The Galactic Center In Motion

Philipp M. Plewa

SIGILL: UNIVERSITAT: LUDOVIC: MAXIM:

München 2018

The Galactic Center In Motion

Philipp M. Plewa

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Philipp M. Plewa
aus Warendorf

München, den 27.4.2018

Erstgutachter: Professor Reinhard Genzel
Zweitgutachter: Professor Andreas Burkert
Tag der mündlichen Prüfung: 2.7.2018

Contents

List of Figures

List of Tables

List of Acronyms

ADU	Analog-Digital Unit
AO	Adaptive Optics
ELT	Extremely Large Telescope
ESO	European Southern Observatory
FWHM	Full Width at Half Maximum
HST	Hubble Space Telescope
ISAAC	Infrared Spectrometer and Array Camera
LMC	Large Magellanic Cloud
MCMC	Markov Chain Monte Carlo
NACO	Nasmyth Adaptive Optics System Near-Infrared Imager and Spectrograph (NAOS-CONICA)
NIRC	Near-Infrared Camera
PSF	Point Spread Function
SINFONI	Spectrograph for Integral Field Observations in the Near-Infrared
VLT	Very Large Telescope
VLTI	Very Large Telescope Interferometer

Zusammenfassung

Hochauflösende astronomische Beobachtungen von Sternbahnen im infraroten Spektralbereich sind überzeugender Beweis für die Existenz eines (super)massereichen Schwarzen Loches im Galaktischen Zentrum, assoziiert mit der kompakten Radioquelle Sgr A*. Das astrometrische Bezugssystem für diese Beobachtungen ist an eine Gruppe roten Riesensterne gebunden, die aufgrund von SiO Maseremission in ihrer Hülle auch im Radio-Spektralbereich detektierbar sind. Durch Angleichen der im Infrarot- und Radiobereich gemessen Sternpositionen lässt sich (die Radioquelle) Sgr A* näherungsweise im Ursprung eines Infrarot-Bezugssystem platzieren, mit einer zu erwartenden Genauigkeit von nun ~ 0.17 mas in Position (in 2009) und ~ 0.07 mas pro Jahr in Eigenbewegung, wie ich in Kapitel 2 der vorliegenden Arbeit zeige. Die signifikante Verbesserung der langfristigen Stabilität dieses Koordinatensystems, im Vergleich zu vorhergehenden Studien, ist das Ergebnis von Modellierung und Korrektur optischer Bildverzeichnungen im NACO-Kamerasystem, Berücksichtigung neuer Radio-Messungen, und einer erneuten Analyse früherer Infrarot-Beobachtungen mittels robuster numerischer Methoden. Zukünftige Beobachtungen werden eine weitere Verbesserung zur Folge haben, und voraussichtlich die Detektion post-Newtonscher Effekte auf Sternbahnen kurzer Periode ermöglichen.

Da die Sterndichte im Galaktischen Zentrum sehr hoch ist, stören häufig auftretende Überlappungen von sich in ihrer Abbildung begegnenden Sternen die Messung einzelner Sternpositionen. Dieser Effekt, zu dem auch undetektierte schwache Sterne beitragen, ist eine wichtige nicht-instrumentellen Fehlerquelle, welche die astrometrische Genauigkeit und Präzision heutiger Bildbeobachtungen beschränkt, sowie die langfristige Verfolgung von Sternbahnen in der näheren Umgebung des zentralen Schwarzen Loches. Ich habe die Bewegungen der bekannten und der noch unidentifizierten Sterne selbstkonsistent simuliert, um die statistischen Eigenschaften der hervorgerufenen Fehler zu charakterisieren. In Kapitel 4 zeige ich dass Sternbegegnungen weiterführende systematische Fehler verursachen können, sowohl in der Bestimmung von stellaren Bahnparametern, als auch in der Abschätzung von Masse und Entfernung des Schwarzen Loches, insbesondere wenn Sterne in kleinem projiziertem Abstand vom Schwarzen Loch beobachtet werden, wie zum Beispiel der Stern S2 im Perizentrum seiner Umlaufbahn. Das Weiteren untersuche ich die Modellierung von Sternbegegnungen als eine über die Zeit korrelierte Fehlerquelle, und demonstriere somit die Notwendigkeit verbesserter Fehlermodelle, um auch in zukünftigen astrometrischen Studien vertrauenswürdige Abschätzungen der interessanten Parameter (und deren Unsicherheiten) zu erhalten.

Die bisher im inneren Bereich des zentralen Sternhaufens detektierten Sterne sind die Leuchtstärksten der existierenden Sternpopulation, und sind hauptsächlich entweder alte Riesensterne späten Spektraltyps, oder massereiche junge Sterne frühen Typs. In Kapitel 5 greife ich das Thema spektraler Klassifikation anhand von engbandiger Photometrie auf, mit dem Hauptziel in der weiteren Umgebung des Schwarzen Loches schwache Sterne frühen Typs zu identifizieren, wo sie selten sind. Ein Random-Forest-Klassifikator, kalibriert mittels einer Stichprobe spektroskopisch klassifizierter Sterne, stellt sich als ähnlich leistungsfähig wie andere Methoden heraus, ohne dass ein Modell spektraler Energieverteilungen der Sterne benötigt wird. Die Vorteile der Verwendung eines solchen Klassifikators liegen in einem geringen Kalibrationsaufwand, einer Vorhersagegenauigkeit die sich mit zusätzlichen Trainingsdaten voraussichtlich noch verbessern wird, und in der einfachen Anwendung auf zukünftige, größere Datensätze. Durch Anwendung dieses Klassifikators auf Archivdaten konnte ich außerdem einige Ergebnisse frühere Studien reproduzieren, hinsichtlich der räumlichen Verteilung und der Helligkeitsverteilung sowohl von Sternen frühen, als auch späten Spektraltyps.

Anfang 2014 erreichte die sich schnell bewegende Infrarotquelle G2 den Punkt größtmöglicher Annäherung an das Schwarze Loch im Galaktischen Zentrum. In Kapitel 3 lege ich die unmittelbar darauffolgende Entwicklung der ionisierten Gaskomponente und der Staubkomponente von G2 dar, die in 2015 und 2016 mit dem Spektrographen SINFONI und dem Kamerasystem NACO beobachtet werden konnte. Die räumlich aufgelöste Dynamik der Brγ-Linenemission kann durch eine ballistische Bewegung und Gezeitenscherung einer Testteilchenwolke erklärt werden, die über die letzten 12 Jahre einer exzentrischen Keplerbahn um das Schwarze Loch gefolgt ist. Die Nichtdetektion einer Abbremsung, oder jeglicher starker hydrodynamischer Wechselwirkungen mit dem heißen Gas im inneren Akktretionsfluss, erlaubt die Abschätzung einer maximalen Umgebungsdichte von einigen wenigen 10^3 cm^{-3} beim Abstand größtmöglicher Annäherung (1500 R_s), unter der Annahme dass G2 eine sphärische Wolke ist, die sich durch eine stationäre und homogene Atmosphäre bewegt. Die Staubkontinuumsemission ist im L'-Band zwar nicht räumlich aufgelöst, die Position der L'-Band-Quelle ist aber konsistent mit der der Brγ-Quelle, und die Gesamtleuchtkraft der Brγ- und L'-Band-Emission hat sich innerhalb der Messgenauigkeit nicht verändert. Die genaue Beschaffenheit und der Ursprung von G2 sind weiterhin unklar, haben aber wahrscheinlich eine Verbindung zur vorherlaufenden Quelle G1, deren Entwicklung ähnlich zu G2 verlief, wenn auch nicht identisch. Beide Objekte haben außerdem mit hoher Wahrscheinlichkeit eine Verbindung zu einer Ausläuferstruktur, die über ein großes Positions- und Geschwindigkeitsintervall eine ununterbrochene Verbindung zu G2 aufweist.

Summary

Near-infrared observations of stellar orbits in the Galactic Center at high angular resolution have provided conclusive evidence for a (super)massive black hole associated with the compact radio source Sgr A*. The astrometric reference frame for these observations is tied to a set of red giant stars, which are also detectable at radio wavelengths due to SiO maser emission in their envelopes. By aligning the positions of SiO maser stars measured at both infrared and radio wavelengths, (radio-)Sgr A* can now be placed in the origin of an infrared astrometric reference frame with an expected precision of ~ 0.17 mas in position (in 2009) and ~ 0.07 mas per year in proper motion, as I show in chapter 2 of this thesis. The significant improvement of the long-term stability of this reference frame, in comparison to previous studies, is the result of modeling and correcting optical distortion in the NACO imager, including new radio measurements, and re-analyzing past infrared observations using robust numerical techniques. A further improvement will follow future observations, and is expected to facilitate the detection of post-Newtonian effects on short-period stellar orbits.

Since the Galactic Center is a highly crowded stellar field, frequent unrecognized events of source confusion, which involve undetected faint stars, are expected to introduce substantial noise in the measurements of individual stellar positions. This confusion noise is the main non-instrumental effect limiting the astrometric accuracy and precision of current near-infrared imaging observations, and the long-term monitoring of stellar orbits in the immediate vicinity of the central massive black hole. I have self-consistently simulated the motions of the known and the yet unidentified stars, to characterize this noise component. In chapter 4, I show that a likely consequence of source confusion is a bias in estimates of the stellar orbital elements, as well as the inferred mass and distance of the black hole, in particular if stars are being observed at small projected separations from it, such as the star S2 at the pericenter of its orbit. Furthermore, I have investigated modeling the effect of source confusion as an additional noise component that is time-correlated, demonstrating a need for improved noise models to obtain trustworthy estimates of the parameters of interest (and their uncertainties) in future astrometric studies.

The stars so far detected in the inner nuclear star cluster are the most luminous members of the existing stellar population, and are generally either evolved late-type giants or massive young, early-type stars. In chapter 5, I revisit the problem of stellar classification based on intermediate-band photometry, with the primary aim of identifying faint early-type candidate stars in the extended vicinity of the central black hole, where they are rare. A random forest classifier, trained on a sample of spectroscopically classified stars, performs similarly well as competitive methods, without involving any model of stellar spectral energy distributions. Advantages of using such a machine-trained classifier are a minimum of required calibration effort, a predictive accuracy expected to improve as more training data becomes available, and the ease of application to future, larger data sets. By applying this classifier to archive data, I was also able to reproduce several results of previous studies regarding the spatial distribution and the luminosity function of both the early- and late-type stars.

In early 2014, the fast-moving near-infrared source G2 reached its closest approach to the Galactic Center black hole. In chapter 3, I report on the evolution of the ionized gaseous component and the dusty component of G2 immediately after this event, revealed by observations obtained in 2015 and 2016 with the SINFONI integral field spectrograph and the NACO imager. The spatially resolved dynamics of the Brγ line emission can be accounted for by the ballistic motion and tidal shearing of a test-particle cloud that has followed a highly eccentric Keplerian orbit around the black hole for the last 12 years. The non-detection of a drag force, or any strong hydrodynamic interaction with the hot gas in the inner accretion zone, limits the ambient density to less than a few 10^3 cm^{-3} at the distance of closest approach (1500 R_s), assuming G2 is a spherical cloud moving through a stationary and homogeneous atmosphere. The dust continuum emission is unresolved in L'-band, but its location has stayed consistent with that of the Brγ emission, while the total luminosity of the Brγ- and L'-band emission has remained constant to within the measurement uncertainty. The nature and origin of G2 remain uncertain, but are likely related to that of the precursor source G1, the orbital evolution of which is similar to G2's, though not identical. Both object are also likely related to a trailing tail structure, which is continuously connected to G2 over a large range in position and radial velocity.

Chapter 1

Introduction: The inner nuclear star cluster and its black hole

As the most nearby galactic nucleus and host of the fastest known stars in the Galaxy, the central region of the Milky Way is a unique and prototypical target for studies of stellar dynamics, star formation, and the interstellar medium in the extreme environment of a (super)massive black hole. For many years already, these studies have benefited from the favorable conditions and growing capabilities of ground-based observations at near-infrared wavelengths (for a comprehensive review see, e.g., Genzel et al., 2010). While being highly extincted at optical wavelengths by interstellar dust along the line of sight, the view toward the inner nuclear star cluster is largely unobstructed in the infrared spectral range (accessible at $\lambda \approx 1 - 4$ μm; see Fig. 1.1), although the extinction is patchy and several regions of sky are almost or completely opaque even in the infrared regime (e.g. Schödel et al., 2010; Fritz et al., 2011). Additional observations in the radio and sub-mm regimes from the ground, as well as in X-rays and gamma rays from space, continue to shape our present-day, multi-wavelength view of the Galactic Center.

The stars abundant in the dense inner nuclear cluster were first revealed individually as a result of several major technological advancements in high-resolution near-infrared observing techniques, such as speckle imaging and the development of AO systems, which mitigate the effect of atmospheric turbulence (e.g. Davies & Kasper, 2012), as well as the development of large and efficient infrared detector arrays (e.g. Rieke, 2007). Routine imaging observations carried out today, at an angular resolution approaching the diffraction limit of the largest (i.e. 8-10 m class) optical telescopes, and a depth limited by stellar crowding, in fact reveal all luminous members of the existing stellar population, within the central parsec[1]. Furthermore, by efficiently combining the means of both imaging techniques and imaging spectroscopy (e.g. Eisenhauer & Raab, 2015), different sub-populations of stars can be distinguished by their age and kinematics (see Fig. 1.2).

The majority of stars visible in the Galactic Center appear to be evolved late-type

---

[1] At a distance to the Galactic Center of approximately 8.3 kpc, a physical distance of 1 pc perpendicular to the line of sight corresponds to an apparent angular separation on the sky of about $25''$.

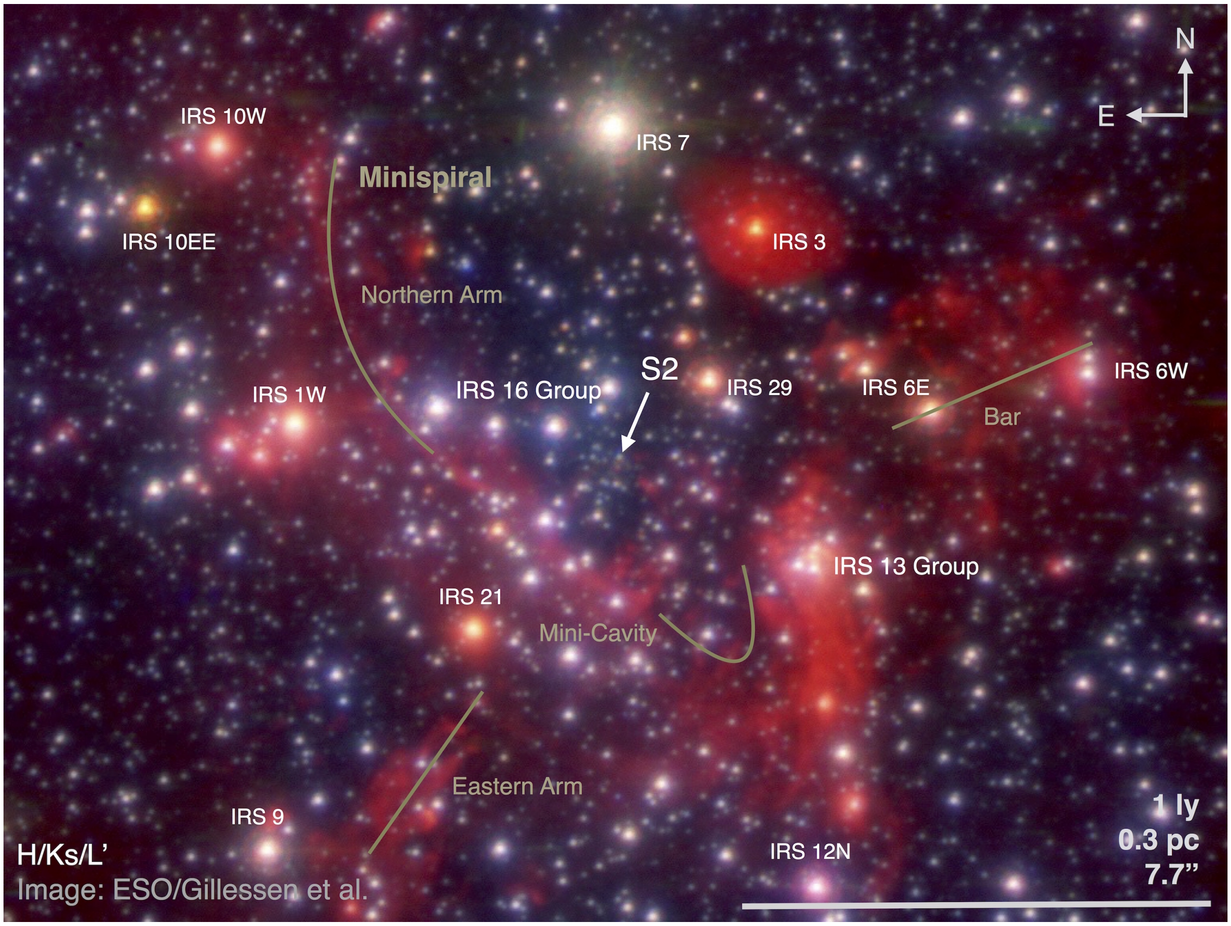

Figure 1.1: Multi-band infrared, adaptive-optics image of the Galactic Center environment, on which some prominent features are labeled. The H-band ($\lambda = 1.66 \pm 0.17$ μm) and Ks-band ($\lambda = 2.18 \pm 0.18$ μm) images reveal the luminous stellar population(s) of the dense inner nuclear cluster, whereas the L’-band ($\lambda = 3.80 \pm 0.31$ μm) image also shows emission from dusty gas filaments spread throughout the region.

giants (of spectral types K and M) with ages > 1 Gry, which likely formed in parallel with the bulk of the Galactic bulge population, and a regular (bulge-like) initial mass function, when the central black hole was still significantly less massive (e.g. Pfuhl et al., 2011). Kinematic modeling suggests that the old cluster observed today is a flattened isotropic rotator on scales of several $10''$, aligned with the Galactic plane, which shows no sign of any large-scale disturbances (Chatzopoulos et al., 2015; Fritz et al., 2016). The central density profile of this cluster could be expected to be a steep power-law cusp (with a stellar surface number density $\Sigma(R) \propto R^{\Gamma}$, where $-0.75 \leqslant \Gamma \leqslant -0.5$; Bahcall & Wolf, 1976), since the cluster could plausibly be old enough to be dynamically relaxed (e.g. Alexander, 2005), yet the observed density profile of the bright giants flattens to a shallow core within a few $1''$ from the cluster center (e.g. Genzel et al., 2003; Buchholz et al.,

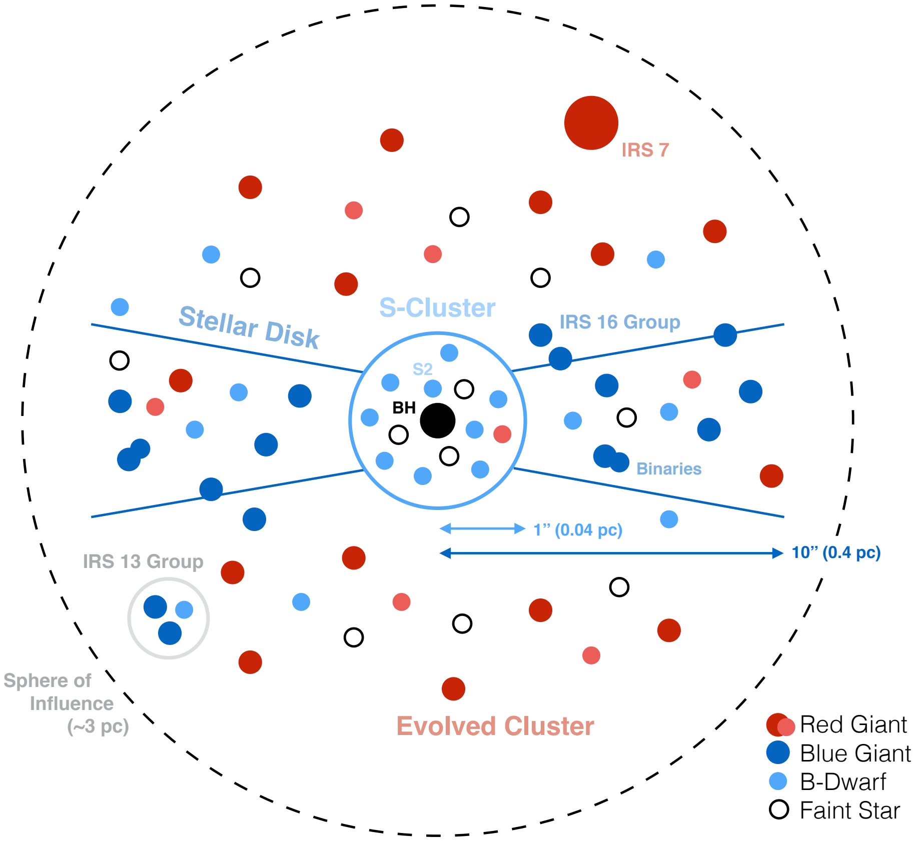

Figure 1.2: Schematic diagram of the stellar populations observed in the Galactic Center environment (adapted, with changes, from Alexander, 2017).

2009; Do et al., 2009; Bartko et al., 2010; Do et al., 2013a). The central density profile of the faintest detectable stars in the K-band (approx. 1-2 magnitudes fainter than the red clump), however, appears to be cusp-like (Gallego-Cano et al., 2018), and is consistent with the distribution of the diffuse background light produced by even fainter, unresolved stars (Schödel et al., 2018). The apparent flattening of this cusp ($\Gamma \gtrsim -0.5$) could possibly be explained by mass segregation, repeated star formation events, and the sporadic in-fall of star clusters (e.g. Baumgardt et al., 2018). If it is indeed predominantly old, as advocated by Gallego-Cano et al. (2018) and Schödel et al. (2018), the existence of a faint cusp implies that the bright giants have either not yet dynamically relaxed (see also Merritt, 2010), or that some process has rendered them undetectable. For instance, their envelopes could have been stripped in collisions with other stars or remnants (e.g. Dale et al., 2009, although high-speed collisions might be ineffective), or, perhaps more likely, in passing through a star-forming gas disk (Amaro-Seoane & Chen, 2014), leaving only very faint stellar cores.

It is such a gas disk from which the young stars also visible in the Galactic Center have likely formed, with a top-heavy initial mass function, during a recent episode of star formation ∼3-6 Myr ago (Paumard et al., 2006; Bartko et al., 2010; Do et al., 2013a; Lu et al., 2013). Despite the presence of strong tidal forces in the vicinity of a massive black hole (e.g. Morris, 1993), it has been shown that stars can form in-situ from an in-falling massive molecular cloud, after it has become bound to the black hole, and been compressed into a disk that quickly fragments (e.g. Bonnell & Rice, 2008). Today, a significant fraction of the most luminous O- and WR-type stars (20-50%; see Naoz et al., 2018) are observed to move coherently inside a remnant stellar disk structure (Paumard et al., 2006; Bartko et al., 2009; Lu et al., 2009; Yelda et al., 2014). This prominent clockwise disk appears to have a sharp inner edge at roughly $1''$ in projection from the center, is truncated at about $10 - 12''$ (Støstad et al., 2015), and has a steep surface density profile in between ($\Gamma \approx -1$; e.g. Do et al., 2013a). There might also exist a warp in this disk, as well as a second, less populated, counterclockwise stellar disk (Bartko et al., 2009). The binary fraction of the disk stars seems to be similar to the binary fraction in comparably young clusters (Pfuhl et al., 2014), but is not well enough constrained to provide conclusive information about the mode of star formation (e.g. Alexander et al., 2008), for example. If surviving long-period binaries can be detected in the future, their number could additionally constrain the properties of a hypothetical dark cusp of compact stellar remnants (e.g. Alexander & Pfuhl, 2014). Less luminous B-type stars also exist in the disk region, of which the brighter ones seem to belong to the disk population, whereas the fainter ones seem to resemble the B-stars existing further inward in their kinematics, where no O- nor WR-stars are found (e.g. Bartko et al., 2010; Madigan et al., 2014).

The central ‘S-star’ cluster with a radius of roughly $1''$ consists of more than a hundred known fast-moving stars, most of which appear to be genuine massive B-type dwarfs with ages of less than 15 Myr (e.g. Ghez et al., 2003; Eisenhauer et al., 2005; Martins et al., 2008; Habibi et al., 2017, ruling out the collisional stripping of giants as a formation scenario). As a result of the long-term monitoring of stellar motions in the Galactic Center, orbital elements have now been determined for 40 of the S-stars in total, of which 25 are early-type (candidate) stars and not members of the stellar disk (Gillessen et al., 2017). Their orbital planes are isotropically oriented, and the distribution of their orbital eccentricities is approximately thermal. Assuming that the young S-stars have also formed in the most recent star-formation event, Kozai-Lidov-like resonances could explain both the necessary redistribution of the orbital elements, as well as the depletion of O- and WR-stars in the central region (by tidal disruption; see Chen & Amaro-Seoane, 2014), thus resolving the ‘paradox of youth’ (Ghez et al., 2003). However, other plausible scenarios of origin for the young S-stars have also been proposed, for example involving stellar migration and the disruption of binaries by the black hole (in a multi-body gravitational interaction; see e.g. Gould & Quillen, 2003; Perets et al., 2007, but also Madigan et al., 2009), efficient relaxation processes (e.g. induced by stellar- or intermediate-mass black holes; see Perets et al., 2009; Merritt et al., 2009), as well as an explosive mode of concentrated star formation (e.g. Nayakshin & Zubovas, 2018). Several of the hyper-velocity stars discovered to be escaping from the Galactic halo (e.g. Brown et al., 2014, 2015) could have possibly been ejected from the Galactic Center via the Hills (1988) disruption mechanism, the anisotropic distribution of their distribution, however, suggests an alternative origin for most of them (e.g. being runaways from the LMC; see Boubert et al., 2017).

The first convincing evidence for the existence of a massive black hole in the Galactic Center was provided by the discovery of the S-stars' large proper motions (Eckart & Genzel, 1996; Genzel et al., 1997; Ghez et al., 1998), and was further strengthened by the discovery of the star S2 moving on a bound orbit (Fig. 1.3; Schödel et al., 2002; Ghez et al., 2003). The fact that the observed orbit of S2 is very close to Keplerian already implies that the gravitational potential is dominated by a large point-like mass, where the amount of enclosed extended mass is limited to a percent-level fraction of the central mass (Gillessen et al., 2009b, 2017). Similarly, the observed motion of S2 constrains the maximum mass of other dark objects allowed to exist nearby. The location of the central dynamical mass also coincides with the location of the compact radio source Sgr A*, which had been suspected to be a massive black hole before (see Genzel et al., 2010), as well as with the locations of recorded flares of the near-infrared counterpart source (see e.g. Gillessen et al., 2017). No convincing dynamical evidence has yet accumulated for the existence of an intermediate-mass black hole in the stellar association IRS 13 (e.g. Fritz et al., 2010a), or anywhere else in the surrounding region, although claims to a detection have been restated repeatedly (e.g. recently by Tsuboi et al., 2017).

The star S2 is a young main-sequence star of spectral type B0-B3 (Habibi et al., 2017), and the brightest star for which an orbit has been determined, within the central 1″. Its orbital period of approximately 16 years is also the second-shortest one measured to date (see Meyer et al., 2012), and it has been fully covered by observations once (see e.g. Gillessen et al., 2009a). For these reasons, the well-measured orbital motion of S2 provides some of the best available constraints on the mass and distance of the central black hole today, and the ongoing monitoring of the motion of S2 and that of the surrounding stars, i.e. regular observations of their on-sky motions, as well as spectroscopic measurements of their radial velocities, continues to yield steadily improved precision estimates of these fundamental parameters (e.g. Ghez et al., 2008; Gillessen et al., 2009b; Boehle et al., 2016; Gillessen et al., 2017). The latest estimates, based on a multi-star analysis by Gillessen et al. (2017), are:

$$R_0 = (8.32 \pm 0.07|_{\mathrm{stat}} \pm 0.14|_{\mathrm{sys}})\ \mathrm{kpc} \tag{1.1}$$

$$M_0 = (4.28 \pm 0.10|_{\mathrm{stat}} \pm 0.21|_{\mathrm{sys}}) \times 10^6 M_\odot \tag{1.2}$$

The achievable precision of these estimates is now limited by systematic uncertainties in the distance determination, which include uncertainties related to the astrometric reference frame, as well as source confusion, among other astrophysical and instrumental effects (see e.g. Fritz et al., 2010b). The mass-distance degeneracy persisting in the analysis of stellar orbits (approx. $M_0 \propto R_0^{2.82}$; Gillessen et al., 2017) can be partially broken by taking into account independent mass or distance constraints, for example based on kinematic modeling of the stellar disk ('orbital roulette'; e.g. Levin & Beloborodov, 2004) or the larger nuclear star cluster (e.g. Do et al., 2013b; Chatzopoulos et al., 2015, see also Fig. 1.3), as well as various other methods (see Bland-Hawthorn & Gerhard, 2016).

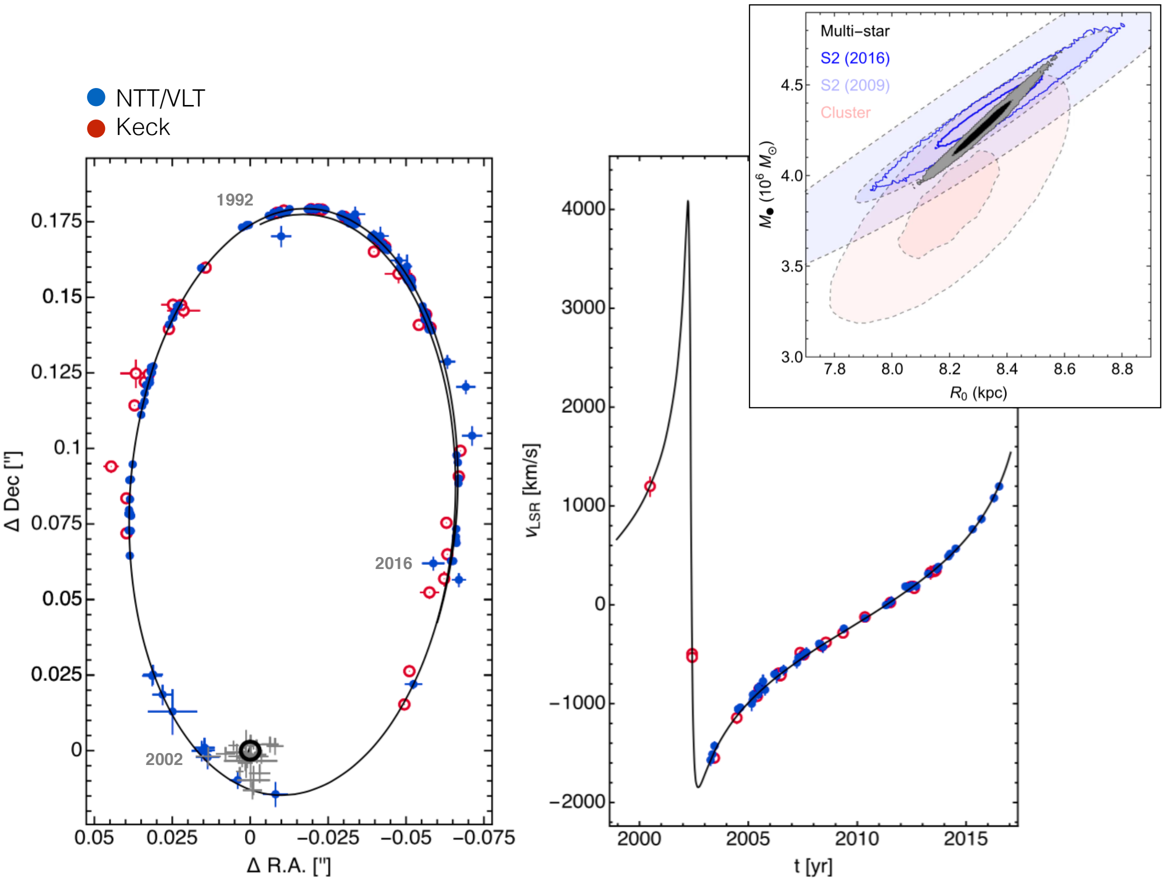

Figure 1.3: The orbital motion of the star S2, seen as a change of its on-sky position (left panel) and radial velocity (right panel). The black line indicates the best-fit Keplerian orbit, which does not close to a perfect ellipse in this figure because of a residual drift of the astrometric coordinate system. The gray points indicate the positions of recorded near-infrared flares of Sgr A*. The inset on the top-right shows the resulting constraints on the black hole mass and distance (in blue), in comparison to several other measurements (reproduced from Gillessen et al., 2017).

In chapter 2 of this thesis, I describe the construction of an astrometric reference frame for near-infrared observations of stellar motions in the Galactic Center, in which the radio source Sgr A* is fixed at the origin. In comparison to previous studies, the significant improvement of the precision and long-term stability of this reference frame, which are of fundamental concern in the analysis of stellar orbits, is largely the result of modeling and correcting optical distortion in the NACO imager, as well as re-analyzing past imaging observations using robust numerical techniques. In appendix 2.A, I present additional distortion corrections for a second camera of NACO, which have since been made available to the community, and I provide some additional information on the instrument characteristics in appendix 2.B. In chapter 3, I discuss the interaction of the peculiar Galactic Center source G2 with the black hole, having modeled it in the four-dimensional space of the available data (i.e. taking into account spectro-spatial information at multiple epochs). Due to our ability to observe its spatially resolved internal gas dynamics, G2 represents a unique probe of the inner accretion zone around the black hole, even though its nature and origin, and their possible implications, remain the subject of ongoing discussion and speculation (see also Chap. 6). In chapter 4, I discuss additional limitations for inferring stellar motions that arise from source confusion in a field as crowded as the Galactic Center, and I derive and demonstrate the application of an improved astrometric noise model, based on Gaussian processes, to observations of S2, with the aim of enabling future astrometric studies to obtain more trustworthy estimates of the parameters of interest and their uncertainties. Finally, in chapter 5, I test a machine-learning approach for efficient classification of early- and late-type candidate stars in the Galactic Center, using limited photometric information only, in preparation for proposed spectroscopic observations. I conclude by discussing further opportunities for studying the nuclear star cluster and the black hole at its center in chapter 6, focusing on the prospective detection of post-Newtonian effects on the orbit of S2, and the future means of high-precision astrometry in particular, using the GRAVITY instrument at the VLTI, next-generation instruments at ELTs, as well as advanced methods of data analysis, in a complementary way.

Chapter 2

A Sgr A* rest frame for high-precision Galactic Center astrometry

Original publication: P. M. Plewa, S. Gillessen, F. Eisenhauer, T. Ott, O. Pfuhl, E. George, J. Dexter, M. Habibi, R. Genzel, M. J. Reid and K. M. Menten, 2015, MNRAS, *Pinpointing the near-infrared location of Sgr A* by correcting optical distortion in the NACO imager*, doi:10.1093/mnras/stv1910

Abstract: Near-infrared observations of stellar orbits in the Galactic Center have provided conclusive evidence for a massive black hole associated with the compact radio source Sgr A*. The astrometric reference frame for these observations is tied to a set of red giant stars, which are also detectable at radio wavelengths due to SiO maser emission in their envelopes. We have improved the precision and long-term stability of this reference frame, in which Sgr A* is localized to within a factor 5 better than previously: ~ 0.17 mas in position (in 2009) and ~ 0.07 mas yr^{-1} in proper motion. This improvement is the result of modeling and correcting optical distortion in the NACO imager to a sub-mas level, and including new infrared and radio measurements, which now both span more than a decade in time. A further improvement will follow future observations, and is expected to facilitate the detection of post-Newtonian orbital effects.

2.1 Introduction

The observability of the Galactic Center in the near-infrared and its relative proximity at a distance of ~ 8 kpc allow the study of a galactic nucleus in unparalleled detail. AO-assisted imaging observations of the Galactic Center with NACO at the ESO VLT, as well as NIRC2 at the Keck observatory, routinely achieve a resolution close to the diffraction limit, and a depth limited by a high level of stellar crowding. Since the discovery of stars with large proper motions (Eckart & Genzel, 1996; Ghez et al., 1998), the continued monitoring of individual stellar orbits has produced strong evidence for a compact mass of $\sim 4 \times 10^6 M_\odot$ that can be convincingly attributed to a quiescent massive black hole (e.g. Ghez et al., 2008; Gillessen et al., 2009b) at the location of the radio source Sgr A* (e.g. Menten et al., 1997; Reid et al., 2003, 2007). A further increase in the astrometric precision would permit measuring the full orbits of more stars than presently possible, permit refined estimates of the black hole's mass and distance, and eventually permit the detection of relativistic effects on the orbits of short-period stars, most immediately post-Newtonian effects on the orbit of the star S2 (e.g. Jaroszynski, 1998; Fragile & Mathews, 2000; Rubilar & Eckart, 2001; Weinberg et al., 2005; Zucker et al., 2006; Nucita et al., 2007; Will, 2008; Merritt et al., 2010; Angélil et al., 2010; Eisenhauer et al., 2011).

While the motions of the (S-)stars within the central few $1''$ are measured from images taken with the S13 camera of NACO (at a scale of ~ 13 mas px^{-1}, see Rousset et al., 1998; Lenzen et al., 1998), less frequent observations with the S27 camera (at a ~ 27 mas px^{-1} scale) are used to efficiently set up the astrometric reference frame (e.g. Trippe et al., 2008). However, the S27 camera is affected by optical image distortion on a non-negligible level, as for example is the narrow-field camera of NIRC2 (Yelda et al., 2010). By systematically altering the relative positions of stars on the detector, this effect becomes a substantial source of uncertainty when positions measured on widely dithered images are combined into a common coordinate system, or when they are aligned with astrometric reference sources.

In our earlier studies of stellar orbits in the Galactic Center (e.g. Schödel et al., 2002, 2003; Eisenhauer et al., 2003a, 2005; Gillessen et al., 2009b), we have used high-order coordinate transformations to mitigate the effect of image distortion and relied primarily on a reference frame based on the assumption that a sufficiently large number of evolved stars in the nuclear cluster should show no net motion. The stellar motions in this 'cluster rest frame' were then placed in an astrometric coordinate system centered on Sgr A* by aligning the positions of several SiO maser stars at a certain epoch with their positions as predicted from radio observations. The comparison of the S-stars' positions over several years has still been limited by a systematic uncertainty of ~ 2 mas associated with the definition of the coordinate system, likely because of residual distortion in the S27 camera. Yet the typical uncertainty of a single position is as low as ~ 0.3 mas, limited by residual distortion in the S13 camera, uncertainty in modeling the PSF, and ultimately source confusion (Fritz et al., 2010b). The long-term stability of the cluster rest frame is also fundamentally limited to $\sigma/\sqrt{n} \approx 0.07$ mas yr^{-1} by the intrinsic velocity dispersion of the selected stars in the plane of the sky ($\sigma \approx 3$ mas yr^{-1}; Trippe et al. 2008) and the number of available stars ($n \approx 2000$; Gillessen et al. 2009b).

In this paper, we construct an alternative astrometric reference frame, which has now become more advantageous, by relating many well-measured stars to the SiO maser stars directly at multiple epochs, and then using their motion with respect to Sgr A* for reference. The details on the construction of this 'Sgr A* rest frame' (Yelda et al., 2010) are described in Sec. 2.4, and the precision of localizing (radio-)Sgr A* in the infrared frame has important implications for the analysis of stellar orbits (see Sec. 2.5). To improve this precision, we have implemented a more accurate distortion correction for the S27 camera of NACO. In Sec. 2.3, the apparent distortion is measured at four epochs between 2004 and 2012 from images of globular clusters in the ESO archive. This is possible by matching

	ESO Prog. ID	Date	Target	Filter	DIT (s)	NDIT	Images	Sample Size	Indiv. Sources
(a)	074.D-0151(A)	2004-10-12	47 Tuc.	Ks	2.5	24	20	19826	1698
(b)	482.L-0793(A)	2009-07-03	ω Cen.	Ks	0.3454	100	22	3910	534
(c)	60.A-9800(J)	2010-02-07	ω Cen.	Ks	1.	10	20	3658	388
(d)	089.C-0638(A)	2012-07-22	47 Tuc.	Ks	1.	30	10	9671	1824

Table 2.1: Summary of globular cluster observations.

them to HST catalogs, which can be considered distortion-free in comparison. Based on the resulting distortion models, it is then possible to also correct other images taken close in time. A per-night self-calibration approach (as used for calibration of the HST, see e.g. Anderson & King, 2003) is not feasible for our Galactic Center observations, because the requirement that many of the same stars are placed in many different regions on the detector is typically not met. To start with, we outline the processing procedure we have applied to our data, in Sec. 2.2.

2.2 Image processing

2.2.1 Observations

We make use of two sets of observations with NACO for different purposes, which target the Galactic Center, as well as the densely populated cores of the two globular clusters ω Centauri and 47 Tucanae. The globular cluster observations that result in viable distortion corrections are summarized in table 2.1. The instrument was configured to use the S27 camera, the Ks-band filter, the Aladdin 3 detector in the high dynamic range mode, and the readout mode ‘read-reset-read’. This is also the typical setup for several epochs of Galactic Center observations using the same camera, which together establish the astrometric reference frame, and are summarized in table 2.2. A small number of images are omitted, for example those compromised by severe saturation of the relevant SiO maser stars, or those affected by PSF artifacts caused by a technical failure in the AO system. The dithering scheme is typically a wide square pattern, such that the total field of view is $42'' \times 42''$ and the central $14'' \times 14''$ are present in all images. At each pointing position, either two, four or six images are taken, before moving on. A small offset of a few pixels is applied on revisiting a pointing position, to minimize the effect of detector artifacts on mosaic images.

We also make use of high-quality HST proper motion catalogs available for ω Centauri (Anderson & van der Marel, 2010) and 47 Tucanae (Bellini et al., 2014, priv. comm.), which provide us with a nominally distortion-free reference system. The HST astrometry profits from the stable observing conditions in space, and thanks to a well-characterized PSF and distortion correction for the instrument channels used, the measurement uncertainty of stellar positions on a single HST image is as low as ~ 0.01 px, or ~ 0.5 mas (e.g. Anderson & King, 2006). More details on the HST observations can be found in the description of the HSTPROMO program by Bellini et al. (2014), and references therein.

ESO Prog. ID	Date	Filter	DIT (s)	NDIT	Pointings × Images	PSF FWHM (px)	Maser Stars	Reference Stars	Distortion Correction
60.A-9026(A)	2002-04-01	Ks	0.5	8	10×1	3.8	7	73	model (a)
71.B-0077(A)	2003-05-09	Ks	0.5	120	19×1	2.7	8	91	model (a)
073.B-0085(E)	2005-05-12	Ks	0.5	60	16×6	2.9	8	91	model (a)
077.B-0014(A)	2006-04-28	Ks	2.0	28	8×4	3.3	7	87	model (a)
077.B-0014(E)	2006-08-27	Ks	2.0	28	8×4	4.5	7	79	model (a)
078.B-0136(B)	2007-03-16	Ks	2.0	28	7×4	2.6	8	91	model (b)
179.B-0261(A)	2007-03-31	Ks	5.0	6	16×6	2.6	8	91	model (b)
179.B-0261(M)	2008-04-04	Ks	2.0	28	8×4	2.8	8	91	model (b)
179.B-0261(N)	2008-08-04	Ks	1.0	57	8×4	3.0	8	91	model (b)
179.B-0261(U)	2008-09-15	Ks	1.0	57	8×4	2.7	8	91	model (b)
179.B-0261(X)	2009-03-28	Ks	1.0	60	8×4	2.9	8	91	model (b)
179.B-0261(X)	2009-03-30	Ks	1.0	60	8×4	2.7	8	91	model (b)
183.B-0100(J)	2009-09-19	Ks	1.0	60	16×4	2.7	8	91	model (b)
183.B-0100(J)	2009-09-20	Ks	1.0	60	8×4	2.8	8	91	model (b)
183.B-0100(T)	2010-05-08	Ks	0.5	126	8×4	3.2	8	90	model (c)
183.B-0100(V)	2010-09-27	Ks	1.0	126	8×2	3.4	8	86	model (c)
183.B-0100(X)	2011-04-01	Ks	1.0	66	8×4	3.1	8	86	model (c)
183.B-0100(V)	2011-05-16	Ks	2.0	9	8×2	2.6	8	91	model (c)
088.B-1038(A)	2012-03-14	Ks	1.0	30	41×1	3.3	8	86	model (d)
088.B-0308(B)	2012-05-03	Ks	0.9	33	8×6	3.4	8	86	model (d)
089.B-0162(D)	2012-08-08	Ks	1.0	60	8×4	2.7	8	91	model (d)
091.B-0081(F)	2013-05-13	Ks	0.9	33	16×6	2.6	8	90	model (d)

Table 2.2: Summary of Galactic Center observations.

2.2.2 Image reduction

The standard steps of sky subtraction, division of a flat-field, and a bad pixel correction are applied to every raw NACO image. The nightly sky background is estimated by taking the median image of either a set of dedicated sky exposures, or else the randomly dithered science exposures. The flat-field is created by stacking and normalizing lamp exposures taken during daytime calibrations. The unusable hot or dead pixels are identified by comparing the ADU counts in neighboring pixels, from which a replacement value is interpolated, if necessary. Individual detector integrations were recorded for most images since 2010, so that frames of exceptionally low quality can be rejected, before the remaining frames are averaged to create the final image.

2.2.3 Star lists

With the aim of achieving the highest astrometric precision, we extract the detector positions (and fluxes) of stars in every reduced NACO image using an empirical model of the PSF (see Anderson & King, 2000). To identify the stars, an image is first correlated with the current PSF estimate, with the correlation operator being the normalized cosine distance[1]. Initially, a Gaussian kernel is substituted for the PSF and since this kernel is

---

[1] This distance is defined as $\frac{u-\overline{u}}{|u-\overline{u}|} \cdot \frac{v-\overline{v}}{|v-\overline{v}|}$ between two vectors u and v, which in our case are one-dimensional arrangements of the correlation kernel and the neighborhood of a certain pixel, respectively.

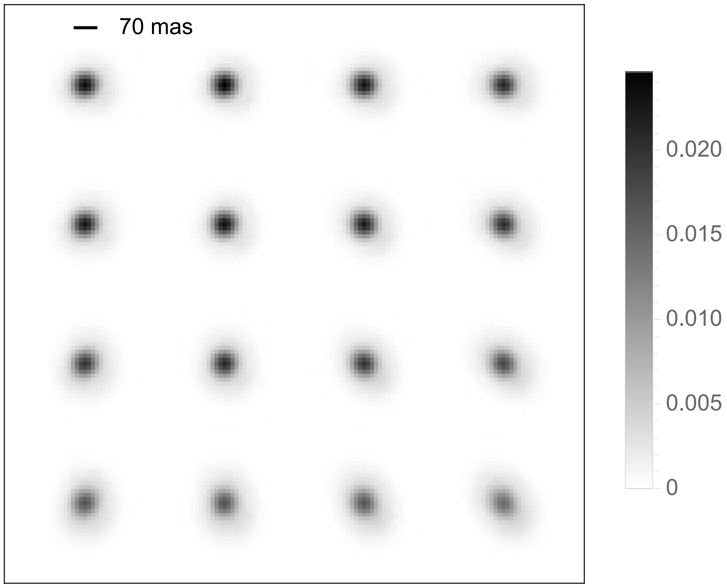

Figure 2.1: Example of a PSF grid obtained from a Galactic Center image with median Strehl ratio. Each PSF is normalized to unit flux. The AO guide star (IRS 7) is located near the top of the image. The typical deviation from the mean PSF in a single pixel is about $10 - 35\%$.

symmetric, all PSFs derived from it are thus centered. The result is a map of coefficients between -1 and 1, which measure at every pixel how well the image matches the PSF locally. The location of a candidate star is the peak of a connected pixel region in this map, for which the correlation coefficient exceeds a threshold value of 0.7 and the ADU counts fall in the range between the noise level and the full-well capacity of the detector. At each candidate position, the PSF is fit to a roughly circular, top-hat image region by method of least squares. We allow for a free position offset and a free scaling factor, as well as three additional parameters, that describe a tilted planar background, to minimize a potential bias in crowded regions. The radius of the fit region, i.e. the number of pixels surrounding a central pixel at all sides, is typically 4 px, but is tied to the FWHM of the PSF, such that the entire PSF core is included.

The actual PSF not only changes with time, but also varies spatially. For example, the PSF gets significantly broader and elongated at distances of $10''$ to $20''$ from the AO guide star (IRS 7), as the correction of atmospheric turbulence degrades due to anisoplanatism. To take into account such variations, we use one 4×4 grid of PSFs per image, each of which is derived from up to fifty, ideally bright and isolated stars, within cells of 256 px × 256 px size spanning the detector (Fig. 2.1). By using a bicubic interpolation to resample the surrounding image regions, the selected stars are magnified by a factor 2 and centered in sub-images typically 35 px wide, depending on the FWHM of the PSF, and also normalized to zero background and unit flux. The median superposition of these sub-images, weighted by the square root of the original fluxes, is then used to estimate the PSF at the center of a grid cell, and a modified cosine window function is used to taper the extended PSF halos. The PSF at a certain detector position is finally estimated by means of a bilinear inter- or extrapolation on the whole grid, using spline interpolations of third order to evaluate each grid PSF at sub-pixel offsets.

The two iterative steps of extracting a PSF grid and creating a star list are repeated three times. In the case of mildly saturated stars, pixels with an ADU count above 80% the full-well capacity are ignored throughout, as long as the fraction of excluded pixels is small ($< 10\%$). Heavily saturated stars are never identified as candidate stars in the first place. The final outcome of this process are lists that contain the positions of all detected stars in the pixel coordinate system of each particular image. They also contain the fluxes of the stars, which are converted into instrumental magnitudes. Only stars brighter than the distribution's peak magnitude are kept ($m_K \lesssim 16$). These are likely to be genuine detections, for which positions can be measured with similarly good precision.

2.3 A distortion correction for NACO

2.3.1 Star matching

To measure the distortion in the globular cluster images, the respective star lists (Sec. 2.2.3) are first matched to a proper motion catalog obtained with the HST. A reference star list is created for each observing night, by propagating all catalog positions to the correct date. Since the images from a single night overlap, it is convenient to determine the relative offsets between the NACO star lists by cross-correlating the original images. Initial transformations that shift and scale the HST star list to roughly align with each NACO star list are determined manually. Matching pairs of stars are then found using a nearest-neighbor search. We require that for each detected star, to make a match with a catalog star, there has to exist exactly one star within a certain radius around it on the transformed HST list, and no other star on its own NACO star list. The latter requirement effectively excludes stars with close neighbors. Additionally, the match needs to be unique. Once matching pairs of stars are known, they can be used to find better transformations for the original HST star list, and another matching process follows. The cutoff radius for the nearest-neighbor search is set to half the FWHM of the PSF, but five times larger for the first iteration. A last quality criterion is the number of matched stars per image. It should be greater than twenty, to allow the combination of multiple star lists later on. Typically, a few thousand detector positions of a few hundred individual stars are accepted in total, which cover the detector reasonably well.

2.3.2 The distortion model

The image distortion we aim to correct is expected to be static, but might still change sporadically due to differences in the optical alignment before and after interventions on the instrument, for example. All our Galactic Center observations, spread over the whole

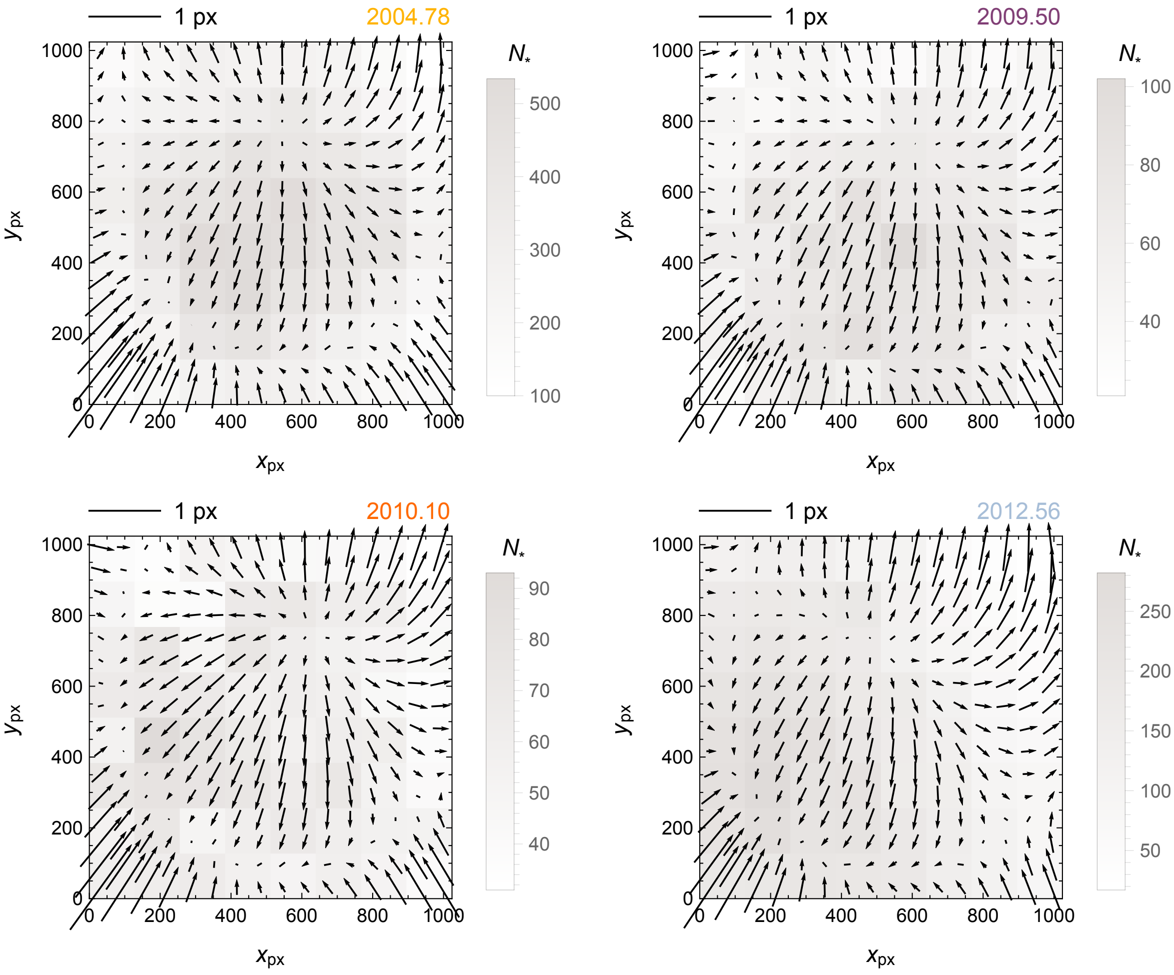

Figure 2.2: Evolution of the image distortion in the S27 camera of NACO. Shown is the shift required to move a measured pixel position into a pixel coordinate system approximately free of distortion (black arrows). Also shown is the number of reference sources in different regions on the detector (grey-shaded bins). Each model is only valid for a limited period of time, as shown in Fig. 2.3.

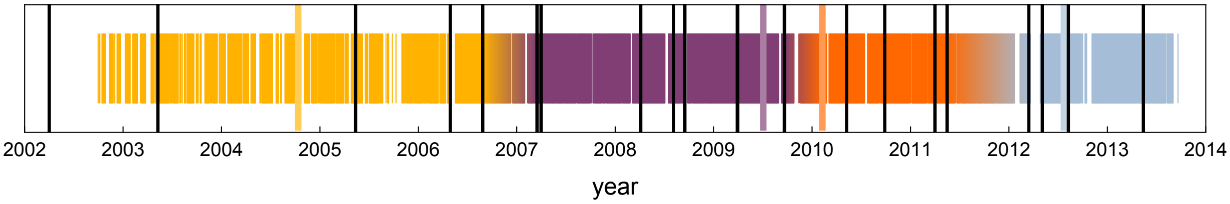

Figure 2.3: Timeline of records in the NACO archive, on which are marked the different epochs of globular cluster observations and the periods of time during which the corresponding distortion corrections are valid (in color), as well as the epochs of Galactic Center observations that together establish the astrometric reference frame (in black). The exact dates of significant changes in the image distortion might coincide with interventions on the instrument, but can not be determined exactly from our data. However, it is straightforward to decide whether a certain distortion correction is applicable, by comparing the resulting image registration accuracy.

time-range of NACO operations from 2002 to 2013, are in the end covered by four different distortion corrections (Fig. 2.2 & Fig. 2.3).

For each of the four epochs of globular cluster observations (Table 2.1), a linear transformation is determined for every NACO star list that aligns the matched reference positions from the HST catalog with the stellar positions as measured on the detector, in a least squares sense. More precisely, it is an affine transformation (with six parameters), which already combines translation, rotation, scaling and shear operations. The map of transformation residuals, i.e. the remaining differences between aligned positions, is a vector field that represents the image distortion. However, to create a single map from multiple images, it is necessary to apply a preliminary distortion correction to the NACO star lists at least once, recalculate the transformations, and refit a model to the pooled residuals. This iteration reduces a potential bias that occurs if stellar positions from individual lists are not distributed uniformly over the detector.

The distortion model we adopt is a linear combination of vector fields from a complete orthonormal basis that can describe arbitrary vector fields on the unit circle (Table 2.3). Such a basis can be constructed from the union of a set of vector polynomials with zero curl and another set with zero divergence (see Zhao & Burge, 2007, 2008). Both sets are needed to account for aberrations of the optical system and the arrangement of the detector. This model is fit to each of the four distortion maps using a least squares technique, after the detector coordinates have been rescaled accordingly and the maps have been smoothed with a median filter to clean them from outliers (e.g. falsely matched stars). We pick a smoothing length of 64 px to roughly correspond to the scale at which the median absolute deviation between the smoothed and the original vector fields becomes approximately constant.

Twenty free parameters are needed to fully capture the spatial variability of the image distortion. This number can be inferred by inspecting the fit residuals, which stop decreasing when enough basis fields of higher order have been included. Even so, most

V	$V_x(x,y)$	$V_y(x,y)$
S_2	a_2	0
S_3	0	a_3
S_4	$\sqrt{2}a_4x$	$\sqrt{2}a_4y$
S_5	$\sqrt{2}a_5y$	$\sqrt{2}a_5x$
S_6	$\sqrt{2}a_6x$	$-\sqrt{2}a_6y$
S_7	$\sqrt{6}a_7xy$	$\sqrt{\frac{3}{2}}a_7(x^2+3y^2-1)$
S_8	$\sqrt{\frac{3}{2}}a_8(3x^2+y^2-1)$	$\sqrt{6}a_8xy$
S_9	$2\sqrt{3}a_9xy$	$\sqrt{3}a_9(x-y)(x+y)$
S_{10}	$\sqrt{3}a_{10}(x-y)(x+y)$	$-2\sqrt{3}a_{10}xy$
S_{11}	$2a_{11}x(3x^2+3y^2-2)$	$2a_{11}y(3x^2+3y^2-2)$
S_{12}	$2\sqrt{2}a_{12}x(2x^2-1)$	$2\sqrt{2}a_{12}y(1-2y^2)$
S_{13}	$2\sqrt{2}a_{13}y(3x^2+y^2-1)$	$2\sqrt{2}a_{13}x(x^2+3y^2-1)$
S_{14}	$2a_{14}(x^3-3xy^2)$	$2a_{14}y(y^2-3x^2)$
S_{15}	$-2a_{15}y(y^2-3x^2)$	$2a_{15}(x^3-3xy^2)$
T_4	$\sqrt{2}b_4y$	$-\sqrt{2}b_4x$
T_7	$\sqrt{\frac{3}{2}}b_7(x^2+3y^2-1)$	$-\sqrt{6}b_7xy$
T_8	$\sqrt{6}b_8xy$	$-\sqrt{\frac{3}{2}}b_8(3x^2+y^2-1)$
T_{11}	$2b_{11}y(3x^2+3y^2-2)$	$-2b_{11}x(3x^2+3y^2-2)$
T_{12}	$2\sqrt{2}b_{12}y(1-2y^2)$	$2\sqrt{2}b_{12}x(1-2x^2)$
T_{13}	$2\sqrt{2}b_{13}x(x^2+3y^2-1)$	$-2\sqrt{2}b_{13}y(3x^2+y^2-1)$

Table 2.3: Explicit form of the distortion model in terms of its basis vector fields (for a derivation, see Zhao & Burge, 2007, 2008).

of the distortion can be attributed to the low-order aberrations y-tilt, defocus, coma and spherical aberration. In terms of the number of parameters, an equivalent model could be composed of two third-order polynomials in two variables, each of which is used to describe the distortion along one dimension[2].

The characteristic shift required to move a measured position into a pixel coordinate system approximately free of distortion is ~ 0.2 px (~ 5.4 mas), but the actual shift varies strongly with location on the detector and can be as high as ~ 0.7 px (~ 19 mas) in the lower left corner. The distortion pattern changes distinctly as the pairwise rms deviation between subsequent models is on the order of ~ 0.1 px (~ 2.7 mas), but the general pattern is static.

---

[2] However, more robust fit results are achieved using the vector field basis functions, since these are orthonormal.

2.4 The Sgr A* rest frame

2.4.1 Image registration

To be able to measure the motions of stars later used as astrometric reference sources, individual images from each night of Galactic Center observations (Table 2.2) must first be combined into a common coordinate system.

We start by combining the star lists of consecutive images belonging to the same pointing position, which share similar observing conditions. Stars are iteratively matched between those lists, allowing for small offsets between the images ($\lesssim 0.2$ px). New star lists are then created from the mean positions, but, to further purge the lists of spurious detections, only stars detected in all images are kept. The measurement uncertainty of the new positions is estimated by the standard error of the mean, and has a typical value of about ~ 0.3 mas (mode of histogram).

The single-pointing star lists are then combined into one master star list per epoch. Affine transformations should suffice to register the star lists after they have been corrected for image distortion, since any remaining non-linear displacements should be negligible (considering for example atmospheric refraction, aberration due to the motion of the Earth, light deflection in the gravitation field of the Sun, or curvature of the celestial sphere). Nevertheless, we explicitly correct the star lists for differential achromatic refraction as well, to minimize anisotropy of the pixel scale in the final master star list (see Gubler & Tytler, 1998).

The initial master star list is the first star list containing the maximum number of SiO maser stars. We proceed by iteratively matching to it the list with the next smallest pointing offset, finding an affine transformation between the matched positions and creating a star catalog that contains all aligned positions, grouped by star (unmatched positions are simply appended). The new master list is a collapsed version of this catalog, i.e. it contains the weighted average positions and the mean uncertainties. In this way, both the catalog and the master star list are updated, until all lists are merged. The whole stitching process is repeated once, with the only change being that every star list is matched to the same intermediate master star list.

A few epochs are treated differently in some ways. For observations before 2004, some image reduction parameters have to be adjusted for the different characteristics of the Aladdin 2 detector, which was replaced during that year. For observations with either a random dithering scheme, or less than three images per pointing, the measurement uncertainty of stellar positions can not be estimated reliably from the images themselves. In these cases, we instead apply a mean error model based on the other observations by fitting a power law to the one-dimensional uncertainties of consistently detected stars as a function of flux (Fig. 2.4). The photometric zero-point is calibrated using the 91 primary astrometric reference stars (Sec. 2.4.3; Gillessen et al., 2009b).

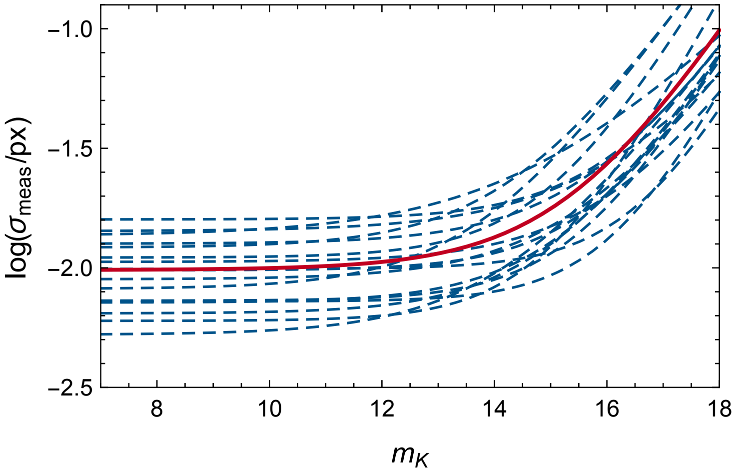

Figure 2.4: The best-fitting models for the measurement uncertainty of stellar positions as a function of K-band magnitude (dashed lines) and the mean model (solid line), which is applied whenever the uncertainty can not be estimated reliably from the images themselves (see text).

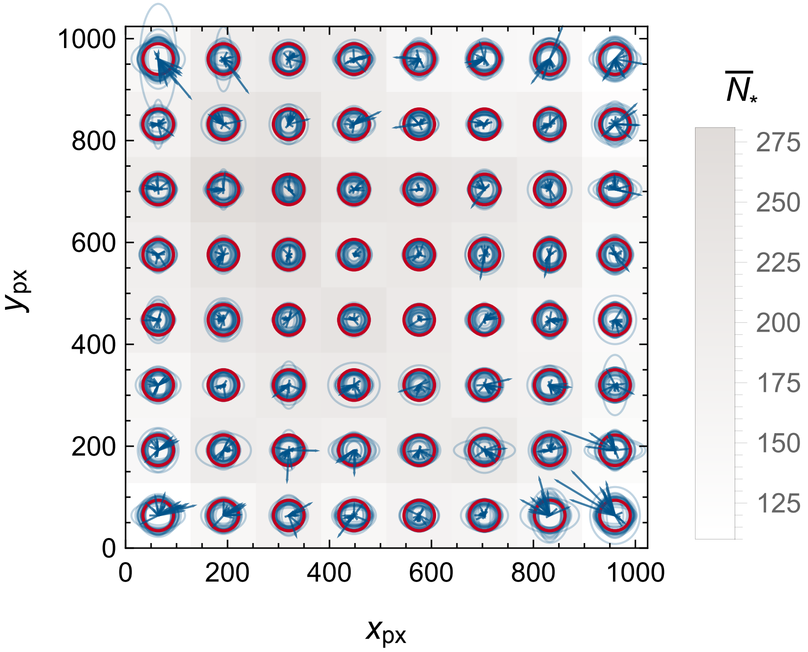

Figure 2.5: The mean (blue arrows) and standard deviation (blue circles) of the standardized registration residuals in bins on the detector, overlaid for all epochs. A systematic deviation from a standard normal distribution (red circles) indicates uncorrected distortion, or unrealistic uncertainty estimates.

2.4.2 Testing the distortion correction

After the image registration step, any uncorrected distortion becomes apparent in non-random structure of the final transformation residuals. This is because the positions of the same stars, measured in different locations on the detector, should align nevertheless. We can thus check the quality of the distortion corrections and identify approximately the periods during which a particular solution is valid (see Fig. 2.3).

Every transformation residual is divided by its expected uncertainty, i.e. the combined uncertainty of the two respective positions, and sorted into bins of 128 px × 128 px size on the detector (Fig. 2.5). The uncertainties of the transformation parameters are negligible, given that a few hundred well-measured stars can always be matched between two star lists. An additional statistical uncertainty of ~ 0.1 mas needs to be added (in quadrature) to each stellar position to make the overall distribution of the so-standardized residuals approximate a standard normal distribution. However, the agreement is best at the center of the detector and worse at the edges, where the remaining time-averaged distortion is usually at most ~ 0.2 mas. We therefore add a different uncertainty in each bin, and then average over a star's detector positions. The distortion model is inherently less constrained at the edges of the detector, but the actual distortion is also expected to vary on a low level even during the night, for example due to unstable performance of the AO system.

2.4.3 The astrometric reference stars

The proper motions of several late-type giant stars have been measured relative to Sgr A* directly at radio wavelengths, by observing their circumstellar SiO maser emission (Menten et al., 1997; Reid et al., 2003, 2007). The results have since been updated with the addition of more data and an improved analysis, including in particular a correction of differential precession, which leads to an apparent rotation[3] (Reid et al., 2015, priv. comm.). Eight of these maser stars lie in the central $7''$ to $24''$ and are typically inside the NACO field of view. Not included is the exceptionally bright star IRS 7, because saturation of the detector prevents an accurate measurement of its infrared position. As the only red supergiant in the sample, its radio position would also have a considerably larger uncertainty than the ~ 0.5 mas intrinsic to red giants, because it is expected to have a larger SiO maser shell ($\gg 4$ AU), and the expected variation of the maser emission on time-scales of ~ 1 yr is not resolved.

At each of the 22 epochs of Galactic Center observations, the infrared positions of the maser stars, taken from the master star list (Sec. 2.4.1), are aligned with the propagated radio positions using a weighted affine transformation, thus taking into account the uncertainties in both sets of positions. The pixel coordinates $(x_{\mathrm{px}}, y_{\mathrm{px}})$ of all detected stars are thereby converted into angular offsets (x, y) from Sgr A*[4]. The positional uncertainties,

---

[3] It has since been shown that this correction was applied in error. To be consistent with the ICRS (J2000) reference frame, the astrometric positions and proper motions reported in this chapter must be rotated by an angle $\theta(t) \approx -0.00645674^\circ(t-2000)$ east of north around (radio-)Sgr A*.

[4] $+x$ increases to the west and $+y$ to the north. The conversion to celestial coordinates is given to good

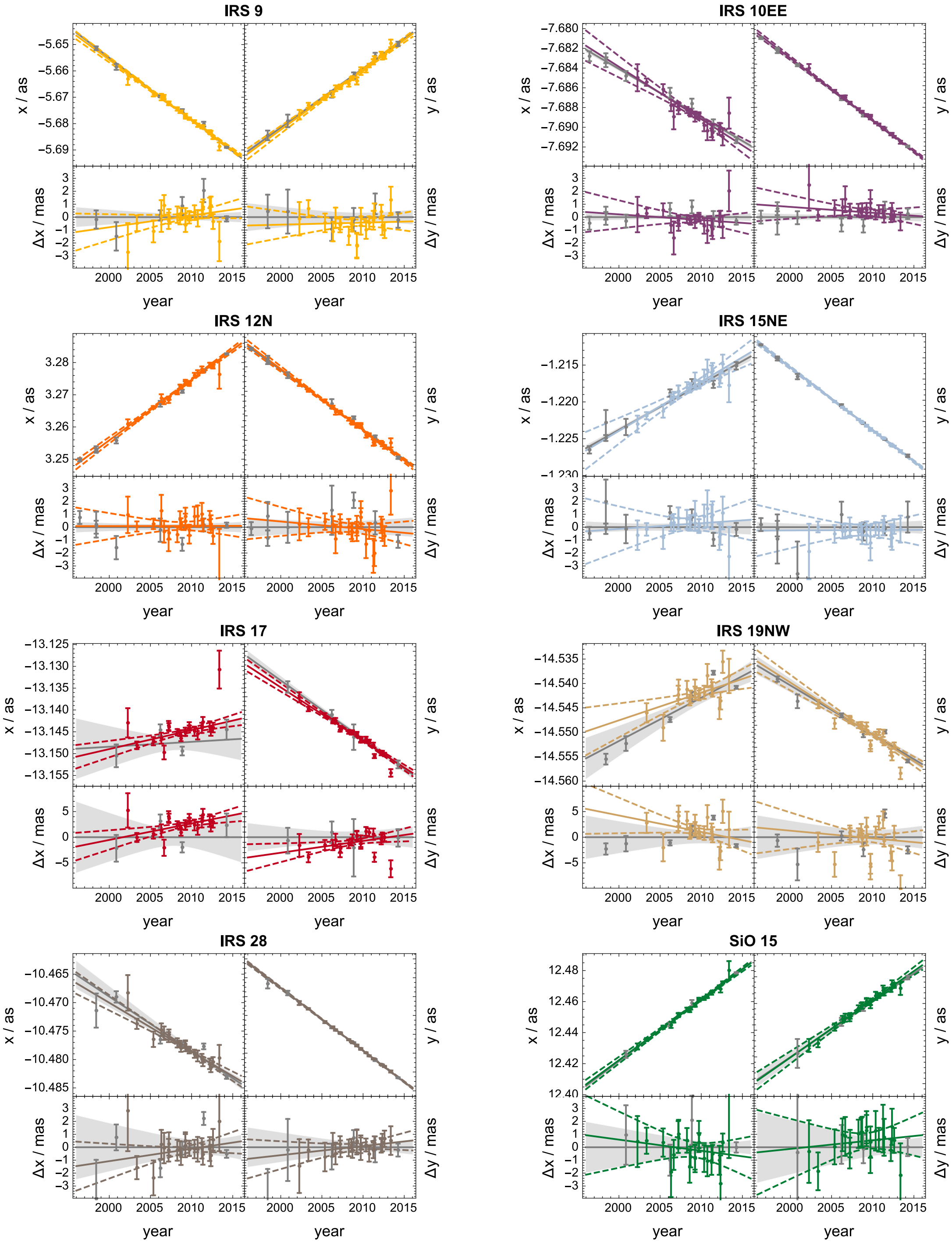

Figure 2.6: The infrared positions of the SiO maser stars, and the best-fitting proper motion models (in color, see also sup. Fig. 2.11). In each panel, the top plots show the motion relative to Sgr A*, and the bottom plots show the motion relative to the radio reference measurements (in gray).

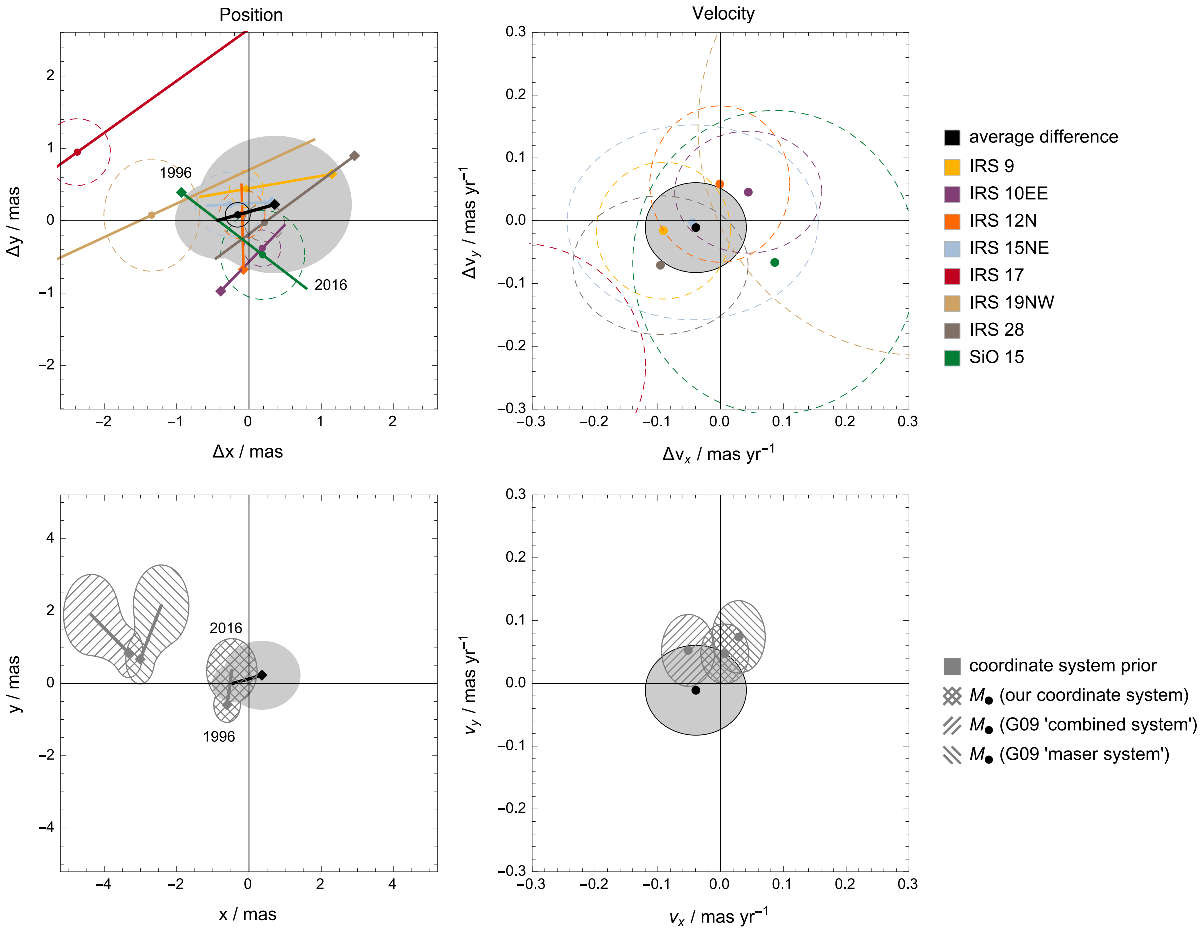

Figure 2.7: Top row: Differences between the radio- and infrared-measured linear motions of the SiO maser stars in position (colored lines on the left) and velocity (colored dots on the right), shown with their 1σ-uncertainties (dashed ellipses). The position differences are evaluated between 1996 (marked by diamond symbols) and 2016 (unmarked endpoints), with the average reference epoch being 2009.02 (marked by dots in between). Bottom row: The drift motion of the central mass required by the best-fitting orbit of the star S2, using our coordinate system (cross-hatched region) compared with the 'combined system' and the 'maser system' from the study by Gillessen et al. (2009b, single-hatched regions). The average deviation shown above (in black and gray) indicates how precisely (radio-)Sgr A* is localized in the infrared reference frame, and is used in the fit as prior information on this drift motion, as indicated here (see Sec. 2.4.4).

ID	t_0 (year)	x_0 (mas)	y_0 (mas)	v_x (mas yr^{-1})	v_y (mas yr^{-1})
IRS 9	2009.17	-5676.73 ± 0.23	-6333.35 ± 0.26	-2.29 ± 0.11	1.62 ± 0.11
IRS 10EE	2009.00	-7688.65 ± 0.26	4186.87 ± 0.25	-0.53 ± 0.12	-2.87 ± 0.10
IRS 12N	2008.56	3271.96 ± 0.31	-6912.22 ± 0.34	1.88 ± 0.11	-2.49 ± 0.12
IRS 15NE	2008.43	-1218.31 ± 0.46	11260.96 ± 0.43	0.67 ± 0.20	-5.90 ± 0.16
IRS 17	2009.04	-13145.03 ± 0.46	5542.86 ± 0.46	0.44 ± 0.20	-2.35 ± 0.20
IRS 19NW	2009.55	-14542.15 ± 0.63	-18483.37 ± 0.75	0.57 ± 0.36	-2.38 ± 0.36
IRS 28	2009.44	-10478.15 ± 0.27	-5853.57 ± 0.26	-0.86 ± 0.14	-6.64 ± 0.11
SiO 15	2009.28	12458.47 ± 0.58	-11060.20 ± 0.61	3.90 ± 0.23	2.15 ± 0.24

Table 2.4: Proper motions of the SiO maser stars in the Sgr A* rest frame.

including both measurement uncertainty and the extra uncertainty due to residual distortion (Sec. 2.4.2), are propagated using a Monte Carlo Bootstrap technique, along with the uncertainties of the transformation parameters themselves.

A total of 10^4 realizations of one transformation are generated by applying random displacements to the infrared and radio positions of the maser stars, according to their known uncertainties, while simultaneously resampling the pairs of corresponding positions with replacement. The artificial transformations are applied to the other stars as well, the pixel positions of which are repeatedly perturbed in the same way. A small number of potential realizations are excluded ($< 1\%$), since at least three unique pairs of positions are needed to define an affine transformation. The uncertainty of each star's astrometric position is finally estimated by the standard deviation of the sample of transformed positions, which in turn is estimated by the median absolute deviation. This statistic is more robust against outliers than the rms deviation, for example. Likewise, the uncertainty of the detector position of Sgr A*, typically ~ 0.6 mas, can be found by inverting the transformations and finding the pixel positions that map on to the origin. Although the typical uncertainty of the maser stars' detector positions is ~ 0.3 mas, that of the astrometric positions is ~ 0.8 mas and evidently dominated by uncertainty in the alignment. The main reason is the thinly scattered distribution of the maser stars across the field of view, but also the additional uncertainty of the radio positions.

We fit each maser star's astrometric positions as a function of time, using both a linear and a quadratic proper motion model, separately in x and y (Fig. 2.6 & Table 2.4). For simplicity, we use one reference epoch per star, instead of one per fit, and accept that correlations between fit parameters are thus not entirely eliminated. The uncertainties of the fit parameters are estimated from a Monte Carlo sample, created by performing fits to the many positions of each maser star generated at every epoch, which are additionally resampled with replacement in time. The mean reduced χ^2 of 0.97 suggests that the positional uncertainties have possibly been slightly overestimated.

Whether a detected acceleration is genuine is decided by two criteria. First, the direction of the acceleration must be towards Sgr A*, i.e. the negative radial component must be statistically significant, while the tangential component must be insignificant, at

---

approximation by $(x, y) = (-\Delta\alpha\cos(\delta), \Delta\delta)$.

the 5σ level. Second, the magnitude of the acceleration must be smaller than the upper limit imposed by the gravitational force that the massive black hole associated with Sgr A* can exert, assuming that this is the dominant force, and given that the observed projected separation from Sgr A* is the minimum true separation. None of the best-fitting quadratic models satisfy both criteria, and we therefore keep the linear models.

Another step of calibration is necessary to measure the astrometric positions of the S-stars on images taken with the S13 camera, because the maser stars are either saturated in the deeper exposures, or fall outside the smaller field of view. The connection between the two image scales is made by a sample of 91 astrometric reference stars, which can always be observed together with either the S-stars (in images taken with the S13 camera) or the maser stars (in images taken with the S27 camera). These reference stars lie in the central $0.8''$ to $5''$ (see sup. Fig. 2.11) and were chosen to be relatively isolated from other known sources (Gillessen et al., 2009b). We fit proper motion models for the reference stars analogously to the maser stars. Since also none of them show significant plausible accelerations (at least so far), we discard the quadratic models. Lastly, we inspect the fit residuals separately for individual stars and epochs. In particular in 2013, there is a systematic pattern originating from a misalignment of the maser stars, but it is consistent with the statistical uncertainties.

The overall detection rate of both the reference and the maser stars is about 97%. Some fainter reference stars were sometimes not detected in a pointing when they could have been, due to a combination of varying image quality and source confusion. Non-detections of the maser stars were caused by saturation of IRS 17 and IRS 9, as well as the placement of IRS 12N close to the edge of the images with a north-east pointing.

2.4.4 The location of Sgr A*

The best-fitting proper motion models for the SiO maser stars can be compared with the predicted radio motions, to assess the precision and stability of the Sgr A* rest frame over time (Fig. 2.7 & Table 2.5). We find that the radio- and infrared-measured motions of the maser stars are consistent with each other, and that the average difference motion is, in units of mas:

$$\Delta x(t) \approx (-0.16 \pm 0.17) + (-0.04 \pm 0.08)(t - 2009.02) \tag{2.1}$$

$$\Delta y(t) \approx (+0.08 \pm 0.17) + (-0.01 \pm 0.07)(t - 2009.02) \tag{2.2}$$

We therefore conclude that (radio-)Sgr A*, i.e. the origin of the radio coordinate system, can be expected to be localized in the infrared reference frame with a precision of ~ 0.17 mas in position (in 2009) and ~ 0.07 mas yr^{-1} in proper motion (~ 2.7 km s^{-1}). Moreover, the average velocity differences in radial and tangential direction indicate that the infrared reference frame shows neither pumping (v_r/r) nor rotation (v_ϕ/r) relative to the radio system, to within ~ 7.5 μas yr^{-1} as^{-1} and ~ 7.0 μas yr^{-1} as^{-1}, respectively.

To calculate the average of the position and velocity differences, we have used as weights the uncertainties of the infrared quantities, which implicitly contain the propagated uncertainties of the radio motions. To estimate the uncertainties of the average values, we

ID	Δx_0 (mas)	Δy_0 (mas)	Δv_x (mas yr^{-1})	Δv_y (mas yr^{-1})
IRS 9	-0.04 ± 0.23	0.44 ± 0.26	-0.09 ± 0.11	-0.02 ± 0.11
IRS 10EE	0.18 ± 0.26	-0.38 ± 0.25	0.04 ± 0.12	0.05 ± 0.10
IRS 12N	-0.09 ± 0.31	0.09 ± 0.34	-0.00 ± 0.11	0.06 ± 0.12
IRS 15NE	-0.26 ± 0.48	0.23 ± 0.44	-0.04 ± 0.20	-0.00 ± 0.16
IRS 17	-2.37 ± 0.46	0.95 ± 0.46	-0.32 ± 0.20	-0.23 ± 0.20
IRS 19NW	-1.35 ± 0.66	0.08 ± 0.78	0.32 ± 0.36	0.15 ± 0.36
IRS 28	0.21 ± 0.28	-0.03 ± 0.26	-0.10 ± 0.14	-0.07 ± 0.11
SiO 15	0.19 ± 0.58	-0.47 ± 0.62	0.09 ± 0.23	-0.07 ± 0.24
Average	-0.16 ± 0.17	0.08 ± 0.17	-0.04 ± 0.08	-0.01 ± 0.07

Table 2.5: Alignment of the SiO maser stars (positions are compared at epoch 2009.02).

have reused the Monte Carlo sample of the proper motion fits. The epoch 2009.02, at which positions are compared, is the average fit reference epoch, with the weights being the uncertainties of the absolute velocities. Because we have also used weighted alignment transformations to begin with, the absolute average values are expected to be non-zero.

For comparison, Gillessen et al. (2009b) were able to localize Sgr A* to within ∼ 1.80 mas and ∼ 0.33 mas yr^{-1} in their 'maser system', applying in particular a distortion correction for the S27 camera that only allowed for barrel or pincushion distortion, and a free optical axis (3 parameters, see Trippe et al., 2008). However, by using a more complex 'combined system' that involves a velocity calibration in the cluster rest frame (Sec. 2.1), the latter value reduces to ∼ 0.06 mas yr^{-1}. Yelda et al. (2010) were able to localize Sgr A* to within ∼ 0.57 mas and ∼ 0.09 mas yr^{-1} in their coordinate system, which is also solely based on observations of maser stars, and constructed from independent observations with the NIRC2 imager at the Keck observatory.

The largest deviations between infrared and radio measurements occur for the maser stars IRS 17, IRS 19NW and SiO 15. In the radio, IRS 17 has the least certain proper motion and was only observed at four of eight epochs, as was SiO 15. In the infrared, IRS 19NW is the faintest maser star, and its bright neighbor IRS 19 might bias its position. It is also the star farthest from the AO guide star IRS 7, with the next closer one being SiO 15.

The fact that the uncertainty of the infrared location of Sgr A* is well defined can be exploited in the analysis of stellar orbits, by setting priors on the allowed drift motion of the central mass[5]. The star S2 is currently the most important contributor to constraints on the gravitational potential and the distance to the Galactic Center. By following Gillessen et al. (2009b) in fitting a Keplerian orbit to the available astrometric and spectroscopic data of S2, and using our updated proper motion models for the astrometric reference stars, we were able to reduce this drift motion significantly (Fig. 2.7), without needing a cross-calibration to the cluster rest frame.

---

[5] Additional prior information at hand is a limit on the radial velocity of Sgr A*.

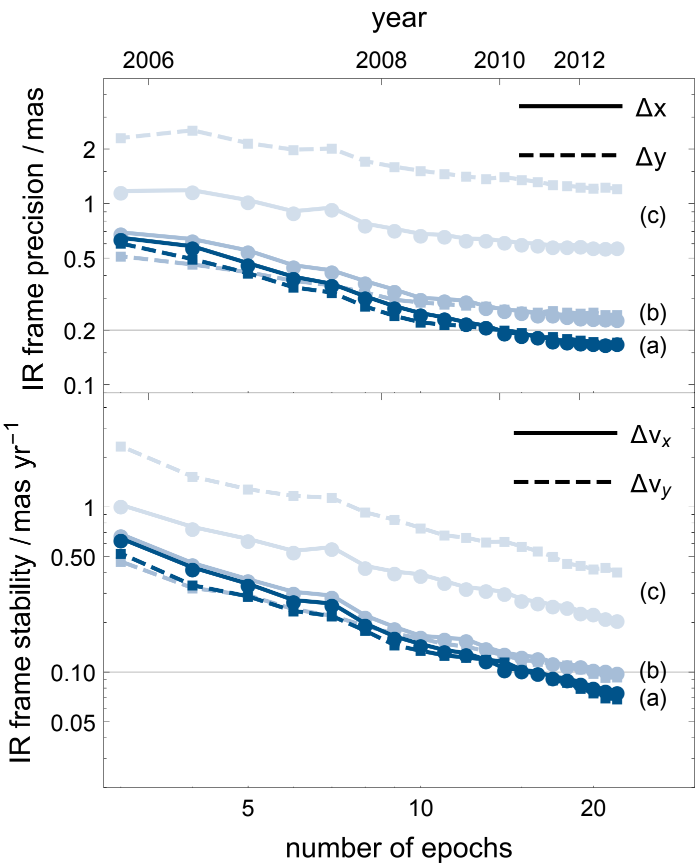

Figure 2.8: Precision and stability of the Sgr A* rest frame as a function of time, when excluding past infrared observations (a) and also either excluding radio observations after 2010 (b), or not applying a distortion correction (c).

2.5 Discussion & Conclusions

The radio source and massive black hole Sgr A* at the Galactic Center can now be placed in the origin of an infrared astrometric reference frame with an expected precision of ~ 0.17 mas in position (in 2009) and ~ 0.07 mas yr^{-1} in proper motion, by aligning the positions of SiO maser stars measured at both infrared and radio wavelengths.

Besides the new data collected in recent years, the factor 5 improvement over the comparable 'maser system' from Gillessen et al. (2009b) is mainly the result of a more accurate, 20-parameter correction for optical distortion in the S27 camera of NACO, which we have derived by comparing archive images of globular clusters to astrometric reference catalogs obtained with the HST. The shifts of stellar positions on the detector due to image distortion are on the level of a few tenths of a pixel (several mas), and the general distortion pattern is complex, but static. The origin of the observed minor, yet distinct and apparently abrupt changes is likely a difference in the optical alignment before and after interventions on the instrument. We make our four individual distortion corrections publicly available online, at `http://www.mpe.mpg.de/ir/gc/distortion`, for further use in other studies.

The precision and stability of the Sgr A* rest frame will continue to improve steadily, under the condition that future infrared and radio observations yield more positions of previously observed, or newly identified maser stars[6] (Fig. 2.8). A fundamental limit arises, should the maser emission not track the centroid of the photospheric emission, owing to the spatial distribution of maser spots in the extended stellar envelopes. Nevertheless, the already gained improvement of the infrared reference frame enables a reanalysis of the S-star astrometry, to be presented in a forthcoming paper (Gillessen et al., 2017).

Precision astrometry at the Galactic Center holds great scientific potential and might eventually lead to the detection of relativistic effects on stellar orbits. Given an orbit with eccentricity e around a black hole of mass $M_\bullet$ at a distance R_0, the dominant post-Newtonian effect with an impact on astrometry is Schwarzschild precession. This effect causes an apparent apocenter shift per revolution of (Weinberg, 1972):

$$\Delta s \approx \frac{6\pi G}{c^2}\frac{M_\bullet}{R_0(1-e)} \tag{2.3}$$

An extended mass distribution would cause a counteracting Newtonian precession, but the amount of extended mass is still uncertain (e.g. Genzel et al., 2010). In the case of the star S2, a drift of the coordinate system of ~ 0.05 mas yr^{-1} would amount to the magnitude of the former effect ($\Delta s \approx 0.8$mas) after one orbital period. The ability to detect a relativistic precession of the orbit of S2 therefore hinges on a very stable reference frame, but also depends crucially on the astrometric data obtained at and around the time of the next pericenter passage, in 2018.

Acknowledgements

We are grateful to Jay Anderson and Andrea Bellini for helpful discussions, and for contributing HST data that proved to be extremely valuable to our calibration efforts.

---

[6] In the near future, a further improvement of the astrometric reference frame for infrared observations of the Galactic Center might be gained by making use of precise Gaia measurements (Gaia Collaboration et al., 2016) of the proper motions of 50 additional astrometric reference stars, which are foreground stars found within $15-60''$ from Sgr A* in projection.

2.A A distortion correction for NACO (II)

Original publication: P. M. Plewa, S. Gillessen, M. Bauböck, J. Dexter, F. Eisenhauer, S. von Fellenberg, F. Gao, R. Genzel, M. Habibi, A. Jimenez-Rosales, T. Ott, O. Pfuhl, I. Waisberg, F. Widmann, 2018, RNAAS, *Optical distortion in the NACO imager*, doi:10.3847/2515-5172/aab3df

2.A.1 A distortion correction for the S13 camera

Here we present a set of distortion solutions (Fig. 2.9) that may be used to correct geometric optical distortion in images taken with the S13 (narrow-field) camera of the NACO AO imager (Rousset et al., 1998; Lenzen et al., 1998), complementing the distortion corrections we have previously derived for the S27 (wide-field) camera (Plewa et al., 2015). As a result of the S13 camera distortions, observed positions on the detector are offset from the true positions by about $1.2 - 2.3$ mas on average, and as much as ~ 7 mas in the corners of the detector, at a pixel scale of ~ 13 mas. The overall distortion pattern is complex but generally static, yet appears to have changed abruptly at least 9 times over the time that NACO has been in regular operation at the VLT, between 2002 and 2017.

2.A.2 Mapping the image distortion

We derive the new distortion corrections from near-infrared, high-angular resolution imaging observations of the Galactic Center, which continue to be routinely performed with NACO in the H- and Ks-bands, using both the S27 and the S13 camera, with the primary goal of monitoring stellar motions around the central massive black hole (e.g. Gillessen et al., 2009b, 2017). The existing set of distortion corrections for the S27 camera has allowed us to create a catalog of about 2000 astrometric reference stars in the inner nuclear star cluster, for which we can measure precise proper motions in an approximately distortion-free reference frame (see Plewa et al., 2015). For every suitable image taken with the S13 camera, we determine the optimal affine transformation (i.e. translation, rotation, scaling and shear operations; accounting for differential refraction, as well as other linear effects) to align the astrometric positions of these reference stars, at the respective times of observations, with their pixel positions as measured on the detector. The map of transformation residuals, i.e. the remaining difference between aligned positions, is a vector field that represents the higher-order distortions still present in each image. Due to the high stellar density in the Galactic Center, we can match as many as a few hundred stars per image, to create a combined distortion map for each observing night that is sampled at several thousand positions across the detector. A self-calibration approach (e.g. as used by Anderson & King, 2003) is not feasible, despite the high stellar density, because the requirement that many of the same stars are placed (at various orientations) in many different regions on the detector is almost never met, due to the specific dithering scheme employed for the past Galactic Center monitoring observations.

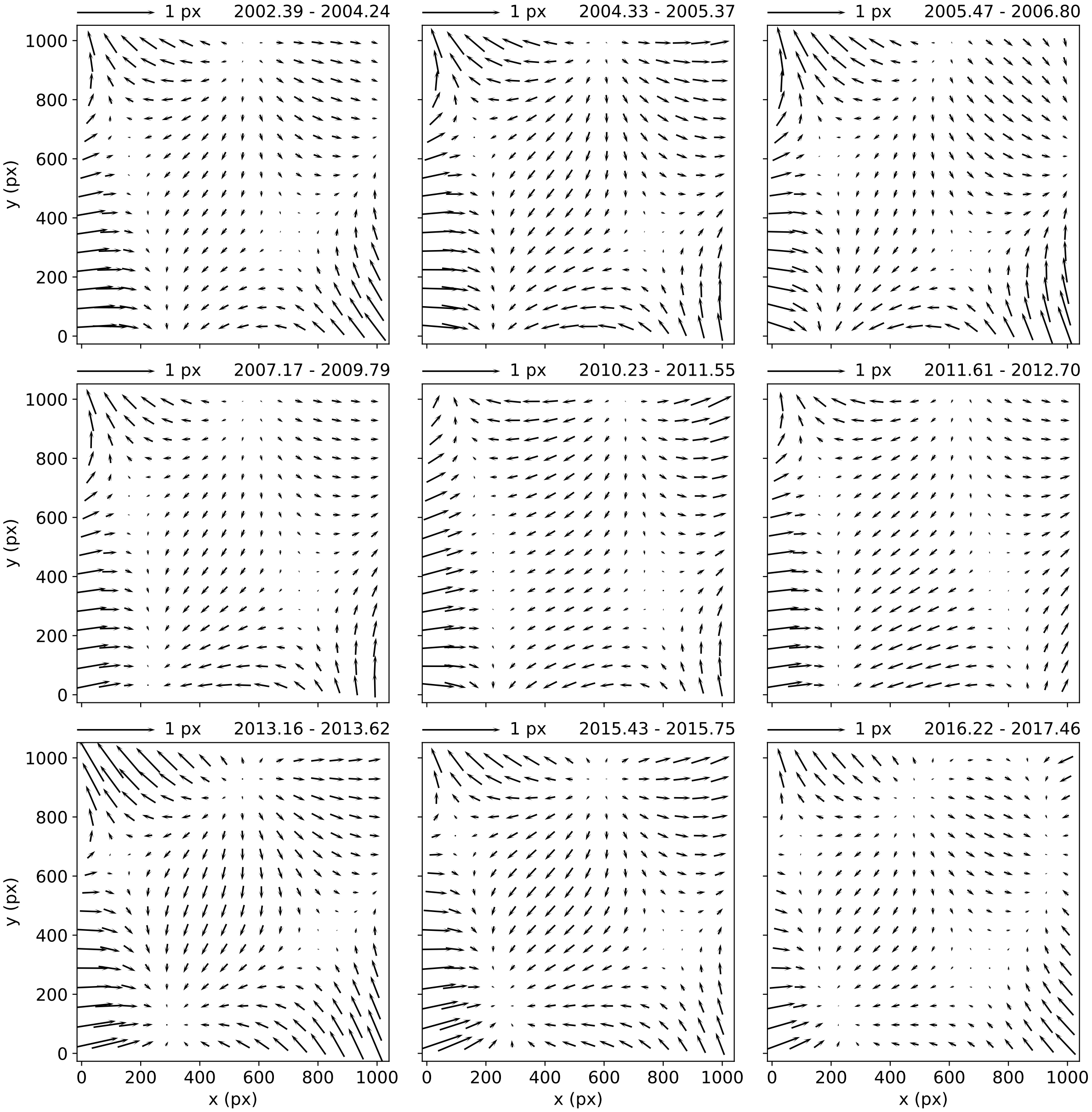

Figure 2.9: Distortion corrections for images taken with the S13 camera of NACO, for different periods between 2002 and 2017.

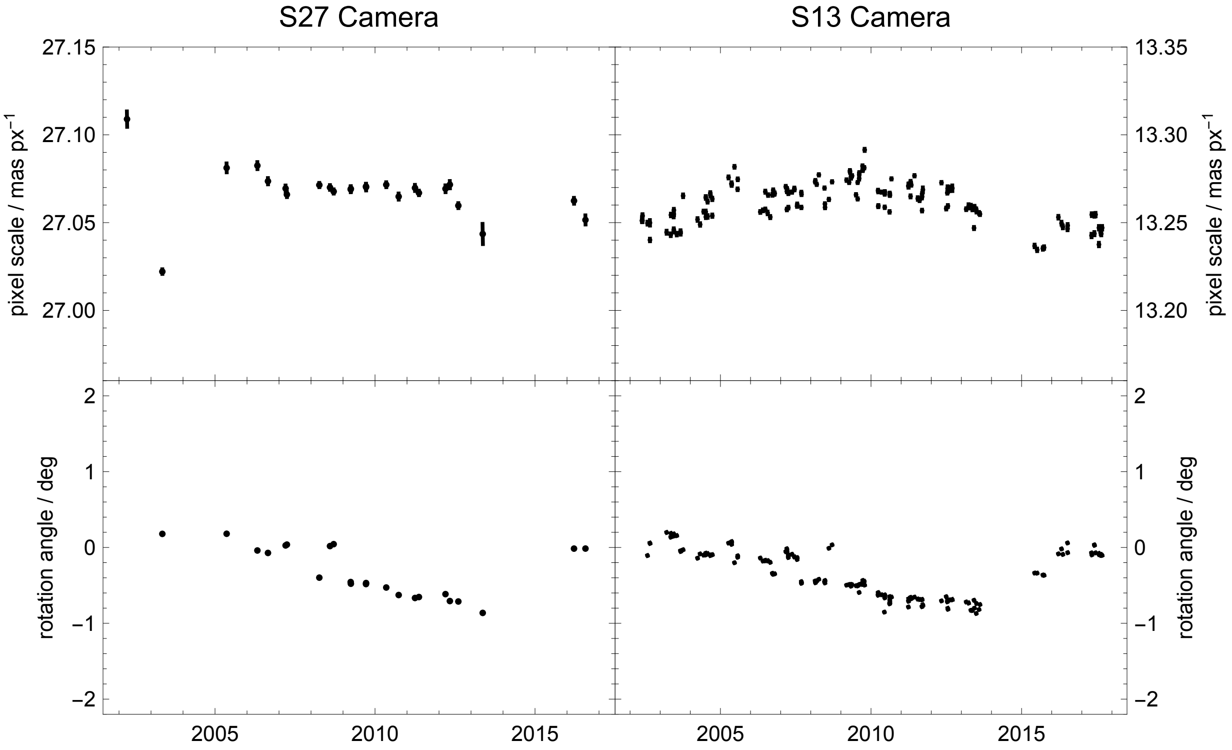

Figure 2.10: The evolution of the pixel scale and position angle zero point, for the S27 and S13 cameras of NACO, as derived from observations of the Galactic Center.

2.A.3 The distortion model

The distortion maps, after smoothing, are fit by a model that fully captures the spatial variability of the image distortion in 20 free parameters, and describes the distortion maps in terms of a basis set of orthonormal vector polynomials (Zhao & Burge, 2007, 2008). In addition to enabling the explicit correction of optical distortion in other NACO images, the fit results justify using cubic transformations to place stellar positions measured on widely dithered images of the Galactic Center into a common (relative) astrometric reference frame, which has been the standard practice to account for unknown image distortions in several previous studies (e.g. Gillessen et al., 2009b, 2017; Fritz et al., 2010b). By a visual inspection assisted by clustering algorithms (e.g. *k*-means), we can identify several abrupt changes in the image distortion that lead us to create 9 distinct, averaged distortion solutions, each of which is applicable for periods of time ranging from 4 to 31 months, over which the image distortion appears to be static, to within 10%. There are no indications for additional systematic changes in these solutions, for instance due to the different filters used. All solutions can be found online, at `http://www.mpe.mpg.de/ir/gc/distortion`, in the form of FITS files.

2.B Pixel scale and rotation angle

From the astrometric transformations that map pixel coordinates into the Sgr A* rest frame (Sec. 2.4), we can measure both the pixel scale of our Galactic Center images and the absolute orientation of the field of view (Fig. 2.10). For images taken with the S27 camera, these values can be derived directly from the astrometry of the SiO maser stars, whereas one additional step of cross-calibration is typically necessary to be able to characterize the S13 camera (see Sec. 2.4.3). The mean pixel scales, since the installation of the Aladdin 3 detector in 2004, are:

$$\langle s \rangle_{\mathrm{S27}} = 27.068 \pm 0.008 \tag{2.4}$$

$$\langle s \rangle_{\mathrm{S13}} = 13.262 \pm 0.012 \tag{2.5}$$

Because the measured values appear to vary with time, if only slightly, we give as an estimate of the uncertainty the observed scatter around the mean values, which is somewhat larger than the average statistical uncertainty of the individual measurements. An additional systematic uncertainty presumably exists on the order of at least 0.01%, considering, for example, the change of scale caused by differential aberration resulting from Earth's motion, or small changes in the telescope focus. To within these uncertainties, our values broadly agree with independent scale measurements based on Pluto's motion against field stars (see the NACO user manual, VLT-MAN-ESO-14200-2761, Issue 101), as well as those based on HST observations of stellar binaries (e.g. Chauvin et al., 2006, 2015). However, we find a slightly larger scale for the S13 camera in comparison (but see Masciadri et al., 2003), which could be due to crowding-related difficulties in determining accurate proper motions for some of the 91 secondary astrometric reference stars, from images taken with the S27 camera. The zero point of the position angle, i.e. the default orientation of the field of view with respect to the direction of north, has been drifting over several years by about 1°, yet since the move of NACO from UT4 to UT1 and another change of detector in 2014 (to an Aladdin 2 detector previously used in the ISAAC instrument) the calibration of the position angle appears to be accurate again. The pixel scales have apparently not changed notably following these latest changes.

2.C Supplementary figures

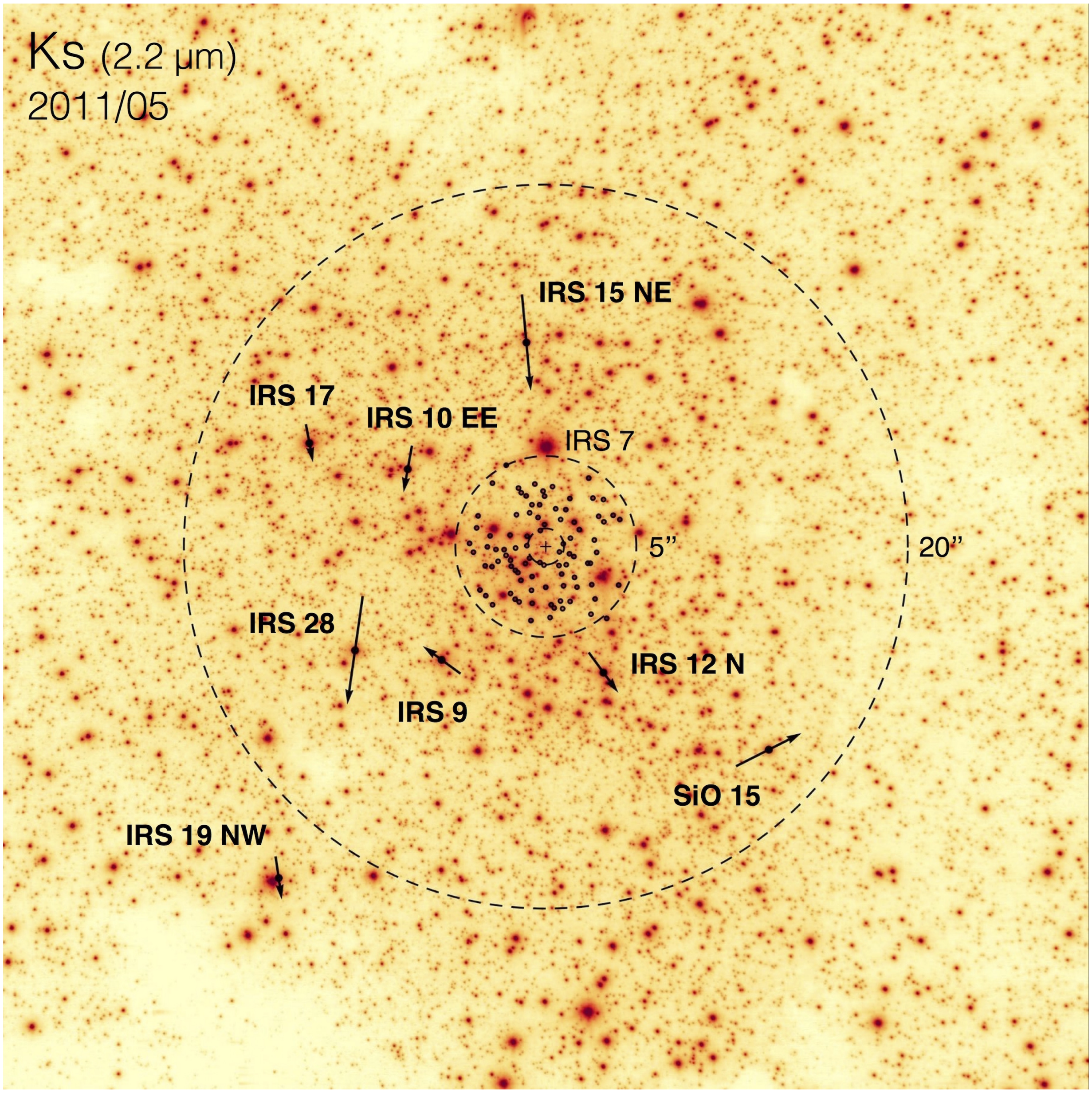

Figure 2.11: A near-infrared Ks-band image of the larger nuclear star cluster surrounding Sgr A* (black cross), on which are marked the proper motions of the SiO maser stars (black arrows), as well as the positions of the 91 secondary astrometric reference stars (small black circles).

Chapter 3

The post-pericenter evolution of the Galactic Center source G2

Original publication: P. M. Plewa, S. Gillessen, O. Pfuhl, F. Eisenhauer, R. Genzel, A. Burkert, J.Dexter, M. Habibi, E. George, T. Ott, I. Waisberg, S. von Fellenberg, 2017, ApJ, *The Post-Pericenter Evolution of the Galactic Center Source G2*, doi:10.3847/1538-4357/aa6e00

Abstract: In early 2014, the fast-moving near-infrared source G2 reached its closest approach to the massive black hole Sgr A* in the Galactic Center. We report on the evolution of the ionized gaseous component and the dusty component of G2 immediately after this event, revealed by new observations obtained in 2015 and 2016 with the SINFONI integral field spectrograph and the NACO imager, at the ESO VLT. The spatially resolved dynamics of the Brγ line emission can be accounted for by the ballistic motion and tidal shearing of a test-particle cloud that has followed a highly eccentric Keplerian orbit around the black hole for the last 12 years. The non-detection of a drag force, or any strong hydrodynamic interaction with the hot gas in the inner accretion zone, limits the ambient density to less than a few 10^3 cm^{-3} at the distance of closest approach (1500 R_s), assuming G2 is a spherical cloud moving through a stationary and homogeneous atmosphere. The dust continuum emission is unresolved in L'-band, but its location has stayed consistent with that of the Brγ emission, while the total luminosity of the Brγ- and L'-band emission has remained constant to within the measurement uncertainty. The nature and origin of G2 are likely related to that of the precursor source G1, the orbital evolution of which is similar to G2's, though not identical. Both object are also likely related to a trailing tail structure, which is continuously connected to G2 over a large range in position and radial velocity.

3.1 Introduction

The development of high-resolution near-infrared observing techniques has enabled ever more detailed studies of the Galactic Center black hole, identified with the compact radio source Sgr A*, and its immediate environment. Almost all of the detectable infrared sources within this environment are stars. However, a few sources show both continuum emission in L-band and recombination emission in hydrogen and helium lines, but are missing a counterpart source in K-band (Clenet et al., 2005; Ghez et al., 2005; Eckart et al., 2013; Sitarski et al., 2014). These sources do not resemble the surrounding stars (see sup. Fig. 3.12), but appear to be concentrations of gas and dust that are being ionized by the radiation of the nearby massive young stars. One of these objects, called G2, is of special interest, due to its proximity to the central massive black hole. G2 was discovered in 2011 (Gillessen et al., 2012) and was found to be on a highly elliptical orbit, which would bring it to a point of closest approach to Sgr A* in early 2014 (Phifer et al., 2013). We have been able to follow the orbital evolution of G2 closely by means of imaging spectroscopy, conducting regular monitoring observations and using archive data from as early as 2004 (Gillessen et al., 2013a,b; Pfuhl et al., 2015). Most notably, in position-velocity (pv) space, G2 became ever more elongated along its orbit as it moved toward Sgr A*, and its radial velocity started to change from being strongly red-shifted to blue-shifted in 2013. The next year's observations captured G2 in the very middle of its extended pericenter passage.

During this time, G2 has brought a substantial amount of mass to a distance of just a few thousand times the Schwarzschild radius of Sgr A* (R_S). The estimated gas mass of G2 is 3 $M_\oplus$ (Gillessen et al., 2012; Shcherbakov, 2014), which is comparable to the mass of gas in the hot accretion flow at this distance (Yuan et al., 2003; Xu et al., 2006). G2 therefore represents a unique probe of the inner accretion zone around the black hole. A possible sign of an interaction between G2 and the ambient gas could have been the detection of X-ray or radio emission originating from a shock front (Gillessen et al., 2012; Narayan et al., 2012; Sądowski et al., 2013b). However, the level of observable emission depends sensitively on the poorly known physical parameters of G2 and the accretion flow, for example the densities and volume filling factors. It is therefore not surprising that no such emission has been detected (e.g. Bower et al., 2015). Any direct accretion of material onto Sgr A* would set in only years after the time of pericenter, as earlier simulations have shown (e.g. Schartmann et al., 2012).

The astrophysical nature of G2 is still debated. Many models and formation scenarios have been proposed, including a clumpy stream of gas that formed from a collision of stellar winds (Cuadra et al., 2006; Burkert et al., 2012; Gillessen et al., 2012; Schartmann et al., 2012, 2015; Calderón et al., 2016) or from the partial tidal disruption of a star (Guillochon et al., 2014), different kinds of evaporating circumstellar disks (Murray-Clay & Loeb, 2012; Miralda-Escudé, 2012) or (proto-)planets (Mapelli & Ripamonti, 2015), different kinds of enshrouded stars (Meyer & Meyer-Hofmeister, 2012; Scoville & Burkert, 2013; Ballone et al., 2013; De Colle et al., 2014; Valencia-S et al., 2015; Ballone et al., 2016) and a binary star merger product (Witzel et al., 2014; Prodan et al., 2015). The data so far only provides direct evidence for an extended cloud of gas and associated dust (Sec. 3.3). A central star embedded within such a dusty gas cloud might stay undetected, if it is sufficiently small, and either very hot or very cool, but the presence of such a compact central source would be irrelevant for the dynamics of the observed gas. The gas dynamics at pericenter are dominated by the gravitational force of the black hole, and possibly affected by the ambient medium.

In this paper, we report on new near-infrared observations of both the ionized gaseous component and the dusty component of G2, obtained in 2015 and 2016 with the SINFONI integral field spectrograph and the NACO imager at the ESO VLT. First, we summarize the observations and describe the data reduction and processing procedures, as well as the modeling techniques (Sec. 3.2). We then analyze the post-pericenter evolution of these two components in the context of our previous observations, which allows us to draw the most comprehensive picture of G2 to date (Sec. 3.3). We also determine a revised orbit solution by using an improved numerical approach. Finally, we discuss how observations of G2 can constrain properties of the accretion flow, how it is related to the precursor source G1 and the trailing tail emission, as well the implications for its nature, origin and fate (Sec. 3.4).

3.2 Observations & Methods

We have obtained new, deep near-infrared observations of the Galactic Center at the VLT with SINFONI (Eisenhauer et al., 2003b,c; Bonnet et al., 2004) in April/May 2015 and April/July 2016, as well as with NACO (Lenzen et al., 1998; Rousset et al., 1998) in July/September 2015 and April 2016.

3.2.1 SINFONI data reduction

As previously (e.g. Pfuhl et al., 2015), we have used the narrow-field camera (25 mas px^{-1}) of SINFONI with the H+K grating and the AO system in natural guide star mode, or, if available, the laser guide star, in combination with an object-sky-object observing pattern, and a detector integration time of 600 s. By using a quadratic dither pattern, the total field of view is extended from originally $0.8'' \times 0.8''$ to $1.2'' \times 1.2''$. The central region containing Sgr A* and G2 is still covered at every dither position, and ultimately sampled at 12.5 mas px^{-1}, with a spectral resolution of about 1500.

We use the data reduction pipeline *SPRED* (Abuter et al., 2006) to perform flat-fielding and sky subtraction, wavelength and telluric calibration, as well as a bad pixel and distortion correction. We also refine the default wavelength calibration based on the observed atmospheric OH lines, and we select only the high-quality data for which the FWHM of the star S2 is smaller than 7.5 px at a wavelength of 2.2 μm. The total integration time spent on source in 2015 is 810 minutes, out of which 600 minutes pass this quality cut. In 2016 the total is 520 minutes, out of which we are able to use 400 minutes. Due to poor weather conditions, these exposures are not as deep as those obtained in previous years. In a final step, we reconstruct a single combined data cube for each year, from which it is possible to extract channel maps, regionalized spectra and pv-diagrams to isolate the recombination line emission from G2. The orbital evolution of G2 is slow enough so that combining observations from a few consecutive months does not skew the data, but on the contrary maximizes the signal-to-noise ratio. Adding the two data cubes from 2015 and 2016 to our existing data set increases the total number of epochs to 10 since 2004.

For each data cube, we manually choose an extraction map that defines source and background pixels. The line emission from G2 has never been strongly confused with other sources (due to G2's trajectory and in particular its large radial velocity), but to attain a clean signal it is still necessary to choose the source- and background regions with care, and to remove atmospheric and stellar emission. Atmospheric features make individual spectral channels appear brighter or fainter, whereas stellar (continuum) emission leads to brightness variations across all spectral channels at a given position. Subtracting the appropriate median value at each position, and then subtracting the average background in each channel, removes these artifacts reliably. Even so, the choice of source and background pixels remains the most likely source of a systematic uncertainty. For the astrometric and photometric calibration, we use a changing set of relatively isolated stars, the motions and (K-band) magnitudes of which are well know from the long-time monitoring of stellar motions in the Galactic Center (Gillessen et al., 2009b).

If taking a certain on-sky trajectory as given is justified (see e.g. sup. Fig 3.13), the complexity of the spectro-spatial information in a three-dimensional data cube can be reduced in an optimal way into a two-dimensional pv-diagram, to visualize the low surface-brightness emission close to the detection limit. In practice, we construct pv-diagrams of G2 by placing virtual slits (8 px wide) along the trajectory of the best-fit orbit, and aligning the resulting spectra as a function of distance along this path. To increase the signal to noise ratio further, we create noise-weighted, coadded diagrams using both the Brγ and HeI lines.

3.2.2 NACO image processing

Because of a malfunction of NACO's camera wheel in 2015, we have used the L' filter ($\lambda = 3.80 \pm 0.31$ μm) in combination with the S13 camera, which according to specification only supports a wavelength range of $1.0 - 2.5$ μm. A longer detector integration time and total exposure time than usual were necessary to compensate for the reduced sensitivity, but the smaller pixel scale (13 mas px^{-1}) and the low level of image distortion of the S13 camera were advantageous, in comparison to the normally used L27 camera. In 2016, we were able to use the standard configuration again.

Because of another issue with the readout electronics, every eighth column of one detector quadrant showed no signal. This problem is easily mitigated by replacing the missing values in every frame with interpolated values based on the neighboring detector columns, since the PSF is well sampled. After this correction, we apply standard flat-field and bad pixel corrections. A sky image is constructed individually for each object image, by calculating the median image of several other randomly offset images taken close in time. Any images of exceptionally bad quality are rejected, for example those affected by a thin cloud cover. We finally use the position of a common bright star as a reference position, to combine the remaining images into a mosaic image. Despite the different instrumental setups, the mosaic images from 2015 and 2016 reach a similar depth in the central region around Sgr A*.

We extract an empirical PSF from the mosaic images using the tool StarFinder (Diolaiti et al., 2000), and use this PSF in a Lucy-Richardson algorithm to deconvolve the images to a moderate depth. From the resulting images, we measure the positions and instrumental magnitudes of several stars that are suitable as astrometric and photometric reference sources (Gillessen et al., 2009b). We are thus able to subtract from the images the bright B-stars that are at the times of the observations confused with G2 in L-band (in particular S2, S19, and S31), assuming a constant K-L color for all stars of this common spectral type. Apart from Sgr A*, there are no unaccounted sources in the central region brighter than about 17th to 18th magnitude in K-band.

The subtraction of the infrared source Sgr A* is not as straightforward, because it is intrinsically variable on a time scale of tens of minutes, during flaring events that occur sporadically. However, only the 2015 data shows significant emission at the position of Sgr A*, and at this time G2 is already spatially distinguishable from Sgr A* due to its orbital motion, even though the images of the two sources still partly overlap in projection. This allows us to measure the L-band light curve of Sgr A* by performing small-aperture photometry on the series of single images, and measure accurately the average brightness that we have to subtract.

3.2.3 Orbit fitting

We let G2 be represented by an ensemble of independent massless particles to fit the orbital evolution of the Brγ-emitting gas cloud over a time span of 12 years, from well before to right after the pericenter passage. Such a test-particle cloud has already proven to be a good description of past observations (Gillessen et al., 2012, 2013a,b; Pfuhl et al., 2015) and has several benefits as a model. A test-particle simulation is far less computationally demanding than a hydrodynamic simulation, for instance, but still offers a way to account for other external forces in addition to gravity. Even more importantly, it allows a direct comparison to the observations in the (position-velocity) space of the data.

A possible interaction with the ambient gas in the accretion zone is accounted for by including a distance- and velocity-dependent drag force on the particles in the simulation, in the form of a ram pressure:

$$\boldsymbol{a}_D = -c_D \left(\frac{r}{\mathrm{as}}\right)^{\alpha} \left(\frac{v}{\mathrm{as/yr}}\right)^2 \frac{\boldsymbol{v}}{v} \frac{\mathrm{as}}{\mathrm{yr}^2} \tag{3.1}$$

The radial dependency is characterized by the exponent α, which has a value of $\alpha = -1.5$ in a spherical Bondi accretion model (Yuan et al., 2003), but has been measured from X-ray observations to be somewhat closer to $\alpha = -0.5$ in the Galactic Center (Wang et al., 2013). However, G2 probes the accretion zone at much smaller distances than the X-ray observations can, and we therefore consider both extreme cases. The free parameter c_D determines the strength of the drag force, and is assumed to be a constant non-negative coefficient. In the absence of a drag force ($c_D = 0$), a particle in the point mass potential generated by the central black hole moves in a fixed orbital plane, and on an ellipse if it is bound. Such a Keplerian orbit is fully described by six parameters: the semi-major axis a and the eccentricity e, which describe the shape of the orbit, the three angles inclination i, longitude of ascending node Ω and argument of pericenter ω, which describe the orientation of the orbital plane, as well as the time of pericenter t_p. In the presence of a drag force ($c_D > 0$), the orientation of the orbital plane is unchanged, but the particles experience an instantaneous deceleration oriented against their direction of motion, and deviate from elliptical orbits. Because the trajectory of G2 is almost radial, and the ambient density increases strongly toward the center, the effect of the drag force is concentrated around the time of pericenter.

We use the *MultiNest*[1] implementation (Feroz et al., 2009) of the nested sampling algorithm (Skilling, 2004) to calculate the marginalized likelihood[2] to be able to compare different models, and at the same time sample from the posterior probability distribution of the model parameters. To evaluate the likelihood once for a certain set of orbital parameters, we run a complete simulation starting at epoch t_0 (well before the first observations), using 10^3 independent test particles that are initially placed on the given orbit and then perturbed randomly in position, where the offsets are sampled randomly from a spherically symmetric normal distribution. Both the starting time and the size of the initial particle distribution are included as two additional parameters in the model. The number of test particles should be chosen as large as possible to limit the noisiness of the likelihood, which is introduced by the inherent randomness of the initial conditions, and diminishes the efficiency of the sampling process[3].

The particle orbits are numerically integrated using the *REBOUND* multi-purpose n-body code (Rein & Liu, 2012), and specifically the *whfast* integrator (Rein & Tamayo, 2015) and the high-accuracy *ias15* integrator (Rein & Spiegel, 2015). The former algorithm is optimized for Keplerian orbits, whereas the latter algorithm adaptively reduces the time step of the simulation to resolve close encounters of particles with the black hole, during which they can experience very strong accelerations. We assume a black hole mass of $4.31 \times 10^6\ M_\odot$, a distance to the Galactic Center of 8.33 kpc and uninformative priors for all sampled parameters, which are chosen either uniform or uniform in log-space for scaling parameters.

Upon reaching each epoch of observation as the orbit integration proceeds, the test-particle cloud is convolved with a Gaussian model PSF to match the resolution of the observations (approx. 80 mas and 120 km s^{-1} FWHM along the spatial and spectral dimensions, respectively), and then compared directly to the relevant volume of the data

---

[1] `https://ccpforge.cse.rl.ac.uk/gf/project/multinest/`

[2] The (fully) marginalized likelihood of a model is the expected value of the likelihood over the prior volume. It is also commonly known as the Bayesian evidence (MacKay, 2003).

[3] In effect, the exact initial conditions appear as additional (hyper-)parameters to the sampler, which have to be marginalized over.

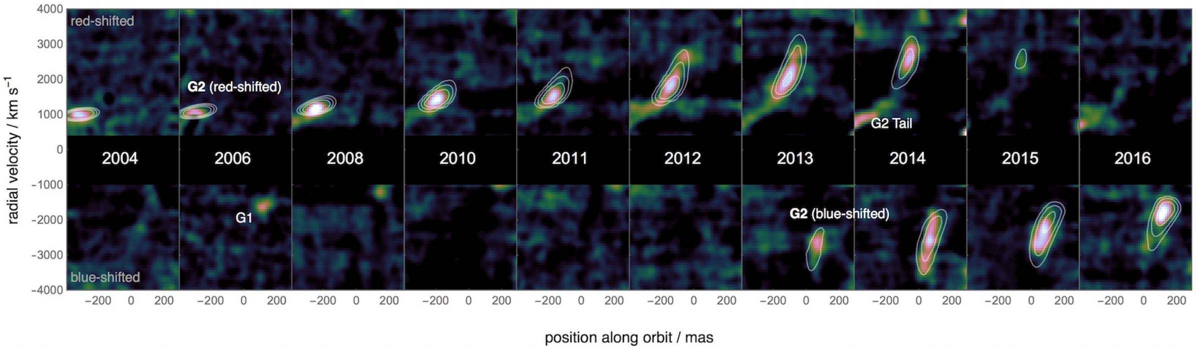

Figure 3.1: Position-velocity diagrams of the coadded Brγ and HeI line emission from G2 between 2004 and 2016 (shown in color, noise-normalized), extracted along the updated orbit and superimposed with a projection of the best-fit test-particle cloud (shown in contours). The region around zero velocity is masked because it is affected by background emission and stellar absorption features. The placement of the slit is optimized for G2 but also partially covers G1, which appears as a blue-shifted source at early times, and the tail of G2, which appears close to zero velocity at later times (but see fig. 3.9). In these diagrams, a velocity gradient appears as a tilt (see also fig. 3.4).

cube around the Brγ line, which is stored as a background-subtracted table of source pixels, assuming zero-mean constant Gaussian noise to define the likelihood function (i.e. a background-dominated regime).

A major advantage of this specific fitting technique is that the entire available spectro-spatial information is made use of, which, for example, includes velocity gradients that manifest themselves as changes of radial velocity across the cloud. In this way, projection effects are naturally accounted for as well, which can have a strong impact on the appearance of G2 in the observations, depending on the orientation of the orbital plane. For instance, the peak or centroid of the cloud emission does not in general correspond to that of the center of mass if the cloud is extended, along any dimension of the data, as might be implicitly assumed. This effect could cause a bias when only these data points were used to fit the orbit and not the data cubes.

3.3 Results

3.3.1 The gas emission

The new observations from 2015 and 2016 reveal the continued evolution of G2 immediately after its closest approach to Sgr A*. Figure 3.1 shows how the complete pericenter passage has unfolded over the last 12 years, in a time series of 10 pv-diagrams (Sec 3.2.1), which we have extracted along the updated best-fit orbit (Table 3.1). Figures 3.2, 3.3 and 3.4 show corresponding sets of integrated maps, velocity distributions and velocity maps of the gas

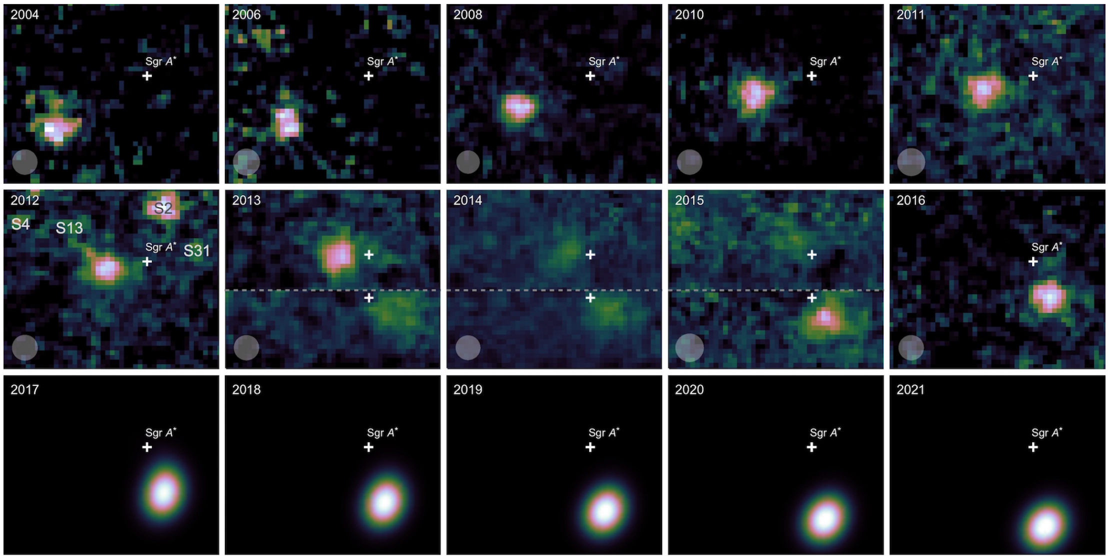

Figure 3.2: Top rows: Velocity-integrated maps of the Brγ line emission from G2 between 2004 and 2016 (shown in color, flux-normalized). In 2013, 2014, and 2015 the blue- and red-shifted components are extracted from different regions and shown separately. The dimensions of each image are $0.6''$ by $0.5''$ (north is up and east is left), and the approximate size of the PSF (FWHM) is indicated by the opaque circle. Bottom row: Predicted appearance of G2 over the next five years according to the test-particle simulation.

emission, which are complementary visualizations of the same underlying three-dimensional data cubes.

The gas emission has changed from being still partially red-shifted in 2015 to being entirely blue-shifted by 2016. This confirms that, as predicted, the bulk of cloud material has passed the black hole in 2014 at high velocity, assuming the emission is a valid proxy for gas mass throughout the evolution. The blue-shifted Brγ line is detected at a radial velocity of -2400 km s^{-1} in 2015 and redetected at -1700 km s^{-1} in 2016, with a respective line width of 320 km s^{-1} and 190 km s^{-1}.

The total Brγ luminosity of G2 continues to be comparable to previous measurements, while the ratio of blue- to red-shifted emission has changed rapidly over just a few years (Fig. 3.5). The fraction of blue-shifted emission increased from 30% to 70% between 2013 and 2015, before reaching 100% by 2016. During this time, the total luminosity has remained constant, and could have at most weakly increased during pericenter.

The leading gas in 2015 must be spatially extended, as it continues to show a pronounced velocity gradient due to tidal shear that is spatially well-resolved (Fig. 3.1 & 3.4). This extent might not be obvious from just a map of the integrated line emission (Fig. 3.2), since the motion of G2 is mostly seen in radial velocity and only one other on-sky coordinate, due to the almost edge-on orientation of G2's orbit. A velocity gradient on the order of 10 km s^{-1} mas^{-1} is still apparent in 2016, even though the gas cloud now appears more compact again in velocity space (Fig. 3.1). This is expected if the cloud is purely interacting

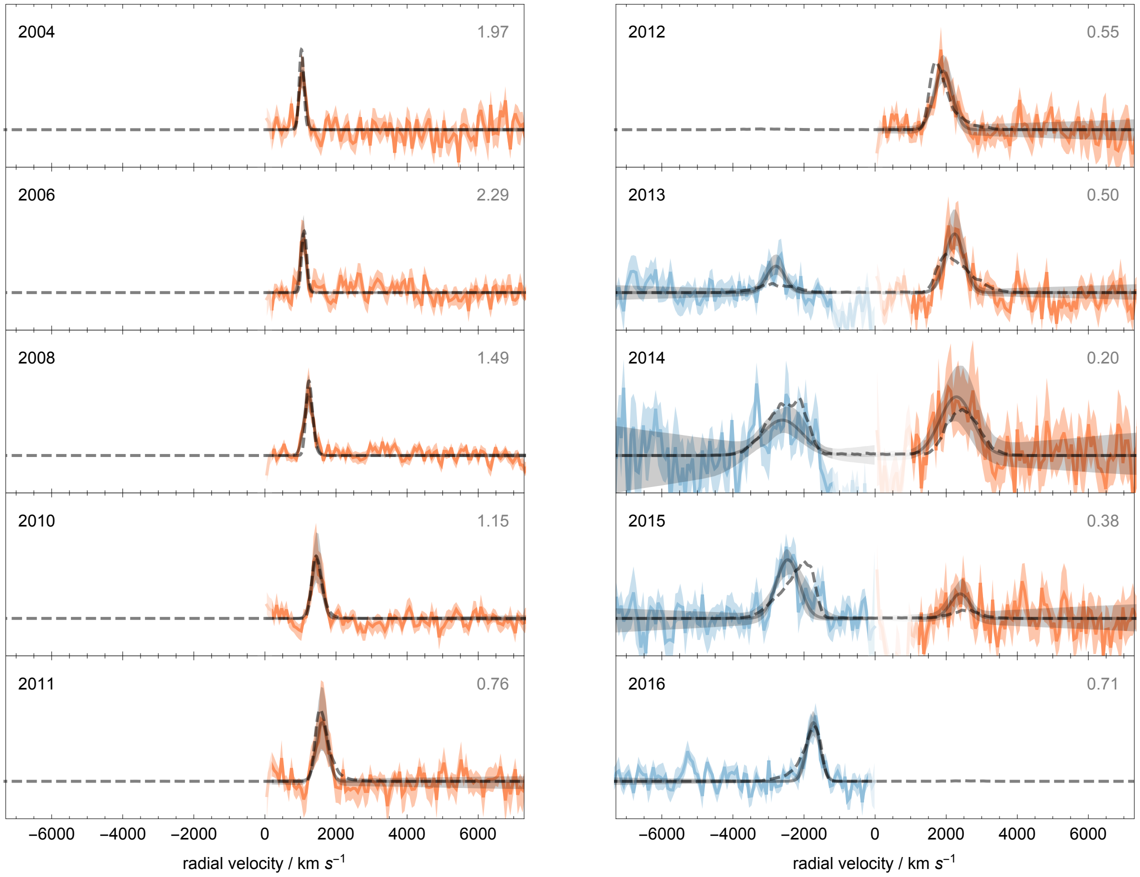

Figure 3.3: Radial velocity distribution of the Brγ line emission from G2 between 2004 and 2016 (colored lines, peak-normalized), superimposed with that of the best-fit test-particle cloud (dashed line, scaled to the mean total luminosity). In 2014 and 2015, the blue- and red-shifted components are extracted from different regions but shown together in this graphic. Also shown is the fit of a Gaussian profile to the Brγ lines (solid black lines with error bands). The region around zero velocity is excluded from this fit, because it is affected by background emission and stellar absorption features. The relative peak value from the fit is indicated by the number in the top right corner.

gravitationally with the black hole, and hardly at all with the ambient gas. Over the next few years the appearance of the cloud will not markedly change, assuming the test-particle model stays valid (Fig. 3.2, bottom row).

The post-pericenter Brγ data (Sec. 3.2.1) and the improved methods of analysis (Sec. 3.2.3) provide a new set of orbital parameters that typically have half of their previous statistical uncertainty or less. The updated best-fit parameters broadly agree with those found by Pfuhl et al. (2015) to within these uncertainties, except that we find a different orientation of the orbital plane ($\Delta\Omega \approx 20°$). The estimates of all orbital parameters are summarized in table 3.1, and their joint posterior distribution is presented in more detail in supple-

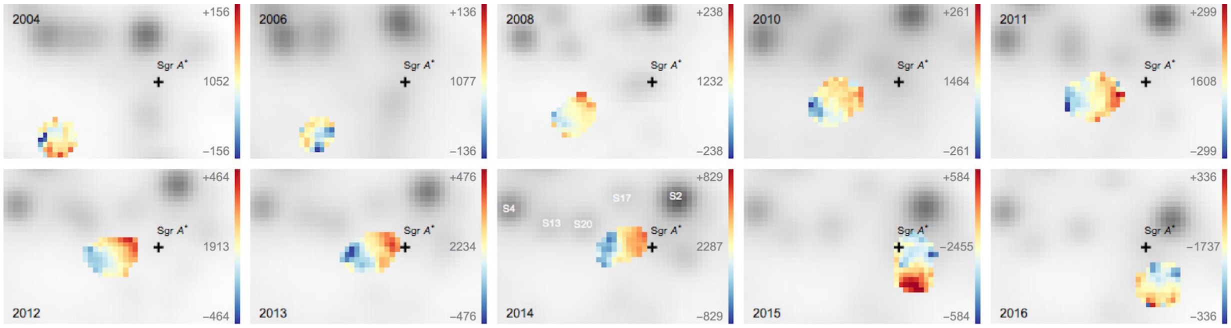

Figure 3.4: Radial velocity maps of the Brγ line emission from G2 between 2004 and 2016, where a change in color indicates a velocity gradient across the gas cloud. The color scale in each map is adjusted according to the measured bulk velocity and line width (Fig. 3.3). A spatially resolved velocity gradient is most clearly detected after 2010. This is also evident from the pv-diagrams (Fig. 3.1), but can be seen in this visualization without making any assumptions about the cloud trajectory. For reference, a K-band continuum image is shown in the background, on which nearby stars are marked. The dimensions of each image are $0.6''$ by $0.4''$ (north is up and east is left).

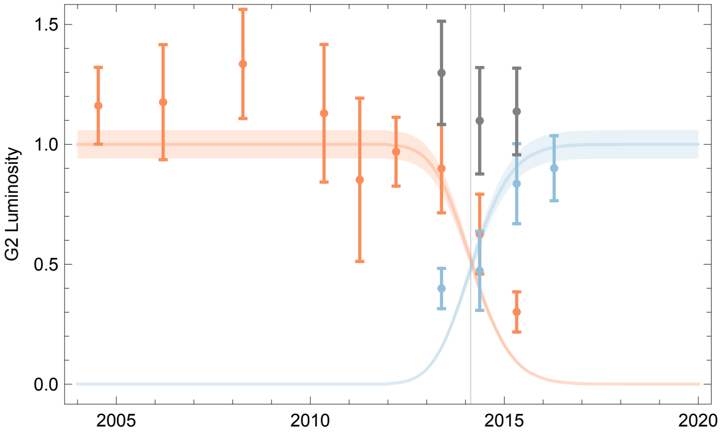

Figure 3.5: The luminosity of the Brγ line emission from G2 as a function of time (colored points). Also shown is the relative amount of red- and blue-shifted material, based on the best-fit test-particle simulation (colored lines, scaled to the mean total luminosity). If red- and blue-shifted components are detected simultaneously, the total luminosity is shown as an additional data point (in black). A strict comparison of the data points to the curves is only valid if the observed emission is a direct proxy for mass throughout the evolution. The closest approach of the center of mass, by definition at the time of pericenter (vertical line), happens shortly before half of the mass has changed from being red- to blue-shifted. The entire pericenter passage is a prolonged process stretching over several months from 2013 to 2016. The error bars stem from resampling the selected source and background pixels, and the error bands reflect the uncertainty of the mean luminosity.

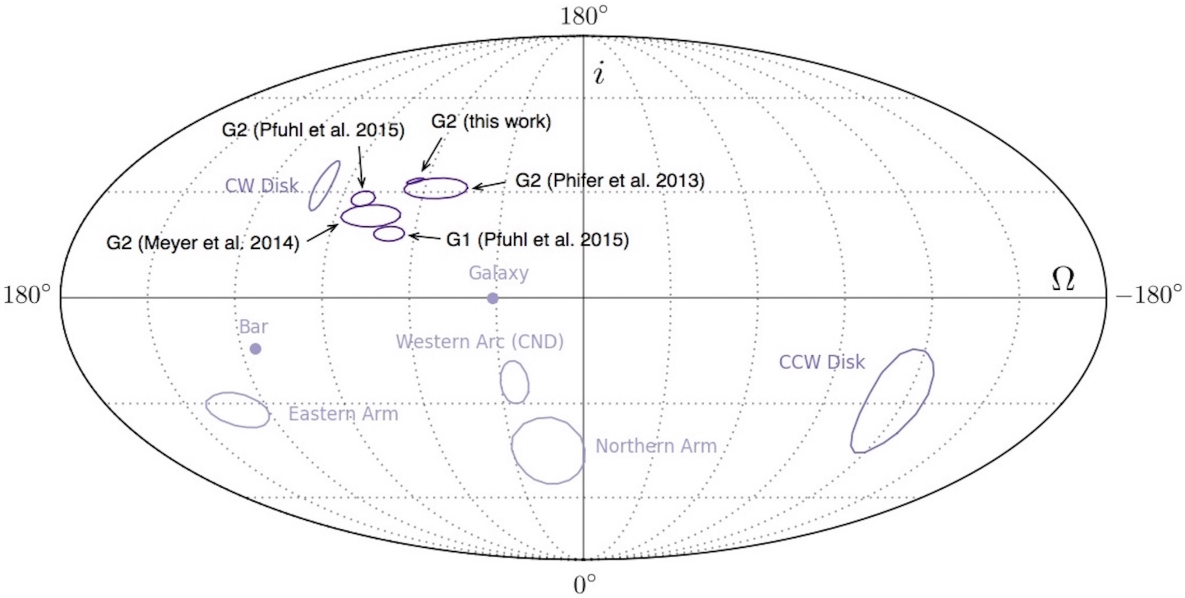

Figure 3.6: Estimates of the orientation of G2's orbital angular momentum vector in comparison to other orbit solutions from the literature (Phifer et al., 2013; Meyer et al., 2014; Pfuhl et al., 2015), and the orientations of other planar structures in the Galactic Center environment (Genzel et al., 2010, tab. 1 and references therein).

mentary Fig. 3.11. The angular offset between the orientation of G2's orbital plane and that of the clockwise stellar disk amounts to 30° (Fig. 3.6), while the apocenter of G2 at about $1''$ projected distance from Sgr A* still falls within the inner disk region (Paumard et al., 2006; Bartko et al., 2009; Yelda et al., 2014). The main characteristic of the orbit is unchanged, which is the almost radial trajectory. The best-fit pericenter distance is 18.4 ± 0.9 light hours, or $1560 \pm 70\ R_S$. In terms of a marginalized likelihood ratio (i.e. the Bayesian evidence ratio), none of the models that we tested in which G2 is affected by an additional drag force (Sec. 3.2.3) is significantly favored over the Keplerian model.

However, none of the fits can accurately describe the pericenter observations in detail (see Fig. 3.1 & 3.3). The gas seems to overshoot at the beginning of the pericenter passage, when in 2013 a substantial fraction of the emission already appears on the blue-shifted side, while as much gas should not have had time to arrive there. The gas also lags behind at the end, when in 2015 a substantial fraction of emission stays present on the red-shifted side, while the gas should have almost finished passing to the other side.

3.3.2 The dust emission

In the cleaned L-band images, from which confusing sources have been removed (Sec. 3.2.2), we clearly detect a residual source in both 2015 and 2016 that we attribute to thermal continuum emission from dust associated with G2 (Fig. 3.7). The location of this source is

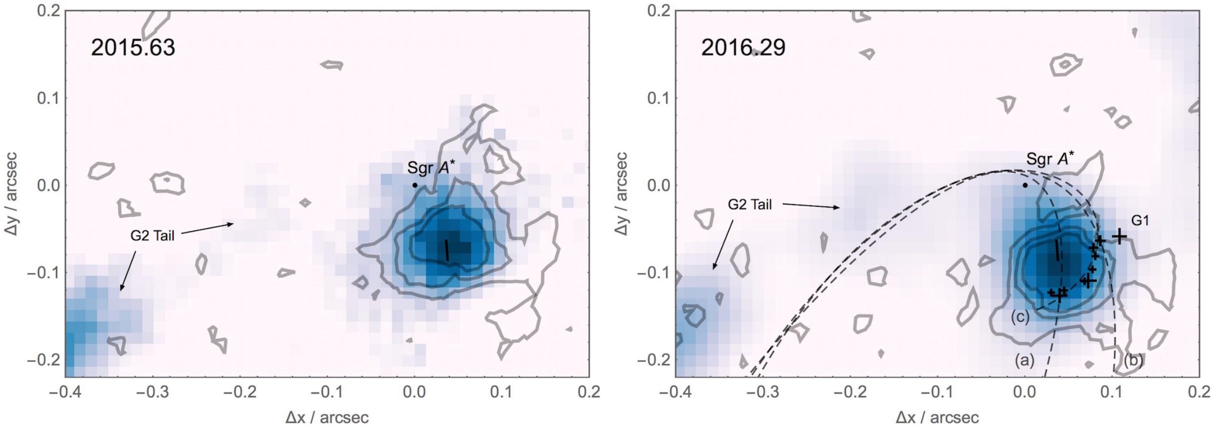

Figure 3.7: Images of the L-band emission from G2 in 2015 and 2016 (shown in color), superimposed with the blue-shifted Brγ line emission at similar times (shown in contours), which appears at the same location. All stellar sources in the region and any significant emission from Sgr A* have been subtracted (Sec. 3.2.2). For ease of comparison, the on-sky motion of the centroid between the two epochs is marked (solid line). The faint emission trailing G2 might be part of the larger tail structure (Sec. 3.4.3). In the bottom panel, the on-sky trajectory of (a) the updated best-fit orbit of G2, (b) the orbit of G2 from Pfuhl et al. (2015) and (c) their combined orbit of G2/G1 is overplotted (dashed lines). The indicated positions of G1 (crosses) were measured between 2003 and 2010.

in very good agreement with that of the blue-shifted Brγ emission at similar times. While the Brγ emission is resolved in position-velocity space (Sec. 3.3.1), the L-band emission is unresolved in our images. That is to say, it is not significantly elongated with respect to the estimated PSF, which has a FWHM on the order of 120 mas. However, due to the high background emission in this crowded region, we are only able to observe the very concentrated part of the dust distribution, which could intrinsically still be very similar to the gas distribution. We estimate the L-band magnitude of G2 to be 13.8 ± 0.5 in 2015 and 13.2 ± 0.5 in 2016, which is consistent with a constant or possibly weakly increased luminosity compared to previous measurements before and around the time of pericenter (Witzel et al., 2014).

3.4 Discussion

3.4.1 G2 as a probe of the accretion zone

The non-detection of a strong drag force on G2 during its pericenter passage (but see Sec. 3.3.1) allows us to constrain internal properties of the gas cloud, as well as external properties of the ambient gas in the inner accretion zone. To first approximation, the drag force acting on G2 as it moves along on its orbit, assuming a stationary and homogeneous

Parameter		Value
semi-major axis	a (arcsec)	0.96 ± 0.07
eccentricity	e	0.983 ± 0.002
inclination	i (deg)	123 ± 1
longitude of ascending node	Ω (deg)	65 ± 3
argument of pericenter	ω (deg)	92 ± 2
time of pericenter	t_p (year)	2014.18 ± 0.09
initial cloud size (FWHM)	w_r (mas)	50 ± 6
starting time	t_0 (year)	1975 ± 8
orbital period	P (years)	346 ± 37

Table 3.1: Summary of the model parameters, which describe a test-particle cloud on a Keplerian orbit that fits the observed evolution of G2 in Brγ emission between 2004 and 2016. (Given is the median value of the marginal posterior distribution of each parameter; see sup. Fig. 3.11.)

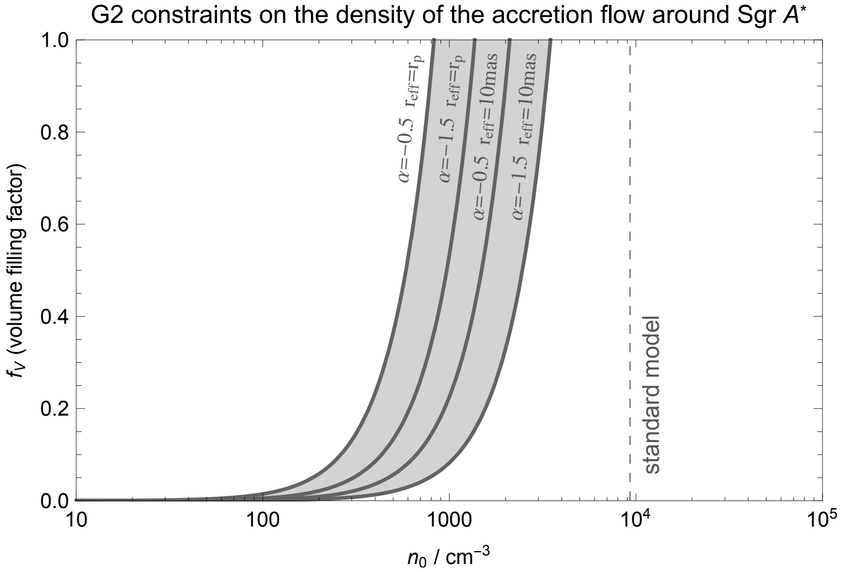

Figure 3.8: The upper limit on the drag force acting on G2 constrains the density of the ambient gas at the pericenter distance and the volume filling factor of G2 to the shaded region, assuming G2 is a spherical cloud with an effective radius between 10 mas and 16 mas ($\approx r_p$) that is moving through a stationary and homogeneous atmosphere. The assumed slopes of the ambient density profile $n(r) = n_0 \times (r/r_p)^\alpha$, where $\alpha = -1.5$ or $\alpha = -0.5$, yield similar exclusion regions in this graphic. The density indicated by the vertical line is derived from the accretion model by Yuan et al. (2003) and Xu et al. (2006) that was used by Gillessen et al. (2012) and others (e.g. Burkert et al., 2012; Schartmann et al., 2012) to predict the pericenter and post-pericenter evolution of G2.

atmosphere, is expected to have the form of a ram pressure:

$$\boldsymbol{F}_D = M_{\mathrm{cl}}\boldsymbol{a}_D \approx -\frac{1}{2}A_{\mathrm{eff}}\rho v^2\frac{\boldsymbol{v}}{v} \tag{3.2}$$

The accretion flow is fully ionized and charge neutral, so the ambient mass density ρ converts into an electron number density $n = \rho/m_p$. We parameterize the density profile $n(r)$ as a radial power law, normalized to the density at the pericenter distance ($r_p \approx 1500\ R_s$).

$$n(r) = n_0\left(\frac{r}{r_p}\right)^{\alpha} \tag{3.3}$$

G2 was estimated to have a mass of $M_{\mathrm{cl}} \approx 3\ M_{\oplus}\sqrt{f_v}$, with f_V being the volume filling factor, and an effective cross-section of $A_{\mathrm{eff}} = \pi r_{\mathrm{eff}}^2$ with $r_{\mathrm{eff}} \approx 15$ mas in 2011 (Gillessen et al., 2012; Shcherbakov, 2014). Additionally, we find from the test-particle simulations that a reasonable minimum size of a cloud resembling G2 at pericenter is 10 mas. Based on these assumptions, we are able to relate the numeric drag force coefficient c_D to the ambient density n_0 and the volume filling factor f_V, by comparing equations 3.1 and 3.2. In all models that we tested, the parameter c_D initially converges toward the edge of the prior volume at zero (assuming a strictly resistive drag force). The final best-fit value is non-zero but small, and the resulting drag force has a negligible impact on the orbital evolution. However, this value represents an upper limit on c_D that yields a constraint on the ratio $n_0/\sqrt{f_V}$ (Fig. 3.8). In particular, it must be $n_0 \lesssim 10^3\ \mathrm{cm}^{-3}$ for filling factors close to unity. The assumed slope of the density profile has only a minor influence on the order of magnitude of this limit.

X-ray observations constrain the ambient density to about $10^2\ \mathrm{cm}^{-3}$ at the scale of the Bondi radius ($10^5\ R_s$), assuming Bremsstrahlung is the origin of the observed gas emission (e.g. Baganoff et al., 2003). At a scale of just a few R_s, radio observations constrain the density to about $10^7\ \mathrm{cm}^{-3}$, via measurements of the polarization (i.e. Faraday Rotation) and the size of the emission from Sgr A* (e.g. Quataert & Gruzinov, 2000; Bower et al., 2003; Marrone et al., 2007; Doeleman et al., 2008). Both of these constraints can be brought into agreement if the density profile is a smooth power law with a slope of $\alpha \approx -1$. In comparison to this prediction, our upper limit at around $10^3\ R_s$ is smaller by roughly an order of magnitude. If the density of the inner accretion zone is indeed this small, it would imply a broken power law with a slope that is shallower at larger radii ($\alpha \approx -0.5$), and steeper at smaller radii ($\alpha \approx -1.4$). It would also explain the non-detection of excess radio or X-ray emission above the quiescent level of Sgr A*, since the expected observable emission from a shock front would not exceed this threshold (following Gillessen et al., 2012). A possible co-rotation of the accretion zone could additionally reduce the velocity of the gas cloud relative to the local gas (i.e. the Mach number), to a similar effect (e.g. Narayan et al., 2012). A lower ambient density would also allow proportionally larger cloud sizes, without the need for additional mechanisms of (e.g. magnetic) confinement (Shcherbakov, 2014; McCourt et al., 2015).

The remaining discrepancy between the test-particle model and the observations (see Sec. 3.3.1) might be explained by additional material behind or in front of the cloud, or by a more complex interaction between G2 and the accretion flow than the drag force we have assumed in our model. For example, the shape of G2 might also not be spherical, or there might be a rotation of the accretion flow (e.g. Sądowski et al., 2013a; McCourt & Madigan, 2016), an in- or outflow (e.g. Madigan et al., 2017), or dynamically important magnetic fields (e.g. McCourt et al., 2015). These effects could alter both the strength and the direction of a drag force. It can also not be excluded that the properties of the emission process are variable. In addition, the data analysis is most susceptible to systematic uncertainties around the time of pericenter, due to the low surface brightness of the observed emission. The test-particle model still performs surprisingly well at describing the overall evolution of the gas cloud between 2004 and 2016.

3.4.2 The relationship between G2 and G1

A relationship between G2 and G1 was first proposed and investigated by Pfuhl et al. (2015) and later studied in more detail by McCourt & Madigan (2016) and Madigan et al. (2017), but challenged by Sitarski et al. (2015) on the basis that their orbital elements are inconsistent. The G1 object, which was first recognized in observations by Clenet et al. (2004a,b, 2005) and Ghez et al. (2005), is a source apparently much alike G2 that passed the black hole slightly more than a decade earlier, and was only observed after its pericenter passage. It started to fade soon afterwards, so that follow-up detections could not be confirmed.

The post-pericenter trajectory of G2 roughly matches the one of G1 (Fig. 3.7), but goes on in a more southern direction. However, the orbits of G2 and G1 both have large eccentricities, times of pericenter that coincide to within 13 years (a small fraction of the 350 yr orbital period), and orbital angular momentum vectors that are aligned to within 20° (i.e. a chance alignment has a probability of less than 3%). It is therefore very likely that G1 and G2 are related objects, even though the individual orbital elements differ significantly at the level of statistical uncertainty, possibly simply because the initial conditions at the times of formation were different.

3.4.3 The relationship between G2 and the tail

The tail of G2 is an extended filament of gas and dust trailing the main G2 source (e.g. Gillessen et al., 2013a; Pfuhl et al., 2015), which appears to be the densest knot in this superstructure. To examine this tail in greater detail, we make use of a widely dithered set of SINFONI data that we were able to obtain as part of our 2014 observing campaign (Pfuhl et al., 2015). From this data, we are able to extract an extended pv-diagram, along a trajectory that follows the apparent tail emission, which is only tentatively detected in our regular monitoring observations (Fig. 3.9 & 3.10). No particular object is known to have moved along an orbit describing such a trajectory, rather the entire tail is visible along this path at this moment in time (see also sup. Fig. 3.14).

The tail is continuously visible in Brγ emission over a large range of velocities, and can be traced back at least into the group of stars surrounding the eclipsing binary IRS 16 SW

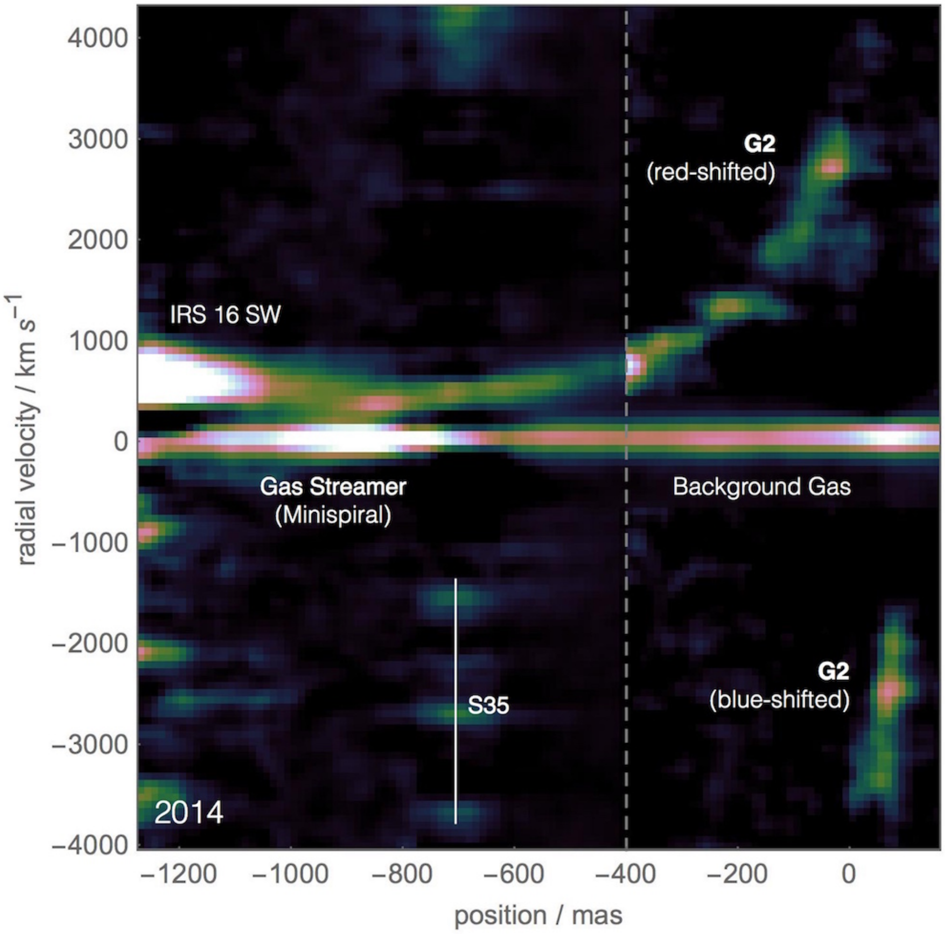

Figure 3.9: Extended pv-diagram showing the Brγ emission trailing G2 in 2014, which is continuously connected to the leading cloud over a large range in position and radial velocity. The color scaling on the left and right side of this graphic is adjusted separately (dashed line), to emphasize the structure of the tail rather than changes in luminosity. Where the tail intersects the star S35 in projection, the noise is increased due to the necessary subtraction of stellar continuum emission.

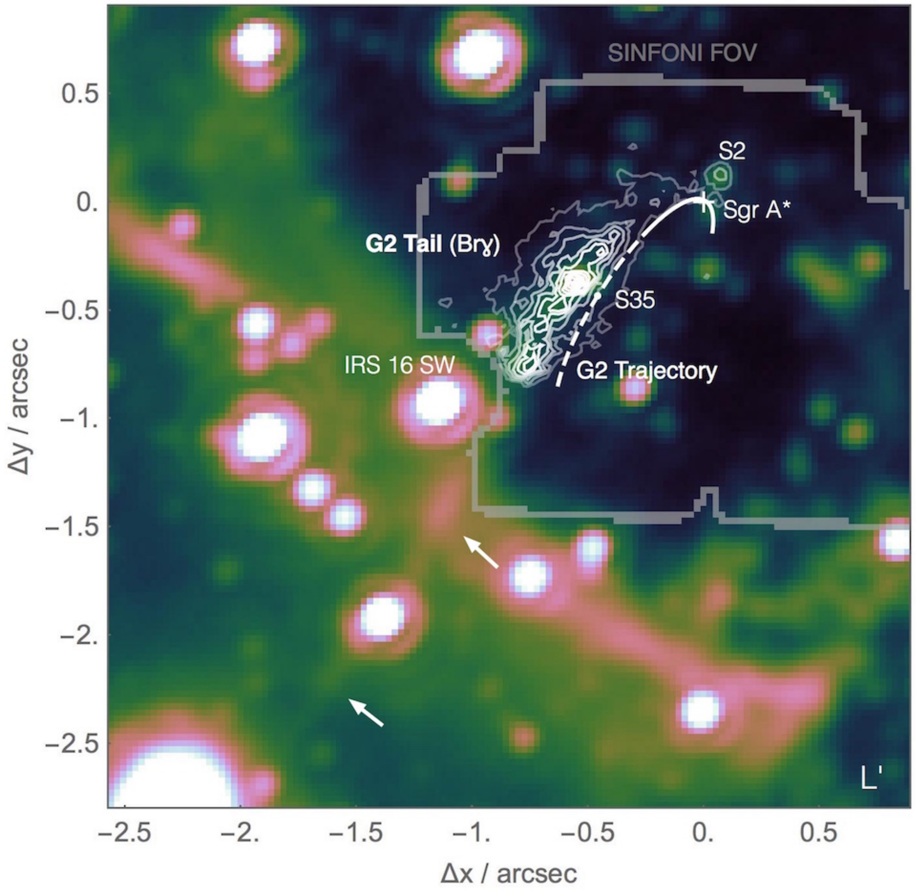

Figure 3.10: Extended L-band image showing the dusty filament that coincides with the tail of Brγ emission connected to G2 (shown in contours), and which possibly continues beyond the field of view of the SINFONI observations. To visualize the full spatial extent of the Brγ tail, which has a variable radial velocity, the data cube is collapsed along the spectral dimension using a maximum filter (see also sup. Fig. 3.14). Also shown is the observed trajectory of G2 (solid line) and an extrapolation of the best-fit orbit to the year 1900 (dashed line).

(Martins et al., 2006), located about 1″ toward the southeast of Sgr A*, from where the tail appears to channel gas toward the center. There is also a dusty filament visible in L-band that coincides with the Brγ tail, and which possibly continues even beyond the field of view of the SINFONI observations (Fig. 3.10). By comparison to estimates for G2, the whole tail must have a gas mass of at least several Earth masses.

The massive contact binary IRS 16 SW is a member of the clockwise stellar disk, and has a broad Brγ emission feature at a wavelength of 2.17 μm. As a colliding-wind system, it could be an effective site of clump formation (Calderón et al., 2016), and thus a candidate origin of G2[4]. However, the part of the tail at the very edge of the field of view does not point directly toward IRS 16 SW, and it is still unclear how much further the tail might continue and whether or not it blends with the background gas.

At its other end, the tail smoothly and seamlessly connects to the red-shifted part of G2 in position-velocity space, as well as in the plane of the sky (Fig. 3.7 & 3.10), and even its velocity gradient matches that of G2 (Fig. 3.9), which in combination is strongly suggestive of a common origin. The tail does not lie exactly on the extrapolated orbit of G2, but a few 0.1″ away in projection. At these larger distances from Sgr A*, the dynamics of the tail could well be different from those inferred for G2, as it would be expected for a genuine tidal tail (e.g. Guillochon et al., 2014). We conclude that it is very likely that both G2 and G1 are not just related to each other, but to the tail as well, since again a chance occurrence would be a rare coincidence.

3.5 Summary

During 2015 and 2016, we have observed G2 completing the pericenter passage of its orbit around the massive black hole Sgr A* in the Galactic Center. By putting these new data sets in the context of our previous observations, we can confidently establish the following observational facts:

- *G2 is firmly detected as a dusty gas cloud in Brγ line emission and L-band continuum emission* (Fig. 3.1, 3.2, 3.3 & 3.7). The spectroscopic data set, obtained with SINFONI at the VLT, spans 12 years and covers the entire pericenter passage of G2.

- *The gas cloud moves on a highly eccentric, ballistic orbit around Sgr A* and tidally evolves approximately like a cloud of test particles between 2004 and 2016.* The gas dynamics are dominated by the gravitational force due to the black hole. The non-detection of a drag force, or any strong hydrodynamic interaction with the hot gas in the inner accretion zone, limits the ambient density to less than a few 10^3 cm^{-3} at the pericenter distance (1500 R_s), assuming G2 is a spherical cloud moving through a stationary and homogeneous atmosphere (Fig. 3.8). This limit is in good agreement with the non-detection of a bow shock at radio or X-ray wavelengths.

---

[4] It has since been shown that a stellar wind collision in IRS 16 SW is unlikely to be able to eject G2-like clumps into G2-like orbits (Calderón et al., in prep.).

- *The gas cloud shows a pronounced spatially resolved velocity gradient both before and after pericenter* (Fig. 3.1 & 3.4). This implies that the gas must be spatially extended throughout the pericenter passage.

- *The dust emission stays unresolved but co-spatial with the gas emission* (Fig. 3.7). We note that this is consistent with the earlier observed compactness of the dust emission around the time of pericenter (Witzel et al., 2014).

- *The Brγ and L-band luminosities remain constant to within the measurement uncertainty* (Fig. 3.5). There is no need to replenish gas or dust from a potential central source.

- *G2 is likely related to G1 and the tail structure.* Although the orbital evolution of G1 and G2 is not identical, the probability of a chance occurrence is small (Sec. 3.4.2). The same is true for the tail, which is continuously connected to G2 in position-velocity space (Sec. 3.4.3).

To further study how G2 could be used as a more powerful probe of the inner accretion zone around Sgr A*, it will be necessary to consider more complete models of the possible interactions between G2 and the ambient medium. If they are computationally feasible, more comprehensive (hydrodynamic) simulations based on such models might be able to explain the detailed features of the observational data, which is rich in information[5]. Moreover, it will be worth continuing to monitor the further evolution of G2, since the gas cloud might eventually interact with the ambient medium and disrupt or break up, as some hydrodynamic simulations have predicted (e.g. Schartmann et al., 2015). At this point a test-particle model would break down, but also give up more information about the properties of G2, as well as the accretion flow.

The continuing post-pericenter evolution of G2 over the next few years might also hold decisive clues about its nature and origin. There has been no direct detection of a stellar photosphere (in K- or H-band), but dynamic evidence for the existence of a massive central source could arise if, for instance, a compact dusty source would maintain Keplerian motion, while the bulk of disrupting gas does not. However, in consequence G2 might also fade as a Brγ source, resembling G1 in this respect, and become unobservable in the near future.

Acknowledgements

For helpful discussion, we are grateful to A. Ballone and M. Schartmann, as well as J. Cuadra and D. Calderón.

---

[5] A recent example are the simulations performed by Steinberg et al. (2018).

3.A Supplementary figures

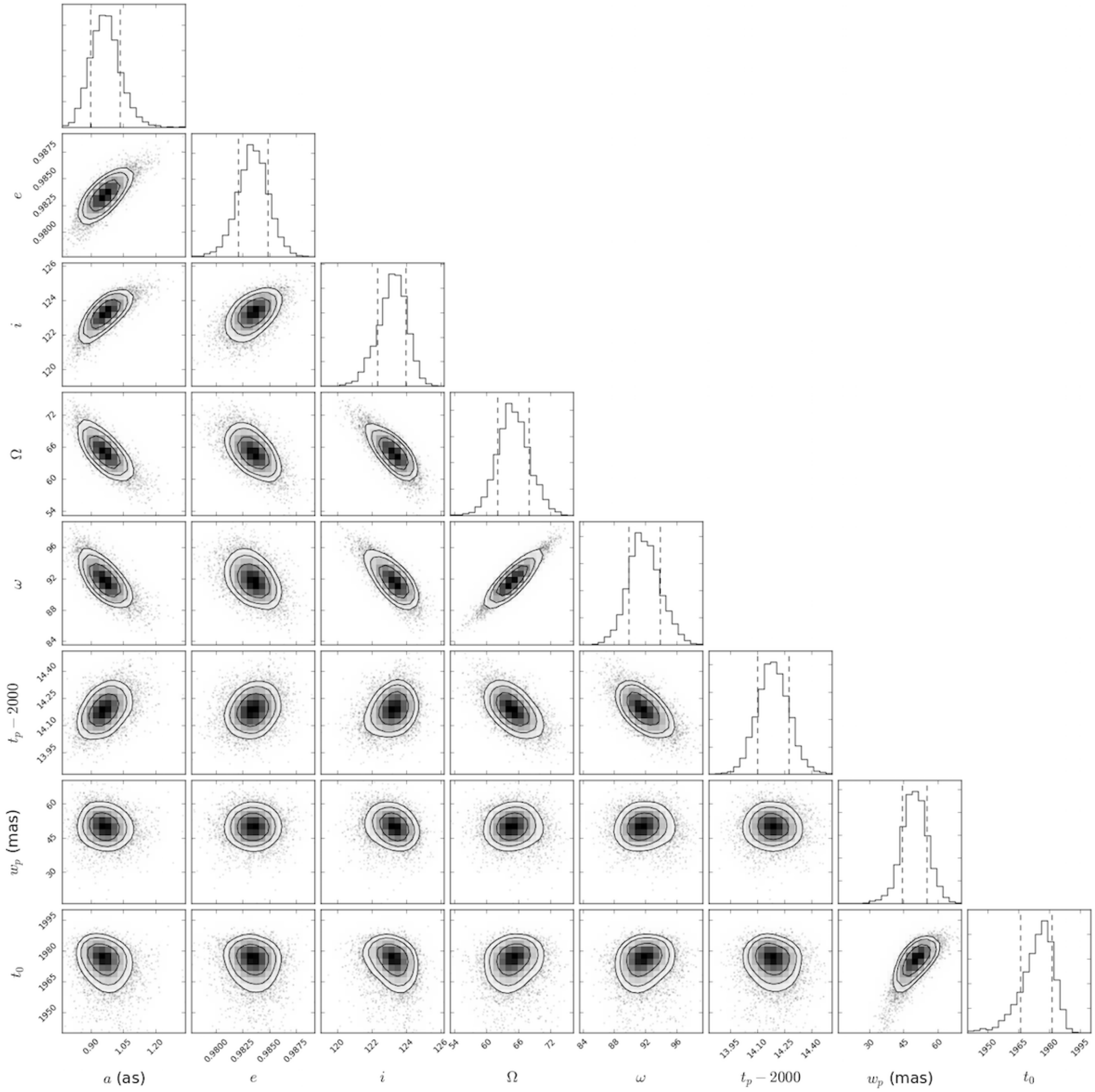

Figure 3.11: Visualization of the joint posterior parameter distributions for the Keplerian test-particle model of G2. The dashed lines in the one-dimensional histograms indicate the 16% and 84% percentiles (1σ), which correspond to the innermost contour in the paired histograms.

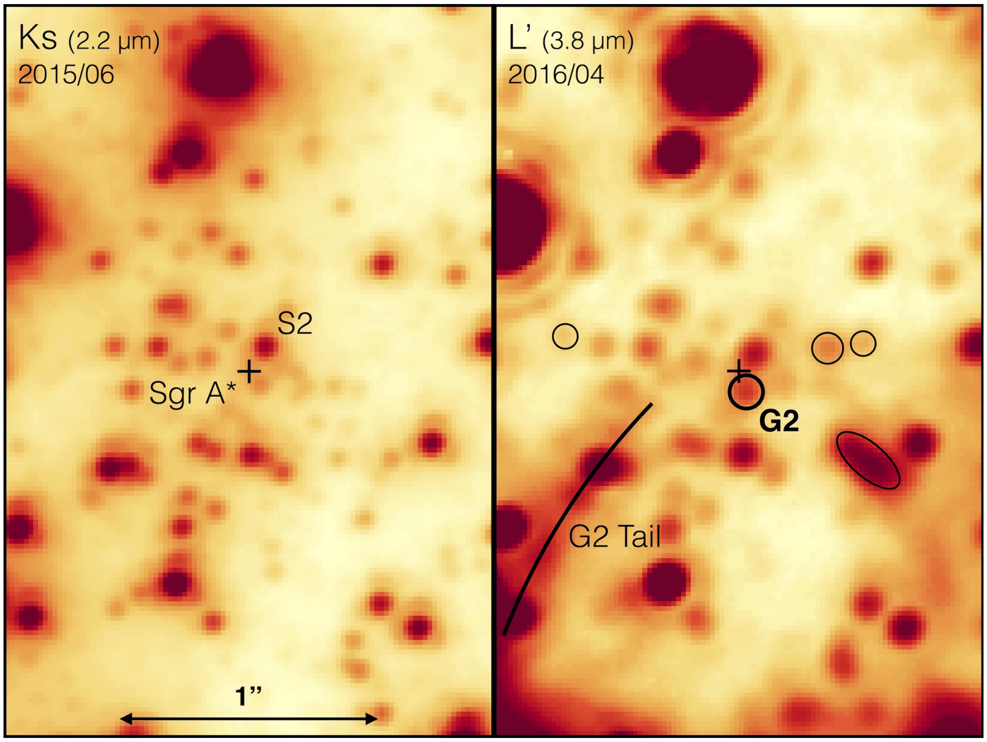

Figure 3.12: A direct comparison of Ks- and L'-band images, on which G2 and other sources with a similar appearance are marked.

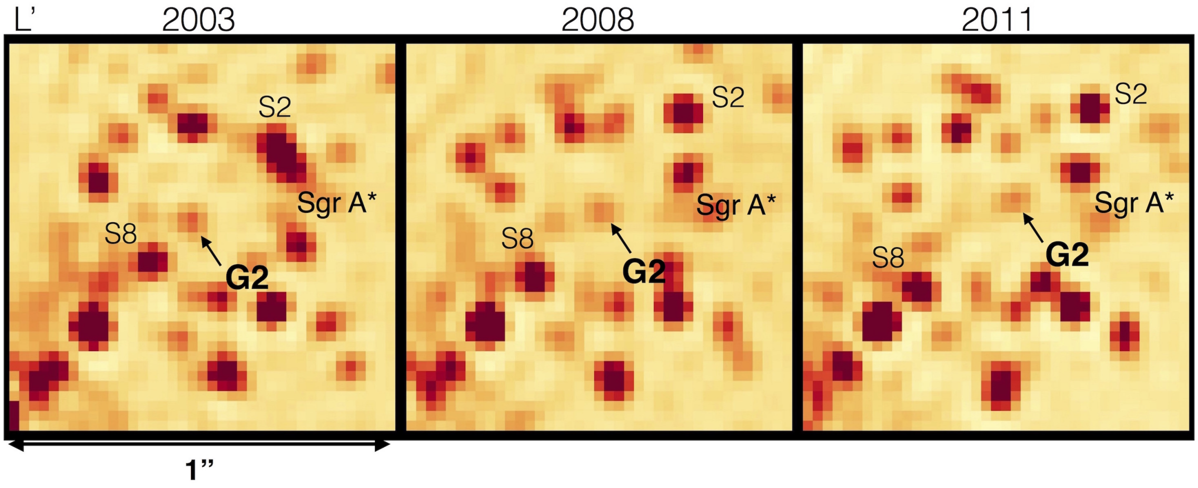

Figure 3.13: The motion of G2 between 2003 and 2011, as seen in deconvolved L'-band archive images (see Gillessen et al., 2013a,b).

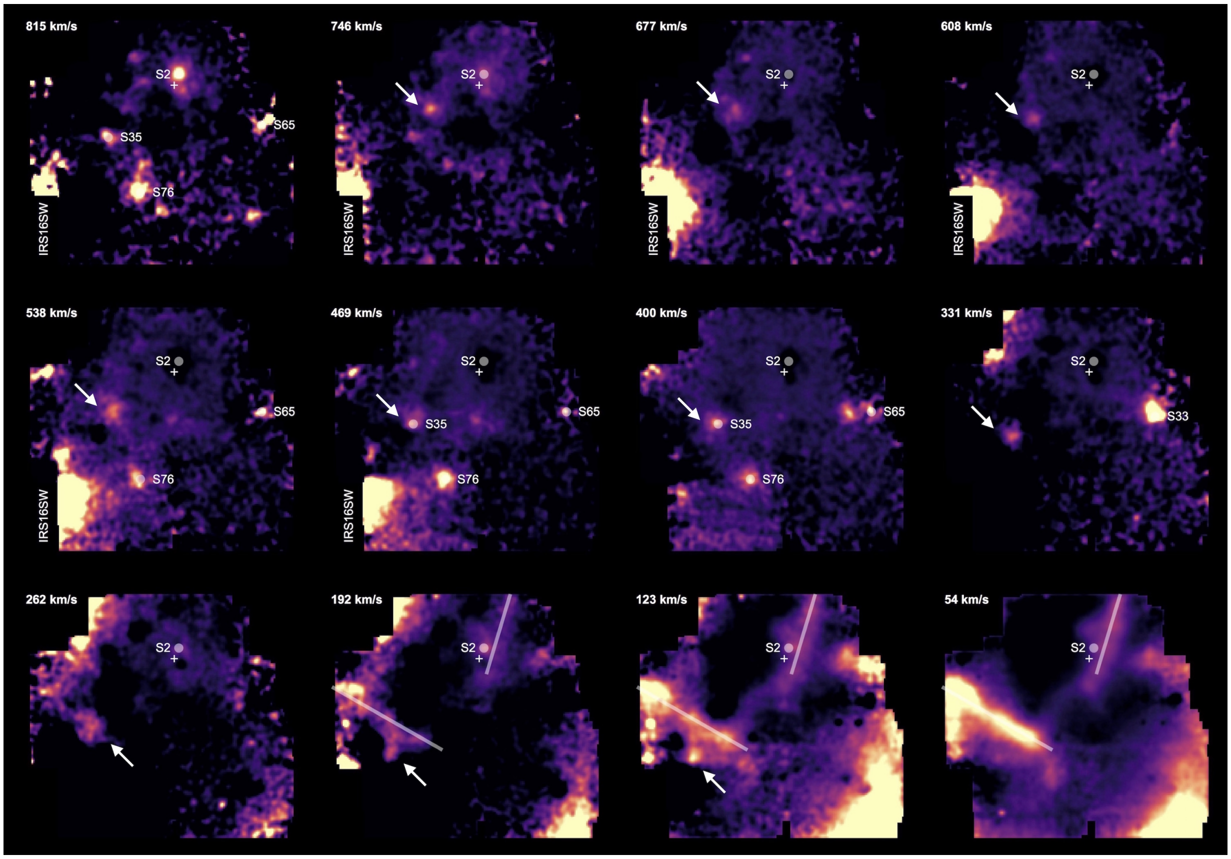

Figure 3.14: Maps of the Brγ tail emission in different velocity channels (white arrows).

Chapter 4

The properties and effects of unrecognized astrometric confusion

Original publication: P. M. Plewa and R. Sari, 2018, MNRAS, *Unrecognized Astrometric Confusion in the Galactic Centre*, doi:10.1093/mnras/sty512

Abstract: Since the Galactic Center is a highly crowded stellar field, frequent unrecognized events of source confusion, which involve undetected faint stars, are expected to introduce astrometric noise on a sub-mas level. This confusion noise is the main non-instrumental effect limiting the astrometric accuracy and precision of current near-infrared imaging observations, and the long-term monitoring of individual stellar orbits in the vicinity of the central massive black hole. We self-consistently simulate the motions of the known and the yet unidentified stars to characterize this noise component, and show that a likely consequence of source confusion is a bias in estimates of the stellar orbital elements, as well as the inferred mass and distance of the black hole, in particular if stars are being observed at small projected separations from it, such as the star S2 during pericenter passage. Furthermore, we investigate modeling the effect of source confusion as an additional noise component that is time-correlated, demonstrating a need for improved noise models to obtain trustworthy estimates of the parameters of interest (and their uncertainties) in future astrometric studies.

4.1 Introduction

The long-term monitoring of stellar motions in the Galactic Center has enabled detailed studies of stellar dynamics in the extreme environment of a galactic nucleus, which have benefited from the favorable conditions and growing capabilities of ground-based observations at near-infrared wavelengths (for a review, see Genzel et al., 2010). Most notably, large proper motions of individual stars (Eckart & Genzel, 1996; Ghez et al., 1998) and the Keplerian orbital motion of the star S2 in particular (Schödel et al., 2002; Ghez et al., 2003) have provided convincing evidence for the existence of a central massive black hole, identified with the compact radio source Sgr A*. As part of the ongoing monitoring effort, more than 100 stars in total are being tracked within the inner few $1''$ (e.g. Ghez et al., 2008; Gillessen et al., 2009b; Boehle et al., 2016; Gillessen et al., 2017). These stars in the immediate vicinity of the black hole were revealed as a result of several major technological advancements in high-resolution observing techniques, such as speckle imaging and the development of AO systems, which mitigate the effect of atmospheric turbulence, while there were parallel advancement in methods of data processing (e.g. deconvolution algorithms) and data analysis (e.g. Bayesian inference). However, even at an angular resolution close to the diffraction limit of today's largest optical telescopes, both the astrometric accuracy and precision of imaging observations of the Galactic Center, as well as their depth, are fundamentally limited by a high level of stellar crowding.

Most importantly, stellar crowding limits the accuracy with which the extended seeing halo of the PSF can be determined from images. The stray halo light that is (at least partially) unaccounted for in empirical PSFs, mainly produced by a small number of luminous (WR/O-)stars, results in astrometric noise (halo noise), which represents the dominant contribution to the noise budget of all but the brightest stars in the inner nuclear star cluster (Fritz et al., 2010b). However, also important is the related effect of source confusion, as well as, for instance, residual image distortion, and ultimately a limited signal to noise ratio. While it is justified to exclude astrometric measurements that are firmly identified as outliers attributable to recognized confusion events involving single known stars, unrecognized confusion events, involving undetected stars, are by definition not separately identifiable and have to be modeled statistically.

Here we present a statistical study of the astrometric effect of unrecognized source confusion in isolation, which is observationally indistinguishable from the instrumental halo noise, to better understand this effect and its impact on the monitoring of stellar motions, in anticipation of future observations using next-generation instruments, and to help further explore the limits of high-precision Galactic Center astrometry.

4.2 Methods

4.2.1 Astrometric confusion

If, while they are moving on the sky, two stars approach each other in projection so closely that their images overlap, a measurement of their astrometric position can be significantly biased, depending on their projected (two-dimensional) on-sky separation $\mathbf{d}$ and flux ratio L'/L. If one of the stars is too faint to be detected, such a confusion event is called unrecognized. However, the position of the other star will still be offset by a small amount $\Delta\mathbf{d}$, and it will appear slightly brighter by an amount ΔL.

To estimate the induced changes in position and brightness in such a scenario, it is practical to minimize the total squared residuals between the true image of two overlapping sources, which are the test star and a perturbing star, and a naive model image of only a single source. This minimization can be regarded a maximum likelihood estimation

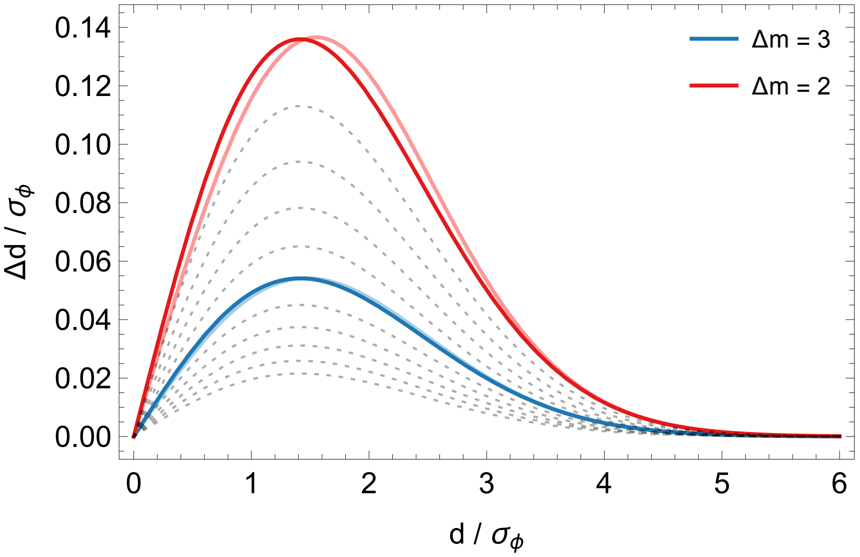

Figure 4.1: The astrometric offset of a test star induced by a single perturbing star, in units of the PSF width σ_ϕ, as a function of the projected separation and magnitude difference of the two stars. Both the accurate numeric solution (to Eq. 4.1) and the analytic approximation (Eq. 4.3) are shown in light and dark color, respectively, for magnitude differences of $\Delta m = 2$ (red lines) and $\Delta m = 3$ (blue lines). The dotted gray lines indicate how the approximate solution changes as the assumed magnitude difference varies over the range $2 \leqslant \Delta m \leqslant 4$ in steps of 0.2.

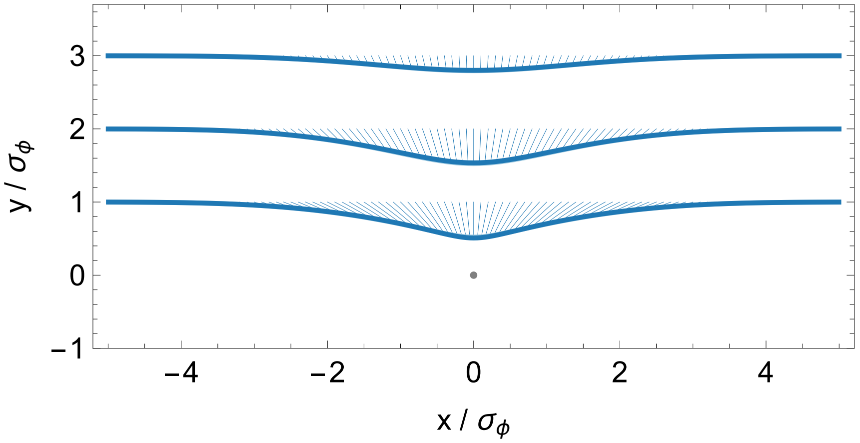

Figure 4.2: The two-dimensional astrometric offset of a test star as it passes a single, stationary perturbing star in projection (gray dot), for different separations at closest approach ($3\sigma_\phi$, $2\sigma_\phi$ and $1\sigma_\phi$, where σ_ϕ is the width of the PSF). In all cases, the assumed magnitude difference is $\Delta m = 3$ and the astrometric offsets are multiplied by a factor 10 for better visibility.

if the noise in the image is Gaussian and constant over the region of interest. Both assumptions should hold for the near-infrared observations of the Galactic Center, since they are background-dominated over any sufficiently small field of view (Fritz et al., 2010b). The log-likelihood function to be maximized is, up to a constant:

$$\log \mathcal{L} = -\frac{1}{2}\int_{-\infty}^{\infty} d^2r \left(L\phi(\mathbf{r}) + L'\phi(\mathbf{r}-\mathbf{d}) - (L+\Delta L)\phi(\mathbf{r}-\mathbf{\Delta d})\right)^2 \tag{4.1}$$

The PSF $\phi(\mathbf{r})$ is assumed to be a well-sampled symmetric Gaussian function with a fixed width σ_ϕ, which, for a first approximation, is a reasonably accurate description of the diffraction-limited core of a realistic PSF under typical observing conditions:

$$\phi(\mathbf{r}) = \frac{1}{\sqrt{2\pi\sigma_\phi^2}} \exp\left[-\frac{\mathbf{r}^2}{2\sigma_\phi^2}\right] \tag{4.2}$$

The exact shape of the instantaneous PSF is variable, both spatially and over time, and can depend quite sensitively on the performance of the AO system.

If the perturbing star is faint enough ($L'/L \ll 1$), maximizing the likelihood function with respect to Δd and ΔL yields:

$$\frac{\Delta d}{d} \approx \frac{\Delta L}{L} \approx \frac{L'}{L} \exp\left[-\frac{d^2}{4\sigma_\phi^2}\right] \tag{4.3}$$

The direction of the test star's offset is toward the perturbing star, along the line connecting the two sources. This approximate solution is sufficiently good if the magnitude difference between the two stars is $\Delta m \gtrsim 2$ (see Fig. 4.1), where by convention:

$$\Delta m = m' - m = -\frac{5}{2}\log_{10}\left(\frac{L'}{L}\right) \tag{4.4}$$

A perturbing star thus has the largest influence if it is located at a projected separation of $d = \sqrt{2}\,\sigma_\phi$ from the test star, while at separations $d \gtrsim 5\,\sigma_\phi$ its influence becomes negligible (see Fig. 4.1 & Fig. 4.2). In the case of more than one perturbing star, the total offset of the test star is the vector sum of the individual offsets toward each of the perturbers, and the increase in brightness is cumulative.

4.2.2 The stellar background

A large number of yet unidentified stars is expected to exist concentrated around the Galactic Center black hole. These are sub-giant and main-sequence stars, which constitute the faint component of the existing stellar population. Despite being hidden below the detection limit of current instruments, this background (or foreground) population has an effect on the measured motions of the known, brighter stars in the central (S-)star cluster, by being a source of astrometric noise due to unrecognized source confusion.

We endeavor to study the properties and implications of this confusion noise by analyzing statistically the motion of a test star in an artificial field of background stars, as shown in Figure 4.3, while accounting for their dynamics and considering many possible realizations of their distribution. For this purpose, we assume that the stellar background population can be described by a spherically symmetric Bahcall-Wolf cusp with an isotropic velocity field (Bahcall & Wolf, 1976; Binney & Tremaine, 2011), which has a power-law distribution function $f(\epsilon) \propto \epsilon^{\gamma}$ that only depends on the stars' specific energy $\epsilon = -v^2/2 - \Phi(r)$, where $\Phi(r)$ is the gravitational potential. The single free parameter $\gamma = M_*/4M_{\max}$ depends on the stellar mass partitioning, where the stars of interest are assumed to have a given mass M_* in the range $0 \lesssim M_{\min} \leqslant M_* \leqslant M_{\max}$, so that $0 \lesssim \gamma \leqslant 1/4$. Such a Bahcall-Wolf cusp is the expected state of a dynamically relaxed stellar population surrounding a massive black hole. The stellar population in the Galactic Center has a complex history (e.g. Pfuhl et al., 2011), is in part young and splits into at least three distinct components, when stars are grouped by their spectral types and kinematics (e.g. Genzel et al., 2010, sec. II). We choose to examine the limiting case of a Bahcall-Wolf cusp with $\gamma = 0$, which is analytically tractable and roughly matches the observed, slightly flatter density profile of the entire visible stellar population (e.g. Genzel et al., 2003; Bartko et al., 2010), as well as the observed distribution of diffuse light (e.g. Schödel et al., 2018), emanating from stars that could plausibly be old enough to be dynamically relaxed.

In the inner region of the nuclear cluster, stars move to good approximation like test particles in a point mass potential $\Phi(r) = -GM_0/r$ created by the central black hole, which has a canonical mass of $M_0 \approx 4 \times 10^6 M_{\odot}$, and is located at a distance of $R_0 \approx 8\,\mathrm{kpc}$ from the Sun. To be in a stable configuration, these stars must be on bound, effectively Keplerian orbits, which can be defined by six orbital elements. The required distribution of the semi-major axes $a = GM_0/2\epsilon$ of the background stars, to generate a realization of a Keplerian Bahcall-Wolf cusp, is (Schödel et al., 2003; Alexander, 2005):

$$p(a \mid \gamma) \propto a^{\frac{1}{2}-\gamma} \tag{4.5}$$

The corresponding number density profile[1] is $n(r) \propto r^{-\frac{3}{2}-\gamma}$. To ensure numerically that in our simulations the correct surface density profile is reproduced within the central few $1''$, an upper cutoff of this distribution at $a = 2''$ is sufficient. Any stars with larger semi-major axes are unlikely to be found very close to the central black hole, even in projection. We do not impose a strict lower cutoff, but exclude very few stars with a minimum (pericenter) distance $r_p = a(1-e) < r_t$ from the black hole, where r_t is the approximate tidal radius of a typical observed B-type star in the Galactic Center (e.g. Habibi et al., 2017):

$$r_t \approx 1.7\,\mathrm{AU} \left(\frac{R_*}{5R_{\odot}}\right) \left(\frac{M_*}{10M_{\odot}}\right)^{-\frac{1}{3}} \left(\frac{M_0}{4 \times 10^6 M_{\odot}}\right)^{\frac{1}{3}} \tag{4.6}$$

The required distributions of the eccentricity e, the inclination of the orbital plane i, the angle of the ascending node Ω, and the argument of pericenter ω, the specific forms of

---

[1] The surface number density profile is: $\Sigma(R) = 2\int_R^{\infty} n(r) \frac{r\,dr}{\sqrt{r^2-R^2}} \propto R^{-\frac{1}{2}-\gamma}$.

which encode the assumption of isotropy for the stellar velocities and positions, are the following:

$$p(e) = \begin{cases} 2e & 0 \leqslant e \leqslant 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.7}$$

$$p(\cos i) = \begin{cases} 1 & 0 \leqslant \cos i \leqslant 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

$$p(\Omega) = \begin{cases} \frac{1}{2\pi} & 0 \leqslant \Omega \leqslant 2\pi \\ 0 & \text{otherwise} \end{cases} \tag{4.9}$$

$$p(\omega) = \begin{cases} \frac{1}{2\pi} & 0 \leqslant \omega \leqslant 2\pi \\ 0 & \text{otherwise} \end{cases} \tag{4.10}$$

The time of pericenter has to be sampled uniformly over the duration of one orbital period, which in turn depends on the semi-major axis:

$$p(t_p \mid T(a),\, t_0) = \begin{cases} \frac{1}{T} & t_0 - \frac{T}{2} \leqslant t_p \leqslant t_0 + \frac{T}{2} \\ 0 & \text{otherwise} \end{cases} \tag{4.11}$$

$$T(a) = 2\pi\sqrt{\frac{a^3}{GM_0}} \tag{4.12}$$

The reference epoch $t_0 = 2000$ is arbitrary, but chosen to coincide roughly with the beginning of Galactic Center observations using advanced high-resolution observing techniques, specifically AO-assisted imaging and integral field spectroscopy.

Magnitudes are assigned to the background stars such that the observed K-band luminosity function (KLF) of the nuclear star cluster is reproduced, which has a power-law slope of $\beta \approx 1/5$ in the magnitude range $10 \lesssim m_K \lesssim 18$ (e.g. Genzel et al., 2003; Bartko et al., 2010). We use the same slope to extrapolate beyond the detection limit to a maximum magnitude of 22, independent of other parameters (e.g. the spatial distribution):

$$p(m_K \mid \beta) \propto 10^{\beta m_K} \tag{4.13}$$

Stars are drawn randomly until a surface density of 60 detectable stars within the central square arcsecond is reached, the magnitudes of which fall in the range $14 < m_K < m_{\text{lim}}$, where typically $m_{\text{lim}} \approx 17$. Afterwards, only the fainter stars with magnitudes larger than m_{lim} are kept, which, in the central region, are the normally undetected perturbing stars. Any stars fainter than magnitude $m_K = 22$ have a negligible astrometric effect even on the faintest observable stars, even though they are predicted to be numerous. Higher angular resolution and sensitivity would be necessary to be able to study these fainter stars as well.

In our simulations, we do not account for any possible deviations from Keplerian orbits, for example due to a not point-like central mass distribution, encounters between stars or with remnants, or post-Newtonian effects, none of which have been detected so far (e.g.

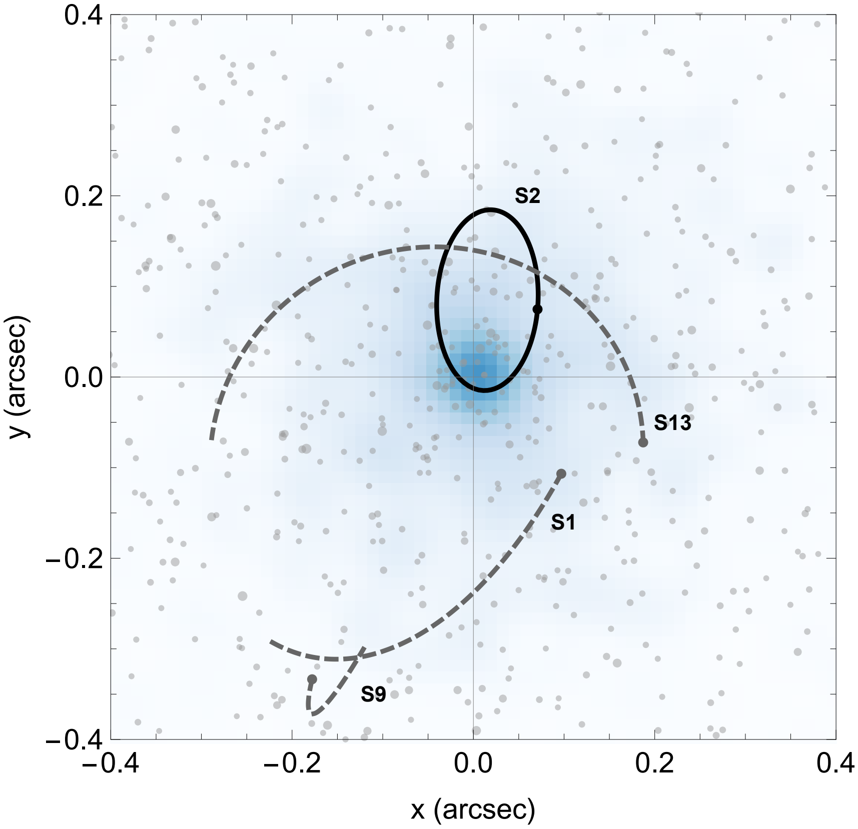

Figure 4.3: The on-sky trajectories of the stars S1, S2, S9, and S13 over the time of simulated observations (solid and dashed lines), starting in the year 2000 and covering exactly one orbital period of S2 (approx. 16 years, see Gillessen et al. (2017) for measurements of the orbital elements), overlaid on a snapshot of a sample realization of the distribution of background stars (shown as points) and their average light distribution (shown in color).

Sabha et al., 2012; Gillessen et al., 2017). These effects are either small enough to be negligible here, or work on a different timescale than the effect of source confusion in the way it affects the star S2 or other fast-moving S-stars in the vicinity of Sgr A* (see Sec. 4.3.2), and would also not significantly affect the expected average spatial distribution of the background stars.

4.2.3 Noise modeling & Orbit fitting

We choose the star S2 as our fiducial test star. S2 is the brightest star for which an orbit has been determined, within the central 1″. Also, the orbital period of S2 is the second-shortest one measured to date (approx. 16 years; see also Meyer et al., 2012) and its orbit has been fully covered by observations once (e.g. Gillessen et al., 2009a). For these reasons, the well-measured orbital motion of S2 provides some of the best constraints on the mass and distance of the central black hole today. The orbital trajectory of S2 is shown in

Parameter		Value
black hole mass	M_0	$4 \times 10^6 M_\odot$
black hole distance	R_0	$8\,\mathrm{kpc}$
BW slope	$-3/2-\gamma$	$3/2$
KLF slope	β	$1/5$
limiting magnitude	m_lim	17
PSF width	σ_ϕ	$13.8\,\mathrm{mas}$
semi-major axis	a	$0.13''$
eccentricity	e	0.88
inclination	i	134°
longitude of the ascending node	Ω	227°
argument of pericenter	ω	66°
time of pericenter	t_p	2002.3
test star magnitude	m_K	14.0

Table 4.1: Summary of main simulation parameters.

Figure 4.3, together with that of the stars S1, S9 and S13, as well as a realization of the simulated cluster of background stars.

As the first step in studying the background stars' effect on astrometric measurements, it is necessary to simulate observations of S2's motion, in a simplified but realistic way. Making use of the measured orbital elements (Tab. 4.1, see also Gillessen et al., 2017), the angular astrometric offsets x and y relative to Sgr A*, and the absolute radial velocity v_z, are evaluated at uniformly sampled times over exactly one orbital period, to avoid any possible biases due to only fractional coverage of the orbit. The average frequency of useful imaging and spectroscopic observations is assumed to be 4 and 3 per year, respectively. For simplicity, neither seasonal nor technical constraints on observability are taken into account. To achieve a more uniform orbit coverage, the pericenter passage would need to be sampled at a higher rate, since S2 moves fastest during that time.

Second, the astrometric positions of the background stars are determined at the times of observations and the expected offsets due to unrecognized source confusion are added to the true positions of S2, which has a K-band magnitude of $m_K \approx 14$ (so that $\Delta m_K \geqslant 3$). The FWHM of the PSF is assumed to be a constant 2.5 px at a pixel scale of 13 mas/px, which yields $\sigma_\phi \approx 13.8\,\mathrm{mas}$ (or a FWHM of 32.5 mas). This is the usual width of a restored PSF in a NACO image after it has passed the reduction pipeline (Fritz et al., 2010b). Since the exact dynamical configuration of the cluster of background stars is unknown, we perform separate orbit fits for many plausible, random realizations of our Bahcall-Wolf cluster model (see Sec. 4.2.2).

Finally, additional offsets are added that are due to a generic measurement uncertainty, in the form of Gaussian noise with a standard deviation of $\sigma = 0.3\,\mathrm{mas}$ in each of the spatial dimensions, to account mainly for halo noise, but also any other sources of astrometric noise, and 30 km/s in the radial velocity dimension. For a second test case, we choose a reduced astrometric uncertainty of $\sigma = 0.1\,\mathrm{mas}$, which approaches the signal to noise limit for an S2-like star under ideal conditions (Fritz et al., 2010b).

We then perform Bayesian parameter estimation on the simulated data sets, using different models to test different approaches of accounting for the effect of astrometric source confusion. In general, we use a likelihood function with three separate Gaussian components, corresponding to the two spatial dimensions and the one radial velocity dimension, which are assumed to be measured independently:

$$\log \mathcal{L} = \sum_{u=x,y,v_z} -\frac{1}{2}\boldsymbol{r}_u^T K_u^{-1} \boldsymbol{r}_u - \frac{1}{2}\log\det K_u - \frac{N_u}{2}\log 2\pi \tag{4.14}$$

The residual vector $\boldsymbol{r}_u$ $(u = x, y, v_z)$ describes the deviation of the individual measurements u_i $(i = 1 \ldots N_u)$ from the prediction of a (Keplerian) orbit model $f_{u,\theta}(t_i)$ evaluated at time t_i, for a certain set of model parameters θ, and an additional mean function $m_{u,\theta}(t_i)$ allows for noise with non-zero mean:

$$(\boldsymbol{r}_u)_i = (u_i - f_{u,\theta}(t_i)) - m_{u,\theta}(t_i) \tag{4.15}$$

The minimum set of model parameters includes the black hole mass M_0 and distance R_0, the 6 orbital elements of S2, as well as 5 additional nuisance parameters, that allow for a small drift motion of the dynamical black hole mass in the astrometric reference frame, described to first approximation by an offset in position (x_0, y_0) at the reference epoch t_0, a proper motion (v_x, v_y) and an offset in radial velocity (v_z). These additional degrees of freedom are introduced to account for the uncertainty in constructing a coordinate system that is stable over the entire time of observations (see Yelda et al., 2010; Plewa et al., 2015). The priors for M_0, R_0, the orbital parameters and the coordinate system parameters are chosen to be uniform, within reasonably broad limits.

K_u is the covariance matrix of the measurements, which is symmetric and has shape $N_u \times N_u$. In our standard model, K_x, K_y and K_{v_z} only contain the (true) measurement uncertainties along their diagonal as non-zero elements:

$$(K_u)_{i,j} = \sigma_u^2\, \delta_{i,j} \tag{4.16}$$

This implies or implicitly assumes that the underlying noise is known and uncorrelated over time. In reality, the true measurement uncertainties are of course unknown and have to be estimated, preferably from the data itself, if that is feasible. To account for any uncertainty in the error estimation itself and the possibility that errors may have been systematically underestimated, the likelihood function can be modified, for example by adding in quadrature another term on the diagonal of K_x and K_y (and possibly also K_{v_z}), that becomes an additional model parameter:

$$(K_{x,y})_{i,j} = \left(\sigma_{x,y}^2 + \eta^2\right)\, \delta_{i,j} \tag{4.17}$$

This implies that there is some additional, intrinsic Gaussian noise with standard deviation η. However, astrometric confusion gives rise to noise that is not quite constant, but correlated over time from measurement to measurement, if only slightly. For some extended duration the test star will be confused with the same set of perturbing stars, which will bias its position in a certain direction. We attempt to model this correlated noise by making use of Gaussian processes (GPs) and inferring three additional (hyper-)parameters instead of one, which also affect the off-diagonal elements of the covariance matrices K_x and K_y (yet without altering K_{v_z}, since the radial velocity is unaffected by astrometric confusion):

$$(K_{x,y})_{i,j} = \sigma_{x,y}^2 \, \delta_{i,j} + k_{x,y}(t_i, t_j) \tag{4.18}$$

There exists a variety of valid kernel functions k, which can be used to encode different assumptions about noise properties (see e.g. Rasmussen & Williams, 2006). We have tested various functions from the Matérn class of kernels, which are widely used because they generate noise with a certain roughness that is appropriate for describing many real-world noise processes. However, confusion events cause perturbations that are intrinsically smooth, which are therefore better described by their limiting function, an exponential squared kernel:

$$k_{x,y}(\Delta t = |t_i - t_j|) = \eta^2 \exp\left(-\frac{\Delta t^2}{2\tau_{x,y}^2}\right) \tag{4.19}$$

The extra hyper-parameters are a common amplitude $\eta = \eta_x = \eta_y$ and two characteristic time scales τ_x and τ_y. We choose to fit two independent timescales because stars on certain trajectories may encounter background stars more frequently in one direction than the other (e.g. if orbits are only partially covered by observations). The priors for these parameters are chosen to be scale-free (i.e. uniform in $\log \eta$ and $\log \tau$). The added-noise model and the GP model are fit to the same data set as the standard model, for each of many random realization of the background star population, and, for a first analysis, assuming zero mean noise ($m_{u,\theta} = 0$). We also fit a control model that is identical to the standard model, except that confusion noise is not simulated, so that the assumed generic measurement uncertainty is the only noise component. In this case, all models actually yield the same result; the fitted noise amplitudes converge to zero, and the time scales are unconstrained.

For each model, if we collectively denote its underlying set of assumptions as $\mathcal{M}$ (e.g. the form of the covariance matrices), and consider a simulated data set $\mathcal{D}' = \bigcup_{u=x,y,v_z} \{u_i\}_{i=1}^{N_u}$, the posterior probability distribution of the model parameters θ is, in the conventional notation (e.g. Hogg, 2012), given by:

$$p(\theta \mid \mathcal{D} = \mathcal{D}', \mathcal{M}) = \frac{1}{\mathcal{Z}_\mathcal{M}} \, p(\mathcal{D} \mid \theta, \mathcal{M})|_{\mathcal{D}=\mathcal{D}'} \, p(\theta \mid \mathcal{M}) \tag{4.20}$$

$\mathcal{L}_\mathcal{M}(\theta) \equiv p(\mathcal{D} \mid \theta, \mathcal{M})|_{\mathcal{D}=\mathcal{D}'}$ is the likelihood function, defined to be the probability of obtaining the data as a function of the parameters, $p(\theta \mid \mathcal{M})$ is the joint prior probability distribution of the parameters, and $\mathcal{Z}_\mathcal{M}$ is the the fully marginalized likelihood, which is also commonly called the Bayesian evidence:

$$\mathcal{Z}_\mathcal{M} = \int d\theta \, \mathcal{L}_\mathcal{M}(\theta) \, p(\theta \mid \mathcal{M}) = p(\mathcal{D} \mid \mathcal{M})|_{\mathcal{D}=\mathcal{D}'} \tag{4.21}$$

We obtain samples from the 13 to 16-dimensional posterior probability distribution using *emcee*[2] (Foreman-Mackey et al., 2013), a software package that implements an affine-invariant MCMC sampler (Goodman & Weare, 2010). The GP model is implemented using the package *george*[3] (Ambikasaran et al., 2015). Summary information about the posterior distribution, which is generally compact and approximately described by a multidimensional normal distribution, is stored in the form of a table of the sample means, the sample covariance matrices, and several quantiles of the one-dimensional marginal distributions (accounting for the burn-in phase of the sampler). For the purpose of model comparison, the Bayesian evidence is calculated as well, using the *MultiNest*[4] implementation (Feroz et al., 2009) of the nested sampling algorithm (Skilling, 2004) and its *PyMultiNest*[5] interface (Buchner et al., 2014). For finding best-fit (i.e. maximum posterior) parameter values, we use optimization routines from the *scipy*[6] package, specifically the *L-BFGS-B* algorithm.

4.3 Results

4.3.1 The flares of Sgr A*

Without yet assuming a particular noise model, the overall amount of astrometric noise due to unrecognized source confusion in the Galactic Center can be characterized by the effect on a test source located at the position of the black hole, at the center of the stellar cluster, where the expected rate of confusion events is highest. We choose a magnitude $m_K = 15$ source to represent a typical flare of Sgr A*, which are known to occur sporadically (e.g. Genzel et al., 2010). These near-infrared flares are also known to show significant astrometric jitter on a mas-scale around the location of the counterpart radio source, when they are by chance captured by observations (Gillessen et al., 2009b, 2017).

The apparent astrometric offset of the test source caused by source confusion in our simulations is the superposition of many separate offsets caused by the individual background stars, which may amplify each other, but may also cancel out partially. In Figure 4.4, we show how the total astrometric offset of the test source can be decomposed, for a number of different realizations of the background star population. The size of the induced individual offsets is also shown in Figure 4.5, as a function of the magnitude and separation of the background stars, in comparison to their number density in this parameter space. Even though the faintest stars dominate by number, the less numerous stars with magnitudes $17 \lesssim m \lesssim 19$ at separations $1 \lesssim d/\sigma_\phi \lesssim 2$ from the test source are almost solely responsible for the total induced offset, the overall probability distribution of which is shown in Figure 4.6. This total offset is typically on the order of 1 mas (or $0.07\,\sigma_\phi$), and only rarely is

---

[2] `https://github.com/dfm/emcee`

[3] `https://github.com/dfm/george`

[4] `https://ccpforge.cse.rl.ac.uk/gf/project/multinest/`

[5] `https://github.com/JohannesBuchner/PyMultiNest`

[6] `https://www.scipy.org/`

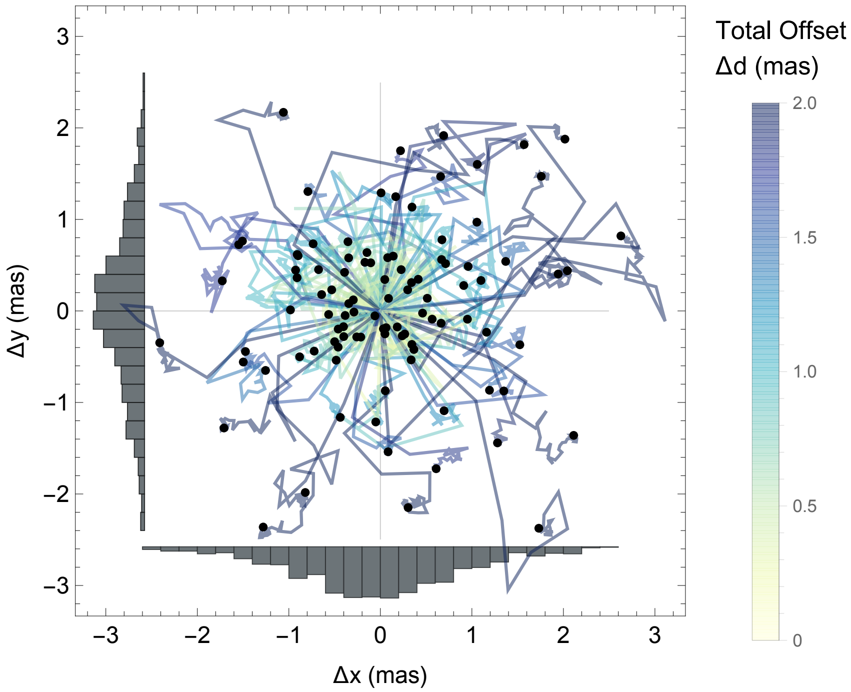

Figure 4.4: The total astrometric offsets of a magnitude $m_K = 15$ test source located at the position of Sgr A*, decomposed into successively smaller offsets caused by the individual background stars, where each line corresponds to a different realization of the background star population. Shown are 100 randomly chosen cases, together with the distributions of the total offsets along each axis, for a 10 times larger sample.

it more than a factor few larger, considering again many possible realizations of the background star population. Pushing the detection limit as deep as $m_{\mathrm{lim}} \approx 17.5$ would make these large, unrecognized offsets even more unlikely, but would only reduce the typical offset to 0.7 mas. In any case, these numbers only describe the minimum expected astrometric scatter, since recognized confusion events involving known stars are also common at the cluster center, and would cause additional offsets.

If we place the same magnitude 15 test source at increasing separation from Sgr A*, now representing one of the S-stars, astrometric offsets caused by unrecognized source confusion can be expected to become smaller and less frequent, as the density of background stars decreases. However, in agreement with the study by Fritz et al. (2010b), we find that at a separation of $0.2''$, the typical total offset of the test source is only slightly less than 0.4 mas, which is still comparable to the usual measurement uncertainty for a source of this brightness. Furthermore, the offsets occur preferentially in the direction of Sgr A*, due to the central over-density of background stars. In our test case, the average overall offset of the measured source position toward Sgr A* is itself as large as 0.4 mas at a source separation of $0.1''$, as shown in Figure 4.7.

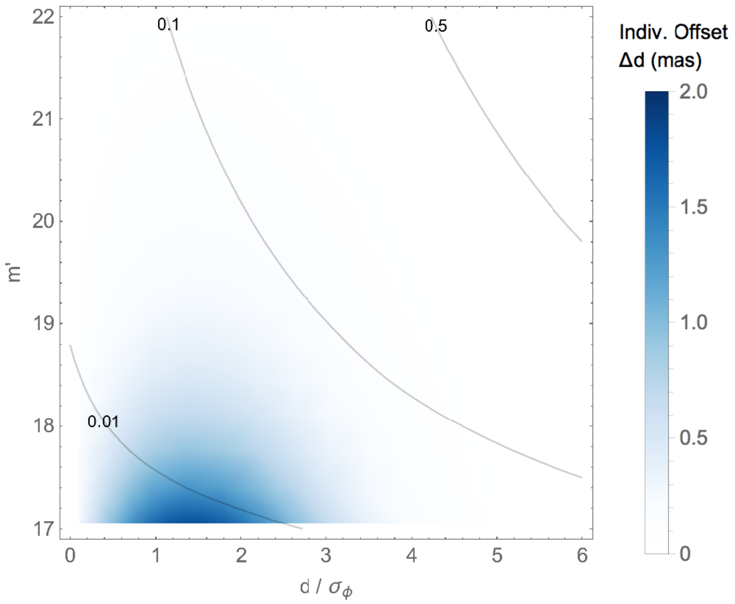

Figure 4.5: The size of individual astrometric offsets of a magnitude $m_K = 15$ test source located at the position of Sgr A* as a function of the magnitude and separation of the background stars (shown in color), and the cumulative fraction of background stars in this parameter space (shown in contours). The effect of the brightest stars dominates ($17 \lesssim m'_K \lesssim 19$), even though these stars are much less numerous than their even fainter counterparts ($19 \lesssim m'_K \lesssim 22$, see also Fig. 4.6).

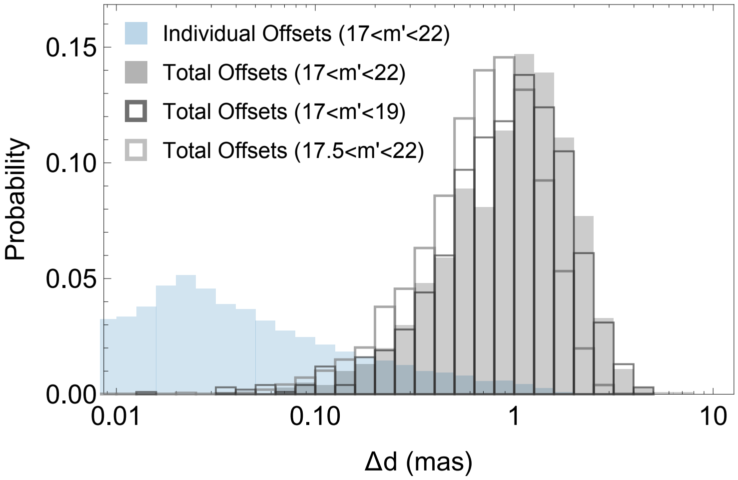

Figure 4.6: The probability distribution of individual astrometric offsets (shown in blue) and the total offsets (shown in gray) of a magnitude $m_K = 15$ test source located at the position of Sgr A*, considering many realizations of the background star population. The mode of the individual offsets is set by the maximum magnitude of the background stars ($m_{\rm lim}$). However, the total offset is almost solely determined by the brightest background stars, and has a typical value on the order of 1 mas in our simulations.

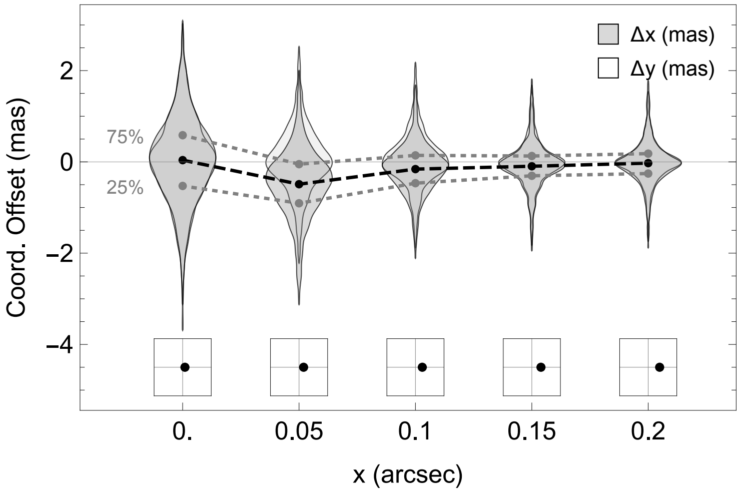

Figure 4.7: The distribution of total astrometric offsets of a magnitude $m_K = 15$ test source located at different projected (on-sky) distances x from Sgr A*, where $y = 0$, along the separation axis (Δx) and perpendicular to it (Δy), for many realizations of the background star population. The insets show the true position of the test source in a $1'' \times 1''$ field centered on Sgr A*. At small distances, both the amount of scatter and the overall offset toward the center are significant compared to the usual measurement uncertainty for a source of this brightness, but both decrease notably at distances beyond $0.1''$, due to the decreasing density of background stars.

4.3.2 The orbits of the S-Stars

The astrometric noise due to unrecognized confusion may thus contribute significantly to the noise budget of the S-stars and cause a bias in estimates of the orbital parameters, as well as the inferred mass and distance of the black hole, depending on the magnitude of each star, its on-sky trajectory, and the sampling of its orbit. The astrometric offsets for the star S2, resulting from our simulations, are shown in Figure 4.8 over a full orbital period, and the results of one example orbit fit are shown in more detail in Figure 4.9. During its pericenter passage, S2 approaches Sgr A* closely ($R_{\mathrm{min}} \approx 0.01''$), due to the large eccentricity of its orbit ($e \approx 0.88$, see also Fig. 4.3). At this time, the probability of confusion events is highest, and the astrometric offset of S2's measured position from its true position is likely to exceed the usual measurement uncertainty. The total offset at pericenter has a 50% chance of being in the range from 0.27 mas to 0.66 mas, and there is a remaining 7% chance of an offset to occur that is greater than 1 mas. These offsets can be persistent for a few months, but can also change quickly over the same time frame, since they are usually caused by only a few important confusion events, and all stars involved in these events are likely to have large proper motions.

However, the pericenter passage is relatively short and S2 spends the majority of its orbital period at considerably larger distances from the black hole ($R_{\mathrm{max}} \approx 0.19''$). At those times, the effect of confusion can be masked by the measurement uncertainty, due to

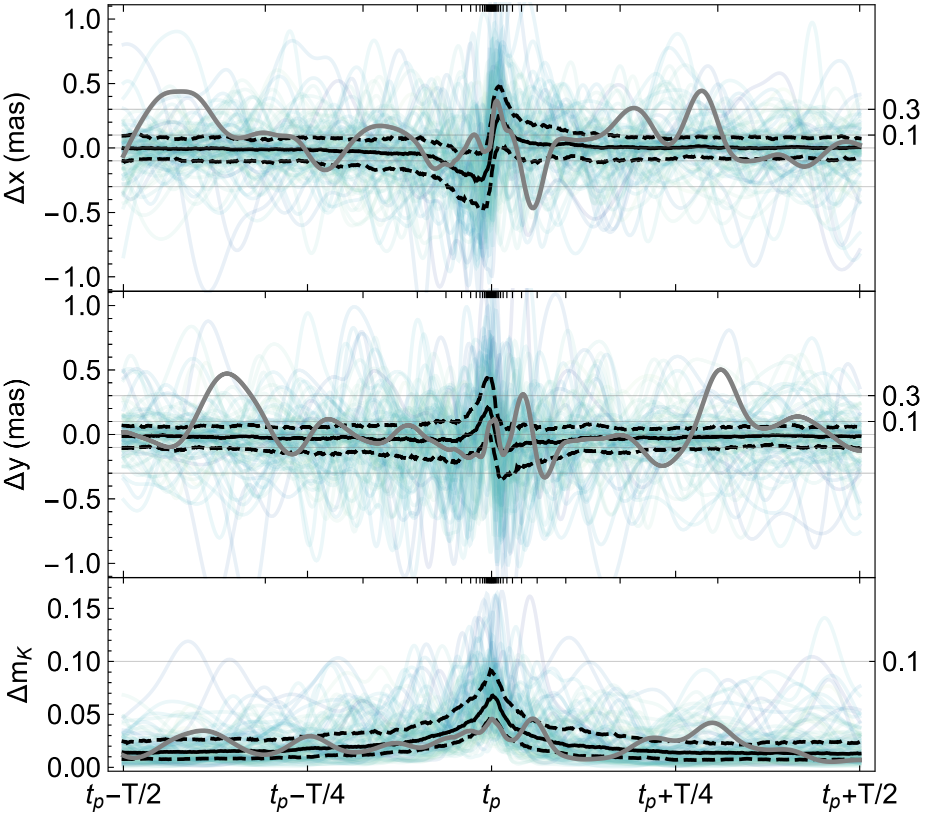

Figure 4.8: The expected offset in the astrometric coordinates of the star S2 due to unrecognized source confusion as it moves along its orbit (top panels), and its light curve (bottom panel). Each line corresponds to one of 100 different realization of the background star distribution, where line color is used as a proxy for the maximum total astrometric offset (see Fig. 4.4 for the color scale). As an example, one randomly chosen line is highlighted in gray. The solid black lines indicate the respective median values, and the dashed lines the 25% and 75% quantiles. During the pericenter passage, the astrometric offset is likely to exceed the usual measurement uncertainty for a considerable amount of time, and a simultaneous increase in brightness could potentially also be detectable.

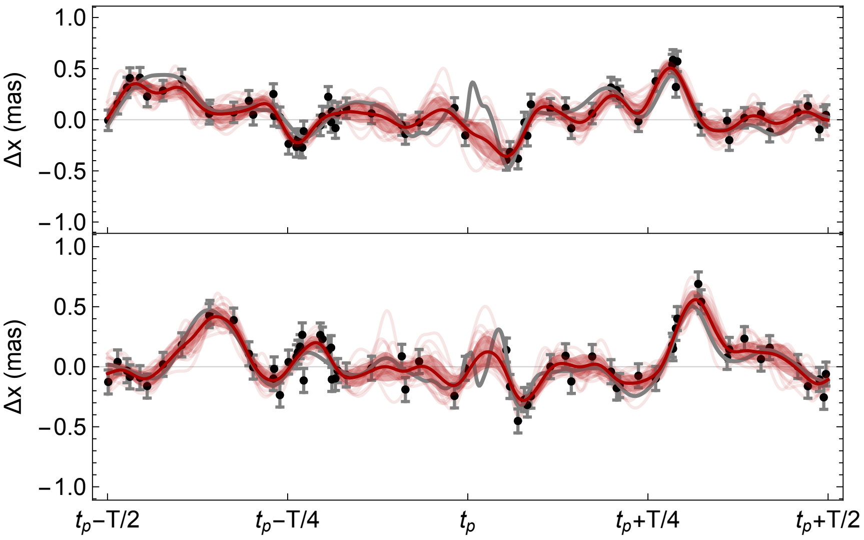

Figure 4.9: Results of a Gaussian process model fit to the motion of the star S2, based on a simulation corresponding to the example case highlighted in Fig. 4.8 (gray line). The black points show the generated mock measurements as residuals with respect to the true Keplerian orbit, where the assumed measurement uncertainty is $\sigma = 0.1\,\mathrm{mas}$, so that (unrecognized) source confusion dominates the astrometric noise budget. The perturbations caused by the confusion events are well accounted for by the noise model (red lines, same as in Fig. 4.13), except at pericenter, where there are too few measurements in this example.

the rapidly decreasing density of background stars with increasing distance from Sgr A*, combined with the exceptional brightness of S2. If relatively strong confusion events occur, however, they are then generally well sampled by observations. The overall confusion noise inferred using the added-noise or GP models in the simulations is around 0.2 mas on average, which is of similar order of magnitude as the typical measurement uncertainty, and is significantly correlated on a typical time scale of a few 0.1 years (see Fig. 4.10), which appears to be the average timescale of unrecognized confusion for S2.

The simultaneous brightening of S2 could be a potentially useful indicator for sufficiently strong confusion events, since it is likely to amount to a magnitude difference of a few 0.01 magnitudes around the time of pericenter, and a change as large as 0.13 magnitudes in cases of offsets as extreme as 1 mas. The measured K-band magnitudes of S2 show a typical scatter of about 0.1 magnitudes in the data set of Gillessen et al. (2009b, fig. 8), but an improved precision may still be achieved by optimizing the image analysis for photometry instead of astrometry (see e.g. Fritz et al., 2010b; Schödel, 2010), at least for a subset of high-quality images.

As a direct result of even the unrecognized source confusion, the inferred mass and distance of the black hole can be systematically biased, and the probability of an outlier measurement based on a single star and orbit is not negligible. Both parameters

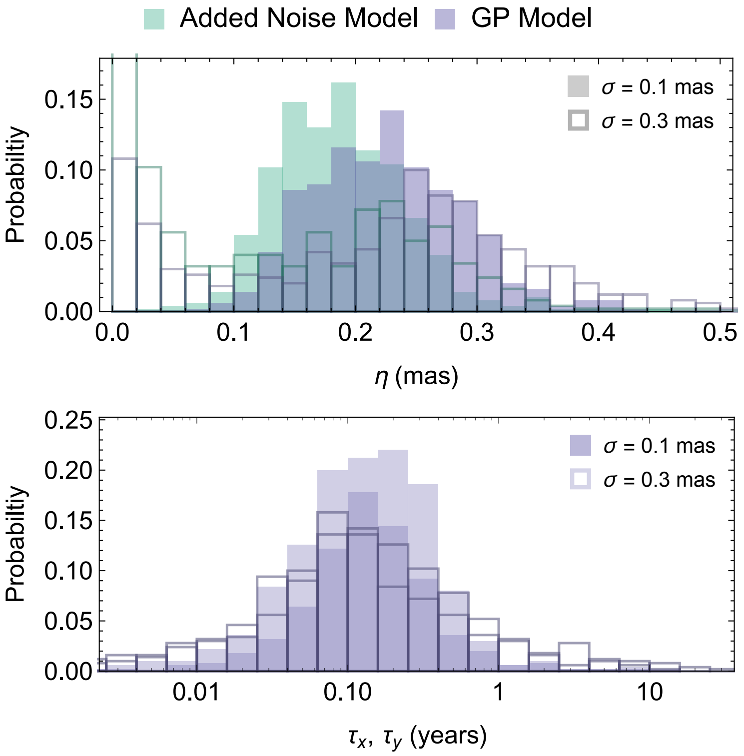

Figure 4.10: The mean values of the marginal posterior probability distributions for the hyper-parameters that describe the confusion noise, when fitting the orbit of the star S2 for many realizations of the background star population. Depending on the amount of measurement noise σ, the effect of confusion can sometimes be masked, resulting in small values of the noise amplitude η, but the estimated additional noise typically amounts to about 0.2 mas, and is correlated on time scales τ_x and τ_y of a few 0.1 years.

are strongly correlated because the data mainly constrains the gravitational parameter μ, through a measurement of the semi-major axis a in angular units and the orbital period T:

$$\mu \equiv \frac{M_0}{R_0^3} = \frac{4\pi^2}{G}\frac{a^3}{T^2} \tag{4.22}$$

The mass-distance degeneracy would be complete in the case of a purely astrometric data set, but it can be broken by measuring a star's radial velocity spectroscopically (in physical units) and comparing it to the star's proper motion (measured in angular units), in the context of an appropriate model (e.g. a Keplerian orbit). The typical joint posterior probability distribution of the black hole mass and distance, inferred from our simulations of the motion of S2, is shown in Figure 4.11, together with the distribution of the mean parameter values, considering once more many different realizations of the background star population. The properties of these distributions allow us to quantify the parameter uncertainties related to both the overall noise in the observations and our limited knowledge of the configuration of the background stars. The average bias in M_0 and R_0 is toward

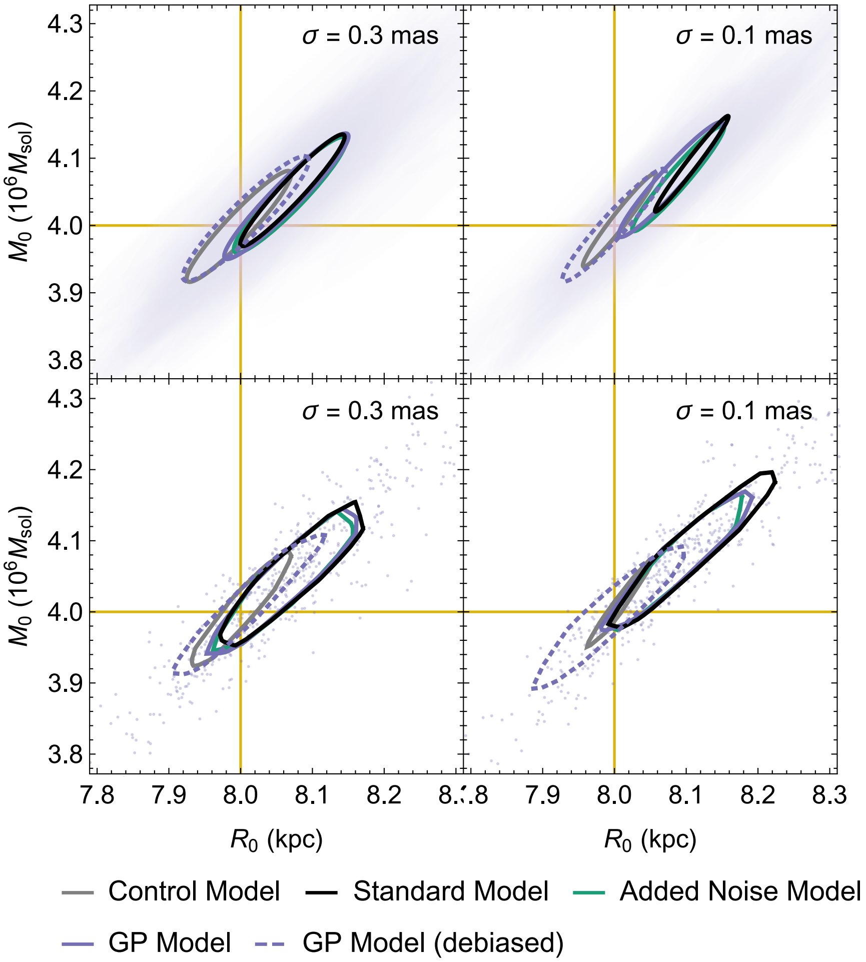

Figure 4.11: The typical joint posterior probability distribution of the black hole mass M_0 and distance R_0 (top panels), and the sampling distribution of the mean (bottom panels), when fitting the orbit of the star S2 for many realizations of the background star population, assuming the standard (left panels) or a reduced (right panels) measurement uncertainty σ. To highlight the effect of confusion, we have assumed zero-mean noise, except in the case marked as debiased, in which the expected average offset of S2 is used as the mean function (see Fig. 4.8). The colored contours indicate the 39.3% quantile of each distribution (1σ) for the different models tested in our simulations, and the straight horizontal and vertical lines indicate the assumed true parameter values.

larger values, and amounts to about $0.06 \times 10^6 M_\odot$ and 0.08 kpc. An estimate of the distance deviating from the true value that would, given only a measurement uncertainty of $\sigma = 0.3$ mas, have a probability of less than 1% to occur ($\Delta R_0 \gtrsim 0.2$ kpc), is predicted to be observed with a probability of about 15%, if source confusion is affecting the observations, but not accounted for in the noise model.

The jointly inferred orbital parameters are also affected by the source confusion (see Fig. 4.12). Perhaps most intuitively, the semi-major axis of S2's orbit is on average slightly underestimated ($\Delta a \approx -0.6$ mas), since astrometric offsets toward Sgr A* are somewhat more likely to occur at any time, due to the positive gradient of the surface density of background stars in that direction. This also explains the normally negative bias in the gravitational parameter ($\Delta\mu/\mu \approx -1.3\%$), since the orbital period must be unaffected ($T^2 \propto a^3/\mu$, see Eq. 4.22), as is on average the time of pericenter. The eccentricity of S2's orbit is on average slightly underestimated as well ($\Delta e \approx -0.001$) and the three orbital angles are affected similarly ($\Delta i \approx 12'$, $\Delta\Omega \approx -12'$, $\Delta\omega \approx -14'$). If an apparent shift of apocenter persists from one orbit to the next, it could complicate the prospective measurement of Schwarzschild precession, the dominant post-Newtonian effect on the orbit of S2 with an impact on astrometry. Of the coordinate system parameters, the location of the central mass is biased in the north-south direction, along which the orbit is elongated ($\Delta y_0 \approx 0.3$ mas).

The average bias in all parameters is reduced by using the expected average astrometric offset of S2 as the mean function in the noise model (see Fig. 4.8), which changes along the orbit but has to be computed only once from simulations (assuming a certain average distribution of background stars), since the orbital trajectory can be sufficiently well determined. The added-noise and GP models yield larger parameter uncertainties as well, which would otherwise be underestimated, as would be the total noise, so that any bias is further reduced in units of the parameters' standard errors.

For the purpose of model comparison, we calculate the evidence ratios for different combinations of models (see Eq. 4.21), assuming that a priori $p(\mathcal{M}_1) = p(\mathcal{M}_2)$, so that:

$$\frac{\mathcal{Z}_1}{\mathcal{Z}_2} = \left.\frac{p(\mathcal{D} \mid \mathcal{M}_1)}{p(\mathcal{D} \mid \mathcal{M}_2)}\right|_{\mathcal{D}=\mathcal{D}'} = \frac{p(\mathcal{M}_1 \mid \mathcal{D}')}{p(\mathcal{M}_2 \mid \mathcal{D}')} \tag{4.23}$$

In terms of these ratios, the GP model is favored by the simulated data in at least 90% of cases, but with respect to fitting the orbit of S2 the practical difference between the GP and the added-noise model appears to be minor. However, including the additional parameters can be justified statistically, since the average evidence ratio is $1 \lesssim \log_{10}(Z_{\mathrm{GP}}/Z_{\mathrm{add}}) \lesssim 10$, depending on the assumed measurement uncertainty ($0.3\,\mathrm{mas} > \sigma > 0.1\,\mathrm{mas}$). In comparison to the standard model, both the added-noise and the GP models are strongly favored.

Several other stars, for example S1, S9 or S13, also have well-measured orbits that constrain the gravitational potential independently, at least in principle (Gillessen et al., 2017). These stars are 0.7 to 1.8 magnitudes fainter than S2, and therefore potentially more affected by source confusion. However, at the moment, these stars are likely unaffected by unrecognized confusion, since, in contrast to S2, they have not entered the very central

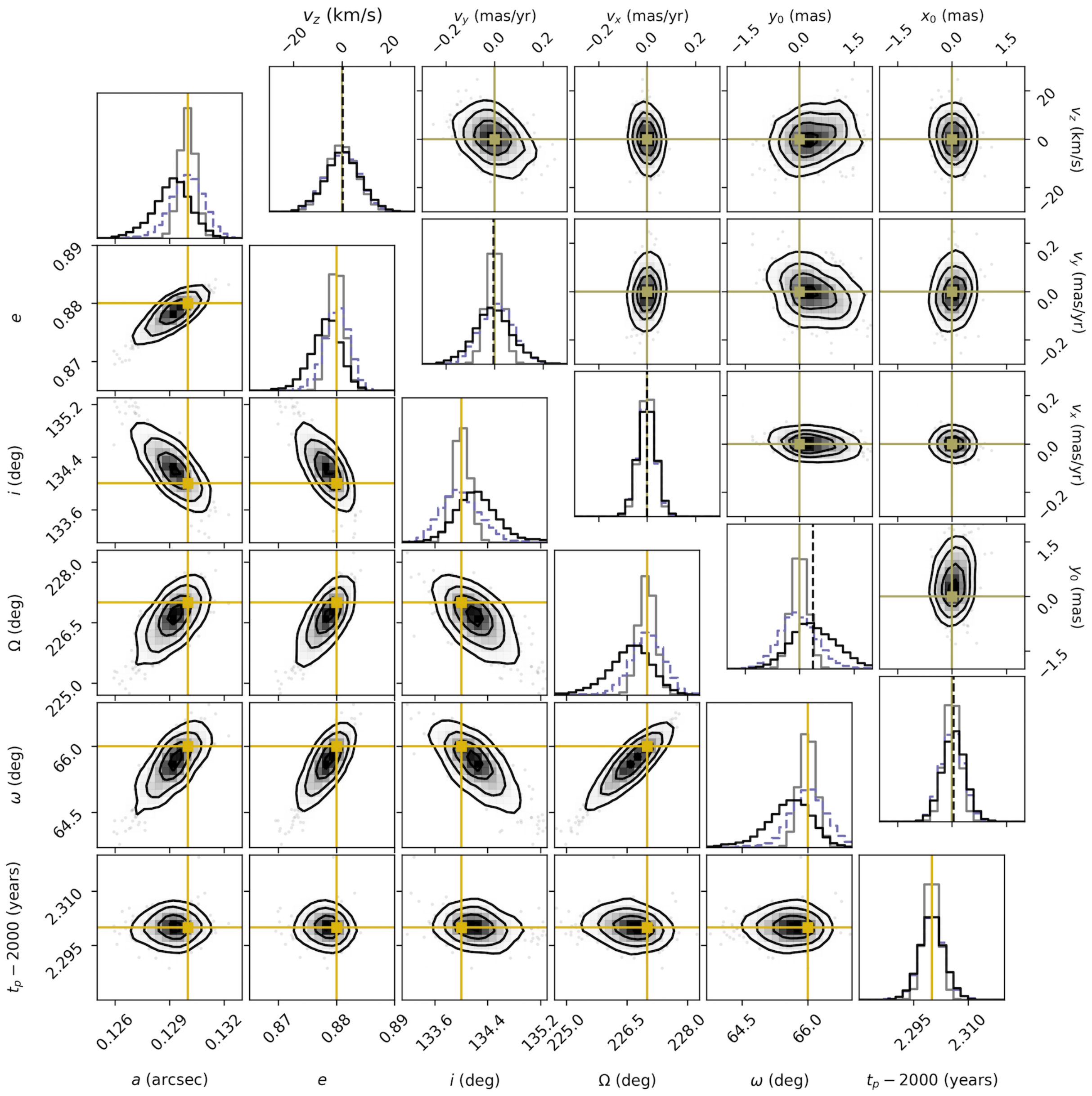

Figure 4.12: The mean values of the joint posterior probability distribution of S2's orbital parameters (lower left) and the coordinate system parameters (upper right) inferred from our simulations, for many possible realizations of the background star population, when using the standard orbit model to fit the motion of S2, i.e. when source confusion is not accounted for in the noise model (black lines and contours). Also shown in the one-dimensional histograms are results from the fits of the debiased GP model (dashed purple lines), as well as the control model (gray lines, same as in Fig. 4.11). For all orbit fits shown here, we have assumed a generic measurement uncertainty of $\sigma = 0.1$ mas, so that (unrecognized) source confusion dominates the astrometric noise budget. The assumed true values of the parameters are indicated by the straight horizontal and vertical lines.

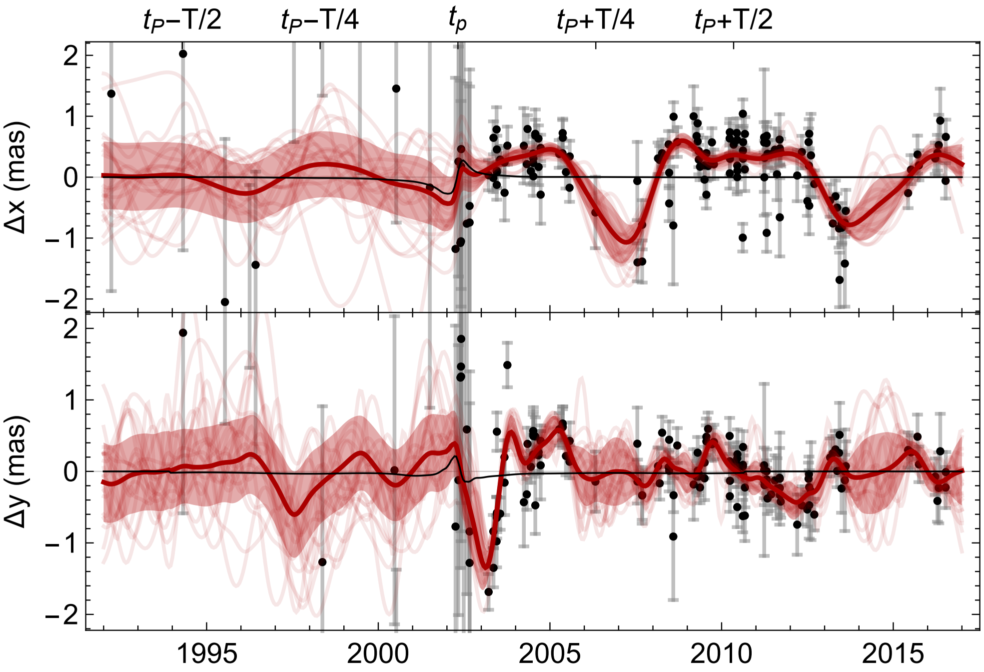

Figure 4.13: Results of a Gaussian process model fit to the motion of the star S2, based on the observational data set provided by Gillessen et al. (2017, tab. 5). The black points show the astrometric residuals with respect to the best-fit Keplerian orbit. The thin red curves are randomly sampled from the posterior distribution of the inferred noise parameters, and show how the model accounts for correlations in the astrometric noise due to potential confusion events. The thick red line and error band corresponds to the posterior mean and its uncertainty (1σ). The thin black lines show the expected average perturbation due to unrecognized source confusion, which is included in the orbit fit.

region around Sgr A* during the time over which they have been observed, but have been confined to a region where the density of background stars is much lower (see Fig. 4.3). The orbits of even fainter stars, for example S38, S55 or S175 are the most uncertain, and usually more strongly affected by recognized confusion events than unrecognized confusion, whenever they are close to Sgr A*.

4.3.3 Application to observations

As a proof of concept, we also fit our models to available observational data. The extensive S-star data set provided by Gillessen et al. (2017, tab. 5) includes 145 astrometric and 44 radial velocity measurements for S2, which have been collected as part of different observing programs, but mostly at the VLT. For a detailed description of these observations and the data reduction procedure, we refer the reader to Gillessen et al. (2017) and the references therein. Prior to the orbit fitting, we have reversed the error upscaling applied by Gillessen et al. (2009b, 2017), to retrieve the astrometric uncertainties originally estimated from the

	Added-Noise Model	GP Model (preferred)
M_0 ($10^6 M_\odot$)	4.50 ± 0.18	4.45 ± 0.21
R_0 (kpc)	8.44 ± 0.16	8.53 ± 0.19
a (arcsec)	0.1249 ± 0.0009	0.1228 ± 0.0012
e	0.882 ± 0.002	0.879 ± 0.003
i	$134.8^\circ \pm 0.4^\circ$	$135.4^\circ \pm 0.5^\circ$
Ω	$225.8^\circ \pm 0.5^\circ$	$226.6^\circ \pm 0.7^\circ$
ω	$64.5^\circ \pm 0.5^\circ$	$64.8^\circ \pm 0.7^\circ$
t_p	2002.32 ± 0.01	2002.32 ± 0.01
x_0 (mas)	-0.15 ± 0.40	0.26 ± 0.74
y_0 (mas)	-1.11 ± 0.55	0.44 ± 0.82
v_x (mas/yr)	-0.078 ± 0.041	-0.084 ± 0.051
v_y (mas/yr)	-0.024 ± 0.061	0.050 ± 0.076
v_z (km/s)	32 ± 7	24 ± 8
η (mas)	0.34 ± 0.03	0.50 ± 0.13
τ_x (yr)		0.78 ± 0.28
τ_y (yr)		0.35 ± 0.08

Table 4.2: Best-fit values of the added-noise and the (debiased) Gaussian process model parameters, when fitting the orbit of the star S2 using the data provided by Gillessen et al. (2017, tab. 5, but see Sec. 4.3.3).

AO imaging data. We assume that any additional astrometric scatter is due to source confusion, and as such accounted for by our noise models.

The best-fit parameters are summarized in Table 4.2, and the astrometric residuals for the fit of the (debiased) GP model are shown in Figure 4.13. The additional confusion noise inferred from this fit has a fairly large amplitude of $\eta \approx 0.50\,\mathrm{mas}$, and an especially long timescale of $\tau_x \approx 0.8$ years. These values indicate that the model is accounting for astrometric perturbations due to a few pronounced, separately identifiable confusion events, rather than a series of weaker events involving undetected stars, as can be seen in Figure 4.13. In 2002-2003 the astrometric measurements of S2 could have likely been perturbed by S19, in 2006-2007 by S13, and in 2013 by S56, while other fainter stars and the variable infrared source Sgr A* could have contributed as well to the overall confusion noise. The fit of the added-noise model results in a similarly large noise amplitude, yet even without explicitly accounting for the motions of any other stars (known or undetected), the GP model is able to describe the observed outlier measurements of S2 more convincingly, and is also favored in terms of the evidence ratio ($\log_{10}(Z_{\mathrm{GP}}/Z_{\mathrm{add}}) \approx 10$). The inferred values for the black hole mass and distance are larger by about $0.10 - 0.35 \times 10^6 M_\odot$ and $0.20 - 0.40\,\mathrm{kpc}$, respectively, when comparing the GP model fit to the various best-fit values of Gillessen et al. (2017, tab. 1). This deviation is significant relative to the statistical uncertainties, and underlines the importance of the systematic effect that any kind of source confusion can have on measurements of stellar motions in a field as crowded as the Galactic Center.

4.4 Conclusions

We have simulated long-term imaging monitoring observations of individual stellar motions in the Galactic Center under idealized conditions, to study in isolation the properties of astrometric noise arising from unrecognized source confusion, and specifically the ramifications for estimating the mass and distance of the central black hole by fitting stellar orbits. We emphasize that it is critical to understand the data generation process and the noise properties in particular, to be able to confidently measure any astrometric signatures of deviations from Keplerian orbits, for instance the predicted post-Newtonian deviations following the pericenter passage of the star S2 in 2018.

As the main non-instrumental source of astrometric noise, unrecognized source confusion accounts for a fundamental part of the S-stars' noise budget, and can even have a significant effect on the apparent motion of a star as bright as S2. Due to the nature of source confusion in the inner nuclear star cluster, the resulting noise is temporally and spatially variable, the background stars being in constant motion themselves, yet concentrated around Sgr A*. Exceptionally large astrometric offsets can occur in particular during pericenter passages of S2 (or during any close enough approach of another star), and can bias the black hole mass and distance inferred from orbital motions. Recognized confusion events, as well as potential confusion with the variable infrared source Sgr A*, add further complication to the data analysis.

The bias induced by unrecognized confusion in the black hole mass and distance can be reduced by excluding or down-weighting astrometric measurements made at and around the time of pericenter. To otherwise reduce this bias, it is necessary to account for the non-zero mean of the confusion noise, which arises from the expected central over-density of background stars. We would advocate to incorporate even a non-specific noise model, parameterizing additional noise in some justified form, if there is a possibility that the measurement uncertainty has been underestimated. Thus estimating any additional uncertainty directly from the data will generally yield more trustworthy estimates of the parameter uncertainties, since the model for the data itself (a Keplerian orbit) is very well motivated, at least so far. This approach is also a statistically robust way to account for systematic uncertainties unrelated to confusion, for example residual image distortion. With respect to confusion specifically, a noise model based on GPs has the advantage of being able to describe time-correlated noise, and is usually favored over a simpler added-noise model in our simulations, demonstrating a need for improved noise models.

Since S2 is expected to be involved in confusion events more frequently the closer it approaches Sgr A*, it could be advantageous or even necessary to use a generalized kernel function with a varying timescale $\tau_{x,y}(t)$ to describe confusion noise in a GP model, e.g. Gibbs' function (Rasmussen & Williams, 2006), if the motion of S2 is well sampled by observations during the time of closest approach:

$$k_{x,y}(t_i, t_j) = \eta^2 \left(\frac{2\tau_{x,y}(t_i)\, \tau_{x,y}(t_j)}{\tau^2_{x,y}(t_i) + \tau^2_{x,y}(t_j)} \right)^{\frac{1}{2}} \exp\left(-\frac{(t_i - t_j)^2}{\tau^2_{x,y}(t_i) + \tau^2_{x,y}(t_j)} \right) \tag{4.24}$$

For a constant timescale $\tau_{x,y}(t) = \tau_{x,y}$, Gibbs' function reduces to the exponential squared kernel (Eq. 4.19), but $\tau_{x,y}(t)$ may be any positive function. For instance, to model a decrease in the confusion timescale around the time of pericenter, i.e. from $t_p - \Delta t_p$ to $t_p + \Delta t_p$, it would be straightforward to set

$$\tau_{x,y}(t) = \tau^{(\max)}_{x,y} + \left(\tau^{(\min)}_{x,y} - \tau^{(\max)}_{x,y}\right) \exp\left(-\frac{(t - t_p)^2}{2\Delta t_p^2} \right), \tag{4.25}$$

where $0 < \tau^{(\min)} < \tau^{(\max)}$, or to choose a function directly proportional to the projected separation of S2 from Sgr A* (or some power of it). GPs may also be used to model specific, marginally recognized confusion events that apparently dominate the observed astrometric scatter of S2, so that fewer data points would have to be excluded manually (and perhaps in doubt), and to reduce any additional biases caused by these events. However, a GP model could eventually become prohibitively expensive computationally, if the number of data points will continue to grow steadily.

Although it could be possible in principle, we conclude that it is currently not feasible to deduce properties of the background star population by analyzing the properties of confusion noise, even if recognized confusion events are perfectly accounted for, unless the instrumental noise can be suppressed and all other noise processes can be thoroughly characterized. Amongst other effects, real perturbations to the orbit of S2 due to the gravitational influence of the background stars would then need to be considered as well (e.g. Sabha et al., 2012). Also, if one would want to infer the parameters of the background star population as part of the orbit fitting, using a direct forward-modeling approach, the uncertainty about the exact dynamical configuration of the background stars would make a (very inefficient) marginalization over many hyper-parameters necessary, namely the initial conditions of the background stars.

Finally, as the monitoring of stellar motions in the Galactic Center continues, we would advocate not to rely on a single orbit of a single star for inference, but to combine measurements of many stars over many years that are affected differently by confusion and fit their orbits simultaneously, to obtain more accurate results (see e.g. Boehle et al., 2016; Gillessen et al., 2017).

Acknowledgements

We are grateful to D. Foreman-Mackey for helpful discussion at the 11th IMPRS Summer School on Astrostatistics and Data Mining in Heidelberg, as well as to S. Gillessen for providing feedback on the paper draft.

Chapter 5

A machine-learning approach to photometric stellar classification

Original publication: P. M. Plewa, 2018, MNRAS, *Random Forest Classification of Stars in the Galactic Centre*, doi:10.1093/mnras/sty511

Abstract: Near-infrared high-angular resolution imaging observations of the Milky Way's nuclear star cluster have revealed all luminous members of the existing stellar population within the central parsec. Generally, these stars are either evolved late-type giants or massive young, early-type stars. We revisit the problem of stellar classification based on intermediate-band photometry in the K-band, with the primary aim of identifying faint early-type candidate stars in the extended vicinity of the central massive black hole. A random forest classifier, trained on a subsample of spectroscopically identified stars, performs similarly well as competitive methods ($F_1 = 0.85$), without involving any model of stellar spectral energy distributions. Advantages of using such a machine-trained classifier are a minimum of required calibration effort, a predictive accuracy expected to improve as more training data becomes available, and the ease of application to future, larger data sets. By applying this classifier to archive data, we are also able to reproduce the results of previous studies of the spatial distribution and the K-band luminosity function of both the early- and late-type stars.

5.1 Introduction

The dense nuclear star cluster of the Milky Way has been observed and monitored for many years at near-infrared wavelengths, being highly extincted in the visible spectral range. Today's routine, ground-based observations at high angular resolution, assisted by AO systems, reveal all luminous members of the existing stellar population, the composition of which has been the focus of numerous previous studies (see e.g. Genzel et al., 2010).

Within the central parsec, the majority of detected stars are old, evolved giants that likely formed at the same time as the Galactic bulge (e.g. Pfuhl et al., 2011). However, more recently formed main-sequence stars are also detected, which are furthermore split into different sub-populations. A significant fraction of the most massive young (WR/O-)stars reside in a disc structure (Paumard et al., 2006; Lu et al., 2009; Bartko et al., 2009; Yelda et al., 2014), while an apparently isotropically distributed 'S-star' cluster of less massive (B-)stars is concentrated around the central massive black hole (e.g. Ghez et al., 2008; Gillessen et al., 2009b; Boehle et al., 2016; Gillessen et al., 2017), identified with the compact radio source Sgr A* (e.g. Reid et al., 2007; Plewa et al., 2015). So far, only few B-stars have been identified further out from the black hole, and it is unclear whether these belong to the stellar disk, the S-stars, or form a distinct population (e.g. Madigan et al., 2014).

To better understand the complex history of the nuclear star cluster in general, and that of the young stars in particular, it is of fundamental interest to study and characterize these different stellar populations, for example by their luminosity function, spatial distribution or kinematics. In this study, we revisit the key problem of identifying stars as members of the young or old population, by exploring alternative, machine-learning techniques for determining their spectral types from images.

The high level of stellar crowding in the Galactic Center demands the use of integral field spectroscopy to achieve a definitive spectral classification of individual stars (e.g. Ghez et al., 2003; Eisenhauer et al., 2005; Do et al., 2009; Bartko et al., 2010; Do et al., 2013a; Støstad et al., 2015; Feldmeier-Krause et al., 2015). However, with respect to covering a large field of view to a sufficiently high depth in a reasonably small amount of observing time, this technique remains inefficient in comparison to imaging techniques. It is therefore of practical interest to develop accurate methods of photometric classification (e.g. Genzel et al., 2003; Buchholz et al., 2009; Nishiyama & Schödel, 2013), in particular to identify the rare young, early-type candidate stars at faint magnitudes in the extended vicinity of the massive black hole. For confirmation, these stars may later be targeted sequentially in deep spectroscopic follow-up observations, of which coverage is still lacking in off-center fields.

In section 5.2, we present the intermediate-band imaging observations that allow us to determine spectral types of a few thousand stars in the nuclear cluster, before we describe the specific classification method we have used in section 5.3. In section 5.4, we discuss the achieved classification performance, estimate surface density profiles of the early- and late-type stars, as well as their luminosity functions, and compare our results to those of other studies. Finally, we present our conclusions and discuss future opportunities in section 5.5.

5.2 Observations & Data reduction

The data set we make use of is a subset of the one previously analyzed by Buchholz et al. (2009), which was obtained in the year 2004 with the NACO imager at the ESO VLT, and is publicly available in raw form through the ESO archive (see Tab. 5.1). We have re-reduced all available images using methods described in detail by Plewa et al. (2015).

Date (UT)	λ_c (μm)	$\Delta\lambda$ (μm)	N	NDIT	DIT (s)	Importance
2004-07-08	2.00	0.06	8	4	36	4%
2004-06-11	2.06	0.06	96	1	30	5%
2004-06-11	2.24	0.06	99	1	30	26%
2004-07-08	2.27	0.06	8	4	36	21%
2004-07-08	2.30	0.06	8	4	36	5%
2004-06-12	2.33	0.06	120	1	30	27%
2004-07-08	2.36	0.06	8	4	36	13%

Table 5.1: Summary of NACO observations: For each date, we list the central wavelength (λ_c) and width ($\Delta\lambda$) of the filter used, the number of frames (N) obtained and their exposure times (DIT), as well as the estimated Gini importance of the respective photometry (see Sec. 5.4.1).

The image reduction process includes a standard sky-subtraction, flat-field and bad pixel corrections, as well as corrections for optical distortion and differential refraction, and a precise (sub-pixel) image registration.

The end result of this process are seven combined images, one for each of the intermediate-band filters used, with a field of view of $40'' \times 40''$ roughly centered on Sgr A*. The average FWHM of the PSF is 90 mas. The two shortest filters provide a continuous coverage between wavelengths of 1.97μm and 2.09μm, and the remaining five filters are interleaved to cover a spectral range from 2.21μm to 2.39μm (see Fig. 5.1). We have also inspected additional images taken with filters centered on wavelengths of 2.12μm and 2.18μm as part of the same observing program, but decided not to use them. The images taken with the former filter are of lower quality, due to poor weather conditions, whereas those taken with the latter filter are affected by recombination line emission of ionized gas in the mini-spiral streamer (Brγ). Unlike Buchholz et al. (2009), we do not use any H-band data, which is only available for a significantly more restricted field of view.

5.3 Methods

The main spectral signature in the near-infrared K-band that facilitates a distinction between the late- and early-type stars detectable in the Galactic Center, and allows their classification based on limited photometric information only, are distinct CO absorption features (see Fig. 5.2 & 5.3). These features start to appear in the spectra of late G-type giants and become increasingly pronounced in giants of spectral types K and M. In contrast, O- and B-type main-sequence stars show an almost featureless black-body spectrum with only a few narrow, weaker absorption lines.

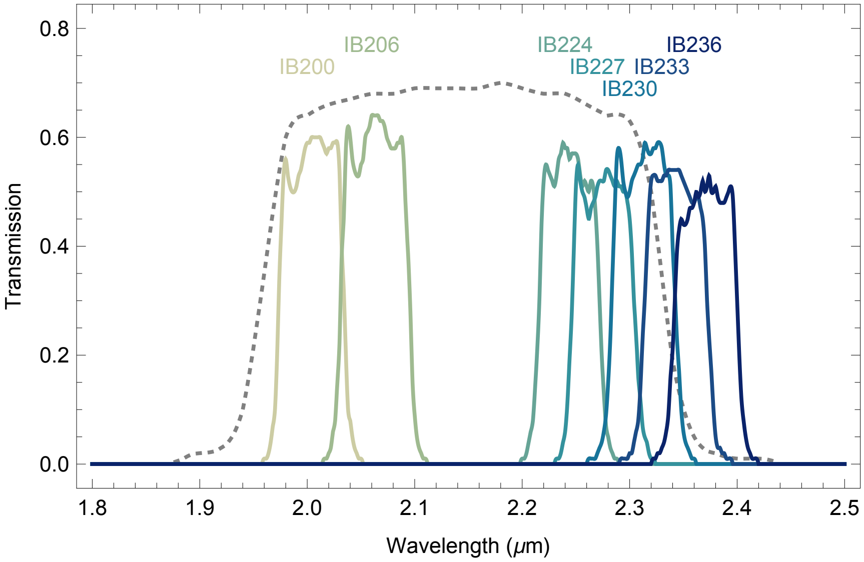

Figure 5.1: Transmission curves of the seven intermediate-band filters used in this study, in comparison to that of the broad-band Ks filter (see `http://www.eso.org/sci/facilities/paranal/instruments/naco/inst/filters.html`).

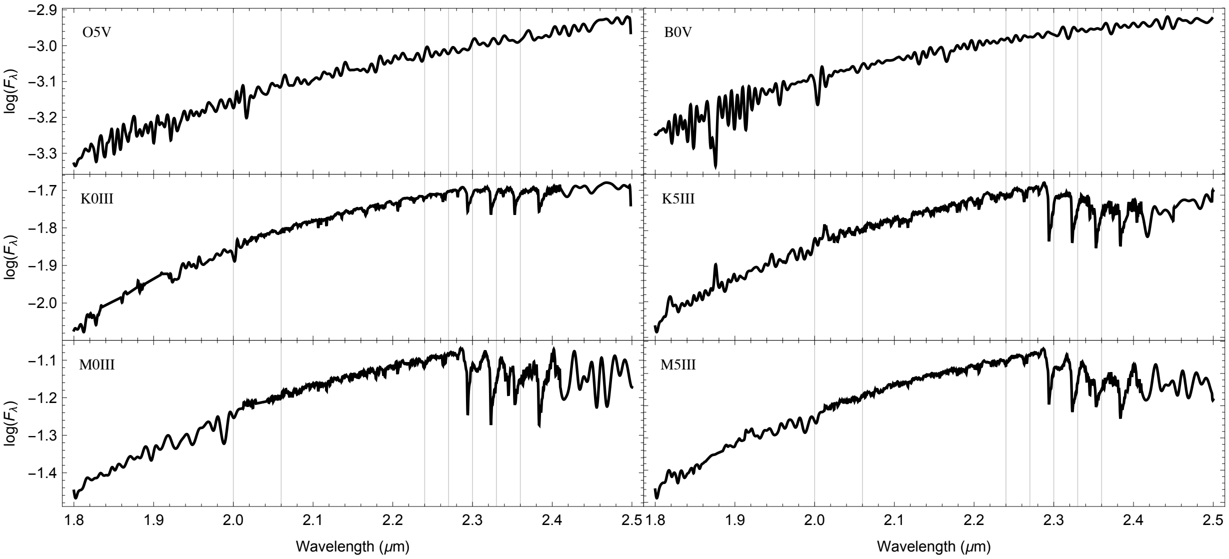

Figure 5.2: Examples of high-resolution K-band spectra of early-type (top row) and late-type (bottom rows) stars (taken from the Pickles (1998) Atlas), which are detectable in near-infrared observations of the Galactic Center. The characteristic CO absorption features that appear in the spectra of the late-type stars allow a distinction between the two classes based on intermediate-band photometry, by sampling the stellar spectra at a few discrete points only (vertical lines, see also Fig. 5.3). To account for reddening, we have used the extinction law for the Galactic Center derived by Fritz et al. (2011).

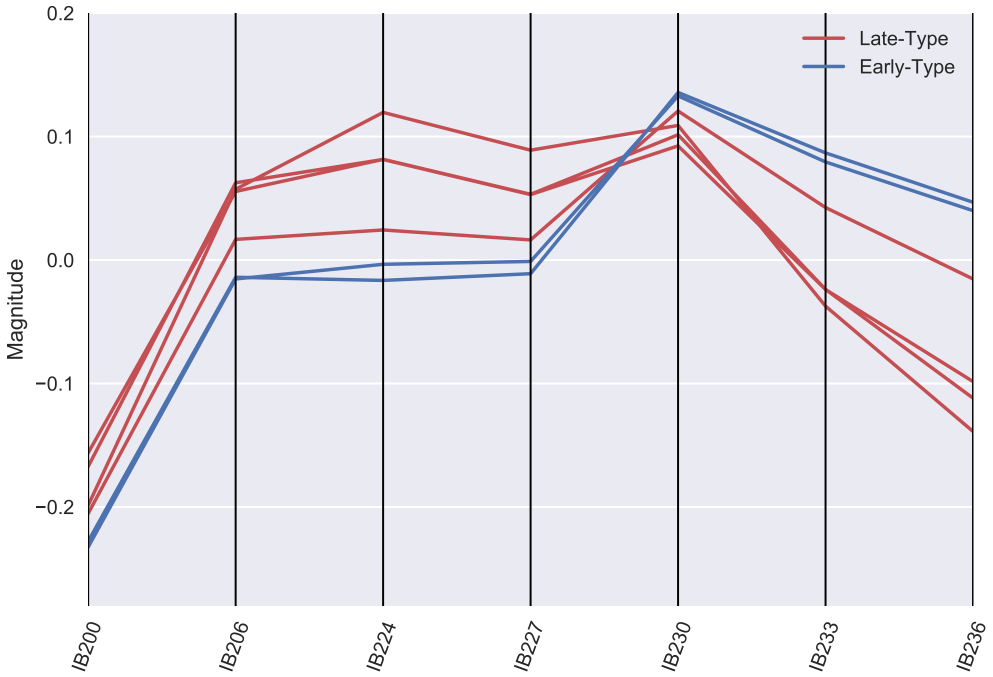

Figure 5.3: A characteristic difference in the shape of the K-band spectrum mainly caused by CO absorption features allows a distinction between late- and early-type stars in observations of the Galactic Center, based on intermediate-band photometry in the K-band using the seven indicated filters, instead of a high-resolution spectrum (see also Fig. 5.2).

5.3.1 Photometry

The quality of the photometry can be expected to have a strong impact on the ultimate performance of our stellar classifier. To determine as accurately as possible the shape of each star's spectrum over the K-band, we perform PSF photometry on the seven reduced intermediate-band images using the *StarFinder* tool (Diolaiti et al., 2000). The *StarFinder* algorithm is optimized to detect stars in a crowded field and provide accurate astrometry and photometry, when aperture photometry would fail, for instance. Besides an image, the required input for the algorithm are the positions of a number of manually chosen PSF reference stars, which should ideally be relatively bright, isolated and scattered across the field of view. For each image, the final output of the algorithm is an estimate of the PSF and a star list containing the positions and instrumental magnitudes of all detected stars. We cross-match these star lists and keep only the sources detected in all seven bands, thereby removing spurious detections of faint stars. In total, we identify 3165 sources in the field of view, with K-band magnitudes ranging from 9.2 to about 16.1, and a high completeness for sources brighter than magnitude 15.7, at least outside the very crowded central 1″.

The distinguishing features between the early- and late-type stars that we aim to isolate are imprinted in the spectral shape of a star, and not necessarily dependent on its overall brightness. Therefore, to remove the latter information, we subtract the average value from the measured magnitudes of each star (i.e. we divide by the average flux), and in the following refer to the so-standardized multi-band photometry as a star's spectral energy distribution (SED).

A few tens of extremely bright stars in the field of view are affected by saturation in one or several bands, and a few of the prominent super-giants and WR-stars in the field are affected severely. A repair of saturated PSF cores is implemented in the *StarFinder* algorithm, but the missing data may result in an increased, possibly systematic uncertainty of the SED of any saturated star. Such stars are nevertheless suitable PSF reference stars, since they provide valuable information about the extended PSF wings.

We expect another systematic uncertainty in the stellar SEDs, which is in part specific to AO-assisted observations and an inevitable consequence of the spatial and temporal variability of the PSF. The spatial variability arises due to anisoplanatism, which causes the AO correction to deteriorate at separations larger than about $10''$ to $20''$ from the AO natural guide star (IRS 7), as a result of differential tilt jitter (e.g. Fritz et al., 2010b). The temporal variability arises due to changing observing conditions and performance of the AO system, from one night to another, as well as within a night. In our photometric data, the resulting effect appears similar to an intrinsic variation of the extinction across the field of view, which also exists (e.g. Schödel et al., 2010; Fritz et al., 2011). To mitigate these effects, while continuing to use a constant PSF model, we derive an empirical, local photometric calibration following a strategy similar to that of Buchholz et al. (2009).

This local calibration relies on the fact that the early-type stars are rare in comparison to the late-type stars. We can therefore select a group of nearest neighbors around each star and use their average SED for reference at that position, such that a typical late-type star will have an approximately flat SED everywhere in the field of view. To further avoid selecting early-type stars, we only consider stars in the magnitude range from 14.5 to 16 in K-band with a minimum projected distance of $1''$ from Sgr A* as reference sources, which are predominantly red clump stars (that produce a bump in the luminosity function at these magnitudes). Most of the excluded stars inside the central region are known to be members of the predominantly young S-star cluster (e.g. Gillessen et al., 2017).

The necessary magnitude correction for each wavelength band does not affect the classification of any specific star directly, since the features of its particular SED are preserved, if only relative to the local average SED. We find that selecting a number of 20 neighboring reference stars is sufficient, which are typically distributed over a $1.7''$ circular surrounding area. The reference stars are generally not distributed evenly within that area, but any discontinuity in the calibration maps is avoided (see Fig. 5.4). After this calibration, the classification accuracy should not depend on a star's position in the field of view.

5.3.2 Classification

We choose a machine-trained random forest classifier, which is a meta-classifier based on an ensemble of decision trees. For in-depth information about the algorithm and details about the implementation, we refer the reader to Breiman (2001) and Pedregosa et al.

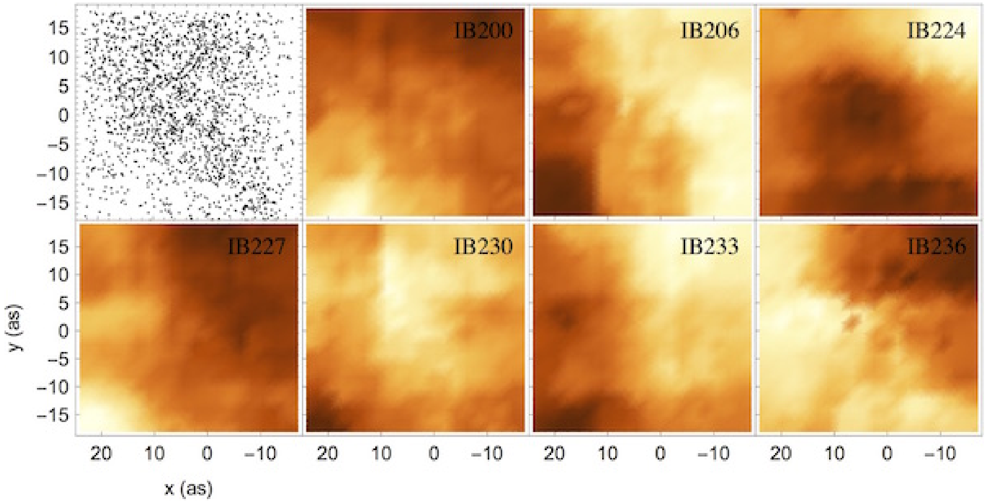

Figure 5.4: Local calibration maps for the photometry are used to account for residual systematic features, which are reminiscent of the four-point dither pattern, as well as the spatial variability of the extinction. The first panel shows the surface density of the reference sources (see Sec. 5.3.1).

(2011, see also `http://scikit-learn.org/`).

A decision tree has several advantages as a method of classification. It can essentially be reduced to a set of rules that, in our case, can be used to predict the class (i.e. spectral type) of a star from its SED, or to obtain class membership probabilities, which are straightforward to interpret. Generally, there is also little data preparation required, for example neither rescaling nor feature selection, which is performed implicitly. In our case, neither an absolute nor even a relative magnitude calibration for the individual wavelength bands is strictly necessary, due to the mentioned scaling invariance, and because the training and test sets used for fitting and evaluating the classifier are subsets of the same data set. Furthermore, the computational cost of performing a classification using a decision tree does not depend strongly on the size of the training set (but scales logarithmically). The main disadvantage of a decision tree is a susceptibility to over-fitting and instability with respect to changes in the training set. This is mitigated by constructing an ensemble of decision trees, for instance a random forest, where multiple trees are fit to random subsamples of the data and results are averaged, to improve the robustness and the overall predictive accuracy of the classification.

It is important to ensure that the stars included in the training set have representative SEDs for each class and that their classifications are indeed correct. We join a sample of early-type stars reported by Yelda et al. (2014), which includes data originally published by Paumard et al. (2006); Bartko et al. (2009); Do et al. (2009) and Do et al. (2013a), and a sample of late-type stars reported by Maness et al. (2007). These stars were manually classified using high-resolution near-infrared spectroscopic data (obtained with the SINFONI and OSIRIS spectrographs), based on the presence of CO absorption features or narrow

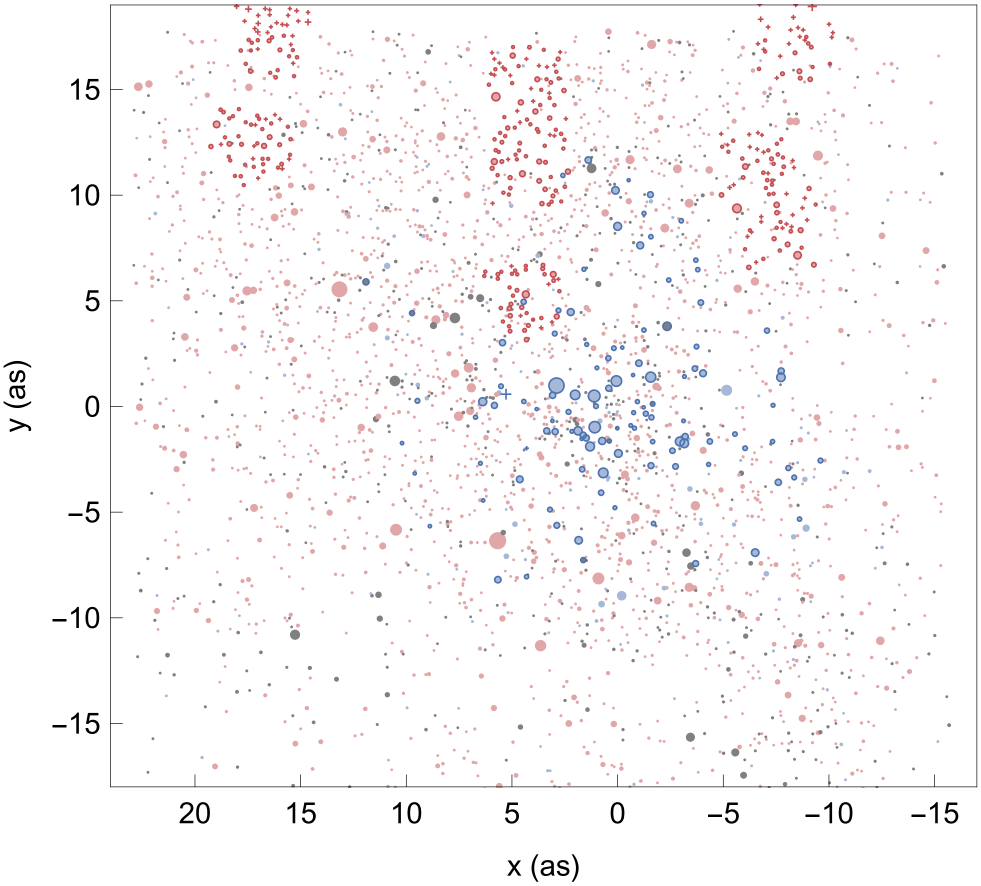

Figure 5.5: A visualization of the classification results, where early- and late-type candidate stars are shown in blue and red color, and unclassified stars are shown in gray, at their angular separations from Sgr A*. The stars that comprise the training set are highlighted in a darker color.

absorption lines of HI (Brγ), HeI or NaI in the stellar spectra. However, we find that the two samples have two sources in common, for which we trust the more recent classification as early-type stars. The late-type sample is widely distributed within the field of view, yet all observed fields are located north of Sgr A*. The early-type sample is concentrated around Sgr A*, being limited in size by the coverage of deep spectroscopic observations. In total, we were able to match 114 of the 116 early-type stars and 215 of the late-type stars to our star list, that comprise our training set. One of the missing early-type stars is extremely saturated, the other is extremely faint.

Type	x (as)	y (as)	IB200	IB206	IB224	IB227	IB230	IB233	IB236	P(E)
	5.676	-6.351	-13.84	-14.39	-15.39	-15.71	-15.56	-14.51	-14.74	0.13
	13.161	5.543	-13.89	-14.42	-15.21	-15.43	-15.27	-14.79	-14.59	0.06
E	2.892	0.989	-14.26	-15.03	-15.07	-15.23	-15.27	-14.89	-15.05	0.89
	0.908	-8.138	-13.13	-14.08	-14.46	-14.70	-14.49	-14.02	-14.05	0.01
	10.486	-5.832	-12.74	-13.71	-14.46	-14.62	-14.35	-14.12	-13.99	0.20
E	1.108	0.496	-13.49	-14.39	-14.40	-14.38	-14.38	-14.18	-14.22	0.76
	3.645	-11.318	-12.84	-13.87	-14.33	-14.26	-14.36	-13.82	-13.61	0.05
E	1.083	-0.978	-13.29	-14.20	-14.31	-13.92	-14.12	-14.15	-14.02	0.76
	-5.160	0.767	-13.00	-13.75	-14.26	-14.50	-14.58	-14.27	-14.44	0.95
	10.541	1.209	-13.05	-13.92	-14.15	-14.18	-14.12	-13.93	-13.77	0.34

Table 5.2: Classification results (abridged[2]). Type: Type of the star, if it is part of the training set, i.e. if it has been classified spectroscopically (E: early-type star, L: late-type star). x/y: Angular separation of the star from Sgr A* in standard coordinates, where x and y increase in the directions of East and North. IB200 - IB236: Instrumental (*StarFinder*) magnitudes in the respective bands (see also Fig. 5.1). The K-band magnitude of each star can be calculated approximately as K ≈ IB224 + 24.63 (see Rafelski et al., 2007, for the photometric standards used). P(E): Estimated probability for the star being an early-type star, where $P(L) = 1 - P(E)$. For stars in the training set, cross-validated estimates are given. The table is sorted by apparent brightness in the IB224 filter.

5.4 Results

5.4.1 Classification performance

Due to the small total number of spectrally classified stars, it is infeasible to keep back a dedicated, representative validation set without compromising the ability to train our classifier. We instead evaluate the classifier's performance by conducting (10-fold) stratified cross-validation on the training set. This means splitting the training set into complementary subsets, accounting for class imbalance, and repeatedly training and testing the classifier using one of the subsets for validation, while using the others combined for training. Each time, the (hyper-)parameters of the classifier are re-optimized as well, using a random search to minimize cross-entropy loss, which is estimated by conducting a second round of cross-validation on the provisional training sets. This whole nested cross-validation procedure is repeated multiple times, to obtain unbiased estimates of the classifier's average performance metrics and their uncertainty (e.g. Cawley & Talbot, 2010). The main parameters to be optimized are the number of features to consider when splitting a tree ($N_{\text{features}} \approx \sqrt{N_{\text{filters}}}$), the depth of the trees ($N_{\text{samples, min.}} \gtrsim 1$ at each leaf), and the number of trees in the random forest ensemble ($N_{\text{trees}} \approx 300$). For completeness, we have also checked that the out-of-bag error rates for selected parameter combinations are reasonable (approx. 7%), which can be estimated already at the training stage.

---

[2] The full table is available online, at `https://doi.org/10.1093/mnras/sty511`.

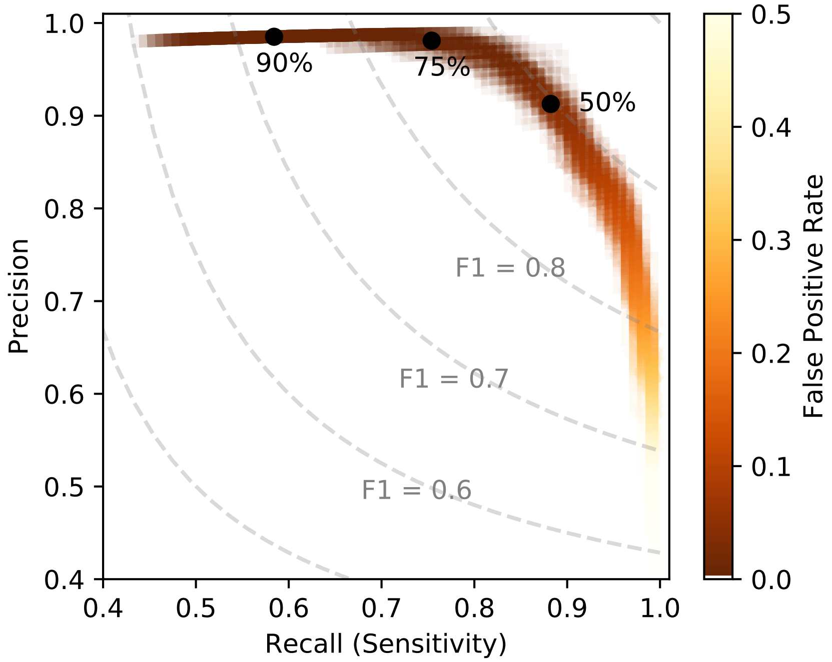

Figure 5.6: Performance of the random-forest classifier with respect to identifying early-type candidate stars, estimated using cross-validation. The labels of the highlighted points indicate the respective thresholds for the class membership probability (see text).

In total, 274 sources in the field of view are classified as early-type stars (class E) and 2216 as late-type stars (class L), each with an estimated class membership probability of at least 75% (see Fig. 5.5 & Tab. 5.2). Of the former, 60 are B-type candidate stars to be confirmed spectroscopically (with $14.5 \lesssim K \lesssim 15.5$). The remaining 675 candidate sources could not be classified reliably using this probability threshold (i.e. have class membership probabilities $P(E) < 75\%$ and $P(L) = 1 - P(E) < 75\%$). The classification of late-type stars is more reliable overall, since roughly 80% of them have a class membership probability exceeding 90%, compared to 60% of the early-type stars.

Based on the cross-validated confusion matrices, we find that the classifier has a high expected overall accuracy of 84%. With respect to identifying early-type stars, the sensitivity (or recall) is 75% and the precision is 98% ($F_1 = 0.85$). Regarding late-type stars, the respective numbers are 89% and 97% ($F_1 = 0.93$). Again, we have required a minimum class membership probability of 75%. This probability threshold could be adjusted downwards to trade precision for better sensitivity, but at the cost of increasing the false positive detection rate for early-type stars to above the 1% level, which we specifically try to avoid (see Fig. 5.6). When using a relatively high threshold value, the early- or late-type stars that are not identified as such are only rarely assigned the wrong spectral type, but instead remain unclassified.

To enable a comparison with the study by Buchholz et al. (2009), we have also cross-matched their star list with our training set. We find that their method of classifying of early-type stars, which involves a direct modeling of the stellar SEDs, appears to be somewhat more reliable, judging on an achieved sensitivity of 85% and a precision of 100% ($F_1 = 0.92$), when considering this common subset of stars. However, the full star lists, when cross-matched, differ in as many as 28% of cases (739 sources). The majority of these different predictions involve stars that are not confidently classified as early- or late type stars in either list, but the star list of Buchholz et al. (2009) notably contains 59 stars labeled late-type that we have classified as early-type stars, and 18 stars labeled early-type that we have classified as late-type.

Of the seven intermediate-band filters, the ones centered on wavelengths of 2.33μm, 2.24μm and 2.27μm prove to be the most critical for the purpose of classifying early- and late-type stars in terms of the Gini importance (see Tab. 5.1), which can be estimated as part of the classifier's training process. The 2.36μm filter provides valuable information as well, but the other filters are less essential. This empirical ranking matches our expectations regarding CO absorption being the distinguishing feature (see Fig. 5.3), but also accounts for variance in the photometric quality across filters.

5.4.2 The stellar population

Apart from the identification of candidate early-type stars for follow-up spectroscopic observations, or promising fields, the large-scale spatial distribution and the luminosity function of the early- and late-type stars are of immediate interest, which we are able to re-estimate using our stellar classification (see Fig. 5.7).

We can reproduce and confirm the results of Buchholz et al. (2009) and other studies of the spatial distribution of stars in the Galactic Center (e.g. Paumard et al., 2006; Lu et al., 2009; Bartko et al., 2009; Do et al., 2009, 2013a). Following Do et al. (2013a, Appendix E), we determine the respective surface density profiles by means of Bayesian inference, using a power-law profile as a model ($\Sigma(R) \propto R^{\Gamma}$). This approach does not require binning the star counts, and allows taking into account the estimated class membership probabilities as weights. For the late-type stars, we find a rather flat density profile ($\Gamma = -0.33 \pm 0.05$), which appears essentially core-like towards the very center. This still presents a puzzle, because these stars are old enough to be dynamically relaxed, and would be expected to form a significantly steeper cusp in that case ($-0.75 < \Gamma < -0.5$; Bahcall & Wolf, 1976, but see Gallego-Cano et al., 2018, and Schödel et al., 2018). For the early-type stars, we find a broken power-law profile with a turnover radius at about $10''$, which we interpret as the outer edge of the young stellar disc (see Støstad et al., 2015). Within that radius, the surface density profile has a well-defined index of $\Gamma = -0.83 \pm 0.10$, or $\Gamma = -1.00 \pm 0.15$ if we exclude the central S-star cluster (at $R < 1''$). Beyond that radius, the number of early-type stars drops rapidly.

We can also reproduce the results of previous studies focused on the stellar K-band luminosity function (KLF). For the the late-type stars with $K < 14.5$, i.e. if we exclude red clump stars, we find a KLF well-described by a relation $N \propto 10^{\beta K}$ with an index of $\beta = 0.36 \pm 0.02$, which is similar to that of the Galactic bulge (e.g. Genzel et al., 2003;

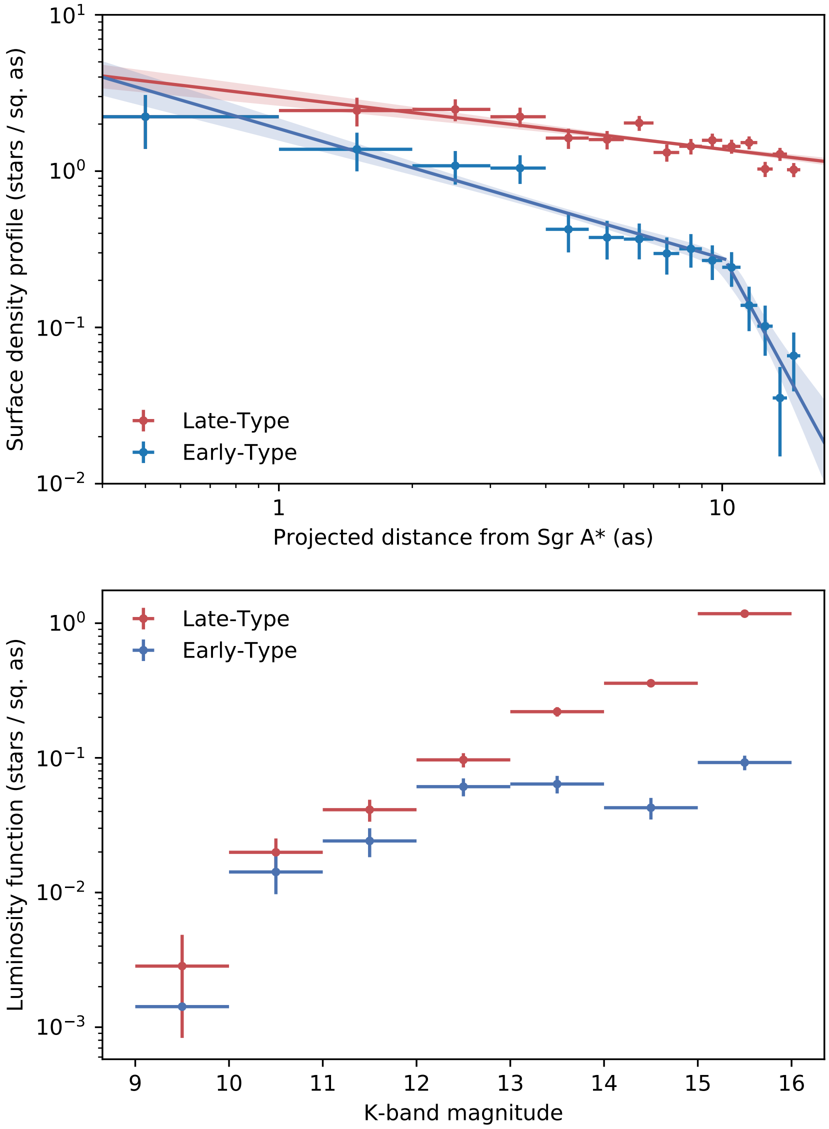

Figure 5.7: The surface density profile (top panel) and the K-band luminosity function (bottom panel) of early- and late-type stars in the Galactic Center. The points mark the stellar number counts in distance and magnitude bins, as indicated by the horizontal error bars, and the vertical error bars indicate Poisson standard errors of the bin counts. For model fitting, the unbinned values are used (see Sec. 5.4.2).

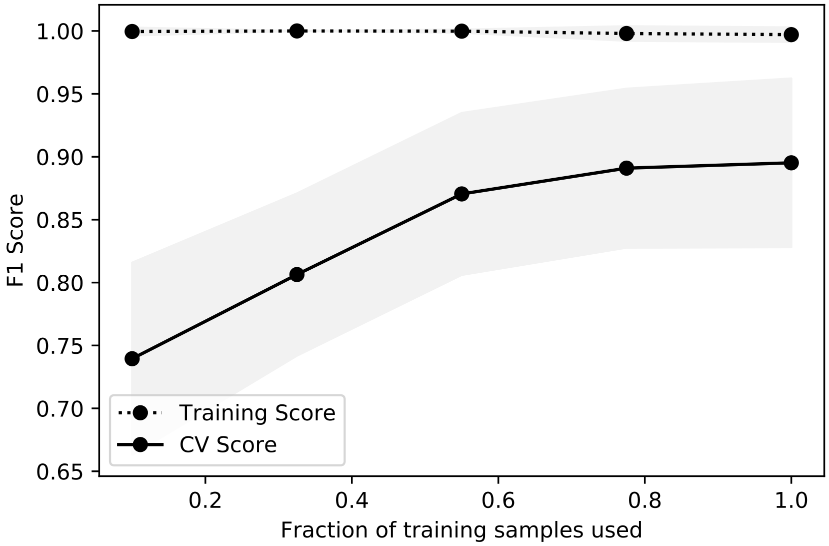

Figure 5.8: Learning curve of the random forest classifier. The performance as measured by the cross-validated F1 score (with respect to identifying early-type stars, and assuming a class membership probability threshold of 50%; see also Fig. 5.6) could probably still be increased by using a larger or higher-quality training set.

Buchholz et al., 2009). The KLF of the early-type stars appears to be top-heavy in the disc region ($1'' \lesssim R \lesssim 10''$), as also found by Paumard et al. (2006), Buchholz et al. (2009) and Bartko et al. (2010). The question of whether the KLF of these stars is indeed flat or somewhat steeper at the faint end, as found by Do et al. (2013a), can likely be answered conclusively only on the basis of spectroscopic observations with a high completeness for B-stars over the entire region, which is currently only reached by deep imaging observations (e.g. $\gtrsim 90\%$).

5.5 Conclusions

For a proof of concept, we have constructed a machine-trained, random forest classifier to identify early- and late-type stars in the environment of the Galactic Center black hole, based on intermediate-band photometry using seven filters in the near-infrared K-band.

With respect to identifying early-type candidate stars in particular, we have demonstrated that our classifier performs similarly well as competitive methods ($F_1 = 0.85$), and we have identified 60 favorable stars in the field of view for follow-up study. The classifier requires an existing training set of spectroscopically classified stars, but requires neither a model of stellar SEDs, nor a calibration of stellar magnitudes beyond accurate instrumental photometry. In principle, a machine-trained classifier will also exploit even subtle class-distinguishing features in the SEDs, i.e. spectral features other than CO absorption, which could make it highly accurate, although such features would not be interpretable in the context of some physically motivated SED model. However, to improve on methods based on direct modeling (e.g. Buchholz et al., 2009), a larger or higher-quality training set would probably be necessary (see Fig. 5.8). Increasing the spectral coverage further, by using additional filters, could be beneficial as well, as would be improvements in the photometric precision and accuracy.

In particular, we find that the brightness of the early- and late-type stars differs in the important filters by only 0.05 to 0.1 magnitudes (see also Fig. 5.3). Even though it is the combined multi-band photometry that makes a spectral classification feasible, to achieve the best possible accuracy, the photometric precision must be sufficiently high in comparison (considering also the intrinsic photometric scatter due to the diversity of spectral types), and a similarly high photometric accuracy must be guaranteed over the whole field of view. While the necessary levels of precision and accuracy have been demonstrated repeatedly (e.g. Lu et al., 2010; Schödel, 2010), capabilities for high-precision photometry, specifically in crowded fields, will be greatly improved when the first instruments on the next generation of large optical telescopes will commence routine operations. We expect that the MICADO imager (Davies et al., 2016), for example, exploiting the light-collecting and resolving power of the ESO ELT, will provide excellent photometry for a much larger sample of stars in the nuclear cluster than is presently possible to obtain. The currently limiting uncertainties in determining the PSF will be reduced by employing an advanced multi-conjugate AO system (MAORY, see Diolaiti et al., 2016), as well as developing complementary PSF reconstruction and modeling techniques as part of the instrument design process. Also, despite not having an angular resolution as high, JWST will likely provide high-quality photometry of the nuclear star cluster as well, due to the PSF stability resulting from the stable conditions of its space environment.

As monitoring observations of the Galactic Center continue, several more early-type stars are likely to be identified spectroscopically over the next few years already, and any further improvements in spectroscopic sensitivity will also help to grow and clean photometric training sets for stellar classification. We are therefore confident that a machine-trained classifier will be useful when applied to future, larger data sets of the Galactic Center, i.e. deep wide-field imaging observations, even if it is only for an efficient preliminary stellar classification.

Chapter 6

Conclusion & Outlook

Near-infrared observations of the Galactic Center have revealed an existing stellar population with a complex history, that formed and evolved in an extreme environment surrounding a (super)massive black hole. While there are still many open questions about the physics of galactic nuclei in general, and the dynamics of nuclear star clusters in particular, the high-resolution observations of the Galactic Center present another, specific set of unsolved puzzles: What processes have shaped the Galactic nuclear star cluster at sub-parsec scales? In particular, why are the bright giants missing, and how did the S-stars form? What are the properties of the faint, unresolved stellar population(s)? Is there a dark cusp? The answers to these questions are likely related, since the dynamics of the extremely dense inner nuclear cluster must have been affected by interactions between stars, stellar remnants, as well as possibly other, exotic objects. The presence of the central black hole must have also had a significant impact throughout its growth, for example on the course of star formation in the Galactic Center, as well as on the dynamical evolution of the formed stars, as it has on the motions of these stars observed today.

Historically, the scientific exploration of the Galactic Center has been an important driving force behind many advances in astronomical instrumentation, for example the development of the first-generation AO-assisted VLT instruments NACO (e.g. Rousset et al., 1998; Lenzen et al., 1998), the near-infrared imager, and SINFONI (e.g. Eisenhauer et al., 2003b,c; Bonnet et al., 2004), the integral field spectrograph. Both of these instruments, still in high demand, have been in almost constant use since the early 2000s, and their capabilities will soon be combined into the re-designed ERIS instrument (see e.g. Amico et al., 2012; Kuntschner et al., 2014; George et al., 2016). In the more recent past, the precise monitoring of the motion of the star S2 during its close approach to the black hole in 2018, and in particular the search for faint stars on even closer orbits, have been two of the main motivations for building the second-generation VLTI instrument GRAVITY, an optical interferometer combining the beams of all four VLT unit telescopes, to effectively create a ∼130 m-diameter telescope with a collecting area of ∼200 m^2 (e.g. Eisenhauer et al., 2008, 2011). Already shortly after seeing first light in late 2016, GRAVITY demonstrated its ability to make minute-long coherent exposures on targets in the Galactic Center, and achieved a differential astrometric precision on the order of a few 10 μas, for S2-like stars

(Gravity Collaboration et al., 2017). Another substantial leap in angular resolution and sensitivity will be brought about by the next generation of extremely large, $30-40$ m-class optical telescopes (ELTs), that are planned to go into operation as early as in the mid-2020s. Assisted by multi-conjugate AO systems, these telescopes will be able to overcome today's limitations imposed by source confusion, make significantly deeper observations of the whole inner nuclear star cluster, and for instance deliver for the first time an astrometric precision on the order of a few 10 μas for a large number of fainter stars in the S-cluster (e.g. Trippe et al., 2010; Davies et al., 2016). Within a few years, the measurement of these stars' orbits could allow determining the black hole mass and distance to a precision better than 0.1%, an order of magnitude better than currently possible (e.g. Weinberg et al., 2005; Do et al., 2017).

Beyond allowing the determination of its mass and distance at increasingly higher precision, the continuous tracking of stellar motions around the Galactic Center black hole will also presently offer the opportunity to probe special and general relativity in a still largely unexplored regime (e.g. Jaroszynski, 1998; Fragile & Mathews, 2000; Zhang et al., 2015; Hees et al., 2017; Grould et al., 2017). The detection of a second-order (in $\beta = v/c$) post-Newtonian term in stellar radial velocities, due to the relativistic Doppler effect and gravitational redshift (in combination with the Rømer effect), could already be feasible during the 2018 pericenter passage of S2, which is imminent at the time of writing, using current-generation instruments (e.g. Zucker et al., 2006; Angélil & Saha, 2010). At this time, S2 is predicted to reach a minimum distance of $\sim$0.7 light days from the black hole ($\sim 1400 R_S$) and a maximum velocity of $\sim$7700 km s^{-1}, i.e. 2.6% of the speed of light. The detection of a prograde Schwarzschild precession of S2's orbit could also be feasible, but is more challenging and likely requires prolonged observations (over at least a few more years), due to the required stability of the astrometric reference frame (e.g. Yelda et al., 2010; Plewa et al., 2015, see also Chap. 2), possible astrometric perturbations caused by source confusion (see Chap. 4), as well as because any extended mass distribution would cause a counteracting Newtonian precession (e.g. Rubilar & Eckart, 2001; Nucita et al., 2007), and interactions with intermediate-mass black holes could induce additional shifts of the orbital elements (e.g. Gualandris et al., 2010), for example. A realistic chance of detecting Lense-Thirring precession due to frame dragging, a third-order effect depending on the black hole spin, or even testing the 'no-hair theorem' (e.g. Will, 2008; Johannsen, 2016), probably only exists if fainter stars are eventually discovered on eccentric short-period orbits, for example with GRAVITY (e.g. Waisberg et al., 2018) or an ELT (e.g. Do et al., 2017), and tracked with high astrometric (and spectroscopic) precision. The motions of these stars are however more likely to be significantly perturbed by interactions with stellar remnants, masking relativistic effects, but also making it possible to probe the dark mass distribution (e.g. Merritt et al., 2010; Zhang & Iorio, 2017). GRAVITY in particular might also be able to probe the physics of the black hole itself in greater detail, by tracking the astrometric centroid motion of Sgr A*'s near-infrared flares in future observations (e.g. Eisenhauer et al., 2011). A complementary effort to study the accretion onto Sgr A* at radio wavelengths, and possibly image the black hole shadow, is the Event Horizon Telescope (e.g. Falcke & Markoff, 2013; Johannsen et al., 2016).

Further observations are also needed to settle the open question of the nature and potentially violent origin of G2, which could have interesting implications for the understanding of star formation and dynamical processes in galactic nuclei. Specifically, it is a current matter of debate whether G2 is a pure gas cloud or host of a massive compact object. The near-infrared observations of G2 itself provide direct evidence for an extended, dusty gas cloud only, the dynamics of which are fully determined by the gravitational force of the black hole (e.g. Gillessen et al., 2013b; Pfuhl et al., 2015; Plewa et al., 2017, see also Chap. 3), in comparison to which the self-gravity of an enshrouded star, for instance, would be negligible. It is therefore to be expected that hydrodynamical simulations of either pure gas clouds (e.g. Schartmann et al., 2015) or mass-loosing stars (e.g. Ballone et al., 2017), for example, can match the observations reasonably well. It is also not particularly surprising that G2 has not yet been disrupted by the black hole, that its infrared luminosity has not dropped, and that no bow shock has been detected, if the density of the ambient medium is in fact lower than previously assumed at the pericenter distance of G2, or even just along its path (e.g. Steinberg et al., 2018, see also Sec. 3.4.1). Any convincing explanation of the nature and origin of G2 would also need to take into account the apparent connection between G2 and G1 (Sec. 3.4.2), and the extended tail (Sec. 3.4.3). The perhaps most plausible explanation, that also accounts for the uniqueness of the observed superstructure, is that G2 is one of many (starless) clumps in a long, fragmenting gas streamer produced by the tidal stripping of a star by the central black hole, several centuries ago (Guillochon et al., 2014). This would imply that many such clumps, in addition to G1, have already preceded G2. These clumps could have cleared a low-density channel for G2 through the ambient medium, and, if they circularized and were subsequently accreted onto the the black hole, could have also caused an increase in the activity of Sgr A*, consistent with other observations (i.e. X-ray echoes; e.g. Ryu et al., 2013). The observed tail should connect G2 to the star that was disrupted, along a trajectory enclosing G2's Keplerian orbit, as observed (see Sec. 3.4.3). To be disrupted by tidal forces at the pericenter distance of G2, this star would need to be large, and have a compatible proper motion today, if it is still observable. Several candidate stars have been identified, but additional kinematic data will be necessary to allow any definitive conclusions (Guillochon et al., 2014). The other supposedly G2-like objects found the vicinity of Sgr A* (e.g. Sitarski et al., 2014) neither have orbits as eccentric as G2's, nor associated tails, and therefore probably originate from different, more common processes, such as the collision of stellar winds (e.g. Cuadra et al., 2006; Calderón et al., 2016)[1].

While innovative technology developments obviously have the potential to bring about breakthrough discoveries, and next-generation telescopes and instruments are all but expected to facilitate such discoveries in the Galactic Center, parallel improvements in the methods of data analysis will be needed to conclusively answer many of today's open questions. Firstly, this is because data sets in general are growing substantially in both size and complexity, and the exploration and analysis of such data sets requires advanced com-

---

[1] A peculiar one among the known objects is X7 (Muzic et al., 2010), a dusty gas structure that is clearly extended towards Sgr A*, but appears to move in an almost perpendicular direction on-sky.

putational methods and tools, in addition to increased computing power. For example, software such as *vaex*[2] (Breddels & Veljanoski, 2018) and *glue*[3] (Beaumont et al., 2015; Robitaille et al., 2017) allows efficient visualization of large, high-dimensional, and interrelated data sets[4], while *ESASky*[5] (Baines et al., 2017), for instance, aims to simplify access to data products from multiple facilities and archives, and *astropy*[6] (Price-Whelan et al., 2018) aims to provide commonly needed tools to the astronomy community, as part of an openly-developed and well-tested software package. At the same time, the construction and operation of new facilities, and in turn the data they provide, are generally becoming more expensive. It is therefore of fundamental concern to fully exploit the available data, in particular by understanding and accounting for both statistical and systematic uncertainties, as best as possible[7]. This requires dedicated modeling of the underlying noise processes, which can be complex to understand in detail, since they typically arise from a combination of effects, both instrumental and astrophysical. An ideal, 'real-world' model should take into account all relevant aspects of the data generation process (see e.g. Hogg et al., 2010), necessarily in a probabilistic manner (since any measurement is by nature stochastic), and could thus be tested against observational data in a meaningful, 'data-driven' way in the framework of Bayesian inference (e.g. Sharma, 2017), with explicit priors (e.g. Gelman et al., 2017), to infer accurate, precise, and, most importantly, trustworthy estimates of the model parameters, together with their uncertainties. In practice, however, such models can be challenging to implement.

For example, even though the challenge of accurately measuring the positions and fluxes of stars in the crowded inner S-star cluster can be addressed successfully, at least to some extent, by modeling the astrometric noise resulting from source confusion (see Chap 4), addressing this challenge fundamentally would require inferring the stellar motions from quasi-raw measurements, i.e. not from source positions, but directly from measurements of light intensity on the detector, which in principle have well-understood noise characteristics. To this general effect, Brewer et al. (2013) proposed to use a generative image model to infer all parameters necessary to create realistic mock images, including the stellar positions and fluxes, their distribution, as well as the total number of stars. In the framework of such a model, it would also be possible to marginalize over parameters describing an arbitrarily complex PSF. The result of this modeling is not a single star list or catalog, but an ensemble of plausible star lists that provides a comprehensive, probabilistic representation of the knowledge about the presence and properties of stars in the underlying imaging data. Any features similar among lists can be considered credible, and features that frequently differ

---

[2] `http://vaex.astro.rug.nl/`

[3] `http://glueviz.org/`

[4] Technology developments in 3D printing and other fabrication techniques (e.g. Madura, 2017; Diemer & Facio, 2017), as well as virtual- and artificial reality (e.g. Rosenfield et al., 2018), are of course offering further prospects for the exploration and communication of complex data, too.

[5] `http://sky.esa.int/`

[6] `http://www.astropy.org/`

[7] The same concern is also providing an additional incentive to quickly make (raw) data products publicly available, to enable and encourage their timely analysis (e.g. as done by the team of the Kepler K2 mission).

should be deemed respectively more uncertain (e.g. see also Daylan et al., 2017; Portillo et al., 2017). With respect to the issue of source confusion specifically, a large fraction of star lists indicating that two stars are blended would imply a rather high confidence in the presence of two stars, instead of one. When testing other models against a probabilistic catalog, which in essence is a closer approximation of the original imaging data than an ordinary star list, this level of confidence would be taken into account immediately. Yet for the Galactic Center science case in particular, it would furthermore be necessary to take into account the stellar proper motions, as well as additional radial velocity measurements.

In comparison to a generative probabilistic model, common machine learning models can often be implemented with less effort, if good training data is available. However, without further input, many such models with large numbers of hidden parameters (e.g. 'deep learning' models) are by construction unsuited for inference tasks because they lack interpretability, in the sense that it should be possible to understand what each parameter of a model represents physically, if scientific conclusions are to be drawn from it (see e.g. Lipton, 2016, for a detailed discussion of interpretability). However, many machine learning models, if carefully tested, can still be useful tools for prediction tasks (e.g. with respect to classification, clustering and outlier detection, regression, or density estimation, and even image processing), due to their flexibility in situations in which probabilistic models are impractical to apply, often because of issues related to data size or complexity, as well as time criticality (see Chap 5, and e.g. Dieleman et al., 2015; Kuntzer et al., 2016; Baron & Poznanski, 2017; George & Huerta, 2017). Machine learning techniques are therefore likely to find application in a wide variety of future studies, in particular when they can be advantageously integrated into the framework of statistical inference (e.g. Leistedt & Hogg, 2017), or other existing workflows (e.g. Beck et al., 2018).

Bibliography

Abuter, R., Schreiber, J., Eisenhauer, F., et al. 2006, New Astronomy Reviews, 50, 398, doi: `10.1016/j.newar.2006.02.008`

Alexander, R. D., Armitage, P. J., & Cuadra, J. 2008, MNRAS, 389, 1655, doi: `10.1111/j.1365-2966.2008.13706.x`

Alexander, T. 2005, Physics Reports, 419, 65, doi: `10.1016/j.physrep.2005.08.002`

—. 2017, ARAA, 55, 17, doi: `10.1146/annurev-astro-091916-055306`

Alexander, T., & Pfuhl, O. 2014, ApJ, 780, 148, doi: `10.1088/0004-637X/780/2/148`

Amaro-Seoane, P., & Chen, X. 2014, ApJ Letters, 781, 18, doi: `10.1088/2041-8205/781/1/L18`

Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O'Neil, M. 2015, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, 252, doi: `10.1109/TPAMI.2015.2448083`

Amico, P., Marchetti, E., Pedichini, F., et al. 2012, Proceedings of SPIE, 8446, doi: `10.1117/12.925170`

Anderson, J., & King, I. R. 2000, PASP, 112, 1360, doi: `10.1086/316632`

—. 2003, PASP, 115, 113, doi: `10.1086/345491`

—. 2006, HST Instrument Science Reports, ACS 2006-01

Anderson, J., & van der Marel, R. P. 2010, ApJ, 710, 1032, doi: `10.1088/0004-637X/710/2/1032`

Angélil, R., & Saha, P. 2010, ApJ, 711, 157, doi: `10.1088/0004-637X/711/1/157`

Angélil, R., Saha, P., & Merritt, D. 2010, ApJ, 720, 1303, doi: `10.1088/0004-637X/720/2/1303`

Baganoff, F. K., Maeda, Y., Morris, M. R., et al. 2003, ApJ, 591, 891, doi: `10.1086/375145`

Bahcall, J. N., & Wolf, R. A. 1976, ApJ, 209, 214, doi: `10.1086/154711`

Baines, D., Giordano, F., Racero, E., et al. 2017, PASP, 129, 028001, doi: `10.1088/1538-3873/129/972/028001`

Ballone, A., Schartmann, M., Burkert, A., et al. 2013, ApJ, 776, 13, doi: `10.1088/0004-637X/776/1/13`

—. 2016, ApJ Letters, 819, 28, doi: `10.3847/2041-8205/819/2/L28`

—. 2017, arXiv, 1706.08547

Baron, D., & Poznanski, D. 2017, MNRAS, 465, 4530, doi: `10.1093/mnras/stw3021`

Bartko, H., Martins, F., Fritz, T. K., et al. 2009, ApJ, 697, 1741, doi: `10.1088/0004-637X/697/2/1741`

Bartko, H., Martins, F., Trippe, S., et al. 2010, ApJ, 708, 834, doi: `10.1088/0004-637X/708/1/834`

Baumgardt, H., Amaro-Seoane, P., & Schödel, R. 2018, A&A, 609, 28, doi: `10.1051/0004-6361/201730462`

Beaumont, C., Goodman, A., & Greenfield, P. 2015, Proceedings of ADASS (XXIV), 495, 101

Beck, M. R., Scarlata, C., Fortson, L. F., et al. 2018, MNRAS, 476, 5516, doi: `10.1093/mnras/sty503`

Bellini, A., Anderson, J., van der Marel, R. P., et al. 2014, ApJ, 797, 115, doi: `10.1088/0004-637X/797/2/115`

Binney, J., & Tremaine, S. 2011, Galactic Dynamics, Second Edition (Princeton University Press)

Bland-Hawthorn, J., & Gerhard, O. 2016, ARAA, 54, 529, doi: `10.1146/annurev-astro-081915-023441`

Boehle, A., Ghez, A., Schödel, R., et al. 2016, ApJ, 830, 17, doi: `10.3847/0004-637X/830/1/17`

Bonnell, I. A., & Rice, W. K. M. 2008, Science, 321, 1060, doi: `10.1126/science.1160653`

Bonnet, H., Abuter, R., Baker, A., et al. 2004, The ESO Messenger, 117, 17

Boubert, D., Erkal, D., Evans, N. W., & Izzard, R. G. 2017, MNRAS, 469, 2151, doi: `10.1093/mnras/stx848`

Bower, G. C., Wright, M. C. H., Falcke, H., & Backer, D. C. 2003, ApJ, 588, 331, doi: `10.1086/373989`

Bower, G. C., Markoff, S. B., Dexter, J., et al. 2015, ApJ, 802, 69, doi: `10.1088/0004-637X/802/1/69`

Breddels, M. A., & Veljanoski, J. 2018, arXiv, 1801.02638

Breiman, L. 2001, Machine Learning, 45, 5, doi: `10.1023/A:1010933404324`

Brewer, B. J., Foreman-Mackey, D., & Hogg, D. W. 2013, AnJ, 146, 7, doi: `10.1088/0004-6256/146/1/7`

Brown, W. R., Anderson, J., Gnedin, O. Y., et al. 2015, ApJ, 804, 49, doi: `10.1088/0004-637X/804/1/49`

Brown, W. R., Geller, M. J., & Kenyon, S. J. 2014, ApJ, 787, 89, doi: `10.1088/0004-637X/787/1/89`

Buchholz, R. M., Schödel, R., & Eckart, A. 2009, A&A, 499, 483, doi: `10.1051/0004-6361/200811497`

Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, 125, doi: `10.1051/0004-6361/201322971`

Burkert, A., Schartmann, M., Alig, C., et al. 2012, ApJ, 750, 58, doi: `10.1088/0004-637X/750/1/58`

Calderón, D., Ballone, A., Cuadra, J., et al. 2016, MNRAS, 455, 4388, doi: `10.1093/mnras/stv2644`

Cawley, G., & Talbot, N. 2010, Journal of Machine Learning Research, 11, 2079

Chatzopoulos, S., Fritz, T. K., Gerhard, O., et al. 2015, MNRAS, 447, 952, doi: `10.1093/mnras/stu2452`

Chauvin, G., Lagrange, A.-M., Udry, S., et al. 2006, A&A, 456, 1165, doi: `10.1051/0004-6361:20054709`

Chauvin, G., Vigan, A., Bonnefoy, M., et al. 2015, A&A, 573, 127, doi: `10.1051/0004-6361/201423564`

Chen, X., & Amaro-Seoane, P. 2014, ApJ Letters, 786, 14, doi: `10.1088/2041-8205/786/2/L14`

Clenet, Y., Rouan, D., Gratadour, D., et al. 2005, A&A Letters, 439, 9, doi: `10.1051/0004-6361:200500146`

Clenet, Y., Rouan, D., Gendron, E., et al. 2004a, A&A Letters, 417, 15, doi: `10.1051/0004-6361:20040031`

Clenet, Y., Rouan, D., Gratadour, D., et al. 2004b, A&A Letters, 424, 21, doi: `10.1051/0004-6361:200400045`

Cuadra, J., Nayakshin, S., Springel, V., & Di Matteo, T. 2006, MNRAS, 366, 358, doi: `10.1111/j.1365-2966.2005.09837.x`

Dale, J. E., Davies, M. B., Church, R. P., & Freitag, M. 2009, MNRAS, 393, 1016, doi: `10.1111/j.1365-2966.2008.14254.x`

Davies, R., & Kasper, M. 2012, ARAA, 50, 305, doi: `10.1146/annurev-astro-081811-125447`

Davies, R., Schubert, J., Hartl, M., et al. 2016, Proceedings of SPIE, 9908, doi: `10.1117/12.2233047`

Daylan, T., Portillo, S. K. N., & Finkbeiner, D. P. 2017, ApJ, 839, 4, doi: `10.3847/1538-4357/aa679e`

De Colle, F., Raga, A. C., Contreras-Torres, F. F., & Toledo-Roy, J. C. 2014, ApJ Letters, 789, 33, doi: `10.1088/2041-8205/789/2/L33`

Dieleman, S., Willett, K. W., & Dambre, J. 2015, MNRAS, 450, 1441, doi: `10.1093/mnras/stv632`

Diemer, B., & Facio, I. 2017, PASP, 129, 058013, doi: `10.1088/1538-3873/aa6a46`

Diolaiti, E., Bendinelli, O., Bonaccini, D., et al. 2000, A&A Supplements, 147, 335, doi: `10.1051/aas:2000305`

Diolaiti, E., Ciliegi, P., Abicca, R., et al. 2016, Proceedings of SPIE, 9909, doi: `10.1117/12.2234585`

Do, T., Ghez, A., Morris, M. R., et al. 2009, ApJ, 703, 1323, doi: `10.1088/0004-637X/703/2/1323`

Do, T., Hees, A., Dehghanfar, A., Ghez, A., & Wright, S. 2017, arXiv, 1711.06389

Do, T., Lu, J. R., Ghez, A., et al. 2013a, ApJ, 764, 154, doi: `10.1088/0004-637X/764/2/154`

Do, T., Martinez, G. D., Yelda, S., et al. 2013b, ApJ Letters, 779, 6, doi: `10.1088/2041-8205/779/1/L6`

Doeleman, S. S., Weintroub, J., Rogers, A. E. E., et al. 2008, Nature, 455, 78, doi: `10.1038/nature07245`

Eckart, A., & Genzel, R. 1996, Nature, 383, 415, doi: `10.1038/383415a0`

Eckart, A., Muzic, K., Yazici, S., et al. 2013, A&A, 551, 18, doi: `10.1051/0004-6361/201219994`

Eisenhauer, F., & Raab, W. 2015, ARAA, 53, 155, doi: `10.1146/annurev-astro-082214-122442`

Eisenhauer, F., Schödel, R., Genzel, R., et al. 2003a, ApJ Letters, 597, 121, doi: `10.1086/380188`

Eisenhauer, F., Abuter, R., Bickert, K., et al. 2003b, Proceedings of SPIE, 4841, doi: `10.1117/12.459468`

Eisenhauer, F., Tecza, M., Thatte, N., et al. 2003c, The ESO Messenger, 113, 17

Eisenhauer, F., Genzel, R., Alexander, T., et al. 2005, ApJ, 628, 246, doi: `10.1086/430667`

Eisenhauer, F., Perrin, G., Rabien, S., et al. 2008, Proceedings of SPIE, 431, doi: `10.1007/978-3-540-74256-2_50`

Eisenhauer, F., Perrin, G., Brandner, W., et al. 2011, The ESO Messenger, 143, 16

Falcke, H., & Markoff, S. B. 2013, Classical and Quantum Gravity, 30, 244003, doi: `10.1088/0264-9381/30/24/244003`

Feldmeier-Krause, A., Neumayer, N., Schödel, R., et al. 2015, A&A, 584, 2, doi: `10.1051/0004-6361/201526336`

Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601, doi: `10.1111/j.1365-2966.2009.14548.x`

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: `10.1086/670067`

Fragile, P. C., & Mathews, G. J. 2000, ApJ, 542, 328, doi: `10.1086/309495`

Fritz, T. K., Gillessen, S., Dodds-Eden, K., et al. 2010a, ApJ, 721, 395, doi: `10.1088/0004-637X/721/1/395`

Fritz, T. K., Gillessen, S., Trippe, S., et al. 2010b, MNRAS, 401, 1177, doi: `10.1111/j.1365-2966.2009.15707.x`

Fritz, T. K., Gillessen, S., Dodds-Eden, K., et al. 2011, ApJ, 737, 73, doi: `10.1088/0004-637X/737/2/73`

Fritz, T. K., Chatzopoulos, S., Gerhard, O., et al. 2016, ApJ, 821, 44, doi: `10.3847/0004-637X/821/1/44`

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1, doi: `10.1051/0004-6361/201629272`

Gallego-Cano, E., Schödel, R., Dong, H., et al. 2018, A&A, 609, 26, doi: `10.1051/0004-6361/201730451`

Gelman, A., Simpson, D., & Betancourt, M. 2017, Entropy, 19, 555, doi: `10.3390/e19100555`

Genzel, R., Eckart, A., Ott, T., & Eisenhauer, F. 1997, MNRAS, 291, 219

Genzel, R., Eisenhauer, F., & Gillessen, S. 2010, Rev. Mod. Phys., 82, 3121, doi: `10.1103/RevModPhys.82.3121`

Genzel, R., Schödel, R., Ott, T., et al. 2003, ApJ, 594, 812, doi: `10.1086/377127`

George, D., & Huerta, E. A. 2017, arXiv, 1711.07966

George, E. M., Gräff, D., Feuchtgruber, H., et al. 2016, Proceedings of SPIE, 9908, doi: `10.1117/12.2231285`

Ghez, A., Klein, B. L., Morris, M. R., & Becklin, E. E. 1998, ApJ, 509, 678, doi: `10.1086/306528`

Ghez, A., Duchêne, G., Matthews, K., et al. 2003, ApJ Letters, 586, 127, doi: `10.1086/374804`

Ghez, A., Hornstein, S. D., Lu, J. R., et al. 2005, ApJ, 635, 1087, doi: `10.1086/497576`

Ghez, A., Salim, S., Weinberg, N. N., et al. 2008, ApJ, 689, 1044, doi: `10.1086/592738`

Gillessen, S., Eisenhauer, F., Fritz, T. K., et al. 2009a, ApJ Letters, 707, 114, doi: `10.1088/0004-637X/707/2/L114`

Gillessen, S., Eisenhauer, F., Trippe, S., et al. 2009b, ApJ, 692, 1075, doi: `10.1088/0004-637X/692/2/1075`

Gillessen, S., Genzel, R., Fritz, T. K., et al. 2012, Nature, 481, 51, doi: `10.1038/nature10652`

—. 2013a, ApJ, 763, 78, doi: `10.1088/0004-637X/763/2/78`

—. 2013b, ApJ, 774, 44, doi: `10.1088/0004-637X/774/1/44`

Gillessen, S., Plewa, P. M., Eisenhauer, F., et al. 2017, ApJ, 837, 30, doi: `10.3847/1538-4357/aa5c41`

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65, doi: `10.2140/camcos.2010.5.65`

Gould, A., & Quillen, A. C. 2003, ApJ, 592, 935, doi: `10.1086/375840`

Gravity Collaboration, Abuter, R., Accardo, M., et al. 2017, A&A, 602, 94, doi: `10.1051/0004-6361/201730838`

Grould, M., Vincent, F. H., Paumard, T., & Perrin, G. 2017, A&A, 608, 60, doi: `10.1051/0004-6361/201731148`

Gualandris, A., Gillessen, S., & Merritt, D. 2010, MNRAS, 409, 1146, doi: `10.1111/j.1365-2966.2010.17373.x`

Gubler, J., & Tytler, D. 1998, PASP, 110, 738, doi: `10.1086/316172`

Guillochon, J., Loeb, A., MacLeod, M., & Ramirez-Ruiz, E. 2014, ApJ Letters, 786, 12, doi: `10.1088/2041-8205/786/2/L12`

Habibi, M., Gillessen, S., Martins, F., et al. 2017, ApJ, 847, 120, doi: `10.3847/1538-4357/aa876f`

Hees, A., Do, T., Ghez, A., et al. 2017, PRL, 118, 211101, doi: `10.1103/PhysRevLett.118.211101`

Hills, J. G. 1988, Nature, 331, 687, doi: `10.1038/331687a0`

Hogg, D. W. 2012, arXiv, 1205.4446

Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv, 1008.4686

Jaroszynski, M. 1998, Acta Astronomica, 48, 653

Johannsen, T. 2016, Classical and Quantum Gravity, 33, 124001, doi: `10.1088/0264-9381/33/12/124001`

Johannsen, T., Broderick, A. E., Plewa, P. M., et al. 2016, PRL, 116, 031101, doi: `10.1103/PhysRevLett.116.031101`

Kuntschner, H., Jochum, L., Amico, P., et al. 2014, Proceedings of SPIE, 9147, doi: `10.1117/12.2055140`

Kuntzer, T., Tewes, M., & Courbin, F. 2016, A&A, 591, 54, doi: `10.1051/0004-6361/201628660`

Leistedt, B., & Hogg, D. W. 2017, ApJ, 838, 5, doi: `10.3847/1538-4357/aa6332`

Lenzen, R., Hofmann, R., Bizenberger, P., & Tusche, A. 1998, Proceedings of SPIE, 3354, doi: `10.1117/12.317287`

Levin, Y., & Beloborodov, A. M. 2004, ApJ, 613, 224, doi: `10.1086/422908`

Lipton, Z. C. 2016, arXiv, 1606.03490

Lu, J. R., Do, T., Ghez, A., et al. 2013, ApJ, 764, 155, doi: `10.1088/0004-637X/764/2/155`

Lu, J. R., Ghez, A., Hornstein, S. D., et al. 2009, ApJ, 690, 1463, doi: `10.1088/0004-637X/690/2/1463`

Lu, J. R., Ghez, A., Yelda, S., et al. 2010, Proceedings of SPIE, 7736, doi: `10.1117/12.859737`

MacKay, D. J. C. 2003, Information Theory, Inference and Learning Algorithms (Cambridge University Press)

Madigan, A., Levin, Y., & Hopman, C. 2009, ApJ Letters, 697, 44, doi: `10.1088/0004-637X/697/1/L44`

Madigan, A., McCourt, M., & O'Leary, R. M. 2017, MNRAS, 465, 2310, doi: `10.1093/mnras/stw2815`

Madigan, A., Pfuhl, O., Levin, Y., et al. 2014, ApJ, 784, 23, doi: `10.1088/0004-637X/784/1/23`

Madura, T. I. 2017, PASP, 129, 058011, doi: `10.1088/1538-3873/129/975/058011`

Maness, H., Martins, F., Trippe, S., et al. 2007, ApJ, 669, 1024, doi: `10.1086/521669`

Mapelli, M., & Ripamonti, E. 2015, ApJ, 806, 197, doi: `10.1088/0004-637X/806/2/197`

Marrone, D. P., Moran, J. M., Zhao, J.-H., & Rao, R. 2007, ApJ, 654, L57, doi: `10.1086/510850`

Martins, F., Gillessen, S., Eisenhauer, F., et al. 2008, ApJ Letters, 672, 119, doi: `10.1086/526768`

Martins, F., Trippe, S., Paumard, T., et al. 2006, ApJ, 649, L103, doi: `10.1086/508328`

Masciadri, E., Brandner, W., Bouy, H., et al. 2003, A&A, 411, 157, doi: `10.1051/0004-6361:20031428`

McCourt, M., & Madigan, A. 2016, MNRAS, 455, 2187, doi: `10.1093/mnras/stv2232`

McCourt, M., O'Leary, R. M., Madigan, A., & Quataert, E. 2015, MNRAS, 449, 2, doi: `10.1093/mnras/stv355`

Menten, K. M., Reid, M. J., Eckart, A., & Genzel, R. 1997, ApJ Letters, 475, 111, doi: `10.1086/310472`

Merritt, D. 2010, ApJ, 718, 739, doi: `10.1088/0004-637X/718/2/739`

Merritt, D., Alexander, T., Mikkola, S., & Will, C. M. 2010, Phys. Rev. D, 81, 062002, doi: 10.1103/PhysRevD.81.062002

Merritt, D., Gualandris, A., & Mikkola, S. 2009, ApJ Letters, 693, 35, doi: 10.1088/0004-637X/693/1/L35

Meyer, F., & Meyer-Hofmeister, E. 2012, A&A Letters, 546, 2, doi: 10.1051/0004-6361/201220145

Meyer, L., Ghez, A., Schödel, R., et al. 2012, Science, 338, 84, doi: 10.1126/science.1225506

Meyer, L., Ghez, A., Witzel, G., et al. 2014, IAU Symposium 303, 264, doi: 10.1017/S1743921314000714

Miralda-Escudé, J. 2012, ApJ, 756, 86, doi: 10.1088/0004-637X/756/1/86

Morris, M. R. 1993, ApJ, 408, 496, doi: 10.1086/172607

Murray-Clay, R. A., & Loeb, A. 2012, Nature Communications, 3, 1049, doi: 10.1038/ncomms2044

Muzic, K., Eckart, A., Schödel, R., et al. 2010, A&A, 521, 13, doi: 10.1051/0004-6361/200913087

Naoz, S., Ghez, A., Hees, A., et al. 2018, ApJ Letters, 853, 24, doi: 10.3847/2041-8213/aaa6bf

Narayan, R., Özel, F., & Sironi, L. 2012, ApJ Letters, 757, 20, doi: 10.1088/2041-8205/757/2/L20

Nayakshin, S., & Zubovas, K. 2018, arXiv, 1804.02914

Nishiyama, S., & Schödel, R. 2013, A&A, 549, 57, doi: 10.1051/0004-6361/201219773

Nucita, A. A., De Paolis, F., Ingrosso, G., Qadir, A., & Zakharov, A. F. 2007, PASP, 119, 349, doi: 10.1086/517934

Paumard, T., Genzel, R., Martins, F., et al. 2006, ApJ, 643, 1011, doi: 10.1086/503273

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Perets, H. B., Gualandris, A., Kupi, G., Merritt, D., & Alexander, T. 2009, ApJ, 702, 884, doi: 10.1088/0004-637X/702/2/884

Perets, H. B., Hopman, C., & Alexander, T. 2007, ApJ, 656, 709, doi: 10.1086/510377

Pfuhl, O., Alexander, T., Gillessen, S., et al. 2014, ApJ, 782, 101, doi: `10.1088/0004-637X/782/2/101`

Pfuhl, O., Fritz, T. K., Zilka, M., et al. 2011, ApJ, 741, 108, doi: `10.1088/0004-637X/741/2/108`

Pfuhl, O., Gillessen, S., Eisenhauer, F., et al. 2015, ApJ, 798, 111, doi: `10.1088/0004-637X/798/2/111`

Phifer, K., Do, T., Meyer, L., et al. 2013, ApJ Letters, 773, 13, doi: `10.1088/2041-8205/773/1/L13`

Pickles, A. J. 1998, PASP, 110, 863, doi: `10.1086/316197`

Plewa, P. M., Gillessen, S., Eisenhauer, F., et al. 2015, MNRAS, 453, 3234, doi: `10.1093/mnras/stv1910`

Plewa, P. M., Gillessen, S., Pfuhl, O., et al. 2017, ApJ, 840, 50, doi: `10.3847/1538-4357/aa6e00`

Portillo, S. K. N., Lee, B. C. G., Daylan, T., & Finkbeiner, D. P. 2017, AnJ, 154, 132, doi: `10.3847/1538-3881/aa8565`

Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, arXiv, 1801.02634

Prodan, S., Antonini, F., & Perets, H. B. 2015, ApJ, 799, 118, doi: `10.1088/0004-637X/799/2/118`

Quataert, E., & Gruzinov, A. 2000, ApJ, 545, 842, doi: `10.1086/317845`

Rafelski, M., Ghez, A., Hornstein, S. D., Lu, J. R., & Morris, M. R. 2007, ApJ, 659, 1241, doi: `10.1086/512062`

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

Reid, M. J., Menten, K. M., Genzel, R., et al. 2003, ApJ, 587, 208, doi: `10.1086/368074`

Reid, M. J., Menten, K. M., Trippe, S., Ott, T., & Genzel, R. 2007, ApJ, 659, 378, doi: `10.1086/511744`

Rein, H., & Liu, S. F. 2012, A&A, 537, 128, doi: `10.1051/0004-6361/201118085`

Rein, H., & Spiegel, D. S. 2015, MNRAS, 446, 1424, doi: `10.1093/mnras/stu2164`

Rein, H., & Tamayo, D. 2015, MNRAS, 452, 376, doi: `10.1093/mnras/stv1257`

Rieke, G. H. 2007, ARAA, 45, 77, doi: `10.1146/annurev.astro.44.051905.092436`

Robitaille, T., Beaumont, C., Qian, P., Borkin, M., & Goodman, A. 2017, Zenodo, doi: `10.5281/zenodo.293197`

Rosenfield, P., Fay, J., Gilchrist, R. K., et al. 2018, arXiv, 1801.09119

Rousset, G., Lacombe, F., Puget, P., et al. 1998, Proceedings of SPIE, 3353, doi: `10.1117/12.321686`

Rubilar, G. F., & Eckart, A. 2001, A&A, 374, 95, doi: `10.1051/0004-6361:20010640`

Ryu, S. G., Nobukawa, M., Nakashima, S., et al. 2013, PASJ, 65, 33, doi: `10.1093/pasj/65.2.33`

Sabha, N., Eckart, A., Merritt, D., et al. 2012, A&A, 545, 70, doi: `10.1051/0004-6361/201219203`

Sądowski, A., Narayan, R., Sironi, L., & Özel, F. 2013a, MNRAS, 433, 2165, doi: `10.1093/mnras/stt879`

Sądowski, A., Sironi, L., Abarca, D., et al. 2013b, MNRAS, 432, 478, doi: `10.1093/mnras/stt495`

Schartmann, M., Burkert, A., Alig, C., et al. 2012, ApJ, 755, 155, doi: `10.1088/0004-637X/755/2/155`

Schartmann, M., Ballone, A., Burkert, A., et al. 2015, ApJ, 811, 155, doi: `10.1088/0004-637X/811/2/155`

Schödel, R. 2010, A&A, 509, 58, doi: `10.1051/0004-6361/200912808`

Schödel, R., Gallego-Cano, E., Dong, H., et al. 2018, A&A, 609, 27, doi: `10.1051/0004-6361/201730452`

Schödel, R., Najarro, F., Muzic, K., & Eckart, A. 2010, A&A, 511, 18, doi: `10.1051/0004-6361/200913183`

Schödel, R., Ott, T., Genzel, R., et al. 2003, ApJ, 596, 1015, doi: `10.1086/378122`

—. 2002, Nature, 419, 694

Scoville, N., & Burkert, A. 2013, ApJ, 768, 108, doi: `10.1088/0004-637X/768/2/108`

Sharma, S. 2017, ARAA, 55, 213, doi: `10.1146/annurev-astro-082214-122339`

Shcherbakov, R. V. 2014, ApJ, 783, 31, doi: `10.1088/0004-637X/783/1/31`

Sitarski, B., Do, T., Witzel, G., et al. 2014, Proceedings of the AAS, 223, 238.05

Sitarski, B., Ghez, A., Morris, M. R., et al. 2015, Proceedings of the AAS, 225, 102.07

Skilling, J. 2004, AIP Conference Proceedings, 735, 395, doi: `10.1063/1.1835238`

Steinberg, E., Sari, R., Gnat, O., et al. 2018, MNRAS, 473, 1841, doi: `10.1093/mnras/stx2438`

Støstad, M., Do, T., Murray, N., et al. 2015, ApJ, 808, 106, doi: `10.1088/0004-637X/808/2/106`

Trippe, S., Davies, R., Eisenhauer, F., et al. 2010, MNRAS, 402, 1126, doi: `10.1111/j.1365-2966.2009.15940.x`

Trippe, S., Gillessen, S., Gerhard, O., et al. 2008, A&A, 492, 419, doi: `10.1051/0004-6361:200810191`

Tsuboi, M., Kitamura, Y., Tsutsumi, T., et al. 2017, ApJ Letters, 850, 5, doi: `10.3847/2041-8213/aa97d3`

Valencia-S, M., Eckart, A., Zajacek, M., et al. 2015, ApJ, 800, 125, doi: `10.1088/0004-637X/800/2/125`

Waisberg, I., Dexter, J., Gillessen, S., et al. 2018, MNRAS, doi: `10.1093/mnras/sty476`

Wang, Q. D., Nowak, M. A., Markoff, S. B., et al. 2013, Science, 341, 981, doi: `10.1126/science.1240755`

Weinberg, N. N., Milosavljević, M., & Ghez, A. 2005, ApJ, 622, 878, doi: `10.1086/428079`

Weinberg, S. 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity (Wiley)

Will, C. M. 2008, ApJ Letters, 674, 25, doi: `10.1086/528847`

Witzel, G., Ghez, A., Morris, M. R., et al. 2014, ApJ Letters, 796, 8, doi: `10.1088/2041-8205/796/1/L8`

Xu, Y.-D., Narayan, R., Quataert, E., Yuan, F., & Baganoff, F. K. 2006, ApJ, 640, 319, doi: `10.1086/499932`

Yelda, S., Ghez, A., Lu, J. R., et al. 2014, ApJ, 783, 131, doi: `10.1088/0004-637X/783/2/131`

Yelda, S., Lu, J. R., Ghez, A., et al. 2010, ApJ, 725, 331, doi: `10.1088/0004-637X/725/1/331`

Yuan, F., Quataert, E., & Narayan, R. 2003, ApJ, 598, 301, doi: `10.1086/378716`

Zhang, F., & Iorio, L. 2017, ApJ, 834, 198, doi: `10.3847/1538-4357/834/2/198`

Zhang, F., Lu, Y., & Yu, Q. 2015, ApJ, 809, 127, doi: `10.1088/0004-637X/809/2/127`

Zhao, C., & Burge, J. H. 2007, Optics Express, 15, 18014, doi: `10.1364/OE.15.018014`

—. 2008, Optics Express, 16, 6586, doi: `10.1364/OE.16.006586`

Zucker, S., Alexander, T., Gillessen, S., Eisenhauer, F., & Genzel, R. 2006, ApJ Letters, 639, 21, doi: `10.1086/501436`

Acknowledgements

I thank my colleagues in the IR group for their support, in particular Michi Bauböck, Sirio Belli, Leonard Burtscher, Peter Buschkamp, Angela Cortes, Rebecca Davies, Ric Davies, Susanne Dengler, Jason Dexter, Frank Eisenhauer, Sebastiano von Fellenberg, Helmut Feuchtgruber, Tobias Fritz, Feng Gao, Vincent Garrel, Reinhard Genzel, Liz George, Stefan Gillessen, Javier Graciá-Carpio, Maryam Habibi, Susanne Harai-Ströbel, Michael Hartl, Rodrigo Herrera-Camus, Annemieke Janssen, Alejandra Jimenez-Rosales, Ming-yi Lin, Lena Lippa, Dieter Lutz, Thomas Ott, Oliver Pfuhl, Sedona Price, Adriane Richter, Taro Schimizu, Andreas Schruba, Raphael Stock, Eckhard Sturm, Linda Tacconi, Hannah Übler, Idel Waisberg, Felix Widmann, Eva Wuyts, Stijn Wuyts, as well as Alessandro Ballone, Linda Baronchelli, Andi Burkert, Diego Calderón, Damien Coffey, Jorge Cuadra, Korbinian Huber, Darshan Kakkad, Karl Menten, Mark Reid, Malin Renneby, Re'em Sari, Marc Schartmann, Christine Schulz, Vlas Sokolov, Graham Wagstaff, and Philip Wiseman.