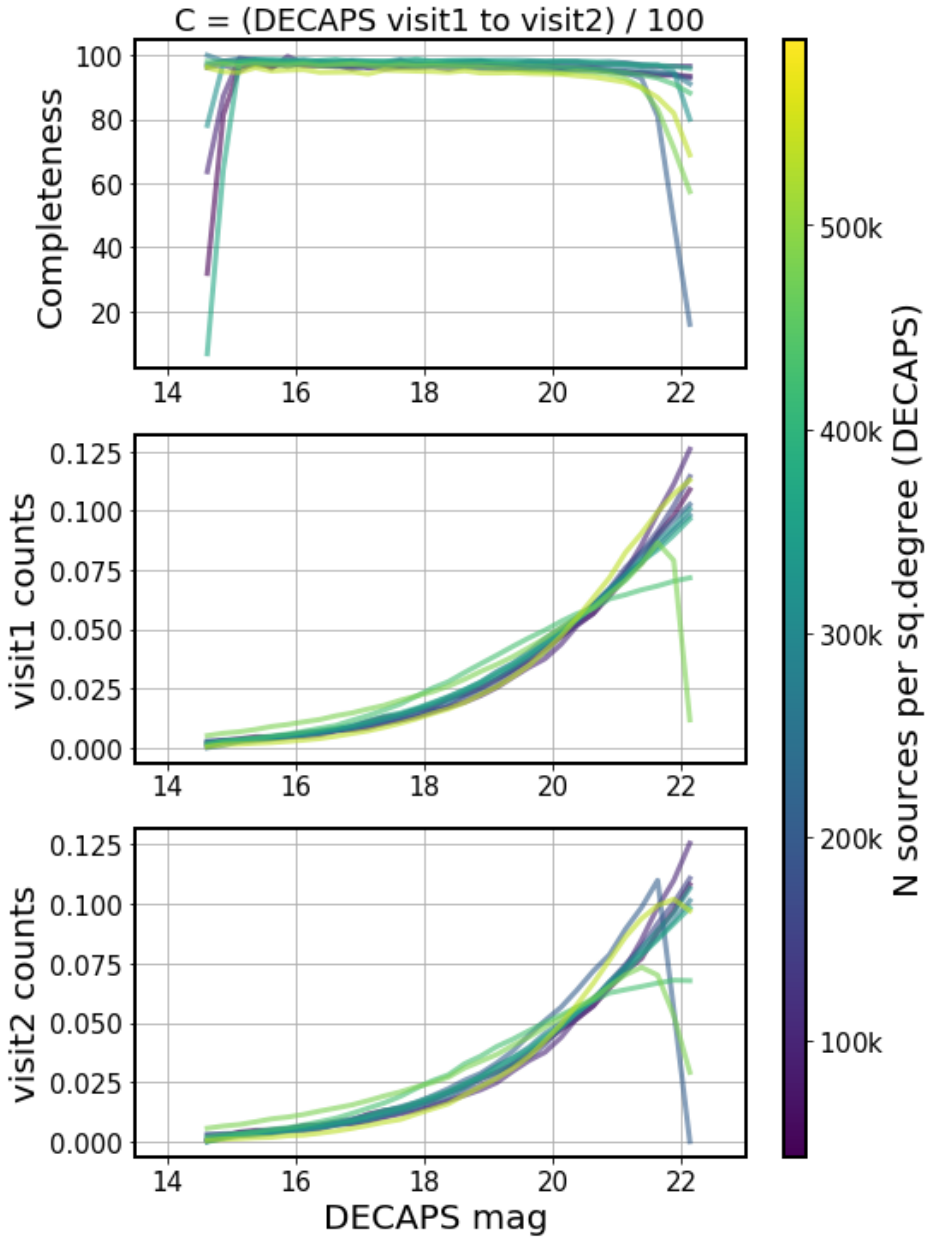

Figure 4.11: The same quantities as on Fig. 4.9, but corresponding to two different visits at the same location, testing the repeatability of DECAPS detections. The two visits (visit1, visit2) were chosen in the same filter and at the same location, and as in tests for completeness, we match source-by-source and consider the number of sources per magnitude bin in visit1 that do have a matching source in visit2.

4.5.2 Photometry

We compare the photometric accuracy between LSST and DECAPS, and photometric repeatability within each pipeline.

Cross-matching LSST and DECAPS source catalogs, we find that the median offset is stable as a function of magnitude in range 16-20 mag, on the level < 0.1 mag (top panel of Fig. 4.13). The spread of magnitude difference increases as a function of magnitude (Fig. 4.12, and middle panel of Fig. 4.13), < 0.05 mag at 20 mag.

We also test the photometric repeatability (LSST-LSST, Fig. 4.14 or DECAPS-DECAPS, Fig. 4.15) cross-matching source catalogs for visits at exactly the same location, with matching exposure time and filter, representing different epochs.

On Fig. 4.16 and 4.17 we compare the spread of photometric scatter between the two pipelines and the empirical measurement of noise from repeatability for a pair of visits at the same location.

The median error reported by either pipeline for either epoch is a measure of Poisson noise - the expected uncertainty in a repeated measurement. If e_{1L} and e_{2L} are LSST-reported error measurements for a given source for the two epochs, the estimated single-image uncertainty is

$$e_{12} = \sqrt{e_{1L}^2 + e_{2L}^2}/\sqrt{2} \tag{4.1}$$

The scatter between the two pipelines calculated either for epoch1 $\sigma_G(D1, L1)$, or epoch2 $\sigma_G(D2, L2)$. We find the estimated single-epoch photometric spread as:

$$\sigma_G(DL) = \sqrt{\sigma_{D1,L1}^2 + \sigma_{D2,L2}^2}/\sqrt{2} \tag{4.2}$$

The scatter between the two epochs within the same pipeline: $\sigma_G(D1, D2)$, or $\sigma_G(L1, L2)$, consists of reported Poisson noise σ_E and an additional systematic uncertainty σ_S :

$$\sigma_{LL}^2 = \sigma_S^2 + \sigma_E^2 \tag{4.3}$$

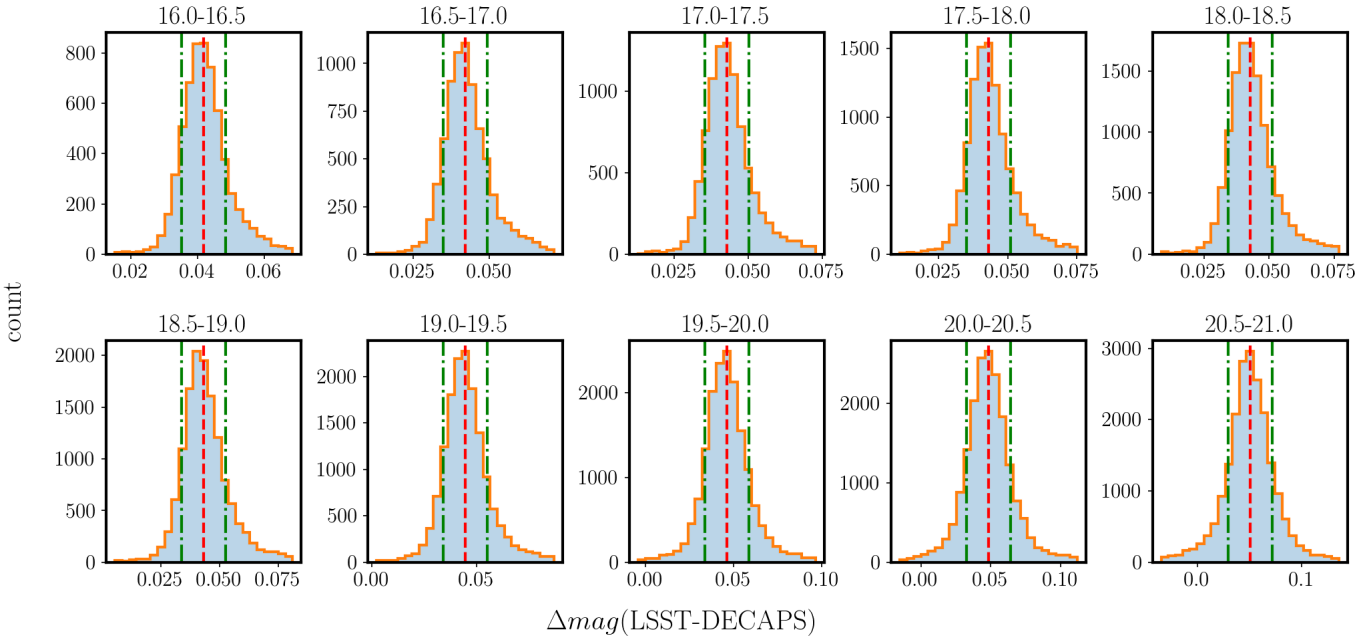

Figure 4.12: Cross-section of a difference in magnitudes between DECAPS and LSST for a visit 611970. Each panel contains the histogram of Δmag per DECAPS magnitude bin. The vertical red line corresponds to the median value of Δmag in that bin, and each histogram is limited between $\pm 4\,\sigma_G$. The vertical dot-dashed green lines mark the median $\pm\sigma_G$.

Thus we find the additional systematic uncertainty for LSST as:

$$\sigma_S = \sqrt{\sigma_{LL}^2 - \sigma_E^2} \tag{4.4}$$

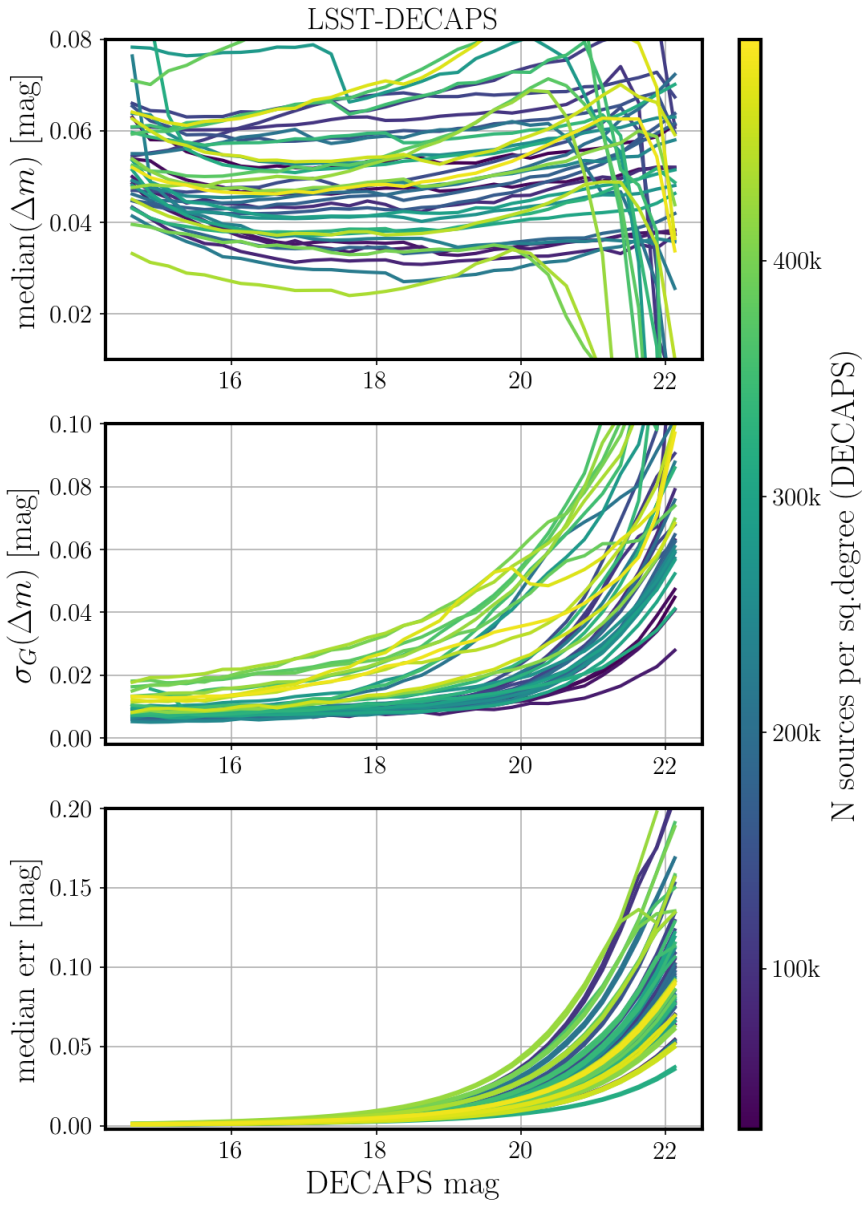

Figure 4.13: The measurement of photometric offset between DECAPS and LSST pipelines. For each visit we cross-matched source catalogs corresponding to LSST and DECAPS processing; Δm is the difference in magnitude reported between DECAPS and LSST for the same source. For each visit we bin sources according to their DECAPS magnitude. On three panels we plot the binned statistics : median(Δm), $\sigma_G(\Delta m)$, and median photometric uncertainty.

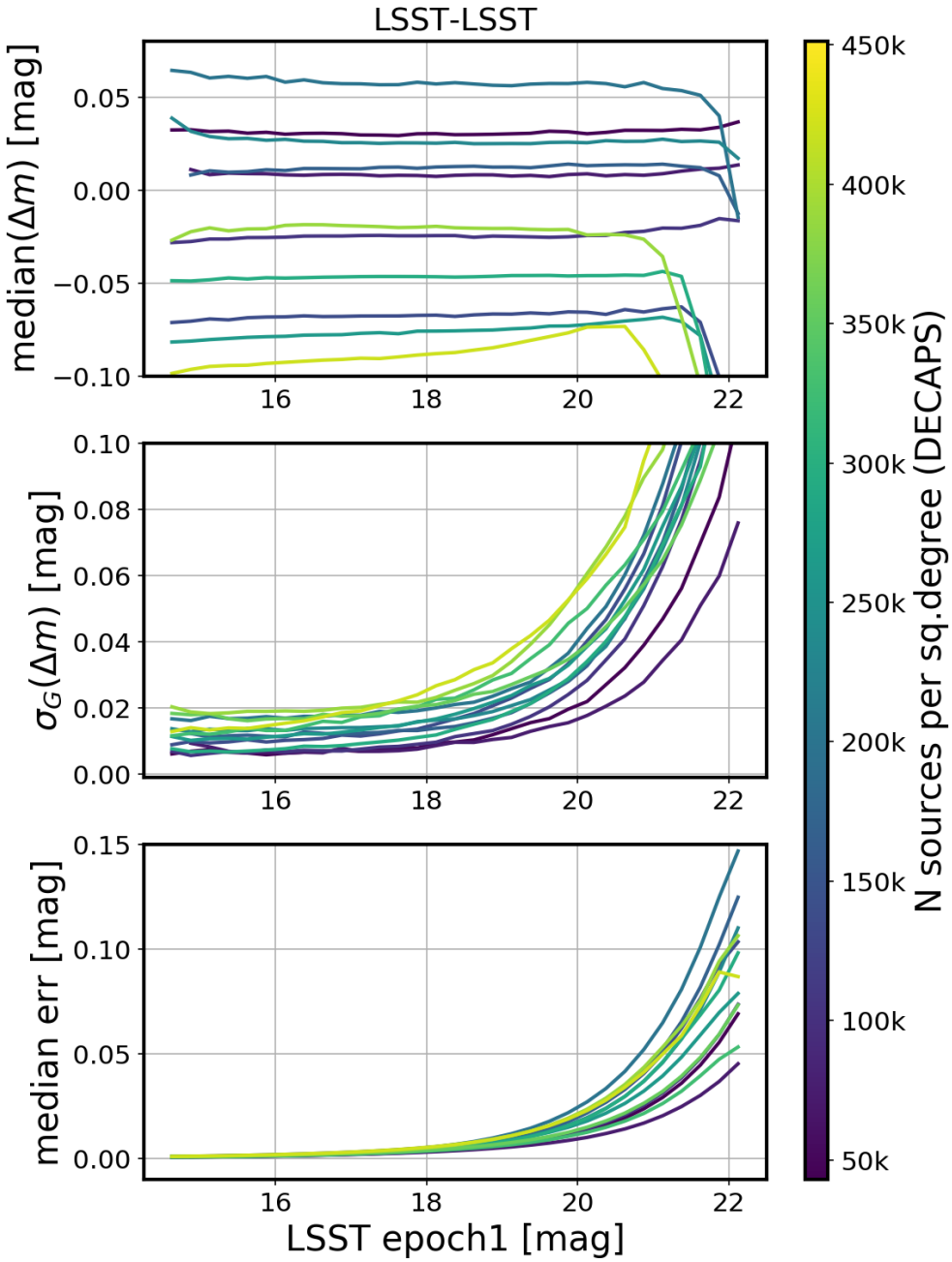

Figure 4.14: The repeatability test of the LSST pipeline. We cross-match the source catalogs for each visit. These two brightness measurements for the same source are akin to a two-epoch light curve. Since inherently variable sources constitute a small fraction of all stellar objects, and the majority of stars are not variable, the spread in the difference of measured magnitudes would correspond to the empirical measure of noise. On the panels we plot, from top to bottom: median photometric offset, the robust interquartile-based measure of standard deviation σ_G, and the median reported measurement uncertainty.

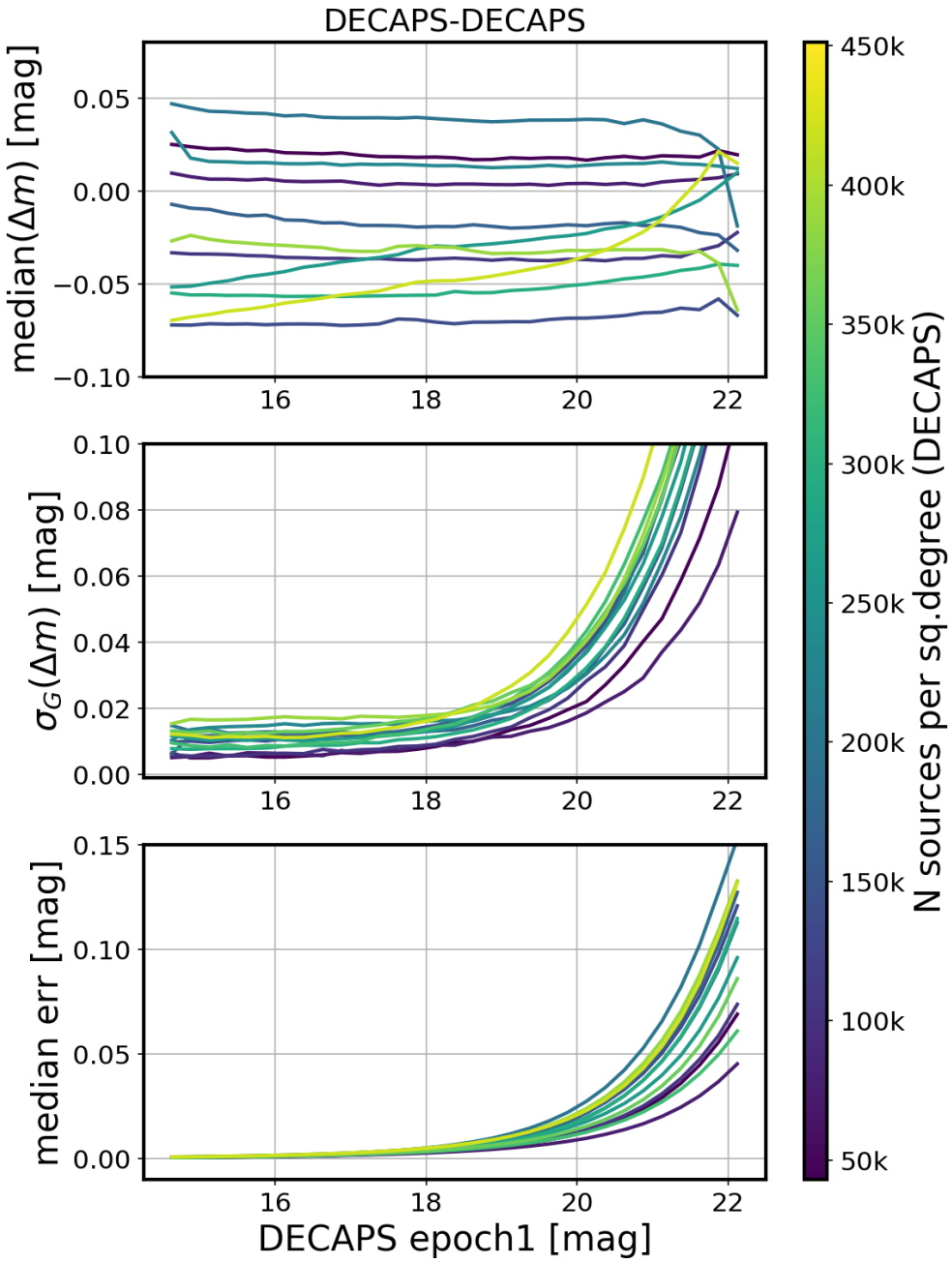

Figure 4.15: The repeatability test of the DECAPS pipeline, as Fig. 4.14.

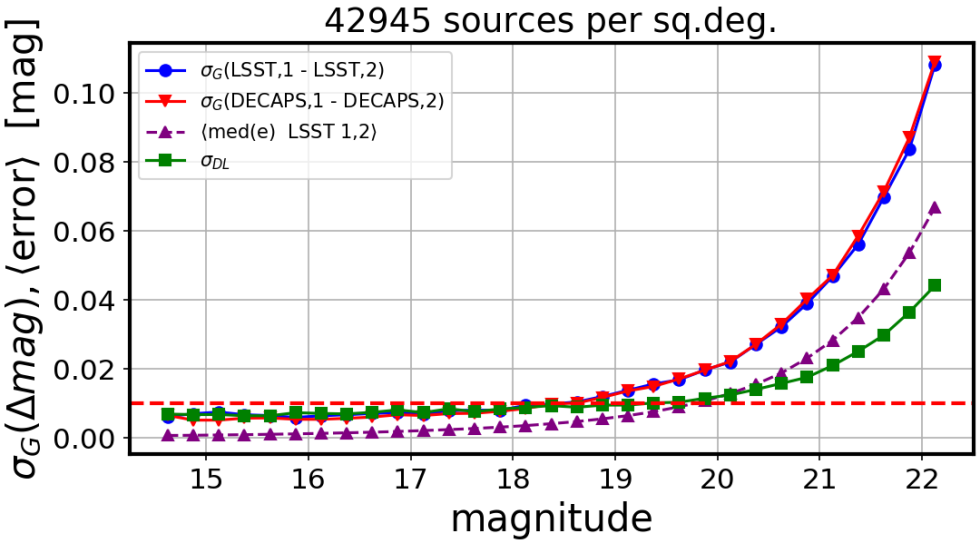

Figure 4.16: Analysis of the photometric spread with two visits in a low density region: 525846 (epoch1) and 530012 (epoch2). The solid blue and red lines represent the spread in photometry within each pipeline ($\sigma_G(L1, L2)$, $\sigma_G(D1, D2)$) . The purple dashed line in the middle traces the average reported error between the two epochs, e_{12} (Eq. 4.1), which is a measure of the Poisson noise. Finally, the bottom solid green line with square markers is the spread in photometry between the two pipelines, $\sigma_G(DL)$ (Eq. 4.2). The difference between the purple dashed and red/blue lines is a measure of an additional systematic uncertainty (see Eq. 4.4, and Fig. 4.17).

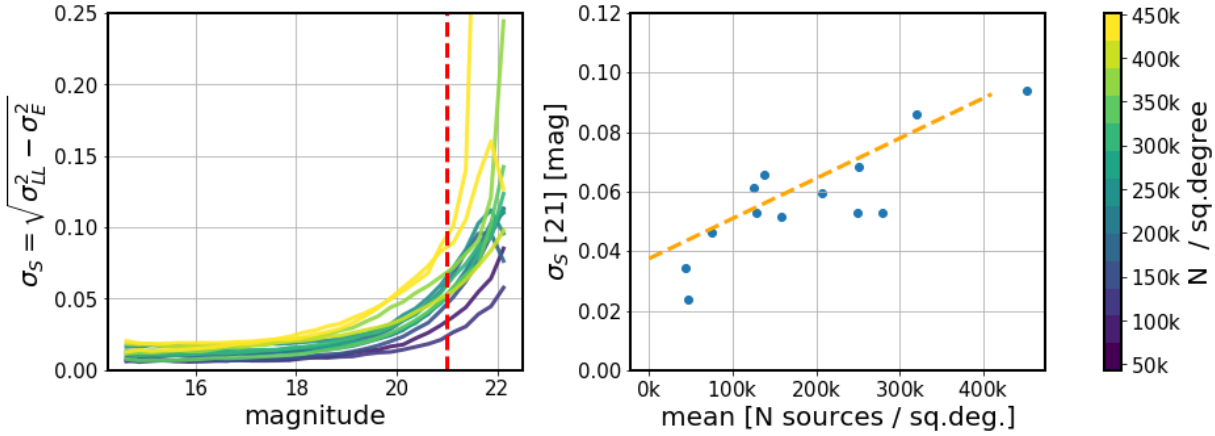

Figure 4.17: The left panel shows the measure of an additional systematic uncertainty σ_S - an excess between reported Poisson noise σ_E, and the estimated single-epoch photometric spread σ_{LL} for the LSST pipeline as a function of magnitude (Fig. 4.16). The vertical line marks the level of 21 mag. The right panel shows σ_S at 21 mag as a function of the measured DECAPS stellar density.

4.6 Astrometry

Astrometry pertains to the measurement of the position of sources in the absolute World Coordinate System (WCS). Accurate and precise astrometry enables eg. catalog cross-matching, and over long term measurement of proper motions.

To measure the repeatability of astrometric measurement within each pipeline, we consider pairs of visits at the same location, exposure time, and filter, but observed at different epochs. For both pipelines we estimate the spread in astrometric differences: $\Delta\alpha$, $\Delta\delta$, by robust interquartile-based measure of standard deviation, $\sigma_G = 0.7413 * (q75 - q25)$ where $q75$,$q25$ are 75th and 25th percentiles, respectively. As can be seen on Figs. 4.19 and 4.20, σ_G is the width of the distribution along dimensions of $\Delta\alpha$, $\Delta\delta$. There is a slight increase in spread of astrometric offset as a function of magnitude (Fig. 4.18), and to avoid including faint (and therefore, more difficult to measure) sources, we limited the object brightness at 19th mag. We also considered pipeline-to-pipeline offset, but since both DECAPS and LSST use GAIA for astrometric calibration, the information contained would be due to details of implementation (see Fig. 4.22)

To investigate the possible dependence on stellar density, we measure the spread in epoch-to-epoch astrometric offset for pairs of visits at increasing levels of crowdedness. The dependence turns out to be not very strong, with LSST astrometric repeatability on the level of 10-30 miliarcsec (Fig. 4.21).

4.7 Conclusions

We performed pipeline comparison tests with DECAPS and LSST pipelines, comparing source counts, photometry, and astrometry.

The LSST pipeline easily handles regions of density up to 200 thousand per sq.deg., and then there is a gradual degradation, mostly in completeness, when progressing towards higher densities. The mean 18-20 mag completeness of LSST to DECAPS detection is 85% at the edges of the Galactic plane region (top 5% density, assuming the single-visit LSST depth

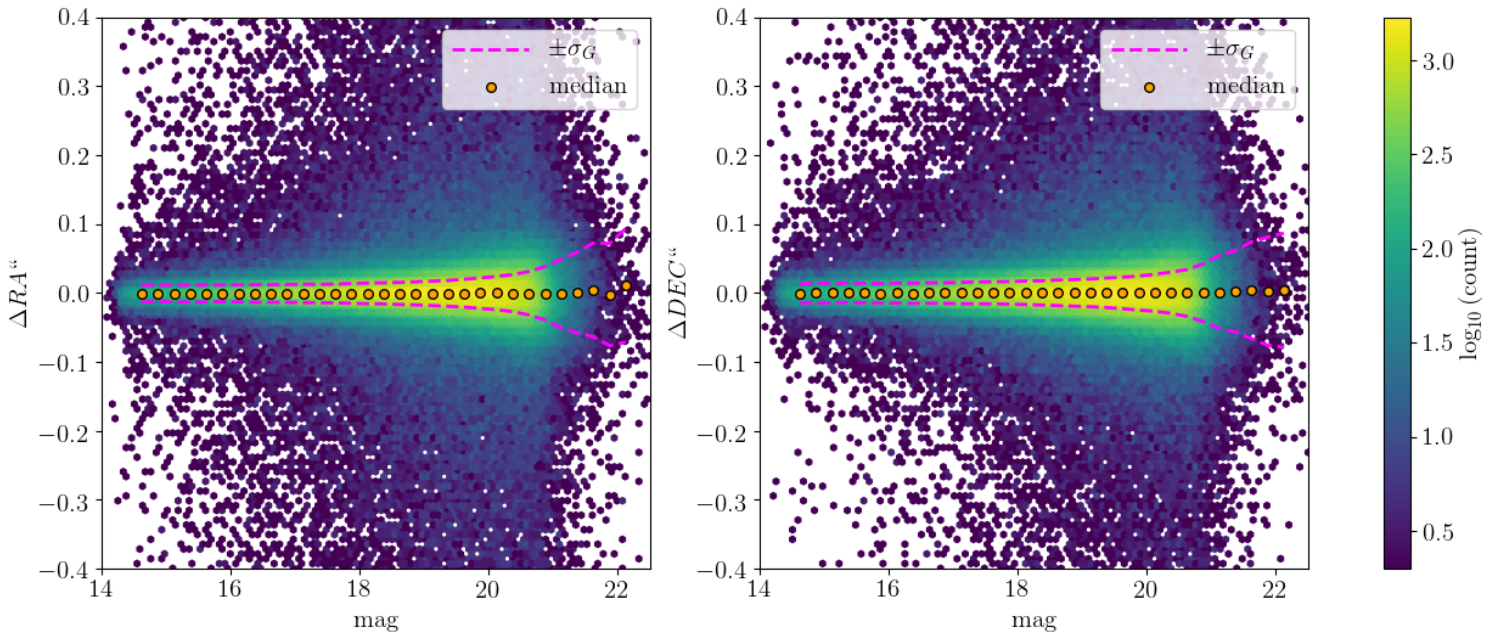

Figure 4.18: The difference in RA,DEC for visits 644074,644070 processed by LSST. According to Galfast simulation this location is a top 1% stellar density region, with DECAPS measured 590 704 sources per sq.deg. The same visits are compared on Figs. 4.19 and 4.22 with mag $<$ 19 cutoff.

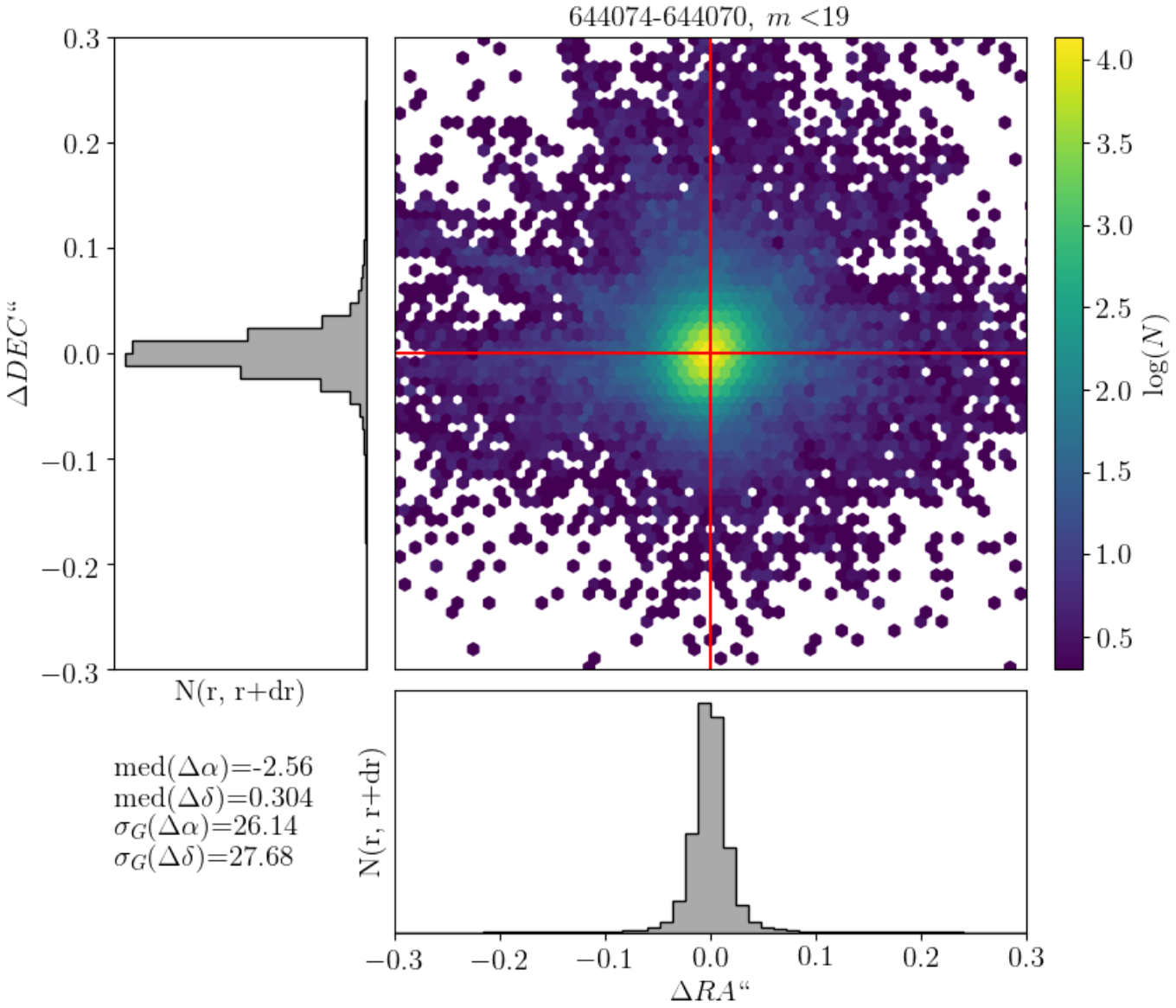

Figure 4.19: The difference of LSST processing for RA,DEC for visits 644074,644070 : a top 1% region according to Galfast, where DECAPS measured 590 704 sources per sq.deg. We select sources brighter than 19 magnitude. For all other pairs the offsets are also centered on zero with a similar spread (Fig. 4.21)

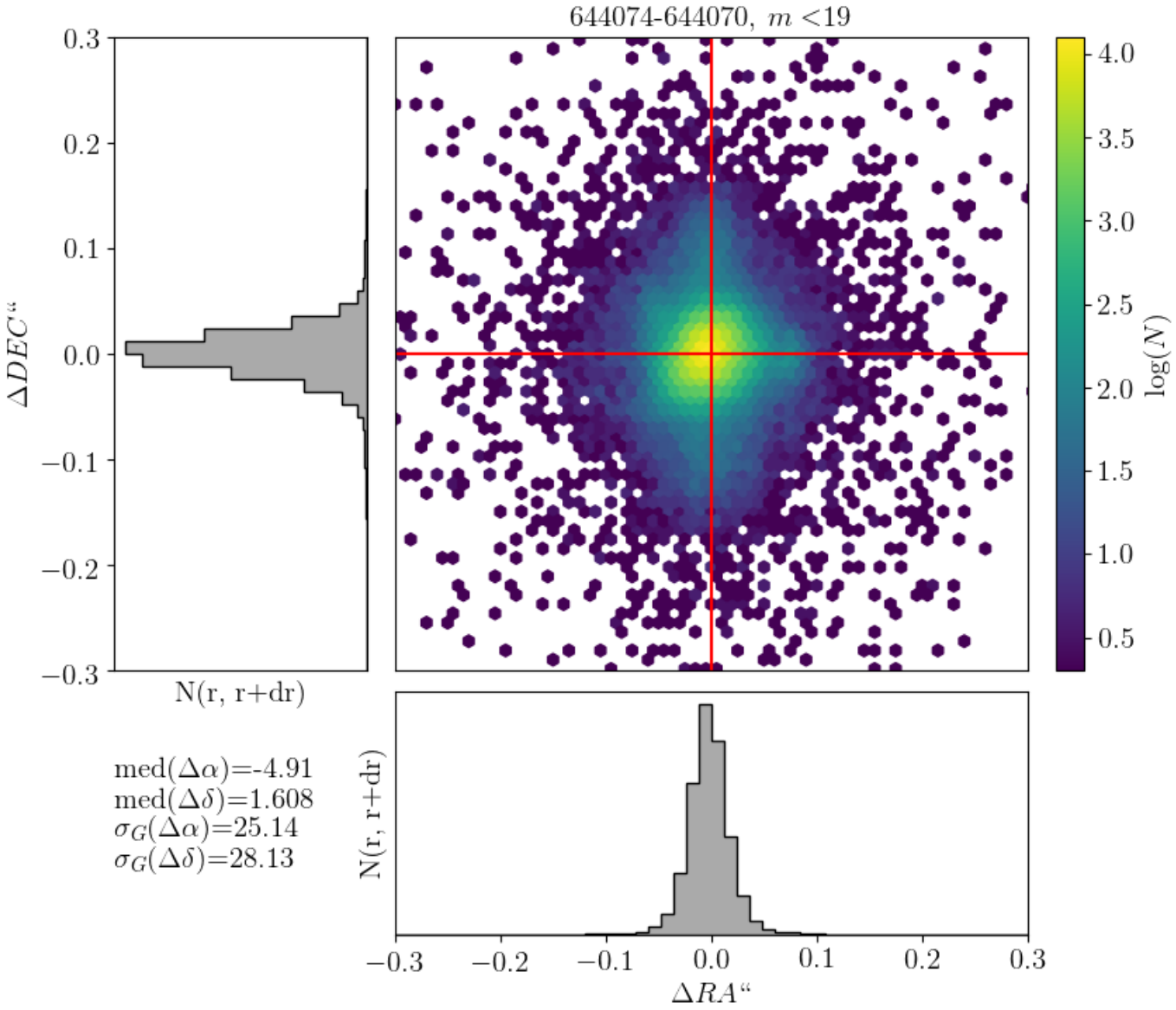

Figure 4.20: The difference in RA,DEC for the same visits as in Fig. 4.19, but comparing DECAPS single-epoch catalogs. The spread of $\Delta\alpha$, $\Delta\delta$ is wider than for equivalent visit pairs processed by the LSST Science Pipelines.

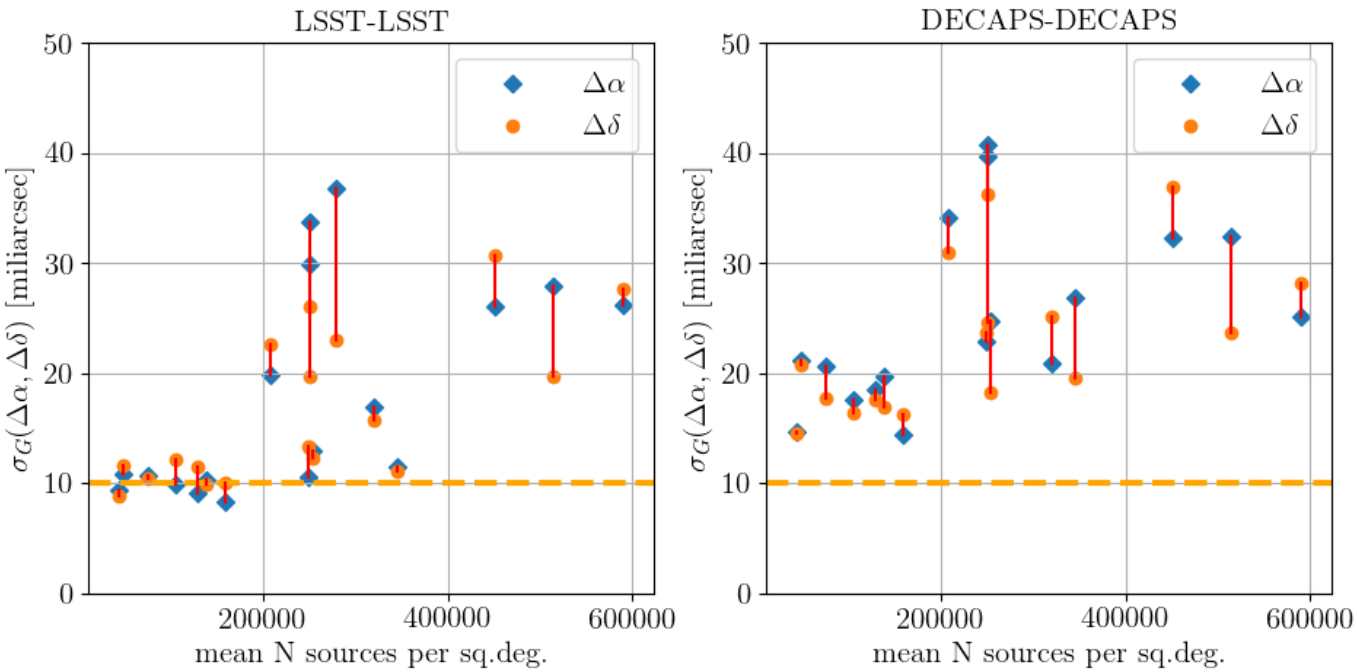

Figure 4.21: Summary of LSST-LSST and DECAPS-DECAPS astrometric repeatability, with magnitude cutoff at 19 mag.

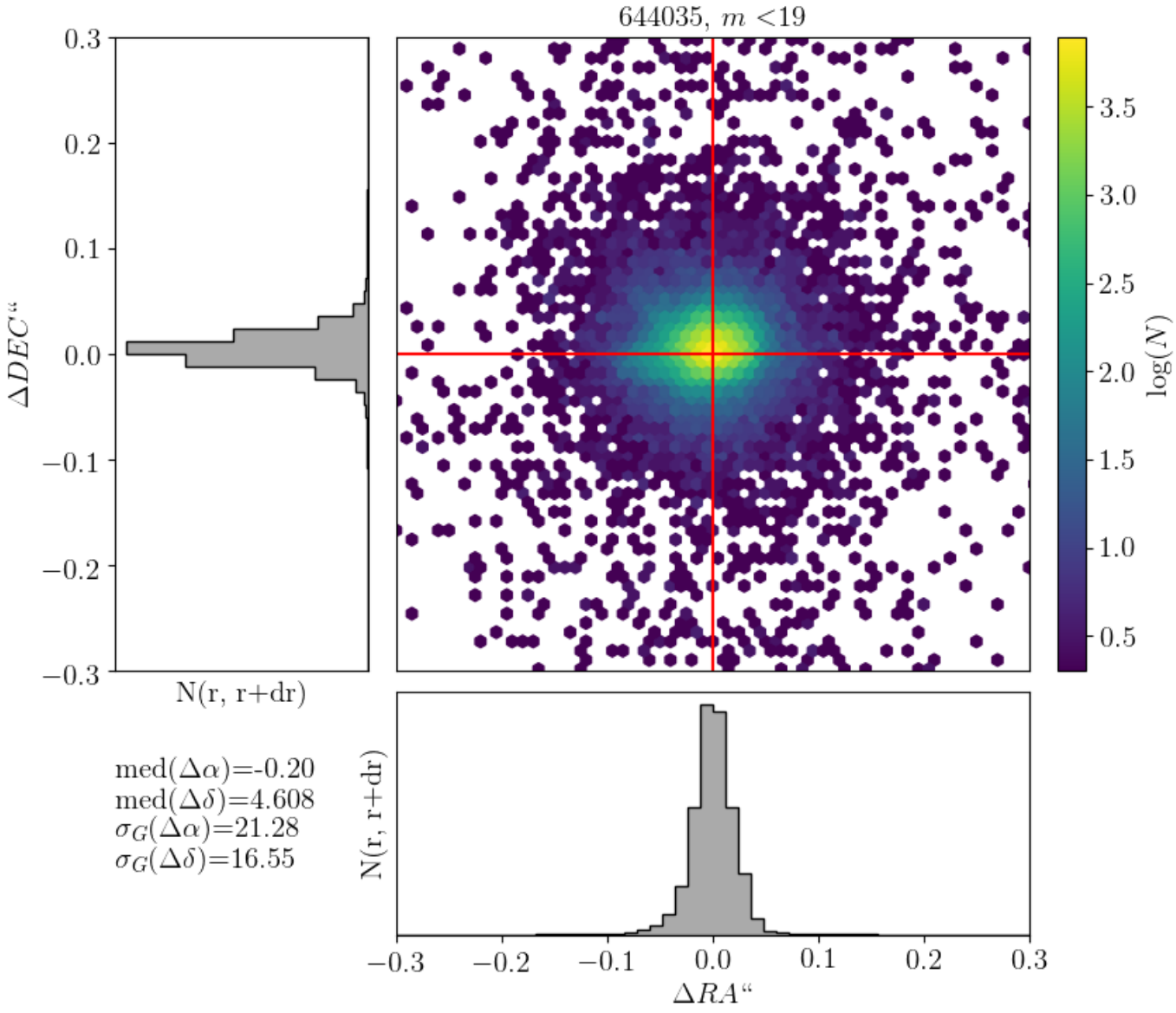

Figure 4.22: The LSST-DECAPS astrometric offset for visit 644035, with 200 000 sources per sq.deg. measured by DECAPS, in a Galfast 10% density region. In DECAPS pipeline the astrometry was tied to 2MASS-GAIA or GAIA depending on the visit number (see Fig.12 in Schlafly et al. (2018)). LSST pipeline on the other hand used solely GAIA for astrometric calibration. For this visit the offset is due to precision, rather than different astrometric standards, since both DECAPS and LSST used GAIA data for astrometric calibration.

of 24.5 mag). Astrometric repeatability within the LSST pipeline comparable to DECAPS, without a very strong density dependence.

However, this leaves roughly 8% of the Wide-Fast-Deep area with degraded completeness in single visit processing, and 15% of the area once the data are coadded. Improvements to the software to support densities of 500k per sq. deg. would be required in order to obtain high completeness over all of WFD; support for even higher densities would be required for the Galactic plane region.

Future work:

- using the simulated sky images with StarFast image simulator[13] where the true source count and position are known, rather than measured

- considering the width of the stellar locus (w-color) on the g-r vs. r-i diagram - this would be helpful if photometry gets corrected for extinction (see eg. Fig.2 in Ivezić et al. (2004))

- exploring the magnitude difference Δm - separation Δd in a catalog cross-matched with self (i.e. for each source, finding the nearest neighbor). This $\Delta m - \Delta d$ space for DECAPS objects that do / do not have an LSST match may yield interesting insights into the nature of mismatches.

[13] https://dmtn-012.lsst.io

Chapter 5
SUMMARY AND CONCLUSIONS

We summarize the main results of each chapter of this thesis, and draw conclusions, describing possible future work in the field of quasar variability.

5.1 Summary

5.1.1 Solving the puzzle of excess quasar variability on monthly timescales

Quasars have been known as variable extragalactic sources for over 50 years (Matthews & Sandage, 1963; Schmidt, 1963). In Chapter 2 we addressed the issue of anomalously large variability for Catalina Real-time Transient Survey (CRTS) quasars found by Graham et al. (2014). The predicted 0.12 mag rms variability on week-to-month rest-frame timescales, was at odds with the earlier Sloan Digital Sky Survey (SDSS)-based results of approximately 0.06 mag rms (eg. MacLeod et al. 2010). This put into question the viability of employing the well-tested damped random walk (DRW) model to describe quasar light curves, and since observed timescales are directly related to physical processes, it affected the hitherto accepted explanations for observed variability (Graham et al., 2017; Kozłowski, 2016). Re-analyzing the SDSS and CRTS light curves for quasars as well as non-variable stars, we found that using the CRTS data for non-variable standard stars, we produce a non-zero structure function. This was caused by the incorrect estimate of photometric uncertainties returned by the CRTS image reduction pipeline (as later confirmed by Graham et al. 2017). When the uncertainties were corrected with factors derived from standard stars with similar colors to quasars, we were able to reconcile the levels of quasar variability as inferred from the CRTS and the SDSS data. Additionally, using the PTF data for the same objects we confirmed that the DRW is an appropriate model for the underlying variability, and that

prior CRTS-based claims of 54 day characteristic timescale (Graham et al., 2014) were due to inaccurate photometric uncertainties.

5.1.2 Extending quasar light curves with PS1

Using the DRW model, in Chapter 3 we repeated the analysis of Kozłowski, Szymon (2017), proving with simulations of quasar light curves that by extending their baselines, we could improve the MacLeod et al. (2010) (hereafter M10) SDSS-based analysis. Although various survey data was available to complement the SDSS observations of Stripe 82 taken between 1998-2004, including CRTS, Palomar Transient Factory (PTF), or Zwicky Transient Facility (ZTF), we found that only Pan-STARRS1 (PS1) offers sufficiently similar bandpasses to avoid photometric color transformations. Such color offsets would unnecessarily introduce additional uncertainty, given that for the majority of light curves there is no temporal overlap between SDSS and other surveys, and thus no accurate way of determining the offsets on the object-by-object basis (apart from CRTS). Furthermore, PTF, CRTS and ZTF are shallower than SDSS and PS1 (by 1.5 mag), which results in larger photometric uncertainties for quasars, which are generally dominating the faint end of the source distribution. Utilizing PS1 DR2 data taken between 2009-2014 we were able to extend the SDSS baselines by over 50%. Similarly to M10, we found that the amplitude of variability is anti-correlated with quasar luminosity with power law slope 0.118 ± 0.003 dex, and correlated with black hole mass (slope 0.118 ± 0.008 dex). Also, the damping timescale is primarily dependent on the black hole mass, with a power-law index of 0.141 ± 0.019 dex. This is consistent with the driving mechanism being in part the Eddington ratio, which is anti-correlated with the amplitude of variability (slope 0.208 ± 0.03 dex). With the available data we cannot distinguish between various proposed models of quasar variability, eg. X-ray reprocessing (Kubota & Done, 2018), or clumpy disk inhomogeneities (Ruan et al., 2014). By comparing the SDSS and SDSS-PS1 based DRW fit parameters, we selected 40 quasars for which SDSS-PS1 timescales are over ten times longer, and amplitudes are twice as large, as for the SDSS-based model fits. Visual inspection and calculating median offsets between the SDSS

and PS1 portions confirmed that these exhibit large (over 0.5 mag) variations in brightness, not unlike the changing-look quasars (CLQs). Indeed, 5 of the candidates are confirmed in the literature as CLQSO (LaMassa et al., 2015; MacLeod et al., 2019).

5.1.3 Preparing for LSST: crowded fields analysis

Quasars are point sources, and as shown from the CRTS example, accurate photometry and correct handling of all sources of uncertainty is crucial for an appropriate interpretation of observed variability. Therefore, in Chapter 4 we investigated the performance of the Large Synoptic Survey Telescope (LSST) science pipelines in crowded stellar regions. We reprocessed the DECam Plane Survey (DECaPS) data in sky regions chosen to reflect a diversity of crowdedness levels, from 200 000 to over 2 million sources per $\deg^2$. We compared the results of using LSST science pipelines to the cutting edge DECaPS pipeline (Schlafly et al., 2018). We found that although there is a gradual degradation in completeness: for a region of moderate density of 200 000 sources per $\deg^2$, between 18th and 20th mag on average almost 80% of DECaPS sources are also detected by LSST, but this drops to 50% for fainter sources around 21.5 mag. We showed that both LSST and DECaPS are in a regime dominated by systematics above 19th mag. We proved that the astrometric scatter for LSST is in the desired range of 25-30 miliarcsec, and has a weak dependence on stellar densities. We concluded that it is possible to achieve the required performance level if small improvements are implemented with the existing science pipelines.

5.2 Future work

5.2.1 Reanalysis of existing datasets with new and uniform pipeline

Our work on the analysis of photometric uncertainties in Chapter 2 points to a deeper underlying problem in sky surveys targeted for fast transient detection, such as CRTS - the image analysis pipeline might not be ideally suited for analysis of long-term brightness variations. Small effects not accounted for, such as seasonal airmass fluctuation (Drake et al.,

2013), may affect long-term trends inferred from the data (eg. see Appendix B in Suberlak et al. 2017). An improvement in photometric accuracy and repeatability on longer timescales is a key to the proper interpretation of quasar time series data, and future surveys such as LSST will have to take this into consideration when designing the image processing pipeline. Reanalysis of the existing survey data can aid not only in the proper estimates of photometric uncertainties, but also allow to detect fainter sources by performing forced photometry, or applying a better image reduction pipeline to the existing datasets. For example, Baldassare et al. (2018) found that it was possible to identify previously unknown low-luminosity AGNs in galaxies with improved image differencing and forced photometry on PTF data. Also, AlSayyad (2016) reprocessed the SDSS Stripe 82 dataset with the LSST science pipelines and improved the detection threshold by over 0.5 mag. As Sartori et al. (2019) showed with a novel forward-modeling approach, with deeper data more quasars can be detected at only certain epochs (eg. in the bright state). Forced photometry allows to apply a Bayesian method to give an upper estimate on the non-detection of flux, which is more informative than a complete lack of observation (Hogg & Turner, 1998). Thus we propose to use the LSST science pipelines to reanalyze the existing datasets (SDSS, PTF, CRTS, ZTF, PS1), accounting for all long-term effects (such as airmass change and other systematics), perform deep coaddition and forced photometry, apply the Bayesian prior for faint sources (below 5σ detection threshold), and thus provide a homogeneous scientific dataset that could be used to extend quasar light curves.

5.2.2 Exploring the departures from the DRW description

In Chapter 3 we assumed that the DRW model is an appropriate description of quasar variability on months-to-years timescales, but both Zu et al. (2013) and Kozłowski, Szymon (2017) call it into question. While Zu et al. (2013) did not find that a more complicated Gaussian process kernel (such as Matérn or Pareto-exponential) was more appropriate than the pure DRW process, Kozłowski, Szymon (2017) showed that with SDSS sampling, the output timescale or amplitude will be biased if the covariance departs from the DRW. We

propose to repeat this experiment, simulating light curves spanning a range in the exponent β of the power exponential covariance matrix ($\beta = 1$ for DRW). Exploring β from eg. 0.4 to 1.6, with LSST-like baseline and cadence, we would calculate SF for each light curve, thus avoiding the parametric fitting, and recover β from SF. That way we would determine whether changing the baseline (eg. 2,5,10 year segments), or cadence (eg. 0.25, 1, 10, 20 days) affords better recovery of the underlying β. Furthermore, using the method of Kozłowski (2017a), we could estimate the SF-based decorrelation timescale, and find whether it depends on the underlying β.

5.2.3 Multi-band DRW fitting with a Gaussian process engine

In Chapter 3 our analysis of quasar light curves involved combining SDSS r-band with PS1 r-band data. M10 used all five SDSS bands, treating each band as an independent light curve, in effect analyzing ∼35 000 light curves for ∼6500 objects. Resulting DRW parameters were correlated independently per band with quasar physical properties such as black hole mass or quasar luminosity. M10 reported the band-averaged coefficients as the final result. Meanwhile, Hernitschek et al. (2016) showed that fitting the DRW model simultaneously to all PS1 DR1 bands (similarly to VanderPlas & Ivezić 2015 extending the Lomb-Scargle periodogram to multiple-bands) can help utilize all the data. This approach has never been tested on a combined SDSS-PS1 dataset. Other methods, such as multi-band template fitting, may be appropriate only for objects whose light curve shape is more predictable, such as RR Lyrae (eg. see Sesar et al. 2017; Stringer et al. 2019). Based on comparison of colors for standard stars in Stripe 82, the most similar filter between SDSS and PS1 is the r-band. Considering the difference between SDSS and PS1 magnitudes for a given band (Δm) as a function of SDSS magnitudes for that band object-by-object, for the r-band there is a slight slope in Δm vs m (less than 0.1%), and an average offset below 1% (0.01 mag). In this comparison, the i-band is second best, with a larger slope in $\Delta m - m$ (approximately 5%) and a larger offset (0.1 mag). The PS1 g and z bands are very different from the SDSS counterparts, and could not be employed without photometric transformations. However,

the method used by Hernitschek et al. (2016) brings all photometry to a fiducial band, assuming similar variability in each band, which may not be the case (see eg. Fausnaugh et al. 2016 for reverberation mapping of NGC 5548, showing that each band has a slightly delayed version of the signal). It would be advantageous to test the fiducial-band DRW fitting using a Gaussian process engine, and explore whether this method would improve on the fidelity of DRW parameter retrieval with the existing data.

5.2.4 Quasar variability from structure function analysis using SDSS-PS1 data

Towards the end of Chapter 3, we identified the outliers in the space of SDSS-PS1 versus SDSS DRW fit parameters. These objects had a similar morphology - large changes in brightness in optical passbands (over 0.5 mag) - to previously reported examples of changing-look Quasars (CLQs, eg. MacLeod et al. 2019). As both empirical models (Sartori et al., 2019) and observations (eg. SDSS-POSS, MacLeod et al. 2012) show, the probability of observing a given magnitude change for quasars is a function of time separation between two considered epochs. In particular, MacLeod et al. (2012) employed the structure function description to analyze the combined SDSS-Palomar Observatory Sky Survey (POSS) dataset, which availed an unprecedented 50 year baseline. Since PS1 data is available for the majority of SDSS quasars in the DR9, it would be advantageous to repeat the SDSS-POSS experiment using the PS1, and soon, LSST data, to hunt for extremely variable quasars (EVQs, Rumbaugh et al. 2018).

5.2.5 Spectroscopic follow-up of Changing-Look Quasar candidates

Quasar brightness can depart from the stochastic oscillations about the mean for various reasons. Two most commonly considered phenomena that involve large (over 0.5 mag) changes in observed quasar flux are tidal disruption events (TDEs) and CLQs. TDEs are most likely caused by a sudden outburst of radiation caused by a star presumably passing within a tidal radius of an SMBH, where roughly half of the stellar mass that remains bound to the SMBH feeds the accretion disk (Holoien et al., 2019; Merloni et al., 2015; van Velzen et al.,

2019). TDEs are usually seen as a large optical brightening by several magnitudes, with a characteristic fall-off proportional to $t^{-5/3}$ (eg. Blagorodnova et al. 2017; Hung et al. 2017). On the other hand, CLQs appear as more long-term changes in the optical output of a quasar, accompanied by a sudden (dis)appearance of broad Balmer lines, likely caused by a change of continuum brightness that is linked to variations in the mass accretion rate (Doan et al., 2019; Graham et al., 2019; Guo et al., 2019; Kynoch et al., 2019; Lawrence et al., 2016; MacLeod et al., 2016; Ruan et al., 2016; Rumbaugh et al., 2018; Sheng et al., 2019; Śniegowska & Czerny, 2019; Stern et al., 2018). Although some CLQs are relatively short (eg. PS1-13cbe faded back to an off state within 2 years), others have only been seen as they change from one level of activity to the next over almost a decade (eg. SDSSJ1011+5442 Runnoe et al. 2016, or majority of SDSS CLQs in MacLeod et al. 2019). The best way to distinguish between eg. a very slow TDE and a fast CLQ is detailed spectroscopic study (eg. see discussion in Gezari et al. 2017; Husemann et al. 2016; Kynoch et al. 2019; LaMassa et al. 2017). Therefore to determine the true nature of the CLQ candidates from Chapter 3 we would suggest the spectroscopic follow up of the brightest targets in that sample, and similarly to MacLeod et al. (2019) and Ruan et al. (2016) performing an analysis of spectral energy distribution, identifying the presence of emission lines, and continuum levels.

5.2.6 LSST science validation and verification

In Chapter 4 we analyzed the performance of the LSST science pipelines in the regions of crowded stellar fields, using DECAm Plane Survey (DECaPS) data. To parametrize the fidelity of LSST processing results to those from the custom-written DECaPS pipeline, we defined appropriate metrics, such as the mean completeness between 18th-20th mag $\langle C_{18-20} \rangle$, or additional systematic uncertainty at 21 mag (σ_S). Meanwhile, the LSST-Data Management (DM) Science Quality and Reliability Engineering team (SQuaRE) developed a Science Quality System Harness (SQuaSH) metrics dashboard[1]. SQuaSH allows temporal tracking

[1] `https://sqr-009.lsst.io`

of science pipeline metrics, showing how software alterations affect the data processing results, such as $\langle C_{18-20} \rangle$ or σ_S. Implementing the metrics from LSST-DECaPS comparison would expand the usefulness of SQuaSH, utilize the validation metrics framework[2], and build on the quality assurance requirements outlined in Tyson et al. (2017) to aid validation and verification of science requirements (Ivezić & the LSST Science Collaboration, 2018). Furthermore, Chapter 4 assumed DECaPS pipeline results (Schlafly et al., 2018) as 'truth', but with the availability of realistic sky simulation from GalSim[3](Rowe et al., 2015), the input sources are known, so that the performance of LSST science pipelines can be tested in a more absolute way. Future work would involve producing realistic simulations of the LSST sky in a variety of stellar crowdedness levels, and analyzing the processing results. Current LSST DM plans for performance monitoring include semi-automated reanalysis of certain test datasets[4], such as DECam HiTS survey (Förster et al., 2016), or Hyper Suprime-Cam Subaru Strategic Program dataset (Yasuda et al., 2019). Adding to these datasets the DECaPS data and simulated night sky from GalSim would enhance the possibility of validation and verification of the required performance level for the LSST science pipelines in crowded fields.

---

[2] https://sqr-017.lsst.io/

[3] https://www.lsst.org/scientists/simulations/galsim

[4] https://dmtn-091.lsst.io/

BIBLIOGRAPHY

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

AlSayyad, Y. 2016, PhD thesis, University of Washington

Annis, J., Soares-Santos, M., Strauss, M. A., et al. 2014, ApJ, 794, 120

Aranzana, E., Körding, E., Uttley, P., Scaringi, S., & Bloemen, S. 2018, MNRAS, 476, 2501

Awan, H., Gawiser, E., Kurczynski, P., et al. 2016, ApJ, 829, 50

Babić, A., Miller, L., Jarvis, M. J., et al. 2007, A&A, 474, 755

Balbus, S. A. 2003, ARA&A, 41, 555

Balbus, S. A., & Hawley, J. F. 1998, Reviews of Modern Physics, 70, 1

Baldassare, V. F., Geha, M., & Greene, J. 2018, ApJ, 868, 152

Bartos, I., Kocsis, B., Haiman, Z., & Márka, S. 2017, ApJ, 835, 165

Bauer, A., Baltay, C., Coppi, P., et al. 2009, ApJ, 696, 1241

Begelman, M. C., Blandford, R. D., & Rees, M. J. 1984, Rev. Mod. Phys., 56, 255

Begelman, M. C., & Silk, J. 2017, MNRAS, 464, 2311

Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002

Bentz, M. C., & Katz, S. 2015, PASP, 127, 67

Bentz, M. C., Peterson, B. M., Netzer, H., Pogge, R. W., & Vestergaard, M. 2009, ApJ, 697, 160

Bianchi, S., Piconcelli, E., Chiaberge, M., et al. 2009, ApJ, 695, 781

Bianchini, F., Fabbian, G., Lapi, A., et al. 2019, ApJ, 871, 136

Blagorodnova, N., Gezari, S., Hung, T., et al. 2017, ApJ, 844, 46

Blanchard, P. K., Nicholl, M., Berger, E., et al. 2017, The Astrophysical Journal, 843, 106

Blandford, R. D., & Znajek, R. L. 1977, MNRAS, 179, 433

Blanton, M. R., Lin, H., Lupton, R. H., et al. 2003, AJ, 125, 2276

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

Bosch, J., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S5

Butler, N. R., & Bloom, J. S. 2011, ApJ, 141, 93

Cackett, E. M., Chiang, C.-Y., McHardy, I., et al. 2018, ApJ, 857, 53

Cackett, E. M., Gültekin, K., Bentz, M. C., et al. 2015, ApJ, 810, 86

Cackett, E. M., Horne, K., & Winkler, H. 2007, MNRAS, 380, 669

Cai, Z.-Y., Wang, J.-X., Gu, W.-M., et al. 2016, ApJ, 826, 7

Cai, Z.-Y., Wang, J.-X., Zhu, F.-F., et al. 2018, The Astrophysical Journal, 855, 117

Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2015, ApJ, 811, 148

—. 2017, ApJ, 834, 111

—. 2018, ApJ, 867, 148

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560

Charisi, M., Bartos, I., Haiman, Z., et al. 2016, MNRAS, 463, 2145

Charisi, M., Bartos, I., Haiman, Z., Price-Whelan, A. M., & Márka, S. 2015, Monthly Notices of the Royal Astronomical Society: Letters, 454, L21

Chen, Z.-F., Pan, D.-S., Pang, T.-T., & Huang, Y. 2018, ApJS, 234, 16

Clarke, A. O., Scaife, A. M. M., Greenhalgh, R., & Griguta, V. 2019, arXiv e-prints, arXiv:1909.10963

Collier, S., & Peterson, B. M. 2001, ApJ, 555, 775

Curran, S. J., & Moss, J. P. 2019, A&A, 629, A56

Czerny, B. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 360, AGN Variability from X-Rays to Radio Waves, ed. C. M. Gaskell, I. M. McHardy, B. M. Peterson, & S. G. Sergeev, 265

Czerny, B., Wang, J.-M., Du, P., et al. 2019, ApJ, 870, 84

de Vries, W. H., Becker, R. H., White, R. L., & Loomis, C. 2005, AJ, 129, 615

Dexter, J., & Agol, E. 2011, ApJ, 727, L24

Dexter, J., & Begelman, M. C. 2019, MNRAS, 483, L17

Djorgovski, S. G., Drake, A. J., Mahabal, A. A., et al. 2011, ArXiv:1102.5004, arXiv:1102.5004

Doan, A. N., Eracleous, M., Anderson, S., et al. 2019, in American Astronomical Society Meeting Abstracts, Vol. 233, American Astronomical Society Meeting Abstracts #233, 242.23

Dong, X. Y., Wu, X.-B., Ai, Y. L., et al. 2018, AJ, 155, 189

Drake, A. J., Djorgovski, S. G., Mahabal, A., et al. 2009, ApJ, 696, 870

Drake, A. J., Catelan, M., Djorgovski, S. G., et al. 2013, ApJ, 763, 32

Eardley, D. M., & Lightman, A. P. 1975, ApJ, 200, 187

Edelson, R., Vaughan, S., Malkan, M., et al. 2014, ApJ, 795, 2

Edelson, R., Gelbord, J. M., Horne, K., et al. 2015, ApJ, 806, 129

Eilers, A.-C., Hennawi, J. F., & Davies, F. B. 2018, ApJ, 867, 30

Elitzur, M., Ho, L. C., & Trump, J. R. 2014, MNRAS, 438, 3340

Evans, I. N., & Civano, F. 2018, Astronomy and Geophysics, 59, 2.17

Evans, I. N., Primini, F. A., Glotfelty, K. J., et al. 2010, The Astrophysical Journal Supplement Series, 189, 37

Fausnaugh, M. M., Denney, K. D., Barth, A. J., et al. 2016, ApJ, 821, 56

Fausnaugh, M. M., Starkey, D. A., Horne, K., et al. 2018, ApJ, 854, 107

Ferrarese, L., & Merritt, D. 2000, The Astrophysical Journal, 539, L9

Flewelling, H. 2018, in American Astronomical Society Meeting Abstracts, Vol. 231, American Astronomical Society Meeting Abstracts 231, 436.01

Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2016, arXiv e-prints, arXiv:1612.05243

Foord, A., Gültekin, K., Reynolds, M., et al. 2017, ApJ, 851, 106

Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220

Förster, F., Maureira, J. C., San Martín, J., et al. 2016, ApJ, 832, 155

Frank, J., King, A., & Raine, D. J. 2002, Accretion Power in Astrophysics: Third Edition (Cambridge University Press), 398

Frederick, S., Gezari, S., Graham, M. J., et al. 2019, arXiv e-prints, arXiv:1904.10973

Gardner, E., & Done, C. 2017, Monthly Notices of the Royal Astronomical Society, 470, 3591

Gaskell, C. M., Bartel, K., Deffner, J. N., & Xia, I. 2019, arXiv e-prints, arXiv:1909.06275

Gezari, S., Hung, T., Cenko, S. B., et al. 2017, The Astrophysical Journal, 835, 144

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Graham, M. J., Djorgovski, S. G., Drake, A. J., et al. 2014, MNRAS, 439, 703

—. 2017, MNRAS, 470, 4112

Graham, M. J., Drake, A. J., Djorgovski, S. G., et al. 2013, MNRAS, 434, 3423

Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015a, Nature, 518, 74

—. 2015b, MNRAS, 453, 1562

Graham, M. J., Ross, N. P., Stern, D., et al. 2019, arXiv e-prints, arXiv:1905.02262

Gravity Collaboration, Sturm, E., Dexter, J., et al. 2018, Nature, 563, 657

Grzędzielski, M., Janiuk, A., Czerny, B., & Wu, Q. 2017, A&A, 603, A110

Guo, H., Wang, J., Cai, Z., & Sun, M. 2017, ApJ, 847, 132

Guo, H., Malkan, M. A., Gu, M., et al. 2016, The Astrophysical Journal, 826, 186

Guo, H., Shen, Y., He, Z., et al. 2019, arXiv e-prints, arXiv:1907.06669

Hameury, J.-M., Viallet, M., & Lasota, J.-P. 2009, A&A, 496, 413

Harrison, C. M. 2017, Nature Astronomy, 1, 0165

Hawkins, M. R. S. 2002, MNRAS, 329, 76

—. 2007, A&A, 462, 581

Heckman, T. M., & Best, P. N. 2014, ARA&A, 52, 589

Hennawi, J. F., & Prochaska, J. X. 2007, ApJ, 655, 735

Hernitschek, N., Schlafly, E. F., Sesar, B., et al. 2016, The Astrophysical Journal, 817, 73

Hogg, D. W. 2001, The Astronomical Journal, 121, 1207

Hogg, D. W., & Turner, E. L. 1998, PASP, 110, 727

Holoien, T. W. S., Vallely, P. J., Auchettl, K., et al. 2019, ApJ, 883, 111

Homayouni, Y., Trump, J. R., Grier, C. J., et al. 2019, ApJ, 880, 126

Hopkins, P. F., Hernquist, L., Martini, P., et al. 2005, ApJ, 625, L71

Hung, T., Gezari, S., Blagorodnova, N., et al. 2017, ApJ, 842, 29

Husemann, B., & Harrison, C. M. 2018, Nature Astronomy, 2, 196

Husemann, B., Urrutia, T., Tremblay, G. R., et al. 2016, A&A, 593, L9

Ivezić, Ž., Connolly, A. J., VanderPlas, J. T., & Gray, A. 2014, Statistics, Data Mining, and Machine Learning in Astronomy

Ivezić, Ž., & the LSST Science Collaboration. 2018, The LSST System Science Requirements Document

Ivezić, Ž., Lupton, R. H., Juric, M., et al. 2004, in IAU Symposium, Vol. 222, The Interplay Among Black Holes, Stars and ISM in Galactic Nuclei, ed. T. Storchi-Bergmann, L. C. Ho, & H. R. Schmitt, 525–526

Ivezić, Ž., Lupton, R. H., Juric, M., et al. 2004, in IAU Symposium, Vol. 222, The Interplay Among Black Holes, Stars and ISM in Galactic Nuclei, ed. T. Storchi-Bergmann, L. C. Ho, & H. R. Schmitt, 525–526

Ivezić, Ž., Smith, J. A., Miknaitis, G., et al. 2007, AJ, 134, 973

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Jiang, Y.-F., Davis, S. W., & Stone, J. M. 2016, ApJ, 827, 10

Jiang, Y.-F., Green, P. J., Greene, J. E., et al. 2017, ApJ, 836, 186

Jin, X., Zhang, Y., Zhang, J., et al. 2019, MNRAS, 485, 4539

Jones, M. L., Hickox, R. C., Mutch, S. J., et al. 2019, ApJ, 881, 110

Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864

Jurić, M., et al. 2017, LSST Data Products Definition Document

Kasliwal, V. P., Vogeley, M. S., & Richards, G. T. 2015, MNRAS, 451, 4328

Kawaguchi, T., Mineshige, S., Umemura, M., & Turner, E. L. 1998, ApJ, 504, 671

Kelly, B. C. 2007, ApJ, 665, 1489

Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009, The Astrophysical Journal, 698, 895

Kelly, B. C., Bechtold, J., Siemiginowska, A., Aldcroft, T., & Sobolewska, M. 2007, ApJ, 657, 116

Kelly, B. C., Becker, A. C., Sobolewska, M., Siemiginowska, A., & Uttley, P. 2014, ApJ, 788, 33

Kelly, B. C., Sobolewska, M., & Siemiginowska, A. 2011, ApJ, 730, 52

Kelly, B. C., Treu, T., Malkan, M., Pancoast, A., & Woo, J.-H. 2013, ApJ, 779, 187

Khrykin, I. S., Hennawi, J. F., & McQuinn, M. 2017, ApJ, 838, 96

Kimball, A. E., & Ivezić, Ž. 2008, The Astronomical Journal, 136, 684

King, A. 2014, Space Sci. Rev., 183, 427

King, A., & Pounds, K. 2015, ARA&A, 53, 115

Kokubo, M. 2015, MNRAS, 449, 94

Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv e-prints, arXiv:1711.03234

Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511

Kozłowski, S. 2015, Acta Astron., 65, 251

Kozłowski, S. 2016, MNRAS, 459, 2787

Kozłowski, S. 2016, ApJ, 826, 118

—. 2017a, ApJ, 835, 250

—. 2017b, ApJS, 228, 9

Kozłowski, S., Kochanek, C. S., Ashby, M. L. N., et al. 2016, ApJ, 817, 119

Kozłowski, S., Kochanek, C. S., Udalski, A., et al. 2010, ApJ, 708, 927

Kozłowski, Szymon. 2017, A&A, 597, A128

Krolik, J. H., Horne, K., Kallman, T. R., et al. 1991, ApJ, 371, 541

Kubota, A., & Done, C. 2018, MNRAS, 480, 1247

Kulkarni, G., Worseck, G., & Hennawi, J. F. 2019, MNRAS, 488, 1035

Kynoch, D., Ward, M. J., Lawrence, A., et al. 2019, MNRAS, 485, 2573

Labita, M., Decarli, R., Treves, A., & Falomo, R. 2009, MNRAS, 399, 2099

LaMassa, S. M., Georgakakis, A., Vivek, M., et al. 2019, ApJ, 876, 50

LaMassa, S. M., Yaqoob, T., & Kilgard, R. 2017, ApJ, 840, 11

LaMassa, S. M., Cales, S., Moran, E. C., et al. 2015, ApJ, 800, 144

LaMassa, S. M., Urry, C. M., Cappelluti, N., et al. 2016, ApJ, 817, 172

Läsker, R., Greene, J. E., Seth, A., et al. 2016, ApJ, 825, 3

Lasota, J.-P. 2016, in Astrophysics and Space Science Library, Vol. 440, Astrophysics of Black Holes: From Fundamental Aspects to Latest Developments, ed. C. Bambi, 1

Lawrence, A. 2016, in Astronomical Society of the Pacific Conference Series, Vol. 505, Astronomical Surveys and Big Data, ed. A. Mickaelian, A. Lawrence, & T. Magakian, 107

Lawrence, A. 2018, Nature Astronomy, 2, 102

Lawrence, A., Bruce, A. G., MacLeod, C., et al. 2016, MNRAS, 463, 296

Li, S.-L., & Cao, X. 2008, MNRAS, 387, L41

Li, Z., McGreer, I. D., Wu, X.-B., Fan, X., & Yang, Q. 2018, ApJ, 861, 6

Lira, P., Arévalo, P., Uttley, P., McHardy, I. M. M., & Videla, L. 2015, MNRAS, 454, 368

Liu, T., Gezari, S., Heinis, S., et al. 2015, ApJ, 803, L16

Liu, T., Gezari, S., Burgett, W., et al. 2016, ApJ, 833, 6

Lupton, R. 2005, in prep.

Lupton, R., Gunn, J. E., Ivezić, Z., Knapp, G. R., & Kent, S. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Harnden, Jr., F. A. Primini, & H. E. Payne, 269

Lupton, R. H., Ivezić, Z., & Gunn, J. 2005, in prep.

Lupton, R. H., Ivezić, Z., Gunn, J. E., et al. 2002, in SPIE Proceedings, Vol. 4836, Survey and Other Telescope Technologies and Discoveries, ed. J. A. Tyson & S. Wolff, 350–356

MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., et al. 2010, The Astrophysical Journal, 721, 1014

MacLeod, C. L., Brooks, K., Ivezić, Ž., et al. 2011, The Astrophysical Journal, 728, 26

MacLeod, C. L., Ivezić, Ž., Sesar, B., et al. 2012, The Astrophysical Journal, 753, 106

MacLeod, C. L., Ross, N. P., Lawrence, A., et al. 2016, MNRAS, 457, 389

MacLeod, C. L., Green, P. J., Anderson, S. F., et al. 2019, ApJ, 874, 8

Marchese, E., Braito, V., Della Ceca, R., Caccianiga, A., & Severgnini, P. 2012, Monthly Notices of the Royal Astronomical Society, 421, 1803

Martini, P., & Schneider, D. P. 2003, ApJ, 597, L109

Marziani, P., Dultzin, D., Sulentic, J. W., et al. 2018, Frontiers in Astronomy and Space Sciences, 5, 6

Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, PASP, 131, 018003

Matthews, T. A., & Sandage, A. R. 1963, ApJ, 138, 30

McGreer, I. D., Fan, X., Jiang, L., & Cai, Z. 2018, AJ, 155, 131

McGreer, I. D., Jiang, L., Fan, X., et al. 2013, ApJ, 768, 105

McHardy, I. M., Connolly, S. D., Horne, K., et al. 2018, MNRAS, 480, 2881

McLure, R. J., & Dunlop, J. S. 2004, MNRAS, 352, 1390

McLure, R. J., & Jarvis, M. J. 2002, MNRAS, 337, 109

Merloni, A., Dwelly, T., Salvato, M., et al. 2015, MNRAS, 452, 69

Morganson, E., Burgett, W. S., Chambers, K. C., et al. 2014, ApJ, 784, 92

Mosquera, A. M., Kochanek, C. S., Chen, B., et al. 2013, ApJ, 769, 53

Mudd, D., Martini, P., Zu, Y., et al. 2018, ApJ, 862, 123

Murphy, M. T., & Bernet, M. L. 2016, MNRAS, 455, 1043

Mushotzky, R. F., Edelson, R., Baumgartner, W., & Gandhi, P. 2011, The Astrophysical Journal, 743, L12

Netzer, H. 2013, The Physics and Evolution of Active Galactic Nuclei (Cambridge University Press)

Noda, H., & Done, C. 2018, MNRAS, 480, 3898

Oke, J. B., & Sandage, A. 1968, ApJ, 154, 21

Olsen, K. A. G., Blum, R. D., & Rigaut, F. 2003, AJ, 126, 452

Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017, A&A Rev., 25, 2

Palanque-Delabrouille, N., Magneville, C., Yèche, C., et al. 2013, A&A, 551, A29

—. 2016, A&A, 587, A41

Panda, S., Czerny, B., Done, C., & Kubota, A. 2019a, ApJ, 875, 133

Panda, S., Marziani, P., & Czerny, B. 2019b, arXiv e-prints, arXiv:1905.01729

Pâris, I., Petitjean, P., Ross, N. P., et al. 2017, A&A, 597, A79

Pâris, I., Petitjean, P., Aubourg, É., et al. 2018, A&A, 613, A51

Parks, D., Prochaska, J. X., Dong, S., & Cai, Z. 2018, MNRAS, 476, 1151

Peters, C. M., Richards, G. T., Myers, A. D., et al. 2015, ApJ, 811, 95

Prochaska, J. X., Lau, M. W., & Hennawi, J. F. 2014, ApJ, 796, 140

Rakshit, S., & Stalin, C. S. 2017, ApJ, 842, 96

Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press)

Rau, A., Kulkarni, S. R., Law, N. M., et al. 2009, PASP, 121, 1334

Reichert, G. A., Rodriguez-Pascual, P. M., Alloin, D., et al. 1994, ApJ, 425, 582

Ricarte, A., Tremmel, M., Natarajan, P., & Quinn, T. 2019, MNRAS, 489, 802

Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, AJ, 131, 2766

Richards, G. T., Myers, A. D., Peters, C. M., et al. 2015, ApJS, 219, 39

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in Proc. SPIE, Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, 914320

Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, ApJ, 876, 85

Risaliti, G., Miniutti, G., Elvis, M., et al. 2009, The Astrophysical Journal, 696, 160

Rosen, S. R., Webb, N. A., Watson, M. G., et al. 2016, A&A, 590, A1

Ross, N. P., McGreer, I. D., White, M., et al. 2013, ApJ, 773, doi:10.1088/0004-637X/773/1/14

Ross, N. P., Ford, K. E. S., Graham, M., et al. 2018, MNRAS, 480, 4468

Rowe, B. T. P., Jarvis, M., Mandelbaum, R., et al. 2015, Astronomy and Computing, 10, 121

Ruan, J. J., Anderson, S. F., Dexter, J., & Agol, E. 2014, The Astrophysical Journal, 783, 105

Ruan, J. J., Anderson, S. F., Eracleous, M., et al. 2019, ApJ, 883, 76

Ruan, J. J., Anderson, S. F., Cales, S. L., et al. 2016, ApJ, 826, 188

Rumbaugh, N., Shen, Y., Morganson, E., et al. 2018, ApJ, 854, 160

Runnoe, J. C., Cales, S., Ruan, J. J., et al. 2016, MNRAS, 455, 1691

Rybicki, G. B., & Press, W. H. 1992, ApJ, 398, 169

Sánchez-Sáez, P., Lira, P., Mejía-Restrepo, J., et al. 2018, ApJ, 864, 87

Sánchez-Sáez, P., Lira, P., Cartier, R., et al. 2019, ApJS, 242, 10

Sartori, L. F., Schawinski, K., Trakhtenbrot, B., et al. 2018a, MNRAS, 476, L34

Sartori, L. F., Trakhtenbrot, B., Schawinski, K., et al. 2019, ApJ, 883, 139

Sartori, L. F., Schawinski, K., Koss, M. J., et al. 2018b, MNRAS, 474, 2444

Scaringi, S., Maccarone, T. J., Kording, E., et al. 2015, Science Advances, 1, e1500686

Schawinski, K. 2012, ArXiv e-prints, arXiv:1206.2661

Schawinski, K., Koss, M., Berney, S., & Sartori, L. F. 2015, MNRAS, 451, 2517

Schlafly, E. F., Green, G. M., Lang, D., et al. 2018, ApJS, 234, 39

Schmidt, K. B., Marshall, P. J., Rix, H.-W., et al. 2010, ApJ, 714, 1194

Schmidt, M. 1963, Nature, 197, 1040

Schmidt, T. M., Hennawi, J. F., Worseck, G., et al. 2018, ApJ, 861, 122

Schneider, D. P., Hall, P. B., Richards, G. T., et al. 2007, AJ, 134, 102

—. 2008, VizieR Online Data Catalog, 7252

Schneider, D. P., Richards, G. T., Hall, P. B., et al. 2010, VizieR Online Data Catalog, 7260

Sesar, B., Ivezić, Ž., Lupton, R. H., et al. 2007, The Astronomical Journal, 134, 2236

Sesar, B., Hernitschek, N., Mitrović, S., et al. 2017, AJ, 153, 204

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Shaw, R. A. 2015, NOAO Data Handbook

Shen, Y. 2013, Bulletin of the Astronomical Society of India, 41, 61

Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., & Schneider, D. P. 2008, ApJ, 680, 169

Shen, Y., & Ho, L. C. 2014, Nature, 513, 210

Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45

Shen, Y., Hall, P. B., Horne, K., et al. 2019, ApJS, 241, 34

Sheng, Z., Wang, T., Jiang, N., et al. 2017, The Astrophysical Journal, 846, L7

Sheng, Z., Wang, T., Jiang, N., et al. 2019, arXiv e-prints, arXiv:1905.02904

Simm, T., Salvato, M., Saglia, R., et al. 2016, A&A, 585, A129

Smith, K. L., Mushotzky, R. F., Boyd, P. T., et al. 2018, ApJ, 857, 141

Śniegowska, M., & Czerny, B. 2019, arXiv e-prints, arXiv:1904.06767

Stern, D., Graham, M. J., Arav, N., et al. 2017, ApJ, 839, 106

Stern, D., McKernan, B., Graham, M. J., et al. 2018, ApJ, 864, 27

Stetson, P. B. 1987, PASP, 99, 191

Stringer, K. M., Long, J. P., Macri, L. M., et al. 2019, AJ, 158, 16

Suberlak, K., Ivezić, Ž., MacLeod, C. L., Graham, M., & Sesar, B. 2017, MNRAS, 472, 4870

Sun, J., & Shen, Y. 2015, ApJ, 804, L15

Sun, M., Xue, Y., Wang, J., Cai, Z., & Guo, H. 2018, ApJ, 866, 74

Tie, S. S., Martini, P., Mudd, D., et al. 2017, AJ, 153, 107

Trakhtenbrot, B., Arcavi, I., Ricci, C., et al. 2019, Nature Astronomy, 3, 242

Tyson, T., et al. 2017, LSST Data Quality Assurance Plan

van Velzen, S., Gezari, S., Cenko, S. B., et al. 2019, ApJ, 872, 198

Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, AJ, 122, 549

Vanden Berk, D. E., Wilhite, B. C., Kron, R. G., et al. 2004a, ApJ, 601, 692

—. 2004b, ApJ, 601, 692

Vanderplas, J., Connolly, A., Ivezić, Ž., & Gray, A. 2012, in Conference on Intelligent Data Understanding (CIDU), 47 –54

VanderPlas, J. T., & Ivezić, Ž. 2015, ApJ, 812, 18

Vaughan, S., Uttley, P., Markowitz, A. G., et al. 2016, MNRAS, 461, 3145

Vestergaard, M. 2002, ApJ, 571, 733

Vestergaard, M., & Peterson, B. M. 2006, ApJ, 641, 689

Wang, T., Brinkmann, W., & Bergeron, J. 1996, A&A, 309, 81

Weigel, A. K., Schawinski, K., Caplar, N., et al. 2017, ApJ, 845, 134

Wilhite, B. C., Brunner, R. J., Grier, C. J., Schneider, D. P., & vanden Berk, D. E. 2008, MNRAS, 383, 1232

Wisotzki, L. 2000, A&A, 353, 861

Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, ARA&A, 43, 861

Yang, G., Brandt, W. N., Alexander, D. M., et al. 2019, MNRAS, 485, 3721

Yang, J., Fan, X., Wu, X.-B., et al. 2017, AJ, 153, 184

Yang, Q., Wu, X.-B., Fan, X., et al. 2018, The Astrophysical Journal, 862, 109

Yasuda, N., Tanaka, M., Tominaga, N., et al. 2019, PASJ, 71, 74

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Zheng, Z.-Y., Butler, N. R., Shen, Y., et al. 2016, ApJ, 827, 56

Zhu, F.-F., Wang, J.-X., Cai, Z.-Y., et al. 2018, The Astrophysical Journal, 860, 29

Zu, Y., Kochanek, C. S., Kozłowski, S., & Udalski, A. 2013, ApJ, 765, 106

Zu, Y., Kochanek, C. S., & Peterson, B. M. 2011, ApJ, 735, 80

Zuo, W., Wu, X.-B., Liu, Y.-Q., & Jiao, C.-L. 2012, The Astrophysical Journal, 758, 104