Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Diplom-Physiker: Adrian Kaminski
Born in: Opole, Poland
Oral examination: 30th October, 2013

The Search for Extrasolar Planets by the Means of High Precision Astrometry with the Dual-Feed Interferometer PRIMA

Referees: **Prof. Dr. Andreas Quirrenbach**

Prof. Dr. Joachim Krautter

Abstract

Several hundred extrasolar planets have been discovered until today. The most commonly applied detection methods are indirect and therefore typically limited in determining the whole set of the orbital parameters and in characterizing the specific planets. Among these, astrometry serves as an important complementary technique, as the orbital inclination and the planet's true mass can be derived from the two-dimensional measurements of the host star's reflex motion due to the presence of a companion. The VLTI's dual-feed interferometer PRIMA has been developed to provide high precision differential astrometry at the level of tens of microarcseconds, by which Saturn-like planets around nearby stars would become detectable. This is accomplished by simultaneous phase-referenced fringe tracking on two sources within the isoplanatic angle. Within this thesis the instrument's overall performance during commissioning is examined. Stable fringe tracking on both sources is generally achieved, and the effect of phase referencing is evident. However, phase jumps during the observations are reported, which might be due to phase shift offsets between the spatially modulated reference signals, from which the phases are recovered. These are investigated. An analysis of the discrepancy between the achieved astrometric precisions on short and long time scales reveals hidden but evident systematic errors.

Zusammenfassung

Einige hundert extrasolare Planeten wurden bis heute bereits entdeckt. Die am häufigsten angewendeten Nachweismethoden sind indirekter Natur und deswegen für gewöhnlich darin begrenzt, den kompletten Satz an Bahnelementen zu bestimmen und die vermeintlichen Planeten vollkommen zu charakterisieren. Astrometrie eignet sich insbesondere als eine ergänzende Technik, da aus der zweidimensionalen Vermessung der Reflexbewegung des Sterns, aufgrund der Anwesenheit eines Begleiters, die Bahnneigung und die wahre Planetenmasse abgeleitet werden können. Das doppeltgespeiste VLTI-Interferometer PRIMA wurde entwickelt, um präzise differentielle Astrometrie im Bereich deutlich unter 100 Mikrobogensekunden zu ermöglichen, wodurch saturnähnliche Planeten um nahe Sterne nachweisbar würden. Dies wird durch simultanes und phasenreferenziertes „Fringe-Tracking“ an zwei individuellen Quellen innerhalb des isoplanaren Winkels ermöglicht. In der vorliegenden Arbeit wird allgemein die Leistung des Instruments während der Testphase untersucht. Stabile Nachführung des Interferenzmusters von beiden Quellen wird üblicherweise erzielt, und die Auswirkungen der Phasenreferenz sind ersichtlich. Dennoch werden Phasensprünge während der Messungen beobachtet. Diese könnten durch die hier untersuchten Phasenverschiebungen der modulierten Referenzsignale hervorgerufen werden, aus welchen die Phasen letztlich bestimmt werden. Eine Analyse der Diskrepanz zwischen der erreichten astrometrischen Präzision auf kurzen und langen Zeitskalen legt verborgene systematische Fehler nahe.

Our imagination is stretched to the utmost, not, as in fiction, to imagine things which are not really there, but just to comprehend those things which are there.

Richard P. Feynman

Contents

List of Figures

List of Tables

Conventions, Units and Acronyms

- Vectors are denoted by bold characters. If a defined vector quantity appears as not bold in the subsequent context, then this is referring to its magnitude.
- The complex conjugate of a complex quantity is denoted by the superscript *.
- Some frequently used non-SI units are listed in Tab. 0.1.
- Important acronyms that can be found commonly throughout the document are given in Tab. 0.2.

Table 0.1. – Frequently used non-SI units

Symbol	Name	Equivalent SI unit
as	Arcsecond	$1\,\mathrm{as} = 4.8481 \times 10^{-6}\,\mathrm{rad}$
AU	Astronomical unit	$1\,\mathrm{AU} = 149.598 \times 10^{9}\,\mathrm{m}$
ly	Light year	$1\,\mathrm{ly} = 9.46073 \times 10^{15}\,\mathrm{m}$
$\mathrm{M}_\oplus$	Earth mass	$1\,\mathrm{M}_\oplus = 5.9722 \times 10^{24}\,\mathrm{kg}$
$\mathrm{M}_{\mathrm{Jup}}$	Jupiter mass	$1\,\mathrm{M}_{\mathrm{Jup}} = 1.8986 \times 10^{27}\,\mathrm{kg}$
$\mathrm{M}_{\mathrm{Sat}}$	Saturn mass	$1\,\mathrm{M}_{\mathrm{Sat}} = 5.6846 \times 10^{26}\,\mathrm{kg}$
$\mathrm{M}_\odot$	Solar mass	$1\,\mathrm{M}_\odot = 1.9884 \times 10^{30}\,\mathrm{kg}$
pc	Parsec	$1\,\mathrm{pc} = 30.856776 \times 10^{15}\,\mathrm{m}$

Table 0.2. – Frequently used and important Acronyms.

Acronym	Meaning
2MASS	Two Micron All-Sky Survey
ADI	Angular Differential Imaging
ADRS	Astrometric Data Reduction Software
AMBER	Astronomical Multi-BEam combineR
ASTRID	AStrometric Target & Reference stars Interactive Database
AT	Auxiliary Telescope
AO	Adaptive Optics
BC	Beam splitter Cube
BL	BaseLine
CCD	Charge-Coupled Device
CCF	Cross-Correlation Function
CMB	Cosmic Microwave Background
CoCo	Correction Collection
CTTS	Classical T-Tauri Star
DFT	Discrete Fourier Transform

Acronym	Meaning
DL	Delay Line
DDL	Differential Delay Line
dOPD	Differential Optical Path Difference
dOPDC	Differential Optical Path Delay Controller
ESA	European Space Agency
ESO	European Southern Observatory
ESPRI	Exoplanet Search with Prima
FOV	Field Of View
FSM	Field Selection Mirror
FSU	Fringe Sensor Unit
FWHM	Full Width at Half Maximum
GD	Group Delay
GI	Gravitational Instability
GMC	Giant Molecular Cloud
GTO	Guaranteed Observation Time
HST	Hubble Space Telescope
IAU	International Astronomical Union
ICRF	International Celestial Reference Frame
ICRS	International Celestial Reference System
IERS	International Earth and Reference System Service
ITRS	International Terrestrial Reference System
IR	InfraRed
IRIS	InfraRed Image Sensor
KBO	Kuiper Belt Object
LSW	Landessternwarte Königstuhl
MARCEL	Multi-beam Alignment, Reference and Calibration Emitter
MCF	Mutual Coherence Function
MDL	Main Delay Line
MHD	MagnetoHydroDynamics
MIDI	MID-infrared Interferometric instrument
MMSN	Minimum Mass Solar Nebula
MPIA	Max Planck Institute for Astronomy
MRI	MagnetoRotational Instability
MSDF	Mutual Spectral Density Function
NAB	Narrow-Angle Baseline
NS	Nearby Star
NPRO	Non Planar Ring Oscillator
NTT	New Technology Telescope
OMC	Orion Molecular Cloud
OPD	Optical Path Difference
OPDC	Optical Path Delay Controller
PA	Position Angle

Acronym	Meaning
PBA	Projected Baseline Angle
PBS	Polarizing Beam Splitter
PRIMA	Phase Referenced Imaging and Microarcsecond Astrometry
PRIMET	PRIMA Metrology
PS	Primary Star
PSF	Point-Spread-Function
PTI	Palomar Testbed Interferometer
RMS	Root Mean Square
RT	Real-Time
RTO	Real-Time Offset
RV	Radial Velocity
SED	Spectral Energy Distribution
SNR	Signal-to-Noise Ratio
SPD	Spectral Photon Distribution
PS	Secondary Star
SSSB	Small Solar System Body
STS	Star Separator
TMC	Taurus Molecular Cloud
TNO	Trans-Neptunian Object
TOA	Time-Of-Arrival
TTV	Transit Timing Variation
UT	Unit Telescope
UTC	Coordinated Universal Time
UV	UltraViolet
VCM	Variable Curvature Mirror
VINCI	VLTI Commissioning Instrument
VLTI	Very Large Telescope Interferometer
WDS	Washington visual Double Star catalog
YS	Young Star
YSO	Young Stellar Object
ZAMS	Zero Age Main Sequence

Chapter 1

Motivation and Thesis Outline

Approaching questions concerning the universe and its understanding is an ancient tradition among mankind. Its bilateral characteristic of being elusive on the one hand, but coevally being one of the most prominent features of the human's environment during nighttime is the source not only for curiosity and motivation to scientists since the Renaissance, but also has been for natural philosophers of the Antiquity.
What has begun quite naively with pure descriptions of the celestial bodies' motions and shapes centuries ago evolved to a branch of science, which in order to cope with the universe's diversity and complexity today incorporates different fields of physics and not only makes use of diverse mathematical concepts, but was also a trigger for the elaboration of some of those in the past.
The achievements in modern astronomy and astrophysics establish our conception of the universe and its nature. Its formation (around 13.7 billion years ago) and expansion are widely accepted and well-founded cosmological concepts. Within this framework measurements of the fluctuations of the *cosmic microwave background* (CMB) allow determinations of its density and composition. As far as can be concluded, only 4–5 % of the energy density in the universe is assigned to baryonic matter that in principle can be observed directly. Since the first telescopes in the beginning of the 17^{th} century instrumentation technology has evolved so that today a wide range of the electromagnetic spectrum (from radio frequencies to X-rays) can be covered by observation.
By that we have learned a lot about the observable structures and their hierarchy. Starting in the nearby neighborhood of our own solar system, the nearest star system is identified as *Alpha Centauri* in a distance of about 4.3 light years. The center of our galaxy the *Milky Way*, a disk-shaped structure harboring 100–300 billion stars and ca. 30 kpc in diameter, is determined to lie in a distance of about 7.6 kpc.
Galaxies in turn, which usually can be classified as elliptical, spiral, dwarf or irregular, form associations called clusters and even superclusters arranged in gravitationally bound filaments, the largest known cosmic structures. All in all the number of galaxies forming within the currently observable universe[1] can be estimated to the order of about 170 billion [Gott05].

Research on those various objects procure an insight into physical processes in nature and certainly in many cases can be regarded as fundamental, but they are also of philosophical importance regarding the self-conception of mankind. A significant issue is and always has been the question "How is it that we are here?" This consequently leads to the question of life in the universe and how it forms itself. In order to study this question one needs to agree on how life can be defined. It is justifiable to limit the problem to life basing upon carbon chemistry

[1] This term denotes the radius of 14×10^9 pc out to the CMB-photons originating at the recombination epoch.

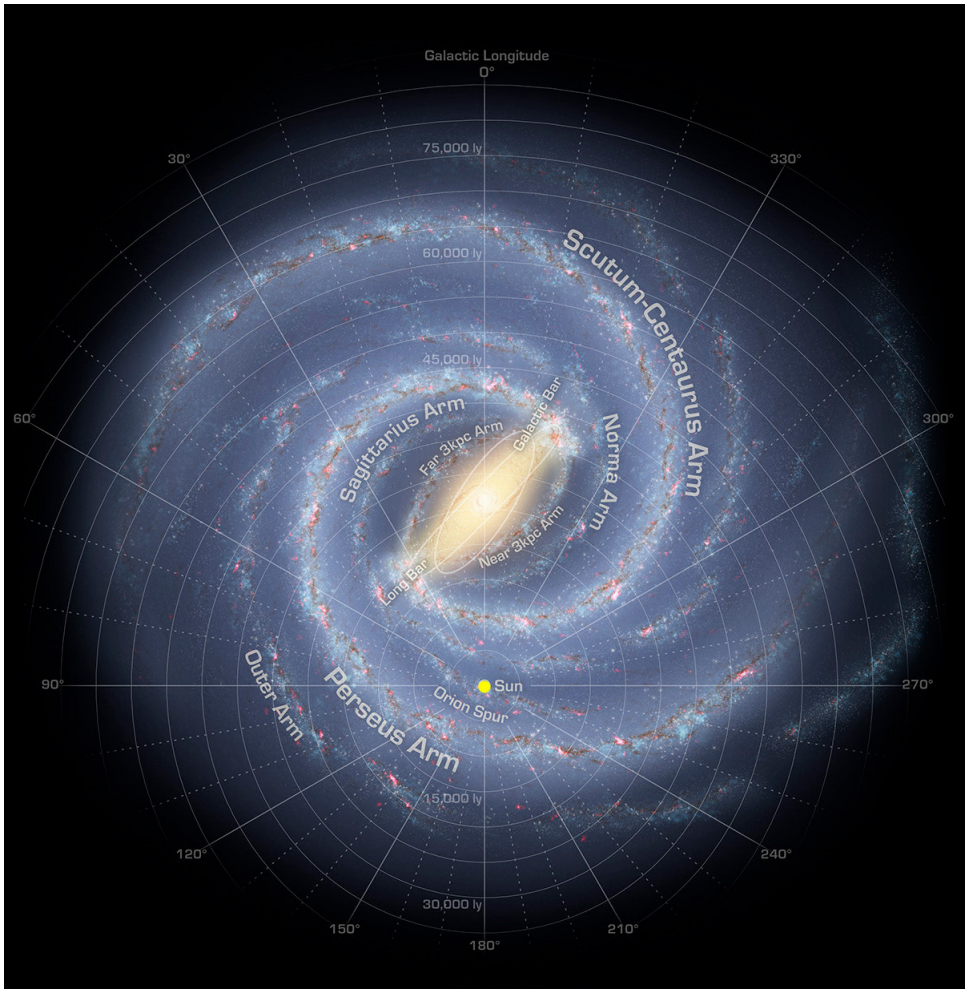


Figure 1.1. – Artistic chart of the Milky way showing the galactic center, the position of the sun (yellow dot) and the most prominent arms. (Credit: NASA/JPL-Caltech/R.Hurt (SSC/Caltech) [NASA08])

as it has evolved on the only place known, namely the earth. Investigations must still clarify which conditions need to be satisfied within an environment where life is to be supported, but also how such an environment itself develops.
Therefore, it is essential to understand, how planetary systems around stars form and how representative the solar system is. Unfortunately, especially this field is hardly accessible in observational astronomy. Comparing to the bright stars, evident to anyone on the night sky, planets are relatively small, cold and dim. This makes the detection and analysis of planets quite challenging and indirect methods, predicated on the understanding of physical interactions within such systems, become inevitable.
Gravitational forces between massive objects for example act as indicator for their existence. *Isaac Newton's* formulation of classical mechanics and the law of gravitation (published 1687 in *Philosophiae Naturalis Principia Mathematica*) provided a tool to understand and to predict the motion of objects interacting gravitationally. The history of the exploration of our own

solar system has shown the power of this concept. Although observed and recorded before, Neptune was not discovered and classified as a planet of the solar system until September 1846. This was achieved only since irregularities between Uranus'[2] observed and expected orbit have been interpreted as a possible reaction to another object farther away. Precise predictions by astronomers of this time (i.a. *U. Le Verrier (1811–1877)* and *J. C. Adams (1819–1892)*) allowed *Johann G. Galle (1812–1910)* to search for it at the correct positions and were crucial for the planet's detection and confirmation.
In this case the mathematical predictions were 'only' a hint leading to the right interpretations, but when looking for extrasolar planets indirect methods become absolutely essential. The spatial proximity of possible planets to their host stars and the enormous contrast in luminosity so far prevented direct observations of such objects. Due to those difficulties, detections of extrasolar planetary systems could not be provided until the 1990s. Precise timing measurements on the millisecond radio pulsar PSR1257 + 12, performed at the 305 m Arecibo radio telescope, revealed the existence of two planet-sized bodies revolving around the neutron star in almost circular orbits [Wol1992]. This discovery in 1992 was the first one of that kind and can be regarded as the begin of a new era within this field in astronomy.
During the subsequent years, many detections followed. Especially radial velocity surveys proved to be effective at finding extrasolar planets. From periodic variations in the star's radial velocity, resulting from reflex motions to a companion in that system, the orbit parameters and a minimum mass of the orbiting object were derived and published in 1995 as the first detection of an exoplanet around a solar-type star [May1995].
The characteristics of the early discovered exoplanetary systems with high masses on short and eccentric orbits certainly have been highly biased by the detection technique, but still provided the first empirical data on the evolution of planetary systems besides the solar system. Therefore, understanding planet formation mechanisms is an origin for the motivation to search for extrasolar planets.
Since then, several hundreds planets and also multiplanetary systems around host stars have been detected. Most of them have been discovered by radial velocity measurements but other methods become more and more successful recently. Especially the photometric technique of measuring transits of companions in front of their host stars (with the most prominent space based transit surveys *COROT* and *Kepler*) has proven its capability in this field.

The oldest of the methods, and one of the more challenging at the same time, is Astrometry, which so far lacks in confirmed detections. In order to find a companion, the host star's position over time needs to be monitored at a high precision level so that its motion, with respect to the system's barycenter, can be measured. The magnitude of the astrometric signature of such a reflex motion naturally depends on the system's distance to the observer, as well as on the properties of the orbit and of the objects themselves. Figure 1.2 illustrates the astrometric signature of our sun due to all solar system objects as it would be measured by an observer at 10 pc distance. This indicates which accuracies are needed to be obtained so that systems similar to our own can be detected by astrometric methods.

[2]Discovered by *William Herschel* in 1781 already.

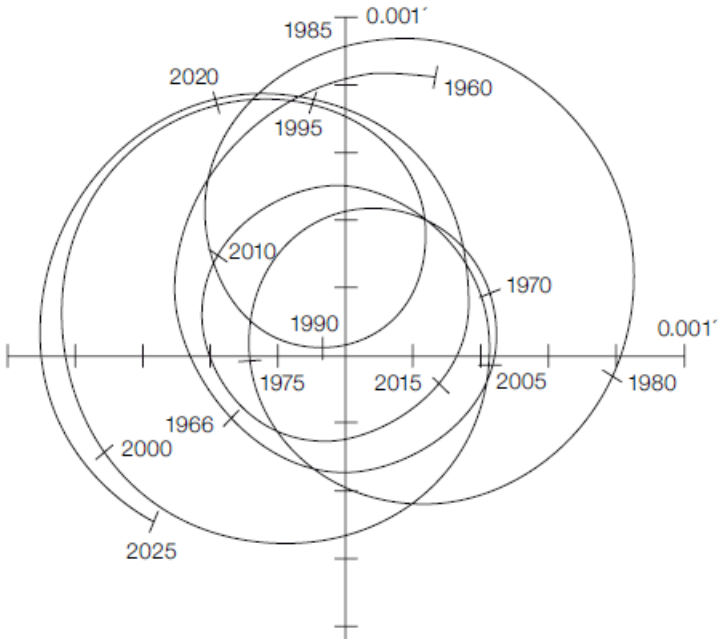


Figure 1.2. – Sun's reflex motion to all solar system planets as astrometric signature in time that would be measured at a distance of 10 pc. The axes lie in the observer's sky plane and span an area of 2×2 milliarcseconds. As a reference, at this distance the sun's diameter would cover an angle of around 1 masec. This at the same time is roughly the magnitude of the astrometric signature as induced by the dominating solar system body Jupiter. (Credit: [vBelle08])

This thesis is devoted to the discussion of one particular technique for performing high precision astrometry, namely double star interferometry, and tries to cover both, the requirements for designing an appropriate instrument and the prospects for its functionality and efficiency.
The new interferometric facility PRIMA (Phase Referenced Imaging and Microarcsecond Astrometry), installed as part of the VLTI (Very Large Telescope Interferometer) infrastructure and currently being commissioned, serves as an exemplification in this context. Its technical components, as well as its conceptional layout, are described in chapter 4, where also the results from the performed analysis of commissioning data regarding the instrument's functionality and preliminarily achievable astrometric accuracies are presented. Chapter 2 shall give an introduction to the whole topic in general. Formation models of planets and planetary systems are summarized and the different detection techniques for the search for exoplanets, with an emphasis on astrometry, are described in more detail.
Chapter 3 provides the therotecial background for the understanding of interferometry in general and in particular the concept of double star interferometry as a technique for obtaining astrometric measurements.
As already mentioned, the design of PRIMA is motivated by the aim of finding exoplanets by the means of astrometry. Consequently, the ESPRI (Exoplanet Search with PRIMA) survey is being prepared under collaboration of the consortium members and will be presented in chapter 5, after which the thesis is concluded by a short summary in chapter 6.

Chapter 2

Introduction to the Topic

2.1. Planets, Extrasolar Planets and Planetary Systems

As outlined already, the search for planets outside the solar system is a new field in astronomy. However, the research on planetary systems in general is not. It is no surprise that the prominence of our system's planets as relatively bright light sources on the sky made them inherent elements in human history of science back from Antiquity through the Middle Ages and Renaissance until now. They played a role in mysticism and religion, but moreover revolutions of the human view of the world, as the foundation of the heliocentric system by *Nicolaus Copernicus* in his *De revolutionibus orbium coelestium* in 1543, have been triggered by the study of celestial mechanics. Long time has passed since the beginning of the transition from the geocentric

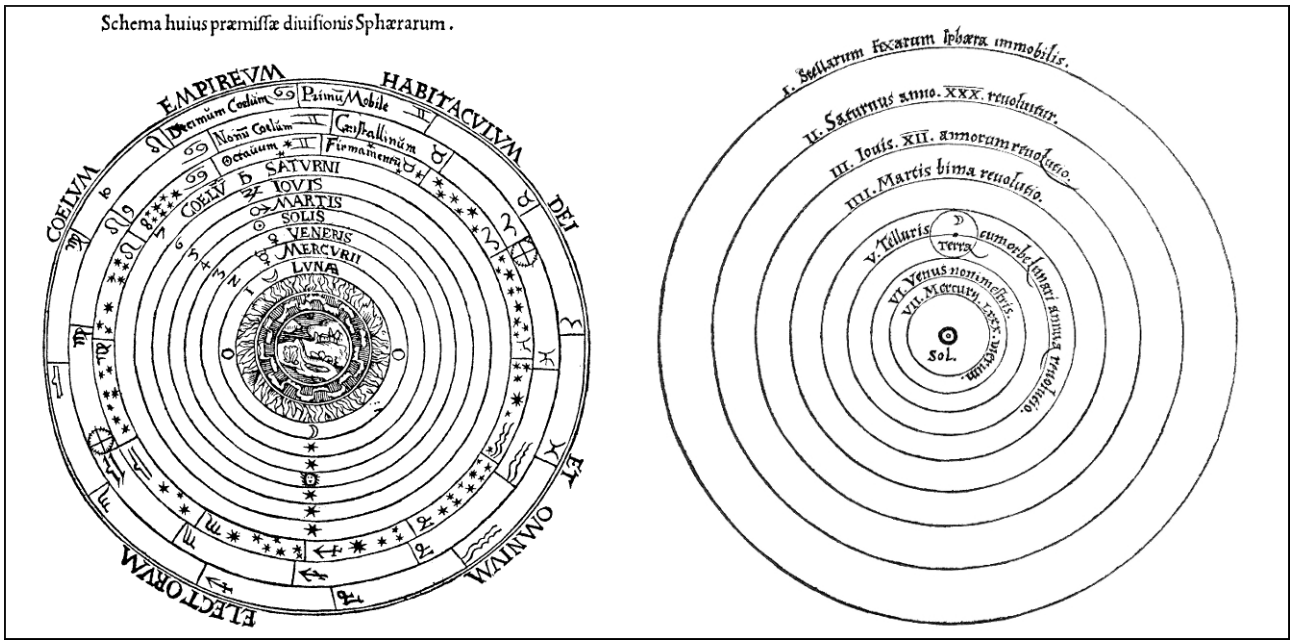


Figure 2.1. – *Left:* Schematic illustration of the geocentric model as depicted in *Peter Apian's Cosmographia, 1539* (credit: [Gra1987]) showing the earth in the center surrounded by outer shells harboring stars as well as the solar system bodies. *Right:* Illustration of the heliocentric system from N. Copernicus' *De revolutionibus orbium coelestium.*

to the heliocentric model in the 16$^{\text{th}}$ century until today, and although modern astronomy achieved a lot since then, the origin of our solar system still is not satisfactorily understood. The process of planetary systems' formation exposes itself as a complex combination of various

physical mechanisms on different spatial scales. The rough basic concept, known as *the nebular hypothesis*, has been established already by *Emanuel Swedenborg (1688–1772)*, *Immanuel Kant (1724–1804)* and *Pierre-Simon Laplace (1749–1827)*. It describes the development of the solar system, stating that at an early stage a dense and slowly rotating molecular gas cloud collapsed and while cooling down formed the sun, surrounded and orbited by a disk of gas and dust particles (see also Fig. 2.2). Consequently, the planets formed from the disk material.

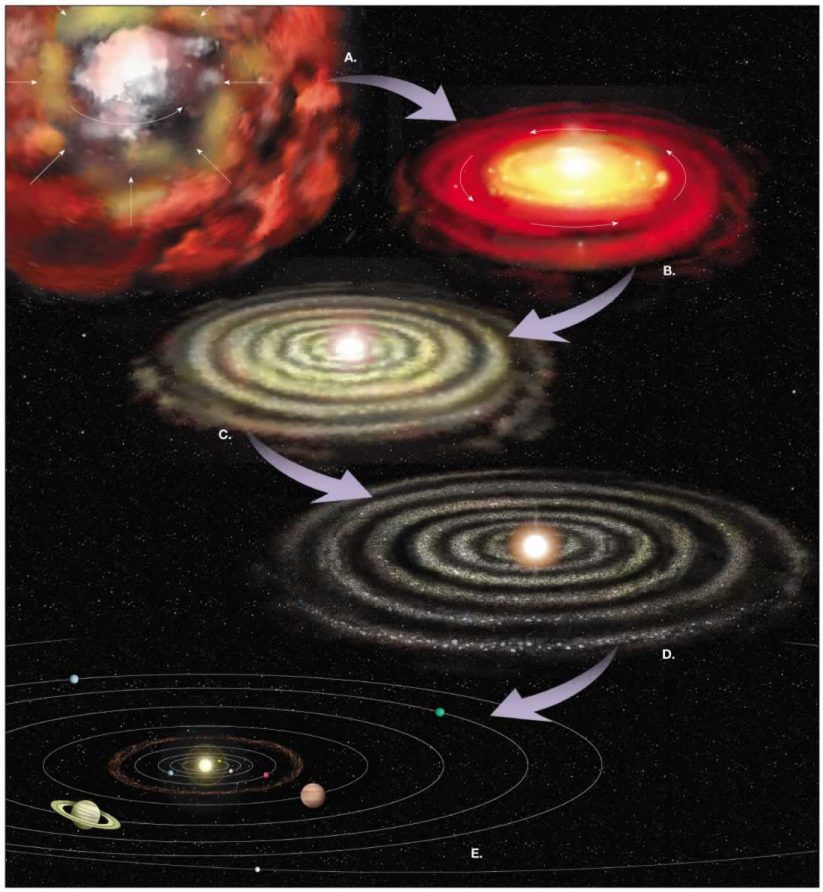


Figure 2.2. – A historical picture of planet formation. The idea of a collapsing molecular cloud cooling down and forming a flat disk is illustrated by the steps A, B and as a basic concept is still considered to be correct. On the other hand, the simplistic portrayal of the disk breaking up into rings and subsequently forming planets (C–E) is outdated. (Credit: [Joh11])

This rather pragmatic and result-oriented model explains by simplistic reasoning the prograde planet orbits and the existence of the ecliptic plane. Modern theories on planet formation expand the fundamental hypothesis. Physical processes are studied in detail in order to model each formation step on the way from star formation itself to a stable planetary system. The goal is to comprehend thoroughly each stage of the whole planetary system evolution and to combine them to a fully coherent theory. As will be outlined here, the main mechanisms are fairly understood, but some open questions still need to be answered. Until then, the whole

picture remains incomplete.
As common in astronomy, one of the major difficulties for the research in this field is the absence of an appropriate testbed. Typically, observations of the objects in question provide a basis for verification or falsification of any modelled theories. In the case of planet formation, for a long time only one instance of a system served as a sample, namely our solar system. Any model can be checked against the properties of planets found in our neighborhood, but it is still not clear, if our solar system is to be regarded as ordinary, or as rare of its kind due to exceptional initial conditions and evolution.
Actually, the discovery of the first exoplanets in the 90's seemed to indicate the latter. Massive bodies with relatively high eccentricities on short orbits have not been expected to be found in such an amount, but it should be realized that the properties of detected bodies are highly influenced by the technique used to look for them. Consequently, since most of the early discovered extrasolar planets have been found by Doppler spectroscopy (see Sec. 2.3.2), which is mostly sensitive to high masses with short semi-major axes, it is clear that those were the first to find. Nevertheless, by the early findings the diversity of planetary systems became evident and justified the study of exoplanetary systems.

2.1.1. Planet Properties and Orbit Parametrization

Before discussing the properties and formation of planets, it is useful to clarify which bodies fall into that category. In 2006, the *International Astronomical Union (IAU)* resolved and published a nomenclature for bodies in our solar system. Such a definition became necessary due to recent discoveries of exoplanetary systems on the one hand, but also due to the exploration of our own solar system. In the 1990's, objects, other than Pluto, on transneptunian[1] orbits have been discovered. During the years since then, our knowledge of these bodies increased, and with successive spottings of such, the region populated by them turned out more and more crowded and is now known as the Kuiper belt. Comparing Pluto's low mass and rather untypical orbital properties[2] to those objects of the new class, a definite nomenclature got inevitable.
The following excerpt from the IAU resolution B5 classifies all solar system bodies, other than the natural satellites, into three categories and can also be regarded as a guideline for extrasolar planetary systems:

1. A planet is a celestial body that
   a) is in orbit around the Sun,
   b) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape, and
   c) has cleared the neighborhood around its orbit.
2. A "dwarf planet" is a celestial body that
   a) is in orbit around the Sun,

[1] Bodies on orbits with semi-major axes greater than Neptune's (≈ 30AU) are usually called *Trans-Neptunian objects (TNOs)*.
[2] Untypical compared to the other solar system's planets. Especially its high inclination of about $i \approx 17^\circ$ to the ecliptic plane emphasizes Pluto's differentness.

b) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape,
c) has not cleared the neighborhood around its orbit, and
d) is not a satellite.

3. All other objects, except satellites, orbiting the Sun shall be referred to collectively as "Small Solar System Bodies".

The main consequence of this reclassification was that Pluto is now categorized as a *dwarf planet*.

With regard to modelling planet formation, the knowledge of specific characteristics of solar system objects is crucial. These have the power to constrain certain aspects of theories and can give clues to the right answers. Generally, planet properties can be divided into two categories. On the one hand, there are intrinsic properties describing the body itself, such as its mass, size and surface structure as the most prominent ones, but also others, which are affected by the planet's evolution at a higher level, need to be taken into account: The existence of a magnetic field is coupled to the composition of the body's interior. The rotation rate, direction and the orientation of the rotation axis are not only results of the initial conditions during the formation stage, but definitely are determined by gravitational or even physical interactions between bodies in the system, as well.
On the other hand, characteristics of the planet's orbit are important. In particular, when trying to characterize extrasolar planets, indirect detection techniques (see Sec. 2.3) mainly aim at, besides the mass, determining the orbital parameters.
The motion of planets is described by the solution of the two-body problem. *Johannes Kepler (1571–1630)* deduced three fundamental laws of planetary motion from observations. In principle those are also derivable from Newton's laws of motion and his universal law of gravity, defining the attractive gravitational force by $\boldsymbol{F}_g = -\frac{Gm_1m_2}{r^2}\boldsymbol{r}$ between two bodies with masses m_1 and m_2 resp., the connection vector $\boldsymbol{r}$ and the gravitational constant G. Consequently, the equation to be solved for the relative motion of the two bodies becomes:

$$\mu_r \frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} = -\frac{G\mu_r M}{r^2}\left(\frac{\boldsymbol{r}}{r}\right) \quad , \tag{2.1}$$

where $\mu_r = \frac{m_1m_2}{(m_1+m_2)}$ is the reduced mass and $M = m_1 + m_2$ the total mass. In principle, the behavior of the system's motion, defined by the relative orbit of the two bodies, is purely determined by its constant total energy:

$$E = \frac{1}{2}\mu_r v^2 - \frac{GM\mu_r}{r} \quad , \tag{2.2}$$

with the kinetic energy expressed by the first term and the potential by the second. As solutions to the problem, the resulting curves of relative motion can be described by conic sections, depending on the sign of the total energy. For $E > 0$ the system is unbound and the relative orbit mathematically would be a hyperbola. Motion on a parabola results from the kinetic and potential energy being equal in magnitude expressed by $E = 0$. Bound orbits are consequences to the case $E < 0$ and are described by elliptical orbits, as stated by the first of the following

three Kepler's laws on planetary motion:

1. Two gravitationally bound bodies move along elliptical paths, where one focal point of each ellipse is located at the center of mass.
   Since the description of the relative motion is completely equivalent to that of a particle with the reduced mass μ_r in orbit around a fixed central mass M (compare Eq. (2.1)), the historic statement of Kepler that planets move on elliptical paths with their host stars at one of the focii becomes also equivalent.[3] In polar coordinates the curve describing the planet's motion in time can be parametrized by its distance to the star:

$$r(t) = \frac{a(1-e^2)}{1 + e\cos\nu(t)} \quad . \tag{2.3}$$

   The shape of the ellipse (and therefore the orbit) is fully described by the semi-major axis $a = \frac{r_{\max}-r_{\min}}{2}$ (the mean of the maximum and minimum distance) and its eccentricity e. The time-dependent *true anomaly* ν gives the angle between the planet's periapsis[4] to its instantaneous position (see Fig. 2.3 a)).

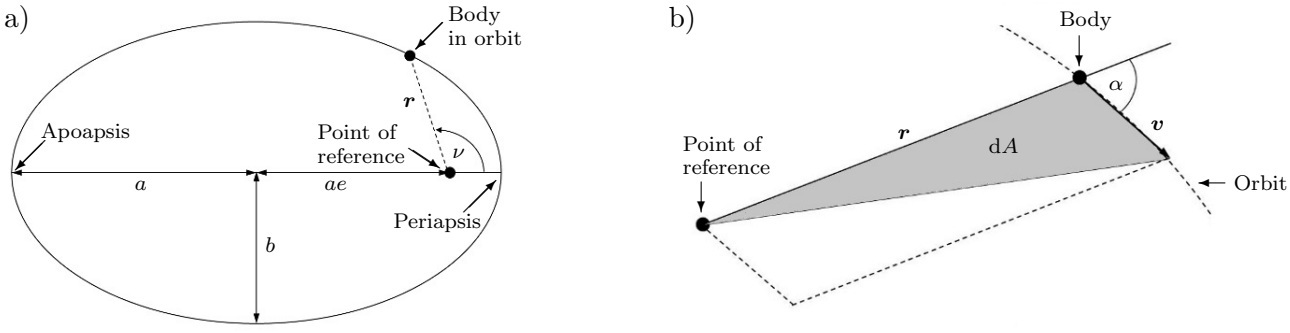


Figure 2.3. – a): Geometry of elliptic orbits. The point of reference is located at one focus of the ellipse, which is defined by its semi-major axis a, semi-minor axis b and eccentricity e. The instantaneous position of the body in orbit is characterized by the magnitude of the position vector $\boldsymbol{r}$ and the true anomaly ν. b): Sketch depicting Kepler's second law of planetary motion. Within an infinitesimal unit time $\mathrm{d}t$ the joining line from the center of reference to the body in orbit sweeps out the area $\mathrm{d}A = \frac{1}{2} rv\sin\alpha$.

2. Kepler's second law follows directly from the conservation of angular momentum $\boldsymbol{L} = \boldsymbol{r} \times m\boldsymbol{v} = \boldsymbol{r} \times \boldsymbol{p}$. Looking at its derivative with respect to time, we see that

$$\begin{aligned} \frac{\mathrm{d}L}{\mathrm{d}t} &= \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \times \boldsymbol{p}\right) + \left(\boldsymbol{r} \times \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}\right) \\ &= \quad 0 \quad + \quad \boldsymbol{r} \times \boldsymbol{F} \quad = 0 \quad . \end{aligned} \tag{2.4}$$

[3] In practice the choice of reference, either the center of mass or one of the bodies (usually but not necessarily that of higher mass), is determined by the choice of what is to be studied, the absolute or the relative orbit. In the gravitational two-body system it is common to call the heavier object the *primary* and the other in orbit the *secondary*.

[4] Point of minimum distance to the point of reference.

The first term on the right hand side vanishes trivially due to $\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{v} = \frac{\boldsymbol{p}}{m}$. The second term remains and describes the resulting torque exerted by the force $\boldsymbol{F}$. Since gravitation is a central force and $\boldsymbol{r} \parallel \boldsymbol{F}$, it vanishes also, and we see that $\boldsymbol{L}$ is constant in time. Consequently, the motion stays within the orbital plane.
As can be seen in Fig. 2.3 b), the area $\mathrm{d}A$ that it swept out by the vector $\boldsymbol{r}$ per unit time $\mathrm{d}t$ can be written as:

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2}\, r \cdot v \sin\alpha = \frac{1}{2}\, |\boldsymbol{r} \times \boldsymbol{v}| = \frac{L}{2m}\,. \tag{2.5}$$

Therefore, the line connecting two bodies (lines from each body to the center of mass respectively) sweeps out area at a constant rate.

3. The last of Kepler's laws gives a relation that allows mass determinations from orbit observations and is valid for any systems, which are dominated by the gravitational two-body interaction. It states that for a relative orbit the square of the period P is proportional to the cube of the semi-major axis. In its generalized form the expression reads:

$$P^2 = \frac{4\pi^2 a^3}{G(m_1 + m_2)} \quad . \tag{2.6}$$

Especially in the limit of $m_1/m_2 \to 0$, when one of the two bodies' masses is negligible, knowledge of a and P directly yields the mass of the host object. This for example can be exploited for the determination of planetary masses, if they are orbited by moons (satellites in general), or the masses of host stars within planetary systems.

In order to entirely characterize the Keplerian orbit in space, one needs to determine the intrinsic geometric properties of the elliptical motion (2 parameters) and the orbit's orientation in space (3 parameters). In addition to that, by the specification of an initial condition of motion, which could be one measurement of the body's time-dependent position and the orbit's period, the trajectory of the body can be determined. Therefore, the following seven orbital elements[5] (comp. Fig. 2.4 and Fig. 2.3 a)) are sufficient to solve for an unperturbed fixed orbit and to predict the body's position for any arbitrary time:

- **Semi-major axis** a as the average of the minimum and maximum distances to the point of reference. For circular orbits the distance $r(t)$ becomes constant and $a = r$.
- **Eccentricity** e as a measurement for the departure from circular geometry is given by

$$e = \sqrt{1 - \left(\frac{b}{a}\right)^2} \quad , \tag{2.7}$$

where b is the semi-minor axis. A circular orbit is defined by $e = 0$.

- The angular distance between the orbital plane and the plane of reference is given by the **inclination** i. When dealing with exoplanets, the plane of reference usually is considered the tangential plane to the celestial sphere at the observed object's coordinates

[5] In literature the given number of orbital elements often varies between five and eight. This is due to different conventions as of what should be counted as an orbital element and depends also on, if the masses of the two bodies are known.

and accordingly perpendicular to the line of sight to the object. The range from 0° to 180° covers each possible orientation. Secondaries on prograde orbits are conventionally defined to have inclinations between 0° to 90°, whereas those, which orbit in the opposite direction to as the primary rotates (retrograde), are assigned to inclinations in the range $90° < i \leq 180°$.

- **Longitude of ascending node** Ω describes the angle between a fixed direction in the plane of reference and the direction to the orbit's ascending node. The ascending node is the location, where the body in orbit passes through the reference plane, in the opposite direction to the observer.[6]
- **Argument of periapsis** ω gives the angle between the line to the ascending node and the line to the periapsis.
- **True anomaly** ν is a time-dependent entity and specifies the angle between the secondary's position to a given time and the periapsis in the plane of orbit.
- **Period** P defines the time for the completion of one full orbit. Furthermore, as stated by Eq. (2.6), it serves as a proportionality factor between the size of the ellipse and the masses involved.

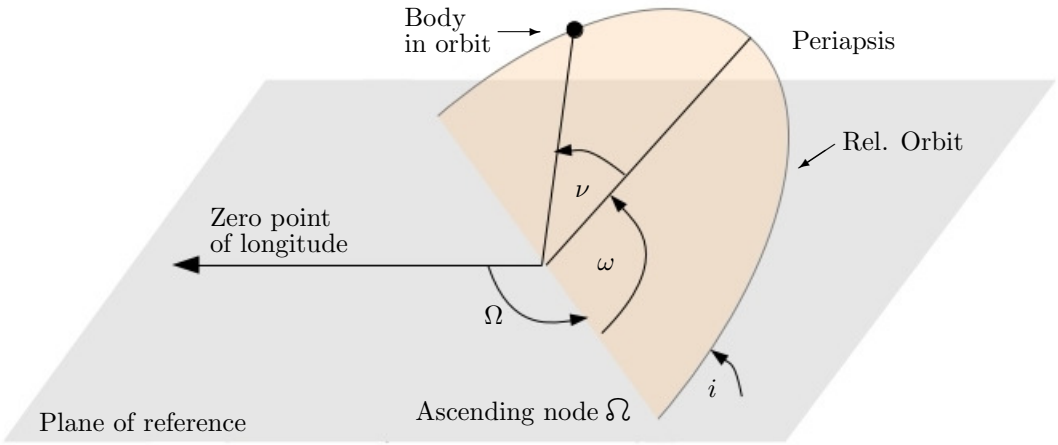


Figure 2.4. – Keplerian orbit in space. The orbital plane intersects the plane of reference along the *line of nodes*, whereas the *ascending* node is the location where the body passes the reference plane, away from the observer. Its direction in reference to a specified direction of zero longitude is called the *longitude of ascending node* Ω. Together with the *inclination* i, the *argument of periapsis* ω and the *true anomaly* ν, the fixed orbit and the position of the object at a given time are uniquely specified.

[6]For orbits in the solar system, the ascending node defines the location of the object passing northwards through the plane of reference.

2.2. Formation of Planets and Planetary Systems

2.2.1. The Solar System

For a long time the solar system was the only one of its kind that could have been studied and, although exoplanetary systems have been discovered by now, it still serves as the best known probe to support theories about the evolution of planetary systems. Therefore, in order to derive assumptions and first ideas on such theories, it is crucial to take a closer look and to summarize the main solar system's components and their properties.
The solar system consists of the sun, as its central star, and various bodies, which are gravitationally bound to it. Despite the high number of bodies in orbit, nearly the whole mass (about 99.8 %) is concentrated in the center, namely in the sun itself with a mass of $M_\odot = 1.9884 \times 10^{30}$ kg, which corresponds to about 333,000 times the mass of the earth.
So, in a first approximation, the system can be regarded as a massive central object surrounded by debris material in orbit. In contrast to the mass distribution, the angular momentum in the system is dominated by the contributions from the planets. In particular the orbital angular momentum of the giant planets Jupiter, Saturn, Neptune and Uranus alone makes up 98 % of the total. Such a distribution is not necessarily typical for gravitationally bound systems. In the planet-moons systems for example the angular momentum is mostly covered by the rotation of the planets. Consequently, there must be an efficient mechanism that allowed mass accretion towards the central body during early formation stages of the system, but on the other hand led to transport of angular momentum outwards in the opposite direction.
The eight planets, which of some properties are listed in table 2.1, are the next most prominent constituents in the system. Their collective mass lies at about $447\,M_\oplus$ and they can be divided into two or even rather three classes. The four terrestrial planets, to which Mercury, Venus, Earth and Mars are counted to, are bodies with relatively high densities. Mostly, they can be characterized by a shell structure with a metallic central core surrounded by a thick mantle and a solid crust that essentially consist of silicates. This kind of differentiation, development into compositionally distinct layers, implies high temperatures and a melted state of the material at some point of time during the formation process. These four rocky planets of the solar system possess atmospheres, whose composition and densities vary widely.
The major fraction of the mass in the debris around the sun is contributed by the four giant planets. Jupiter's and Saturn's main components are hydrogen and helium. Their inner structure can be derived from the analysis of their gravitational field, as determined from flyby space missions *Pioneer*, and *Voyager* and their rotation rate. Both are believed to have a heavy element core of $\gtrsim 10\,M_\oplus$ enclosed by two outer layers, an outer hydrogen-helium molecular atmosphere layer and an inner envelope of liquid metallic hydrogen ([Gui05],[Mil08]). Due to their composition, these two are also called *gas giants*, whereas Uranus and Neptune are often referred to as *ice giants*. Besides a hydrogen-rich atmosphere, covering about one-fifth of the outer planetary radius, the latter two are mainly composed of fluids like water (H_2O), ammonia (NH_3), methane (CH_4) and high temperature condensates of silicates and metals ([dePat01]).
It is worth mentioning that the giants and rocky planets not only differ in their composition and structure, but also can be distinguished by their spatial positions.

The four terrestrial planets are distributed within an inner region of the solar system at distances below 2 AU, while the giants' semi major axes span a region of roughly $2\,\mathrm{AU} \lesssim a \lesssim 30\,\mathrm{AU}$. Obviously, there seems to be a boundary separating regions of terrestrial planet formations from those where giants can form. The term *snow line* (or *ice line*) is often found in literature and describes the point in the protoplanetary disk, by which those two spatial domains are separated. Outside the snow line, whose position depends on the radial temperature profile within the disk and also on the type of the host star (or rather protostar), the disk surface density is increased due to the condensation of hydrogen compound. This situation favors faster and more efficient growth of protoplanetary bodies and provides an explanation for the difference in composition between the terrestrial planets and the outer giants ([Ken08]).

Table 2.1. – Overview of some important orbital and physical properties of the solar system planets. Inclination is given with respect to the ecliptic, the geometric plane of the earth's orbit around the sun. (Data from [dePat01] and [Uns1999]).

	Mercury	Venus	Earth	Mars	Jupiter	Saturn	Uranus	Neptune
Semi-major axis a [AU]	0.387	0.723	1.000	1.524	5.203	9.543	19.192	30.069
Sidereal orbit period P [yr]	0.241	0.615	1.000	1.881	11.857	29.424	83.747	163.723
Eccentricity e	0.206	0.007	0.017	0.093	0.049	0.056	0.046	0.009
Inclination i [deg]	7.005	3.394	0.0	1.850	1.303	2.489	0.773	1.770
Equatorial radius R [$\mathrm{R}_\oplus$]	0.38	0.95	1.00	0.53	11.2	9.41	4.01	3.81
Mass M [$\mathrm{M}_\oplus$]	0.055	0.82	1.00	0.11	317.8	95.2	14.6	17.1
Mean density ρ [$\mathrm{kg\,m^{-1}}$]	5430	5240	5520	3930	1330	690	1260	1640
Obliquity to orbit [deg]	2.0	3.0	23.5	23.9	3.1	26.7	97.9	28.8

Leaving the planets' natural satellites (moons) aside, the next group down the object hierarchy in the solar system are the Asteroids.[7] There are two regions that harbor numerous irregularly shaped *small Solar System bodies* (SSSBs). The *asteroid belt* is the region between the orbits of Mars and Jupiter. According to statistical models and IR observations, it is occupied by an estimated number of around 1.2×10^6 objects with diameters greater than 1 km ([Ted02]) and a vast number of bodies down to the size of dust particles. Most of them are carbona-

[7] Although terminologically distinct from asteroids, dwarf planets are not depicted here separately. So far only five objects (Ceres, Pluto, Haumea, Makemake and Eris) have been categorized as members of this class and basically just due to their sufficient mass to be in hydrostatic equilibrium they set themselves apart from asteroids or TNOs.

ceous but a non-vanishing fraction of about 25 % is counted to silicate and metal-rich. Despite their high numerical number the cumulative mass of the asteroid belt objects lies below 1 ‰ of the earth's mass. In average, the orbits are nearly circular and show only small inclinations. Due to interactions with other bodies within the solar system, gaps within the distribution of the semi-major axes can be found. Those can be explained by orbital resonances, periodically recurring orbital perturbances, which occur when the orbital period is an integer fraction of the disturbing body's orbital period. Within the asteroid belt, the main gaps of this kind are caused by the presence of Jupiter.
Apart from that, regions of stabilized orbits are also found. Many objects gather in groups and populate certain regions of motion, as the co-orbital *Trojan asteroids* (see App. A), located at the gravitationally stable *Lagrangian points* L_4,L_5 of Jupiter, or the *Hilda* group in a $2:3$ resonance with Jupiter.
Beyond the orbit of Neptune, at distances in the regime of 30–50 AU, another reservoir of SSSBs, the so called Kuiper belt can be found. The *Kuiper Belt Objects* (KBOs) may be divided into two major dynamical classes, namely those in interaction with Neptune (in orbital resonance with) and those decoupled from the motion of the planet. The second population is often referred to as the *classical* objects. Within this group there are subsets of KBOs on nearly circular orbits at low inclinations (*dynamically cold*) on the one hand, but also objects with a wider spread distribution of eccentricities and orbits stronger inclined against the ecliptic ([Daw12]). Compared to the inner asteroid belt, the Kuiper belt is larger, more massive and the composition is dominated by volatile ices. Still, even counting in the objects that have been classified as dwarf planets, the total mass of the material within the Kuiper belt does not exceed $0.1\,M_\oplus$ ([Wei09]).
The region beyond that point in the solar system is relatively unknown. Comets, entering the inner parts of the solar system, spend most of their time during an orbital period in those outer regions. The analysis of *short period comets'*[8] orbital characteristics imply that their origin is a place also called the *scattered disk*. That region overlaps with the Kuiper belt but harbors mostly objects on orbits strongly perturbed by Neptune.
Long period comets exhibit a fairly isotropic distribution of orbit orientations, which agrees well with the hypothesis of a quasi-spherical reservoir of icy bodies at long distances up to 10,000 AU or even 100,000 AU. That is known as the *Oort cloud* ([Mor05]).

Collecting the information on the solar system, the aspects constraining the picture of planet formation can be summarized as follows:

- Distribution of mass and angular momentum imply transport of both in opposite directions during early evolution stages.
- Mostly circular, prograde orbits and concentration of objects in the ecliptic plane as a result to a flattened circumsolar disk.
- Spacings in the distribution of the semi-major axes of bodies in orbit as evidence to gravitational interactions (system stability).
- Radially depending properties within the primordial disk may manifest themselves in

[8] The distinction between short and long is rather arbitrary here at about 200 yr.

present regimes of planet composition (terrestrial vs. giants).

- The age of the system may be constrained by radioscope dating of chondritic meteorites with dating of the oldest up to ca. 4.56 Gyr. Comparison to the age of terrestrial rocks may indicate time scales of formation processes. The duration for the final assembly of the earth is believed to be of the order of $\sim$100 Myr ([Bak05], [Arm07]).
- Differentiation of planet material.
- Analyses of the isotopic composition of solar system material reveals, to a first approximation, a well-mixed environment over large distant scales within the protoplanetary disk, although deviations to this homogeneity have been detected and can be used to estimate the age of the system and provide indications to the origin of the material ([dePat01]).
- ...

Without claiming this list to be exhaustive, explaining these listed features is a necessary condition for a complete theory of planet formation. Therefore, those should be kept in mind during the following subsections, where mechanisms and processes, most important in scenarios of planetary system evolution, are outlined.

2.2.1.1. Minimum Mass Solar Nebula

As to the origin of the system, it might be noticed that the comparison of elemental abundances between carbonaceous chondrites to that of the sun's photosphere reveals a remarkable similarity. Mainly, only the most volatile elements show lower concentrations in the meteorites' material ([dePat01], [And1982]), whereas lithium is depleted in the solar photosphere, which might be due to nuclear burning enhanced by diffusion processes under solar-type stars' surface convection layers.[9] Hence, it appears convincing that the sun and also the planets formed from the same primordial material. Following this assumption, the concept of the *Minimum Mass Solar Nebula* (MMSN), as elaborated by [Wei1977], is widely accepted. This protoplanetary disk serves as a starting point to planet formation in the solar system. The original mass distribution of the MMSN in principle is reconstructed by smearing out the planets. To get an idea of the initial total mass, it is assumed that the dust component of the disk has been incorporated efficiently into planets, while on the contrary a huge amount of gas got lost due to the solar wind, photoevaporation or accretion onto the sun. Therefore, typically solar abundances are used to estimate the quantity of this missing ingredient associated with each planet.
In a simple approach, these resulting planet masses are redistributed onto annuli centered on each planet's current orbital distances so that the surface densities can then be found by dividing by the surface areas of the individual annuli. With assumptions on, at what distances different solids, with main focus on water ice, would condense [Hay1981b], the following surface

[9] For interested readers it might be worth mentioning that the surface depletion of lithium in solar-type stars might also be related to the presence of orbiting planets as indicated by ([Isr09]). Still, how mixing and deepening the stars' convective zones could be affected by present planetary systems in detail, is still to be answered.

densities are derived, considering the snow line at the solar distance of $r = 2.7\,\mathrm{AU}$:

$$\begin{aligned}
\Sigma_{\mathrm{rock}}(r) &= 7\left(\frac{r}{1\mathrm{AU}}\right)^{-3/2}\mathrm{g/cm^2} \qquad \text{for:} \quad 0.35 < r/\mathrm{AU} < 2.7 \\
\Sigma_{\mathrm{rock+ice}}(r) &= 30\left(\frac{r}{1\mathrm{AU}}\right)^{-3/2}\mathrm{g/cm^2} \qquad \text{for:} \quad 2.7 < r/\mathrm{AU} < 36 \\
\Sigma_{\mathrm{gas}}(r) &= 1700\left(\frac{r}{1\mathrm{AU}}\right)^{-3/2}\mathrm{g/cm^2} \qquad \text{for:} \quad 0.35 < r/\mathrm{AU} < 36 \quad .
\end{aligned} \tag{2.8}$$

Evaluating the sum of the integrals over the whole range between 0.35 AU and 36 AU yields the disk's total mass of $M_{\mathrm{tot}} \approx 0.013\,\mathrm{M}_\odot$.
Given the radial density profile and assuming axial symmetry for the disk composed from matter on nearly Keplerian orbits so that $v_{\mathrm{K}} = (GM_\odot/r)^{1/2}$, the vertical structure needs to be elaborated in order to get a scheme of the entire disk geometry. This can be achieved by consulting the equation of motion in vertical direction of a mass element $\mathrm{d}m$ at height z over the disk's mid-plane:

$$\mathrm{d}m\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = \mathrm{d}m\, g_z - \mathrm{d}m\frac{1}{\rho}\frac{\mathrm{d}P}{\mathrm{d}z} \quad , \tag{2.9}$$

where g_z is the vertical component of the acceleration due to gravity and the second term represents the opposite force resulting from the pressure within the disk.
Postulating a thin disk so that the distance of the mass element to the sun is of the same order and can be approximated by its radial distance, the first term can be expressed in terms of the Keplerian orbital frequency $\Omega_{\mathrm{K}} = \sqrt{GM_\odot/r^3}$.
Further presuming a temperature dependence only in radial direction ($T(z) = \mathrm{const}$), the isothermal equation of state $P = c_{\mathrm{s}}^2\rho$, with a constant sound speed c_{s}, can be exploited to formulate the equation for hydrostatic equilibrium in direction perpendicular to the mid-plane:

$$0 = -\Omega_{\mathrm{K}}^2 z - c_{\mathrm{s}}^2\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}z} \quad . \tag{2.10}$$

Consequently, the vertical density profile can be derived solving Eq. (2.10).
The temperature profile within the solar nebula is also of high significance as an input feature towards modelling the disk's evolution, but unfortunately is not simple to evaluate, since general various heat sources may play an important role.
However, relying on the simplest case, only considering solar irradiation of an optically thin nebula and leaving aside other processes as viscous heating or secondary irradiation by nearby stars, a radial dependence of the temperature can be derived.
Summing up, the following set of equations can be utilized in order to define a model of the

initial MMSN ([Hay1981b]):

$$\Sigma(r) \approx 1700 \left(\frac{r}{1\mathrm{AU}}\right)^{-3/2} \mathrm{g/cm^2} \tag{2.11}$$

$$T(r) = 280 \left(\frac{r}{1\mathrm{AU}}\right)^{-1/2} \mathrm{K} \tag{2.12}$$

$$\rho(r,z) = \rho_0(r) \exp\left[-\frac{z^2}{2H(r)^2}\right] = \frac{\Sigma(r)}{\sqrt{2\pi}H(r)} \exp\left[-\frac{z^2}{2H(r)^2}\right] \tag{2.13}$$

$$\frac{H(r)}{r} = \frac{c_\mathrm{s}}{\Omega_\mathrm{K}}\frac{1}{r} = \frac{c_\mathrm{s}}{v_\mathrm{K}} = 0.033 \left(\frac{r}{1\mathrm{AU}}\right)^{1/4} \tag{2.14}$$

$$c_\mathrm{s} = 9.9 \times 10^4 \left(\frac{2.34}{\mu}\frac{T}{280\mathrm{K}}\right)^{1/2} \mathrm{cm/s} \quad , \tag{2.15}$$

with the radially dependent mid-plane density ρ_0, the scale height $H(r)$ and the average molecular mass μ. The density profile, as indicated here, is shown in Fig. 2.5.
Due to the density power law in the radial component, the mid-plane density varies over several orders of magnitude between the terrestrial planet regime and the outer regions. This might be important when modelling planet formation in those different regimes and constraining time scales on growth of planetisimals, before the disk becomes depleted in its gaseous component. The ratio of the disk thickness to the radial distance is described by $H(r)/r$ (Eq. (2.14)).
An increase with r is a characteristic to flaring disks and results in the fact that any point on the disk surface receives a substantial amount of stellar radiation. One part of the intercepted energy is radiated away, but another part is re-radiated down into the disk's deeper layers and provides additional heating to the material in the disk interior.

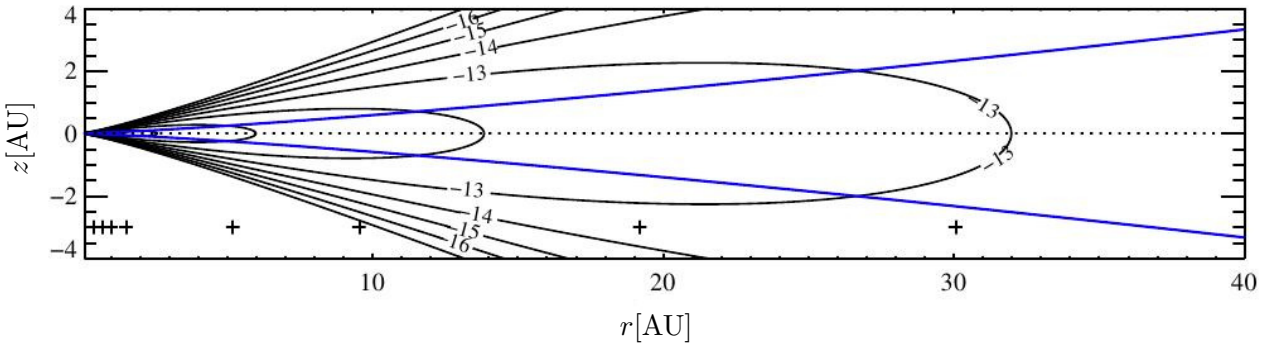


Figure 2.5. – Cross-sectional view of the density profile within the solar nebula as described by Eq. (2.13). Contour lines are denoted by their exponents so that the shown contour line of lowest density equals $\rho = 10^{-17}\mathrm{g/cm^3}$. The blue line highlights the locations where the vertical coordinate z equals the radially depending scale height H. Today's radial solar distances of the eight planets are indicated by the cross symbols. (Credit: [Joh11])

Although the MMSN model seems a reasonable approach to the characterization of typical protoplanetary disks, it should be clear that its simplistic foundation is limited.
As [Des07] argues, the total mass is derived by the assumption that the planets have accreted all the solid material in their neighborhood and that if they did not manage to do so within the disk lifetime, the total mass would be considerably underestimated. Besides, possible migration scenarios of planetisimals during their growing phase are not taken into account.
Nevertheless, with those limitations in mind, the model can be used as a first starting point in studying the properties of protoplanetary disks.

2.2.2. From Cold Clouds to Planetary Systems

2.2.2.1. Star Formation and T-Tauri Phase

Considering all implications from our understanding of the present solar system, a complete model describing planet formation needs to cover the origin of the system as a whole, as well as its development in time. Fig. 2.6 depicts a comprehensive scheme for the formation of low mass stars, as commonly found in literature, e.g. in [Hay1981a] or [Shu1987]. The entire process can be divided into subsequent evolutionary stages. During the first period, the progenitor of the system emerges from the gravitational collapse of accumulated matter. Such star forming regions can be found in *Giant Molecular Clouds* (GMCs) within the galactic plane. With typical masses of $10^4 - 10^6 \mathrm{M}_\odot$ and sizes up to 100 pc those structures mainly consist of cold H_2 molecules with a low mass fraction (about 1 %) of embedded micrometer-sized interstellar dust particles, which also can serve as condensation nuclei for the accumulation of more complex molecules. Prominent regions of star formation are the *Orion Molecular Cloud* (OMC), the *Rho Ophiuchi cloud complex* or the *Taurus Molecular Cloud* (TMC).

Centrally condensed and slowly rotating cores build within molecular clouds ([Liz1987]). As pointed out by [Cha1939], the density distribution of such a core can be approximated by a state of a singular isothermal sphere, in which the self-gravity is compensated by a thermal pressure gradient

$$\rho = \frac{c_\mathrm{s}^2}{2\pi G r^2} \quad , \tag{2.16}$$

with the isothermal sound speed $c_\mathrm{s} = (kT/m)^{1/2}$, where m is the mean molecular weight of the gas and T the constant temperature.
This configuration is unstable to gravitational contraction and has a self-similar[10] collapse solution ([Shu1977]). Various trigger factors could brake the marginal stability of the core. Local deviations in temperature, shock waves due to supernovae or strong stellar activity in the neighborhood are able to result in local density fluctuations which would lead to a self-energizing collapse. The critical state of a mass distribution, unstable to its own gravity, has

[10] In this context self-similarity describes the feature that solutions to any instant of time differ only in a scaling factor.

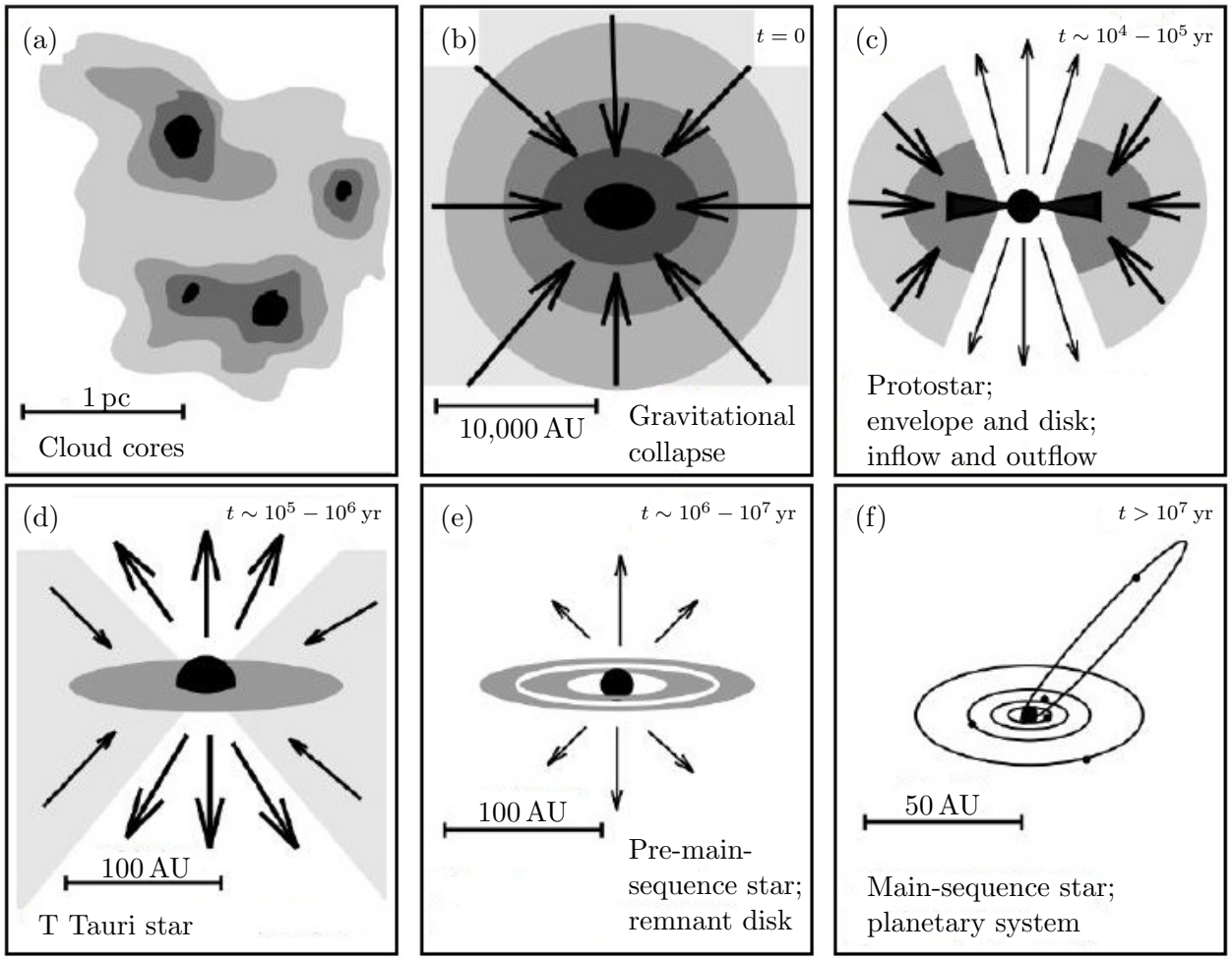


Figure 2.6. – The sequence describes the process of star formation and divides it into consecutive stages. (a): Cores of increased density within molecular clouds eventually collapse due to gravitational instability (b). A protostar forms in the central region, embedded in an envelope of gas and dust (c), evolving into an axisymmetric structure due to initial angular momentum of the contracting material. During further evolution, which is affected by an active phase of matter accretion and outflows, the system develops into a young stellar object with a circumstellar disk. The SEDs of such T-Tauri stars (d) show characteristic features resulting from an approximated black body spectrum from the stellar source overimposed by dust emission at longer wavelengths and observed in the IR. Following this scheme, the system's gas content is being dissipated, leaving behind a star evolving to the main-sequence and a planar remnant protoplanetary disk (e), within which companions can form, if dust material is accumulated efficiently enough (f). (Original illustrations by *M. Hogerheijde* from [Joh11])

been investigated by *J. Jeans* ([Jea1902]). Assuming a spherical, isothermal and homogeneous gas configuration with a total mass M, criteria, under which the equilibrium between gas pressure and gravitation breaks, can be formulated. For this purpose the virial theorem for

systems under gravitational attraction, stating that the time average of the system's kinetic energy equals half of the average negative potential energy, is applied:

$$\langle E_{\mathrm{kin}} \rangle = -\frac{1}{2} \langle E_{\mathrm{pot}} \rangle \quad . \tag{2.17}$$

The kinetic energy is expressed by the total thermal energy of the ideal gas E_{T} and the potential energy by the gravitaional energy E_{G} of a homogeneous sphere:

$$E_{\mathrm{T}} = \frac{3}{2} nkT = \frac{3}{2} \frac{kT}{\mu m_{\mathrm{u}}} M \tag{2.18}$$

$$E_{\mathrm{G}} = -\frac{3}{5} \frac{GM^2}{R} \quad . \tag{2.19}$$

The situation, where thermal pressure cannot compensate self-gravity, is correlated to the condition $2E_{\mathrm{T}}/-E_{\mathrm{G}} < 1$. With Eq. (2.18) and (2.19) this yields limiting stability criteria for the properties of the mass distribution ([Uns1999]):

$$R_{\mathrm{J}} = \frac{1}{5} GM \frac{\mu m_{\mathrm{u}}}{kT} \tag{2.20}$$

$$M_{\mathrm{J}} = 5.46 \left(\frac{kT}{\mu m_{\mathrm{u}} G} \right)^{3/2} \rho^{-1/2} \quad . \tag{2.21}$$

When the Jeans-mass M_{J} is exceeded or the spatial extension falls below the Jans-radius R_{J}, the collapse is initiated. The infall of the matter is triggered by an expansion wave propagating at sound speed from inside out. Within around $0.4\, r_h$, where r_h is the radius of the expansion wave at a given time, matter inflow happens at a supersonic level, and the density in this region can be described by the free-fall form ([Shu1987])

$$\rho(r) = \frac{\dot{M}}{4\pi\sqrt{2GM}} r^{-3/2} \quad , \tag{2.22}$$

where $\dot{M}$ is the rate, at which matter is accumulated by the central object.
Hence, the time scale of the molecular cloud cores collapse is relatively short. For typical cloud densities the duration of the free-fall phase is estimated to some $10^4 - 10^5\,\mathrm{yr}$, which makes it difficult to find and observe objects in this stadium.
The simple approach to the situation does not take rotation and conservation of angular momentum into account. Rotation can be included as a small perturbational effect and leads to the formation of a rotating nebular disk around the central object ([Ter1984]). The initial specific angular momentum is stored in the disk, while the core in the inner region can maintain contraction. The flatness of the disk and the outcome of the axisymmetric structure is a direct consequence to the non-vanishing angular momentum. Centrifugal forces damp matter infall in directions perpendicular to the system's rotation axis, while vertical collapse is not influenced. The free-fall contraction can be sustained, inasmuch as heat is dissipated effectively, so that the process remains predominantly isothermal. Dust particles, observed in infrared, but also carbon monoxide, water or molecular oxygen can absorb the increasing thermal energy during

the contraction and re-radiate it. These cooling mechanisms work sufficiently well, and energy is depleted from the system, as long as the surrounding material is optically thin. At some point the density in the bulk region becomes critical, and due to the increasing opacity gas pressure stabilizes the collapse of the inner core region. During the following adiabatic phase, further contraction can only proceed, if any other cooling mechanisms become effective.
When temperatures of several 1,000 K are reached, endothermic dissociation of H_2-molecules begins and drives the next phase of contraction until the entire hydrogen of the central region is dissociated and even ionized. By the end of this stage, the system consists of a compact protostar, nearly in hydrostatic equilibrium, embedded in a surrounding envelope of gas and dust showing axially symmetric features due to the initial rotation. The masses of those protostars lie at about $0.001\,\mathrm{M}_\odot$ and the core temperature around 10,000 K.
In order to start nuclear burning the temperatures within the center of the protostar must increase at least beyond around 10^6 K, which is a condition to deuterium burning. The central temperature for a convective protostar depends on its mass and radius and is estimated to $T_\mathrm{C} = 0.54 G M_* \mu H / k R_*$ ([Cha1939]). To reach this condition more matter needs to be accreted. At this point the progression past this stage becomes more difficult.
The protostellar disk features differential rotation and is dynamically stable with a radially increasing angular momentum profile $\mathrm{d}L/\,\mathrm{d}r > 0$. In other words, disk material can only be transported inwards towards the center when losing angular momentum. Accordingly, the evolution of the system and its *Young Stellar Object* (YSO) in this era is highly influenced and regulated by the efficiency of transferring angular momentum away from the accreted matter. This problem is not solved in theory and various mechanisms have been suggested. In the following subsection (see p. 22) the problem and one of the approaches towards understanding will be outlined. In principle, angular momentum must be redistributed within the disk itself or be drained from the system at all. This might be related to observed energetic outflows from YSOs. Observations strongly imply that simultaneous matter infall (accretion) and outflow are linked together and mark a typical intermediate state during YSO evolution. Outflows usually appear as well collimated and bipolar structures, breaking through at the regions of the rotational poles of the accreting protostar. There the total column of infalling matter is at its minimum, and therefore these regions provide outflow channels of least resistance. Bipolar outflow is encountered in two most prominent manifestations. Molecular high velocity CO outflows (e.g. [Bal1983]) have been detected at YSOs of different masses and luminosities and are believed to consist of ambient gas swept up by underlying wind. Narrow jets of ionized material, often referred to as *Herbig-Haro objects*, are observed in optical and centimeter-wavelengths (e.g. [Mun1983]), and are sometimes associated with molecular outflow. An overview of outflows around young protostars can be found in [Bac1996].
During the phase of deuterium burning, which is highly sensitive to temperature, the protostar is largely convective so that the generated heat can be carried away into outer shells or even the surface, while the core is supplied with unconsumed, newly accreted material so that a continuous energy production is ensured. In this way, the process acts as a thermostat to the system so that hydrogen burning is delayed ([Sta1988]). Once deuterium is exhausted, or radiative energy transport establishes a boundary layer so that the core is cut off from the outer regions, the nuclear fusion process ceases and the last phase of contraction is initiated. This lasts until temperatures sufficient for hydrogen burning are reached and the star, finally

reestablishing hydrostatic equilibrium again, appears at the *Zero Age Main Sequence* (ZAMS) in the Hertzsprung-Russell diagram.

It is reasonable to argue that after the ignition of deuterium burning and with an established matter outflow the outflow opening angle will widen with time and the protostar becomes optically visible. This marks the beginning of the star's *T-Tauri* phase. YSOs that fall into this class, firstly defined by [Joy1945] and later revised (e.g. [Bas1983]), show quite a broad spectral energy distribution with excesses in UV and particularly in the infrared. Irregular and quasi-periodic variations at various wavelengths and the spectral properties in general support the picture of a recently ($< 10^7$ yr) formed, usually active low mass ($< 3\,M_\odot$) stellar object with strong indications for the presence of circumstellar matter and complex accretion and outflow features ([App1989]). Modelled star, respectively disk spectra, which imply that an optically thick disk intercepts and re-radiates fractions of the original star luminosity, agree with observed T-Tauri spectra, and the disk's equilibrium temperature profile can be approximated by $T \propto r^{-3/4}$, which corresponds to a power law form of $\nu L_\nu \propto \nu^{-4/3}$ at long wavelengths of the SEDs. The latter component arises from the emission of dust within the disk.
The dust contingent makes around 1% of the disks' total masses, which typically lie in the range of $10^{-4} - 10^{-1}\,M_\odot$. Spectral profiles, flatter or even constant in the long wavelength regime, are believed to be outcomes from more extended, massive disks of $0.1 - 1.03\,M_\odot$. Such circumstellar matter distributions, with intrinsic disk luminosities of the order of the stellar luminosity, possess flatter rotation curves than that of a classical Keplerian accretion disk and a disk temperature profile less steep than $T \propto r^{-3/4}$.
During the active T-Tauri phase, with a lifetime of the order of $10^5 - 10^6$ yr and thereafter, while the stellar object accretes material at variable rates and traverses its pre-main-sequence phase before reaching the main-sequence, the surroundings are steadily cleared. Altogether, due to accretion processes onto the protostar, stellar winds and photoevaporation of stellar flux from stars in the neighbourhood, the gas component of the disk dissipates on timescales of up to 10^6 yr. This is an important temporal constraint on mechanisms driving planet formation, especially when regarding the gas giants, which need to have accumulated their material during this 'short' period of time.

2.2.2.2. Circumstellar Disks

When studying the physics of circumstellar disks around young stars it is substantial to consider the accretion rate onto the central object. Disks, mainly powered by the release of gravitational potential energy of inflowing matter, are usually called *active*, whereas the term *passive* marks disks, whose luminosities are primarily driven by reprocessed starlight. The different energy sources need to be considered when modelling temperature profiles or SEDs from those objects. The estimation of accretion rates from observations is a complicated task. Irradiation from the central star is a larger heat source than the disk's self-luminosity for the most T-Tauri objects. Consequently, the accretion rates are usually determined from the hot continuum excess radiation that is produced when matter reaches the stellar surface. In order to measure that fraction of radiation, one needs to distinguish between stellar photospheric emission and the

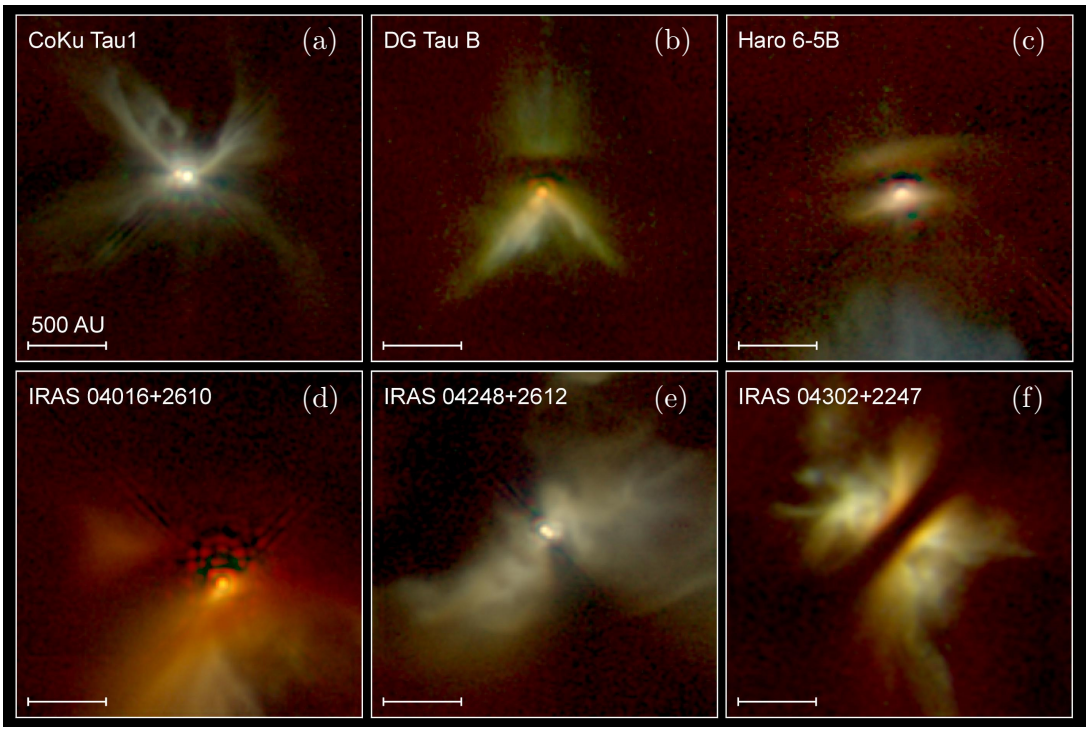


Figure 2.7. – Young stars at distances of around 140 pc in the constellation Taurus taken by NASA Hubble Space Telescope's Near-Infrared Camera and Multi-Object Spectrometer (NICMOS). The central YSOs (binary stars in (a) and (d)) are surrounded by nebulae of gas and dust. Dense parts of the disk reveal themselves as dark bands in the edge-on images of the systems, and surrounding dust particles are partially illuminated by the host star. (Credit: D. Padgett (IPAC/Caltech), W. Brandner (IPAC), K. Stapelfeldt (JPL) and NASA; [Hub1999])

emission induced by accretion, which can be a source of systematic errors. Accretion rates, spectrophotometrically derived from measurements of strong emission *classical T-Tauri stars* (CTTSs) vary between $10^{-9}\,M_\odot\,yr^{-1}$ and $10^{-7}\,M_\odot\,yr^{-1}$ ([Gul1998]). Critical accretion rates, defining the limiting boundary between passive and active disks, are of the same order, which makes pure distinctions complicated.
The intrinsic properties of circumstellar disks, as consequences of the kind of the dominant energy source, are rather expected to be depending on time and spatial distance to the center. At small orbits and at early epochs internal heating should dominate, while for more evolved systems and/or at large radii reprocession might be more significant ([Arm07]).

Actively Accreting Disks

As already stated before, accretion within protoplanetary disks is an important mechanism, but also demands for transport of angular momentum. Following the summary of [Arm07], an overview of standard considerations regarding this problem is outlined here.

In the context of accretion disks it is often stated and assumed that the matter's orbital velocity can be approached as Keplarian to the first order and that the specific angular momentum $l = L/m$ can analogously be described by the increasing function of radius:

$$l(r) = rv = r^2\Omega = \sqrt{GM_* r} \quad . \tag{2.23}$$

This simplification is justified by considering the radial force balance for the gas given by

$$\frac{v_\phi^2}{r} = \frac{GM_*}{r^2} + \frac{1}{\rho}\frac{\mathrm{d}P}{\mathrm{d}r} \quad . \tag{2.24}$$

With the orbital velocity v_ϕ, the centrifugal force (left-hand side) is compensated by gravity and the radial pressure gradients of the gas. The pressure gradient term can be further estimated by:

$$\begin{aligned} \frac{1}{\rho}\frac{\mathrm{d}P}{\mathrm{d}r} &\sim -\frac{1}{\rho}\frac{P}{r} \\ &\sim -\frac{1}{\rho}\frac{\rho c_\mathrm{s}^2}{r} \\ &\sim -\frac{GM_*}{r^2}\left(\frac{H}{r}\right)^2 \quad , \end{aligned} \tag{2.25}$$

where the relation $H = c_\mathrm{s}/\Omega$ between the scale height and the sound speed, as was given by Eq. (2.14), is used. Now, the velocity can be expressed by the Keplerian velocity v_K and a contribution from the pressure gradients, which can be considered as negligible for geometrically thin disks with $h \ll r$:

$$v_\phi^2 = v_\mathrm{K}^2\left[1 - \mathcal{O}\left(\frac{H}{r}\right)^2\right] \quad . \tag{2.26}$$

Based on the Keplerian differential rotation profile, one can think of mechanisms that would lead to a redistribution of the angular momentum within the disk. Molecular viscosity between fluid elements on slightly different orbit distances r and $r + \delta r$ might be a first guess mechanism to accomplish this. While rubbing against each other, the outer particles, slower at first according to the velocity profile, are accelerated in their moving direction, while on the contrary the inner are decelerated. In result, the outer ring particles are raised to a higher orbit, by which they are slowed down again, where the inner ring speeds up by falling to a lower orbital distance. This process obviously supports transport of mass as well as angular momentum and is a source to heating in the disk. Depending on the opacity of the disk material, the released energy can either be radiated away efficiently so that the disk becomes cool and thins out, or if the heat is stored, then it rather puffs up into a thick torus.
The temporal evolution of the disk, being subject to viscous forces, has been studied as a possible model, driving accretion onto the central objects. Utilizing the continuity equation for axisymmetric flows and the conservation of angular momentum under the effect of torques acting on the fluid due to viscous stresses, [Pri1981] derived an evolution equation for the surface density $\Sigma(r,t)$ of a thin Keplerian accretion disk:

$$\frac{\partial\Sigma}{\partial t} = \frac{3}{r}\frac{\partial}{\partial r}\left[r^{1/2}\frac{\partial}{\partial r}\left(\nu\Sigma r^{1/2}\right)\right] \quad , \tag{2.27}$$

where ν stands for the kinematic viscosity.
By changing the variables to $X \equiv 2r^{1/2}$ and $f \equiv 3/2\,\Sigma X$ and presuming ν to be constant, Eq. (2.27) takes the standard form of a diffusion equation:

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial X^2} \quad , \tag{2.28}$$

where $D = 12\nu/X^2$ is the diffusion coefficient.
From Eq. (2.28) the characteristic diffusion time scale can be estimated to $\tau = X^2/D$ and becomes

$$\tau_{\mathrm{vics}} \simeq \frac{r^2}{\nu} \quad , \tag{2.29}$$

when expressed by the original physical variables.
In order to solve Eq. (2.27) the viscosity needs to be defined. Although it is expected to be somehow dependent on local disk properties, such as temperature, surface density or other, analytical solutions, which already give indications to the system's evolution, can be found when the viscosity is approximated by a power-law in the orbital distance $\nu \propto r^{\gamma}$ ([Lyn1974]). Fig. (2.8) shows solutions for two simplifications of this problem. On the left, the viscosity is fixed to be constant and the graph shows the evolution of the surface density to an initially thin ring of gas situated at radius r_0 at $t = 0$. The different curves describe the evolved density distributions for subsequent instants of time, defined by the dimensionless variable $\tau = 12\nu r_0^{-2} t$ and given by the solution:

$$\Sigma(r,\tau) = \frac{m}{r_0^2}\frac{1}{\tau}\left(\frac{r}{r_0}\right)^{-1/4} \exp\left[-\frac{1+(r/r_0)^2}{\tau}\right] I_{1/4}\left(\frac{2(r/r_0)}{\tau}\right) \quad , \tag{2.30}$$

where $I_{1/4}$ is a modified Bessel function of the first kind and m is the mass of the gas in the ring. One can see that with increasing time the matter spreads and the curves' barycenters move towards smaller radii. In fact, as t approaches infinity, most of the mass flows towards $r = 0$, while the angular momentum, coupled to an negligible fraction of the matter, flows towards $r = \infty$.

The second set of curves (on the right) depicts the situation, where the viscosity is set to be linear in radius ($\nu \propto r$). As an initial condition, the surface density profile is defined by a steady-state solution with $\Sigma \propto r^{-1}$ out to the radius r_1 and an exponential cut-off for larger distances:

$$\Sigma(t=0) = \frac{C}{3\pi\nu_1}\left(\frac{r_1}{r}\right) e^{-r/r_1} \quad , \tag{2.31}$$

with ν_1 the viscosity at radius r_1 and a normalization constant C. The distribution's progress is illustrated by the initial profile at $T(t=0) = 1$ and successive curves for $T = 2$, $T = 4$ and $T = 8$, where $T \equiv 1 + 3(r_1^2/\nu_1)^{-1} t$ is a variable of scaled time. The corresponding solution reads as:

$$\Sigma(r,T) = \frac{C}{3\pi\nu_1 T^{-5}}\left(\frac{r}{r_1}\right)^{-1} \exp\left[-\frac{(r/r_1)}{T}\right] \quad . \tag{2.32}$$

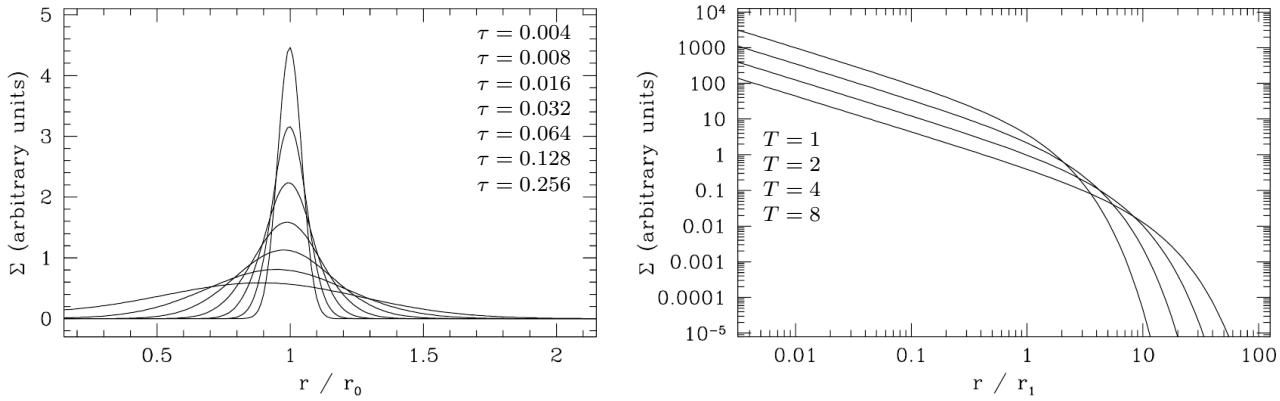


Figure 2.8. – Analytical solutions to the evolution of the surface density profile of a thin accretion disk due to molecular viscosity. Net mass transport towards the center and transfer of angular momentum to larger radii is implied. *Left:* The curves (from up to bottom) represent the temporal redistribution of mass, initially placed in a thin ring structure at radius $r = r_0$ for $\nu = \text{const}$, as described by the solution in Eq. (2.30). *Right:* Self-similar evolutionary profiles in surface density (from up to bottom), as given by Eq. (2.32). Here the viscosity was defined as linear in radial distance and the initial distribution follows $\Sigma \propto r^{-1}$ to a characteristic scale radius r_1 and is exponentially cut-off at larger distances. (Credit: [Arm07])

The most important features to note here are that, while temporally evolving, the disk's mass decreases and the angular momentum is conserved as the characteristic distribution scale increases.
These analytical considerations can be beneficial for deriving disk properties, such as masses or the viscosity from observations (e. g. [Har1998]), but more importantly suggest that molecular viscosity processes in principle are able to drive accretion and the required transport of angular momentum within circumstellar disks.
Unfortunately, considering the time scale for molecular viscous processes, as implied by Eq. (2.29), the prospect that this mechanism could play a key role in this matter is less likely. The viscosity for a gas with the mean free path $\lambda = 1/n\sigma$ is of the order $\nu \sim \lambda c_\mathrm{s}$. With appropriate estimates ([Arm07], [Joh11]) on a typical protoplanetary disk's scale, density, sound speed and collision cross-sections of the molecules within, the viscosity can be approximated to the magnitude of the order of $\nu \sim 10^3\,\mathrm{m^2\,s^{-1}}$, implying a time scale for the disk evolution of about $10^{14}\,\mathrm{yr}$, which definitely is out of observational indications or even realistic considerations.
Hence, molecular viscosity is regarded negligible in protostellar disks.

Seeking for alternative mechanisms, [Sha1973] proposed turbulences within the disk as a source to providing sufficient viscosity.
Within this framework of the so called α-*disks*, the expression for the turbulent viscosity takes the parametrized form:

$$\nu_\mathrm{t} = \alpha c_\mathrm{s} H = \alpha \Omega_\mathrm{K} H^2 \quad , \tag{2.33}$$

and the typical evolution time scale due to turbulent viscosity consequently becomes:

$$\tau_{\mathrm{t}} \simeq \frac{r^2}{\nu_{\mathrm{t}}} = \frac{1}{\alpha \Omega_{\mathrm{K}}} \left(\frac{H}{r} \right)^{-2} \quad . \tag{2.34}$$

The parametrization in Eq. (2.33) is founded by the assumptions that the dimension of any isotropic turbulence cells would be similar to the scale height H and that the velocity of turbulent flows would not exceed the sound speed, since supersonic motions would evoke shocks and dissipation of the kinetic energy into heat.
When α is treated as a constant, which is rather an illustrative and probably oversimplified academic approach, an expression for the viscosity as a function of the disk parameters r, Σ and α can be derived, and, together with the evolution equation (2.27), a complete system for the time dependence of the disk can be established, where α is the only unknown parameter.
Although this method delivers a simple way of modelling, the underlying physics are completely hidden within the parameter α, which in general could also be a more complicated function. Nevertheless, taking that into account and treating the whole approach with caution, values for α can be estimated. With similar disk characteristics as assumed above for the estimation of the viscous time scale, and by constraining the turbulent evolution time to a realistic limit of about 10^6 yr, the order of magnitude $\alpha \simeq 0.01$ is obtained. Those estimations are passably confirmed by attempts to constrain α from observations ([Har1998]).

With these arguments in mind, mechanisms leading to turbulences need to be found and have been studied analytically and in experiments during the recent decade (e.g. [Bal06], [Afs05], [Ji2006]). Substantially to note is the fact that Keplerian disks fulfill the condition for hydrodynamical stability of differentially rotating fluids, the Rayleigh criterion:

$$\frac{\mathrm{d}}{\mathrm{d}r}(r^2 \Omega) > 0 \quad . \tag{2.35}$$

Hence, when the specific angular momentum increases with radius, axisymmetric waves are linearly stable and fluid displacements in the disk plane exhibit oscillations around their equilibrium circular orbits at a characteristic frequency. Analysis of hydrodynamic perturbances that could lead to nonlinear instabilities imply that those would eventually decrease rather than drive steady turbulent viscosity.
Frequently also discussed, as another possible source to turbulent motions within the circumstellar matter, is the disk's self-gravity ([Too1964]). Non-axially symmetric modes, such as trailing spiral arms, permit accretion flow by providing sufficient dissipation of work performed by produced torques and drive redistribution of angular momentum ([Cla09]). The major conclusion is that self-gravity might be important during the early stages, when the disk is still massive enough, but not in more evolved systems.
The most promising candidates for providing an efficient way of producing turbulences and transport of angular momentum respectively result from interactions between the disk matter and magnetic fields.
As shown by [Bal1991], in presence of an magnetic field the stability criterion (compare to Eq. (2.35)) becomes stricter:

$$\frac{\mathrm{d}}{\mathrm{d}r}(\Omega^2) > 0 \quad . \tag{2.36}$$

In a magnetohydrodynamic (MHD) flow the angular velocity must increase with radial distance, a condition not fulfilled by a Keplerian velocity profile. Initially small perturbations amplify and generate self-sustaining turbulent motions. The underlying mechanism, known as *magnetorotational instability* (MRI) due to matter-field coupling can be understood by the following simplified picture ([Bal1998]). Given a magnetic field infiltrating the disk, two mass elements at the same orbital radius r and joined by a vertical field line are considered. The illustration in Fig. 2.9 shows the outcome from an initial perturbation, displacing one of the elements slightly inwards ($m_{\rm in}$) and the other ($m_{\rm out}$) in the opposite direction. When the coupling is sufficient enough, a magnetic tension between those two is induced. Consequently, due to $v_{\rm K} \propto r^{-1/2}$, angular momentum is continuously transferred from the inner element to the outer one. This leads to $m_{\rm in}$ dropping to an even lower orbit, whereas $m_{\rm out}$ is excited to higher r. Following that scenario, the separation between the particles increases as well as the tension and the effect is magnified. This growing nature of the weak-field instability assures sufficient coupling and

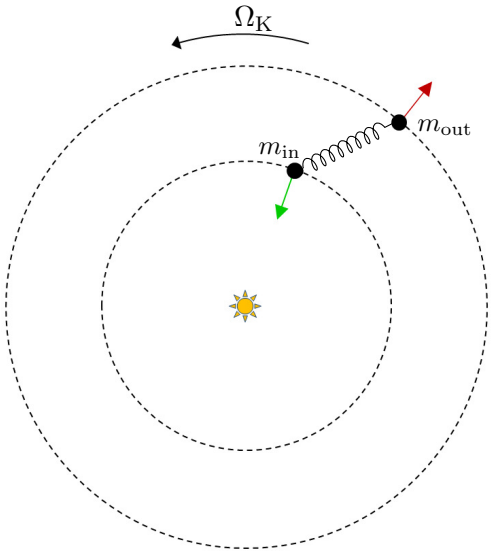


Figure 2.9. – The schematic picture of the mechanism inducing turbulent viscosity by the magnetorotational instability depicts the face-on view of a Keplerian accretion disk with two mass elements, initially at same orbital distances, shortly after a small displacement in radial direction. The particles are connected by magnetic field lines coupling to the gas. Shearing occurs due to differential rotation ($v_{\rm in} > v_{\rm out}$) so that tension is induced by the field lines dragged along with the fluid elements. Consequently, angular momentum is transferred from $m_{\rm in}$ to $m_{\rm out}$, raising the outer element to a higher orbit and making the inner one move inwards. The tension increases and the effect is amplified.

its effectivity, even for low ionization fractions and a non-vanishing resistivity, which in general would damp the MRI. Neverthelss, with indications of magnetic activity in YSOs (e.g. [Gla00]), non-thermal ionization processes are required to drive this mechanism.
Stellar X-ray radiation from the host star (e.g. [Hay1981b], [Fei07]), or even cosmic rays pen-

etrating the disk ([Ume1981]), could contribute to that.
While the MRI is potentially successful in redistributing angular momentum within the accretion disk, magnetically induced outflows ([Pud07], [Sal11]) from the disk surface could cause its depletion from the system itself. All in all, magnetic fields and their interaction with the circumstellar matter are most probably critical to the evolution of protoplanetary disks and other systems, in which accretion is conveyed.

2.2.2.3. Formation of Planetisimals

With the development of a circumstellar disk around protostars, as described above in the context of star formation, the main ingredients for the formation of planets are provided. Still the whole remaining process is only partially understood to a satisfying level. This results from its complexity and diversity. In order to obtain planetary companions, they must successfully grow over many orders of magnitude, at which different types of interactions with the environment take place. Furthermore, the evolution is bound to timing constraints. Planetary embryos and especially gas giants need to form considerably before the host star enters its active T-Tauri phase and a vast amount of disk material is depleted.

Condensation, Sedimentation and Coagulation

As a starting point for building solid bodies within the environment of the protostar, one usually considers the disk material to consist mainly of gas (hydrogen and helium) with a small fraction of dust around $m_{\mathrm{dust}}/m_{\mathrm{gas}} \approx 0.01$. The μm-sized dust particles partly originate from the initial molecular cloud itself, but also from condensations within the disk. Although the disk material is well-mixed primarily, differentiation processes in radial direction should occur due to the radius dependent temperature profile. Whereas mineral aggregates can form down to short central distances, more volatile molecules such as water, ammonia or methane condense only at outer regions, where the temperatures are moderate.
At this level the dust is tightly coupled to the gas due to drag forces F_D in opposing direction to its relative motion velocity v ([Wei1977b]):

$$F_D = -\frac{1}{2} C_D \cdot \pi s^2 \cdot \rho_s v^2 \quad , \tag{2.37}$$

where C_D is a dimensionless drag coefficient and s, ρ_s are the solid particles' size and density. For particle sizes within the *Epstein regime*, where those are small compared to the mean free path of the gas molecules, the drag coefficient can be expressed as a function of the mean thermal velocity $\bar{v}$ in the gas $C_D = \frac{8}{3}\frac{\bar{v}}{v}$. Using this, the settling speed and time scale of dust in vertical direction can be estimated, which yields that at μm scale the dust is able to sediment to a thin mid-plane layer in the disk ([Arm07]).
The efficiency of this sedimentation in presence of turbulences, providing diffusion and dust trapping in turbulent cells, are still a matter of discussion and analysis (e.g. [Joh05], [Car06]). However, as a consequence from sedimentation or any other processes that lead to relative motions (turbulences, Brownian motion etc.), collisions between particles are inevitable and provide the basic principle for growth.

At this point, van der Waals forces allow dust to agglomerate after colliding. Unfortunately, estimating growth rates is not that trivial. Basic approaches (see [Dull05]), neglecting any turbulences and assuming ideal particle adhesion following encounters, show that, while settling towards the mid-plane, coagulation is largely effective and can generate small particles of up to even mm-sizes quickly enough. In fact, this simple ansatz presents itself as too ideal, since the rapid runaway growth occurs at the expense of digesting small grains and leads to a lack of those in the distribution of particle sizes, inconsistent with the observed IR-excess in the SEDs of young T-Tauri objects. This is of no surprise, as in general the outcome of collisions is highly sensitive to input parameters and can lead to fragmentation rather than sticking. An extensive review on experimental results from collisions under various conditions regarding particle sizes, porosities and velocities can be found in [Gütt10]. A schematic overview of possible outcomes of collision experiments is given by Fig. 2.10. In principle, sticking of particles is expected for moderate velocities below $\sim 1\,\mathrm{m/s}$ and for high speeds only, if the collision partners considerably differ in size and the larger one is not porose.
This introduces a difficulty in the coagulation models. While particles of small sizes grow after low-velocity encounters, high speeds tend to disrupt the building bodies. When considering possible relative velocities of even above $30\,\mathrm{m/s}$ ([Wei1977b]) for bigger (meter-sized) bodies in protostellar disks, this is a hard constraint, commonly referred to as the *meter-barrier* of planetisimal growth.

Another complication to successfully modelling the planetisimal formation arises from the interaction between the dust (bodies at posterior stages) and the gas flow. The effect, known as *radial drift*, describes the spiralling down of solids towards shorter radii and is examined in two different regimes ([Wei1977b]). To both mechanisms the radial pressure gradients in the disk is decisive. The resulting forces on the gas act radially outwards and lead to a reduced net force, compared to, if only gravitation would be considered. Consequently, the orbital frequency Ω of the gas is slightly reduced to that of a Keplerian orbit at the same radius ([Joh11]):

$$\Omega^2 r - \Omega_\mathrm{K}^2 r = \frac{1}{\rho}\frac{\partial P}{\partial r} \quad . \tag{2.38}$$

Small solids, still coupled to the gas, orbit at the same sub-Keplerian velocity but do not experience the same radial pressure so that they drift towards smaller orbital distances due to a net inward force.
While bodies coagulate and grow in size, they also slowly, but continuously, decouple from the gas. These objects orbit at Keplerian frequencies and accordingly encounter a transverse gaseous headwind forcing them to decay in orbit at the radial drift speed v_drift ([Joh11]):

$$v_\mathrm{drift} = -\frac{\Delta v}{\Omega_\mathrm{K} t_\mathrm{fric} + (\Omega_\mathrm{K} t_\mathrm{fric})^{-1}} \quad , \tag{2.39}$$

where Δv resembles the difference between the velocity on a Keplerian orbit and that of the gas, and $t_\mathrm{fric} = mv/|F_D|$ (compare to Eq. (2.37)) is the friction time scale, on which drag forces on the particle would change the relative velocity by the order of unity.
The determination of maximum radial drift speeds yields that bodies of sizes in the range of

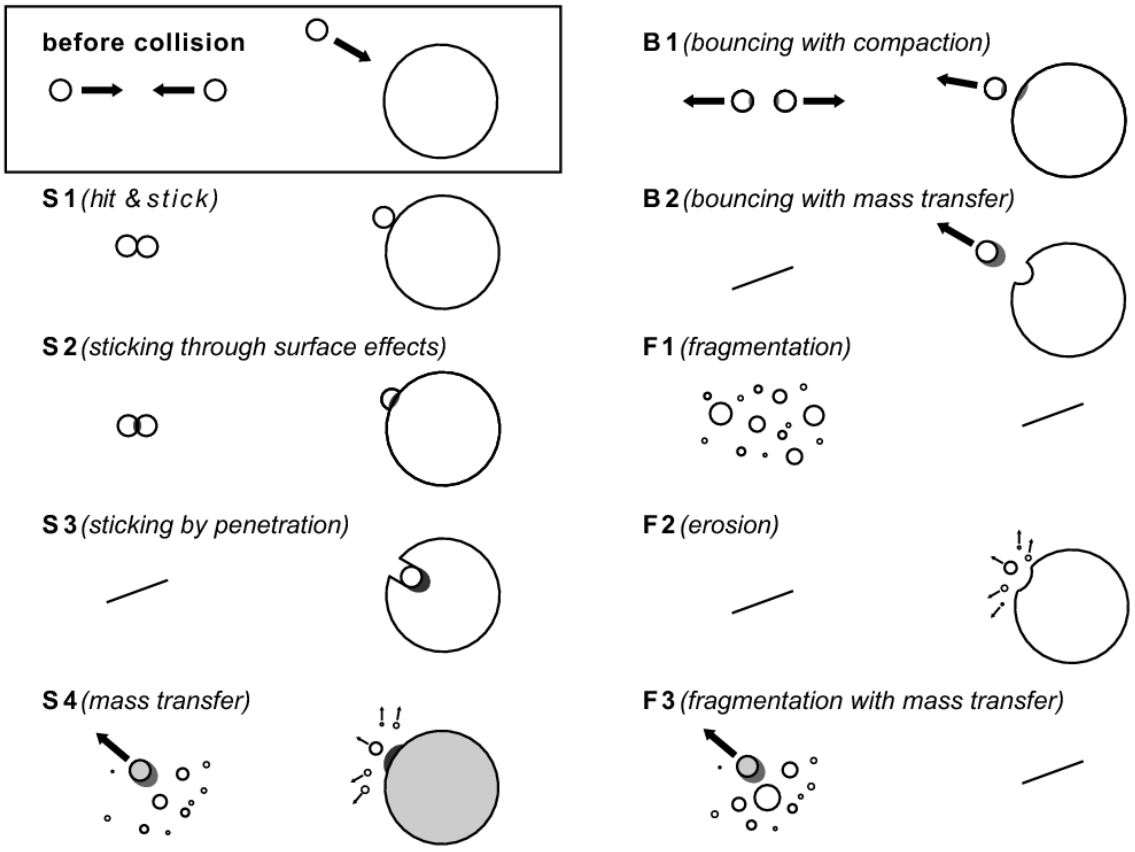


Figure 2.10. – Classification of collisional outcomes to laboratory experiments. Depicted are three kinds of results: sticking (S), bouncing (B) and fragmentation (F). (Credit: [Gütt10])

0.1–1 m are mostly affected by this mechanism and reach drift velocities up to 100 m/s, resulting in drift time scales much smaller than the disk lifetime. Hence, if the growth does not advance rapidly enough beyond those most vulnerable size scales, a critical mass loss, due to inward drift and evaporation in zones of higher temperature, would be the consequence. These indications are assured by full numeric solutions to the coagulation problem, as performed by [Bra08]. The analysis confirms the meter-barrier and shows that, when particle fragmentation and radial drift is ignored, objects can grow due to coagulation enhanced by Brownian motion, differential settling and turbulent motions up to km-sizes, whereas, if the loss mechanisms are taken into account, only particle sizes of not more than 1 cm in scale are generated (see Fig. 2.11).

Hence, destructive high-speed collisions and radial drift somehow need to be bypassed in order to grow planetary embryos. One possible way to accomplish this would be the gravitational collapse within a local region of enhanced particle densities. The development of regions, in which the matter concentrations would be high enough to allow such a scenario, is the matter of ongoing studies, and turbulences seem to be the most promising factors in such considerations. Numerical simulations ([You05], [Joh09]) show that hydrodynamical *streaming instabilities*, arising in the coupled movement of particles and gas, are capable of initiating spontaneous

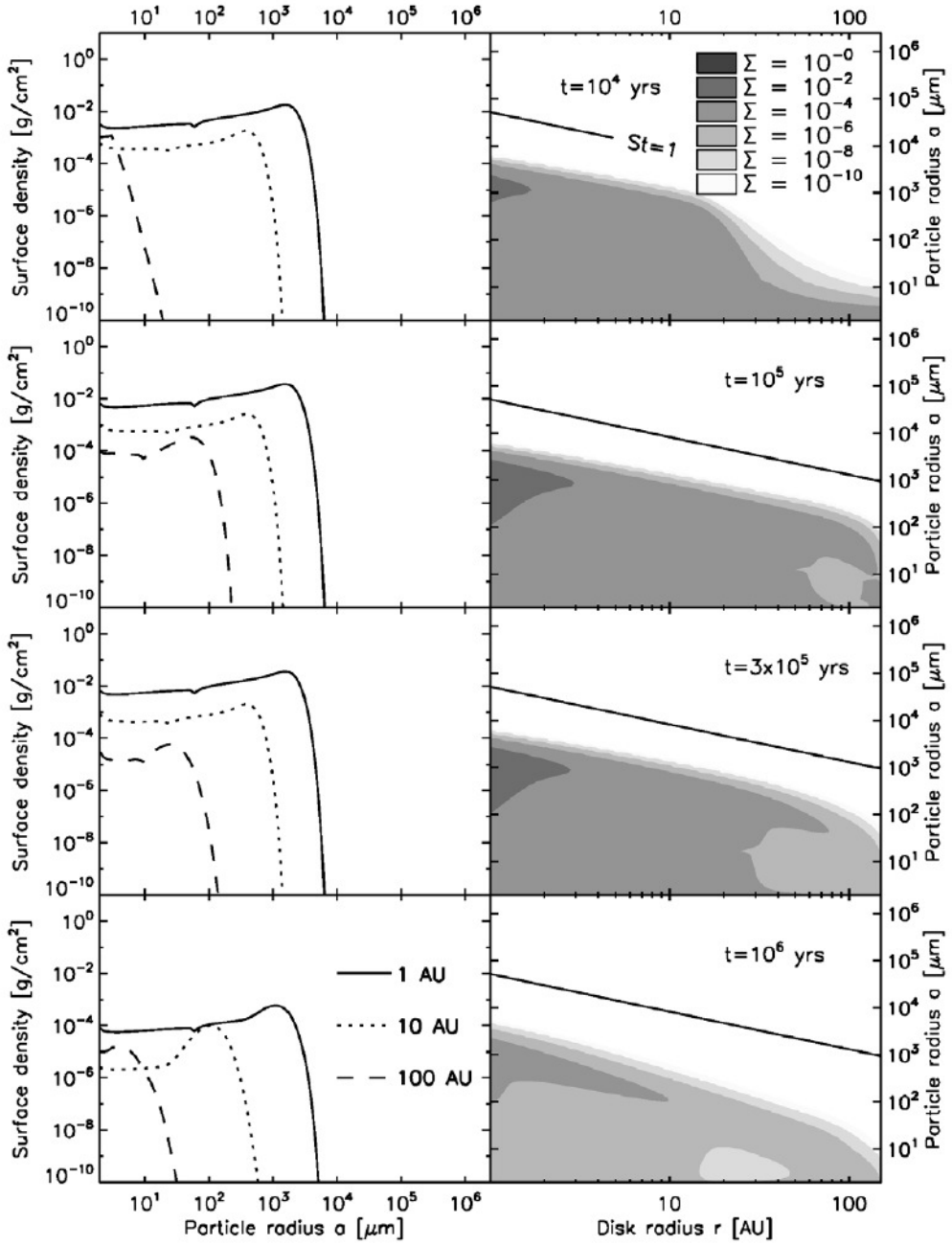


Figure 2.11. – Particle size distribution in the disk as result to numerical analysis of coagulation. Brownian motion, differential settling and turbulent coagulation lead to object growth, whereas collisional fragmentation combined with inward radial drift result in material loss and limit the resulting particle sizes intensely. Different times during the evolution are ordered vertically. *Left panel:* Surface density as a function of particle radius at three different radial distances (1 AU, 10 AU, 100 AU). *Right panel:* Corresponding contour plots of surface density as function of radial distance and particle radius. (Credit: [Bra08])

clumping. Within those zones of increased solid to gas ratio the headwind from the gas component, locally accelerated by the particles towards Keplerian speeds, is reduced, which slows down the radial drift so that further particles approaching from higher orbits can pile up and in turn increase the local density. The efficiency of this mechanism benefits from higher metallicities so that this could also be a contribution to understanding the strong dependency of the actual indicated exoplanet probability increasing with the host star's metallicity.
In addition to producing particle densities high enough to allow gravitational contractions in these regions, the relative velocities between particles within those clumps, created by the streaming instabilities, are derived to be quite moderate. The typical collision speeds lie in the range of 2–5 m/s and only a small fraction (5 %) of encounters happen at velocities above 10 m/s.
In spite of being analyzed only recently, turbulences as induced by the streaming instabilities, offer a promising concept for solving the problems in coagulation modelling and could provide a mechanism for a successful creation of planetisimals of 10–100 km in radius.

2.2.2.4. Progress beyond Planetisimals

Although the processes are not completely understood, considering the solar system, where planets clearly have formed at a broad range of orbital distances, it is obvious that nature has found a way to leave the difficulties of initial growth behind.
Given 10–100 km-sized solid bodies, the underlying mechanisms of further growth slightly change. At this scale the aerodynamic interaction between the objects and the gas are negligible. Stronger coupling is expected to be restored not until gravitational coupling becomes effective at object sizes $\gtrsim 10^3$ km.
Therefore, for further considerations, the formed planetisimals are regarded as a N-body system with constituents on definite orbits.[11]

Gravitational focusing

Growth of solids beyond the planetisimal stage is supposed to be mainly an outcome of collisional contacts. In contrast to coagulation on small scales, at this point gravitational interactions become increasingly important. The collisional cross section is enhanced by an effect called *gravitational focusing*, and hence collisional and consequently growth rates are increased. From basic treatments of physical encounters, including gravitational forces between two solid bodies with masses m, a limiting value for the impact parameter can be derived by conserving energy and angular momentum between the time of the collision and before ([Arm07]):

$$b_{\max}^2 = R_s^2 + \frac{4GmR_s}{v_{\mathrm{rel}}^2} \quad , \tag{2.40}$$

[11] This starting point approximation does not exclude further orbital evolution. In fact, modelling planet migration (see Sec. 2.2.2.6) is substantial for consistency with properties of detected exoplanets.

where R_s is the sum of the objects' radii and v_{rel} the initial relative velocity. Encounters with impact parameters below $b_{\max}$ will result in a collision, and the resulting cross section becomes:

$$\Gamma = \pi R_s^2 \left(1 + \frac{4Gm}{R_s v_{\mathrm{rel}}^2}\right) \quad . \tag{2.41}$$

The second term inside the bracket gives the deviation to the pure geometrical cross section. This effect influences the growth rate of massive bodies significantly, but obviously relative velocities prior to collisions and consequently the velocity dispersion of objects within the disk are not only essential for the focusing but in fact determine the collisional outcome.
As for collisionally induced coagulation of small particles, an encounter does not inevitably lead to mass accretion, but can also result in fragmentation, studies are performed to derive the parameter spaces, regarding impact speeds, masses, sizes and composition of the collision partners, at which the encounters are not disruptive. In this framework it has become convenient to parameterize the collisions by their specific impact energy, which is defined by the collision speed and the mass ratio between the two colliding partners. In this manner [Ben1999] investigated the outcomes from simulated collisions of bodies with sizes in the range of 1–10^5 m and for different impact energies. An evident distinction between two regimes, regarding the objects' sizes, is one of their main results. The *strength dominated regime* includes events between small bodies and is characterized by the feature that the specific energy threshold for catastrophic collisions is decreasing with increasing size of the target object. This resembles the decrease in material strength for growing bodies in that regime. Objects in sizes between 100 m and 10,000 m show to be the weakest compositions, whereas for larger bodies the threshold again increases with size. For those, located in the *gravity dominated regime*, self-gravity becomes more and more effective and prevents them from disruption. While disruptive impacts in the strength regime lead to fragmentation of the target and usually to mass dispersal, catastrophic disruptions become more difficult in the gravity dominated regime. High energy collisions might result in shattering, but the fragmented pieces tend to remain gravitationally bound.

All in all, it should be kept in mind that growth rates of planetisimals depends strongly on the gravitational enhancement and on the velocity distribution. The latter evolves in time and is mainly modified by gravitational interactions and physical collisions. During early growing stages, when gas drag is still effective, eccentricities and inclinations should be damped due to the coupling. While gravitational scattering conserves the kinetic energy of the bodies, collisions usually lead to dissipation and consequently reduce the relative velocities of larger objects compared to smaller bodies within zones, where mass concentration occurs.
Estimations of growing rates of planets due to planetisimal accretions show that the magnitude of gravitational focusing makes the difference between slow initial growing rates at linear time scales and a runaway growth when the gravitational effect increases with growing masses ([Lis1993]).
Although not surprising, but still worth to be stressed out, is the fact that the rate scales linearly with the local planetisimal surface density. This conveys faster growth in disks of more mass in general or in those with a higher solids to gas ratio. Furthermore, with increasing radial distances, growth becomes slower due to the initial disk's radial density profile and lower

orbital frequencies, which correlate with the action time scale, at which accretion takes place. This conclusion becomes particularly relevant for the formation of giant planets in outer regions within a reasonable time.
Numerical approaches to the question about growth rates involve the analysis of the evolution of the mass spectrum of an initial particle distribution and often base on the solution of the *coagulation equation* (e.g. [Lee2000]):

$$\frac{\mathrm{d}n_k}{\mathrm{d}t} = \frac{1}{2}\sum_{i+j=k} A_{ij} n_i n_j - n_k \sum_{i+j=k}^{\infty} A_{ki} n_i \quad . \tag{2.42}$$

In this discrete formulation masses of particles are assumed as integral multiples of some small mass m_1 and so n_k gives the number density of bodies with masses $m_k = km_1$ at a specific time. By that Eq. (2.42) is an expression for the temporal evolution of the number of objects with a specific mass due to successive merging events. The first term on the right-hand side accounts for the gain of particles with the mass m_k, while the second term defines the loss due to merging to higher masses. The coefficients A_{ij} stand for the rate of mergers between bodies of masses, implied by the indices, and hence cover all physical mechanisms, such as gravitational focusing that might be incorporated. If disruptive collisions are to be accounted for, this can be done by adding another terms to the coagulation equation.
Solutions to Eq. (2.42), as well as similar approaches, imply that protoplanets are formed by accretion of planetisimals at first in the regime of an *orderly growth*, with the mass distribution evolving moderately, followed by an epoch of gravitationally enhanced *runaway growth*, when a finite number of larger bodies grows even bigger at the expense of smaller ones, which is characterized by a power-law tail of the evolved mass distribution.
At some point the interaction of the large bodies results in increasing the velocity dispersion of planetisimals in their vicinity due to scattering processes, and therefore, in combination with the depletion of smaller objects through accretion, the growth rates decrease. This post-runaway epoch is widely referred to as *oligarchic growth* (e.g. [Kok1998]).
The mass, reached by the protoplanet by accretion of planetisimals at the time when growth slows down, is called the *isolation mass*. Its value can be estimated by relying on the assumption that only planetisimals within a limited orbital distance to the orbital radius of the protoplanet can be efficiently accreted. This is justifiable insofar, as rapid growth requires low relative velocities of the collision partners, which is most probably provided by planetisimals on not excessively disturbed Keplerian orbits. Since the distance, at which the gravitational influence by the protoplanet is able to perturb the particles' trajectories and to induce collisions, is limited, so is the supply of objects that can be incorporated. Hence, commonly the isolation mass is estimated by the summation of all planetisimal material within the so called *feeding zone* of the planet at orbital radius r_0 ([Lis1993]):

$$M_{\mathrm{iso}} = \int_{r_0-\Delta r}^{r_0+\Delta r} 2\pi r \Sigma_p(r)\,\mathrm{d}r \approx 4\pi r_0 \Delta r \Sigma_p(r_0) \quad . \tag{2.43}$$

With a given surface density of the planetisimal material in the disk $\Sigma_p(r)$ and the radius of the feeding zone Δr, commonly approximated by the planet's *Hill radius* (see App. A), the isolation

mass can be derived and can be used to constrain density profiles and masses of protoplanetary disks. For the typical MMSN the estimation gives masses of $M_{\mathrm{iso}} \simeq 0.07\,\mathrm{M}_{\oplus}$, $M_{\mathrm{iso}} \simeq 1.36\,\mathrm{M}_{\oplus}$ at orbital distances of 1 AU and 5 AU respectively.
However, this derivation is quite strict, since it does not take any diffusion of additional mass into the feeding zones into account. Scattering events, perturbations by protoplanets in neighboring accretion zones, gas drag of small objects, but also migration of planets (see Sec. 2.2.2.6) violate the simple concept of the isolation mass and should drive protoplanets eventually towards higher resultant masses.

2.2.2.5. Formation of Giant Planets

The discussion so far outlined the processes leading to formation of protoplanets, but also pointed out some difficulties and strict time constraints. This becomes particularly severe when concerning the formation of giant planets at larger radial distances. Obviously, gaseous giants have managed to incorporate a large amount of gas from the protoplanetary nebula, so in order to account for those compositions these planets need to form before the gas component is expelled from the disk, limiting models to formation time scales below 10^7 yr.
Two fundamentally competing models have been developed to describe the successful formation of giant planets. The formation by direct collapse, which bases on gravitational instabilities within the disk and the core accretion model, describing the process by successive growth and gas accretion.
Both models are subjects to ongoing controverse discussions, whether they can generate giants within the time limits and whether they are consistent with indications from observations and our knowledge about giant planets. This is why, here at this point, only the underlying principles are presented, while more detailed and critical reviews can be found in e.g. [Dur07], [Mos10], etc.

Direct Collapse by Disk Instabilities

Instabilities within a matter distribution can result in a gravitational collapse, as it happens prior to star formation. For a similar scenario in an accretion disk local perturbations from the smooth and regular disk model are required to trigger destabilizing processes. Matter infall onto the disk, irregular cooling, mass accretion and accumulation or perturbations from neighboring systems all are expected to induce instabilities and turbulences on different spatial scales. Whether those can lead to mass fragmentation and contraction into cores of posterior planets is another question.
Commonly, the occurrence of a gravitational instability (GI) is described by a condition for the *Toomre parameter* Q ([Too1964]):

$$Q \equiv \frac{c_{\mathrm{s}}\kappa}{\pi G \Sigma} < 1 \quad , \tag{2.44}$$

with the local surface density Σ, the sound speed c_{s} and the epicyclic frequency κ, at which fluid elements, disturbed from circular motion, oscillate. For nearly Keplarian disks this can be approximated by the rotational angular frequency $\kappa \approx \Omega$.

Pressure and rotation tend to stabilize the system, whereas for low values of Q the local region becomes unstable. Estimations show that, in order to introduce GIs in a thin disk, relatively high surface densities, as compared to the MMSN models, are required ([Arm07]).
Nevertheless, with decreasing sound speed for increasing radii, the disk, when large enough, should provide regions of sufficiently low stability. Due to such considerations GIs are most probable at early stages when the disk mass is still high.
Still, there is another control parameter decisive for the outcome of an instability scenario. As supported by numerical simulations, non-axisymmetric instabilities, like spiral waves, can be triggered due to decreased values of Q, which do not necessarily drive fragmentation, but rather transport of angular momentum and dissipation. Within those zones gas is adiabatically compressed by self-gravity and leads to heating. This energy needs to be depleted in order for a gravitational collapse to occur, which results in a criterion for the cooling time. Since gas compression acts on time scales of the order of an orbital period, cooling must be effective on similar times or shorter. A criterion for the cooling time is given by [Gam01] to:

$$t_{\mathrm{cool}} \leq 3\Omega^{-1} \quad . \tag{2.45}$$

If the cooling time is too large, the generated heat prevents contraction and the instability leads to a steady state, in which cooling is balanced by dissipation of the gravitational turbulences. Energy transport in vertical direction and the opacity determine the cooling time in the disk. In general, disk models imply that gravitationally induced fragmentation and self gravity can generate giant planets, but preferably at outer orbital distances and early epochs, where and when enough mass and sufficiently high densities are provided. With a tempting rapidness of this mechanism it should result in a considerable reservoir of exoplanetary giants with high semi-major axes. Until today the most successful methods for exoplanet detection (see Sec. 2.3) are mostly sensitive to short orbits, so future progress in observational techniques towards the parameter space of high radial distances will definitely be more meaningful for answering the question, whether the disk instability model is realistic ([Arm07]).

Core Accretion

Modelling the formation of gas giants by core accretion in principle is conservative in the sense that it is founded on growing the planet from accumulating planetisimals under interaction with the gaseous component of the disk. The formation process can be divided into several stages and is depicted in the scheme in Fig. 2.12. During the first and longest epoch a solid protoplanet, later called the core of the giant, is formed by planetisimal accretion. Once the growing core is massive enough to hold a gaseous envelope, gas accretion from the disk commences slowly. For some time growing of both, the solid core and the gaseous envelope, takes place at moderate rates, while the envelope is expected to sustain quasi-hydrostatic equilibrium. The growth of the core will be supported until the isolation mass is reached and the supply of planetisimals ceases, although this might be delayed since the growing envelope should simultaneously increase the feeding zone of the protoplanet.
However, considering the release of potential energy from the accretion of incoming planetisimals as the main source of support for the envelope in maintaining hydrostatic equilibrium and taking energy transport, such as radiative diffusion and convection in the envelope into account,

one finds that at a critical core mass the hydrostatic equilibrium cannot be sustained. When this $M_{\mathrm{core}}^{\mathrm{crit}}$ is reached, the outer envelope itself releases energy by contraction, which results in a rapid runaway gas accretion at the boundary between the gas bound to the planet and the gas in the protoplanetary disk. Computations (e.g. [Iko2000]) imply a power law dependency of the critical core mass on the core accretion rate $\dot{M}_{\mathrm{core}}$ and the opacity κ, induced by gas and grains within the envelope and governing its thermal response to energy transport:

$$M_{\mathrm{core}}^{\mathrm{crit}} \sim 7\left(\frac{\dot{M}_{\mathrm{core}}}{1\times10^{-7}\,\mathrm{M}_{\oplus}\,\mathrm{yr}^{-1}}\right)^{q'}\left(\frac{\kappa}{1\,\mathrm{cm}^2\,\mathrm{g}^{-1}}\right)^{s}\mathrm{M}_{\oplus} \quad , \tag{2.46}$$

where both exponents q' and s can be approximated by 0.2–0.3.

This runaway growth process occurs at rates as fast as the released gravitational potential energy can be radiated away until the gas supply ceases either due to gap opening in the feeding zone of the planet or due to overall clearing of the disk.
Whether the stage of runaway gas accretion is reached within the lifetime of the disk, is again mainly determined by the growth rate during core accretion. Higher accretion rates are supported by increased surface densities and therefore also by higher metallicities, which not only results in longer times for the core growth process at larger distances to the star, but also to larger final core masses in general. Hence, increasing the accretion rates to stay within the time window for giant formation may lead to final core masses significantly higher than such, which are implied by derived core masses, as e.g. for Jupiter.
Still, the dependency on the opacity in Eq. (2.46) is capable to speed up the process without the expense of high core masses by reducing the critical core mass itself. Models to grain opacities in protoplanetary envelopes (e.g. [Mos10]) derive sufficiently low values, which enhance heat loss from the envelope and in turn lead to an earlier contraction. The calculations show that time scales of the order 1 Myr together with final core masses around $10\,\mathrm{M}_{\oplus}$, at radii compared to Jupiter's position, are reasonable, although the surface density in the disk needs to be increased, compared to the MMSN approximations.
It should be noted that even if core accretion ceases before the critical mass is reached, the envelope will contract due to the lack of energy support and the break of hydrostatic equilibrium. The contraction triggers gas accretion, during which the envelope increases mass at moderate rates.

With the main mechanism outlined, there are still certain uncertainties that need to be tackled, before the core accretion model can be regarded as well established. As simulations show, the magnitude of the opacity can be of great importance for triggering runaway growth, when its values are highly reduced, compared to the interstellar opacity, but at the same time the magnitude is highly uncertain.
Furthermore, it is obvious from the findings of giant companions at short orbits (*hot Jupiters*) in exoplanetary systems that radial migration processes occur. In which extent those might influence the outcome and timing constraints for the core accretion scenario, remains to be studied.
Regarding the distribution of giants in the solar system, the change in composition towards higher radii, with Jupiter as the innermost with the largest gas envelope, could mirror the

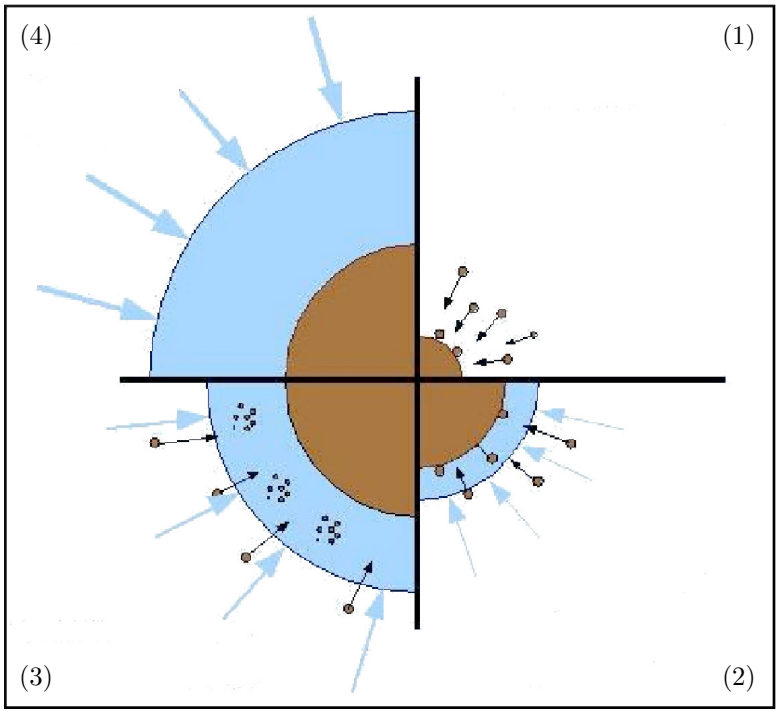


Figure 2.12. – The sequence of giant planet formation, following the core accretion model, is depicted counter-clockwise starting at upper-right. A solid core is formed by planetisimal accretion (1). When it becomes sufficiently massive, slow accretion of nebular gas onto the growing core is triggered (2). The core reaches its final mass, either the isolation mass $M_{\rm iso}$ due to gap opening, or the critical core mass $M_{\rm core}^{\rm crit}$, while gas is still accreted (3). After core growth ceases, the gas envelope contracts and leads to further rapid runaway accretion if $M_{\rm core} \geq M_{\rm core}^{\rm crit}$, or else at moderate rates (4). (Original image from [Ben11])

increasing time scales for core formation and runaway gas accretion with increasing distances to the sun.
On the other hand, time scales for the building of Uranus and especially Neptune seem to be too long to be realistic for formations in situ at those long orbits, again demanding for radial motion during the planet's evolution.

So, at this point it is not possible to exclude one of these approaches towards giant planet formation, as discussions between supporters of both are ongoing.
In principle, a hybrid scenario, in which core accretion is triggered and accelerated by gravitationally induced instabilities, is also imaginable.

2.2.2.6. Orbital Evolution in Young Planetary Systems

In the previous sections the main focus was laid on the formation of solid bodies growing to planets. However, during this process interactions between the constituents of the young and still evolving system might influence the orbital parameters of the planets.

In the following, mechanisms affecting the dynamical evolution of the planetary system, with an emphasis on radial migration of massive solid bodies, are summarized.

Radial Migration

When solid bodies grow, they eventually decouple from the gas and become less vulnerable to gas drag. Nevertheless, theoretical studies of the interaction between planets and protoplanets with the gaseous component of the disk result in the conclusion that orbital migration takes place in young planetary systems. This is not only consistent by implications from detections of hot Jupiters, massive giant exoplanets at short radial distances, but would also yield a possible explanation for giants at outer orbits like Neptune in the solar system.

The basic idea behind migration mechanisms is founded on the fact that an orbiting planet breaks symmetry and induces a non-axisymmetric and time dependent gravitational potential into the system. The gas reacts to that by configuring in density waves, which in turn are perturbations to the potential. This results in a feedback on the planet and in radial migration, as its angular momentum can be changed due to torques exerted by the perturbations.

This response of the disk and the transport of angular momentum due to the excitation of density waves can be analyzed by approaching the resulting torques by summing over partial torques at certain resonant locations within the disk ([Gol1979]).
Two kinds of resonances are to be distinguished. *Corotational resonances* exist in the disk in the vicinity of radii with angular frequencies $\Omega = \Omega_p$, equal to that of the planet at radius r_p, whereas the *Lindblad resonances* occur when

$$m(\Omega = \Omega_p) = \pm\kappa \quad , \tag{2.47}$$

whith the integer m and the epicyclic frequency $\kappa = \Omega$ in a Keplerian potential. When the gas is approximately considered in keplerian motion, the radial locations of the Lindblad resonances are given by:

$$r_{LR} \approx \left(1 \pm \frac{1}{m}\right)^{2/3} r_p \quad . \tag{2.48}$$

The positive sign resembles outer Lindblad resonances, where gas rotates slower than the planet, and the negative sign the inner, faster rotating resonances.
The planet induces positive torques on the gas at radii corresponding to outer resonances and respectively negative torques on gas at inner resonances. This leads to the planet gaining angular momentum from the inner gas and in turn transporting angular momentum to the outer disk. Consequently, gas tends to be repelled in both directions from the vicinity of the planet's orbit.
In order to derive the total effect on the planet, the torques exerted at all resonances need to be summed up. In fact, [War1997] states that on contrary to implications from Eq. (2.48) the positions of the outer resonances lie slightly closer to the planet's radius than the inner resonances of the same order m, and due to further asymmetries the outer torques result to be of higher magnitudes, which leads to a differential net torque that drives the planet towards

orbital decay.
For the quantification of these considerations, [Tan02] studied the collective effect from disk material librating on horseshoe orbits (see App. A) in the planet's corotational region, as well as from gas at the Lindblad resonances. Their three-dimensional, linearized calculations of the interaction within an isothermal disk yield the total torque exerted on the planet:

$$\Gamma = -(1.364 + 0.541\alpha) \left(\frac{M_p}{M_*} \frac{r_p \Omega_p}{c_\mathrm{s}} \right)^2 \Sigma_p r_p^4 \Omega_p^4 \quad , \tag{2.49}$$

and the corresponding timescale for inward migration:

$$\tau_\mathrm{mig} = (2.7 + 1.1\alpha)^{-1} \frac{M_*}{M_p} \frac{M_*}{\Sigma_p r_p^2} \left(\frac{c_\mathrm{s}}{r_p \Omega_p} \right)^2 \Omega_p^{-1} \quad , \tag{2.50}$$

where α is given by the assumed power law disk's surface density $\Sigma_p = \Sigma_0 (r/r_0)^{-\alpha}$ and M_* is the stellar mass.
The inward motion, as described by these considerations, is referred to as the *Type I* migration. It is particularly important to notice that the efficiency of this mechanism is increasing with the planet's mass.
Furthermore, the time scale results to be critically short. Considering reasonable masses of a growing planet of a few $M_\oplus$, placed in a typical low-mass protoplanetary disk at distances of a few AUs, yields migration times below 1 Myr and due to the mass dependency the time scale lies even below 10^5 yr for planets approaching Jupiter sizes.
This high migration rate implies that growing planets are strongly affected by radial inward migration and particularly the formation of giant planets within the given time becomes difficult, since at such a time scale those would probably fall into the star before reaching the critical core mass.

On the other hand, it must be stressed that the above derivation of the total torques and the time scales relies on linear approximations to the perturbations and on smoothly behaving surface density profiles within presumably isothermal disks. In reality, thermodynamics of the disk appear to be of great importance. Models of Type I migration in evolving, radiative disks with adiabatic regions ([Par06], [Lyr10], [Mor11]) indicate that migration is highly affected by disk properties, such as the opacity, the temperature or the gradient of the surface density so that sub-regimes of migration in opposite directions might develop within the disk.
As a consequence, planetary bodies tend to be pushed into convergence zones of zero-torque, rather than experience the catastrophic infall onto the central star.

As stated above, the torques, which are induced by the planet on the gas within its orbital radius, act in opposite direction to those that affect the gas at larger radii. This results in a repulsion of the gaseous component away from the planetary orbit. If this effect is strong enough, the disk's surface density is disturbed and is not able to relax quickly enough so that the planet is able to clear its vicinity and the torques diminish.
At this point, the planet will be locked to the disk and follows the motion of the gas. Radial migration occurs with the disk's viscous evolution at timescales, as given by Eq. (2.34).

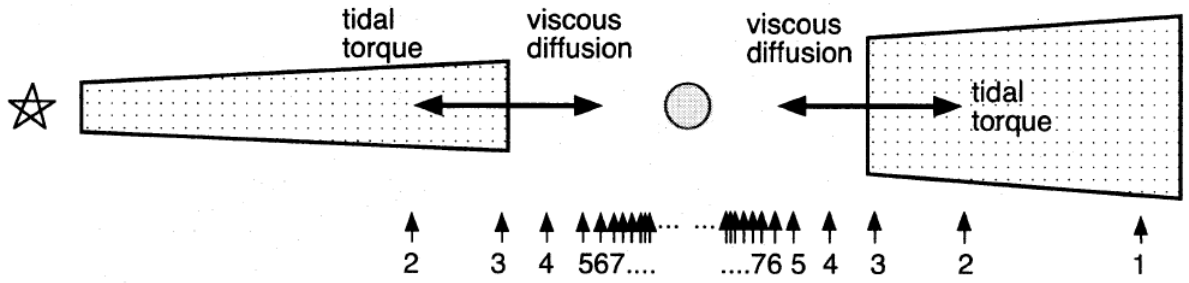


Figure 2.13. – The picture illustrates a model to the mechanism of gap formation leading to radial migration of Type II. A planet is able to open a gap in its vicinity, if the exerted torques, which result in gas repulsion, are more effective than the counteracting viscous diffusion in the disk. Positions of the Lindblad resonances are indicated by the arrows, labeled by the resonance order m. (Credit: [Tak1996])

Consequently, the migration rate in this case of the so called *Type II* migration is independent of the planet mass and is only affected by the disk characteristics and the mass of the central star. In order for the protoplanet to be able to open a gap within the gas distribution, generally two conditions need to be fulfilled. Since the planet's Hill sphere defines the region and scale length, at which the body's gravitational effect considerably influences its environment, the Hill radius must be comparable to the scale height H of the disk. This implies a condition on the mass ratio q between the planet and the star:

$$q \equiv \frac{M_p}{M_*} \gtrsim 3 \left(\frac{H_p}{r_p} \right)_p^3 \quad , \tag{2.51}$$

where the index p means that the scale height ratio is evaluated at the planet's orbital radius. For typically thin protoplanetary disks ($H/r \approx 0.05$) and a solar type star Eq. (2.51) is satisfied for masses above $\sim \mathrm{M_{Jup}}$ or equivalently for $q \gtrsim 1.25 \times 10^{-4}$.
Still, while the torques exerted by the planet tend to gap opening, viscous diffusion of the disk matter acts on filling the gap. This situation is depicted in Fig. 2.13 and results in a second condition for gap formation, which can be derived by postulating that the time scale for viscous gap closing must be longer than the time scale, at which the gap opening mechanism is effective ([Tak1996]):

$$q \gtrsim \left(\frac{H_p}{r_p} \right)_p^2 \alpha^{1/2} \quad , \tag{2.52}$$

with the viscosity characteristics covered by the disk parameter α (comp. Eq. (2.33)). In a typical disk, as defined before, and with $\alpha = 10^{-2}$ this viscous condition is fulfilled for values of $q \gtrsim 2.5 \times 10^{-4}$.
Regarding both conditions altogether, it seems convincing that small protoplanets are affected by orbital evolution of Type I, and while growing beyond a critical threshold mass, which

depends on the disk parameters, Type II migration eventually takes over, involving a characteristic decrease in drift velocity of one or two orders of magnitude ([War1997]).
The planet's lock-in to the gas within the disk will indeed lead to inward directed migration at small radii, but, as [Ver04] have shown, at larger orbital distances, and within disks that experience a significant mass loss during the system's evolution, even outward migration might occur.
Therefore, the combination of both mechanism under the right conditions can result in differential and oppositely directed orbital motions within protoplanetary disks, and so this theoretical treatment not only provides means to comprehend the presence of giants at outer radii, but simultaneously provides an explanation to the numerous detections of hot Jupiters in exoplanetary systems.
Those most probably represent planets, whose formation has started at orbital distances that provided suitable conditions for giant formation and which experienced inward migration thereafter, until the orbital decay ceased due to disk dispersion or truncation or due to tidal interactions with the central star.

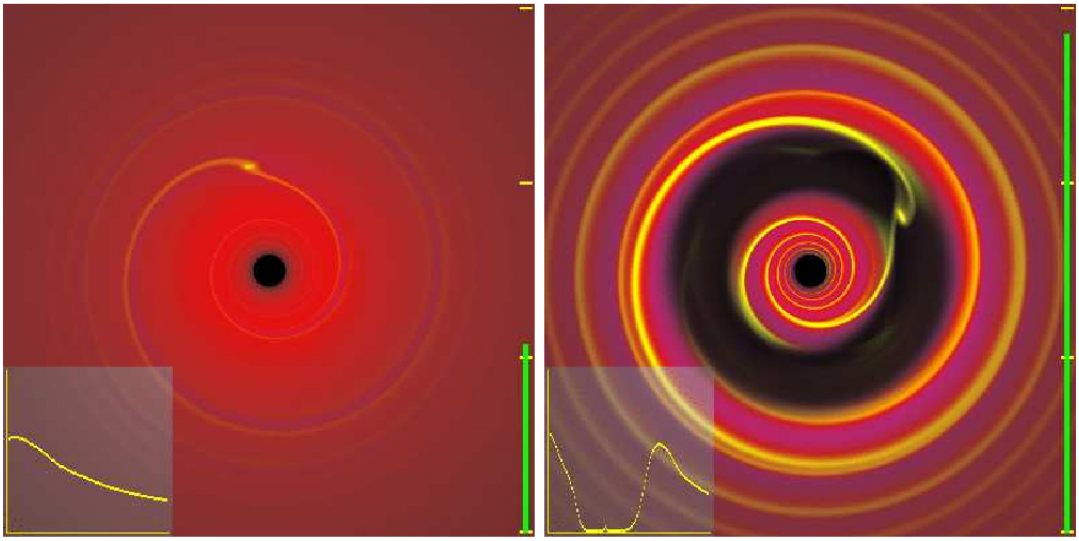


Figure 2.14. – Illustrative comparison of the planet-disk interaction in both migration regimes. *Left:* A low-mass planet undergoes Type I migration. A density wave excited within the disk is noticeable, but the surface density profile (azimuthally averaged and depicted by the graphs in the left corners) is not significantly disturbed. *Right:* In the regime of Type II migration an annular gap is cleared by the massive planet, which then is locked to the gas and follows its orbital evolution on a viscous timescale. (Credit: [Arm05])

Gravitational Scattering

In the previous discussions, orbits of the developing planets were assumed to be nearly circular and in plane with the initial protoplanetary disk. Although these simplifications seem necessary and justifiable for the early epoch during planet formation, dynamical evolution must be taken into account at later stages. The diversity in excentricities and inclinations among detected exoplanets (see 2.3.7) are non-deniable indicators for gravitational interaction during the evolution of young planetary systems.
Analytical studies towards system stability can be a first approach to this problem. In the framework of the three body problem, it can be shown (e.g. [Gla1993]) that a system of two bodies with masses $m_{1,2}$ at initial circular orbits around a central stellar mass M with $m_1 + m_2 \ll M$ remains stable if the initial relative separation $\Delta = (a_2 - a_1)/a_1$ between the planets' radial distances does not fall below a critical value of:

$$\Delta_c \simeq 2.40 \left(\frac{m_1 + m_2}{M} \right)^{1/3} \quad , \tag{2.53}$$

also known as the Hill stability criterion.
When considering complex systems of more than two co-orbiting planets, one might extend this stability condition to any pair of objects, but in reality no absolute statements about the system stability can be made for $N > 3$, and so N-body simulations must be performed to initial configurations.
In principle, gravitational scattering of two planets can result in different outcomes. Particularly interesting are those scenarios, which might lead to evolved distributions of orbital parameters, such as the semi-major axis, the eccentricity, or the inclination.
While collisions tend to result in rather small eccentricities, it has been found that in the case, when close encounters lead to ejections of one body out of the system or to weakly bound states characterized by high semi-major axes, the remainings' radial distances are typically decreased and their eccentricities become significant ([For03]).
Planets, deflected towards their central stars with either a further catastrophic impact or entering an orbit of short period, on which tidal effects might lead to synchronization and circularization, as well as systems with objects at resonant orbital configurations are also imaginable and suggested from analyses.
Recent numerical studies have been performed with the aim to simulate orbital parameter distributions that would be comparable to the actual ones, as extrapolated from the samples of known exoplanetary systems. Regarding the eccentricity, results from integrations of over 3,000 ensembles of randomly constructed initial planetary systems [Jur08] imply a noticeably common final equilibrium distribution that the systems relax to, and which, except for an undersampling at $e \lesssim 0.2$, fits relatively well to the data derived from observations.
Furthermore, they find a general broadening in the distribution of inclinations along the resulting planets, but no significant correlations between the parameters concerned (masses, semi-major axes, inclinations and eccentricities).

All in all, it seems persuasive that, besides migration mechanisms induced by interactions with the gas disk, gravitational scattering between growing planets plays a major role in the planets' orbital evolution and in the dynamical development of the system as a whole.

It might also be the leading candidate for explanations of eccentricity and inclination distributions of the exoplanets, detected until today.

2.3. Extrasolar Planet Hunting Techniques

Although speculations about planets beyond the solar system prevailed amongst philosophers and natural scientists, at least since the times of ancient Greece, significant progress in this field was not achieved before the 1990's. By the end of the 17$^{\text{th}}$ century, *Christiaan Huygens (1629–1695)* estimated stellar distances and is the first one documented to have performed investigations towards exoplanet detection ([Huy1698]), but was destined to fail due to the lack of appropriate means. The first claim for the existence of an unseen companion to a star was made by *W. S. Jacob* in 1855 ([Jac1855]). He has investigated the orbit of the binary star *70 Ophiuchi* and came to the conclusion that a dark planetary companion might be bound to the system. In spite of having been revisited as erroneous (lately by [Hei1988]), this approach resembles the first attempt to the indirect detection of extrasolar planets.
The inevitability of indirect techniques, when searching for exoplanets, is mainly provoked by two aspects related to the subject:

Brightness ratio

According to the Stefan-Boltzmann law $L \propto R^2 \cdot T_{\text{eff}}^4$, the luminosity L of an object, treated like a black-body, is highly regulated by its effective temperature T_{eff} and radius R. This gives approximately $L_\odot / L_\oplus \sim 1.6 \times 10^9$ for the contrast in luminosity when considering the sun and the earth and $L_\odot / L_{\text{Jup}} \sim 6 \times 10^8$ for sun and Jupiter respectively. Consequently, at least in visible wavelengths, direct light from the planet can be neglected and only the reflected component needs to be considered. Imagining an observer at a given distance D to the planetary system, where the spatial separation a between the planet and the host star is negligible comparing to D, the flux ratio from the two objects, as acquired by the observer, can be estimated. Approximating the planet's surface by a plane disk of radius R_p, the incident energy per time reaching the planet is given by

$$L_{\text{in}} = \frac{L_*}{4\pi a^2} \cdot \left(\pi R_p^2\right) \quad , \tag{2.54}$$

with the stellar luminosity L_*. Taking into account that the planet reflects light only into a half sphere, and not considering its intrinsic thermal emission, the flux ratio at the observer's position, when the whole reflecting area is visible, can be estimated to:

$$\frac{s_p}{s_*} = \frac{s_{\text{rfl}}}{s_*} \approx \frac{L_{\text{rfl}}}{L_*} = \frac{L_{\text{in}} / \left(2\pi D^2\right) p}{L_* / \left(4\pi D^2\right)} = \frac{p}{2} \cdot \left(\frac{R_p}{a}\right)^2 \quad . \tag{2.55}$$

Here the planet's reflective characteristics are defined by the planetary albedo p.
Now, considering the system's inclination angle i and the orbital phase, the contrast in brightness becomes

$$\frac{s_p}{s_*} \approx \frac{p}{4} \cdot \left(\frac{R_p}{a}\right)^2 \cdot \left[1 - \sin i \left(\cos \phi(t)\right)\right] \quad , \tag{2.56}$$

where the phase is defined by the orbital period P to $\phi(t) = (2\pi t)/P$ and covers the full orbit with $\phi = 0$ at the time of the conjunction, when the planet lies between the observer and the host. So, for an edge-on view ($i = 90°$) the flux ratio oscillates between 0 and the maximum value of $\frac{p}{2} \cdot \left(\frac{R_p}{a}\right)^2$, while for face-on observers it remains constant at $\frac{p}{4} \cdot \left(\frac{R_p}{a}\right)^2$. However, the dominating term is given by the ratio between the planet's radius and its orbital distance. In the case of the earth-sun system this ratio between the reflected and primarily emitted light is around $\sim 10^{-10}$.
Those high contrast values become more moderate when moving from the visible to infrared wavelengths. In this regime the star appears dimmer and the planet's thermal emission gains significance. Nevertheless, with a ratio of $\sim 10^{-6}$ for the earth-sun configuration carrying out observations at longer wavelengths improves the situation, but the contrast still remains highly demanding to cope with.

Angular separation

For an observer the apparent separation between two distinct objects on the sky decreases with the distance D to the system. Typically, the spatial separations within a planetary system are very low compared to the astronomical distances, at which those systems are to be found, and the angular distance on sky is given by:

$$\theta\,[\text{as}] \approx \frac{a\,[\text{AU}]}{D\,[\text{pc}]} \quad . \tag{2.57}$$

At a distance of 1 pc, which is of the order of the nearest stars, the earth-sun system appears at an angular distance of 1 as (arcseconds), but scales reciprocally with the distance. The angular resolution of an optical system determines, if objects can be identified as distinct sources in the telescope image. Due to the wave character of light, any point source on sky is rendered as an diffraction pattern, the *Point-Spread-Function (PSF)*, in the telescope's image plane. For ideal circular apertures of diameter D the extent of the main maximum is called *Airy-disk* and its angular diameter is:

$$\theta = 1.22\frac{\lambda}{D} \quad , \tag{2.58}$$

or in convenient units:

$$\theta\,[\text{as}] = 0.25\frac{\lambda\,[\mu\text{m}]}{D\,[\text{m}]} \quad . \tag{2.59}$$

Consequently, θ gives the angle between the incoming light and the direction, at which the first minimum occurs. Here it is assumed that the aperture scales are negligible to the emitting source's distance and the incoming light is treated as plane waves.
When observing two sources, each will generate an intensity pattern in the image and in principle these can only be resolved, if the maximum of the one pattern does not fall into the main maximum of the second's PSF. Consequently, Eq. (2.58) gives the ideal diffraction limited minimum angle of objects on the sky, under which they can be recognized as separate sources and is known as the *Rayleigh criterion*. Larger telescope aperture sizes increase the resolution (decrease in θ), while with longer wavelengths the Airy disk increases, which results in a decrease in resolution.

Unfortunately, in reality the situation is not as simple as implied by the diffraction limit. Turbulences within the atmosphere along the lightpath cause deformations to the wavefront and generate multiple images, the speckles, of which the PSFs are diffraction limited. When the exposure time is chosen longer than the time scale, at which the disturbances in the atmosphere change (around 0.02 s), the randomly moving speckles superpose, which results in smearing out the image of the point source to the extent of the so called *seeing disk*. Therefore, the resolution of optical ground-based telescopes with aperture sizes above $\sim$0.5 m is not diffraction limited, but rather defined by the angular diameter of the seeing disk. Depending on the location and weather conditions, the Full Width at Half Maximum of the seeing disk is a varying property of the telescope site and in best cases limits the resolution to 0.5–1 arcseconds.

However, there are ways to avoid or to compensate the atmospheric limitations. Space-based telescopes are mainly diffraction limited and also allow to observe in wavelength regimes not accessible from the earth due to the opaqueness of the atmosphere. According to the Rayleigh criterion, the angular resolution at $\lambda = 500$ nm for the Hubble Space Telescope (HST), with an aperture of 2.4 m in diameter, is about 0.05 as.
On the ground, adaptive optics (AO) can be used to handle atmospheric disturbances. Such systems identify the wavefront distortions in real-time and correct for them. In practice, one part of the incoming light is typically sent by a beam splitter to the wavefront sensor, which measures the deviations from the ideal plane wave, and for compensation controls a deformable mirror that is incorporated into the system. The frequencies, which the control loop needs to operate at, are determined by the time scale of considerable changes to the wavefront. This *coherence time* increases with wavelength and demands for working frequencies in the range of 0.5 kHz–2 kHz.
In order to measure the wavefront distortions, one needs a point source as reference. This can be the observation target itself, if bright enough, or a reference star close to the target. The angular separation of these two objects must lie within the *isoplanatic angle*, which is considered as the angular distance, for which, when not exceeded, the atmospheric distortions to the two light beams can be approximated as the same. In the near infrared the isoplanatic angle is about $20''$–$30''$, but in the visible decreases to $\sim 2''$, which in some cases is a challenging constraint for observation planning (see also Sec. 3.2.2 for the isoplanatic angle in interferometry and its influence on target selection for ESPRI in Sec. 5.1).
If no suitable reference is found, some systems provide the possibility of projecting a laser guide star onto the sky. This artificial source then acts as the reference for wavefront measurements.

The difficulties, arising from the combination of these two effects and technical limitations naturally, bias the sample of target stars, which are usually investigated when looking for exoplanets. Although strongly dependent on the sensitivity and method of each individual detection technique, the challenges can be tackled by the proper choice of host star candidates.
Obviously, favorizing nearby stars moderates the demands on resolution, as the apparent spa-

tial separation scales reciprocally with distance. Furthermore, the high contrast problem can be tackled by choosing target stars of low luminosity or young stars in general, under the assumption that planets, also young in those systems, exhibit higher temperatures.
Still, given our technical limitations, it is of no surprise that efforts on indirect detection methods have been taken and that those led to the first successful results.
The most obvious indication for the existence of a substellar companion to look for, is the host star's reflex motion relative to the system's common center of gravity.
During the last century, first surveys to investigate irregularities in stars' proper motions due to companions have been undertaken. In the 1960's, *Peter van de Kamp* published results from his astrometric study of the nearby high proper motion *Barnard's Star*.
In [vdKam63] he proclaims indications for a companion of 1.5 times the mass of Jupiter. During the following decades van de Kamp partially revised his conclusions from the data, now claiming for a multiple system with two planetary companions. Still, his analysis was never confirmed, but questioned, mainly because of contrary results and concerns about systematics within the telescope (e.g. [Gat1973], [Her1973]).
Despite that, this survey remains a kind of a pioneering work in this field.
Eventually, today commonly accepted as the first discovery of a planetary mass object, is a publication from 1992 by *A. Wolszczan* and *D. A. Frail*. In [Wol1992] the authors present their study of timing measurements of varying pulses from the pulsar PSR B1257+12, from which they conclude the existence of at least two planetary bodies with masses of $\sim 2.8\,\mathrm{M}_\oplus$ and $\sim 3.4\,\mathrm{M}_\oplus$ in orbit around the neutron star. This detection strongly attracted the attention of the scientific community, not only as being the first detection at all, but also due to the system's exotic character. Within the framework of planet formation, the existence of planets around supernova remnants were not even notably considered until then.
For this reason the first discovery of a substellar companion to a main sequence star in 1995 was indisputably a breakthrough in the field of exoplanetary research. In [May1995] *M. Mayor* and *D. Queloz* derive a minimum mass of the planetary body to 0.47 ± 0.02 Jupiter masses. With its short orbit of ~ 4.23 days and a semi-major axis of around 0.052 AU, the companion 51 Peg b also became the first known hot Jupiter. Mayor and Queloz came to those results by analyzing host star's dynamics by the means of spectroscopic Doppler shift measurements. Searching for periodical variations in the star's radial velocity was about to become the most promising and successful technique in the search for exoplanets. With increasing precision, hundreds of exoplanets have been found during the following years by this method, and even though alternative approaches, as surveys to find hints for planetary occultations (transits) in the stars' lightcurves, became also significantly effective, the radial velocity (RV) method remains the standard practice for following-up confirmations.

The extraordinarily quick progress in the search for exoplanets is depicted in Fig. 2.15 and declares the last two decades an era of immense relevance.
Within the following sections the different methods are described, with emphasis on their sensitivity range concerning the planet's intrinsic orbital parameter space and the derived masses.

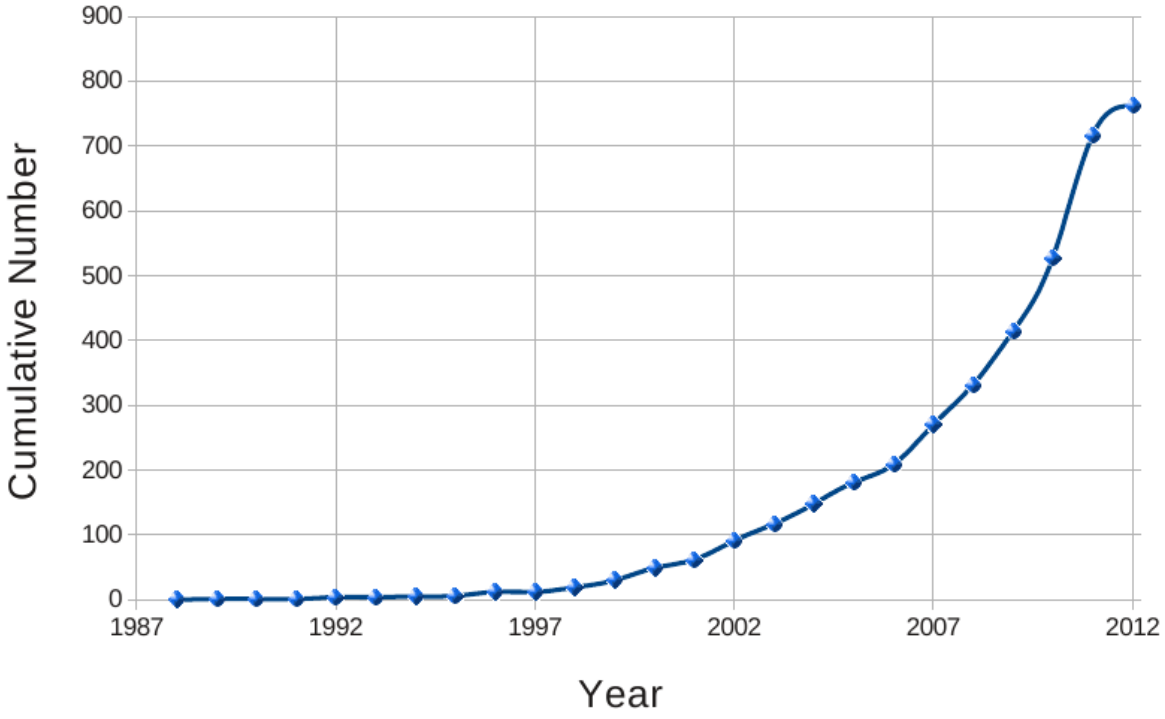


Figure 2.15. – Cumulative number of exoplanetary *candidates* (see [Sch11] as reference) versus time. The data has been exported from the *exoplanet.eu* database in April 2012 and includes 763 planets altogether, discovered by radial velocity, transit detection, microlensing, timing variations and imaging. Note that the candidate HD 114762 b, included in the catalog, is considered as having been discovered as a confirmed companion in 1989, but its nature, whether it is a high mass gas giant or a dwarf star, is still unclear.

2.3.1. Pulsar Timing

Pulsars are fast rotating neutron stars, whose rotational and magnetic field axes are misaligned. Since radio emission from the supernova remnant mostly occurs from its magnetic poles into a narrow light-cone, observers at certain directions do not experience a continuous beam, but rather a pulsed signal. Those signals, with periodicities between milliseconds and seconds for individual objects, are extremely stable in time so that variations, apart from this regularity, mirror irregularities in the system, such as those that would be induced by the presence of a companion in orbit around the neutron star.
If such one exists, both components perform a periodical motion relative to the system's mutual gravitational barycenter, and that of the pulsar results in a periodic variation of the distance to the observer and respectively light travel time. This can be detected by deviations of the by the observer measured signal's periodicity from the pulsar's emitting frequency and becomes evident in a periodic variability of the measured pulse frequency on the time scale of the orbital period. When moving away, the distance and the light travel time increase, leading to higher

spacing between individual pulses and contrarily, when approaching the observer, the time between the measured pulses becomes shorter.
In order to exploit the timing stability of pulsars, one typically folds the pulse data of the observation at the instantaneous pulse period to obtain a pulse profile. This profile, related to the epoch of data recording, is then cross-correlated to an averaged standard pulse profile for this source so that a time dependent phase offset can be derived. This procedure, also called *phase connection*, is based on the assumption that the average pulsar profiles remain stable and that any profile, observed at a certain time, is only a scaled and phase-shifted version of the intrinsic undisturbed profile ([Bai10]).
This allows precise tracking of the rotational phase of the object. From the phase offset at the time of observation to the averaged template profile and the pulse frequency precise pulse times-of-arrival (TOAs) are obtained. When these are compared to predicted TOAs, delays in light travel time, if existent, can be determied. Those can be directly converted to light travel distances, which would occur when the pulsar is in motion in presence of a companion.
The maximum amplitude of the pulses' delay in time for circular Keplerian orbits can be easily derived from the center of mass relation, and with the speed of light c and the inclination of the planet's orbit i is given by:

$$\tau = \sin i \left(\frac{a_*}{c}\right) = \sin i \left(\frac{a_p}{c}\right) \left(\frac{M_p}{M_*}\right) \quad , \tag{2.60}$$

with mass and semi-major axis of the planet represented by M_p, a_p and respectively the pulsar's by M_* and a_*. From the measured variations in the pulses' time of arrival one can obtain the planet's orbital period directly and its eccentricity from the deviation of the signal from a pure sine/cosine function. Furthermore, only the system's projected mass ratio can be derived, because only displacements of the star projected onto the radial direction to the observer contribute to the time delay. Consequently, even if the pulsar's mass can be estimated independently, the planet's mass is still masked by its orbital inclination and only the minimum value ($M_p \cdot \sin i$) can be determined.
The method's efficiency increases with the planet's mass on the one hand, but its sensitivity is also highly dependent on the amount of recorded data over the time span in question, which strongly affects the accuracy of the averaged pulse profile template and by that determines the level of uncertainty regarding the predicted TOAs and consequently the delay measurements.
Fig. 2.16 shows TOA data from the 6.2 ms pulsar PSR B1257+12 taken by the 305 m Arecibo radio telescope at a frequency of 430 MHz over a time span of around 12 years. The data points resemble residuals to the TOAs after the fit of a standard timing model without planets has been performed ([Kon03]). The deviations are dominated by two signals in an orbital resonance of 3:2 with periodicities of 66.5 and 98.2 days. With the pulsar mass estimated to $1.4\,\mathrm{M}_\odot$, the corresponding planet masses have been derived to $4.3 \pm 0.2\,\mathrm{M}_\oplus$ and $3.9 \pm 0.2\,\mathrm{M}_\oplus$ at nearly coplanar orbits with $i \approx 50°$ and semi major axes of 0.36 AU and 0.46 AU respectively.

The relatively low masses derived for this system within the mass regime of terrestrial planets show the efficiency and strength of this method, which unfortunately, due to its unique requirement concerning the host star being a pulsar, is not helpful when studying planetary systems around regular main sequence stars.

As for today, over 15 planet candidates within more than ten planetary systems have been discovered by this technique.

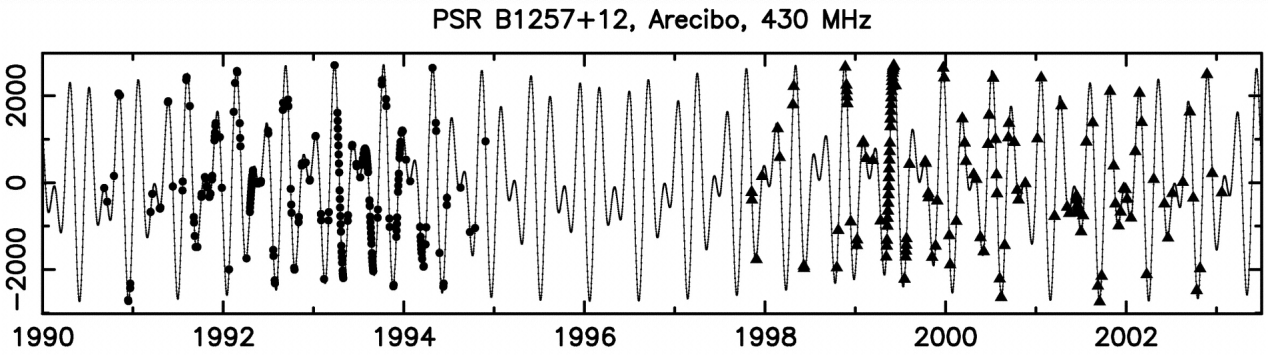


Figure 2.16. – Best fit daily averaged pulse time-of-arrival (TOA) measurements for the pulsar PSR B1257+12, taken by the Arecibo radio telescope at a frequency of 430 MHz. Depicted are the residuals in μm for different epochs after a standard timing model without planets has been fitted to the data. TOA measurements performed with two different instrument back ends are distinguished by different symbols, and the predicted TOA variations by the three planet model, as described in Sec. 2.3.1, is marked by the solid line (Credit: [Kon03]).

2.3.2. Doppler Spectroscopy

The Doppler spectroscopy, or also radial velocity (RV) method, is so far the most successful technique in the search and confirmation of exoplanet candidates. It also exploits the system's dynamics, the star's wobbling due to the gravitational pull of a planet in orbit. The emitted light of a monitored star, in motion around its system's mutual center of mass, exhibits standard Doppler shifts of $\Delta\lambda/\lambda = V_r(t)/c$ due to its varying radial velocity V_r in direction towards the observer.
This component of the host star's velocity, projected onto the line of sight, can be derived from the Kepler laws, and the observed Doppler curve can be described by ([Udr11]):

$$V_r = V_\gamma + K_* \left[\cos(\nu + \omega) + e\cos\omega\right] \quad . \tag{2.61}$$

Here V_γ is the radial velocity of the system's center of mass, ν the true anomaly, ω the argument of periapsis (see Sec. 2.1.1) and K_* is the semi-amplitude of the star's radial velocity oscillations. Directly from the shape of the velocity curve one can obtain the orbital parameters P, e, ω and T_p, the time of periapsis. Alternatively, the orbital period can be converted to the planet's semi-major axis.
Combining the relation between K_* and a_*, the semi-major axis of the star's orbit relative to the barycenter

$$a_* \sin i = \frac{P}{2\pi}\sqrt{1-e^2}K_* \quad , \tag{2.62}$$

together with Kepler' third law (Eq. (2.6)), a relation, referred to as the mass function $f(m)$, can be derived to:

$$f(m) = \frac{(M_p \sin i)^3}{(M_* + M_p)^2} = \frac{P}{2\pi G} K_*^3 \left(1 - e^2\right)^{3/2} \quad , \tag{2.63}$$

where $M_* a_* = M_p a_p$ has been used.

In those cases, where the mass of the planet can be neglected compared to the star, the situation is simplified, and the projected planetary mass, or the minimum mass, can be determined from the semi-amplitude K_* of the velocity curve:[12]

$$M_p \sin i = \left(\frac{P}{2\pi G}\right)^{1/3} K_* M_*^{2/3} \sqrt{1 - e^2} \quad . \tag{2.64}$$

Consequently, if the mass of the star is estimated from models and its spectral characteristics, only the value of the planet's minimum mass can be derived due to the lack of information on the system's orientation.
For all the detection methods that share this ambiguity concerning the inclination, confidence levels for the real companion mass can be constructed from the probability distribution of the inclination angle for random orbit orientations. The probability for the inclination to exceed some angle θ can be expressed by $p(i > \theta) = \cos\theta$ ([Kür1999]), and eventually $i = 60°$ can be considered the median inclination within an unbiased sample.
However, although from a statistical point of view the minimum mass is a passable estimation for the companions' masses, the uncertainty is preserved for individual systems.

The strength of the RV technique is that in principle it is independent of the distance to the analyzed system, provided that the stellar spectra can be obtained with a sufficient signal-to-noise ratio.
Solving Eq. (2.64) for the semi-amplitude, one can see that this method is more sensitive to high planet masses on short orbits and for lower masses of the host star.
The radial velocities, as induced by different planet masses on circular orbits around a solar type star, are plotted for varying semi-major axes in Fig. 2.17 and illustrate the sensitivity regime.

The precision, or accordingly the resolution in radial velocity measurements, determines the minimum of the detectable projected masses. It depends on the involved instruments and on the technique that is used to derive the radial velocities. Two prominent methods are commonly implemented to obtain the radial velocity measurements:

- The **iodine absorption cell** technique makes use of an iodine cell placed at the entrance of the spectrograph, so that the iodine spectrum is superimposed onto the probed stellar spectrum and by that provides a series of absorption lines as reference for the

[12] Clearly, the measured plain Doppler curves need to be related to a common reference frame, such as the solar system's barycenter, and need to be corrected for superimposed velocity components, as induced by the earth's rotation and orbital motion.

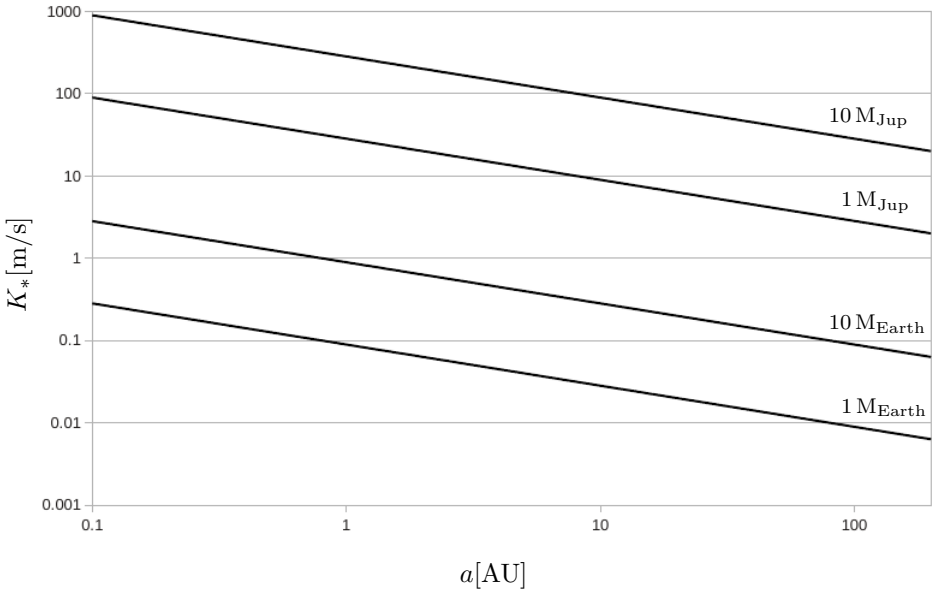


Figure 2.17. – Radial velocity semi-amplitude K_* of the host star against semi-major axis of planets on circular orbits. The four solid lines represent different mass regimes, whereas for the plot the inclination was disregarded ($i = 90°$). In the solar system configuration Jupiter induces a motion of the sun at semi-amplitudes of around $12.5\,\mathrm{m\,s^{-1}}$, and the Earth at $\sim 9\,\mathrm{cm\,s^{-1}}$. This clearly illustrates the high degree of requirements on the precision for detections of earth-like planets.

determination of the Doppler shifts. This design is suitable for slit spectrographs and its incorporation into the system is relatively uncomplicated. The precisions reached by this method lie about 2–3 $\mathrm{m\,s^{-1}}$, but the application of the iodine cell limits the spectral range to 500–630 nm, and the determination of the radial velocities from the combined spectra requires a very high signal-to-noise ratio of the stellar source.

- Simultaneous wavelength calibration within the science exposure is achieved by the **simultaneous thorium** technique, where the emission spectrum from a ThAr-lamp is fed into the spectrograph by one fiber and the stellar spectrum by another. The emission spectrum acts as a separate wavelength reference frame, and the radial velocities are determined by cross-correlating the stellar spectrum with a template mask. The response cross-correlation function (CCF) to that peaks at those shifts in the mask's position, where the correlation to the probe is highest and this position can be translated into the corresponding radial velocity. The exploit of all absorption lines' contribution to the CCF results in high efficiency. In order to reach comparable precisions this method requires about 10–20 times less photons than the iodine cell technique. Resolutions below $1\,\mathrm{m\,s^{-1}}$ are achievable today and the spectral range of usage is about 380–680 nm.

Both described techniques require that the observed target star offers a sufficient number of narrow absorption lines. This constraints the choice of host star candidates to main-sequence

stars in the spectral range of roughly F to M. Hotter stars of earlier type exhibit lines, significantly broadened due to rotation, and on the other side of the range, one is mostly limited by the faintness of the stars. Furthermore, at those precisions ($\sim 1\,\mathrm{m\,s^{-1}}$) intrinsic stellar turbulences limit the resolution in the measurements. Gas motion within the stellar atmosphere, but also magnetic cycles, can obscure or mimic motions due to exoplanets so that stable, non-active stars are highly favorized. K-giants are particularly suitable for this method.
Besides 51 Peg b, the first confirmed exoplanet discovery around a regular type star, the most popular RV surveys e.g. *ELODIE*, *CORALIE*, *HARPS*, *SOPHIE* and others have discovered altogether more than 450 confirmed extrasolar planets.

2.3.3. Photometry and Transits

Despite the fact that high contrast at low angular separations typically prevents us from observing exoplanets directly, although progress has been made in the field of direct imaging (see Sec.2.3.5), precise photometry, at certain conditions, allows indirect detections of sub-stellar companions.
If an exoplanetary system is observed under proper alignment so that for some time during the orbit the planet moves in front of the stellar disk, a characteristic decrease in brightness is measured by the observer (see Fig. 2.18). In order to detect this kind of obscurations, the star and its planet need to be sufficiently aligned along the line of sight of the observer. This limits inclinations of the observed systems to nearly edge-on views.
For specific spatial extents of the bodies and the planet's semi-major axis a relation between the system's inclination and the latitude of the observed transit on the stellar disk can be derived from geometry ([Dee1998]):

$$\cos i = \frac{(R_* + R_p)\sin\delta}{a_p} \quad , \tag{2.65}$$

with R_* and R_p denoting the radii of the star and the planet.
From the limiting case, the planet grazing at the stellar disk at latitudes $\delta = \pm 90°$, the minimum inclination for observable transits is derived to:

$$\cos i_{min} = \frac{(R_* + R_p)}{a_p} \quad . \tag{2.66}$$

When this condition is met and the star is observed during the occultation, whose duration to first order is given by

$$\tau = \frac{P}{\pi}\left(\frac{R_*\cos\delta + R_p}{a_p}\right) \approx 13\left(\frac{M_*}{M_\oplus}\right)^{-1/2}\left(\frac{1}{1\,\mathrm{AU}}\right)^{1/2}\left(\frac{R_*}{R_\oplus}\right)h \quad , \tag{2.67}$$

with P the orbital period, a dip in the lightcurve can be measured.
When assuming a uniform brightness of the star, the depth D of the decrease in the detected flux can be approximated by:

$$D = \frac{\Delta L}{L_*} \approx \left(\frac{R_p}{R_*}\right)^2 \quad . \tag{2.68}$$

Typically, the system is observed over a time interval of multiple transits. This repeatability, on the one hand, is essential for determining the event's period and, on the other hand, for ruling out false-positives, which would occur in cases of background binaries, grazing eclipsing binaries, but also could result from intrinsic stellar variabilities, such as rotationally modulated influences of starspots. Furthermore, data from multiple consecutive transits can be folded at the determined frequency, at which they occur, to increase the sampling of the transit and consequently the precision of the derived parameters.
With sufficient sampling of transit events at high photometric sensitivities one is capable of estimating many characteristics of the exoplanetary system. The orbital period follows trivially from the observations and with knowledge of the stellar mass, the planet's semi-major axis can be estimated by Kepler's third law. In reality, proper modelling of the shape of the transit dip within the lightcurve takes also the star's limb-darkening into account and provides an estimation on the transit latitude δ. In fact, which procedure is actually pursued, also depends on, which information concerning the host star is given from independent sources, such as its spectroscopic classification.
In principle, the stellar radius can also be estimated from the relation ([Udr11]):

$$R_* \approx \frac{a_p \pi W}{\sqrt{\left(1+\sqrt{D}\right)^2 - \sin^2\delta}} \quad , \tag{2.69}$$

where the transit's dimensionless width is parametrized by $W = \tau/P$. Combining the transit depth with Eq. (2.65) then yields the system's inclination and the planetary radius. This is particularly profitable, if transit detections can be combined with radial velocity observations of the systems. In those cases, the mass uncertainty, due to the inclination, vanishes and not only the planet's true mass, but also its density can be determined.
In addition to the access of those primary planet parameters, the analysis of transit lightcurves offers further unique opportunities. If alignment allows transit detections, then, provided sufficient sensitivity, secondary eclipses are also observable. This significantly smaller dip in the measured flux profile results from the planet moving behind the star and yields the star's true brightness. When this is subtracted from the combined flux of the star and the planet, the sole planet contribution can be estimated. Consequently, its effective temperature can be derived under the assumption of a blackbody spectrum.

Apart from this, non-transiting companions in multiple systems can become evident in *transit timing variations* (TTVs), and direct signatures from the planets' atmospheres can be derived from transmission spectra, obtained from combining spectral observations at high signal-to-noise levels during and outside of transits, as was firstly demonstrated by [Bro2000] on the exoplanet HD 209458b, which is also the object, on which the first photometric transit detections have been achieved.

The requirements on the photometric precision can be estimated by the intensity variations, as they would occur to a distant observer of the solar system and its components. For the inner terrestrial planets the transit depth would be in the range between 10^{-5} and 10^{-4}, whereas for Jupiter it lies around 1.1×10^{-2}. For ground-based observations the variability of and the

scintillations in the earth's atmosphere limit the photometric precision to about 1 ‰ so that moving to space is inevitable to detect earth-sized planets.
COROT and *Kepler* are two prominent and successful space survey missions, specifically designed for the search of exoplanets by the transit method. The *Kepler* spacecraft was launched in 2009 and on a heliocentric orbit is presumed to monitor a large amount (about 100,000) stars by continuous pointing at a single star field in the Cygnus-Lyra region. Those high numbers are necessary for transit surveys in order to compensate the low detection probability due to strict constraints regarding the inclination, which is given by $p = R_*/a_p$ for an arbitrary system within a sample of randomly oriented systems ([Dee1998]). Consequently, the method is biased towards large planets on short orbits, although with longer survey duration the detectability of companions on wider orbits also increases.
Apart from the most prominent low mass planets as CoRoT-7 b, the planets around Kepler-20 and Kepler-22 b, a $\sim$2.4 earth-radius planet in the habitable zone of a solar-type star ([Bor12]), more than 200 exoplanets and nearly 30 multiple systems have been detected by the transit technique so far, while more than 1,000 unconfirmed candidates, from the Kepler mission alone, could follow in future.

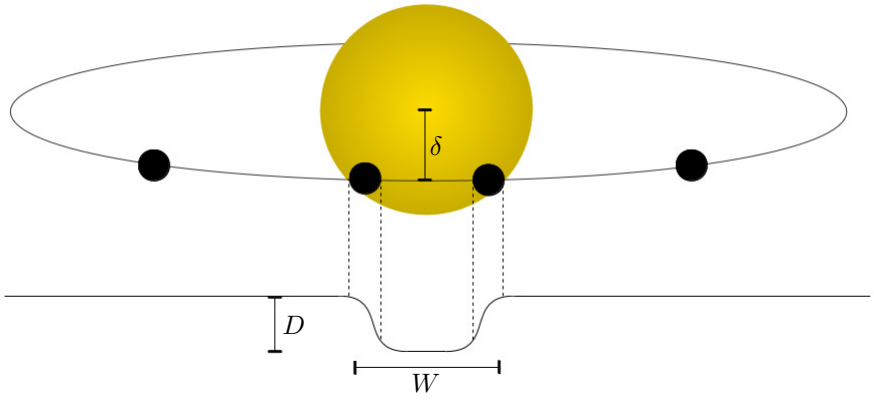


Figure 2.18. – A decrease in brightness is measured when an exoplanetary system is observed at high inclinations so that the planet moves in front of the stellar disk. The shape of the dip in the lightcurve, its depth $D = \Delta L/L_*$ and width $W = \tau/P$, where τ denotes the transit duration and P is the orbital period, are determined by the stellar and planetary masses and sizes, the orbital parameters and by the orientation of the system. The transit's latitude δ is defined within the interval $[-90°, +90°]$, where $\delta = 0°$ characterizes a transit through the center of the host star.

When photometric data with such a precision, as the *COROT* and *Kepler* surveys provide, is analyzed, in addition to the primary and secondary eclipses, more subtle phase modulated variations in brightness can be detected. Three effects are expected to be responsible for this variability and can be modelled in order to exploit the data even further. The gravitation based

ellipsoidal effect leads to tidal distortions of the stellar surface and consequently to a phase dependent flux variation when observing the star. Secondly, due to a varying radial velocity of the host star, the received flux is expected to vary also as a consequence of relativistic beaming, an effect resulting from the Lorentz transformation of the energy, which is radiated away, so that photons are concentrated in forward direction of the source's motion. Eventually, thermal emission from the companion's dayside and brightness modulations, due to reflection of the starlight, are expected to be measured at some level of sensitivity.
Although the magnitude of flux variations due to those effects depends on several system parameters and for realistic assumptions can be estimated to be of the order of tens or hundreds parts-per-million (ppm), [Maz10] have modelled the phase dependencies of the modulations and successfully demonstrated an algorithm capable to measure the modulations of this order and to identify the effects within lightcurves provided by the *COROT* data. The same approach has been taken by [Fai12] on *Kepler* data to discover non-eclipsing binary systems with low-mass companions, which have been confirmed by follow-up RV observations. This indicates the potential of high precision photometry to detect massive exoplanets in even non-transiting systems.

2.3.4. Gravitational Microlensing

Among today's detection methods this unique technique is outstanding in several aspects. Although it depends on random events that require strictly prescribed and generally improbable configurations, it is capable to probe stars at large distances of several kiloparsecs and to detect planets within the terrestrial mass range.
It is based on the behavior of light when passing mass distributions, which is described and quantified by General Relativity. The propagating light beam follows its geodesis in the curved spacetime and bends towards the mass source, which acts like a lens. In the case of a perfect alignment of the background light source and the lens, the light appears symmetrically spread around the foreground object, and its image takes the form of the so called *Einstein-Ring*, with its angular radius in the lens plane given by ([Wam1997]):

$$\theta_E = \sqrt{\frac{4GM_*}{c^2}\frac{(D_S - D_L)}{D_L D_S}} \quad , \tag{2.70}$$

where M_* is the mass of the source and D_S, D_L are the distances to the source and the lens respectively.
In reality, the alignment typically is not perfect, and in those cases several single images of the source are created. When microlensing is used for planet detection, both, the lens and the background object, are stellar sources, and on those scales the Einstein-Ring and consequently the multiple images cannot be resolved. In fact, what can be measured is the total magnification, which is a function of $u(t)$, the angular separation between the source and lens in units of the Einstein radius:

$$\mu(t) = \frac{u^2(t) + 2}{u(t)\sqrt{u^2(t) + 4}} \quad . \tag{2.71}$$

Here the time dependency originates from the objects' proper motions, as well as from any relative motions of the observer, and is essential to this method.
Furthermore, depending on the relative distances within the given configuration, as well as on the lensing mass, such random alignments and microlensing events typically last on time scales between hours and weeks and generate a characteristic lightcurve, which samples any asymmetries of the mass distribution within the lens plane. If the star is accompanied by a dim companion at an angular separation of the scale of the Einstein-Ring (lensing zone), then for adequate alignments the secondary object can act as an additive lens, and this secondary magnification is superimposed over the primary event, resulting in a significant blip in the measured lightcurve (see Fig. 2.19).

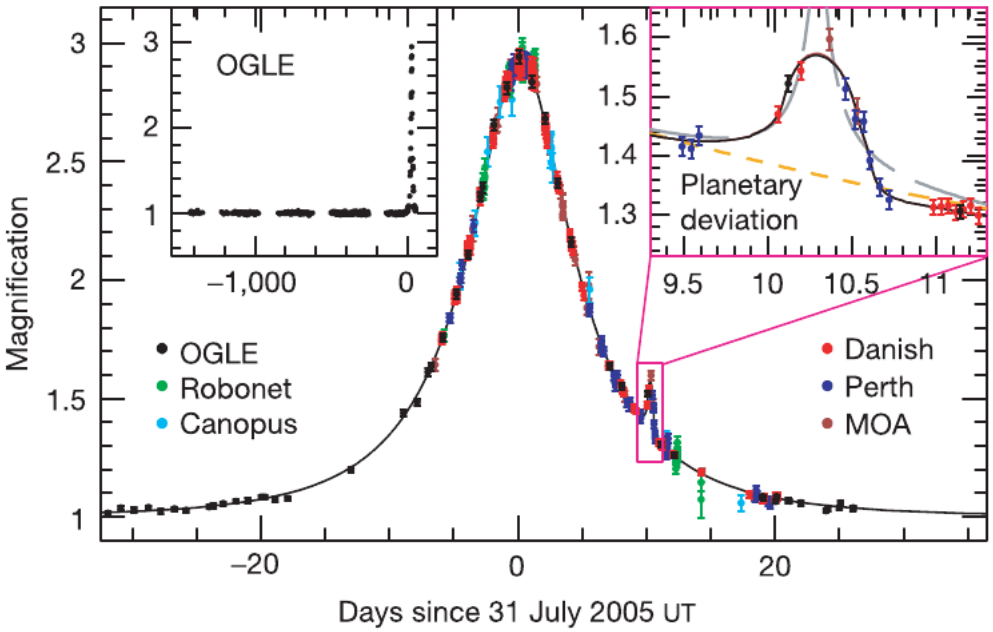


Figure 2.19. – Observed lightcurve of the microlensing event, which had lead to the discovery of the 5.5 earth-mass planet OGLE-2005-BLG-390Lb on an orbit with a projected semi-major axis of 2.1 AU around a M-type star. The data consists of 650 data points from different microlensing programs, as depicted by different colors. The best fit lens model, determining the mass ratio and projected separation, is illustrated by the solid curve. The secondary microlensing event due to the planetary companion is pointed out in the upper right inset, covering a time span of 1.5 days. There, the dashed orange line shows the best single lens model, which the magnification by the planet is superimposed on, and a rejected binary source model is represented by the grey dashed curve. The photometric stability of the source and the significance of the event is pointed out in the top left inset, showing the lightcurve for the source during 4 years, previous to the event. (Credit: [Bea06]).

The projected distance of the companion to its host can in principle be derived from the temporal offset of the secondary peak to the center of the main event, and from the magnification pattern the mass ratio between them can be determined.
Several exoplanet candidates have been discovered by microlensing up to date with estimated masses down to 3–5 $M_\oplus$, on orbital separations comparable to those of terrestrial planets in the solar system.
This capability of probing lower mass regimes at high distances is in some way compensated by the random character of this technique. In order to detect a microlensing event the stellar source and the lens need to be well aligned within the magnitude of the Einstein radius. The occurrence of such configurations is relatively improbable, only about one in 10^6 for observations in directions towards the galactic center, and in addition to that those events usually remain single and unreproducible due to vanishing chances of reoccuring proper alignments.
Also, although this method is able to analyze stars at high distances of several kiloparsecs, in those cases it is more difficult to obtain a sufficiently precise mass estimation on the star from its spectrum than for nearby systems. Currently, the difficulties are successfully approached by several collaborating survey programs, such as *OGLE*, *MOA*, *Robonet* and others, monitoring a high number of stars and by informing each other in cases of event detections so that independent and complement measurements are gathered and good sampling in time is achieved.

Besides the findings concerning regular exoplanetary systems, surprising results have been obtained from a two year microlensing survey towards the Galactic Bulge, regarding a completely different class of objects.
Evidences for the discovery of Jupiter-mass planetary objects, which, after cross-checking with results from direct imaging, appear to be not bound gravitationally to any host star, have been reported by [Sum11]. Those could be some of the first detections of *free-floating* planets, which have been scattered into wide or unbound orbits during the evolution of their protoplanetary disks.

2.3.5. Direct Imaging

Although being successful in detecting extrasolar planetary candidates and determining several orbital parameters of those, the indirect methods altogether suffer from limitations regarding the characterization of the discovered objects, concerning their intrinsic physical properties, such as their chemical composition, temperature, characteristics of atmospheres and others. Analysis of direct radiation received from the planets, either from their thermal emission or the reflected starlight, is required in order to tackle those questions.
Unfortunately, as already discussed above in this section, the high contrast in luminosity and the typically small angular separations between the host star and the exoplanets are hardly solvable problems for today's instrumentations.
On the other hand, with its natural constraints concerning small separations, direct imaging can in turn serve as a suitable and complementary technique towards companions on wide and long orbits, which are hardly detectable by RV measurements or the transit method.
Several imaging techniques are exploited to enhance the images' dynamic range and to cope

with atmospheric seeing. Apart from moving the observatory to space, or the usage of adaptive optics, speckle and spectral, as well as angular differential imaging, have proven to be promising approaches in this field. Also, full or semi-transparent coronography, by which the star's contribution to the image is suppressed, contribute to increase sensitivity in high-contrast observations.
When moving to mid-IR wavelengths, the contrast is decreasing due to thermal brightness of the usually cool exoplanets with temperatures in the range of a few 100 K, but angular resolution decreases also. Still, for observations in this wavelength regime interferometric techniques, such as *masking* or *nulling interferometry*, as an analog to standard coronography, are applicable. A more comprehensive review about the progress in this area can be found in [Opp09].

Considering those limitations, recent results and discoveries of exoplanet candidates by direct imaging are impressive. Despite uncertainties concerning the true nature of some of those objects, with masses in the range from a few Jupiter masses up to the brown dwarf limit and on wide orbits of several hundred AUs, which mirror the natural bias of this method, future developments and projects will definately provide a complementary tool for exoplanet characterization.
Prominent planetary candidates that have been discovered by imaging are the companion 2M1207 b around a young brown dwarf ([Cha04]), the $< 3\,M_{Jup}$ candidate Formalhaut b ([Kal08]), the components of the multiple system around HR 8799 ([Mar08]), which is depicted in Fig. 2.20, but also several isolated planetary mass objects in the σ Orionis star cluster ([Zap2000]).

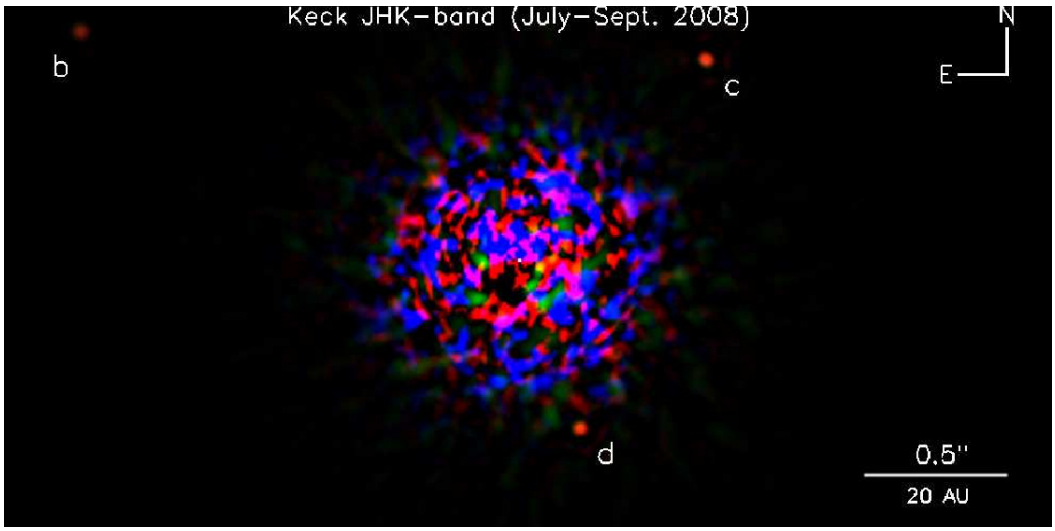


Figure 2.20. – Combined J-, H-, Ks-band multi epoch image composition of three companions (denoted by b, c, d) around the star HR 8799. Angular differential imaging (ADI) has been applied onto the Keck telescope images to remove the light of the host star and to picture the companions at projected separations of about 68, 43 and 27 AU. The orbital motion of component d during the time between the image epochs has been compensated for by a rotation of 1° of the earliest image taken (Credit: [Mar08]).

2.3.6. Astrometry

Among the detection techniques astrometry is the one that is based on the oldest observational method and, leaving imaging aside, is probably the most naive of all the indirect approaches. Classically, astrometry means measuring the positions of celestial objects over time. In the past, celestial object positioning and tracking with increasing precision lead to the discovery and verification of various phenomenons, such as proper motions of stars, aberrational effects, the motion of the earth's rotational axis (precession, nutation), but also to the detection of outer planets within the solar system.
Despite the long history and the conceptional triviality of this method, high precision astrometry is by no means simple.

One complication results from the fact that in principle no static reference points can be found on the sky. Stars move approximately linear due to their proper motions that arise from the relative velocities between them and the solar system. Still, additional periodic variations in their apparent positions due to the earth's orbit and consequently the motion of the observer complicate the situation and need to be considered, when positions on the sky are about to be modelled for any purposes. On the other hand, measuring those effects can yield information about the systems observed, as is done by parallax measurements, from which distances in space can be determined.
However, it should be noted that in this context the angles, at which those motions occur, are typically small. The star with the largest known proper motion of around 10 as/yr is a nearby object at a distance of about ~ 1.8 pc, but for distant stars within the Milky Way proper motions of the order of only several milliarcseconds (mas) per year are recorded. For objects outside our galaxy the values are even smaller.
Those considerations should demonstrate the demand on accuracies in astrometric studies. To achieve such on a large scale, a precise reference system is required. For this purpose several catalogs have been established in the past. Before the *International Celestial Reference Frame* (ICRF) that defines a quasi-inertial reference frame by the positions of over 200 quasars, which due to their enormous distances can be considered as static, was established in 1998, the *catalog of fundamental stars*, version FK5 in those days, provided a standard coordinate system. In 1989 the *European Space Agency* (ESA) launched the instrument *Hipparcos*, which was about to perform space-based high precision astrometry for the following three and a half years. As a result from this survey a catalog of more than 100,000 stars with accuracies around 1 mas has been published ([ESA1997], [vLee07]).
Using subsets of this data the fundamental catalog has been updated (FK6), and the *International Celestial Reference System* (ICRS), with its origin at the solar system barycenter and fixed axes, has been introduced.

Given the implied difficulties of applying an absolute and fixed reference system to astrometric studies, it has become a common procedure to relate the measurements to a local reference frame that can be composed of objects within the field of view (FOV). Usually, higher precisions can be achieved by this approach.
Astrometric measurements have been used to derive orbits and system masses of visual binaries,

but in principle can also be applied to a system, in which only the main component is visible. Exoplanetary systems are typical representatives of such ones, and in those cases the projection of the host star's orbit around the system's center of mass onto the tangential plane at the star's coordinates is observable. With the condition $M_p \ll M_*$, and utilizing $M_* a_* = M_p a_p$, the host star's angular motion on the sky can be estimated by Kepler's third law, and considering the distance D to the observer, which as usual is assumed here to be much larger than the scales within the exoplanetary system, the semi-amplitude of that stellar reflex motion reads as:

$$\theta[\text{rad}] = \frac{M_p}{D} \left(\frac{G}{4\pi^2}\right)^{1/3} \left(\frac{P}{M_*}\right)^{2/3} \quad . \tag{2.72}$$

The orientation of the system was not considered here, while in reality the magnitude of the astrometric signal that really can be measured is determined by the semi-major axis and naturally by the orientation angles of the star's orbit. Regarding the fact that the observed apparent orbit can be measured in two dimensions, the minimum astrometric signal, which is observed at the most unfavorable orientation, when the system is seen edge on and the line of apsides is aligned with the line of sight, will be defined by the semi-minor axis of the true orbit.

In general, the observed projection of the path, which the star follows during its orbit around the system's barycenter, will appear as an apparent elliptical orbit on the sky with the semi-amplitude astrometric signature in radians:

$$\alpha[\text{rad}] \approx \frac{a'_*}{D} = \frac{M_p}{M_*} \frac{a'_p}{D} \quad , \tag{2.73}$$

where a'_*, a'_p are the semi-major axes of the two components' apparent orbits. In more intuitive units this relation can be written as:

$$\alpha[\mu\text{as}] \approx 3 \left(\frac{M_p}{1\, M_\oplus}\right) \left(\frac{D}{1\text{ pc}}\right)^{-1} \left(\frac{a_p}{1\,\text{AU}}\right) \left(\frac{M_*}{1\, M_\oplus}\right)^{-1} \quad . \tag{2.74}$$

Here one can see, which parameters the astrometric detection method is sensitive to. The displacement of the host star increases with the planetary mass and semi-major axis, whereas in turn the stellar mass and the distance to the observed system damp the signal. Furthermore, the relation (2.74) implies the order of magnitude, at which such astrometric signals, as variations in the star's angular position, are expected and indicates the precision requirements of measurements, when detection of those exoplanets is intended.
In Fig. 2.21 astrometric signals are plotted against companion masses for different distances to a virtual exoplanetary system, where the planet is in orbit at $a_p = 3$ AU around a solar-type star. Even in order to detect Jupiter sized planets around fairly low distant stars at between 10 and 100pc, accuracies of around 30–300 μas need to be achieved.
Consequently, this method is mostly efficient for nearby and low-mass stars and is biased towards high planetary masses on wide orbits.
Moreover interesting is the fact that, as long as the photocenter does not shift by the same order as the star itself, astrometry is applicable also to hot or rapidly rotating, as well as young and active stellar sources, which this method particularly complementary to the RV technique

and allows to study a different regime in the systems' parameter space.

In the past, astrometry has been successfully used for determining system masses of visual binaries. When it comes to exoplanetary systems, the situation is a bit different, since only the main component is visible. In order to monitor any variations in the star's position, it is necessary to establish a local frame, which commonly is done by proper choice of reference targets. Given such a reference frame, to which the observations can be related, the star's apparent orbit around the center of mass can be modelled, and in principle each of the seven orbital parameters can be determined, if the orbit is sufficiently sampled.
The measurements consist of two-dimensional positions, assigned to the time of observation. Eventually, the apparent orbit is observed and needs to be converted to the star's true orbit, in order to derive the true semi-major axis. The knowledge of the three orientation parameters i, ω and Ω (see Sec. 2.1.1) is sufficient for transformations between the true coordinates within the orbital plane and the rectangular coordinates within the tangential plane on the sky, and the other way around. With the two coordinate systems' point of origin O, at the focal point of the true orbit, the true orbital coordinates, expressed in seconds of arc, are defined by:

$$\begin{aligned} X &= \varpi r \cos\nu \quad , \\ Y &= \varpi r \sin\nu \quad , \end{aligned} \tag{2.75}$$

where ϖ is the parallax, r is the body's in orbit distance to O in astronomical units and ν is the true anomaly. The coordinates in the tangential plane, also in arcseconds, are given by:

$$\begin{aligned} x &= \rho \sin\theta \quad , \\ y &= \rho \cos\theta \quad , \end{aligned} \tag{2.76}$$

where in this case the body's position is defined by its separation to the system's point of origin ρ and the position angle θ, evaluated eastward from the northern direction of the local celestial meridian. The transformation is then performed by:

$$\begin{aligned} x &= BX + GY \quad , \\ y &= AX + FY \quad , \end{aligned} \tag{2.77}$$

where the orientation parameters are incorporated into the *Thiele-Innes* constants ([Kov04]):

$$\begin{aligned} A &= \cos\omega\cos\Omega - \sin\omega\sin\Omega\cos i \quad , \\ B &= \cos\omega\sin\Omega + \sin\omega\cos\Omega\cos i \quad , \\ F &= -\sin\omega\cos\Omega - \cos\omega\sin\Omega\cos i \quad , \\ G &= -\sin\omega\sin\Omega + \cos\omega\cos\Omega\cos i \quad . \end{aligned} \tag{2.78}$$

It is not crucial, in which of these individual coordinate systems the actual orbit fit is performed, but although the period P is independent of the choice of the system, the other dynamical elements, in this case the eccentricity and especially the semi-major axis, are not. Therefore, the distinction of these two representations should be treated carefully.
After the determining parameters are obtained with respect to the true orbit, the real mass

ratio between the planet and the host star can be estimated so that when the stellar mass is somehow determined separately, this method gives the true planetary mass, without the inclination ambiguity.

Although the principle of this kind of observations seems simple, the demands on accuracy are a great challenge. In order to model the star's orbital motion it is not sufficient to disentangle it from the stellar proper motion and the displacement due to the parallax, but also corrections concerning earth rotation and even relativistic effects, affecting the light on its complete path from the source to the observer, need to be taken into account with increasing demand on precision (see section 4.2.3 on Astrometric Corrections).

Summing up the previous discussion, the strengths of astrometry considering the search of extrasolar planets are:

- The complete set of orbital elements, in particular the **inclination**, under which the system is observed can be derived. With an estimation of the host star's mass this technique yields the **true planetary mass**.
- The studied parameter space can be expanded towards **wide orbits** and consequently **long-period planets**.
- In contrast to Doppler spectroscopy, astrometry is not bound to a certain kind of stars. Although sensitivity increases with decreasing stellar masses, **young pre-main-sequence stars** can be probed, so that hypotheses regarding early stages in the evolution of exoplanetary systems can be tested.
- Multiple systems, in which the signal from different companions can be disentangled, can be examined regarding the planets' **coplanarity**, since the orbital inclination can be determined.

To date, regarding successful discoveries, astrometry cannot compete with the other detection methods, but signals from companions have been derived from HST (Hubble Space Telescope) observations (e.g. [Ben10]) and confirmed earlier RV measurements around several stellar objects. Also, from Hipparcos data mass constraints on substellar companion candidates can be determined and used for confirmation as was done by [Ref11].

In future, the next generation ESA space mission *Gaia* will provide high accuracy astrometric measurements at the level of several tens of microarcseconds, whereas from ground the actually being commissioned instrument PRIMA at the VLTI (see following chapters), will hopefully be able to perform at astrometric accuracies of the same order.
The ability of achieving such accuracies by phase-referenced interferometry has been already demonstrated by reports of the discovery of a ~ 1.5 Jupiter mass companion in the binary system HR 7162 using the *Palomar Testbed Interferometer* (PTI) ([Mut11]).

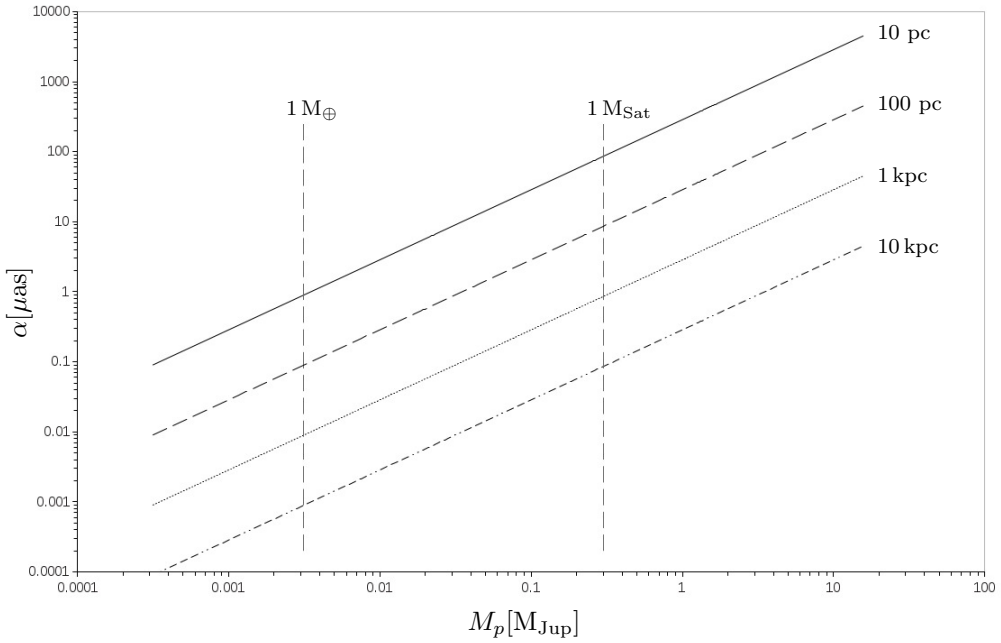


Figure 2.21. – The astrometric semi-amplitude of a solar-like star plotted against the mass of the companion that is placed at an orbital distance of $a_p = 3\,\mathrm{AU}$. The angular magnitude of the measurable stellar reflex motion depends strongly on the distance to the system and is represented by four different lines with distances between 10 and 10,000 pc, as denoted at the right. Vertical lines of one Earth and one Saturn mass are inserted as references. If astrometric accuracies of the order between 10 and 100 μas can be achieved, then detections of Saturn-like planets around nearby stars should become possible.

2.3.7. Review and Prospects

During the last 20 years our understanding regarding exoplanetary systems changed substantially. Although the existence of exoplanets was never really implausible, discussions about their occurrences could become profound only after the era of exoplanet discovery has begun. From then on, properties of the new detected objects could have been studied and in principle implications on planet formation theories could have been finally derived from observations of systems other than the solar system.
Today, each detected exoplanet serves as a probe within the whole sample and significance is obtained statistically. We have learned that formation of exoplanetary systems is not an unusual process, and analyses of orbital and intrinsic planet parameters demonstrate high diversity among those found and studied.
In contrast to the configuration of the solar system, which appears quite regular, with planets on relative coplanar, nearly circular orbits and correlations between the orbital distances and the companions' masses, the whole sample of today's known exoplanets shows a higher diversity, with planets in different mass regimes between several Earth masses and multiple Jupiter masses up to the Brown dwarf limit.

The derived orbital parameters appear also well distributed, with quite extreme cases of short orbits, very close to the stellar source, but also those of high eccentricities.
These findings during the last years not only convey the progress of our understanding in this scientific field, but also arise public interest so that reports on newly classified objects like Hot Jupiters or super-Earths find their way into the media on a regular base.

While it is typically difficult to derive universal conclusion from individual systems, statistical analyses of a larger sample can constrain planet formation theories. Usually, this is performed by the comparison of synthetic exoplanet populations, modelled under certain assumptions on initial conditions regarding the stellar sources and the protoplanetary disks, to the results of real exoplanet search surveys (e.g. [Mor09]). In this context it is important to take the biases into account, which are introduced by the different sensitivity regimes of the various detection techniques.
As is shown in Fig. 2.22, the distribution of the known exoplanets within the plane of semi-major axis and mass is not uniform, which directly mirrors detection probabilities of different methods in different parameter space regimes. Therefore, when survey outcomes are compared to modelled populations, the results need to be corrected for completeness.

Although statements about the significance and validity of conclusions from such studies should be handled carefully, some derived qualitative implications seem convincing (e.g. [May11]). Concerning the studied characteristics, there seems to be a clear distinction of planet populations, regarding their masses, divided by a gap at around $30\,M_\oplus$, which is only hardly explainable by detection biases. The population of gaseous giant planets ($\gtrsim 30\,M_\oplus$) covers a large range of orbital radii, although the occurrence frequency seems to increase weakly with the orbital period.
The eccentricities within this sub-sample feature significant scatter, and high values of up to $e = 0.97$ have been found. A strong and well established correlation, which suggested itself quite early, is the increasing occurrence of those giant planets with the host star metallicity, in literature also known as the *metallicity effect*. This seems to be a clear indicator on giant planet formation dependency on the initial conditions concerning the composition (gas to solid ratio) of the protoplanetary disk.
As significant as this correlation reveals itself for planets in the high mass regime, it is interesting that no clear analog indications can be found within the second population of planets of masses below 30–40 $M_\oplus$. More than this, even an inverse metallicity effect among the objects within the terrestrial mass regime cannot be excluded from the data on the gathered collection. Also, whereas the occurrence rate for giant planets lies around 14 %, about one half of solar-type stars seem to harbor a planet of the lower mass population.
However, those rates are highly influenced by the selection of stellar targets for different surveys and also by the observing limitations towards long period orbits.
Further notable are the increase of the mass distribution towards lower masses, analyses of the semi-major axes indicating higher occurrences for periods between 40 and 80 days, and the distribution of eccentricities among the super-Earth and Neptune population, which also exhibit some scatter, but only below a maximum value of around $e = 0.45$.

Discussions about constraints on physics of planet formation from the population of detected exoplanets is ongoing and fruitful. Still, it is clear that future improvements of the detection techniques towards higher accuracies will provide possibilities to explore exoplanetary systems in parameter domains not accessible today and that those will further expand our understanding of the underlying physical mechanisms.

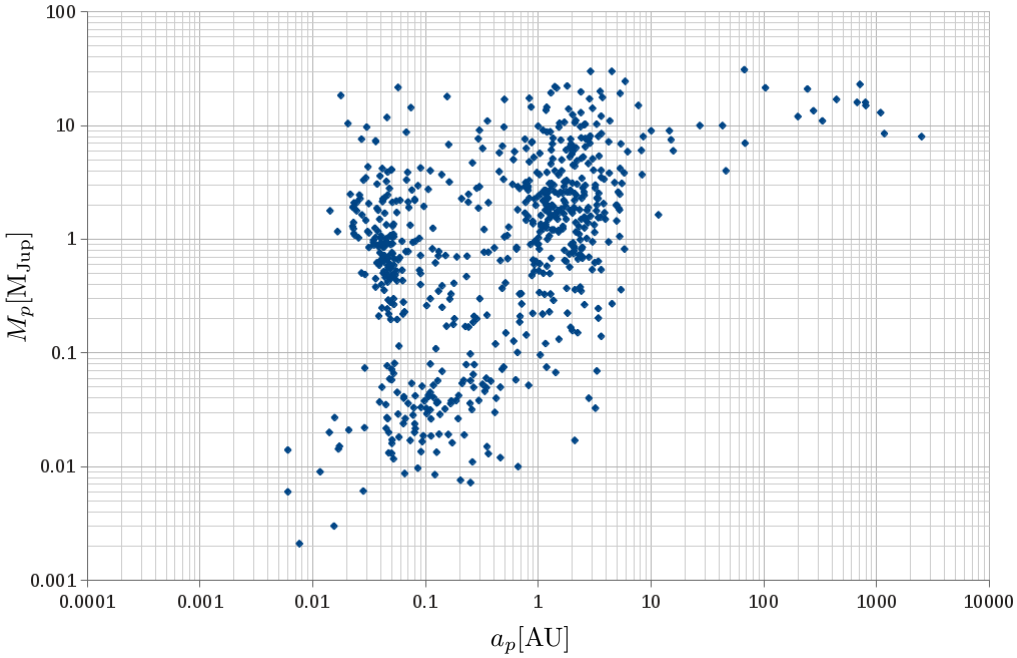


Figure 2.22. – Exoplanetary masses against corresponding semi-major axis. It should be noted that for some companions, such as those detected by the RV technique, only the projected mass $M_p \sin i$ is known and was considered here, whereas the true mass can be significantly larger. The data on these more than 700 exoplanet candidates has been obtained from the *exoplanet.eu* database. It is clearly visible that the distribution is not uniform all over the plane, which not necessarily reflects the true exoplanet occurrences, but is strongly biased by the different detection methods being sensitive within distinct parameter regimes. Conclusions from the specific characteristics of the detected exoplanets on planet formation processes are typically drawn from statistical comparison to synthetically modelled planet populations.

Chapter 3

Interferometry in Theoretical Context

In the last chapter the importance of high precision astrometry, in the context of the search for exoplanetary systems, has been stressed out. Due to technical improvements, such as the replacement of photographic plates by CCDs, the accuracy for astrometric measurements has improved to that point that now the precision achieved is mainly limited by atmospheric refraction and turbulences.
Relocating the instruments into space, just above the source of errors, is one way to overcome the difficulties, as was demonstrated by the success of the spaced-based mission *Hipparcos*.
Another, conceptually different, approach becomes possible by the means of interferometry. Astrometric measurements are usually performed on a target object and one or several reference objects within the field of view. How long baselines can be used to improve the accuracies has been shown by [Sha1992]. Differential positional errors due to the atmosphere are induced since the light from the objects follows different paths in the atmosphere, and consequently the fictitious surfaces of the same phase of the propagating wave, the *wave fronts*, are disturbed separately. In principle, these are governed by the combination of the objects' angular separation θ, their beams' spatial separation in the atmosphere θh and the amount of the overlap of the two beams, as compared to the diameter of the telescope or respectively the length of the interferometric baseline ([Col09]).
Usually, the diameters of single dish telescopes, as used for conventional narrow-angle astrometry, are significantly smaller than the spatial separation of the two objects of interest, projected to the height above the turbulent layers of the atmosphere at about $\sim 10\,\mathrm{km}$ (see Fig. 3.1). In the regime, where the instrument sizes extent the objects' projected separation $d = \theta h$ at height h, the mean square error, in arcseconds, for an astrometric measurement can be given by ([Sha1992]):

$$\sigma_\delta^2 \simeq 5.25 B^{-4/3} \theta^2 \int h^2 C_n^2(h) (Vt)^{-1}\,\mathrm{d}h \quad . \tag{3.1}$$

Here B is the telescope diameter, or the length of the interferometric baseline, $V(h)$ is the height-dependent wind speed in the atmosphere, t the time of integration, and $C_n^2(h)$ the vertical profile of the atmosphere's refractivity power spectrum as induced through temperature fluctuations. As one can see, the accuracy now depends not only on the objects' angular separation, but also on the length of the baseline. By the means of simultaneous interferometric measurements on two objects in the narrow-angle regime with sufficiently long integration times and baselines of the order of 100 m astrometric accuracies in the range of tens of microarcseconds are in principle achievable ([Sha1992]).

This strategy is pursued by the instrument PRIMA, incorporated at the VLTI at the Cerro Paranal in Chile and described in detail in chapter 4.

In preparation for the understanding of the principles, which the implementation of the instrument is based on, this chapter shall provide the theoretical background of interferometric concepts in general and lay out, in which way dual star interferometry can be utilized for astrometric measurements.

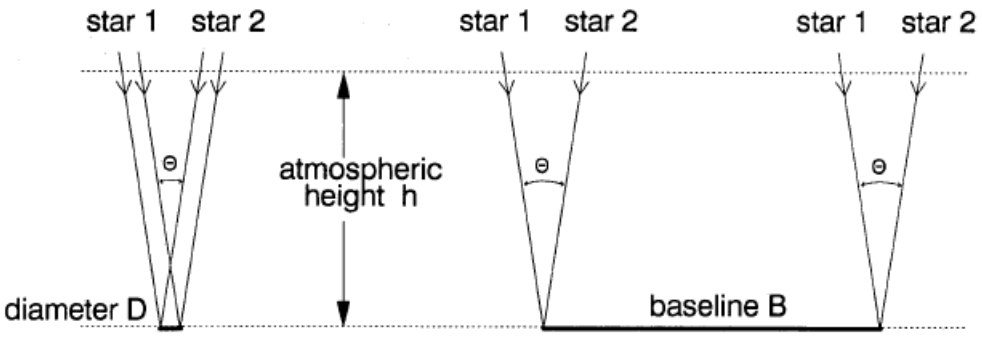


Figure 3.1. – This sketch depicts the difference between narrow-angle astrometry performed in a traditional way with single long-focus telescopes and the situation when a long baseline dual-star interferometer is used for the measurement. On the left, representing the single-dish measurement with an instrument of diameter D, the astrometric error induced by the turbulent atmosphere is only weakly dependent on θ, the angular separation of the two objects, and mostly independent on the instrument's extent. On the contrary, when the instrument extent increases to the regime so that the condition $\theta h \ll B$ is met, which can be achieved by long interferometric baselines, the error's dependency on the separation becomes linear and decreases with increasing B (comp. Eq. (3.1)). (Credit: [Sha1992])

3.1. Principles of Interferometry

Historically, observed phenomenons of interference have been indications to the wave character of light, which formally can be derived from the Maxwell's equations. Deduced from these, visible light and generally electromagnetic waves of any frequency are manifestations of the propagation of a coupled vector field, namely the self-inducing electromagnetic field. Hence, in order to cover the major properties of electromagnetic waves, it is convenient to reduce the matter to the description of a plane parallel monochromatic field of frequency ν, propagating in free space in direction $\hat{\boldsymbol{n}}$:

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}_0 \cos(\boldsymbol{k}\boldsymbol{r} - \omega t + \varphi_0) \quad . \tag{3.2}$$

This wave equation describes the electric field vector at any time t and location $\boldsymbol{r}$ in space, with the angular frequency ω and the wave vector $\boldsymbol{k}$ defined by

$$\begin{aligned} \omega &= 2\pi\nu = 2\pi c/\lambda \quad , && (3.3)\\ \boldsymbol{k} &= k\hat{\boldsymbol{n}} = 2\pi/\lambda \quad , && (3.4)\\ k &= \omega/c = 2\pi\nu/c \quad , && (3.5) \end{aligned}$$

where λ is the wavelength and $c = \lambda\nu$ stands for the wave's phase velocity.
An additional phase shift due to any reasons can be incorporated by the quantity φ_0.
Especially when dealing with interference phenomena, it is mathematically very handy, when the wave is described by a complex function of the form

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{r},t) &= \boldsymbol{E}_0\, e^{i(\boldsymbol{kr}-\omega t+\varphi_0)} \\ &= \boldsymbol{E}_0 e^{i\varphi_0}\, e^{i(\boldsymbol{kr}-\omega t)} \\ &= \tilde{\boldsymbol{E}}_0\, e^{i(\boldsymbol{kr}-\omega t)} \quad , && (3.6) \end{aligned}$$

where $\tilde{\boldsymbol{E}}_0$ generally is a complex entity. However, it should be clear that only the real part of this representation is of physical significance.
The electromagnetic field propagates as an transverse wave with the oscillating and one another inducing electric and magnetic fields E, H being perpendicular to each other and to the direction of propagation. Generally, the direction of the field vectors, as denoted by the vector character of $\boldsymbol{E}_0$, is not static, but is described by the superposition of two orthogonal components with different phases. The phase difference δ between them determines the polarization state of the wave. In the cases when $\delta = 0$ or $\delta = \pm\pi$, the field vectors oscillate in a fixed direction and the electromagnetic wave is called linearly polarized. As far as not stated differently, these waves are considered throughout of the chapter, as this property does not contribute to the principle discussion.

Although in Eq. (3.6) the electric field is used for the description of the wave, this is just an arbitrary choice of convention, since the wave is entirely defined by any of the two field vectors. In fact, as long as electromagnetic field propagation is considered in isotropic media only, the magnetic field could be used instead as well.
As characteristic for waves, the electromagnetic wave is naturally described by an oscillating amplitude and a phase ϕ. In interferometry, which deals with the superposition of electromagnetic waves, it is indeed the phase relation between the two beams that determines the result. Before dealing with that it should be clarified, how the power of incident radiation, detected by optical detectors, can be described. Clearly, it is related to the energy that is transported by the electromagnetic wave. The energy flow density of an electromagnetic wave is usually given by the *Poynting vector* $\boldsymbol{S} \parallel \hat{\boldsymbol{n}}$:

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H} \quad . \qquad (3.7)$$

For linearly polarized waves in vacuum the Poynting vector has only a component in the direction $\hat{\boldsymbol{n}}$ of propagation and its absolute value is ([Gli11])

$$\begin{aligned} S_{\hat{n}} = |\boldsymbol{S}| &= E_0 \cos(\boldsymbol{kr} - \omega t) \cdot H_0 \cos(\boldsymbol{kr} - \omega t) \\ &= E_0 \frac{B_0}{\mu_0} \cos^2(\boldsymbol{kr} - \omega t) \\ &= \frac{1}{\mu_0 c} E_0^2 \cos^2(\boldsymbol{kr} - \omega t) \\ &= c\epsilon_0 E_0^2 \cos^2(\boldsymbol{kr} - \omega t) \quad , \end{aligned} \tag{3.8}$$

where for the vacuum values of the permittivity and permeability μ_0 and ϵ_0 the relation $c^2 = 1/(\epsilon_0 \mu_0)$ has been exploited.
As one can see, at a defined and fixed location the value of the energy flow density oscillates with twice the frequency of the electromagnetic field. The temporal resolution of the optical detectors is far too low to resolve this oscillatory behavior so that usually only a time average is measured. This temporal mean of a time-dependent signal $f(t)$ can be defined by:

$$< f(t) > = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f(t) \, \mathrm{d}t \quad . \tag{3.9}$$

Using this definition, the detected time average of the Poynting vector, usually referred to as flux or irradiance, can be derived to (see App. B):

$$\begin{aligned} F = < \boldsymbol{S}_{\hat{n}} > = < \boldsymbol{E} \times \boldsymbol{H} > &= \frac{1}{2} \mathrm{Re} \left[\tilde{E}_0 \tilde{H}_0^* \right] \\ &= \frac{c\epsilon_0}{2} \mathrm{Re} \left[\tilde{E}_0 \tilde{E}_0^* \right] \\ &= \frac{c\epsilon_0}{2} \tilde{E}_0^2 \quad . \end{aligned} \tag{3.10}$$

Here it is important to note that the observed quantity is proportional to the square of the modulus of the wave's amplitude.[1]
Furthermore, since the flux gives the received energy per unit time and cross-sectional area, the overall detected power depends on the effective area of the instrument. This is typically affected by collecting and also limiting apertures and the illuminated area on the detector. Naturally, the spectrum of the source, as well as the detector's usually wavelength dependent efficiency need to be taken into account.
If two light beams are superposed and combined directly at the detector, interference of the fields takes place at the location of the detection, and the phase difference between the beams is decisive for the outcome. Such two monochromatic signals of the same frequency, as given by Eq. (3.6), but with a well-defined and stable phase shift of $\Delta\varphi$ read then as:

$$\begin{aligned} E_1(\boldsymbol{r}, t) &= A_1 \, e^{i(\boldsymbol{k}_1 \boldsymbol{r} - \omega t)} \\ E_2(\boldsymbol{r}, t) &= A_2 \, e^{i(\boldsymbol{k}_2 \boldsymbol{r} - \omega t + \Delta\phi)} \quad . \end{aligned} \tag{3.11}$$

[1] The superscript * denotes the complex conjugate of the complex quantity.

Here the vector nature of the fields has been neglected for simplicity.
If optical path differences for the two waves are ruled out, and $\boldsymbol{k}_1\boldsymbol{r}_0 = \boldsymbol{k}_2\boldsymbol{r}_0$ so that at the location $\boldsymbol{r}_0$ of the detector the overall phase difference is determined solely by the initial phase shift $\Delta\phi$, the resulting measurable field is given by the superposition:

$$E_\Sigma(\boldsymbol{r}_0, t) = e^{i(\boldsymbol{k}\boldsymbol{r}_0 - \omega t)} \left(A_1 + A_2\, e^{i\Delta\phi}\right) \quad . \tag{3.12}$$

Following the derivation above for the resulting flux we get:

$$\begin{aligned} F_\Sigma &\propto \mathrm{Re}\left[\left(A_1 + A_2\, e^{i\Delta\phi}\right)\left(A_1 + A_2\, e^{i\Delta\phi}\right)^*\right] \\ &\propto A_1^2 + A_2^2 + 2A_1A_2\cos\Delta\phi \\ &= F_1 + F_2 + 2\sqrt{F_1F_2}\cos\Delta\phi \quad . \end{aligned} \tag{3.13}$$

Due to the interference term $\sqrt{F_1F_2}\cos\Delta\phi$, the measured flux depends also on the ratio of the power from the individual waves. In the special case when $F_1 = F_2$, the relation reduces to:

$$\begin{aligned} F_\Sigma &= 2F_1(1 + \cos\Delta\phi) \\ &= 4F_1\cos^2\frac{\Delta\phi}{2} \quad . \end{aligned} \tag{3.14}$$

The outcome is described by a periodic function, where the flux oscillates between its maximum values $F_{\max} = 4F_1$ at phase differences of zero and even multiples of π and its minima $F_{\min} = 0$ respectively at uneven multiples.
Recalling that the flux measurement is a time average, it should be realized that this location dependent total flux pattern can only be detected if the phase relation between the two beams is stable for at least over the time of the integration on the detector. This condition is referred to by the term *temporal coherence* (see Sec. 3.1.1 and 3.1.2).

If by any means the intensity can be measured for different phase differences, either by appending optical path differences (OPDs) to one of the two light paths, or simply by measuring at different locations at a screen or a plane detector, such a fringe pattern can be registered.
The ratio between the extreme values in these intensity patterns provides a measure for the fringe contrast. For this purpose the entity

$$\mathcal{V} = \frac{F_{\max} - F_{\min}}{F_{\max} + F_{\min}} \quad , \tag{3.15}$$

called the *fringe visibility*, is introduced. The interference pattern from above yields perfect visibility of $\mathcal{V} = 1$, but, as one can see from Eq. (3.13), already unequal flux values from the two overlapping beams reduce the fringe visibility. In general, with a linear relation between the intensities, as $F_2 = xF_1$, the total flux becomes $F_\Sigma = F_1(1 + x + 2\sqrt{(x)}\cos\Delta\phi)$, and the visibility reads as:

$$\mathcal{V} = \frac{2\sqrt{x}}{1+x} \quad . \tag{3.16}$$

With this in mind it is clear that, especially when beams from two different apertures following different light paths, where several distinct optical elements are involved, parallelization in the

instrument's design should be pursued in order to keep the decrease in visibility due to the mismatch in intensities as low as possible.
Another reason for keeping the instrument symmetric for the two light paths are polarization effects. Usually, when light is being reflected at a mirror, then phase shifts are induced between the electric vector components that are perpendicular and parallel to the plane of incidence, namely the s and p components. Within a perfect symmetric layout, where the sequence of reflections is the same for the two beams and equivalent mirrors are used, identical phase shifts ϕ_{pol} between the two polarization components can be reached for the two beams.
Consequently, at the location of the beam combination, the s and p components combine independently and generate identical fringe packets. In reality such ideal conditions are not met and the visibility is lowered by the factor V_{pol} ([Tra1999]):

$$V_{\mathrm{pol}} = |\cos\frac{\phi_{\mathrm{pol}}}{2}| \quad . \tag{3.17}$$

3.1.1. Concept of Coherence

Although interference is a natural result from the superposition of electric wave fields, in order to measure fringe patterns by using optical detectors, the interfering light beams need to be at least partially coherent. This means that the phase relation between the beams is well-defined and stable over time intervals compared to the detector integration times. To understand the concept of optical coherency, a formalism, as introduced by e.g. [Gli11], turns out useful and will be followed here.
For this purpose a complex quantity, the *optical disturbance* $v(\boldsymbol{r},t)$ is introduced by a linear relation to one component of the electrical field vector $v = CE$, where C is a constant to make the optical disturbance dimensionless.[2] It is convenient to separate the time dependency of the optical disturbance when describing propagating waves by introducing an, consequently also dimensionless, amplitude $V(\boldsymbol{r})$ by:

$$v(\boldsymbol{r},t) = V(\boldsymbol{r})\, e^{-i2\pi\nu t} \quad . \tag{3.18}$$

The dimensionless intensity $I(\boldsymbol{r})$, proportional to and analogously introduced as the flux, as evaluated by Eq. (3.9) and (3.10), the dimensionless intensity is defined by the time average $< v(\boldsymbol{r},t)v^*(\boldsymbol{r},t) >$ to:

$$\begin{aligned} I(\boldsymbol{r}) &:= < v(\boldsymbol{r},t)v^*(\boldsymbol{r},t) > \\ &= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} v(\boldsymbol{r},t)v^*(\boldsymbol{r},t)\,\mathrm{d}t = |V(\boldsymbol{r})|^2 \quad . \end{aligned} \tag{3.19}$$

When polychromatic light is considered, the optical disturbance can be represented by the linear superposition of individual monochromatic waves. This is done by the not time but frequency dependent *spectral amplitude* $V(\boldsymbol{r},\nu)$ so that:

$$v(\boldsymbol{r},t) = \int_0^\infty V(\boldsymbol{r},\nu)\, e^{-i2\pi\nu t}\,\mathrm{d}\nu \quad . \tag{3.20}$$

[2]Note that here only one arbitrary component E, perpendicular to the direction of propagation, of the field vector $\boldsymbol{E}$ was implicitely chosen.

With this given, also a frequency dependent *spectral intensity* $I(\boldsymbol{r}, \nu)$ can be introduced so that an integration over all frequencies in turn gives the intensity $I(\boldsymbol{r})$ from above:

$$I(\boldsymbol{r}) = \int_0^\infty I(\boldsymbol{r}, \nu)\, \mathrm{d}\nu \quad . \tag{3.21}$$

This approach proves convenient when propagation of light from real, not idealized sources is regarded. In general these sources are thermal radiators and emit electromagnetic waves in a broad spectrum. Moreover, although extended sources, such as stars, can be treated approximately point-like at sufficiently large distances so that the wave fronts are often regarded as plane waves, their light is composed of contributions from many individual and independently radiating points on the source's surface. Consequently, at any time and location on the wave fronts, the optical disturbance (or analog the amplitude of the electric field) must be considered as a quantity of unpredictable values, fluctuating at timescales $1/\nu_0$, where ν_0 is the average frequency. As in reality detector integration times exceed these time scales, the fluctuations average out and a well-defined value of the flux or intensity can be measured.

Hence, the concept of coherence describes the statistics of light propagation, based on the electromagnetic theory. In particular, one is interested in the correlations of the fluctuating amplitude $v(\boldsymbol{r}, t)$ at different points in space and time. It is helpful to depict the propagating waves, as it is done in Fig. 3.2. It shows a sample of wave fronts at a given moment in time $t = t_0$. The fronts are defined here by their values of the optical disturbance along the solid lines. The emission is supposed to have taken place at different moments in time from an extended source at a large distance, whereas after that the illustrated fronts propagated along the z-axis. Consequently, the random fluctuations regarding the shape of and the spacings between the fronts represent the random character of the emission processes and are decisive for the temporal and spatial coherency of the wave field. The temporal coherency describes the phase correlations at a given location $\boldsymbol{x}_0$ for a time interval τ, whereas the correlation at a fixed time t_0 between an ensemble of optical disturbance values is referred to by spatial coherency. In order to analyze the statistics of the wave fronts, *ergodicity* is assumed for the random process. Two conditions need to be fulfilled to justify this:

- Statistical stationarity of the radiation process in time results in the statistical properties being the same over any chosen ensemble. In particular this means that the time averages are independent of the absolute value of t, and the correlations depend only on time differences $\tau = t_i - t_j$.
- Equivalence of time and space concerning the optical disturbance as the random variable. If this is satisfied, statistical quantities, as the time averages at a given location in space can be replaced by the corresponding and equal ensemble averages over a defined wave front at a fixed moment in time. This property can also be expressed by the relation

  $$E\{v_0(\boldsymbol{x}, t_0) v_0^*(\boldsymbol{x}, t_0)\} = < v_{z_0}(\boldsymbol{x}_{01}, t) v_{z_0}^*(\boldsymbol{x}_{01}, t) > \quad , \tag{3.22}$$

  where $E\{v_0(\boldsymbol{x}, t_0) v_0^*(\boldsymbol{x}, t_0)\}$ stands for the ensemble average over a wave front at time $t = t_0$ and a defined distance $z = z_0$ (comp. Fig. 3.2), and the right hand side of the equation represents the time average of wave fronts passing through the fixed location x_{01}.

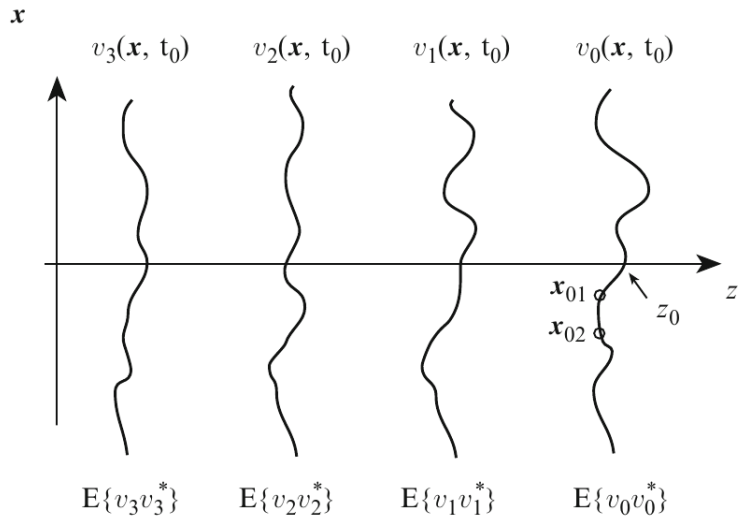


Figure 3.2. – The propagation of an electromagnetic wave field along the z-axis is depicted here as a snapshot at a given moment in time t_0. The illustrated individual wave fronts are defined by their values of the optical disturbance $v_i(\boldsymbol{x}, t_0)$, where the subscript i distincts the individual fronts and indicates their position on the z-axis. The light was emitted by an extended source at a large distance and at different times so that their position at t_0 is determined by the light travel time. Since the emission as a whole is composed of independent and random emission processes on the source's surface, the fronts themselves resemble the random character, where the optical disturbance can be regarded as the random variable. Due to ergodicity of the process the ensemble averages $E\{v_i, v_i^*\}$ of the fluctuating variable v over the wave fronts at the fixed moment in time equal the corresponding time averages at a fixed location in space, where subsequent wave fronts are passing through, if the integration time is significantly longer than the oscillation frequency of the wave field (see Eq. (3.22)). (Credit: [Gli11])

Now, for the exploration of the wave field's coherency the *mutual coherence function* (MCF) is introduced by [Gli11]:

$$\begin{aligned} \Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau) &= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} v(\boldsymbol{r}_1, t+\tau) v^*(\boldsymbol{r}_2, t)\, \mathrm{d}t \\ &= < v(\boldsymbol{r}_1, t+\tau) v^*(\boldsymbol{r}_2, t) > \quad . \end{aligned} \tag{3.23}$$

It describes the correlation of the optical disturbances and respectively the electrical fields at positions $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and different times $t+\tau$ and t.
If one compares this definition to Eq. (3.19), one can see that the evaluation of the MCF at the same moment and two identical locations gives the dimensionless intensity $I(\boldsymbol{r}) = \Gamma(\boldsymbol{r}, \boldsymbol{r}, 0)$.
As the fringe visibility defined above, the MCF can provide a measure for the degree of coherence. The *correlation coefficient* $\gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau)$, in general a complex quantity, yields the

correlation of the fields at different location and times in space:

$$\gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau) = \frac{\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau)}{\sqrt{\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_1, 0)\Gamma(\boldsymbol{r}_2, \boldsymbol{r}_2, 0)}} \quad . \tag{3.24}$$

A normalization has been performed here by the help of the intensities $I(\boldsymbol{r}_1)$ and $I(\boldsymbol{r}_2)$ so that the value of the modulus of the coefficient γ lies in the range between 0 and 1.

A similar approach is followed when the spectral distribution of light is taken into account. As before the spectral amplitude was introduced in Eq. (3.20), now the *mutual spectral density function* (MSDF) is defined by

$$\hat{\Gamma}(\boldsymbol{r}_1, \boldsymbol{r}_2, \nu) \coloneqq \lim_{T\to\infty} \frac{1}{2T} E\{V_T(\boldsymbol{r}_1, \nu) V_T^*(\boldsymbol{r}_2, \nu)\} \quad , \tag{3.25}$$

where the quantity is evaluated for a period of time $2T$, which as usual is assumed to be long compared to the oscillation time scale of the electromagnetic wave ($T \gg 1/\nu$).
Given this definition, by the evaluation of the correlation of the spectral amplitudes between two distinct locations $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, the MSDF describes the spatial coherence at the frequency ν, and consequently forms a Fourier pair together with the MCF:

$$\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau) = \int \hat{\Gamma}(\boldsymbol{r}_1, \boldsymbol{r}_2, \nu)\, e^{-i2\pi\nu\tau}\, \mathrm{d}\nu \quad . \tag{3.26}$$

Just as the MCF at two identical points and the same moment yields the intensity $I(\boldsymbol{r})$, the spectral intensity $I(\boldsymbol{r}, \nu)$ can be defined by the evaluation of the MSDF at $\boldsymbol{r}_1 = \boldsymbol{r}_2$:

$$I(\boldsymbol{r}, \nu) \coloneqq \hat{\Gamma}(\boldsymbol{r}, \boldsymbol{r}, \nu) = \lim_{T\to\infty} \frac{1}{2T} E\{V(\boldsymbol{r}, \nu) V^*(\boldsymbol{r}, \nu)\} \quad . \tag{3.27}$$

Now the relation (3.21) between the spectral intensity $I(\boldsymbol{r}, \nu)$ and the white-light or polychromatic intensity $I(\boldsymbol{r})$ can be expanded to

$$I(\boldsymbol{r}) = \Gamma(\boldsymbol{r}, \boldsymbol{r}, 0) = \int \hat{\Gamma}(\boldsymbol{r}, \boldsymbol{r}, \nu)\, \mathrm{d}\nu = \int I(\boldsymbol{r}, \nu)\, \mathrm{d}\nu \quad , \tag{3.28}$$

where the integral is taken over the whole frequency band.

3.1.2. Temporal and Spatial Coherence

Although at a first glance the introduction of the coherence functions appears academic and abstract, they build a solid and convenient framework for the understanding of basic concepts and observations in interferometry. When it comes to practical realizations of interferometric instruments, there are in principle only two general methods, by which two beams can be extracted from a single light source and thereafter be brought to superposition. These are depicted in the two set-ups in Fig. 3.3.

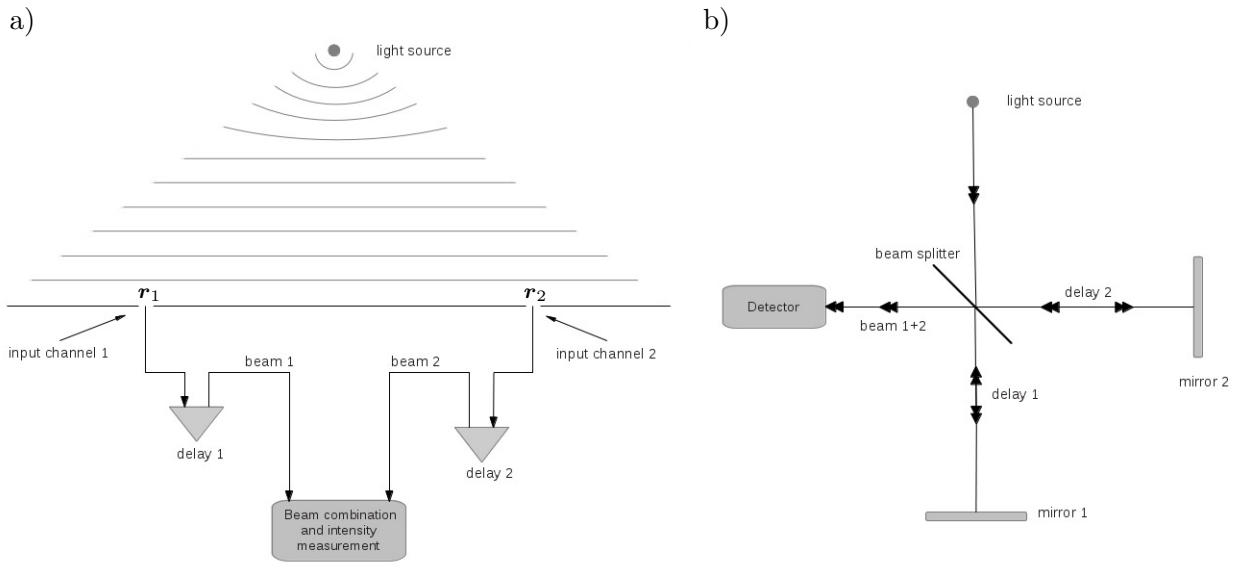


Figure 3.3. – The picture illustrates two general concepts, by which interferometric measurements can be accomplished. In order to measure the interference pattern of light emitted by one single source, the light needs to be divided into two (or even multiple) individual beams, which then again can be superposed in order to probe the resulting interference pattern, which is determined by the beams' relative phase relations. a): A wave field, emitted by a source at a large distance so that in the vicinity of the instrument the waves are assumed to be approximately plane, is extracted and probed by two apertures at different locations separated by the baseline vector $\boldsymbol{B} = \boldsymbol{r}_1 - \boldsymbol{r}_2$. This method is often referred to by the term *division of wave front*. At these locations the two beams are fed into the system. Within the instrument they follow different paths, before they are recombined again, and the interference pattern is observed at the detector. Generally, the distinct paths are different in length, and additional delay lines can be introduced in order to compensate for any phase differences or also to introduce some on purpose. The measured intensity of the interference depends on the one hand on the coherence between the two beams, but on the other hand on the effective phase difference between the beams at the location of the detector. This phase shift is given by $\Delta\phi = \mathrm{OPD} \cdot 2\pi/\lambda$, where OPD is the optical path difference of the two beams and λ the wavelength. Typically, one distinguishes between an internal OPD, which arises within the interferometer itself and an external path difference that results from e.g. unequal distances between the source and the two apertures. In these cases the delay lines can be used to compensate for this initial phase shift at the entrance of the apertures and the interference pattern can then be measured at $\Delta\phi = 0$, also called the zero OPD. b): In this set up one beam from the source is divided by a half reflecting mirror (the beam splitter) and, before recombined again, the two individual beams travel along different paths inside the instrument. This method is called *division of amplitude* and the effective OPD is determined solely by the internal delays. The fringe contrast in the measured interference pattern of both methods is determined by the degree of coherence between the two beams. Set up a) picks up the concept of spatial coherence, while temporal coherence only is picked up by instruments that are designed, as shown in b).

The scheme of the arrangement in Fig. 3.3 a) follows the principle of Young's double-slit interferometer, where two apertures are used to collect light from the same source separately. Since parts of the incoming wave are extracted at different locations $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ of the incoming wave front, this procedure is also called *division of wave front* and the distance vector between the apertures is usually called the baseline $\boldsymbol{B} = \boldsymbol{r}_1 - \boldsymbol{r}_2$. In order to investigate the interference pattern, the inputs at the two different points in space need to be combined again, so that effects, deriving from the superposition of the two beams, can be observed. As described above, the measured intensity of the combined light depends on the phase differences between the two beams at the point of the measurement. Generally, the phase differences arise from and are determined by any path length differences Δd between the two beams through the relation:

$$\Delta\phi = \Delta d \cdot 2\pi/\lambda \quad . \tag{3.29}$$

Here in the illustration, the path length difference is defined by the *optical path distance* (OPD). It is schematically depicted as resulting from two distinct delays $DL1$, $DL2$ to the beams in the two interferometer arms, and is directly evaluated from their values to $\mathrm{OPD} = \mathrm{DL2} - \mathrm{DL1}$. In practice, the delay lines inside of the instrument are used to compensate any external OPDs that are already present at the entrance pupil, the telescope apertures, as naturally would arise at observations of off-zenith sources.
Another method is adopted in the cases when light from a source that is about to be observed by an interferometer is divided into two distinct beams, which follow paths of different length in the system, before being superposed again. This procedure is referred to by the term *division of amplitude*. For the purpose of the light division, a beam splitter is used and in this case the resulting phase differences are solely determined by the internal delays, as illustrated by Fig. 3.3 b).
The beam combination itself again can be achieved by two fundamentally different techniques. In the *image-plane interferometry* each of the two beams is focused separately. The two formed images are then superposed and combined so that a fringe pattern can be observed within the image-plane. In such an interferometer implementation the telescope apertures with diameter D and their distribution are reimaged by the optical system into the exit pupil inside the instrument. Consequently, the distance between the telecopes, the modulus B of the baseline, is also reimaged inside the beam combining instrument. This downscaled entity is sometimes called the *beam-combination baseline* B_i. Now, in the focus plane the intensity distribution on one hand depends on the beams' phase differences, but on the other hand is affected by the finiteness of the used apertures, so basically by their shape and diffraction pattern, which are described by the aperture's PSF function. In the case of circular telescope apertures (see Eq. (2.58)) point-like sources are imaged as Airy disks of a given width. The position of the center of an Airy disk, belonging to the corresponding object in the sky plane, is determined by the angular magnification factor $m_p = \theta_\mathrm{A}/\Delta s = D/D'$, where Δs is the object's angular distance to the telescope pointing direction $\hat{\boldsymbol{s}}_0$ on sky, θ_A the corresponding angular position of the center of the Airy disk in the image and D' the diameter of the downscaled image of the telescope aperture, the exit pupil. If internal delays in the instrument are adjusted in such a manner that the external OPD is compensated for objects in the pointing direction, then a fringe pattern shows within the extent of the corresponding Airy disk, which acts as an envelope to the intensity distribution. For the source that the telescopes are pointed at

the point in the interference pattern, where the OPD between the two beams equals zero, also called the *zero OPD position* or the *central fringe*, is well aligned with the center of the Airy disk. The telescope pointing direction acts as a phase reference on the sky, as for objects, apart from these reference coordinates by the angle[3] $\Delta\boldsymbol{s}$, the external OPD changes by the amount of $\Delta\mathrm{OPD}_{\mathrm{ext}} = \boldsymbol{B} \cdot \Delta\boldsymbol{s}$. Therefore, the position of the central fringe of such objects will shift also in the image plane. Unfortunately, this shift is not equal to the amount of shift of the position of the Airy disk, with regard to the source's position on the sky. While the second is determined by the ratio of the diameters of the entrance and exit pupil, the angular position θ_0 of the zero OPD in the image plane is determined by the baseline demagnification factor $m_b = \theta_0/\Delta s = B/B_{\mathrm{i}}$. Consequently, in general realizations of such instruments, also referred to as *Michelson configurations*, the center of the Airy disk and the central fringe will only coincide for objects in pointing direction, whereas shearing will show for off-axis sources. This becomes an issue particularly when spectral bandwidth (temporal coherence) comes into play. As will be discussed below, interference patterns of sources radiating in a finite spectral band, show maximum contrast only at the central fringe position, while they blur out with increasing distance to the zero OPD. Hence, for polychromatic off-axis sources the fringe pattern shifts under the Airy disk so that fringes can only be detected with decreased contrast, and if the angular distance from the pointing becomes too large, then no interference pattern can be observed at all. For this reasons Michelson interferometers suffer from a highly limited field of view.
Only in the special case, when the system's set-up fulfills the condition $m_p = m_b$, the PSF is shift-invariant and the field of view is not limited. Instruments designed this way are called to be in *Fizeau configuration*.

In contrast to the beam combination in the image-plane, a half-silvered mirror or equivalent can be used as a beam combiner. As indicated in the combination process of the set-up in Fig. 3.3 b), the output is a direct superposition of two parallel and completely overlapping beams before reaching the detector. In this way, the telescope apertures are imaged on top of each other and the resulting Airy disk shows no fringes, but its overall intensity is regulated only by the phase difference of the beams and is fed typically into single detector pixels. In order to investigate the interference pattern in such a set-up, one must scan the OPD space by changing the internal delays $DL1$ and $DL2$. This beam combination and fringe detection method is called *pupil-plane interferometry*, or *co-axial combination*.

Based on the two described and conceptionally different methods of interferometric measurements, division of wave front and division of amplitude, the two distinct aspects of coherence, namely the temporal and spatial coherence, can be introduced. The dependency of the fringe contrast, or to be more precise, the coherence function on the spectral bandwidth of the incoming light can be analyzed when the MCF is applied to the situation, where the beam is split at one single point $\boldsymbol{r}$. This construct of the MCF $\Gamma(\boldsymbol{r}, \boldsymbol{r}, \tau)$, where τ is arbitrary and stands for the measurement/combination at two different times, which is achieved by the delays introduced to the light paths, is a measure for the temporal coherence of the light source and is called the *self-coherence function*. Applying Eq. (3.26) and Eq. (3.27) to this definition, one can see

[3] The vector character of this angular entity emphasizes the importance of the direction of the displacement in relation to the baseline direction.

that the self-coherence function can be constructed by the Fourier transform of the spectral intensity:

$$\Gamma(\boldsymbol{r},\boldsymbol{r},\tau) = \int \hat{\Gamma}(\boldsymbol{r},\boldsymbol{r},\nu)\, e^{-i2\pi\nu\tau}\,\mathrm{d}\nu = \int I(\boldsymbol{r},\nu)\, e^{-i2\pi\nu\tau}\,\mathrm{d}\nu \quad . \tag{3.30}$$

The fact that the self-coherence function and consequently the temporal coherence depends on the light's spectral intensity distribution is not that surprising. Given a finite source spectral bandwidth $\Delta\nu$, it is clear that at the position of zero OPD, or analogous $\tau = 0$, the wave interferes constructively with itself throughout the whole spectrum and the coherence is at its maximum. However, when an OPD is introduced, it resembles the interference of the wave at the moment t with itself, but at the time $t + \tau$. The phase relations between the wave's contributions of different wavelengths change with the OPD and τ respectively due to OPD $= c\tau$. When these are increased, the crests and troughs of the individual wavelengths fail to overlap more and more. This results in the Fourier transform of the spectral intensity distribution acting as an envelope to the fringe pattern. The narrower the spectral bandwidth of the source, the wider the self-coherence function will be. In order to measure this effect, the quantity *coherence time* τ_c is introduced as the value of τ, at which the self-coherence function is significantly reduced.
In interferometry, when working with delays in the instrument, one is also interested in the limitations of the overall OPDs that can be introduced, before the visibility decreases that much so that fringes cannot be detected any more. For these reasons the term *coherence length* is defined by $l_c = c\tau_c$.
It is important to note that the coherence time, as introduced at this point, is a natural characteristic of light and mainly determined by the source's spectral width. Therefore, it defines the light's temporal coherence and acts as a natural limitation on permitted path differences within interferometric instruments. Furthermore, it must not be confused with effective coherence times, which can be even shorter, when additional sources of systematics along the light paths are taken into account, as it is done in the context of atmospheric turbulences (see also Sec. 3.2.2).
The relation (3.30) between the coherence function and the spectral intensity distribution of the observed light can also be exploited for instrument characterization. As will be shown in Sec. 4.1.3, by scanning the fringe pattern, which is achieved by continuously sweeping through the OPD space, the coherence function, with respect to τ, is probed, and from that the intensity spectrum, as arriving and detected by the detector, can be obtained so that the bandpass of the system can be analyzed.

Just as the temporal coherence is affected by the source's spectral distribution, it is interesting to investigate the dependency of the spatial coherence on any source's features. For this purpose the propagation of the MSDF from the source plane to the plane of observation is derived to

$$\hat{\Gamma}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\nu) = \frac{1}{(\lambda z_0)^2} \iint_{\Sigma} \hat{\Gamma}(\boldsymbol{x}_1',\boldsymbol{x}_2',\nu)\, e^{ik(r_1-r_2)}\,\mathrm{d}\boldsymbol{x}_1'\,\mathrm{d}\boldsymbol{x}_2' \quad , \tag{3.31}$$

where $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ stand for coordinate vectors within the plane of observation and respectively $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ for vectors in the source plane Σ (compare with Fig. 3.4).
It has also been assumed here that all angles of interest are small, which in general is justifiable

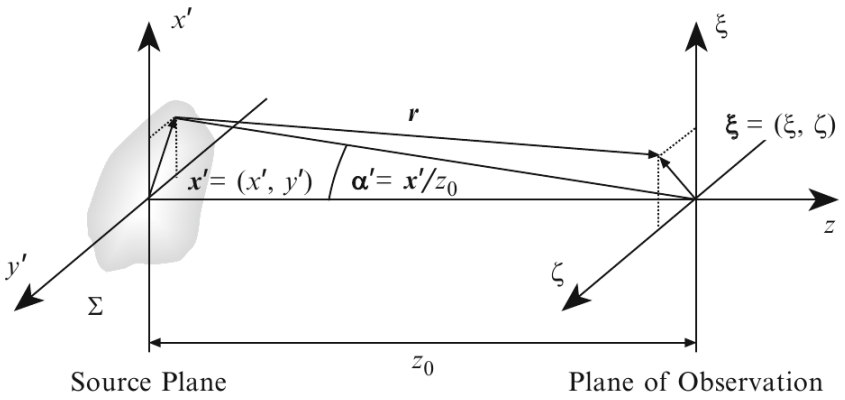


Figure 3.4. – Here the source lies within the source plane Σ, and at a distance of z_0 the instrument's apertures define the plane of observation. The corresponding coordinate vectors are given by $\boldsymbol{x}'$ and $\boldsymbol{\xi}$ respectively. The typically incoherent source is small in comparison to z_0 so that the small angle approximation for the angle of observation $\boldsymbol{\alpha}' = \boldsymbol{x}'/z_0$ is justified, and when this is interpreted as a source coordinate, the source brightness distribution can be described by $I_b(\boldsymbol{\alpha}')$. (Credit: [Gli11])

if the distance z_0 to the source is much larger than the extent of the source, as well as any considered distances within the plane of observation. Given this and applying the *Fresnel approximation*, where the modulus of the distance vector $\boldsymbol{r}$ is described by a power series with higher orders neglected, the relation (3.31) can be simplified to the approximation:

$$\hat{\Gamma}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \nu) = G(\nu) \int I_b(\boldsymbol{\alpha}')\, e^{-ik(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\,\cdot\,\boldsymbol{\alpha}'}\, \mathrm{d}\boldsymbol{\alpha}' \quad . \tag{3.32}$$

Here the angle of observation $\boldsymbol{\alpha}' = \boldsymbol{x}'/z_0$ is used as a source coordinate, and the source is assumed to be spatially incoherent, which is the case for all real celestial bodies of finite size, as the thermal radiation from each point on their surfaces is independent of radiation from adjacent points. In addition to that, Eq. (3.32) holds only for the simplification that over the observed spectrum the source shape is independent of the wavelength. In this situation the expression of the spectral intensity can be split into two independent parts, namely the *source spectrum* $G(\nu)$ and the *source brightness distribution* $I_b(\boldsymbol{\alpha}')$:

$$I_b(\boldsymbol{\alpha}', \nu) = I_b(\boldsymbol{\alpha}')G(\nu) \quad , \tag{3.33}$$

with the source spectrum usually calibrated to unity so that $\int G(\nu)\,\mathrm{d}\nu = 1$.
The most interesting result from this is that the MSDF in the plane of observation and consequently the spatial coherence do not depend on the absolute coordinates $\boldsymbol{\xi}_i$, but only on their difference $(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2) = \boldsymbol{B}$, and that the MSDF is described by the product of the source spectrum and the Fourier transform of the source's brightness distribution.
Evaluating it at $\boldsymbol{\xi}_1 = \boldsymbol{\xi}_2$ gives by definition (Eq. (3.27)) the spectral intensity at frequency ν, which by using Eq. (3.32) yields the constant spectral intensity in the plane of observation $I(\boldsymbol{\xi}, \nu) = \hat{\Gamma}(0, \nu) = G(\nu) \int I_b(\boldsymbol{\alpha}')\,\mathrm{d}\boldsymbol{\alpha}' = G(\nu)I_0$.

The integration of the spectral intensity over the spectral band gives the constant value I_0 for the polychromatic intensity.

The propagation of the MCF must be looked upon, as it gives the propagation of the coherence function in the general polychromatic case. In order to do that, it is a common procedure to evaluate this problem in the *quasi-monochromatic approximation*. Doing so, the MCF in the plane of observation, which is given by

$$\begin{aligned} \Gamma(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2, \tau) &= \int \hat{\Gamma}(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2, \nu)\, e^{-i2\pi\nu\tau}\, \mathrm{d}\nu \\ &= \int G(\nu) \int I_b(\boldsymbol{\alpha}')\, e^{-ik(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\cdot\boldsymbol{\alpha}'}\, \mathrm{d}\boldsymbol{\alpha}'\, e^{-i2\pi\nu\tau}\, \mathrm{d}\nu \quad , \end{aligned} \tag{3.34}$$

can be decomposed in a purely spatial and a purely temporal part, which is accomplished by the assumption that not only the source brightness distribution itself is independent of λ, but that also its Fourier transform is invariant over the regarded spectrum. This in general does not need to be the case, due to the appearance of the wavenumber $k \propto 1/\lambda$ within the exponent of Eq. (3.34). By restricting $G(\nu)$ to a narrow spectrum, where its width is much smaller than the average frequency ($\Delta\nu \ll \nu_0$), the Fourier transform of $I_b(\boldsymbol{\alpha}')$ can be represented by its value at the average positions. Additionally, when only short time differences, with $\tau \ll 1/\Delta\nu \approx \tau_c$ are assumed,[4] also the integration over the frequency in Eq. (3.34) can be replaced by its value at the average frequency, and by that, the term for the MCF simplifies to

$$\begin{aligned} \Gamma_{\mathrm{qm}}(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2, \tau) &= \int I_b(\boldsymbol{\alpha}')\, e^{-ik_0(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\cdot\boldsymbol{\alpha}'}\, \mathrm{d}\boldsymbol{\alpha}'\, e^{-i2\pi\nu_0\tau} \\ &= \Gamma_{\mathrm{qm}}(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2, 0)\, e^{-i2\pi\nu_0\tau} \quad , \end{aligned} \tag{3.35}$$

where the subscript denotes the quasi-monochromatic approximation performed.
Being interested in the correlations at different points, separated by the baseline $\boldsymbol{B}$, but for a measurement at the same time, the MCF is evaluated for $\tau = 0$, simply yielding:

$$\Gamma_{\mathrm{qm}}(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2, 0) = \int I_b(\boldsymbol{\alpha}')\, e^{-ik_0(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\cdot\boldsymbol{\alpha}'}\, \mathrm{d}\boldsymbol{\alpha}' \quad . \tag{3.36}$$

This relation is known as the *van Cittert-Zernike theorem*, stating that the spatial coherence in the plane of observation, described by the MCF at $\tau = 0$, can be derived by a Fourier transform of the source's brightness distribution.
An expression for the degree of coherence, as defined by the correlation coefficient in Eq. (3.24), can be formulated for $\tau = 0$ by normalizing the MCF by the geometric means of the intensities $I(\boldsymbol{\xi}_i) = \Gamma_{\mathrm{qm}}(\boldsymbol{\xi}_i - \boldsymbol{\xi}_i, 0)$ at the positions of the measurements, which leads to the definition of the *complex visibility function*:

$$\begin{aligned} \mu_{\nu_0}(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2) &= \frac{\Gamma_{\mathrm{qm}}(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2, 0)}{\sqrt{I(\boldsymbol{\xi}_1) I(\boldsymbol{\xi}_2)}} \\ &= \frac{\int I_b(\boldsymbol{\alpha}')\, e^{-ik_0(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\cdot\boldsymbol{\alpha}'}\, \mathrm{d}\boldsymbol{\alpha}'}{\int I_b(\boldsymbol{\alpha}')\, \mathrm{d}\boldsymbol{\alpha}'} = \frac{\int I_b(\boldsymbol{\alpha}')\, e^{-ik_0(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\cdot\boldsymbol{\alpha}'}\, \mathrm{d}\boldsymbol{\alpha}'}{I_0} \quad . \end{aligned} \tag{3.37}$$

[4] In particular when the spectral intensity distribution is described by a rectangular function of width $\Delta\nu$, the Fourier transform is given by a sinc-function and the correlation time is often chosen to be $\tau_c = 1/\Delta\nu$.

Again, the modulus of this expression lies in the range between 0 and 1 and acts as a measurement for the spatial coherence, where full coherence is expressed by the maximum value, as e.g. by $\mu_{\nu_0} = 1$ for $\boldsymbol{\xi}_1 = \boldsymbol{\xi}_2$. However, it should be taken caution with the replacement of the integral over the brightness distribution by the constant I_0, as this simplification cannot be done in cases, when the values at the different positions $\boldsymbol{\xi}_i$ differ from each other, as it is generally the case when observing through an atmosphere, where turbulences result in random fluctuations. Still, this approach allows the conceptional separation of the effects of temporal and spatial coherence, which again manifests itself when the MSDF in the plane of observation, as given by Eq. (3.32), is rewritten by the help of the complex visibility:

$$\hat{\Gamma}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \nu) = G(\nu) I_0 \mu_\nu(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2) \quad , \tag{3.38}$$

where the temporal coherence is determined by the spectrum $G(\nu)$ and the spatial by $\mu_\nu(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)$. In analogy to the introduction of the coherence time, now the *coherence width* ω_c can be defined by the coordinate difference, or respectively baseline $\boldsymbol{B} = \boldsymbol{\xi}_1 - \boldsymbol{\xi}_2$, at which the value of the visibility, and so the fringe contrast, is significantly lowered. Just as the coherence time is affected by the spectral bandwidth, the coherence width is inversely proportional to the angular size of the observed source. Due to the two dimensional Fourier transform relation between μ_ν and the source's brightness distribution, the visibility depends on the wavelength, and generally decreases with increasing size of the source and also with longer baselines. This intuitively can be understood, when the complex visibility function (normalized MCF) is interpreted as the portion of coherent radiation within the entire detected radiation power. Then, the size of the baseline determines its ability of resolving source structures along its orientation, where an idealized point source can never be resolved and therefore would provide full coherence. Depending on the intentions of interferometric measurements, this visibility dependence on the baseline can be exploited. By using variable baseline lengths, the visibility measurements, or in particular its decrease, can be directly used to estimate stellar diameters, as it was shown firstly by the means of the star interferometer designed by *A. A. Michelson* (e.g. [Mich1921]). More than that, by testing different baseline sizes and orientations, image synthesis can be performed and through the derived brightness distribution, the source's shape can be probed and determined, as e.g. is done in studies concerning circumstellar disk structures.

The described effects of polychromatism and extended sources on the coherence of interfering beams, and consequently on the contrast of the fringe pattern, are summarized by the images in Fig. 3.5. The illustrations show the measured intensity by the detector after beam recombination in OPD space, thus as a function of the overall optical path distance, which can be converted to the phase difference between the beams at a given wavelength through the relation:

$$\mathrm{OPD} = \Delta\phi \frac{\lambda}{2\pi} \quad . \tag{3.39}$$

In the idealized case, when the source is point-like and monochromatic, the visibility and the fringe contrast are at a maximum, and the fringes show perfect constructive and destructive interference (Fig. 3.5 a)). The situation changes when the radiation exhibits a finite bandwidth, which is indicated by the colored curves, featuring slightly different wavelengths in Fig. 3.5 b).

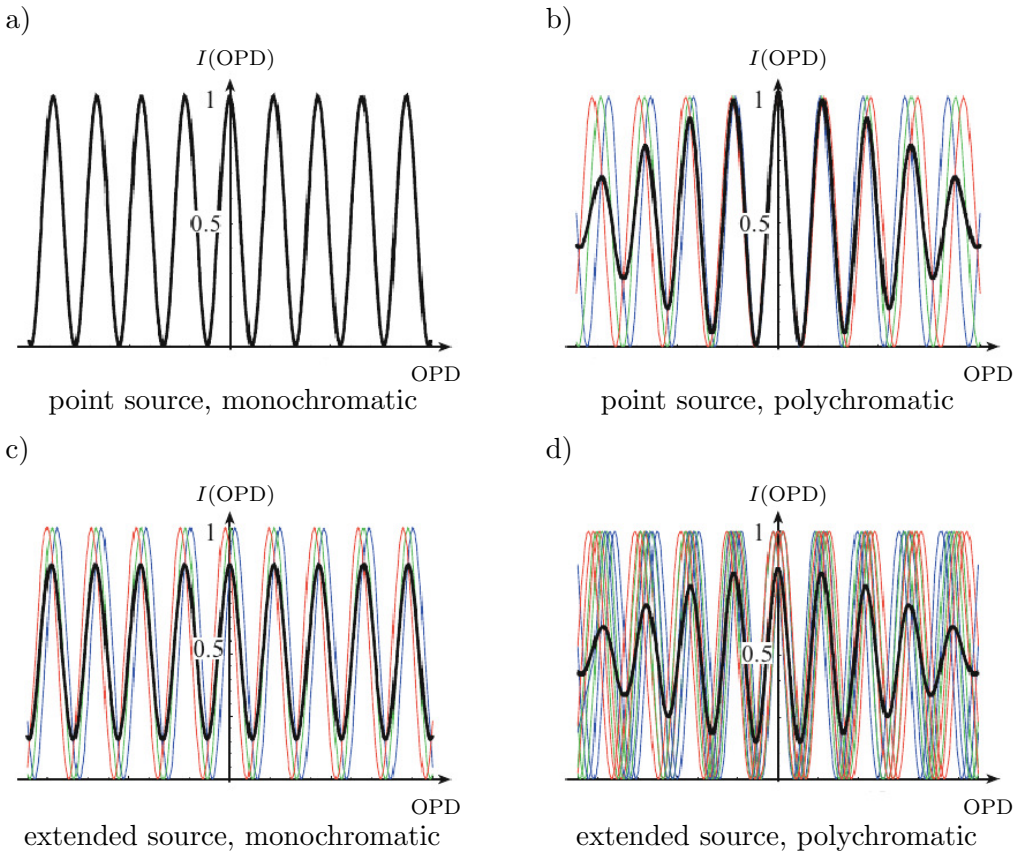


Figure 3.5. – The effects of temporal and spatial coherence on the visibility, and consequently fringe patterns, are illustrated by the detected intensity of the superposed beams as function the OPD. a): An idealized monochromatic point source implies a maximum degree of coherence and perfect fringe contrast throughout the whole OPD space. b): The influence of a finite source spectrum is characterized by the Fourier transform of the source's spectral intensity acting as an envelope to the fringe pattern, since the positions of the maxima and minima from contributions of different wavelengths (color coded) fail to overlap when moving away from the zero OPD. c): An extended source is partly resolved by the interferometer depending on the baseline length. The resulting shifts in positions of the extrema, corresponding to distinct positions on the source's surface (color coded), smear out the interference pattern and decrease the fringe contrast. d) The effects of polychromatism and extended sources are combined. (Original image from [Gli11])

Here, the Fourier transform of the source's spectrum acts as an envelope to the fringe pattern and shows decreased visibility for increasing beam phase differences, while the point of zero OPD, the central fringe, shows the full visibility. Therefore, finite spectral bandwidths result in limitations to observations with interferometers, by limiting the OPDs to a critical coherence length, at which sufficient coherence is still present so that the fringes can be detected.
The effect of partly resolved extended sources can be seen in Fig. 3.5 c), where fringes, arising

from points on the source's surface with distinct angles of observation, are shown. The shift in position on the sky results in a shift of the fringe pattern in the OPD space and the consequence is that any maxima and minima smear out, which is a direct result to the loss of visibility throughout the whole OPD or phase space. This loss of fringe contrast is also triggered by the length of the baseline, as longer aperture separations implicate higher spatial resolution of the source's brightness distribution. The combination of both effects, the temporal and spatial coherence, are illustrated in Fig. 3.5 d), where even at the zero OPD full coherence cannot be achieved.

3.2. Dual Star Interferometry and Astrometry

In the previous chapter the challenges and difficulties, regarding the search for extrasolar planets, have been discussed, and several techniques, some of which were proven to be highly successful in handling those, have been presented. The motivation behind this thesis is an insight into the design of the interferometric instrument PRIMA and the demonstration of, how the task of performing high precision astrometry by the means of dual star interferometry can be approached. For this reason the considerations throughout the following section will be closely tied to the actual design and implementation of PRIMA, which is presented in detail within chapter 4.

The accuracy of astrometric measurements that is needed for the detection of exoplanets has been estimated by Eq. (2.74) and is visualized in Fig. 2.21 in Sec. 2.3.6. It is obvious that the measurable astrometric signature is highly affected by parameters that are determined by the system's properties, such as masses of the planet and the host star and the dimension of the orbit, and also by its orientation and distance in relation to the observer. However, it can be safely stated that in order for the astrometric method to become a substantial and complementary method in this field of interest, precisions of the on sky position measurements of the order of tens of microarcseconds need to be achieved. In order to reach this goal the design of PRIMA exploits the dependence of the external OPD on the observed star's coordinates. When an appropriate internal delay is applied by the interferometer's delay lines, then the external OPD is compensated and the fringes are detectable. Still, the position of the fringes' center must be determined with high accuracy, wherefore the system operates in pupil-plane, and by which the position on sky is mapped onto the length of the applied delay, which in this manner becomes the measurable observable. Unfortunately, at the desired accuracy level it is impossible to relate this measurement to an absolute coordinate system on sky, which, besides of problems arising from turbulences in the atmosphere (see below), is the main reason why the astrometric measurement is performed in a differential way. In this concept a near star in the field of view is chosen to act as a reference and the angular separation between these two is determined by the difference of the applied compensating OPDs, the *differential OPD* (dOPD or Δd). For such a simultaneous observation, in principle two interferometric systems, using the same two telescope apertures, need to be implemented. The principle of the measurement is visualized in Fig. 3.6.

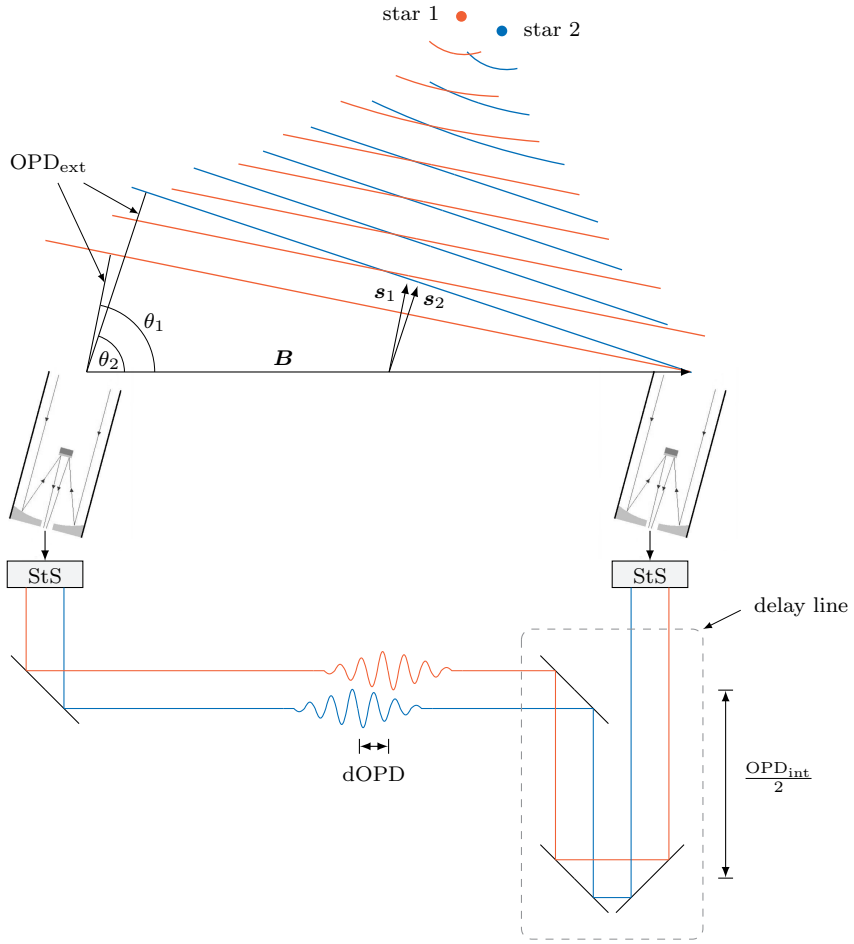


Figure 3.6. – Two stellar sources with coordinates defined by the unit vectors $\hat{\boldsymbol{s}}_1$ and $\hat{\boldsymbol{s}}_2$, separated by only a small angular distance so that both lie within the instrument's field of view, are observed simultaneously by two telescopes. Star separators (STS) are installed in order to select the light from the two distinct sources and to feed them on different paths into the system. On this way the interference patterns from both stars can be analyzed separately. Due to the sources' different positions on sky, their external OPDs, with regard to the two apertures, differ also. Here in this two dimensional illustration they are given by $\text{OPD}_{ext} = B\cos\theta$, and so depend on the distance between the apertures (modulus of the baseline $\boldsymbol{B}$) and the angle θ between the baseline vector and the direction towards the given source. In general, the projection $\boldsymbol{B}\cdot\hat{\boldsymbol{s}} = \text{OPD}_{ext}$ determines the external path difference. Consequently, the differential external OPD for the two sources is given by $\Delta\text{OPD}_{ext} = (\boldsymbol{B}\cdot\hat{\boldsymbol{s}}_2) - (\boldsymbol{B}\cdot\hat{\boldsymbol{s}}_1) = \boldsymbol{B}\cdot\Delta\boldsymbol{s}$. This differential OPD reveals itself again within the interferometer. A delay line is used to compensate the external OPD, and if the same internal path difference, OPD_{int} between the two interferometer arms, is applied to the beams of both stars, the fringes appear at different positions in the OPD space. The distance between the zero OPDs (center of the fringes) of the two distinct sources is called differential OPD (dOPD) and is directly given by the difference of the external OPDs: $\text{dOPD} = \Delta\text{OPD}_{ext} = \boldsymbol{B}\cdot\Delta\boldsymbol{s}$. Therefore, determining the dOPD inside of the system gives a direct measurement of the angular separation between the two stars, projected onto the baseline.

The two telescopes are pointed towards two distinct but close stars on the sky, whose coordinates (right ascension α and declination δ) are related to the unit vectors $\hat{\boldsymbol{s}}_1$, $\hat{\boldsymbol{s}}_2$ in their direction by

$$\hat{\boldsymbol{s}} = \begin{pmatrix} \cos\alpha\cos\delta \\ \sin\alpha\cos\delta \\ \sin\delta \end{pmatrix} \quad , \tag{3.40}$$

and which consequently are separated by the angular distance $\Delta\boldsymbol{s} = \hat{\boldsymbol{s}}_2 - \hat{\boldsymbol{s}}_1$.
The external optical path difference to each of these stars is determined by the unequal distances from the source to the apertures and therefore by the orientation of the incoming plane waves in reference to the interferometer's baseline $\boldsymbol{B}$:

$$\mathrm{OPD}_{\mathrm{ext}} = \boldsymbol{B} \cdot \hat{\boldsymbol{s}} \quad . \tag{3.41}$$

Consequently, during a simultaneous observation the star coordinates are the only variables, and the differential OPD, as determined within the interferometer from the delay line lengths and so by the positions of the fringes in the OPD space, acts as an indirect measurement of the stars' angular separation:

$$\Delta d = (\boldsymbol{B} \cdot \hat{\boldsymbol{s}}_2) - (\boldsymbol{B} \cdot \hat{\boldsymbol{s}}_1) = \boldsymbol{B} \cdot \Delta\boldsymbol{s} \quad . \tag{3.42}$$

The vector character of this equation's right hand side is crucial and indicates that two aperture interferometers are sensitive only to features of the source brightness distribution along the direction parallel to the baseline. Therefore, such a dual-feed interferometer probes the angular separation of the stars only in one dimension. However, when observing from earth, the baseline vector is not static during the time of the observation. Due to earth rotation its orientation changes and so does the measured dOPD. In order to illustrate this relation, it is often more convenient to consider only the non-vanishing and time dependent baseline components in the same plane as the observed stellar sources. This vector $\boldsymbol{B}_p$ is usually called the *projected baseline*, as it is determined from the projection of the baseline vector $\boldsymbol{B}$, which is typically given in the equatorial coordinate system or the ICRS, onto a local tangential plane, defined at the celestial sphere at the coordinates α_0, δ_0 of the star pair's midpoint and spanned by the orthonormal unit vectors

$$\boldsymbol{p}_0 = \begin{pmatrix} -\sin\alpha_0 \\ \cos\alpha_0 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{q}_0 = \begin{pmatrix} -\sin\delta_0\cos\alpha_0 \\ -\sin\delta_0\sin\alpha_0 \\ \cos\delta_0 \end{pmatrix} \quad , \tag{3.43}$$

where $\boldsymbol{p}_0$ lies in the east-west direction and points to increasing right ascension and $\boldsymbol{q}_0$ accordingly points to increasing declination in the north-south direction.
Consequently, the projected baseline and its components B_ξ and B_η along the basis vectors of this local plane are given by:

$$\boldsymbol{B}_p = \begin{pmatrix} B_\xi \\ B_\eta \end{pmatrix} = \begin{pmatrix} \boldsymbol{p}_0 \cdot \boldsymbol{B} \\ \boldsymbol{q}_0 \cdot \boldsymbol{B} \end{pmatrix} \quad . \tag{3.44}$$

The projected baseline can also be expressed in polar coordinates of the tangential plane, where

$$B_p = \sqrt{B_\xi^2 + B_\eta^2} \quad \text{and} \quad \tan\theta_{BL} = \frac{B_\xi}{B_\eta} \tag{3.45}$$

give the length and the projected baseline angle (PBA) that is specifying the angular offset of the baseline vector relative to the celestial north pole when counting over the east direction.
Naturally, the separation vector $\Delta\boldsymbol{s}$ on the sky between the two stars is given by its components $\Delta s_\alpha = \Delta s \cdot \sin\theta_s$ along the right ascension and $\Delta s_\delta = \Delta s \cdot \cos\theta_s$ along the declination respectively, where these are measured in angular distances so that the resulting and detected differential OPD can be expressed by

$$\Delta d = \boldsymbol{B}_p \cdot \Delta\boldsymbol{s} = \begin{pmatrix} B_\xi \\ B_\eta \end{pmatrix} \cdot \begin{pmatrix} \Delta s_\alpha \\ \Delta s_\delta \end{pmatrix} \quad , \tag{3.46}$$

or in polar coordinates by

$$\Delta d = B_p \cdot \Delta s \cdot \cos(\theta_{BL} - \theta_s) \quad . \tag{3.47}$$

Here θ_s denotes the position angle (PA), which is the angular offset of the secondary star to the primary, counted in the same manner relative to the north pole, as described above for the baseline angle (see also Fig. 4.22).
The time dependence of the earth bound baseline and the measurable dOPD is covered by the phase difference $\theta_{BL} - \theta_s = \Delta\theta$ and can be explicitly expressed by $\Delta\theta = \omega t + \varphi_0$, where φ_0 is an arbitrary but fixed phase offset, and $\omega = (2\pi/T)$, the circular frequency of the baseline rotation, is determined by the time period T of one revolution, namely 24 h.
By combining dOPD measurements for varying baseline orientations, different moments in time respectively, one can disentangle the ambiguity regarding the measured observable and the baseline orientation, relative to the separation on the sky, and the real angular distance between the stars can be obtained (see Sec. 4.2.2 and 4.3.3).

3.2.1. Realization of Dual Star Observations

By applying the method, which is characterized by Eq. (3.46)-(3.47), the astrometric measurement on the sky is translated into the measurement of the optical path lengths in the system. In order to be able to perform these kind of simultaneous interferometric observations, the instrument's implementation generally needs to be equipped with specific components that feature the following functionalities:

Light collection:

In stellar interferometry the two partly coherent light beams that are superposed are usually obtained by two apertures, which extract light from the incoming waves at distinct positions defined by the telescope positions. The distance between them and their orientation, the baseline of the interferometer, heavily affect the resolution of the separation measurement. According to Eq. (3.42), the objects' separation on sky is linearly mapped onto the dOPD so that with a given accuracy of the delay measurements the resolution increases with longer baseline lengths. However, the increase of the baseline length must be treated with care, as in principle the source itself will be resolved progressively, which in turn can lead to a loss of spatial coherence and consequently to a reduced fringe contrast.

Object selection from field of view:

In addition to the naturally limited field of view of interferometers in Michelson configuration, for an astrometric measurement it must be ensured that the detected fringes

originate from the sources of interest and not from objects in their vicinity. Besides this, fringe patterns of each of the two stars are detected separately and simultaneously, which means that although entering through the same telescopes, the light paths of the distinct sources must be separated and injected into different channels somewhere within the instrument. This is usually achieved by a special periscope layout, as a whole device called the *star separator* (STS).

Fringe detection:

When recombining the beams co-axially, the intensity measurement of the detector probes the interference pattern only at one defined position, compared to the fringe center, also called the *white-light fringe*. In order to determine the dOPD between the two observed objects, the exact position of the fringe centers needs to be known. This information in principle can be obtained from a fast scan within the coherence time, during which the fringe patterns are not smeared out. Alternatively, this information can also be derived from an estimation of the phase and modulus of the visibility function at the OPD applied at the moment of the measurement. The *fringe sensor untis* (FSUs) of PRIMA perform this estimation by a four channel, spatially modulated, intensity measurement, called the *ABCD method*, as will be described below in Sec. 4.1.3.1.

Fringe tracking:

Generally, the accuracy of the astrometric measurement is increased and the limiting brightness of the observed stars is decreased by higher integration times, which in turn are limited by fringe motions, thus the coherence time. Due to this, the integration times can only be increased by stabilizing the fringe pattern. This can be achieved, if the position of the central fringe is not only determined, but if this information is fed back in real-time into a controlling mechanism that drives the delay lines and consequently controls the OPD, which continuously can be adjusted to reduce the residual fringe motions. By reducing motions at lower frequencies, such a fringe tracking algorithm consequently acts as a filter to the power spectrum of the fringe center position variations, and to OPD fluctuations.

Path length compensation:

The delay lines in an interferometer are supposed to compensate the external path differences in order to detect fringes at all. When the zero OPD positions of two sources are about to be determined at the same time in order to derive the dOPD, it is clear that also the internal delays for both objects need to be controlled separately. This in principle demands for two sets of delay lines, which can be applied appropriately onto the the different light paths.

Monitoring the light path:

With the realization of long baseline interferometers the light paths of the beams inside of the instrument become considerably long (up to several hundreds of meters). When it comes to the measurement, particularly with the goal of precise astrometry, each variations limit the effective accuracy that can be reached by the instrument. Aiming at an astrometric accuracy of $\delta s = 10\,\mu$as, with a baseline length of 100 m, the limiting uncertainty δd in the determination of the dOPD can be estimated to $\delta d \approx \delta s \cdot B \approx 4.8\,$nm.

Usually, laser metrology systems are incorporated for sensing path length variations at these levels. Ideally, a laser beam follows the same path in the instrument as the light of the stellar sources, and, by measuring occurring phase fluctuations of the laser interference signal, any beam path variations can be detected and corrected for.

3.2.2. Atmospheric Disturbances and Phase Referencing

In order to achieve accuracies of the order of 10–100 μas it has been shown that the light path needs to be controlled down to variations of around $\sim 5\,\mathrm{nm}$. Although full stability of the whole interferometer to this precision is rather illusory, in principle a laser metrology system provides the means to monitor intrinsic motions and to correct for them. Unfortunately, any disturbances to the propagating waves, occurring outside of the system, before the beams arrive at the telescopes, are not registered. In particular turbulences in the atmosphere result in wave front distortions on short time scales, which manifest themselves in form of phase shifts and which lead to a decrease of fringe contrast and coherence time.
The basic approach to understand the structure of waves propagating through the atmosphere relies on the description of the statistics of the random motions as described by the *Kolmogorov turbulence* model. Well structured derivation of the principles can be found in the literature (e.g. [Qui99], [Gli11]), whereas here the main results and consequences to the project are summarized.
Following the underlying concept, the atmosphere, with typical wind speeds of several m/s, is considered a turbulent fluid, where the largest scale turbulences build at the *outer scale* L_0, and kinetic energy is transferred successively down to smaller eddies. This breaks down at the *inner scale* l_0, where viscous dissipation processes become significant. The range between these two entities, called the *inertial range*, is still a matter of ongoing discussions. While the inner range is assumed to be of the order of a few millimeters, the estimation of the outer scale, somewhere between a few tens of meters and a kilometer, is quite vague. Although attempts to determine L_0 have been performed (e.g. [Bus95]), uncertainties in the models' parameters and their interaction during the observations prevent definite interpretations and precise estimations.
For the analytic examination it is assumed that the turbulences are statistically stationary and isotropic. Given that, the kinetic energy spectrum, as a function of the spatial wave number $k \propto 2\pi/l$, can be derived from the energy transfer between eddies of different sizes l. As in the context of this discussion, one is mainly interested in the perturbations, as induced on plane waves passing through the turbulent cells of the atmosphere, and therefore the fluctuations regarding the refractive index n must be considered. The turbulent eddies can be regarded as mixing structures of varying temperature in pressure equilibrium and hence with different densities and consequently different indices of refraction. The spatial distribution of the fluctuations of the refractive index obey the same statistics as the energy spectrum and is usually expressed by the power spectral density (power per volume element $\mathrm{d}\boldsymbol{k}$):

$$\Phi_n(\boldsymbol{k}) = 0.033\, C_n^2\, k^{-11/3} \quad . \tag{3.48}$$

This *Kolmogorov spectrum* is homogenous and isotropic, where the structure constant C_n^2 describes the strength of the refractive index fluctuations. The Kolmogorov power law with the $\propto k^{-11/3}$ dependency is valid only within the inertial range, as beyond the outer scale, respectively at low spactial frequencies, the spectrum is assumed to saturate, since the turbulences become more and more uncorrelated. The so called *von Kármán spectrum* (see [Gli11]) serves as an adjustment to the Kolmogorov spectrum in order to account for this and exhibits reduced power outside the inertial range.

With the statistics of the refractive index at hand, the consequences for initially plane wave fronts, travelling through the rapidly fluctuating regions within the atmosphere, can be derived. Due to the distribution of the refractive index $n(\boldsymbol{x}, z)$, where $\boldsymbol{x}$ is the horizontal coordinate vector and the altitude is given by z, a phase shift of

$$\varphi_h(\boldsymbol{x}) = \frac{2\pi}{\lambda} \int_h^{h+\delta h} n(\boldsymbol{x}, z)\, \mathrm{d}z \tag{3.49}$$

is introduced to a wave that propagates through a thin layer of δh at the height h.
As usual, the consequences on the coherence of the wave are evaluated by the computation of a correlation function and in this case are determined by the statistics of the turbulences themselves, which through Eq. (3.49) directly translate onto the phase distribution of the wave, after passing the turbulent thin layer. This correlation function is described by

$$\Gamma_{\varphi,h}(\boldsymbol{x}) = e^{-\frac{1}{2} D_{\varphi,h}(\boldsymbol{x})} \quad , \tag{3.50}$$

where the variance of the phase shifts $D_{\varphi,h}(\boldsymbol{x}) =< [\varphi_h(\boldsymbol{x}') - \varphi_h(\boldsymbol{x}' - \boldsymbol{x})]^2 >$ defines the spatial *phase structure function* of the turbulences of the layer in question.
Temporal correlations have not been taken into account here, since it is assumed that the lifetime of the turbulent cells is generally significantly longer than the time scales, on which the whole structure is blown away by the wind, so that it can be regarded as not evolving. This concept is referred to as Taylor's *frozen turbulence* hypothesis, which justifies the simplification

$$\begin{aligned} D_{\varphi,h}(\boldsymbol{x}, \tau) &= < [\varphi_h(\boldsymbol{x}', t) - \varphi_h(\boldsymbol{x}' - \boldsymbol{x}, t + \tau)]^2 > \\ &= D_{\varphi,h}(\boldsymbol{x} - \boldsymbol{v}\tau, 0) \quad , \end{aligned} \tag{3.51}$$

of the full, temporal and spatial, phase structure function to its spatial characteristics. Therefore, the specification of the instantaneous $D_{\varphi,h}(\boldsymbol{x})$ and the wind velocity $\boldsymbol{v}$, at which the eddies are moving across the line of sight, describe the statistics completely.
Finally, when the correlation scale of the fluctuations is smaller than δh, the phase structure function for a horizontal plane wave passing through the particular layer i at an altitude h is dependent only on the magnitude $x = |\boldsymbol{x}|$ of the distance between two points and is given by ([Gli11]):

$$D_{\varphi,h_i}(\boldsymbol{x}) = 2.91 \left(\frac{2\pi}{\lambda}\right)^2 \delta h_i C_{n_i}^2 x^{5/3} \quad . \tag{3.52}$$

Applying this to the correlation function Eq. (3.50) of the wave, one can see that the perturbations by the turbulent layer correspond to a finite width of the correlation width and therefore

result in a decrease of coherence.

In order to arrive at meaningful results, one needs to take the whole atmosphere into account, which can be regarded as a composition of many, statistically independent turbulent, layers, passed through by the optical waves before reaching the ground. When a continuous distribution of such is applied, the phase structure function, giving the mean-square difference of the phases at two points, separated by $\boldsymbol{x}$, can be derived to:

$$D_{\varphi,h}(\boldsymbol{x}) = 6.88 \left(\frac{x}{r_0}\right)^{5/3} \quad . \tag{3.53}$$

The correlation length r_0, also called the *Fried parameter*, is defined by the integral over the turbulence profile

$$r_0 := \left(0.423 \left(\frac{2\pi}{\lambda}\right)^2 (\cos\zeta)^{-1} \int C_n^2(h)\,\mathrm{d}h\right)^{-3/5} \quad , \tag{3.54}$$

and specifies the diameter of the circle, over which the phase variance is about $1\,\mathrm{rad}^2$ and cannot be neglected, where ζ is the angular zenith distance of the observed source.
Consequently by definition, the resolution that can be achieved by seeing-limited observations, when the atmospheric turbulences are characterized by r_0, is equal to the diffraction-limited resolution of apertures with diameters of the same size. Therefore, the Fried parameter is often referred to as the size of the turbulence cells itself, although strictly speaking the Kolmogorov spectrum resembles a continuous spatial distribution within the inertial range.

As atmospheric seeing is site and weather dependent, it can significantly change between different times of observations. Still, an approximation of r_0 at sites with generally good conditions can be estimated to ([Lan03]):

$$r_0 \simeq 0.1 \left(\frac{\lambda}{0.5\,\mu\mathrm{m}}\right)^{6/5} \mathrm{m} \quad . \tag{3.55}$$

The dependency on the wavelength plays an important role on the instrument design as waves of longer wavelength are less disturbed by the atmosphere. For the infrared K-band ($\simeq 2.2\,\mu$m) one gets Fried parameters of around 0.6 m.

In differential astrometry, or also when using adaptive optics systems, two sources, which are separated by small angles θ, are observed simultaneously. Since one of those targets is used as a reference to correct for the atmospheric disturbances, the correlations of the phase fluctuations between the two objects become important. Due to the slight displacement on the sky, the light from the two objects travels through different turbulence cells and hence the waves are also perturbed differently.
The *isoplanatic angle* θ_0 is introduced as a limitation to on sky separations, for which the variance of the relative phases becomes significant ($1\,\mathrm{rad}^2$).

It can be derived by simplifying the model so that the turbulence profile is replaced by a single layer at a mean effective height H, which is defined by:

$$H := \left(\frac{\int C_n^2(h) h^{5/3} \, \mathrm{d}h}{\int C_n^2(h) \, \mathrm{d}h} \right)^{3/5} \quad . \tag{3.56}$$

When the spatial displacement is parametrized by the separation angle by the means of the $x = \theta H$, the phase variance $\langle \sigma_\theta^2 \rangle$ between the two positions can be estimated from the phase structure function (Eq. (3.53)).
One can use

$$\theta_0 = 0.314 \, (\cos \zeta) \, r_0 / H \tag{3.57}$$

in order to relate the isoplanatic angle to the Fried parameter, and by this one finally obtains the result

$$\langle \sigma_\theta^2 \rangle = \left(\frac{\theta}{\theta_0} \right)^{5/3} \quad , \tag{3.58}$$

where the formal definition of the isoplanatic angle is given by:

$$\theta_0 := 6.88^{-3/5} \left(0.423 \left(\frac{2\pi}{\lambda} \right)^2 (\sec \zeta)^{8/3} \int C_n^2(h) h^{5/3} \, \mathrm{d}h \right)^{-3/5} \quad . \tag{3.59}$$

As indicated by Eq. (3.57), the isoplanatic angle increases with wavelength ($\propto \lambda^{6/5}$) and strongly depends on the turbulence profile, as substituted by the effective altitude H. When observing in K-band and H within the range of 1–7 km, the isoplanatic angle can be estimated to values between 6 and 40 microarcseconds.

The magnitude of θ_0 is a stern constraint on target selection, if real-time controlling mechanisms are used to correct for atmospheric perturbations. One of the two observed objects, typically the brighter source, serves as a reference for the measurement of the phase fluctuations. If the separation between the two targets lies within the isoplanatic patch, the disturbances of the incoming waves can be regarded as sufficiently correlated and therefore, after detecting them on the reference object, corrections can be instantaneously applied to the second target.

During the interferometric observations, which PRIMA is designed for, the phase of the incoming waves, and hence its fluctuations, are directly measured by the determination of the fringe position in the OPD space. Consequently, if the phase shifts of the two observed stars are correlated, the absolute fringe positions respond to the variations, but the influence on the measurement of the differential OPD is highly minimized.
Furthermore, since phase referencing reduces the fringe motion of the second target, it permits longer integration times at the used detector, which in turn provides higher sensitivity and allows fringe detection on fainter targets.
The integration time, and therefore the sensitivity, of an interferometric measurement are limited by the coherence time. This is why the temporal evolution of the atmospheric turbulences needs to be considered. Following the basic assumption of Taylor's frozen turbulence hypothesis

that while being blown in front of the light paths the evolution of the turbulences is negligible, the structure of the turbulences can be regarded as a static screen, moving with the wind velocity v. Therefore, the phase at a given point $\boldsymbol{x}$ and time $t+\tau$ is equivalently given by the phase at point $\boldsymbol{x}-\boldsymbol{v}\tau$ at time t (compare to Eq. (3.51)). In this way the results from the analysis of the spatial characteristics can be translated directly to conclusions on temporal statistics.
In this manner, the atmospheric coherence time τ_{atm}, in analogy to the Fried parameter, gives the time interval, during which the root-mean-square difference of the measured phases reaches 1 rad. It is related to r_0 by

$$\tau_{\mathrm{atm}} = 0.314\,\frac{r_0}{V} \quad , \tag{3.60}$$

where V stands for the effective wind speed, to which variable wind velocities in the different layers contribute, and which is given by:

$$V := \left(\frac{\int C_n^2(h) v^{5/3}\,\mathrm{d}h}{\int C_n^2(h)\,\mathrm{d}h}\right)^{3/5} \quad . \tag{3.61}$$

Both characteristic quantities, the coherence time and the Fried parameter impose constraints on the sensitivity during a measurement, which naturally scales with the number of detected photons. Since only photons, detected within the coherence time and within a circle of diameter r_0 can contribute to a sufficiently coherent measurement so that the fringe contrast is high enough, the sensitivity generally follows $r_0^2 \cdot \tau_{\mathrm{atm}} \propto \lambda^{18/5}$. With increasing wavelength both limiting quantities are increased also, and respectively the measurement is less affected by atmospheric perturbations. In addition to that, phase referencing in combination with fringe tracking in dual-feed interferometers drives the sensitivity limit to fainter objects.

Chapter 4

Astrometry with PRIMA

This main chapter of the thesis is completely devoted to PRIMA, its implementation and its functionality regarding astrometry. It is divided into three sections.
Within the first the instrument's hardware and technical layout will be described.
In order to achieve the anticipated level of accuracy, the raw data gathered during the observations needs to be thoroughly processed into the final scientific products. This usually will be accomplished by the partially automated data reduction software, which has been developed by our exoplanetary group at the *Landessternwarte Königstuhl* in Heidelberg, Germany, and which will be described within the second part.
Following that, the last section engages the current state of the instrument. In this context the emphasis is laid on the description of the procedure, by which the astrometric observables can be derived from the dual-feed interferometric measurements, and at which point of the chain of the complex measurement systematic errors might be introduced.
For this purpose, in the course of this work, the instrument's commissioning data, which has been collected from test observations during the commissioning runs over the last years, has been analyzed.
The results from that shall on one hand demonstrate the system's principle functionality regarding fringe tracking and phase referencing, and on the other hand provide first estimations on the momentarily achievable astrometric accuracies, also in terms of stability and systematic uncertainties.

4.1. PRIMA and its Components

Since the year 2000, in collaboration between ESO, the European Southern Observatory, and the ESPRI consortium, consisting of participators from the Observatoire de Genève, the Max Planck Institute for Astronomy and the Landessternwarte Königstuhl (LSW), both in Heidelberg, a lot of devotion and time has been put into the development and construction of PRIMA, the facility for Phase-Referenced Imaging and Microarcsecond Astrometry.
Using this instrument, which to a given level compensates atmospheric disturbances during interferometric observations, limitations to the brightness of the sources are decreased, and moreover its implementation also provides phase-referenced aperture synthesis imaging.
However, its dual-feed design was highly motivated by the prospect of conducting a search for extrasolar planets by the means of an astrometric instrument that would be capable of achieving precisions at the level of tens of microarcseconds.
In 2010, all subsystems, which have been developed particularly for the dual-feed mode, eventually have been incorporated into, but also considerably upgraded the VLTI infrastructure at

the Cerro Paranal in Chile.
The astrometric method with PRIMA relies on differential measurements, during which fringes of two targets are detected simultaneously, and the difference between the internal delays that need to be applied to the light paths of these two, the differential delay, is being determined. Since atmospheric turbulences induce motions of the fringes' centers on short time scales of several up to tens milliseconds in the infrared, these need to be monitored at sufficiently high frequencies and corrected for by the delay lines.
As long as both of the observed objects are positioned on the sky within the isoplanatic patch, the atmospheric effects can be appropriately considered the same, and the control signals from the bright source can be equally applied to the second target. Consequently, this method results in an extended synthetic coherence time so that fainter sources can be observed.
The four additional components, which provide dual-star interferometry at the VLTI, are introduced in the following, while their position in the PRIMA facility and their influence on the beam paths is depicted in Fig. 4.1.
The dual-feed mode provided by PRIMA can be utilized by both kind of telescopes, operated at the site on Paranal, namely by the stationary 8.2 m Unit telescopes (UT) and by the movable 1.8 m Auxiliary telescopes (AT), which can be repositioned on different stations. Thus, in general the employment of the ATs gives access to measurements with variable baselines and expands the parameter space during observations, which is propitious not only for astrometry, but for phase-referenced imaging as well.

4.1.1. Star Separators

Due to the limited interferometric field of view of the Michelson configuration, the system is equipped with star separators (STSs), which are capable of selecting objects from the telescopes' FOV and to feed them separately into different injection ports of the VLTI system. *TNO Science & Industry* has developed and built such devices for both types of telescopes, the UTs and the ATs. As triggered by the predefined requirements to meet the scientific goals, these opto-mechanical mirror systems are mounted underneath the telescopes and were designed to allow to pick up two subfields from the initial 2 arcminute wide field of view and thereon to lead the two channels into the delay lines. Furthermore, the implementation of the STSs, as depicted in Fig. 4.2, provides means of compensating field rotations, as well as beam tip-tilt stabilization and pupil alignment, both with actuation frequencies of 50 Hz ([Del08]).
The field segmentation is obtained by the placement of a rooftop mirror (M10)[1] at the Coudè focus of the telescope. When observing with the ATs, a reflective K-prism assembly acts as a field de-rotator device, which is installed directly above the STS and compensates mainly for diurnal field motions, as those due to earth rotation.
This unit ensures that the two objects of interest are re-imaged at either opposite sides of the rooftop edge.[2]
Also, switching between the NORMAL and SWAPPED states of observation, during which the

[1] The numbering of the mirrors starts with the telescope's primary mirror M1 and incrementally follows the beam path down to the detector.
[2] Within the star separators for the UTs, the de-rotating mechanism is directly incorporated on the M10 mirror.

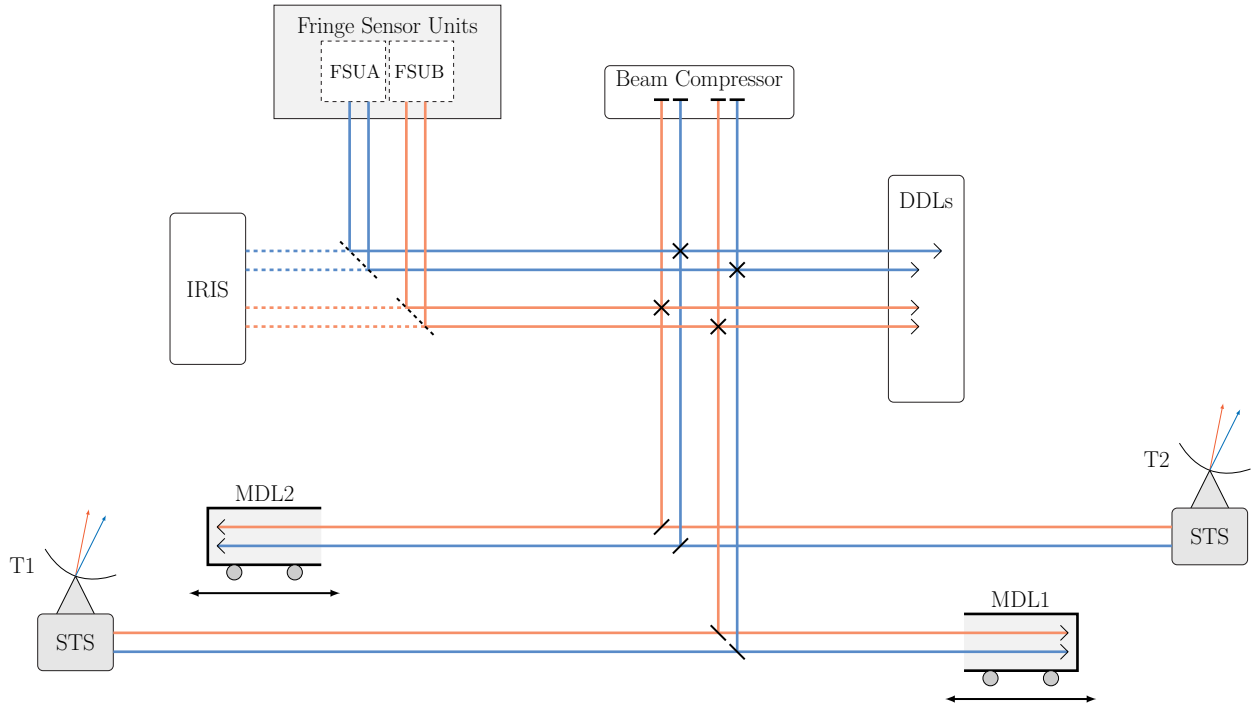


Figure 4.1. – Simplified depiction of light paths within the PRIMA system during a dual-feed observation with the ATs in *NORMAL* mode. Light from both stars, the usually brighter primary (PS) and the secondary star (SS), is selected from the telescope's field of view by the star separators (STS), situated in the Coudè focii and is injected into distinct entry ports of the delay lines in the VLTI tunnel. The main delay lines are used to compensate the external OPD for the PS, while the differential delay lines (DDLs) provide the additional differential delay to one part of the beam of the SS so that fringes of both stars are centered and can be detected after beam recombination in the fringe sensor units. For symmetry reasons each of the four beams are sent to the DDLs, although usually three of them are kept at fixed positions and only one is used for OPD compensation during the fringe tracking process. In this configuration light from the PS is combined in FSUB, and accordingly FSUA is fed by the beams from the secondary. By the means of the STSs, this situation can be reversed. Consequently, in the so called *SWAPPED* mode the light paths of the stars are exchanged in the system, which not only reduces differential effects to the light beams due to air dispersion effects in the light ducts and tunnels, but is a necessary step during an astrometric observation in order to determine the metrology zero point, since it monitors the beam paths not in an absolute, but only in a differential manner (see Sec. 4.1.4). Four laser metrology beams are injected into the systems at the fringe sensor units and follow the star's light paths up to the star separators, where they are reflected. Operating in the H-band, the guiding camera IRIS, which is placed near the fringe sensor units and is fed by a dichroic mirror, extracting the H-band portion from the star beams, provides a fast tip-tilt sensing in order to stabilize the beams in the entrance to the FSUs.

injection of the light beams of the two observed stellar sources is swapped, is carried out by a field rotation of 180° of the de-rotators.

As the pupil coincides with the piezo-actuated M11, rotation of it around two orthogonal axes allows pointing on any of the objects in the Coudè focus on M10. Therefore, M11 acts as a *field selection mirror* (FSM) and is operated in a closed loop with the tip-tilt sensor IRIS for the purpose of fast beam stabilization, which is inevitable for keeping the photon leakage at the fringe sensor units at a low level in order to maintain an acceptable sensitivity.
Moreover, the STSs mark the PRIMET metrology end points in the system. The laser beams at the wavelength of 1319 nm, injected at the fringe sensor units and travelling the beam path in opposite direction to the star light, are being transmitted at the dichroic mirror M9 and by that reach their retroreflectors (see also Sec. 4.1.4).
As M12 re-images the focus on the M14 mirrors, which are actuated by a tip-tilt stage, they, together with the PRIMET pupil tracking system, provide a mean of monitoring and correcting the lateral and axial alignment of the output with the input pupil in order to minimize vignetting effects and to assure a proper overlap for the metrology system ([Nij08]).
Since the M14 mirror's curvature directly affects the pupil's longitudinal position, it is planned to control it at real-time by mirrors of variable curvature (VCM), but is not implemented at present ([Sah13]).

By this described set-up, the STSs ensure a proper pointing and tracking of the science objects and represent the first of the key components that provide PRIMA's dual-feed mode of simultaneous interferometric observations.

4.1.2. Differential Delay Lines

As during ordinary PRIMA operation the main delay lines of the VLTI are adjusted to compensate for the optical path delay of the usually brighter primary source, which is used as the phase reference, the fringes of the secondary star can only be observed at the second detector, if an additional delay, the differential delay, accounting for the objects' separation on the sky, is applied. Compared to the main delay lines that can provide common path delays to both stars with nanometer accuracy at lengths over 200 meters, the differential delay needs to introduce only a short delay. A rough estimate of the maximum dOPD assuming a 200 m baseline and a star separation of 120 as gives the requirement for the additional delay device to about $\Delta d_{\max} \approx 116\,\mathrm{mm}$. For this purpose the differential delay lines have been designed and built by the ESPRI consortium in collaboration with ESO and have been installed on Paranal in 2008. The design of the DDL units, along with results from tests in the laboratory, are presented in [Pep08]. From the operational point of view, the DDLs require to be able to perform fast OPD scans for fringe finding or sampling the fringe envelope within the full range of applied optical delays. They also need to provide means for tracking blind trajectories up to velocities of $200\,\mu$m/s in order to compensate for the changing of the delay due to field rotation and to provide an active tracking mode, during which perturbations due to atmospheric piston or vibrations in the system are corrected for in a controlled closed loop with the detectors, at a bandwidth above 200 Hz.
The implementation of the DDLs, as illustrated in Fig. 4.3, approaches these challenges by the help of a Cassegrain-type retroreflector telescope (cat's eye), consisting of three mirrors so that

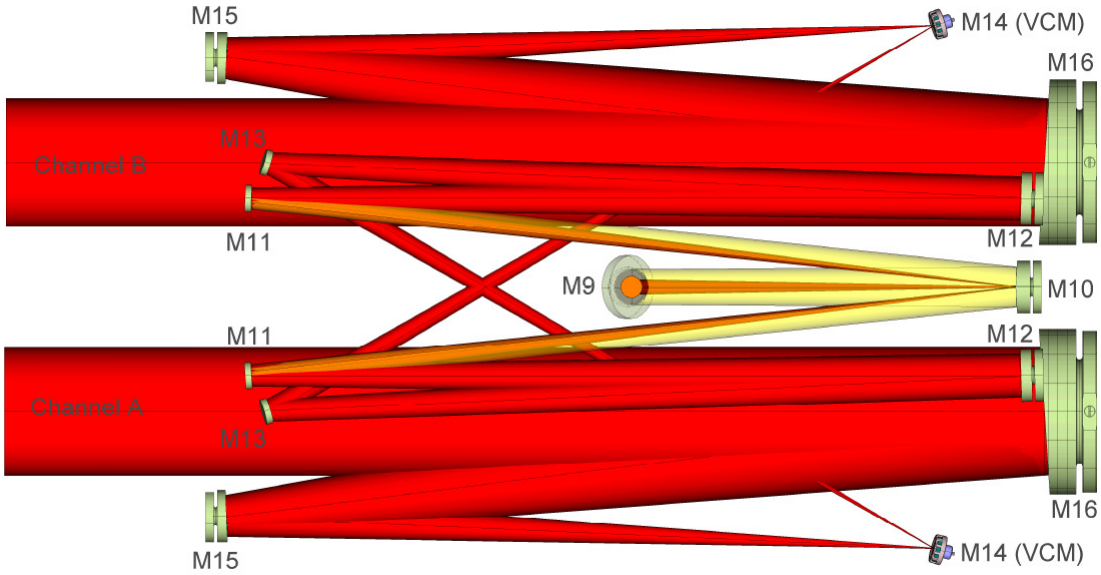


Figure 4.2. – The light collected by the telescopes enters the star separator (STS) devices at the dichroic mirror M9 and the initial field of view is divided into two subsections at the rooftop mirror M10, which is situated in the telescope's Coudè focus. Any object, re-imaged onto the rooftop, can be pointed to by the mirrors M11, which act as the field selectors and provide tip-tilt corrections. The output channels of the STSs are two separated beams of 80 mm in diameter, which are injected then into the main delay lines in the VLTI tunnel. (Credit: [Nij08])

after five reflections (M1→M2→M3→M2→M1)[3] the output beam is perfectly opposed and separated by 120 mm from the input. The cat's eye optical system is mounted on an adjustment plate, which acts as the interface to the stiff main linear translation stage. This conceptionally first stage of actuation, driven by a DC motor, provides a long stroke of 69 mm and allows to follow OPD variations with a velocity of up to 1 mm/s. The fine motion at a high bandwidth of nearly 400 Hz, sufficient for effective fringe tracking, is achieved by a three-axis piezo actuator, connected to the spherical mirror M3. It allows fast stroke adjustments over around 10 μm and provides a tool for corrections of pupil displacement in both, longitudinal and lateral directions. In this manner flatness errors of the translation stage can be accounted for, and the input and output beams are kept co-linear within an uncertainty below 1.5 as in tip-tilt angle.
In order to keep the longitudinal position of the stellar pupils in the beam combination environment at the same position, as at which it would reside without the installation of the DDLs, the input pupil's re-imagination by the DDLs is adjusted by the choice of the M3 curvatures, which consequently is the only varying characteristic between the four different DDL devices, which are installed for symmetry reasons so that each of the four science beams passes through a DDL device.

[3]The mirror numbering inside of the DDL is out of sequence from the numeration, referring to the full system, and must not be confused with the telescope's mirrors.

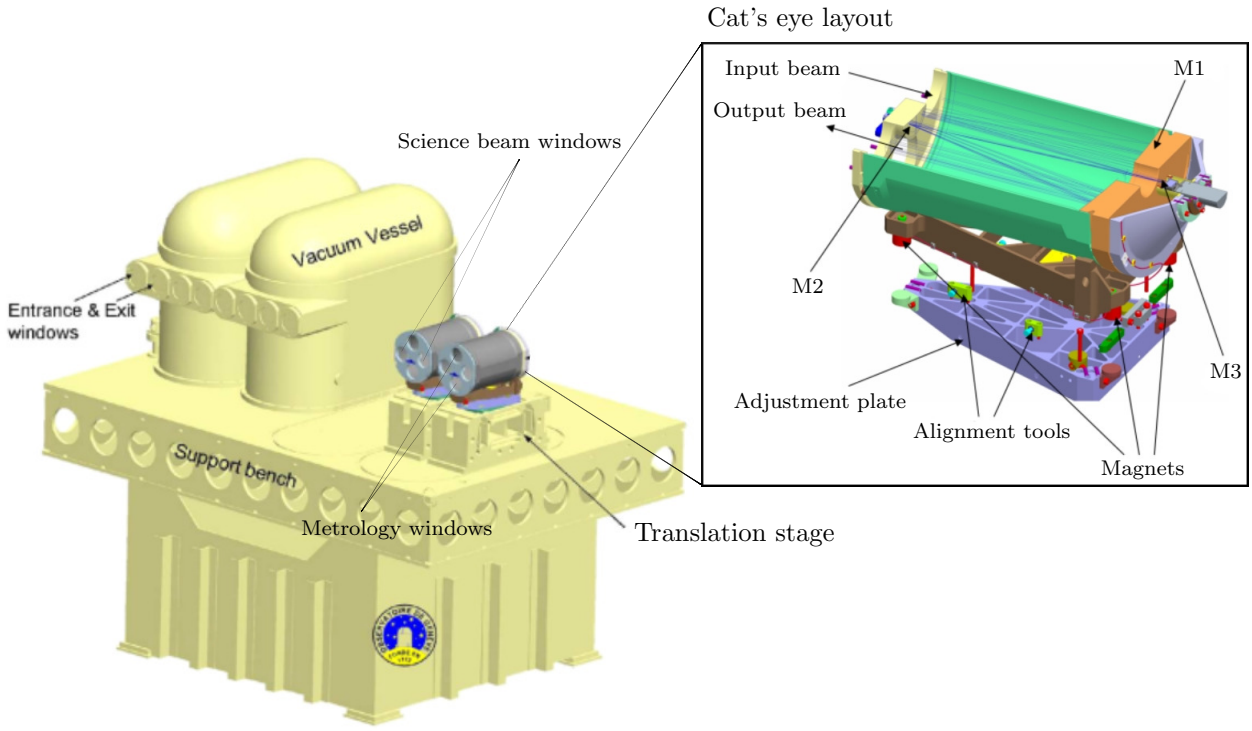


Figure 4.3. – The illustration shows the implementation of the DDLs. Placed within vacuum vessesls, the cat's eye mirror assembly is mounted on a translation stage, which provides a rough displacement of the whole reflecting device over the full stroke of more than 60 mm. The cat's eye itself consists of three mirrors, the parabolic M1, the hyperbolic M2, and the spherical M3 so that after five reflections of the input beam, entering through one of the four windows, the output beam is perfectly directed in the opposite direction and horizontally shifted by 120 mm. The fine and fast motion control of the DDLs is provided by a three-axis piezoelectric actuator, located at the M3. (Original image from [MPIA10])

During typical observations, three of the DDL units will be kept at a fixed position, while only one is used to adjust the applied dOPD.
For the purpose of precisely determining the position of the retroreflector, an internal laser metrology system has been implemented. The laser beams, which for phase measurement feed a mach-Zender type interferometer, are launched and collected outside of the cat's eye and, just as the science beams, enter and exit the retroreflector via the four entrance windows at the front of the tube. The windows, aligned with the horizontal plane are designated for the stellar beam, while the windows in the vertical plane are reserved for the metrology.
Consequently, the laser beams pass the DDLs perpendicular to the science beams, but for symmetry reasons experience the same optical delay, which then is determined with an accuracy of 2.47 nm at up to 1 kHz, and can be measured for each DDL independently, as each of these is equipped with its own metrology receiver.
In order to minimize the effects of differential dispersion on the OPD, the DDLs are placed and

operated in non-cryogenic vacuum vessels. The vessels' entrance and exit windows are aligned with the cat's eye windows so that the science beams can pass after having been downsized in diameter from 80 mm to 18 mm by beam compressors along the lightpaths (comp. Fig. 4.1). Each vessel can harbor up to two DDL devices and together they are mounted on the optical table.

Altogether, by the means of the DDLs, the obervational requirements are met in the sense that differential optical path delays to one interferometric arm of one of the observed sources can be applied up to a maximum of at about $2 \cdot 69\,\text{mm} = 138\,\text{mm}$ and controlled with a resolution of $\approx 2.5\,\text{nm}$, which is sufficient for the astrometric purpose.
Furthermore, the fine actuator allows fast corrections to pupil displacements in three directions (tip, tilt, longitudinal) and, operated in a closed loop with the fringe sensor units, allows effective fringe tracking. Due to the implementation of the DDLs, the light beams experience five additional reflections within the retroreflector and two transmissions at the entrance and exit of the vacuum vessels, but as tests in laboratory environment and after implementation at the VLTI suggest, additional wavefront errors are limited to the RMS value of $16 \pm 2\,\text{nm}$ and the throughput of the optical system is estimated a total transmittance of 0.86 in the K-band and 0.89 for the PRIMET laser wavelength of $1.319\,\mu\text{m}$ ([Pep08], [Sah13]).

4.1.3. Fringe Sensor Units

The key component of the PRIMA system are the detectors, by which after beam combination the fringes are detected, and beyond that, when operated in a closed loop with the delay lines, their position can be locked so that fringe tracking becomes possible.
After their design and fabrication by *Thales Alenia Space Italy*, the fringe sensor units (FSUs) have been tested in a laboratory environment since 2006 ([Abu06]). Eventually, during summer 2008 the FSUs have been successfully integrated into the infrastructure of the VLTI. The requirements on the FSU design arise from the desired modes of operation, by which the scientific goals are approached. In order to provide the access to differential astrometry, which demands for simultaneous interferometric observations of two sources within a narrow angle on the sky, two identical FSU devices have been installed. Fast fluctuating OPD variations due to atmospheric fluctuations and disturbances along the beam paths within the instrument require a quick response of the control system and a solid phase measurement. For these reasons it has been decided to operate the FSU at fairly long wavelengths in the near infrared K-band between $1.95\,\mu\text{m}$ and $2.45\,\mu\text{m}$, and for the fringe sensing a spatial filtering scheme (see Sec. 4.1.3.1) has been implemented along with the beam combination so that the units deliver real-time estimates of the fringe phase and the group delay at sampling rates of the detector up to 2 kHz. This is achieved by the unique opto-mechanical design, which will be outlined in the following, and which is thoroughly described by [Sah09].

With the high bandwidth and throughput of the PRIMA FSUs, the VLTI is finally equipped with its second generation fringe tracking system, which, besides offering the dual-feed mode for astrometry, also can be combined with other VLTI instruments, such as MIDI and AM-

BER. In this mode, one of the FSUs serves as the tracker, stabilizing the fringes, whereas the visibility measurement is performed by the secondary instrument. As [Mül10] have shown for the combination of MIDI with the PRIMA FSU, the limitations in the source brightness for phase referenced imaging observations can be improved significantly.

The first successful fringe detection by these subsystems has been achieved on September 3rd, 2008 ([vBelle08]). During an observation with two of the AT telescopes, separated by a baseline of 32 m, fringes of the source HD 19349, a star with the brightness of $K = 0.44$ and $V = 5.27$ in the visible, have been recovered from a delay scan through the zero OPD position. The signal, as recorded by FSUA during this event, is presented in Fig. 4.4.

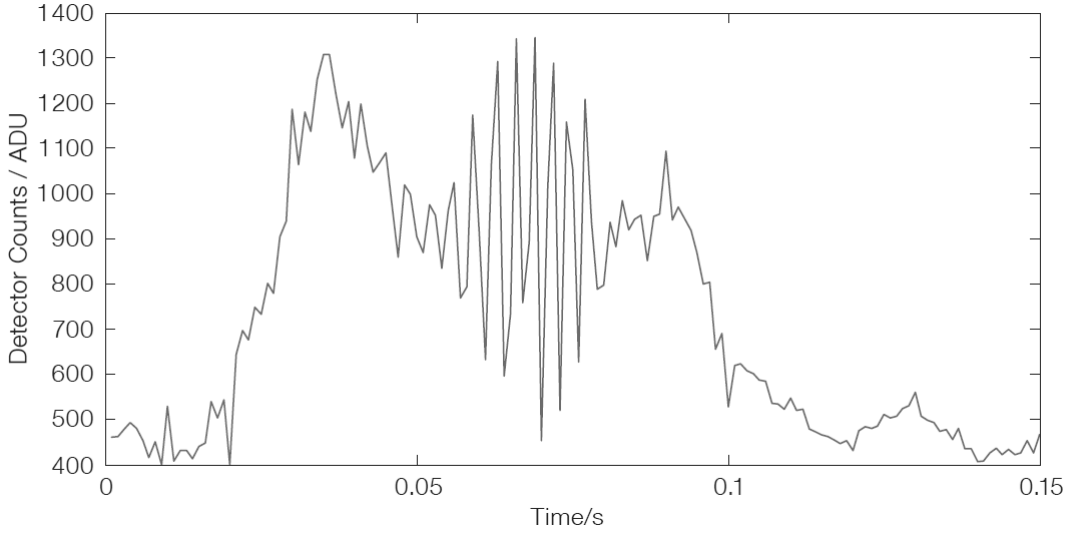


Figure 4.4. – The time modulated signal of the first fringe detection by FSUA on September 3rd, 2008. The fringe center at $t \approx 0.07$ s and its envelope are clearly visible from the data recorded during an OPD scan, where the delay lines moved through the zero OPD position, at which the external path difference is compensated by the equal internal in the beam train of the VLTI. (Credit: [vBelle08])

4.1.3.1. Fringe Sensing

As discussed before, fringe stabilizing and phase referencing can significantly increase the coherence time during an observation. In order to achieve this, the fringe tracking system needs to control the internal delay, imposed by the instrument's delay lines, so that during the measurement the integration over time is performed at a definite position on the fringe. Low-frequency drifts, mainly due to earth rotation, as well as the more challenging high-frequency fluctuations due to variations within the atmospheric layers, manifest themselves as variations in the position of the fringe center, the white light fringe, and tend to smear out the visibility function during the measurement if not compensated. In order to approach this difficulty the fringe

tracking mechanism relies on a fringe sensor that is able to estimate the fringe position by determining the phase of the complex visibility from the fringe pattern at sampling rates higher than the rates, at which the most concerning perturbations occur.

Phases and the ABCD Method

A standard method to measure the phase of the signal, which describes the position, at which the fringe is sampled with regard to its center, is called the *ABCD*-method and was suggested by [Wya1975] and [Sha1977]. This procedure basically is founded on intensity measurements of the combined beams at two or more different points, with a known separation in phase, on the fringe pattern around the white light fringe.
If the measurements are obtained at four distinct points, separated by $\lambda/4$ in the OPD space, or respectively by $\pi/2$ in phase space, which are usually denoted as A, B, C and D, and to be more precise the phase relations between them, omitting any noise contributions for the sake of simplicity, are described by

$$\begin{aligned} I_A &= I_{\text{tot}} \left(\frac{1}{4} + \frac{|\mu_\lambda|}{2\,C} \cos\varphi_\lambda \right) \\ I_B &= I_{\text{tot}} \left(\frac{1}{4} + \frac{|\mu_\lambda|}{2\,C} \sin\varphi_\lambda \right) \\ I_C &= I_{\text{tot}} \left(\frac{1}{4} - \frac{|\mu_\lambda|}{2\,C} \cos\varphi_\lambda \right) \\ I_D &= I_{\text{tot}} \left(\frac{1}{4} - \frac{|\mu_\lambda|}{2\,C} \sin\varphi_\lambda \right) \quad , \end{aligned} \tag{4.1}$$

then the modulus of the complex visibility μ_λ, as introduced in Sec. 3.1.1, and its phase φ_λ can be estimated from:

$$|\mu_\lambda| = C\,\frac{\sqrt{(I_A - I_C)^2 + (I_B - I_D)^2}}{I_{\text{tot}}} \quad , \tag{4.2}$$

$$\varphi_\lambda = \frac{2\pi}{\lambda} D = \tan^{-1}\frac{\text{Im}\,[\mu_\lambda]}{\text{Re}\,[\mu_\lambda]} = \tan^{-1}\left(\frac{I_B - I_D}{I_A - I_C}\right) \quad . \tag{4.3}$$

Here the subscripts indicate the corresponding points (A, B, C and D) on the fringe pattern, whereas the sum of the individual intensity measurements is denoted by I_{tot}, and the sine and cosine terms depict the phase relations between the four bins.
The value of the constant C depends on the specific implementation of this detection method, while the offset of the OPD at the instant of the measurement in reference to the center of the fringe, or the zero OPD, is given by D, which shall stress the relationship between the measured fringe phase and the position in the OPD space, at which the measurement is taken.
In accordance with these relations, the complex visibility itself comes along by

$$\mu_\lambda = \frac{C}{I_{\text{tot}}}\left[(I_A - I_C) + i\,(I_B - I_D)\right] = X + i\,Y \quad , \tag{4.4}$$

with the quadratures X, Y defining its real and imaginary parts.
However, the estimation of the visibility, as indicated by Eq. (4.2) is strongly affected by the

temporal coherence of the signal. With increased OPD distance to the white light fringe, the contrast of the measurement decreases and the estimation of the visibility tends to be underestimated, if not calibrated properly. This is why alternatively the Fourier spectrum of the whole pattern is widely used to derive a proper estimation of the complex visibility.
A similar difficulty is encountered concerning the phase determination, as the result is confined to the range of $[0, 2\pi[$ (or respectively $[-\pi, \pi[$), resulting in phase wrapping, if those limits are exceeded. This strengthens the importance of a secondary indicator that the measurement is taken at the center of the fringe pattern.

There are two common ways, by which this kind of measurement can be implemented into the interferometric observation. In the case of the *temporal phase modulation*, the path length in one of the interferometer arms, and by this the OPD, is modulated so that the whole phase shift range of 2π is covered. In doing so, the optical path distance and consequently the phase become time dependent entities. In order to make usage of the ABCD method, each scan is subdivided into four equal sections, which ideally are swept through at the same time intervals of $T/4$, when T is the time it takes to sweep the whole range. Since the number of the detected photons during the subintervals is proportional to the integrated intensities $I_{A,\dots,D}$ over the time of the subscans, the equations (4.1)–(4.3) can be utilized for the desired estimations, when the value $\pi/\sqrt{2}$ is applied to the constant C.
Fast OPD oscillations at rates of comparable order to the modulation frequency can severely affect the quality of the phase measurement by this technique, as phase relations between the four bins (A,B,C,D) are not stable and vary in an unpredictable way during the time of single scans. Consequently, in order to freeze these variabilities the scans need to be performed at high frequencies, which limits the integration time per single bin and demands for fast detectors.

In contrast to this approach, higher measurement bandwidths can typically be achieved by fringe sensors with an implemented *spatial phase modulation* scheme. In this case the OPD is not modulated to cover the whole phase shift range along one fringe, but static optical elements are used to provide instantaneous and simultaneous intensity measurements at several points with a known phase shift. This can be achieved, as is the case for the PRIMA FSUs, by the implementation of polarization dependent intensity measurements (see details below) so that the four measurements are performed, while for each of them the phase differences between the combined beams are different. If the values of these phase differences have the values 0, $\pi/2$, π and $3\pi/2$, the phase separations between the four signals (A,B,C,D) equal $\pi/2$, and so the provided formulas, in this approach with $C = 2$, yield the estimations for the phase and the visibility.
Unfortunately, although the principle of implementing this scheme is fairly simple, it is difficult to obtain stable phase shifts precisely at the desired nominal values.
As the extraction of the four signals depends on the separation of the beam's polarization states, any polarizing effects along the light path introduce errors to the phase shifts, which eventually affect the phase recovery. Consequently, these sources of error need to be taken into account (see Sec. 4.3.1.1).
Furthermore, calibration routines also need to cover any differential influences between the four signals. Such might be due to differences and instabilities of the injection into the four chan-

nels, detector typical characteristics as pixel responses, but also result from unequal effective wavelengths and differential transmission in the system.

The Group Delay

Due to phase wrapping, OPDs, exceeding the length of the wave, cannot be determined unambiguously, which is problematic as the tracker should stabilize the measurement at the central fringe position. By quantifying the increasing divergence of the wavelength dependent phase estimates, which are equal at the zero OPD, but disperse when moving away, the group delay (GD) is an entity, which qualifies perfectly as a complementary indicator for the position, at which the measurement is taken. It can be defined by

$$\mathrm{GD} = \frac{c}{2\pi}\frac{\mathrm{d}\varphi(\nu)}{\mathrm{d}\nu} \quad , \tag{4.5}$$

where $\varphi(\nu)$ are the frequency dependent wrapped phases, or equivalently phases with respect to the position of the measurement, as determined by the ABCD method for different spectral channels. This relation is illustrated in Fig. 4.5 a). The fringe envelope is centered at the position of zero OPD, at which also the group delay vanishes. The group delay is linear and nonambiguous for OPDs up to the beam's coherence length $\lambda^2/\Delta\lambda$ ([Gli11]) and will be kept by the tracker's control system at its zero position to ensure phase referencing at the central fringe. In principle, once the measurement is centered, sole phase estimates should provide accurate tracking, but additional control of the GD allows the detection of and the correction for any sudden phase jumps exceeding the value of 2π.

Just as the beam's intensity spectrum can be derived by the Fourier transform of the sampled fringe signal, the fringe pattern can be deduced reversely from a channeled spectrum, if the signal is sampled at different wavelengths. As the zero GD position marks the position of the white light fringe, where the fringe contrast is at its maximum, it can be determined from the center of the envelope of the fringe pattern.

Longitudinal Dispersion

The main delay lines of the VLTI are situated in the tunnel and are not evacuated. This becomes a problem when considering the situation that the internal path difference is introduced by the delays in order to compensate the external, also called geometric, OPD that is a result to a path difference above the atmosphere with the refractive index $n = 1$. The first effect becomes apparent when the path length L, within a delay line in vacuum, is compared to the corresponding air-filled situation, where the optical path is then given by $n(\nu)L$, as due to the decrease in wavelength a higher number of phase cycles fits into the spatial length along the path of beam propagation. Due to that, the position of the zero OPD for all off-zenith sources ($\mathrm{OPD}_{\mathrm{ext}} = L \neq 0$), for a given wavelength, shifts by the amount of:

$$\Delta\mathrm{OPD}_0(\lambda) = L\,(n(\nu) - 1) \quad . \tag{4.6}$$

The situation is complicated, as variations of the refractive index $n(\nu)$ along the detector's bandpass result in the wavelength dependency of the zero OPD, which now does not define the

position of the white fringe any more. If the central wavelength of the bandpass is considered as the reference for the OPD_0 position, then its distance to the center of the fringe envelope, and equivalently to zero group delay, can be derived from Eq. (4.5) when substituting the phase by its corresponding wavelength dependent OPD position and is given by:

$$\mathrm{OPD}_{\mathrm{GD}} = L\nu \frac{\mathrm{d}n(\nu)}{\mathrm{d}\nu} \quad . \tag{4.7}$$

This offset needs to be corrected for in the context of astrometric measurements when tracking at GD = 0. As the position of zero group delay depends on the path length through the delay line, which compensates the external OPD ($\mathrm{OPD}_{\mathrm{ext}} = \boldsymbol{B} \cdot \hat{\boldsymbol{s}}$), its offset position becomes naturally a dynamical quantity in the course of an observation due to earth rotation.

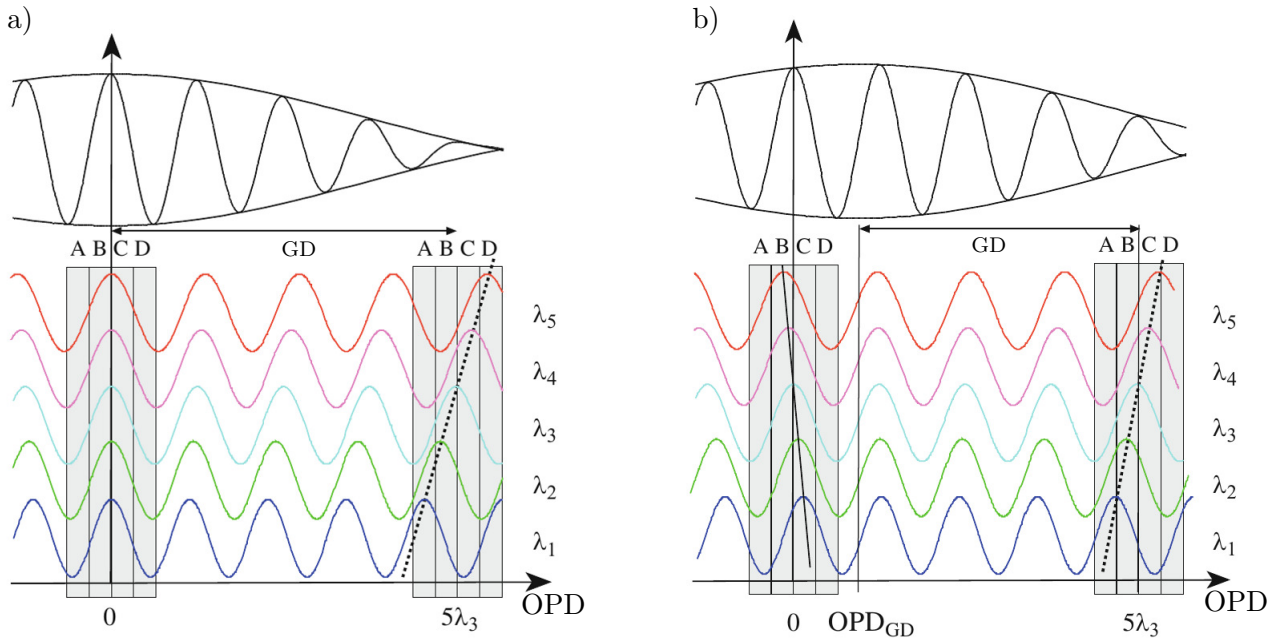


Figure 4.5. – The images illustrate the relations in OPD space between the position of the zero OPD, the group delay and the phases of the ABCD signals at five different wavelengths λ_i over the detector's bandpass for both situations, with delay lines in vacuum (*left*) and affected by longitudinal dispersion (*right*). The grey areas define two positions of measurements, at OPD = 0 and OPD = $5\lambda_3$, whereas the ABCD signals are separated by $\pi/2$ in phase. According to Eq. (4.5), the value of the group delay is determined by the change in phase over the spectral channels ($\frac{\mathrm{d}\varphi(\nu)}{\mathrm{d}\nu}$) and gives the distance GD between the location of the measurement and the position $\mathrm{OPD}_{\mathrm{GD}}$ of zero GD. In the case of dispersion, the zero OPD becomes wavelength dependent due to $\mathrm{d}n(\nu)/\,\mathrm{d}\nu \neq 0$, and the center of the fringe envelope is shifted along with zero group delay. (Credit: [Gli11])

4.1.3.2. Implementation of the FSU

In order to track the position of the fringe pattern during observations, a fast fringe sensor is indispensable. The design of the PRIMA FSUs, thoroughly laid out in [Sah09] and shortly summarized here, makes use of spatial phase modulation for real-time phase measurements, which makes it unique among fringe sensors in stellar interferometry.

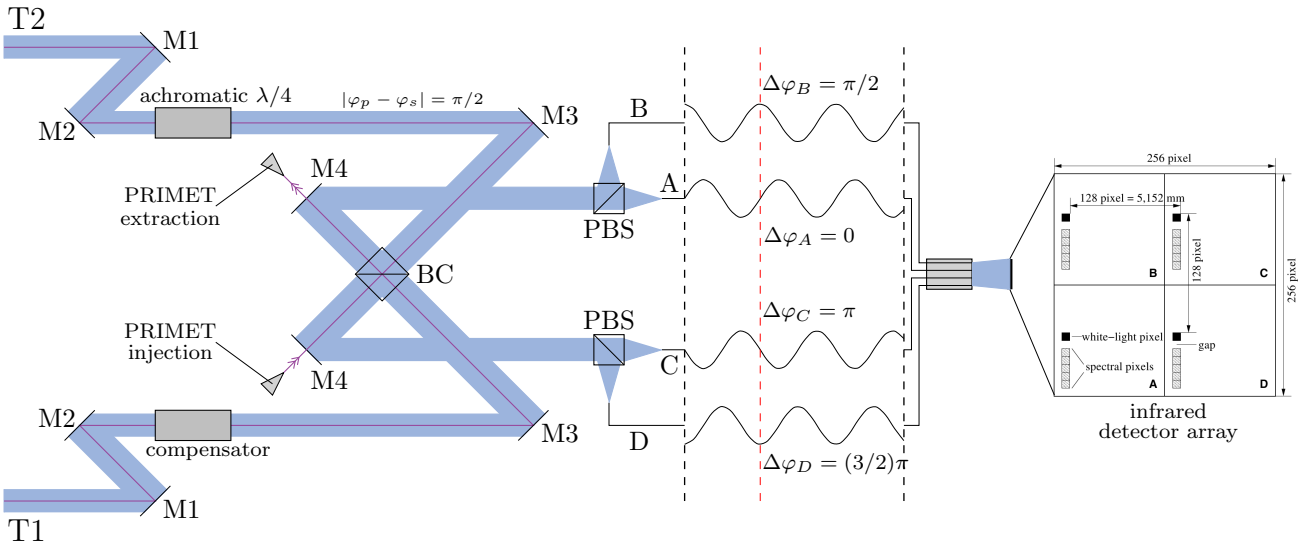


Figure 4.6. – The optical layout, which is identical for both FSUs and which provides the spatial filtering for the ABCD method. The goal is to provide four signals, for which the phase shifts between the two combined beams equal the desired values of 0, $\pi/2$, π and $3\pi/2$. This is achieved by a polarization dependent encoding design. After the beams from the two telescopes (T1, T2) enter the device, an achromatic $\lambda/4$ retarder introduces an initial phase difference of $\pi/2$ between the p and s polarization states of one of the beams, while the other travels through a silica block to compensate the OPD introduced by the phase shifter. The beam combination process at the 50/50 beam splitter cube (BC) results in a further phase shift of $\pi/2$ between the transmitted and the reflected beams so that after this process the resulting phase shift is polarization dependent. After extraction by polarizing beam splitters (PBS), the four signals, separated in phase by the indicated values ($\Delta\varphi_{A,\ldots,D}$), are injected into monomode fibers and then imaged onto the four different quadrants of the detector array. The beams are spectrally dispersed over 5 spectral channels. (Image of detector originally from [Sah09])

As depicted in Fig. 4.6, the spatial filtering is performed after beam combination by polarizing beam combiners, at which the two orthogonal polarizaton states (p, s) are separated, consequently providing four signals with different phase shifts between the combined beams, namely the values $\Delta\varphi_{A,\ldots,D} = 0$, $\pi/2$, π, $3\pi/2$, from which at any time of the measurement the phase of the fringe and the offset from the fringe center in OPD space can be derived. Integrated in a closed servo loop with a real-time control system, those estimations serve as a feedback, and

the delay lines positions are adjusted to keep the measurement as close as possible to the fringe center. In dual-feed mode one of the identical FSUs provides fringe sensing on the primary source, feeding the optical path delay controller (OPDC), which drives the main delay line, whereas the second FSU measures the fringes on the secondary, and the differential delay line is actuated by the differential optical path delay controller (dOPDC). The control signals are obtained at the detector's sampling rate up to 2 kHz.

In order to keep the background noise at a low level, the four beams, hereafter denoted as A, B, C, and D are focused onto the cores of single-mode fibers for spatial filtering, which conduct the light to a cryostat, evacuated and cooled with liquid nitrogen, and where the cold optics, featuring a low resolution spectrograph, of both FSUs are situated. After passing a prism assembly for dispersion, the initial K-band signals are finally focused by a camera doublet lens onto the four quadrants of the infrared detector arrays, where the beams are chromatically dispersed over five spectral channels.[4]
An example of the spectral pixel response for each quadrant and pixel is depicted in Fig. 4.7, and the corresponding effective wavelengths, as derived from the same calibration, are given in Tab. 4.1. The five spectral channels are spread over the entire detector bandpass and are used to derive the group delay. Slight chromatic dispersion is also observed between the four quadrants, which is mostly problematic for calibration purposes when determining the actual phase shifts between them (see also Sec. 4.3.1.1).

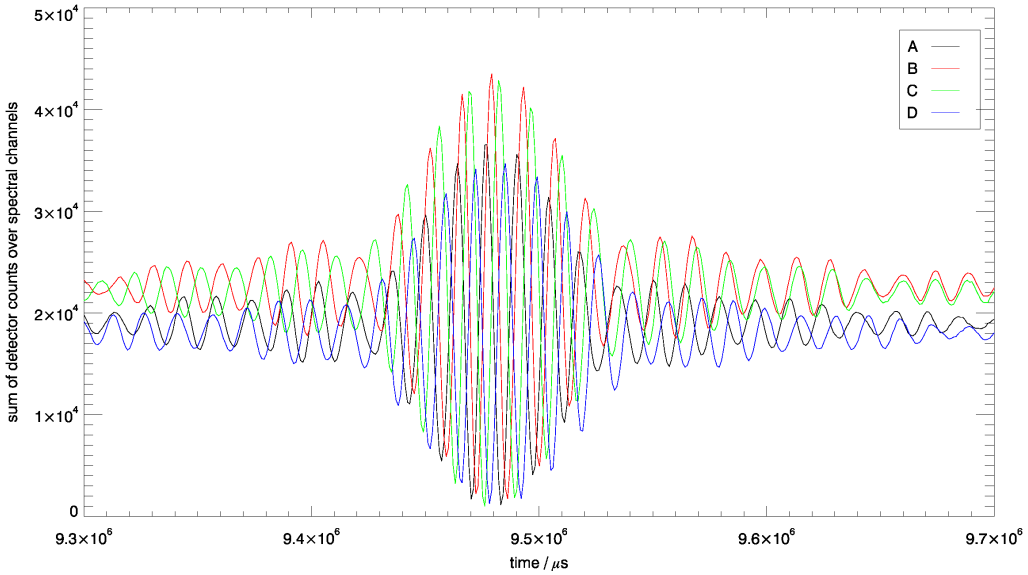
Table 4.1. – The effective wavelengths, given here in μm for each FSU, channel and quadrant, as well as averages over ABCD, have been derived as weighted averages from the FSU responses (comp. Fig. 4.7 and see app. C for file reference) in laboratory conditions.

	$\lambda_{\text{eff.}}[\mu\text{m}]$ for FSUA					$\lambda_{\text{eff.}}[\mu\text{m}]$ for FSUB				
	A	B	C	D	av.	A	B	C	D	av.
ch. 1	2.0084	2.0224	2.0076	2.0123	2.0127	1.9983	2.0205	2.0009	2.0122	2.0079
ch. 2	2.1244	2.1309	2.1252	2.1175	2.1245	2.1192	2.1342	2.1231	2.1227	2.1248
ch. 3	2.2441	2.2519	2.2494	2.2452	2.2476	2.2410	2.2551	2.2540	2.2537	2.2510
ch. 4	2.3623	2.3656	2.3748	2.3654	2.3670	2.3629	2.3743	2.3731	2.3645	2.3688
ch. 5	2.4428	2.4462	2.4677	2.4541	2.4527	2.4566	2.4676	2.4535	2.4537	2.4578

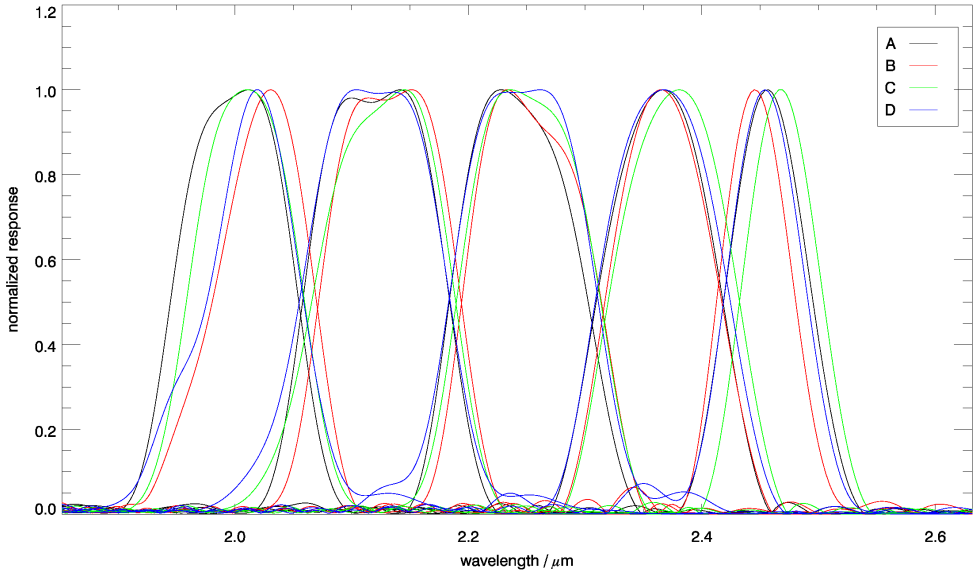
The polarization dependent phase shifts themselves are obtained by the design of the beam combination process. The 50/50 beam splitter cube, as the central element, superimposes the telescope beams and naturally introduces a phase difference of $\pi/2$ between the reflected and

[4]Initially, it was foreseen by the design that by adjustment of the prism assembly some fraction of the light would be dispersed only over one single pixel to provide the white-light signal. This concept has been dismissed after post installation test evaluations have shown an eminent loss of light due to difficulties of the prism assembly alignment and degrading coatings.

transmitted beams. As now an achromatic polarization retarder, realized by a K-shaped silica prism leading to three internal total reflections, is placed into the light path of one of the telescope beams, an additional phase shift is introduced between the p and s polarization states. In order to equalize the OPD induced within the achromatic $\lambda/4$ retarder, the beam from the other telescope travels through a compensating silica block.
This layout as a whole provides the desired phase differences for the ABCD signals.
The beam injection, monitored by the guiding camera IRIS, is controlled and stabilized by the device's alignment system. The active FSU M2 mirrors allow piezo-driven tip-tilt corrections and are mounted on a motorized linear stage for OPD attunement, when the FSUs are used with the laboratory calibration source *MARCEL* (Multi-beam Alignment, Reference and Calibration Emitter for the VLTI Laboratory). This artificial light source provides a black-body spectrum at a default temperature of 700 ℃ as a reference for the VLTI instruments ([ESO09]). The calibration unit additionally provides integrated corner-cube retroreflectors for the metrology beams so that also during calibration procedures in the laboratory PRIMET can be used for the monitoring of the OPD.
The PRIMA laser metrology beams (see Sec. 4.1.4) for both FSUs are injected and extracted by the means of circular dichroic patches of $\varnothing = 2.5\,\mathrm{mm}$ on the FSU M4 mirrors. From the point of injection, the laser beams, 1 mm in diameter, pass the beam combiner and travel in direction to the telescopes until reflected back at their endpoints within the star separators.
Since the main delay lines at the VLTI are filled with air, as discussed above, longitudinal dispersion along the detector bandpass affect the fringe detection by shifting the zero group delay position and the center of the visibility envelope away from the zero OPD and the fringe contrast is reduced ([Tang1990]). Clearly this limits the capability of observations towards fainter targets and the overall reduction of the visibility can be partly compensated by introducing correcting dispersive glass plates into the beam paths ([Lev1996]). For this purpose, motorized stages at the entrances of the FSUs are placed so that plates of various thickness can be introduced into the beam paths.



(a) FSUB raw white-light signals during OPD scan



(b) Response of FSUB in laboratory

Figure 4.7. – Sample of spectral responses as obtained during a laboratory calibration procedure in November 2011. For this purpose, the OPD was temporally modulated around the central fringe. The synthetic white-light signals, the sum of the counts over the spectral channels, are shown in (a). As indicated by Eq. (3.30), the Fourier transform of the sampled fringe pattern yields the spectral intensity distribution. Here the moduli of the Fourier transforms of each of the 20 spectral pixel channels have been corrected for the spectrum of the calibration source, have been normalized to their maximum values and are plotted against the wavelength (b). The distribution of the five spectral channels over the whole band is clearly visible. The four quadrants (ABCD) are color coded and also show a slight chromatic dispersion, particularly at both ends of the overall FSU bandpass.

4.1.3.3. FSU On-Line Operation

The fringe tracking units are designed to provide information on the OPD position with regard to the fringe center and to feed the OPD controller in a closed loop so that the delay line positions can be adjusted in order to keep the observation at the fringe. The delay offsets, processed by the controller, and the controlling algorithms depend on real-time determinations of the fringe phase, the group delay, more precisely the distance to the zero GD, and an estimate on the signal-to-noise ratio (SNR). As a proper estimation of these quantities is crucial for the fringe tracking performance of the system, a careful calibration routine is applied prior to the actual measurements, during which several parameters for correction are derived and stored in the FSU database, from where they can be accessed by the real-time machines. This procedure, as well as the actually implemented methods, described in detail in [Sah09], are summarized in the following.
A typical observation with PRIMA can generally be divided into three stages, namely the calibration in the laboratory, the night calibration on the sky, and the actual science measurements.

- During the **laboratory calibration**, which can be carried out at day time, the artificial light source MARCEL is used to feed both FSUs. The procedure derives standard photometric correction parameters (dark and flat) for each of the 20 FSU pixels (five spectral channels per individual quadrants ABCD). For the estimation of the dark-values time averages are determined from detector signals during measurements with off-pointed beams, while flat-corrections are acquired from averages of uncorrelated integrated signals, obtained by consecutively pointing off one of the two individual beams. These parameters given and stored, further raw detector pixel read-outs $I_{i,\Gamma}(t)$ are processed to the flat-normalized and corrected signals

$$S_{i,\Gamma}(t) = \frac{I_{i,\Gamma}(t) - D_{i,\Gamma}(t)}{F_{i,\Gamma}(t) - 2D_{i,\Gamma}(t)} \quad , \tag{4.8}$$

  where $D_{i,\Gamma}$ and $F_{i,\Gamma}$ stand for the dark and flat values respectively, while the individual pixels are denoted by the subscripts $i = 1, 2, 3, 4, 5$ and $\Gamma = A, B, C, D$.
  As the real-time estimates of phase and GD rely on the signal's wavelength and phase shifts between the four quadrant bins, those parameters are determined from the Fourier transforms of data recorded during several fringe scans, with the metrology for OPD reference. The effective wavelengths $\lambda_{i,\Gamma}$ in each pixel are then computed from the barycenter of the transform modulus (see Fig. 4.7 b) and Tab. 4.1), and the resolution of the computation naturally depends on the sampling, in this case on the combination of the sampling frequency and the delay velocity during the scan. The relative phase shifts between the ABCD signals, in particular their deviations $\psi_{i,\Gamma}$ from the nominal values, are derived from cross-correlations of the fringes.
  Furthermore, for a posterior SNR estimation, a value for the *visibility noise* v_0 is derived by applying the standard formula (Eq. (4.2)) to the synthesized and incoherent white light signals.

- The calibration parameters, which are obtained and stored during the day calibration, can serve as primary auxiliary quantities for the real-time estimates during fringe sensing

on sky, but do not represent the system's characteristics in science mode, when the light is collected by the telescopes instead of being injected in the lab. For this reason, a **night calibration** procedure is executed, at its best on each new source that is observed.
In the course of this operation, the dark and flat values from the calibration in the laboratory are replaced by the corresponding sky-dark and sky-flat values, obtained during the calibration on the sky. Whenever possible, also the fringe-based parameters (wavelengths, phase shift errors, visibility noise) are substituted by the corresponding ones that are determined from the *VLTI response* during OPD scans over the fringes of the stellar sources in question, where one of the FSUs serves as a fringe stabilizer, while the other records the signals during the OPD modulation.
By these routines the FSU database is provided with calibration data that accounts for the instrument's spectral transmission, as well as the individual sources' spectral features.

- During the **science observation** itself, the raw pixel counts can now be corrected in terms of photometry for dark and sky background, and the fringe sensing is driven by the real-time estimates of the phase, the group delay and the SNR. The, during the observations time dependent, fringe phases $\varphi_i(t)$ are computed by the use of the standard approach, as given by Eq. 4.3, after a correction due to the phase shift errors $\psi_{i,\Gamma}$ has been applied (see Sec. 4.3.1.1).
The group delay, in reference to the OPD of the zero GD, is computed from the estimation of the OPD position of the maximum of the complex visibility modulus, which is equivalent to the center of the fringe envelope. Consequently, the offset to the fringe center is estimated from a parabolic fit to the maximum value of the modulus of the discrete Fourier transform of the visibility quadratures

$$\mathcal{F}(x,t) = \sum_{k=1}^{5} \left[X_k(t) + i\, Y_k(t)\right]\, e^{-i2\pi x/\lambda_k} \quad , \tag{4.9}$$

where the sum is taken over the spectral channels $k = 1, 2, 3, 4, 5$, and the x samples an evenly spaced range of $\pm 12\,\mu$m in OPD space with reference to the OPD(t) position at the time of the measurement.
While real-time phases and GD estimates serve to provide offset signals for the delay lines, the controlling routine needs an indicator on the signal coherence. The signal-to-noise ratio is estimated for this purpose by normalizing the modulus of the visibility estimate by the visibility noise:

$$\mathrm{SNR}(t) = \frac{\sqrt{X_0^2(t) + Y_0^2(t)}}{v_0} \quad . \tag{4.10}$$

The subscript indicates that the values are determined from the synthesized white light signal, which is the sum over the spectral pixels.
Depending on the estimated SNR, the OPD controller operates in three different states, where transitions between them are determined by three predefined SNR threshold values (*det*, *close* and *open*). A signal-to-noise value lower than *det* indicates that the FSU is recording aside of the fringe envelope and the OPD controller is in *SEARCH* mode, performing a triangular OPD modulation around the fringe position, as predicted by the applied OPD model. Fringes are regarded as detected, whenever during the searching

procedure the SNR value increases above *det*. In these cases, the OPD controller switches to the *TRACK* mode, in which the fringe tracking loop is closed and OPD offsets are sent to the delay lines to keep the observation stabilized. The controller responds at a higher rate ($\approx 10\,\mathrm{Hz}$) to the phase estimates, in order to compensate atmospheric variabilities, and at a significantly slower rate to the GD estimates for the detection of phase jumps exceeding 2π, in order to sustain tracking on the central fringe.
This state is maintained until the SNR decreases below the threshold *open*, whereupon the controller again switches to *IDLE*. The system stays in this mode with an open tracking loop and the delay line following the predicted trajectory, for either a user-defined limiting interval of time (20 ms by default), after which a fringe search is performed again, or until SNR > *close* is detected and the loop is closed again.
It should be clear that the fringe tracking performance strongly depends on the predefined threshold values and therefore benefits significantly by the instrument user's experience with the system.

4.1.4. Metrology

The PRIMA laser metrology system *PRIMET* has been developed under collaboration of ESO and the Institute of Microtechnology of Neuchâtel (IMT). It has been designed to provide the differential optical path distance between the two PRIMA beam paths, as introduced by the DDLs, with a required accuracy of at least $\sim 5\,\mathrm{nm}$, and is based on super-heterodyne laser interferometry phase measurements ([Lev03]). To obtain the desired information, the beat signals of two heterodyne laser interferometers, shifted in frequency, are mixed, where the individual phases of each of these beat signals are proportional to the corresponding OPD. When the two signals are combined and a bandpass filter around the difference of the two heterodyne frequencies $\Delta\nu = \Delta\nu_2 - \Delta\nu_1$ is applied, one gets a resulting beat signal

$$S(t) \approx \cos(2\pi\Delta\nu t + \phi) \quad , \tag{4.11}$$

where the phase is the difference of the initial beating signals' phases and carries the OPD information ([Schu06]):

$$\begin{aligned} \phi &= \phi_2 - \phi_1 \\ &= 2\pi \left[\frac{\Delta\mathrm{OPD}}{\lambda_1} - (\Delta\nu \cdot \mathrm{OPD}_2) \right] \quad . \end{aligned} \tag{4.12}$$

Assuming that both, the phase of the combined signal and the OPD, or equivalently phase of one of the interferometers, are measured, and the wavelengths λ_1 are known, the dOPD can be determined.
For PRIMET an intrinsically frequency stable NPRO (Non Planar Ring Oscillator) Nd:YAG laser, with $\lambda = 1319\,\mathrm{nm}$, has been chosen to provide the required power and linewidth for a coherence length of the maximum OPD at the VLTI ($\sim 250\,\mathrm{m}$). Within the heterodyne assembly, the generated single-mode wave is split into four linearly polarized beams, shifted in frequency by fiber coupled acousto-optics modulators so that the heterodyne frequencies of the

two interferometers become 650 kHz and 450 kHz respectively. By the means of polarizing beam splitters (PBS), each two beams are then superimposed and injected into the system at the FSU M4 mirrors, resulting in each of the FSUs being fed by two beams with orthogonal polarization states, which in the following are separated at the beam combining cube splitters and sent on different optical paths to the telescopes (see Fig. 4.6). A small amount of light, leaking at the PBSs due to imperfect polarization, is used as the reference signal at the phasemeter. There, the phase shift between the science beams, after having been retroreflected at the telescope's STSs and recombined within, as well as extracted at the FSUs, and the reference is precisely measured by a zero-crossing analyzer in order to deduce the dOPD.
In the course of these measurements, a fringe counter is used to provide an unwrapped phase, not limited to the range of one wavelength.
Before the phasemeter is reached, again some fraction of the light is extracted in order to monitor the beams' lateral displacements, which within a controlling loop can be corrected by the field selecting mirrors M11 at the star separators (see Fig. 4.2).

Eventually, two estimates are provided by PRIMET and stored into the raw data files during data recording, namely the dOPD and OPD_B, as seen by FSUB. Still, as these measurements, due to the fringe counter, are recorded only in an incremental way and do not convey absolute measurements, the metrology zero point needs to be calibrated. For this purpose, within each observation procedure, the two science beams are swapped at the STSs, while the injection of the PRIMET beams is maintained. Consequently, the sign of the dOPD changes formally and the zero point can be derived when fitting the delay (see description on `par_scired2` on p. 123 and the diff. delay model in Sec. 4.2.2).

4.2. Astrometric Data Reduction

Recalling the complexity of the instrument and all the so far described effects that need to be considered regarding fringe sensing in interferometry, it is not surprising that the goal of stellar astrometry at the precision level of tens of microseconds can be only achieved, if the data is carefully and deliberately processed to correct for any possible error sources. At this level of accuracy the ordinary instrument users of the scientific community are not expected to know the system's error sources by the same amount as the instrument's developers and maintainers, wherefore an automated data processing pipeline is required to provide the user with fully calibrated scientific data. For this reason the *Astrometric Data Reduction Software* (ADRS) for the PRIMA astrometric mode has been designed and implemented by the exoplanet group of the ESPRI consortium participant Landessternwarte Königstuhl, Heidelberg.
The raw observation data, taken with regard to astrometry, is recorded usually while fringes of both stars, the target star and the reference, are tracked by the instrument. Consequently, in this dual-feed mode it primarily contains the FSU signals of all detector pixels and the delay information from the laser metrology. In the course of data processing, the ADRS, designed upon the results from the error analysis by [Tubb08], combines this observation data with calibration frames and environmental data and delivers the main observable of interest, the differential

delays, which are completely corrected for instrumental and principal astrometric errors (see Tab. 4.2) and which are converted to star separations on the sky through the knowledge of the projected baseline. The data interpretation, as orbit fitting of possible companions, is not part of the astrometric pipeline and lies in the responsibility of the individual users.

Table 4.2. – Generally, the interferometric raw data recorded by PRIMA needs to be calibrated for effects of several sources. The ADRS pipeline corrects for errors of three different kinds. Instrument specific calibration comprises FSU characteristics, the entire system's spectral response, as well as corrections to the delay controller's mechanism and the calibration of the arbitrary PRIMET constant term. Furthermore, environmental data in the delay line tunnel is collected and corrected for, as well as effects related to the specific observation.

instrument	environment	observation
• relative gains between FSU pixels • relative phase errors between ABCD signals • spectral response of the VLTI system • tracking errors (offset from fringe center) • metrology zero point	• dispersion effects for main delays (temperature, pressure, humidity) • (environmental trends)	• sky background • source spectrum • earth orientation • astrometric corrections

4.2.1. Pipeline Processing and Data Flow

The raw instrument data recorded and delivered by PRIMA is stored within files of the standard FITS format. Observation templates trigger the system's operation, and predefined template keywords, stored to the file headers, define the identity of the recorded file, which might have been created for calibration purposes or contain standard tracking data from an ordinary observation, and drive the ADRS operations. In order to distinguish between the raw data and intermediate products along the data reduction process, an increasing level number is assigned to the data. It directly represents the amount of processing that has been performed by the ADRS. In this manner, the raw data, most importantly including the recordings of all FSU detector pixels and the metrology information about the main delays and the differential delay, is labelled as "Level 0", whereas the final product is called "Level 3" data.
The software is subdivided into two principal modules, the *on-line* and *off-line* part. The principle behind this approach is the strive for a compromise, namely to provide the on the site user with a fundamental collection of procedures, performing already the basic data reduction on the one hand, but on the other to have a reduction tool, which not only delivers completely calibrated scientifically relevant data, but also one that can be updated when knowledge regarding the instrument and the science objects has been gathered so that the data can be reprocessed

and the results improved.
For this reason the on-line part processes files that have been recorded for calibration, either within the laboratory, or during the night on the sky, but in particular performs data "averaging", by which the highly sampled tracking data is compressed to a product of reduced file size, which can be more quickly processed further and which provides all the important information, referred to a single well-defined and fixed time grid.
The recipes of this module are listed and sketched in the following, while the background mathematical concepts are given in [Eli08]:

- `pacma_labdark`:
  Processes dark frame files, recorded with the FSU entrances covered up, in order to calculate the bias and dark current for each of the 20 pixels of both FSUs. The results are later used as detector health checks and for FSU data calibration.

- `pacma_labflat`:
  As part of the calibration process this recipe retrieves relative gains between the FSU detector pixels. For this purpose the signals are recorded, while the fringe sensor units are fed by coherent beams of the artificial light source MARCEL with a known and adjustable black-body spectrum.

- `pacma_fsuresponse`: This recipe analyzes the fringe package, which is obtained from FSU signal recordings during multiple delay scans over the fringes. By using the calibration light source MARCEL, pure spectral response of the fringe sensor unit pixels can be derived (see Fig. 4.7(b)). After removing the dark current and correcting for relative gains between the several scans, the delay dependent fringe package $p(D)$ in general is described by the oscillating fringe function and its envelope, which carries the information about the system's bandpass' shape. The spectral response $\tilde{p}(\kappa)$, as a function of the wavenumber, is then obtained for each pixel and scan individually by applying a DFT to the scan data:

$$\tilde{p}(\kappa) = \frac{1}{N} \sum_{n=0}^{N-1} p(D_n)\, e^{-i2\pi\kappa D_n} \quad , \tag{4.13}$$

  where $n = 0, 1, 2, \ldots, N-1$ indicates the individual data samples within one scan. The resulting responses are corrected for (divided by) the source's spectrum and, for each of the 20 pixels averaged over the individual scans. From this final result the effective wavenumbers $\kappa_{\text{eff.}}$ are derived as weighted averages, and the relative phases between the ABCD signals are calculated from the corresponding cross spectral density spectra (see Sec. 4.3.1.1).

- `pacma_skybackground`:
  As part of the on sky calibration the telescopes are pointed at a dark region near the observation targets to measure the response of the detector pixels to the flux from the sky background. This recipe processes the corresponding recorded data and calculates the background from the means of the signals of the provided frames for each pixel independently.

- `pacma_starflat`:
  To correct for uncorrelated gain errors between the FSU pixels, which are also affected by the source's spectral distribution, since the effective wavenumbers vary from pixel to pixel (by principle for the spectral channels and only slightly between the different detector quadrants), the telescopes are pointed at the stars of interest. After the sky background is removed, the relative gains are obtained from the means of several frames and a normalization to the highest value.

- `pacma_vltiresponse`:
  This recipe behaves in principle just like `pacma_fsuresponse`, different only in the sense that the input data is obtained not with the calibration light source, but on sky with the target stars. Consequently, the resulting spectral responses are corrected by the stellar spectra of the observed stars, instead of the MARCEL spectrum. The sampling of the fringes is achieved by tracking one star and using it as the phase reference, whereas the differential delay line scans over the fringe package by modulating the dOPD. From these scans the spectral responses of the FSUs, as well as the effective wavenumbers and relative phases, respectively their deviations from the nominal values of $\pi/2$ are measured on sky, while the light beams pass through the entire PRIMA system.

- `pacma_starspectrum`:
  In order to calculate the effective wavenumbers and relative phases as precisely as possible, the observed star's *spectral photon distribution* (SPD) needs to be known and provided to the recipe `pacma_vltiresponse`. Usually, these will be provided by the user, but in principle PRIMA can also be used to measure them on the sky directly. For this purpose the same procedure is carried out as for the VLTI response frames, tracking the fringes of one star and scanning over those of the secondary. Consequently, the SPD in question can be derived from the scans by correcting the spectral responses, calculated by Fourier transform spectroscopy as in the case of `pacma_vltiresponse`, but corrected for the system's bandpass (divided by the output of the VLTI response).

- `pacma_sciave`:
  The main purpose of this recipe is the scientific averaging of relevant data and to prepare the output for further processing, which basically means to interpolate all entities onto the same fixed time grid. If the desired sampling ratio is not provided by the user, the recipe will create by default a time grid of timestamps t_i with intervals in time of $\Delta t = 1\,\mathrm{s}$. The averaging process itself depends on the type of the regarded data. Since main delay and PRIMET delay data are determined by the orientation of the projected baseline on the sky, earth rotation contributes significantly even on the time scales of seconds. For this reason the delays are estimated from a linear fit over Δt around each time stamp t_i, where the change of the baseline orientation $\partial\theta/\partial t$ is regarded to be constant on this time scale.
  From the FSU data, which is corrected for relative gains and the sky background, the squared visibilities and phases for all spectral channels are derived by the ABCD method, and an average photon number is estimated from the average of the sum of the ABCD signals.
  Furthermore, for later processing in the chain of data reduction, two additional estimates

are calculated from the signal averages. The *cross visibility* between the primary and secondary signals (at FSUB and FSUA) is derived from

$$\mu_{\times}(\kappa_m) = \frac{F_{\times}(\kappa_m)}{N_1(\kappa_m)N_2(\kappa_m)} \quad , \tag{4.14}$$

where the index m stands for the distinct spectral channels and $N_{1,2}$ for the averaged photon numbers in the FSUs, distinguished by the indices $1, 2$ (primary and secondary). The here underlying complex *cross flux* is given by

$$\begin{aligned} F_{\times}(\kappa_m) &= F_1^*(\kappa_m)F_2(\kappa_m) \\ &= X_1(\kappa_m)X_2(\kappa_m) + Y_1(\kappa_m)Y_2(\kappa_m) \\ &\quad +i\left[X_1(\kappa_m)Y_2(\kappa_m) - X_2(\kappa_m)Y_1(\kappa_m)\right] \quad , \end{aligned} \tag{4.15}$$

with the quadratures $X_{1,2} = (A_{1,2} - C_{1,2})$ and $Y_{1,2} = (B_{1,2} - D_{1,2})$.
From the phase angles $\varphi_{\times}(\kappa_m) = \tan^{-1}\frac{\mathrm{Im}\left[F_{\times}(\kappa_m)\right]}{\mathrm{Re}\left[F_{\times}(\kappa_m)\right]}$, also called the *differential phases*, the differential tracking offset from the fringe center will be estimated during data processing in the following recipes.

Since the most important scientific processing is postponed within this part of data reduction, the on-line module usually needs to be executed only once.

Within the off-line part the delays are corrected for environmental, instrumental and astrometric effects to yield the desired star separations on the sky. For this purpose, additional external information, such as any stellar properties, generally need to be provided as input to the ADRS recipes. This is achieved by a collection of catalogs of relevant data, in this context called the *Correction Collection* (CoCo). The CoCo contains catalogs, which are in some cases compiled by the user and in others created by the ADRS recipes themselves. However, the correct FITS format (preliminary given in [Eli08] but will be updated and provided with the final release and delivery of the ADRS to ESO) must be guaranteed so that it can be exploited during the off-line data processing:

Astrometric catalogs:

These catalogs contain basic information about the observed stellar sources, as positions, reference epochs, proper motions, distances (parallaxes), radial velocities and more. The instrument user himself is responsible for this data collection. As it will be the case usually, the information about the primary source will be of higher precision than of the secondary, in general more distant, sources. With time evolving during a science program, the PRIMA measurements themselves can be used to gather more precise differential information regarding specific pairs of stars, which in turn can be exploited to improve the astrometric data of the secondary objects. The differential catalog, created or, if already present, updated by the recipe `pacma_scired3`, serves as input to previous recipes of the off-line part so that with improved astrometric information the whole module execution can be iterated to reprocess the pipeline results.
A public star catalog, containing stellar sources that have been certified by ESO and the consortium to be used as baseline calibrators within the `pacma_baseline`, will also be provided.

IERS data:
The *International Earth and Reference Systems Service* (IERS) provides data predictions and precise measurements regarding earth orientation, which needs to be collected and converted to a suitable FITS file format that can be processed by the ADRS. This information is needed in the context of baseline calibration itself, as well as when the baseline is converted from the geocentric *International Terrestrial Reference System* (ITRS) to the ICRS to correct for any effects related to the earth orientation, such as precession, nutation or polar motion. Besides that, some of the IERS data is necessary for the calculations of the astrometric corrections (see Sec. 4.2.3).

Ephemerides of Solar System Bodies:
Within the `pacma_scired2` recipe the software estimates the influence of relativistic light deflection on the measured differential delay. For this reason, it needs to be provided with accurate positions of the most prominent solar system bodies, which are estimated from the high-precision ephemerides, made available by the *Jet Propulsion Laboratory*.

Baseline catalog:
In this catalog the baseline vectors for all used VLTI configurations will be stored and updated by the solutions of the baseline recipe.

VLTI responses:
This collection contains the spectral responses of the VLTI system, which are derived from calibration observations on the sky and processed by the recipe `pacma_vltirespose`. Although the results are stored in the CoCo, this routine should probably be considered as part of a standard night, or even object calibration.

SPD collection:
For the calculation of the effective wavenumbers and relative phases, the recorded responses are corrected for the individual stellar spectral photon distributions. Those, which are usually gathered by external means, but also in principle could be created by the on-line recipe `pacma_starspectrum`, are also part of the CoCo.

Environmental trends:
This catalog, whose format still needs to be specified, is conceptionally incorporated as an opportunity to keep track of and store any manifestations of data trends as functions of environmental parameters, such as atmospheric seeing, telescope pointing or others. The idea is that some correlations, of which there is no knowledge of in the beginning, might be found in the data on a long term. These could be investigated and characterized afterwards, yielding the possibility of further data detrending.

With this colletion provided, the "Level 1" data can be processed by the off-line part of the pipeline, which the following recipes belong to:

- `pacma_environment`:
  Within this recipe the environmental sensor data that has been compiled from the ESO database and that provides information on temperature, pressure, etc., measured by the sensors at different positions of the instrument infrastructure, is verified and interpolated onto the same time grid as the observational "Level 1" files, as the sensor readings are

sparsely spaced in comparison to the PRIMET and FSU data. This interpolation is performed by a linear fit after a low-pass filter is applied to the data to suppress occurrences of noise in the otherwise smooth and slow variations.

- `pacma_scired1`:
  This second main scientific reduction recipe performs instrumental and environmental corrections onto the measured delays. The main delay (MDL), as well as the PRIMET FSUB and derived FSUA delay data is corrected for longitudinal dispersion (see p. 107). Two major effects are taken into account, which both result from the circumstance that the main delay lines, supposed to compensate the external OPD, are not evacuated. Consequently, on the one hand the applied and measured delays within the interferometer need to be stretched due to the ratio of the stellar beams' wavelength in vacuum and air ($\lambda_{\mathrm{s,vac}}/\lambda_{\mathrm{s,air}} = n_{\mathrm{s}}$), and on the other hand a correction needs to be applied as the vacuum wavelength of the metrology is used to recover the delays, which leads to overestimated delays by the ratio $\lambda_{\mathrm{met,vac}}/\lambda_{\mathrm{met,air}} = n_{\mathrm{met}}$.
  Altogether the corrected delays are determined from the measured delays d by

$$d_{\mathrm{corr}} = \frac{n_{\mathrm{s,G}}(T,P,H,\kappa_{\mathrm{s}})}{n_{\mathrm{met}}(T,P,H,\kappa_{\mathrm{met}})} \cdot d \quad , \tag{4.16}$$

  where the *group refractive index* $n_{\mathrm{s,G}}(T,P,H,\kappa_{\mathrm{s}}) = n_{\mathrm{s}}(T,P,H,\kappa_{\mathrm{s}}) + \kappa\partial n/\partial\kappa$ is used to account for the fringe control system tracking at zero group delay that shifts in an non-evacuated environment due to chromatic dispersion.
  The appropriate refractive indices, which depend on temperature, pressure, humidity and wavenumber, are derived from models, valid for infrared (FSU bandpass and PRIMET) and visible (main delay line metrology) wavelengths, to be found in [Bir1993] and [Mat06].
  The PRIMET differential delays are corrected for tracking errors only (offsets from the fringe center), as with the DDLs situated in vacuum any differential longitudinal dispersion should be negligible. The tracking offset correction is important due to imperfect fringe tracking, resulting in some delay jitter around the fringe center positions.
  The offsets are not derived and corrected for each FSU individually, but the differential offset term Δd_{off} is estimated from the cross visibility $\mu_{\times}(\kappa_m)$, already calculated beforehand in `pacma_sciave`. In the case of precise on-fringe tracking, the differential phases of $\mu_{\times}(\kappa_m)$ vanish and Δd_{off} becomes zero.
  Typically, this is not the case, and the offset is estimated from the position of the peak of a synthesized differential fringe-packet function, the *differential group delay*, defined by ([Eli08])

$$Q(\Delta d) = \frac{1}{M}\sum_{m=1}^{5} \mu_{\times}(\kappa_m)\, e^{-i2\pi\kappa_m \Delta d} \quad , \tag{4.17}$$

  with m indicating the five FSU spectral channels.

This is also the point in the data reduction process, where environmental trend corrections, if any are provided by the CoCo, are applied to the data. Due to that, in the

case of finding new trends, the data processing needs to be repeated upwards from this recipe on.

- `pacma_baseline`:
  Main delay data is compiled here to solve for the wide-angle baseline (see also Sec. 4.3.2). Although in principle it is foreseen to use only single-feed observations of stars with extraordinary accurate astrometric solutions, as such listed in the FK6 catalog and consequently certified as good baseline calibrators, in general each tracking data can be exploited for this purpose.
  The recipe gathers the required information from the "Level 2" observation frames and, provided with the corresponding star catalog, feeds the data into the *Interferometer Optical-Path-Length Analysis System* (IPHASE ([Wall08])) software to fit for the interferometric baseline and the internal delay constant term following the standard delay equation:
  $$d(t) = \hat{\boldsymbol{s}}(t) \cdot \boldsymbol{B} + C \quad . \tag{4.18}$$
  Here C denotes the constant term, which defines the delay's zero OPD position in the case of no external OPD. The baseline is assumed to be constant in the carthesian geocentric reference system and the time dependence is covered by the star unit vectors, depending on the observation site's local sidereal time t_{LST}:
  $$\hat{\boldsymbol{s}} = \cos\delta\cos(\alpha - t_{\mathrm{LST}})\hat{\boldsymbol{x}} + \cos\delta\sin(\alpha - t_{\mathrm{LST}})\hat{\boldsymbol{y}} + \sin\delta\hat{\boldsymbol{z}} \quad . \tag{4.19}$$
  Naturally, the physical baseline is bound to the individual telescope positions so that individual solutions to distinct VLTI configurations need to be acquired, which are thereon stored in the CoCo baseline catalog.

- `pacma_scired2`:
  This recipe is principally the last chain in the data reduction procedure and delivers the final scientific results in sense of fully corrected differential delays Δd and star separations $\Delta \boldsymbol{s}$ for each time stamp of acquired observations.
  As the system's delay lines, both the MDLs and the DDLs, are adjusted so that the internal delays compensate the path differences for the two stars inside of the instrument, following Eq. (4.18) the individual measured delays for the primary (1) and the secondary (2) source are given by
  $$d_{1,2}(t) = \hat{\boldsymbol{s}}_{1,2}(t) \cdot \boldsymbol{B} + C_{1,2} \quad , \tag{4.20}$$
  and the pure differential delays that should manifest in the measurement are determined by
  $$\begin{aligned} \Delta d(t) &= d_2(t) - d_1(t) \\ &= \Delta \boldsymbol{s}(t) \cdot \boldsymbol{B} + \Delta C + z_0 \quad , \end{aligned} \tag{4.21}$$
  where ΔC is the difference between the internal constant terms and z_0 is the PRIMET zero point, since, as stated before, the laser metrology follows the varying OPD only in an incremental way and needs to be calibrated. This is done by the combination of observations in the NORMAL and the SWAPPED modes respectively.

Switching the light paths in the instrument by the field rotation by 180°, flips the orientation of the baseline in reference to the star separation vector, and the equation for the measured differential delays becomes (see Sec. 4.2.2 for derivation)

$$\begin{aligned} \Delta d(t) &= \pm \Delta \boldsymbol{s}(t) \cdot \boldsymbol{B} + (\Delta C + z_0) \\ &= \pm a_0 \pm a_1 \cos(\omega t) \pm a_2 \sin(\omega t) + Z_0 \quad , \end{aligned} \tag{4.22}$$

where the upper signs refer to the NORMAL state and the bottom to the SWAPPED state, a_i denote the astrometric fit parameters, $Z_0 = (\Delta C + z_0)$ the instrumental zero point, and the circular frequency $\omega = (2\pi/24\,\mathrm{h})$ accounts for the earth rotation.
Due to the sign switching, the astrometric and instrumental fit constants decouple from each other and the zero point can be determined and corrected for if observation data from both modes is available. Consequently, this recipe must be provided with at least one "Level 2" frame from both set ups, where in-between the metrology is not reset, which would change the value of the metrology zero point.
Within the recipe the fit is not performed on the differential delays directly, but on residual values after the harmonic OPD model, based on the baseline and the estimates from the differential catalog, is removed first. When more and more data on individual star pairs is gathered, the updated OPD model, used during the iterative reprocessing of the data, should improve and converge towards the actual star separations, which would be indicated by the fit coefficients a_i approaching zero.
Also, before the fit is performed, the differential delays are corrected for astrometric effects, such as aberration, light deflection, etc. (see Sec. 4.2.3), which themselves are estimated from the knowledge of the observed sources.
After the instrumental zero point is recovered, the fully calibrated delays ideally represent the star separations on the sky and are converted to those scientifically meaningful quantities by the time dependent projected baseline, whereby the data reduction is completed.

- `pacma_scired3`:
  Although the "Level 3" files already contain the scientific results, which are completely corrected for instrument- and sky-based effects, a further refinement is expected when the data is reprocessed with improved input parameters. For this reason this recipe processes the "Level 3" differential delays in order to fit for differential positions, proper motions and parallaxes for individual given star pairs of interest. The results are stored to the differential catalog of CoCo and can serve in turn as input to the recipes earlier in the chain of data reduction, whenever an iteration in form of data reprocessing is performed. In particular, the knowledge about the positions and astrometric quantities of the secondary targets should improve with time, as the collection of PRIMA observations increases.

The full sequence of data reduction is sketched in Fig. 4.8, accentuating the ADRS data products at the distinct levels and the division of the routines to the on-line and off-line parts respectively, allowing a quick and basic data reduction on the site of observation together with a postponed but independent reduction to scientific quantities, which can be iterated with increasing knowledge about the system and the observation targets.

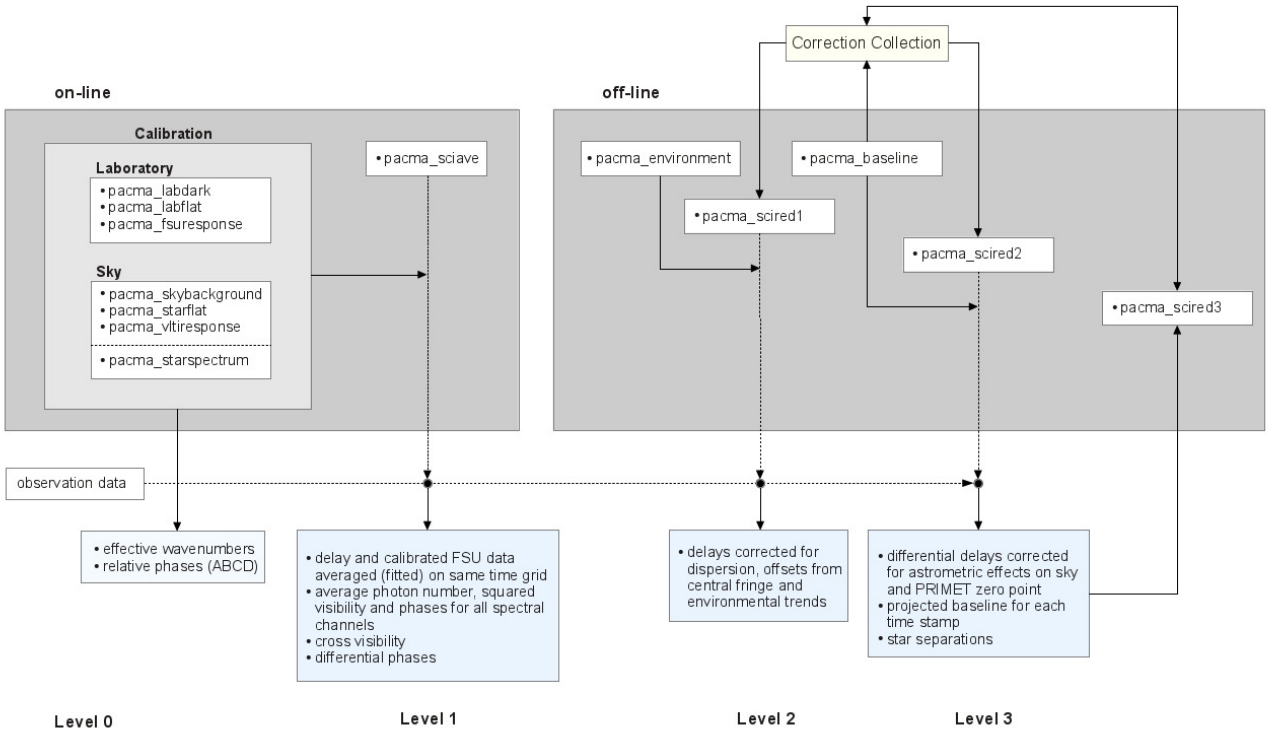


Figure 4.8. – The simplified scheme illustrates the sequence of data processing within the ADRS pipeline. The upper part of the illustration shows the software recipes that act on the recorded data in a well-defined order (interactions illustrated by vertical dashed arrows). Naturally, the data reduction process starts with the raw "Level 0" observation data, recorded by the instrument. At given points in the processing sequence (black knots) the main reduction recipes `pacma_sciave`, `pacma_scired1` and `pacma_scired2` process the information and create preliminary ("Level 1", "Level 2") and final products ("Level 3"). The data flow is represented by the continuous arrows, and the pale blue boxes in the lower part list the products, as created by the pipeline on the different levels and which are accessible by the user. The software package's procedures and actions on the data are conceptionally subdivided into two segments. The on-line part of the software includes recipes for processing calibration frames, as well as the data averaging process, during which the highly sampled raw data is reduced to a less storage space consuming product with all sensor recordings interpolated to a common time grid. At the next level, with the execution of `pacma_scired1`, the data is being corrected for instrumental and environmental effects, whereas the last recipe creates the final differential delays and star separations on the sky.

4.2.2. Differential Delay Model

For the derivation of the model to the measured differential delays, it is convenient to consider the entities in an equatorial reference frame, in which the star's coordinates RA (right ascension) and DEC (declination) do not change over time. Therefore, Δs_α, Δs_δ (separations along the directions of right ascension and declination), as well as the coordinates of the midpoint between the primary and secondary star α_0, δ_0, are constant entities, and the baseline vector, in particular its orientation, becomes a dynamic quantity due to earth rotation.
The physical baseline $\boldsymbol{B} = \boldsymbol{T}_2 - \boldsymbol{T}_1$ in the ideal case is defined by the connecting vector between the positions of the two telescopes' pivot points at $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ and in carthesian coordinates can be denoted by

$$\boldsymbol{B} = \begin{pmatrix} B_L \cos\phi \cos\lambda \\ B_L \cos\phi \sin\lambda \\ B_L \sin\phi \end{pmatrix} \quad , \tag{4.23}$$

where B_L is the baseline length, ϕ is the constant baseline latitude and λ is its time dependent longitude.
When the baseline vector is projected onto the local tangential plane, with its origin placed at the celestial sphere at the star pair's midpoint coordinates and spanned by the orthonormal vectors $\boldsymbol{p}_0$, $\boldsymbol{q}_0$, as defined by Eq. (3.43), its components with reference to this basis become:

$$\begin{aligned} B_\xi = \boldsymbol{p}_0 \cdot \boldsymbol{B} &= B_L \begin{pmatrix} -\sin\alpha_0 \\ \cos\alpha_0 \\ 0 \end{pmatrix} \begin{pmatrix} B_L \cos\phi\cos\lambda \\ B_L \cos\phi \sin\lambda \\ B_L \sin\phi \quad , \end{pmatrix} \\ &= B_L \left[\cos\phi \left(\cos\alpha_0 \sin\lambda - \sin\alpha_0 \cos\lambda\right)\right] \\ &= B_L \cos\phi \sin(\lambda - \alpha_0) \end{aligned} \tag{4.24}$$

and

$$\begin{aligned} B_\eta = \boldsymbol{q}_0 \cdot \boldsymbol{B} &= B_L \begin{pmatrix} -\sin\delta_0 \cos\alpha_0 \\ -\sin\delta_0 \sin\alpha_0 \\ \cos\delta_0 \end{pmatrix} \begin{pmatrix} B_L \cos\phi\cos\lambda \\ B_L \cos\phi \sin\lambda \\ B_L \sin\phi \end{pmatrix} \\ &= B_L \left[-\sin\delta_0 \cos\phi \left(\cos\alpha_0 \cos\lambda + \sin\alpha_0 \sin\lambda\right) + \cos\delta_0 \sin\phi\right] \\ &= B_L \left[-\sin\delta_0 \cos\phi \cos(\lambda - \alpha_0) + \cos\delta_0 \sin\phi\right] \quad , \end{aligned} \tag{4.25}$$

where the standard relations

$$\cos(x+y) = \cos x \cos y - \sin x \sin y \tag{4.26}$$

$$\sin(x+y) = \sin x \cos y + \cos x \sin y \tag{4.27}$$

have been utilized.

According to Eq. (3.47), the OPD and consequently the measured differential delay Δd, when tracking on the fringe center and neglecting instrumental and environmental offsets, are determined by the geometric OPD, the dot product of the projected baseline $\boldsymbol{B}_P$ and the star

separation on sky $\Delta\boldsymbol{s}$. Plugging in equations (4.24) and (4.25) into this relation yields:

$$\begin{aligned}
\frac{\Delta d}{B_L} &= \cos\phi\sin(\lambda-\alpha_0)\cdot\Delta s_\alpha \\
&\quad -\sin\delta_0\cos\phi\cos(\lambda-\alpha_0)\cdot\Delta s_\delta \\
&\quad \underbrace{+\cos\delta_0\sin\phi\cdot\Delta s_\delta}_{C_a} \\
&= C_a+\cos\phi\left[\sin(\lambda-\alpha_0)\Delta s_\alpha-\cos(\lambda-\alpha_0)\sin\delta_0\Delta s_\delta\right] \\
&= C_a+\underbrace{\Delta s_\alpha\cos\phi}_{C_s}\sin(\lambda-\alpha_0)\underbrace{-\cos\phi\sin\delta_0\Delta s_\delta}_{C_c}\cos(\lambda-\alpha_0) \\
&= C_a+C_s\cdot\sin(\lambda-\alpha_0)+C_c\cdot\cos(\lambda-\alpha_0) \quad , \qquad (4.28)
\end{aligned}$$

where C_a, C_s and C_c are constants.

When the term $(\lambda-\alpha_0)$, which represents the time dependency of the problem, is substituted by $(\omega t+\varphi)$, with the circular frequency of the earth rotation $\omega=(2\pi/24\,\mathrm{h})$ and an unknown phase φ, and the relations (4.26), (4.27) are applied repeatedly, one arrives at the following final expression for the differential delays:

$$\begin{aligned}
\frac{\Delta d}{B_L} &= +C_a \\
&\quad +C_s\sin(\omega t)\cos\varphi+C_s\cos(\omega t)\sin\varphi \\
&\quad +C_c\cos(\omega t)\cos\varphi-C_c\sin(\omega t)\sin\varphi \\
&= C_a+\cos(\omega t)\underbrace{\left[C_s\sin\varphi+C_c\cos\varphi\right]}_{A_1}+\sin(\omega t)\underbrace{\left[C_s\cos\varphi-C_c\sin\varphi\right]}_{A_2} \\
\frac{\Delta d}{B_L} &= C_a+A_1\cdot\cos(\omega t)+A_2\cdot\sin(\omega t) \quad . \qquad (4.29)
\end{aligned}$$

If now the instrumental zero point and the sign dependence on the modes of observation (NORMAL vs. SWAPPED) are incorporated, one obtains:

$$\Delta d=\pm B_L\cdot C_a\pm B_L\cdot A_1\cdot\cos(\omega t)\pm B_L\cdot A_2\cdot\sin(\omega t)+Z_0 \quad . \qquad (4.30)$$

This is the same expression as given by Eq. (4.22), when the fit parameters described before and the constants here are linked by:

$$\begin{aligned}
a_0 &= B_L\cdot C_a \qquad (4.31)\\
a_1 &= B_L\cdot A_1 \qquad (4.32)\\
a_2 &= B_L\cdot A_2 \quad . \qquad (4.33)
\end{aligned}$$

The so called *astrometric constant* a_0 gives the portion of the delay due to the projection onto the z-axis, pointing to the northern celestial pole, and the constants a_1 and a_2 describe the rest, originating from the projection onto the xy-plane ($\sqrt{(B_L\cdot A_1)^2+(B_L\cdot A_2)^2}$).

It is worth mentioning that the two fit parameters a_0 and Z_0 decouple only due to their behavior in course of beam swapping and that the instrumental zero point can be determined and corrected for by fitting the differential delays to the model given by Eq. (4.30).

Also interesting is that in principle, in order to determine the zero point, a fit to the delays can be performed without the knowledge of the baseline. However, meaningful star separations can only be recovered when the baseline is estimated beforehand.

4.2.3. Astrometric Corrections

As described in the previous subsection, the fit to the measured differential delays that determines the instrumental zero point is performed by the ADRS recipe `pacma_scired2` on delay residuals δd, after the OPD model has been removed. In addition to that, at this point also sky related corrections are applied. These astrometric effects are consequences to the relative motions between the observer, following the observation site's trajectory through the solar system instead of being fixed in the barycentric coordinate system, and the observed objects moving in space, described by their proper motions, as well as to any other disturbances along the light paths, as is the case for light deflection by massive bodies.
The influences from these effects ultimately should be known to a precision that is approximately one order above the instrument's desired and pursued level of accuracy, thus to the μas level in the case of PRIMA.
In the usual case, one would measure an apparent angular separation, or equivalently the stars' apparent positions, as seen by the observer on earth and correct these for the astrometric effects afterwards to obtain the real geometric star coordinates that could be further investigated. Since the effects corrected for can be considered being sources of disturbances at certain points along the path of the information transporting photons travelling from the star to the observer, also the corrections should be applied sequentially and in a certain order and not arbitrarily. [Kli03] suggests the following order of sequence, starting with effects due to the motion of the observer:

1. Aberration
2. Relativistic light deflection
3. Parallax
4. Proper motion

In the case of PRIMA's data reduction with the pipeline, the situation is a little different. The astrometric effects need to be estimated before the fit to the delays is performed and therefore also before the measurement, which gives the apparent separation, is available. For this reason the astrometric effects are estimated based on the knowledge of the star pair, which is provided by the catalogs in the CoCo and is updated with time when the star pair of concern is observed and the differential catalog is updated. However, in principle the actual star separation vector $\Delta\boldsymbol{s}$ is the starting point, from which the apparent angular separation vector, as would be observed, in this context called $\Delta\boldsymbol{s}'$ from here on, is calculated.

The astrometric corrections can be then considered small perturbations to the actual star positions and are defined by and can be recovered from:

$$\Delta\boldsymbol{\phi} = \Delta\boldsymbol{s} - \Delta\boldsymbol{s}' = \Delta\boldsymbol{s} - g(\Delta\boldsymbol{s}) \quad , \tag{4.34}$$

where the function $g(\Delta\boldsymbol{s}, \hat{\boldsymbol{s}}_1, \hat{\boldsymbol{s}}_2, \ldots)$ describes the transformation from the actual to the apparent separation vector

$$g(\Delta\boldsymbol{s}, \hat{\boldsymbol{s}}_1, \hat{\boldsymbol{s}}_2, \ldots) : \Delta\boldsymbol{s} \mapsto \Delta\boldsymbol{s}' \quad , \tag{4.35}$$

and the contribution of the corrections to the differential delays are generally given by the projection onto the baseline $\boldsymbol{B}$:

$$\delta d_{\mathrm{astro}} = \Delta\boldsymbol{s} \cdot \boldsymbol{B} - \Delta\boldsymbol{s}' \cdot \boldsymbol{B} = \Delta\boldsymbol{\phi} \cdot \boldsymbol{B} \quad . \tag{4.36}$$

Based upon this principle of procedure, it is clear that the order, at which the corrections are calculated and applied, needs to be reversed in comparison to the one stated above, as the effects are taken into account sequentially while following the light path from the stellar source to the observer.
The corrections are applied to the one second averages with time stamps t_j of the "Level 2" differential data by removing their overall contributions, as evaluated by Eq. (4.36), so that the residuals, to which the delay fit is applied, are eventually given by

$$\delta d(t_j) = \Delta d(t_j) - \Delta d_{\mathrm{mod}}(t_j) - \delta d_{\mathrm{astro}}(t_j) \quad , \tag{4.37}$$

with $\Delta d_{\mathrm{mod}}(t_j)$ the geometric delays obtained from the used OPD model.

In the following, the astrometric corrections, which have been derived in the course of this thesis, based on the descriptions found in [Kov04], are presented individually in the order of appliance.

4.2.3.1. Proper Motions and Epoch Transformations

The transformations of star properties, in particular the stars' coordinates, between distinct moments in time are crucial, as usually they are given in reference to a certain epoch T_{ref}. Consequently, in the ADRS they need to be converted to the time of the observation T_{obs}, or even to each time stamp, as they are used as inputs for the calculations of the astrometric corrections, as well as for the estimation of the differential delays, based on the OPD model.
It is trivial that the star positions change in the course of time due to their proper motions, but it should be noted that the proper motion itself, although describing a constant motion in space, is not a constant quantity, which is also the case for other stellar properties describing the object's position in space, such as the parallax and the radial velocity.
This effect is of pure geometrical nature, as the proper motion vector, defined by the time derivatives of the objects' coordinates α and δ at a certain point of time t_0

$$\begin{aligned} \boldsymbol{\mu} &= \left(\frac{\mathrm{d}\alpha}{\mathrm{d}t}\right)_{t_0} \cos\delta \boldsymbol{p}_0 + \left(\frac{\mathrm{d}\delta}{\mathrm{d}t}\right)_{t_0} \boldsymbol{q}_0 \\ &= \mu_\alpha \cos\delta\, \boldsymbol{p}_0 + \mu_\delta\, \boldsymbol{q}_0 \\ &= \mu^*_\alpha\, \boldsymbol{p}_0 + \mu_\delta\, \boldsymbol{q}_0 \quad , \end{aligned} \tag{4.38}$$

describes only the object's motion along the local tangential plane, whereas the radial component is usually separately defined by the radial velocity $V_{\rm rad}$ along the line of sight.
Together, these two quantities specify the object's entire motion in space, which generally can be considered constant and is given in kilometers per second by:

$$\boldsymbol{V} = \frac{A}{\varpi}\,\mu_\alpha^*\,\boldsymbol{p}_0 + \frac{A}{\varpi}\,\mu_\delta\,\boldsymbol{q}_0 + V_{\rm rad}\hat{\boldsymbol{s}} \quad , \tag{4.39}$$

with the annual parallax $\varpi = 1/\rho$, determined by the reciprocal of the distance ρ in units of parsec, the unit vector $\hat{\boldsymbol{s}}$ towards the object's coordinates on the celestial sphere and the constant $A = 4.74047\,\mathrm{km\,yr/s}$.
Considering now the modulus of the proper motion $\mu = (V\sin\theta)/\rho$, where θ is the angle between the direction of the space velocity and $\hat{\boldsymbol{s}}$, then its derivative with time is non-vanishing as long as $V_{\rm rad} \neq 0$ ([Kov04]):

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = -\frac{2\mu}{\rho}\,V_{\rm rad} \quad . \tag{4.40}$$

This results in a perspective acceleration of the celestial objects with respect to motions within their locally defined tangential plane.

The correct transformations to epochs $t = T_{\rm obs} - T_{\rm ref}$ can be recovered when the entire motion in space is taken into account. The time dependent radius vector $\boldsymbol{r}(t)$ of the star, also called the barycentric vector when regarded in the ICRS, is completely determined by its properties at the reference epoch, denoted by the subscript "0"

$$\boldsymbol{r}_0 = \hat{\boldsymbol{s}}_0 \frac{A}{\varpi_0} \quad , \tag{4.41}$$

and the space velocity $\boldsymbol{V}$ through the relation

$$\boldsymbol{r}(t) = \boldsymbol{r}_0 + \boldsymbol{V}t \quad . \tag{4.42}$$

Given this, the temporal evolution of the unit vector towards the star reads as:

$$\hat{\boldsymbol{s}}(t) = \frac{\boldsymbol{r}(t)}{|\boldsymbol{r}(t)|} = [\hat{\boldsymbol{s}}_0(1+\zeta_0 t) + \boldsymbol{r}_0 t]\,f \quad , \tag{4.43}$$

where the radial velocity has been parametrized by $\zeta_0 = \frac{\varpi_0}{A} V_{\rm rad,0}$, and f is introduced as a normalization factor describing the ratio between distances to the star at different epochs:

$$f = \frac{|\boldsymbol{r}_0|}{|\boldsymbol{r}(t)|} = \left[1 + 2\zeta_0 t + (\mu_0^2 + \zeta_0^2)t^2\right]^{-1/2} \quad . \tag{4.44}$$

From the time dependent unit vector, which incorporates all perspective effects due to the complete star's motion, the object's coordinates can be recovered by the standard relation given by Eq. (3.40) and can be fed into the program recipes, whenever needed.

The propagated proper motion vector is calculated using $\boldsymbol{\mu}(t) = \mathrm{d}\hat{\boldsymbol{s}}(t)/\,\mathrm{d}t$ so that after differentiation of Eq. (4.43) and simplification of the result one finally obtains:

$$\boldsymbol{\mu}(t) = [\boldsymbol{\mu}_0(1+\zeta_0 t) - \hat{\boldsymbol{s}}_0\mu_0^2 t]f^3 \quad . \tag{4.45}$$

The transformation of the parallax is recovered simply from the definition of the scale factor f and the initial distance at the reference epoch

$$\varpi(t) = \frac{1}{|\boldsymbol{r}(t)|} = \frac{f}{|\boldsymbol{r}_0|} = \varpi_0 f \quad , \tag{4.46}$$

and the value of the parametrized radial velocity is calculated from:

$$\begin{aligned} \zeta(t) &= \frac{\varpi(t)}{A} V_{\mathrm{rad}}(t) \frac{\varpi(t)}{A} = \frac{\varpi(t)}{A} \cdot \frac{\mathrm{d}|\boldsymbol{r}(t)|}{\mathrm{d}t} \\ &= \left[\zeta_0 + (\mu_0^2 + \zeta_0^2)t\right] f^2 \quad . \end{aligned} \tag{4.47}$$

These time dependent quantities are recovered within the `pacma_scired2` recipe by transformations to the epoch of observation and are used for the estimation of the astrometric corrections.

4.2.3.2. Light Time Delay

This correction is applied to the measured delays due to varying light travel times between the stars and the observer. It is less a problem due to the fact that the information, which arrives simultaneously at the observer, was emitted at different moments in time due to the differential stellar distances, but rather that radial velocities change light travel times in the course of time. For the astrometric search for exoplanets, as is planned to be performed by PRIMA, astrometric signals resulting from host stars' reflex motions due to the presence of a companion can only be detected and characterized if sufficient measurements of relative star positions are obtained over a long period of time. If these measured star separations are not synchronized adequately, the change in light travel times, together with the objects' proper motions, could have an effect and result in spurious accelerations and false signals.
The effect would be of an order of $\approx 20\,\mu$as due to a change in light travel time of about 1.8 h over an observation period of two years, if typical values for the radial velocity (20 km/s) and proper motions (100 ms/ yr) are presumed.
In order to achieve the correct synchronization, one could incorporate the light time delays to the time stamps of the observations, by incorporating offsets to those, which would result to adjusted times of observations, different for the two stars. To avoid the complications of introducing many of such new distinct epochs, for the ADRS it has been decided to correct the separation measurements by transforming them along the timeline to a global epoch of reference t_{ref}.
This reference epoch, which all observations of the same star pair will consequently be referred to, will be stored in and provided by the CoCo.
In order to minimize errors due to uncertainties in the applied proper motions, it might be a

reasonable choice to place t_{ref} in the middle of the period of all observations. However, if the reference should be repositioned at any time, the "Level 3" data must be reprocessed.

For the correction of the angular separation forward or back to the reference epoch, at first the difference in light travel time ΔT between the current observation epoch and the light travel time to the observer at the reference time, needs to be estimated.
According to equation (4.46), the difference in the distance to a star at time t_{obs} and another time t_{ref} is

$$\rho(t_{\mathrm{obs}}) - \rho(t_{\mathrm{ref}}) = \left(\frac{\rho(t_{\mathrm{cat}})}{f_{\mathrm{cat}\to\mathrm{obs}}} - \frac{\rho(t_{\mathrm{cat}})}{f_{\mathrm{cat}\to\mathrm{ref}}} \right) \quad , \tag{4.48}$$

where $\rho(t_{\mathrm{cat}})$ is the distance at the epoch provided by the provided catalog, and the scale factors f are functions of proper motion and radial velocity, introduced in Eq. (4.44), when transforming, as indicated by the subscripts, from t_{cat} to t_{obs} and respectively to t_{ref}.
Thus, the difference in light travel time between the current epoch t_{obs} and the reference epoch t_{ref} for a given star is

$$\Delta T = \frac{\rho(t_{\mathrm{cat}})}{c} \cdot \left(\frac{1}{f_{\mathrm{cat}\to\mathrm{obs}}} - \frac{1}{f_{\mathrm{cat}\to\mathrm{ref}}} \right) \quad . \tag{4.49}$$

A positive sign of ΔT means that the light at the current epoch t_{obs} needs more time to reach the observer than it would at the reference epoch t_{ref}, and equivalently the other way around, when the value is negative.
Consequently, at t_{obs} the observation can be interpreted as delayed by the amount of the light time delay, and the observer sees the scene as it was at the epoch ΔT earlier.
In order to obtain the corresponding apparent unit vector of the primary star 1

$$\hat{\boldsymbol{s}}_1' = \hat{\boldsymbol{s}}_1 - \Delta \boldsymbol{s}_1 \quad , \tag{4.50}$$

as it would be observed at t_{obs}, an epoch transformation from the current epoch t_{obs} to an epoch $t_{\mathrm{obs}} - \Delta T$ needs to be performed. This is achieved by applying Eq. (4.43) so that one obtains

$$\hat{\boldsymbol{s}}_1' = f_{1,\mathrm{obs}\to\mathrm{obs}-\Delta T_1} \left[\hat{\boldsymbol{s}}_1 (1 - \zeta_1(t_{\mathrm{obs}}) \Delta T_1) - \boldsymbol{\mu}_1(t_{\mathrm{obs}}) \Delta T_1 \right] \quad , \tag{4.51}$$

where $\zeta(t_{\mathrm{obs}})$ and $\boldsymbol{\mu}(t_{\mathrm{obs}})$ represent the star's radial velocity and proper motion at the epoch of the observation,[5] and the scale factor for this particular epoch transformation $f_{1,\mathrm{obs}\to\mathrm{obs}-\Delta T_1}$ is again evaluated from Eq. (4.44).
Subtracting this from $\hat{\boldsymbol{s}}_1$, the vector form of the star's displacement is recovered:

$$\Delta \boldsymbol{s}_1 = (1 - f_1) \hat{\boldsymbol{s}}_1 + f_1 \zeta_1(t_{\mathrm{obs}}) \Delta T_1 \hat{\boldsymbol{s}}_1 + f_1 \boldsymbol{\mu}_1(t_{\mathrm{obs}}) \Delta T_1 \quad , \tag{4.52}$$

where the subscript of the scale factor has been shortened for simplicity.

[5] As these quantities are needed here with reference to the time of observation, they need to be evaluated by the corresponding epoch transformation from t_{cat} to t_{obs}.

The same procedure is applied to the secondary star 2, and consequently the star separations, corrected for the light time delay, become:

$$\begin{aligned}
\Delta\boldsymbol{s}' &= \hat{\boldsymbol{s}}_2' - \hat{\boldsymbol{s}}_1' \\
&= \hat{\boldsymbol{s}}_2 - \Delta\boldsymbol{s}_2 - \hat{\boldsymbol{s}}_1 + \Delta\boldsymbol{s}_1 \\
&= \Delta\boldsymbol{s} - \Delta\boldsymbol{s}_2 + \Delta\boldsymbol{s}_1 \\
&= \Delta\boldsymbol{s} + (f_2 - 1)\hat{\boldsymbol{s}}_2 - f_2\zeta_2(t_{\mathrm{obs}})\Delta T_2\hat{\boldsymbol{s}}_2 - f_2\boldsymbol{\mu}_2(t_{\mathrm{obs}})\Delta T_2 \\
&\quad -(f_1 - 1)\hat{\boldsymbol{s}}_1 + f_1\zeta_1(t_{\mathrm{obs}})\Delta T_1\hat{\boldsymbol{s}}_1 + f_1\boldsymbol{\mu}_1(t_{\mathrm{obs}})\Delta T_1\,. \qquad (4.53)
\end{aligned}$$

4.2.3.3. Annual and Diurnal Parallax

If the stars' motion in space is left aside, their equatorial coordinates, which are given in the ICRS with reference to the solar system barycenter, are constant. However, due to their finite distances, any variation of the observer's position translates to an apparent displacement of the observed source on the sky. The earth's orbital motion around the sun is the reason for the most prominent parallactic displacement. The stars' apparent positions over a year reflect the earth's motion and follow elliptical paths on the celestial sphere, whose semi-major axes are the quantities, which are usually referred to by the term (annual) parallax.
The offsets to this effect can be derived from the relative positions of the solar system barycenter, the observer and the star in question. For an object with the barycentric coordinates (α, δ) and the distance ρ, the barycentric three-dimensional position vector $\boldsymbol{r}$ is described by the sum of the vector to the observatory $\boldsymbol{r}_{\mathrm{obs}} = (x_{\mathrm{obs}}, y_{\mathrm{obs}}, z_{\mathrm{obs}})$ and the geocentric position vector to the star $\boldsymbol{r}_{\mathrm{obs}} = \boldsymbol{r}'$, which can be denoted as:

$$\begin{pmatrix} \rho\cos\alpha\cos\delta \\ \rho\sin\alpha\cos\delta \\ \rho\sin\delta \end{pmatrix} = \begin{pmatrix} x_{\mathrm{obs}} \\ y_{\mathrm{obs}} \\ z_{\mathrm{obs}} \end{pmatrix} + \begin{pmatrix} \rho'\cos\alpha'\cos\delta' \\ \rho'\sin\alpha'\cos\delta' \\ \rho'\sin\delta' \end{pmatrix} \quad . \qquad (4.54)$$

The primed quantities stand for the (apparent) geocentric coordinates and the actual distance from the observer to the star. As the vector to the observatory is constructed from the vector to the earth, together with its second contribution, the observer's offset from the earth center, also diurnal effects due to earth rotation are considered.
Expressing the barycentric quantities by their corresponding geocentric counterparts and a small offset

$$\begin{aligned}
\alpha &= \alpha' + \Delta\alpha_{\mathrm{par}} \\
\delta &= \delta' + \Delta\delta_{\mathrm{par}} \\
\rho &= \rho' + \Delta\rho_{\mathrm{par}} \quad , \qquad (4.55)
\end{aligned}$$

these offsets can be approximated under the suppression of second-order effects, which limits the accuracy to about a microarcsecond for the closest stars, by ([Kov04]):[6]

$$\begin{aligned}
\Delta\alpha_{\mathrm{par}}\cos\delta &= -\frac{x_{\mathrm{obs}}}{\rho}\sin\alpha + \frac{y_{\mathrm{obs}}}{\rho}\cos\alpha \\
\Delta\delta_{\mathrm{par}} &= \left(-\frac{x_{\mathrm{obs}}}{\rho}\cos\alpha - \frac{y_{\mathrm{obs}}}{\rho}\sin\alpha\right)\sin\delta + \frac{z_{\mathrm{obs}}}{\rho}\cos\delta \qquad (4.56)\\
\Delta\rho_{\mathrm{par}} &= x_{\mathrm{obs}}\cos\alpha\cos\delta + y_{\mathrm{obs}}\sin\alpha\cos\delta + z_{\mathrm{obs}}\sin\delta \quad . \qquad (4.57)
\end{aligned}$$

Provided by the exact position of the observer at the time of the observation, these equations can be applied to estimate the geocentric and time dependent unit vectors $\hat{\boldsymbol{s}}'$ for both, the primary and the secondary star:

$$\hat{\boldsymbol{s}}'(t) = \begin{pmatrix} \cos(\alpha - \Delta\alpha_{\mathrm{par}})\cos(\delta - \Delta\delta_{\mathrm{par}}) \\ \sin(\alpha - \Delta\alpha_{\mathrm{par}})\cos(\delta - \Delta\delta_{\mathrm{par}}) \\ \sin(\delta - \Delta\delta_{\mathrm{par}}) \end{pmatrix} \quad . \qquad (4.58)$$

Although the apparent angular separation vector $\Delta\boldsymbol{s}'$ can be constructed from this, for the ADRS another approach has been implemented, where the result is expressed by differential quantities between the two stars, of which the accuracy should improve in the matter of time with increasing number of observations.
For this purpose the rephrased equation (4.54) for the position vector, evaluated for both stars

$$\begin{aligned}
\rho_1\hat{\boldsymbol{s}}_1 &= \boldsymbol{r}_{\mathrm{obs}} + \rho_1\hat{\boldsymbol{s}}_1' \quad , \qquad (4.59)\\
\rho_2\hat{\boldsymbol{s}}_2 &= \boldsymbol{r}_{\mathrm{obs}} + \rho_2\hat{\boldsymbol{s}}_2' \quad , \qquad (4.60)
\end{aligned}$$

are subtracted from each other so that with the substitutions $\hat{\boldsymbol{s}}_2 - \hat{\boldsymbol{s}}_1 = \Delta\boldsymbol{s}$ and respectively $\hat{\boldsymbol{s}}_2' - \hat{\boldsymbol{s}}_1' = \Delta\boldsymbol{s}'$ one arrives at the final expression

$$\begin{aligned}
\Delta\boldsymbol{s}' &= \frac{\rho_2}{\rho_2'}\cdot\Delta\boldsymbol{s} + \frac{\rho_2-\rho_1}{\rho_2'}\cdot\hat{\boldsymbol{s}}_1 - \frac{\rho_2-\rho_1}{\rho_2'}\cdot\hat{\boldsymbol{s}}_1 \\
&= \frac{\rho_2}{\rho_2'}\cdot\Delta\boldsymbol{s} + \left(\frac{\rho_2}{\rho_2'} - \frac{\rho_1}{\rho_1'}\right)\hat{\boldsymbol{s}}_1 + \left(\frac{1}{\rho_1'} - \frac{1}{\rho_2'}\right)\boldsymbol{r}_{\mathrm{obs}} \quad , \qquad (4.61)
\end{aligned}$$

where the relation (4.57) is then applied to replace the geocentric quantities by exploiting

$$\begin{aligned}
\rho_1' &= \rho_1 - \Delta\rho_{\mathrm{par}} \qquad (4.62)\\
\rho_2 &= \rho_2 - \Delta\rho_{\mathrm{par}} \quad , \qquad (4.63)
\end{aligned}$$

and the absolute distance of star 2 is substituted by $\rho_2 = \rho_1 + \Delta\rho$, the sum of the distance to the primary star and the differential distance

$$\Delta\rho = \frac{1}{\varpi_2} - \frac{1}{\varpi_1} = \frac{1}{\varpi_1 + \Delta\varpi} - \frac{1}{\varpi_1} \quad . \qquad (4.64)$$

[6]The interested reader should notice the sign errors in the formula given within the reference.

4.2.3.4. Relativistic Light Deflection

As we have learned from observations and the theory of general relativity, photons follow null geodesics in the spacetime, which is being curved by masses. Consequently, since the light paths are influenced by the mass distribution along the way of the travelling bodies, apparent displacements of the observed background sources need to be considered. If only the light deflection due to the spherically symmetric mass distribution of a body is taken into account, then the total deflection angle ϕ_{def}, where both, the light source and the observer are considered to be situated at infinite distances in opposite directions, is determined only by the mass of the disturbing body and the distance R between its center of mass and the point of the light's nearest approach:

$$\phi_{\mathrm{def}} = \frac{4GM}{Rc^2} \quad . \tag{4.65}$$

Calculating the situation in the extreme situation of a sun grazing light path, the deflection angle becomes $\approx 1.75''$. However, although this kind of observation is not realistic, the same calculation for Jupiter gives a value of $\approx 0.16\,\mu$as.
Clearly, this is still an extreme case, and in reality the disturbing object will usually not be situated between the two observed stars so that the differential effect will be even smaller, but still needs to be considered when it comes to accuracies at the microarcsecond level.
In order to get a feeling for the magnitude of the involved effect, the limiting separation angles between the observed star pair and all potentially deflecting objects within the solar system, for which the magnitude of the differential effect would lie below $0.1\,\mu$as, has been roughly estimated, arriving at results between the highest limit of about 4° in the case of Jupiter and $\approx 3.5''$ for Pluto (see Tab. 4.3).
Based on this analysis, it has been also decided that the ADRS would always apply the correction for the deflection by the Sun, whereas for the other considered objects, as Pluto, the Earth's moon and the solar system planets, the corrections would only be calculated if the angular distance between them and the observed stars' midpoint would fall below the limiting angle of the corresponding object.

The general expression for the deflection angle can be derived to ([Kov04]):

$$\begin{aligned} \phi &= \frac{2GM_\circ}{c^2 R_\circ}\left(\frac{\sin\lambda}{1+\cos\lambda}\right) = \frac{2GM_\circ}{c^2 R_\circ}\tan\lambda/2 \\ &= \frac{2GM_\circ}{c^2 R_\circ}\cot\psi/2 \quad , \end{aligned} \tag{4.66}$$

where $M_\circ$ denotes the object's mass and $R_\circ$ its distance with reference to the observer, and λ is the measured angular separation between the light emitting source and the observer, as seen from the deflecting body. This entity can be translated to the angular separation ψ between the star and the disturbing body, as seen by the observer, approximately by $\psi = \pi - \lambda$.
As the deflection usually occurs in slightly different directions for both stars, in the extreme case in the opposite directions, when the deflecting body is positioned between them, it is important to evaluate also the orientation of the deflection, before the differential effect is estimated.

Table 4.3. – Estimated limiting angles of angular separation between the disturbing bodies and the observed star pair, below which the differential effect could impose displacements above 0.1 μas and for which the ADRS is correcting.

deflecting body	limiting angle [deg]
Mercury	0.138
Venus	0.821
Mars	0.211
Jupiter	4.196
Saturn	1.601
Uranus	0.424
Neptun	0.362
Pluto	0.001
Moon	1.076

The direction of the deflection, which is perpendicular to the position vector of the star, is given by the unit vector:

$$\hat{\boldsymbol{d}} = \frac{\hat{\boldsymbol{s}}_\circ}{\sin\lambda} + \hat{\boldsymbol{s}}\cot\lambda \quad , \tag{4.67}$$

with the unit vectors $\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}_\circ$ towards the light source and respectively the deflecting body.
As the massive body always bends the light towards itself, the star appears at a higher angular separation ψ than it actually is. Consequently, in order to estimate the apparent star position $\hat{\boldsymbol{s}}'$, the deflection is applied on the star unit vector by:

$$\begin{aligned}
\hat{\boldsymbol{s}}' &= \hat{\boldsymbol{s}} - \phi\cdot\hat{\boldsymbol{d}} \\
&= \hat{\boldsymbol{s}} - \frac{2GM_\circ}{c^2R_\circ}\frac{\sin\lambda}{1+\cos\lambda}\left(\frac{\hat{\boldsymbol{s}}_\circ}{\sin\lambda} + \hat{\boldsymbol{s}}\cot\lambda\right) \\
&= \hat{\boldsymbol{s}} - \frac{2GM_\circ}{c^2R_\circ}\left(\frac{\hat{\boldsymbol{s}}_\circ + \hat{\boldsymbol{s}}\cos\lambda}{1+\cos\lambda}\right) \\
&= \hat{\boldsymbol{s}} - \frac{2GM_\circ}{c^2R_\circ}\left(\frac{\hat{\boldsymbol{s}}_\circ - \hat{\boldsymbol{s}}\cos\psi}{1-\cos\psi}\right) \\
&= \hat{\boldsymbol{s}} - \frac{2GM_\circ}{c^2R_\circ}\left[\frac{\hat{\boldsymbol{s}}_\circ - \hat{\boldsymbol{s}}\cdot(\hat{\boldsymbol{s}}_\circ\hat{\boldsymbol{s}})}{1-\hat{\boldsymbol{s}}_\circ\hat{\boldsymbol{s}}}\right] \quad .
\end{aligned} \tag{4.68}$$

The apparent separation vector is then constructed to:

$$\begin{aligned}
\Delta\boldsymbol{s}' &= \hat{\boldsymbol{s}}_2' - \hat{\boldsymbol{s}}_1' \\
&= \Delta\boldsymbol{s} + \frac{2GM_\circ}{c^2R_\circ}\left[\frac{\hat{\boldsymbol{s}}_\circ - \hat{\boldsymbol{s}}_1\cos\psi_1}{1-\cos\psi_1} - \frac{\hat{\boldsymbol{s}}_\circ - \hat{\boldsymbol{s}}_2\cos\psi_2}{1-\cos\psi_2}\right] \\
&= \Delta\boldsymbol{s} + \frac{2GM_\circ}{c^2R_\circ}\left[\frac{\hat{\boldsymbol{s}}_\circ - \hat{\boldsymbol{s}}_1\cdot(\hat{\boldsymbol{s}}_\circ\hat{\boldsymbol{s}}_1)}{1-\hat{\boldsymbol{s}}_\circ\hat{\boldsymbol{s}}_1} - \frac{\hat{\boldsymbol{s}}_\circ - \hat{\boldsymbol{s}}_2\cdot(\hat{\boldsymbol{s}}_\circ\hat{\boldsymbol{s}}_2)}{1-\hat{\boldsymbol{s}}_\circ\hat{\boldsymbol{s}}_2}\right] \quad .
\end{aligned} \tag{4.69}$$

Alternatively, the apparent coordinates (α', δ') of a star due to light deflection can also be

calculated from the real coordinates and some offsets $\Delta\alpha, \Delta\delta$, which take the form:

$$\begin{aligned} \Delta\alpha &= \frac{2GM_\circ}{c^2 R_\circ} \cdot \frac{\cos\delta_\circ \sin(\alpha - \alpha_\circ)}{(1 - \cos\psi)\cos\delta} \\ \Delta\delta &= \frac{2GM_\circ}{c^2 R_\circ} \cdot \frac{\sin\delta\cos\delta_\circ\cos(\alpha - \alpha_\circ) - \cos\delta\sin\delta_\circ}{1 - \cos\psi} \quad , \end{aligned} \tag{4.70}$$

where $\alpha_\circ, \delta_\circ$ are the coordinates of the deflecting object and ψ, the elongation between the star from the body, can also be recovered from the coordinates:

$$\cos\psi = \sin\delta\sin\delta_\circ + \cos\delta\cos\delta_\circ\cos(\alpha - \alpha_\circ) \quad . \tag{4.71}$$

4.2.3.5. Aberration

Due to the combination of the finiteness of the speed of light and any transversal component of the observer's velocity, with reference to the incoming beam, the direction, in which the source is observed, is tilted towards the direction of motion. The resulting displacement of observed stars is called stellar aberration and classically can be subdivided into three components: the annual aberration due to the earth's orbital motion, the diurnal aberration as a result from its rotation, and the secular aberration as a consequence from the stars' motion with reference to the solar system barycenter. Leaving the latter aside, which reflects itself in the measured proper motions and is corrected for by the consideration of the light travel time (see Sec. 4.2.3.2), the remaining aberrational effect is calculated considering the relations between the source's real geometric position $\hat{\boldsymbol{s}}$, the light's velocity vector $-c\hat{\boldsymbol{s}}$ and the observer's velocity vector $\boldsymbol{v}$. In the classical approach the apparent direction towards the source is derived from the addition of the velocity vectors and is given by:

$$\hat{\boldsymbol{s}}' = \frac{\hat{\boldsymbol{s}} + \boldsymbol{v}/c}{|\hat{\boldsymbol{s}} + \boldsymbol{v}/c|} \quad . \tag{4.72}$$

The angle of the displacement $\Delta\theta$ can then be estimated from the scalar part of the cross product $\hat{\boldsymbol{s}} \times \hat{\boldsymbol{s}}'$, yielding ([Kov04]):

$$\begin{aligned} \sin\Delta\theta &= \frac{(v/c)\sin\theta}{\sqrt{(1 + 2(v/c)\cos\theta) + (v/c)^2}} \\ &= \frac{v}{c}\sin\theta - \frac{1}{2}\left(\frac{v}{c}\right)^2 \sin 2\theta + \ldots \quad , \end{aligned} \tag{4.73}$$

where θ is the angle between the direction of the observer's velocity and the star's geometric position so that $|\hat{\boldsymbol{s}} \times \boldsymbol{v}/c| = \frac{v}{c}\sin\theta$.
The expansion in terms of v/c is interesting, as we will see that this classical ansatz already deviates from the relativistic treatment at the level of the second order term.

Within the framework of special relativity, where the correct Lorentz transformations are applied for the addition of the velocities, the corresponding expression for the apparent direction

takes the form:

$$\begin{aligned}\hat{\boldsymbol{s}}' &= \frac{\gamma^{-1}\hat{\boldsymbol{s}} + (\boldsymbol{v}/c) + (\hat{\boldsymbol{s}}\cdot\boldsymbol{v}/c)(\boldsymbol{v}/c)/(1+\gamma^{-1})}{1+\hat{\boldsymbol{s}}\cdot\boldsymbol{v}/c} \\ &= \frac{\gamma^{-1}\hat{\boldsymbol{s}} + (\boldsymbol{v}/c)}{1+a} + \frac{a\boldsymbol{v}/c)}{(1+a)(1+\gamma^{-1})} = \frac{\hat{\boldsymbol{s}}\,(\gamma^{-1}+\gamma^{-2}) + \boldsymbol{v}/c\,(1+a+\gamma^{-1})}{(1+a)(1+\gamma^{-1})} \\ &= \frac{\hat{\boldsymbol{s}}\,(\gamma^{-1}+\gamma^{-2}) + \gamma^{-1}\boldsymbol{v}/c}{(1+a)(1+\gamma^{-1})} + \frac{\boldsymbol{v}/c}{1+\gamma^{-1}} \quad , \end{aligned} \tag{4.74}$$

where $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$ is the corresponding Lorentz factor to the magnitude of the observer's velocity and the substitution $a = \hat{\boldsymbol{s}}\cdot\boldsymbol{v}/c$ has been employed.
Again, the angle of displacement can be estimated from the cross product of the apparent position vector with the true position vector, from which the expansion

$$\sin\Delta\theta = \frac{v}{c}\sin\theta - \frac{1}{4}\left(\frac{v}{c}\right)^2\sin 2\theta + \frac{1}{4}\left(\frac{v}{c}\right)^3\sin 2\theta\cos\theta + \ldots \quad , \tag{4.75}$$

can be derived ([Kov04]).
Comparing this result to equation (4.73), one can see that there is a difference of $\frac{1}{4}\left(\frac{v}{c}\right)^2$ in the second order. As this difference yields around 500 μas of displacement in the extreme case of $\theta = 45°$ and the earth's mean orbital velocity of $v \approx 30\,\mathrm{km/s}$, the relativistic formula is implemented into the ADRS pipeline.

The apparent separation vector, which naturally is time dependent and will be evaluated for the time stamps of the "Level 1" data, is finally calculated by applying Eq. (4.74) to both stars, which gives:

$$\begin{aligned}\Delta\boldsymbol{s}' &= \frac{(1+\gamma^{-1})^{-1}}{(1+a_2)(1+a_1)}\left[(\gamma^{-1}+\gamma^{-2})\left[\Delta\boldsymbol{s}(1+a_1) + \hat{\boldsymbol{s}}_1(a_1-a_2)\right] + \frac{\boldsymbol{v}}{c}\gamma^{-1}(a_1-a_2)\right] \\ &= \frac{(1+\gamma^{-1})^{-1}}{(1+a_2)(1+a_1)}\left[(\gamma^{-1}+\gamma^{-2})\left[\Delta\boldsymbol{s}(1+a_1) - \hat{\boldsymbol{s}}_1(\Delta\boldsymbol{s}\cdot\frac{\boldsymbol{v}}{c})\right] - \frac{\boldsymbol{v}}{c}\gamma^{-1}(\Delta\boldsymbol{s}\cdot\frac{\boldsymbol{v}}{c})\right] \end{aligned} \tag{4.76}$$

with $a_1 = \hat{\boldsymbol{s}}_1\cdot\boldsymbol{v}/c$ and respectively $a_2 = \hat{\boldsymbol{s}}_2\cdot\boldsymbol{v}/c$.

Just as the positions of the earth and the other solar system bodies is recovered from the JPL ephemerides, so is the precise velocity of the earth that, together with the observatory's position on the earth's surface, is necessary to assemble the final observer velocity $\boldsymbol{v}(t)$.

4.3. Analysis of Commissioning Data

4.3.1. Fringe Tracking Performance

During the PRIMA commissioning phase, the system's particular modules have been sequentially integrated and tested on the site. Besides the fundamental functionality of the individual components and their cooperation and communication among each other, the instrument's overall performance is defined by two major aspects: firstly and obviously by the resulting overall precision of the astrometric measurements, but more elementary by the fringe tracking capability, regarding on the one hand its efficiency, which directly affects the brightness space of possible science targets, and on the other hand the accuracy of the tracking algorithms, which are limited by systematic variances of the estimated quantities and consequently can limit the astrometric accuracy, if not calibrated or corrected for.
In order to face some aspects of these questions, fringe tracking sequences from commissioning runs 13 until 17, spanning a time period of 13 months between November 2010 and November 2011, have been examined in the course of this work. The emphasis was laid on the tracking performance with respect to environmental parameters, such as the coherence time and the seeing, but also on the dependence on the targets' brightness.
The exact numbers of the analyzed sequences are summarized in Tab. 4.4. Altogether, 934 files have been taken into account, where a limiting lock ratio of 0.01 % on the main delay tracking controller has been used for the preselection from all available data. From these files 451 dual-feed fringe tracking recordings could have been extracted so that summed up 1,385 tracking sequences, from either main delay or differential delay tracking, could have been collected.

Table 4.4. – 1,385 fringe tracking sequences from the commissionings 13 until 17 have been analyzed, of which 934 represent the data from the FSU connected to the main delay tracking controller and 451 represent fringe data from the diff. optical path delay controller, tracking the secondary object. The K magnitude of the observed objects could not be recovered and is not available for 16 tracking sequences.

	Controller		Tracking Unit		K magnitude					
	OPDC	dOPDC	FSU A	FSU B	1–2	2–3	3–4	4–5	5–6	NA
Comm. 13 (Nov. 2010)	343	59	104	298	4	12	337	35	14	-
Comm. 14 (Jan./Feb. 2011)	113	77	93	97	11	-	92	76	-	11
Comm. 15 (Jul. 2011)	143	83	98	128	-	19	118	57	27	5
Comm. 16 (Aug. 2011)	83	54	63	74	-	2	40	14	81	-
Comm. 17 (Nov. 2011)	252	178	210	220	-	11	348	71	-	-
Comm. 13–17	934	451	568	817	15	44	935	253	122	16
	= 1385		= 1385		= 1385					

The targets' brightness, in units of K magnitudes, has been extracted from the fits files headers, whenever provided, and otherwise taken from the corresponding entries of the 2MASS point

source catalog ([Skr06]). Brightness information on the secondary component of the observed binary could not be recovered for 16 sequences in dual-feed mode.

The results of this analysis are displayed in figures 4.9–4.11 and show relations between quantities estimated by the FSUs and the environmental conditions.
Both, the coherence time and the FWHM of the seeing disk are taken from the information of the files themselves, which provide two values for each of them, one for the point in time at the beginning of data recording and respectively one for the end. The means of those have been used here for each sequence.
The lock ratio was derived from the status data of the tracking controller in question and gives the fraction of time, during which the controller is in either of the two states TRACK or IDLE, meaning that the tracking loop is closed or only shortly interrupted (comp. to Sec. 4.1.3.3), compared to the entire time of the observation.
Since by principle the secondary source is tracked only when the phase reference is given by the closed loop on the primary, for the lock ratio of the dOPDC this aspect was accounted for by relating the time of tracking not to the full time of data recording, but only to the time intervals of tracking OPDC.
The variations of the phase measurements, as discussed here, describe the stability of the estimates on short time scales. Therefore, the phase data from the fits files was averaged onto a one second time grid and the standard deviations over these averages have been computed. In turn, the average value of these deviations represents then the mean one second standard deviation σ of the phase estimates for each tracking sequence.
The same routine is applied for the mean 1-second standard deviation of the tracking controller's *real-time offsets*.
Driven by the according FSU, the real-time offset (RTO) is computed by the tracking control routine and gives the corrections that are sent to the delay line to account for atmospheric turbulences or fringe motions of any kind. As it does not include the blind trajectory computed from the used OPD model, it serves as the setpoint of the correcting delay line in relation to the actual position. The RTO variations consequently illustrate the amount of correction that has been applied during the observation to compensate the atmospheric piston.
In Fig. 4.9 the lock ratio, as well as the variations of the measured phases and the RTOs, are plotted against the coherence time, where distinctions have been made for the two distinct FSUs, but also targets of different brightness.
Although fringe tracking with the PRIMA FSUs has been already demonstrated for dimmer sources, the targets of the sequences in this analysis span only the interval in K magnitude between $K_{\min} = 1.329$ and $K_{\max} = 5.751$.
High lock ratios are clearly achieved over the whole interval of recorded coherence times, although it seems that there is a tendency towards higher values, as would be expected with improved conditions and consequently increased τ_c.
Obviously, there is a threshold value ($\tau_c \approx 3\,\mathrm{ms}$), under which good tracking seems to be achieved only with difficulties and cannot be guaranteed.
There is no clear evidence of different behavior towards various source brightness, at least not within this magnitude interval.

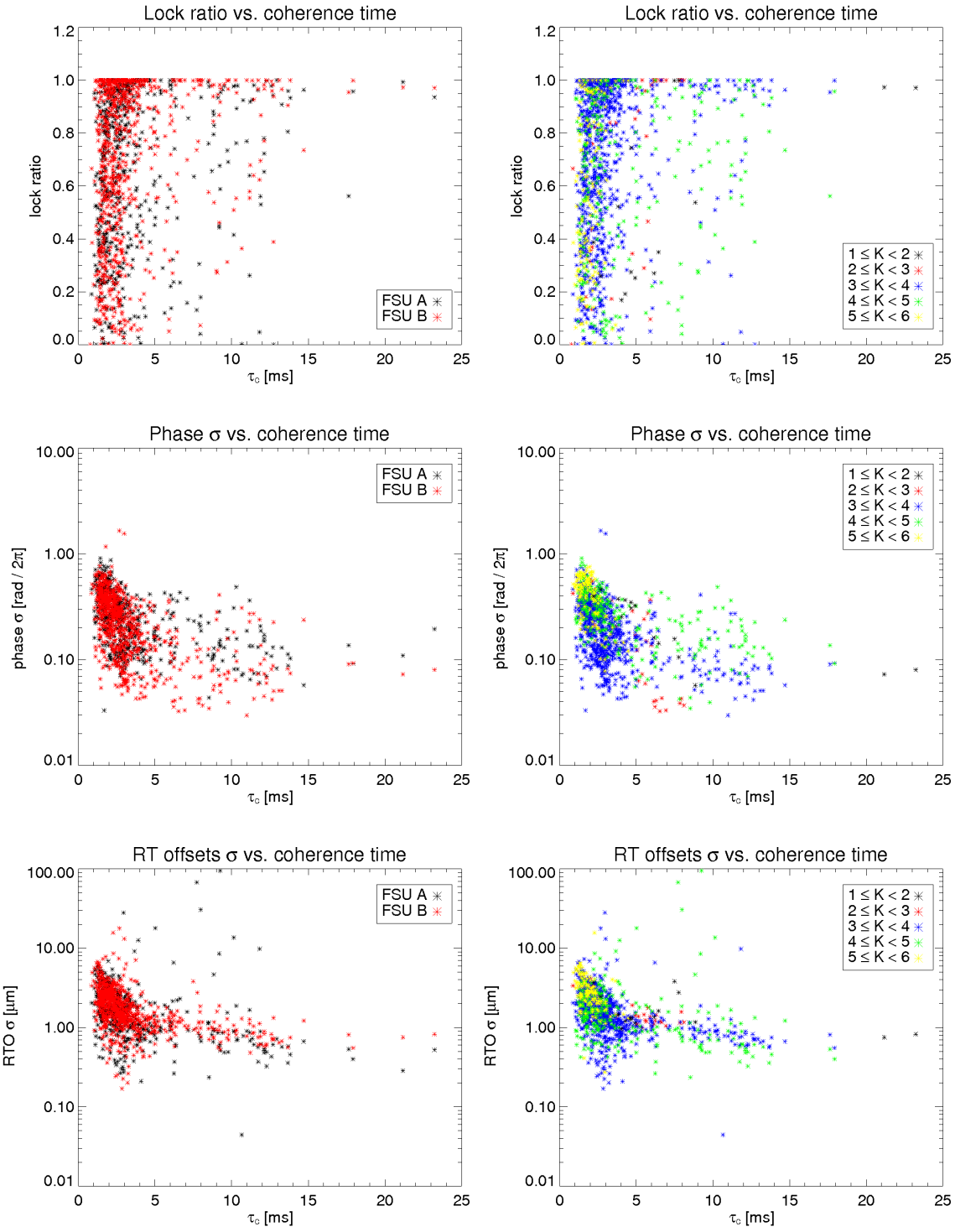


Figure 4.9. – Distribution of 1,385 tracking sequence samples regarding the lock ratio (*top*), the standard deviation (over 1 second) of the estimated phases (*middle*) and the standard deviation (over 1 sec.) of the real-time offsets (*bottom*) against the coherence time. Within the left column a distinction regarding the two FSUs is color coded, whereas the colors in the right hand side plots represent different intervals in the targets' K magnitude.

This is different when investigating the variations of the phases and the RTOs. Although both show also low values at bad conditions, implied by low coherence times, a clear scatter up to maxima of $\sigma_\varphi \approx 2\pi$ and $\sigma_{\mathrm{RTO}} \approx 8\,\mu$m is visible but decreasing with higher τ_c in the interval up to $\tau_c \approx 5\,$ms, where the quantities seem to settle.
This is particularly evident for the RTOs, which for higher coherence times seem to settle around 0.7–0.8 μm, with no more obvious dependency on the sources' brightness, which on contrary is clear at low coherence times.

In bad conditions the star's brightness seems to play an important role for the RTOs, as well as for the estimated phases. In fact, the phase estimates seem to be particularly sensitive to this parameter throughout the whole coherence time spectrum with standard deviations below $\sigma_\varphi = 0.2\pi$ for sources brighter than $K = 4$ and above for dimmer targets.
No clear difference is observed between the different FSUs. While a slight hint of different behavior could be suspected from the plot of the phase variations against the coherence times, with lower values for FSUB and increasing τ_c, a comparison with the brightness dependent plot uncovers the correlation there between the FSUA and dimmer sources.
This is understandable, as in the NORMAL operation mode the brighter source is tracked by FSUB and the secondary source by FSUA, and within this sample 295 of the 451 dOPDC sequences are performed with FSUA, which is nearly double as much as with FSUB.

An interesting feature arises, when the same quantities are plotted with a distinction, whether the source was tracked by the OPDC on the main delay or on the differential delay line by the dOPDC (see Fig. 4.10).
While for the phase variations, which happen to be lower at the OPDC, it is not clear, if this again only mimics the brightness dependency, as a correlation seems plausible from the comparison, the RTO standard deviations seem to be decreased for the dOPDC sequences, especially at rather bad conditions with low coherence times below 5 ms. This is also suggested by the graphs showing the RTO variations against the lock ratio and in particular when plotted against the phase variations.
Besides the obvious correlation between these two quantities, a bimodal structure is clearly apparent and features lower RTO variations when tracking with the differential delay line.
This finding is supported by the plot depicting the different brightness intervals, as the lower RTO variations correlate with dimmer targets, as more often tracked by the dOPDC.
This reflects the principle of the phase reference, provided by the tracking on the primary target, and shows the damping of the corrections that need to be sent to the tracking differential delay line, in this case within time intervals of one second.

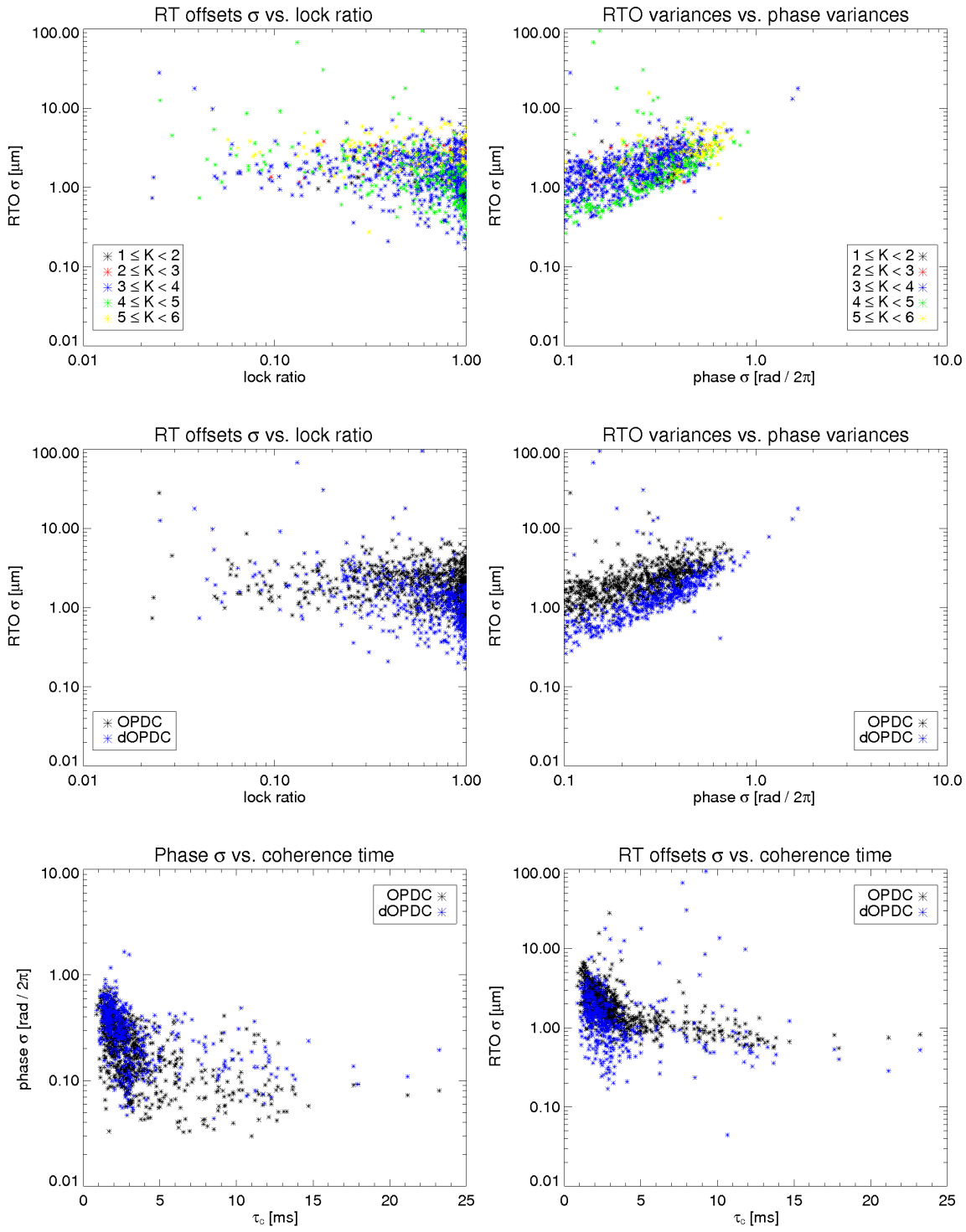


Figure 4.10. – The upper two rows show the relationship between the standard deviations over 1 second of the real-time offsets, the phases and the lock ratio, where the target brightness (*top*) and the corresponding tracking controller (*middle*) are color coded. The same distinction between main delay line tracking and observations of the secondary source, during which the diff. delay line tracking is controlled by the dOPDC, is made for the plots of the variations against the coherence time (*bottom*).

The sample was also explored for clear correlations of the quantities of interest to the observatory seeing during the observation, which is depicted in the plots in Fig. 4.11.

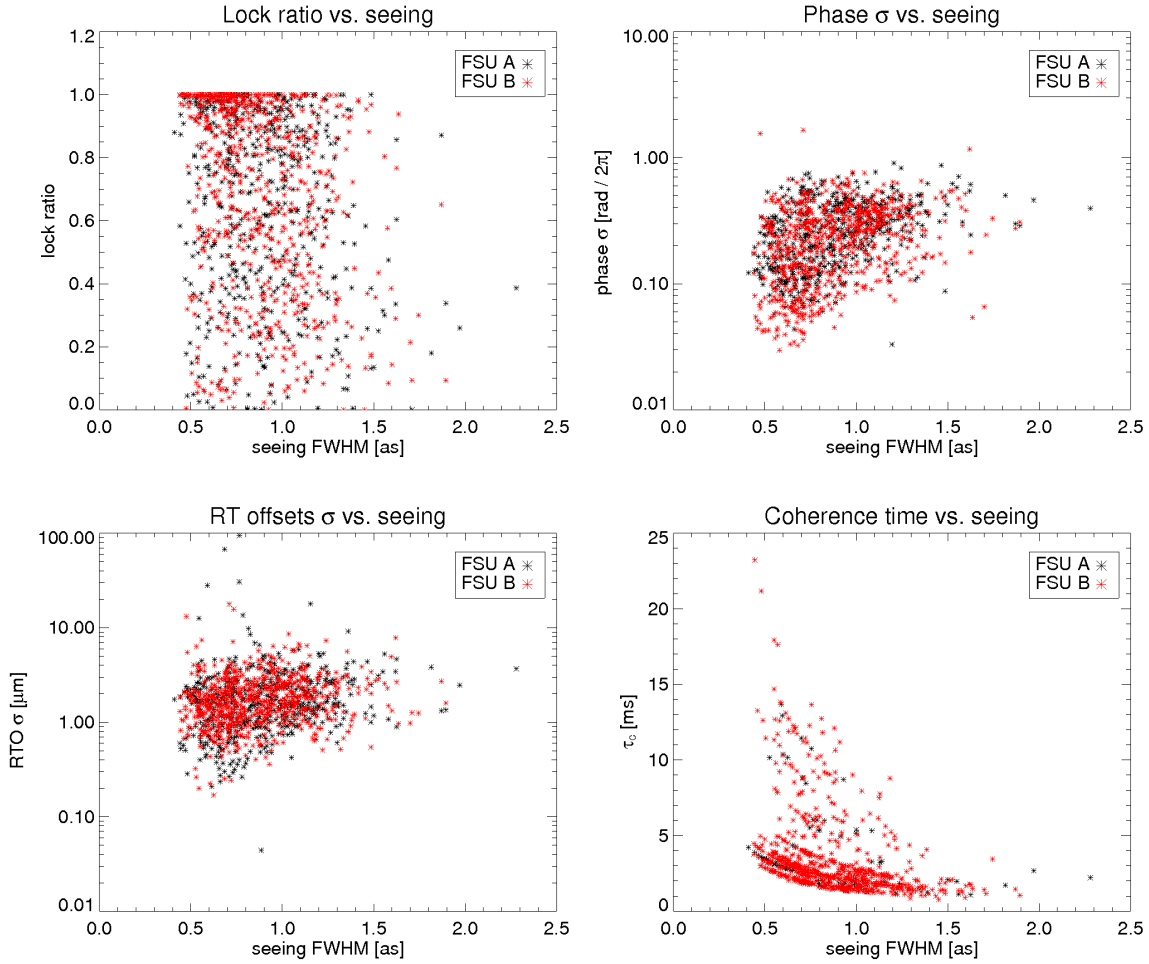


Figure 4.11. – The plots depict the relations between the seeing during the observation and the lock ratio, as well as the 1-second standard deviations of the estimated phases and the real-time offsets as derived from the sample of 1,385 tracking sequences. The correlation between seeing and the coherence time is also illustrated (*bottom right*). Color coding accounts for the distinct fringe sensor units.

The real-time offsets variations, as well as the phase deviations over one second, show a tendency towards higher values for increased seeing, which is more evident for the latter.
At good seeing, down to the seeing disk FWHM of around 0.5 as, phase variations over the whole spectrum between 0.06π and 1.6π are found. With degrading conditions, this scatter is more and more confined to the high end of the initial spread. A similar structure, although not that clear, is observed regarding the RTO variations. An explanation for, why the dependency on the seeing is not as evident, as one might suspect, can be found in the correlation

of the recorded coherence times against the seeing. Although τ_c principally decreases with the seeing increasing, and extraordinary long coherence times are found only in good conditions, it is also evident that quite low values, even below the previously discussed critical threshold of $\tau_c \approx 3\,\mathrm{ms}$ are recorded throughout the entire band of obtained seeing disk diameters from 0.5 as to 2.0 as. This scatter in coherence times is then again reflected in the RTO and phase variations, as those show a clear dependency on τ_c.

The same is supported by the plot of the lock ratios versus seeing. The scatter makes it difficult to give a secure statement, but the graph seems to indicate that observations with maximum or nearly optimal tracking ratios of around 100 % become less frequent above some values of around $\sim$1.1 as, and none is observed above FWHM = 1.5 as.
However, since a major fraction of the sequences have been recorded at better conditions than such that might be implied by these limiting values, these features might also be some artifact due to the sample bias.

Overall, the results on the variations of the estimated phases seem promising, regarding the tracking availability, especially in good conditions and for long coherence times. Within the sample of 1,385 tracking sequences the lowest value obtained is $\sigma_{\varphi,\mathrm{MIN}} = 0.06\pi$ and the median is given by $\sigma_{\varphi,\mathrm{MED}} = 0.48\pi$, which translates to less than 1/4 of the observation wavelength and to $\sim 0.53\,\mu$m in the K-band, in which PRIMA is operated.

Still, as will be shown in the following, this is no implication on the ultimate accuracy of the fringe tracker. Two fringe tracking sequences, controlled by the OPDC and measured by FSUB, with nearly optimal lock ratios above 0.99 are depicted in Fig. 4.12.
Naturally, as the tracking algorithm tries to keep the delay at the zero group delay position, the plotted low-pass filtered GD still shows some fluctuations, but essentially maintains the zero level, and the OPDC state is mostly at 7, which corresponds to the state TRACK.
Contrarily, the estimated phases, also low-passed filtered here for the clarity of the plot, do not keep the same level, but exhibit phase jumps with a phase difference of $\Delta\varphi = \pi$.
More than this, when the phase estimates, as provided by the fringe sensor data only within the interval $[-\pi, \pi]$, are unwrapped, which has been performed here in a standard procedure of tracking and correcting phase jumps of more than 1.2π in any two consecutive readings, one can see that even bigger phase jumps can occur and that the phases settle on different levels, which are all separated by $\Delta\varphi = \pi$ from each other.
These features seem suspicious, especially since they are observed within sequences of good tracking. In those situations the tracking state itself should be sufficiently reliable in the sense that the controlling loop is in fact closed, while the delay follows the fringe position.
This conclusion can be drawn when the signal-to-noise ratio distribution is inspected.
Fig. 4.13 shows normalized distributions of recorded SNR estimates (see Eq. (4.10)) during observation sequences for points in time when the OPDC was tracking, as well as when the loop was open and the system performed a fringe search.
Typically, in good and stable situations, when enough light reaches the FSU, the finding of a fringe can be reliably identified by the SNR measurement.

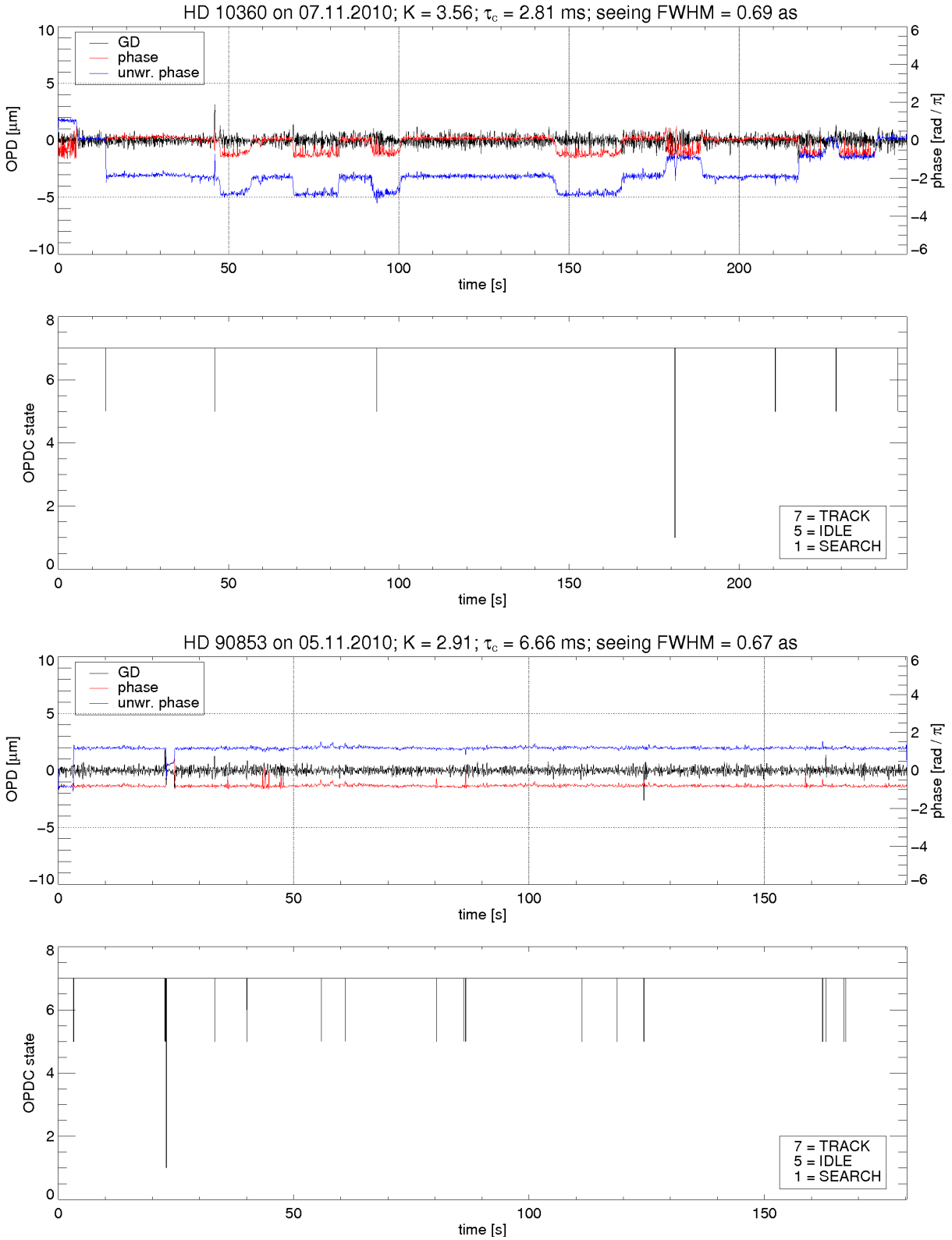


Figure 4.12. – Two tracking sequences with nearly optimal lock ratio illustrate the low-pass filtered real-time estimates of the GD in the OPD space and the phases. Whereas the GD is tracked to zero, the phases and particularly the unwrapped phases exhibit several levels, separated by $\Delta\varphi = \pi$. Also plotted is the state of the OPD controller, where the value 7 represents the closed tracking control loop.

If the thresholds for the OPD controller are chosen so that a high lock ratio is achieved, the estimates during the tracking are normally distributed.
Furthermore, although the distribution can be wide, it is still well distinguishable from the noise dominated readings, when the fringes are lost by the tracker. This situation changes dramatically for observations, during which mostly just low SNR values are obtained.
In those cases the fringes become less identifiable and consequently high lock ratios cannot be achieved. This is also apparent in the relatively high overlap between the distributions of SNR recordings while tracking and while searching, which is the zone, in which the fringes are lost by the instrument.

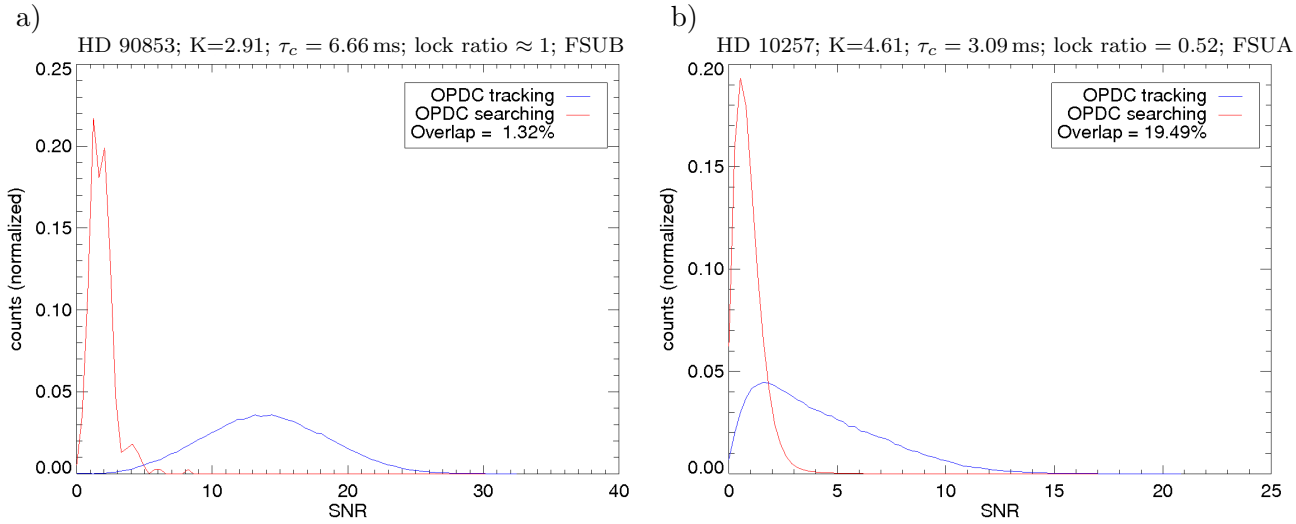


Figure 4.13. – Typical normalized SNR distributions during fringe tracking sequences with nearly optimal lock ratio (*left*) and in conditions, which resulted in significantly lower tracking efficiency (*right*). Colors distinguish estimates while the tracking loop was closed from those while the OPDC was in the SEARCH state. In good observing conditions the fringes are identified by SNR readings that are clearly distinguishable from the noise.

However, although the tracking state of the controller can be trusted, particularly for high lock ratio sequences, the phase jumps described above put the position of the tracker in OPD space in question, and it cannot be guaranteed that the tracking is performed on the central fringe. This problem has been identified and studied also by [Sah10].
In that analysis the nonlinearities of the estimated OPDs, as derived from the measured phases, and of the determined GD have been investigated and found to be significant, periodical and increasing when moving away from the zero OPD. In particular, the nonlinear behavior of the group delay estimation is problematic, as the fringe tracking algorithm, underlying the OPDC, tries to keep the fringes at zero GD.

In theory, the group delay exhibits only one well-defined zero-crossing in the OPD space, but if disturbed strongly enough, several local extrema can arise, and in the worst case, multiple zero

points at distinct OPD positions, which thereby would act as stable attractors to the tracker, could emerge.
The reasons for those cyclic nonlinearity errors can be multifarious, as the real-time estimates are derived from the ABCD signals, and therefore anything that influences those also affects and disturbs the tracking control.
However, two major error sources, namely differential photometric effects and non-nominal phase offsets between the ABCD signals have been identified to be most likely responsible for the observed nonlinearities. Uncalibrated inequalities in the detected intensities between the channels of the different detector quadrants lead to cyclic phase errors with a periodicity in the OPD space that equals that of the light's effective wavelength. Although calibration routines are performed beforehand, any flux variabilities during the observation itself, which could be consequences of changes in the instrument's overall transmission or high-frequency injection fluctuations due to the beams' tip-tilt jitter in combination with non ideal alignment or wavefront perturbations in general, will degrade the measurement and the functionality of the fringe tracker, as no real-time photometric monitoring is implemented.
On the other hand, phase offsets of the ABCD signals from their nominal values lead to periodic phase errors in delay space, which are characterized by half of the effective wavelength, at which the measurement is made. These offsets can be calibrated if known (see 4.3.1.1), but any deviations from those calibrated values, which for example can occur due to polarization variabilities in the course of the observation, during which, amongst others, the telescope pointing and the angle of the star separators vary, will again have an impact on the linearity of the OPD and GD determinations.

4.3.1.1. Phase Offset Deviations

While the applied spatial phase modulation technique and the ABCD method serve particularly well for the fringe tracking efficiency, regarding the rate, at which the fringes and consequently the fringe position can be sampled, the fringe encoding depends on the estimated phases and accordingly is affected by the phase relations between the ABCD signals.
Consequently, any deviations from their nominal values of $\pi/2$ disturb the measurement and result in errors in the phase estimation itself, but also in the determination of the OPD and the group delay, which can also lead to such nonlinear effects, as described in the previous section.
In order to stress the importance of a proper calibration of such phase offset deviations, an initial analysis has been carried out to investigate their occurrences especially during observations on the sky. Just as the effective wavenumber and the relative phases between the ABCD signal of the different detector quadrants can be obtained in laboratory by the means of Fourier spectroscopy when an OPD scan is performed over the entire fringe package from the laboratory source, fringe package scans on the sky offer the opportunity to derive the offsets in more realistic conditions.
For this purpose, the tracked primary source serves as usual as the phase reference, and fringe data is recorded while the differential delay line performs consecutive scans around the fringe center of the secondary source.
Such a signal is shown by Fig. 4.14, where the measured SNR is color coded and depicted with

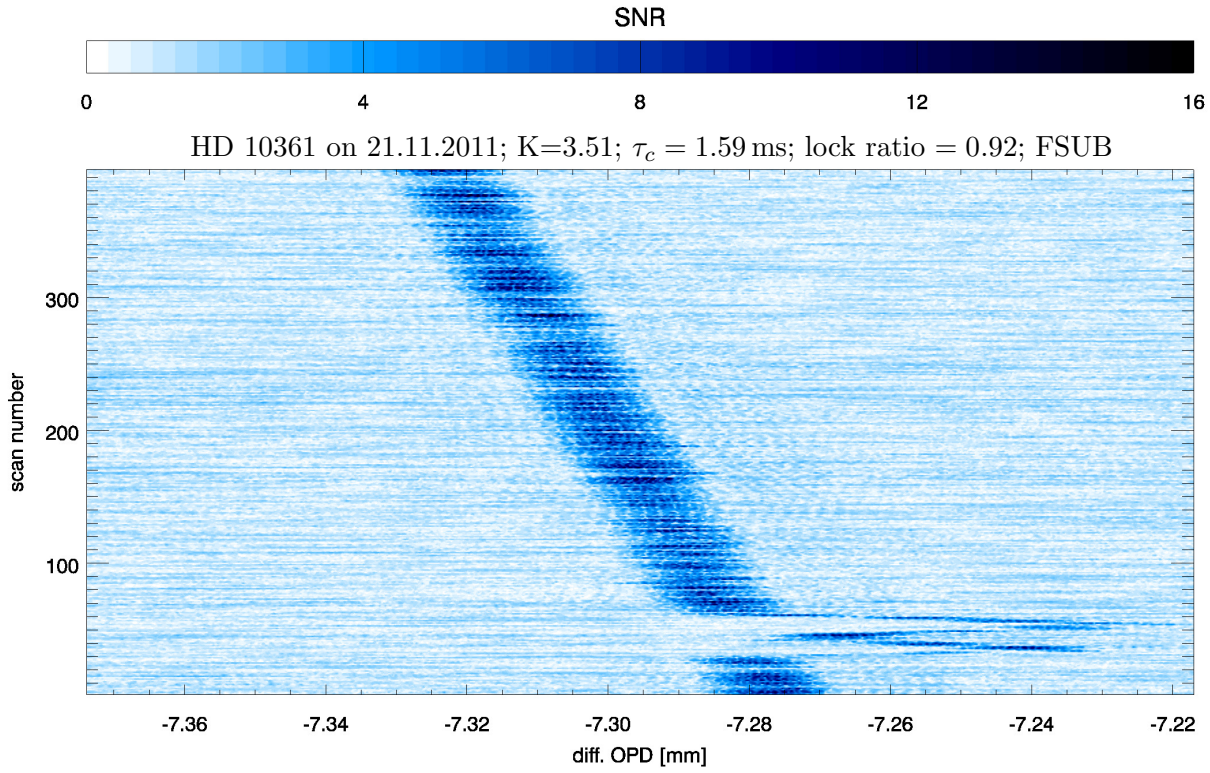


Figure 4.14. – The SNR estimates during an observation in SWAPPED mode with FSUB recording fringe data on the secondary source HD 10361 (K=3.51), while the DDL was performing 400 consecutive scans over a time period of around 243 seconds. A lock ratio of about 90 % on the primary source was achieved and the loss of the fringes is clearly visible, approximately between scans 30 and 60, where the fringe center location is varying or even wiped out. The visible drift in the OPD space is caused by the earth (baseline) rotation over the four minutes of observation.

reference to the individual scans and the applied differential OPD, as given by the differential PRIMET metrology.
In this case data has been recorded over a time span of ∼243 s, during which 400 dOPD scans have been carried out. With the achieved sampling intervals a clear distinction between the fringe signal and the uncorrelated noise can be made by the signal-to-noise ratio estimates. Also apparent is the fringe center's drift over time in the OPD space due to the rotation of the baseline and the accordingly varying projection of the baseline vector onto the separation vector of the two observed sources on the sky, as well as an incidence of a fringe loss on the primary source, approximately during scans 30 to 60.
During this period of time, the fringe tracking loop was open and the main OPD controller performed its fringe search, which resulted in variations of the fringe positions on the secondary source and even partial smearing out of the fringe location.

Data from files, which have been recorded in this manner during the commissionings 11–13, has been processed for the here presented analysis of the phase offsets between the A, B, C and D quadratures. For this purpose the relative phases have been determined for each available scan, wherefore the cross spectral density

$$S_{UV}(\kappa) = U^*(\kappa) \cdot V(\kappa) \quad , \tag{4.77}$$

is computed first, where $U(\kappa), V(\kappa)$ denote the Fourier transforms in OPD space of the two signals, for which the correlation is performed.
From this the relative phase is then obtained by applying

$$\phi = \tan^{-1} \frac{\sum_i \mathrm{Im}\left[S_{UV}(\kappa_i)\right]}{\sum_i \mathrm{Re}\left[S_{UV}(\kappa_i)\right]} \quad , \tag{4.78}$$

which in principle is the standard formula, but with the real and imaginary parts of the spectrum summed over all wavenumbers κ.

Naturally, the conditions during the measurements, as well as the scanning routine in terms of the number of scans, the sampling rate, the scanning velocity, etc., can differ from observation to observation, particularly in the course of those commissioning runs, during which this procedure was rather experimental and still to be optimized. Due to that, the quality of the scan data varies significantly not only between the individual scans, but also between the different observations, which makes an up-front distinction in terms of quality difficult. For this reason, all scanning files have been processed, and all available scans have been taken into consideration. A quality control driven first inspection has been performed before the offset computation in order to discard presumably bad data.
This was achieved by considering the SNR as the first indicator. A gaussian fit to the SNR recordings over each scan has been performed, followed by the exclusion of scans, for which the peak signal-to-noise ratio was less than 0.6 times the maximum peak SNR of all scans within one file.
Furthermore, the scan velocities were determined from the timestamps and the metrology data, and those scans, for which the velocity departed from the average value of the corresponding file by more than 0.1 times the mean velocity, have been discarded.
The relative phases for all combinations of the four quadratures and in each spectral channel have been then evaluated from the remaining scans.
As experience has shown that this approach still leads to quite unexpected results in some cases, those have been compared to the raw signals themselves, as well as to their Fourier transforms, which has revealed that in some cases bad signal sampling undesirably affects the computation of the Fourier transforms. In order to account for this, at least rudimentarily, the effective wavenumbers have been estimated from the spectra of the synthetic white channels for further rejection of scans, for which those wavenumber estimations yielded results outside of the interval in wavenumber between 250,000 m^{-1} and 600,000 m^{-1}.
Although this is a rather loose condition, when compared to the operating waveband of the instrument, it turned out to serve sufficiently well as a control parameter for intercepting scans with obviously misleading products.

The distributions of the results of the ultimately remaining scans have eventually been analyzed, as they might indicate the variability of phase relations during on-sky observations.
The corresponding histograms for phase offsets between the A and B, B and C, as well as respectively the C and D signals, together with the gaussian fits to the distributions for the entire sample (comm. 11–13) are depicted in Fig. 4.15 for the white channel, whereas the complete results, distinguished by the individual commissioning runs and spectral channels, are summarized in Tab. 4.5.

Table 4.5. – The table summarizes the results from the phase offset analysis of scanning data on secondary sources during the commissioning runs 11–13. For each commissioning the number of scans and the peak position (in degrees) of the gaussian fits to the phase offset distributions are listed for all channels and both FSUs separately, as well as combined. The effective wavelength of the spectral channels increases with the channel number.

	FSUA			**FSUB**			**FSUA + FSUB**		
	A-B	B-C	C-D	A-B	B-C	C-D	A-B	B-C	C-D
Comm. 11 (Jul. 2010)	N = 216			N = 3807			N = 4023		
white ch.	99.04	71.36	87.70	111.87	74.10	112.61	111.33	74.01	112.48
ch.1	87.69	92.81	64.94	81.47	113.85	58.48	81.83	113.26	58.56
ch.2	99.93	90.75	94.52	107.83	84.60	101.25	107.31	84.82	101.06
ch.3	109.17	80.62	106.27	119.45	74.12	115.50	118.98	74.33	115.34
ch.4	121.56	68.13	121.70	130.05	66.00	128.21	129.63	66.12	128.06
ch.5	117.09	43.04	76.25	132.75	60.69	137.88	132.48	60.55	137.86
Comm. 12 (Sep. 2010)	N = 43			N = 1626			N = 1669		
white ch.	98.93	81.62	121.51	111.84	69.93	111.35	111.21	70.40	111.68
ch.1	80.65	108.89	79.73	85.18	106.80	57.43	84.91	106.86	57.84
ch.2	96.32	89.26	114.84	111.57	82.40	102.29	111.04	82.71	102.59
ch.3	103.18	84.00	121.35	122.00	73.48	113.88	121.55	73.87	114.11
ch.4	116.70	73.62	133.42	135.22	63.20	127.96	134.88	63.52	128.16
ch.5	122.84	64.88	142.19	136.54	58.17	139.16	135.83	58.41	139.35
Comm. 13 (Nov. 2010)	N = 1063			N = 599			N = 1662		
white ch.	99.50	102.43	96.33	108.49	80.01	112.77	100.71	93.54	100.17
ch.1	110.80	92.12	104.67	81.10	123.91	61.60	110.04	103.12	104.58
ch.2	106.46	100.05	104.39	104.98	92.02	102.26	106.18	96.24	103.79
ch.3	100.71	108.47	97.52	116.32	80.22	116.00	103.32	99.41	101.52
ch.4	97.66	111.70	95.90	128.20	67.48	129.01	97.77	111.64	95.92
ch.5	88.01	115.14	90.62	131.09	59.69	136.84	88.02	115.04	90.64
Comm. 11–13	N = 1322			N = 6032			N = 7354		
white ch.	99.43	100.26	95.92	111.65	74.06	112.45	108.47	74.49	110.53
ch.1	110.19	92.78	104.22	82.08	113.96	58.67	85.71	110.18	58.86
ch.2	105.60	98.28	103.99	108.22	85.09	101.54	107.19	86.81	102.01
ch.3	101.78	106.36	98.25	119.63	74.75	115.25	117.32	75.60	114.39
ch.4	97.74	111.56	96.02	130.88	65.52	128.24	130.45	65.77	128.17
ch.5	88.11	114.98	90.57	133.18	60.01	138.02	132.79	60.01	138.05

Altogether, 7,354 scans (1,322 on FSUA and 6,032 on FSUB) have been taken into account. However, while the results point out some implicating trends, they should also be treated with a proper amount of caution, due to their statistical character on the one hand, but also due to their hardly predictable dependence on some known issues that are not accounted for by the applied way of computation.
For instance, as has been shown before in Sec. 4.1.3.2 (comp. Tab. 4.1 and Fig. 4.7), some chromatic dispersion is apparent between the responses of the detector pixels of different quadrants but same spectral channels, which is not accounted for here, and in principle makes even the definition of a phase relation difficult, as the carrying signals slightly differ in wavelengths. Furthermore, overall effects and biases from varying signal coverage by the data sampling, different sampling rates or not excluded grid irregularities, are difficult to retrace.
For those reasons, statistical perspective onto the data has been chosen here as a proper method, although some features, as for example the widths of the obtained phase offset distributions, probably do not reflect the pure variance of the instrument's measurements, but are also affected by the response of the computational method to the imperfect data sets.
With these constraints in mind, some implicative results can be pointed out. Considerable deviations from the nominal values of 90° in phase shifts between A/B, B/C and C/D throughout the white channel are evident up to in general around ±20°, but also an offset of more than 30° is found for C/D on FSUA in the commissioning 12 data, although the sample might be less reliable due to the low scan number (N=43).
Also significant is the dispersive character of the obtained phase shifts, as is indicated by the data and plotted for both FSUs in Fig. 4.16.

The spectral dependence is more evident on FSUB, with a drift of the mean offsets up to even 80° over the whole waveband, but is also clearly identifiable within the FSUA data.
With increasing channel number, respectively wavelength, the phase offsets between the two quadratures, whose beams differ in their polarization states, as they are separated by design by the polarizing beam splitter (A/B, C/D), increase also for FSUB, whereas a decrease is visible on FSUA.
This might be a clue to the source of these phase offset errors, in particular when compared to the offsets between A, C and B, D, which in principle should not be affected by polarization effects, as their phase relation should be the consequence of the 50/50 beam cube splitter only. Although those also feature significant deviations from their nominal values of 180°, they are stable over the spectral channels.
This is a strong indication that polarization effects somewhere along the beam train, down to the polarizing beam splitters, might be one major reason for the undesirable phase shifts. The validity of this conclusion is also strengthened by the results of a prior offset analysis on fringe scans in the laboratory, which also has shown spectral dispersion, but at a clearly weaker level ([Kam10]).

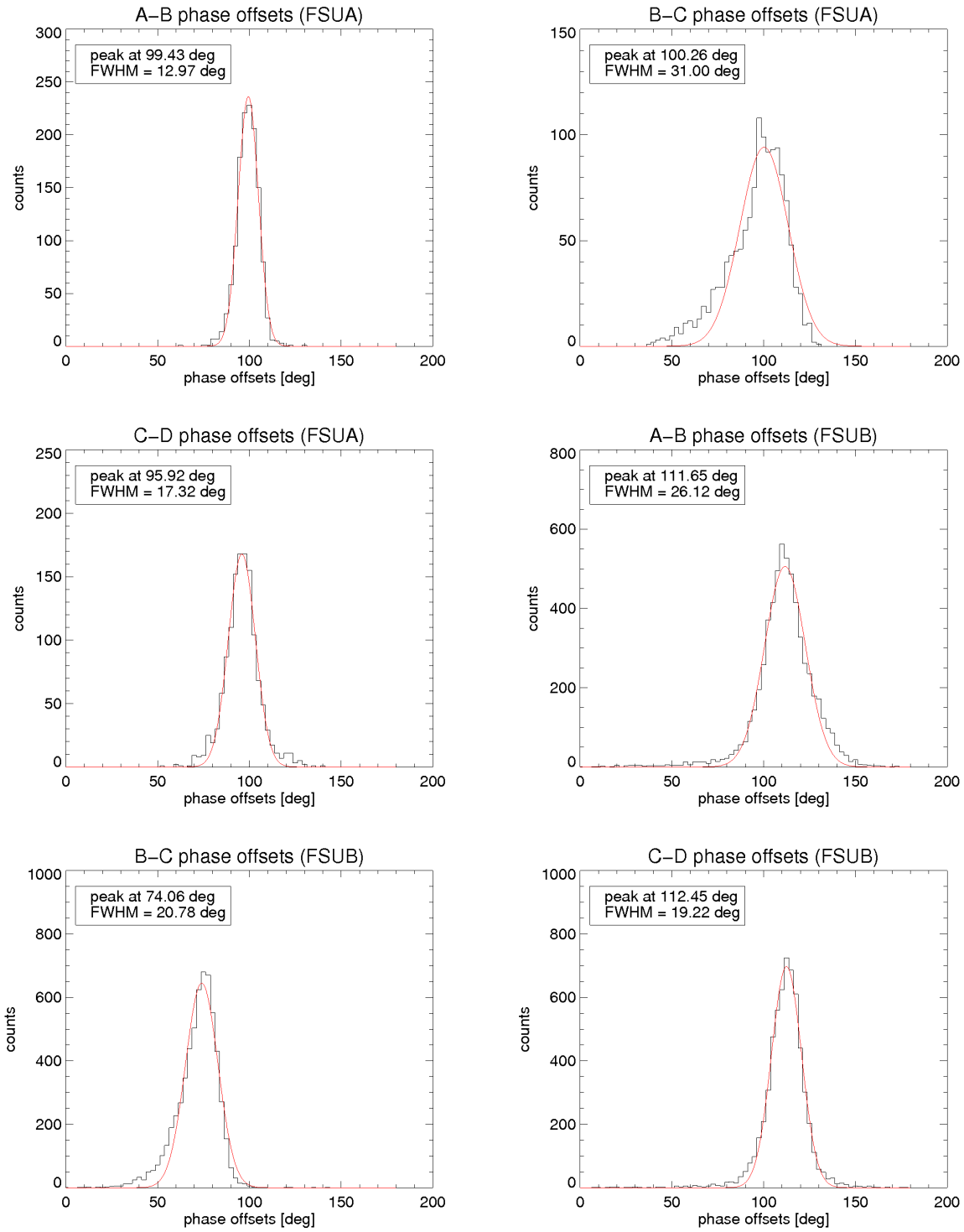


Figure 4.15. – Phase offset distributions on the white channel and the corresponding gaussian fits to the entire sample of 1,322 scans on FSUA and 6,032 on FSUB from the analysis of the scanning data from commissioning runs 11–13. The results have been distributed onto a grid with a binning width of 2.5°. For both fringe sensor units the gaussian fits to the data show considerable variability and not negligible overall deviations from the by design encoded nominal 90°(equiv. $\pi/2$) values.

As a matter of fact, this result is not notably astonishing, as the beams pass many optical elements before recombination within the fringe sensor units.
Changing observational parameters, in particular the source direction on the sky, require changes to the configurations of the entire system, such as the pointing of the telescopes, the position angles of the star separators, the choice of the main delay and differential delay lines and others so that variations of the beams' polarization states cannot only be not ruled out, but on the contrary appear rather probable.

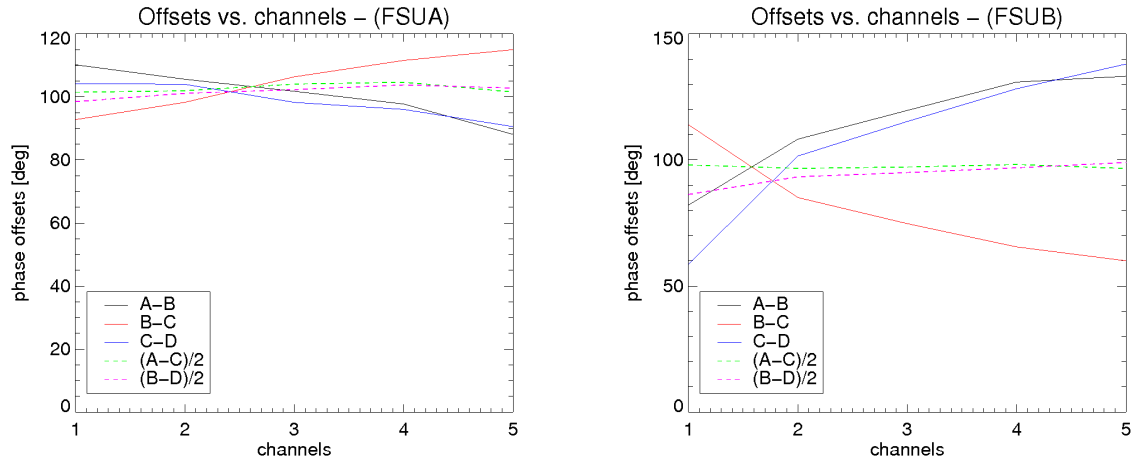


Figure 4.16. – The graphs depict the mean phase offsets between the four quadratures ABCD, as obtained from the gaussian fits to the distributions over the whole sample of data from comm. 11–13. Spectral dispersion is evident for the beams, which differ in their polarization states, but not between those, whose phase offsets result from the separation by the 50/50 beam splitter cube only (A/C, B/D).

Nonetheless, the actual but non-ideal phase relations between the ABCD signals affect phase and group delay estimations and can lead to nonlinear behavior.
As apparently they cannot be prevented, they must be measured and the depending estimates need to be calibrated, either in real-time to increase the tracking efficiency, or at least during post-processing.
This can be achieved by the following routine, suggested by [Sah09]. While in the ideal case, with nominal phase relations between the ABCD signals, the fringe phase can be derived from the correct fringe quadratures $X = (I_A - I_C)$ and $Y = (I_B - I_D)$ by applying Eq. (4.3), this approach needs to be expanded in reality, when phase offset errors are present and described by ψ_B, ψ_C and ψ_D, as the deviations of the corresponding signals' phase difference from their ideal values ($\pi/2$, π, and $3\pi/2$) to the reference signal A.
The corrected quadratures can be constructed then by a linear combination of the non-ideal signals and are given by the expressions

$$\begin{aligned} X(t) &= [(I'_A(t) - I'_C(t))\gamma - (I'_B(t) - I'_D(t))\alpha]\, c_{\rm sc} \quad , \\ Y(t) &= [(I'_B(t) - I'_D(t))\beta - (I'_A(t) - I'_C(t))\delta]\, c_{\rm sc} \quad , \end{aligned} \tag{4.79}$$

where the primed quantities denote the disturbed signals, $c_{\mathrm{sc}} = (\beta\gamma - \alpha\delta)^{-1}$ is a normalization factor and the phase shift errors are hidden in the coefficients:

$$\begin{aligned} \alpha &= \sin\psi_C \\ \beta &= 1 + \cos\psi_C \\ \gamma &= \cos\psi_B + \cos\psi_D \\ \delta &= -\sin\psi_B - \sin\psi_D \quad . \end{aligned} \tag{4.80}$$

With these corrected quadratures the desired quantities can be estimated by the standard equations of the ABCD method, as introduced in Sec. 4.1.3.1.

However, it should be noted that even though the phase shift errors can be in principle calibrated in this manner, they still need to be known. So far, as their occurrences are not understood sufficiently well, the most secure approach would be to determine those from a calibration scan on the source, prior to the actual scientific observation.
Still, it is unclear, how stable the phase offsets are and how frequent such calibration data should be recorded. Nevertheless, further effort should be put into the analysis of the phase shift errors so that eventually some trends, with reference to the system's configuration during observations, may become evident, which could provide the means for proper modelling of the effects.

4.3.1.2. Fringe Motion and Phase Referencing

The concept of fringe tracking in presence of a phase reference extends the coherence time during the measurement and allows observations of dimmer sources. For PRIMA this is achieved by the simultaneous interferometric measurement on two sources with an angular separation within the isoplanatic patch.
As the atmospherically induced phase variations of those two targets are presumed to be sufficiently equal, any corrections to these, as derived from the phase monitoring of the brighter primary source, are applied to both objects equivalently by the adjustment of the main delay line, which affects the delay of both targets. Consequently, this approach damps phase fluctuations on the secondary target, which need to be compensated by the differential delay lines.
As results from the analysis in Sec. 4.3.1 of the RTO variations, which can be regarded as a response to the atmospheric disturbances, already reflect this effect, it might be interesting to take a deeper look into the temporal properties.
This can be done by investigating the power spectra of the fringe motion, as has been for example performed in the past at the Mark III stellar interferometer and discussed in [Col1987] or [Bus95].
According to theory (see [Gli11] for details), the temporal power spectra can be computed by integrating the spatial power spectra, which describe the statistics of the disturbances, as the Kolmogorov spectrum (Eq. (3.48)), introduced in Sec. 3.2.2, over all spatial frequencies.
As the Kolmogorov spectrum statistically specifies the pure phase fluctuations without accounting for the technique of measurement, this approach yields a temporal power spectrum with a $f^{-8/3}$ dependency over the whole frequency band. This situation changes, when in context of

an interferometric observation not the general phase fluctuations on the wave front, but rather the variations of the fringe positions in the instrument are considered.
Since the instrument's configuration is not isotropic, due to a designated direction of the baseline, the symmetry is also broken for the spatial and respectively temporal power spectra.
When the wind velocity, at which the turbulence cells are moving across the atmospheric layer, is also taken into account, the temporal power spectrum of the fringe motion becomes dependent on the wind direction with regard to the baseline orientation. However, a consistent overall expression can be derived by averaging over all possible wind orientations ([Lin1980]).
Doing so, the resulting power spectrum exhibits two ranges of different dependencies on the frequency with a transient frequency $f_{t1} \approx 0.2V/B$, where V is the effective wind speed of the single layer approximation (Eq. (3.61)) and B is the baseline length.
At the high frequency part of the spectrum, equivalently at low time scales, the phase fluctuations at the two apertures are mostly independent from each other, and the power spectrum of the fringe motion follows the power law $\propto f^{-8/3}$ dependency, just as the statistics of phase fluctuations at a single telescope do.
At lower frequencies, which correspond to spatial scales that are comparable to the baseline, the phase variations at the instrument's inputs become more and more correlated, which translates internally to a suppression of the fringe motion at those frequencies and consequently to a shallower asymptote in the power spectrum with $\propto f^{-2/3}$.
When this approach is extended to frequencies that represent spatial extents of the order of the sizes of the single sub-apertures, the phase averaging over the telescope diameter d acts as a high-pass filter and causes a drop in the power spectrum for frequencies above a second transient value $f_{t2} \approx 0.3V/d$, and the power law is described by $\propto f^{-17/3}$.
While the behavior within the two regimes below f_{t2} has been found and verified with several instruments already, such as with the Sydney University Stellar Interferometer ([Dav1995]), the Mark III interferometer ([Bus95]), the Palomar testbed interferometer ([Lan03]), or with the VLTI Commissionning Instrument VINCI ([DiF03]), the steep relation at the very high end has not been confirmed so far.
Furthermore, it should also be noted that the bare $\propto f^{-2/3}$ dependency, which has been derived from the Kolmogorov infinite outer-scale model, exhibits a pole at $f = 0$. This issue is avoided when a finite outer scale L_0 of the turbulent eddies is considered within this context. This leads to a saturation in power on large scales and results in a rather flat asymptote at frequencies below $f_L \approx V/L_0$.
As the transition frequencies in the power spectra depend on the used baselines, as well as on the atmospheric conditions, the transition from the $\propto f^{-8/3}$ region to the flat regime can be fuzzy. However, if a break is observed within the spectrum, the frequency, at which it is observed, can be regarded as an upper limit to f_{t1} and f_L, whichever is greater, and consequently as a lower boundary to the outer scale L_0.
The tracker's responses to the fringe motions in OPD space, due to atmospheric turbulences, have been investigated here on the basis of the RT offsets, recorded by the two OPD controllers, and are illustrated in Fig. 4.17.
It shows processed data of a tracking sequence, acquired with a 91 m baseline on a pair of stars (HD 10360, HD 10361), which are similar in brightness (both around 3.5 in K magnitude), and separated by $\rho \approx 11.5$ as.

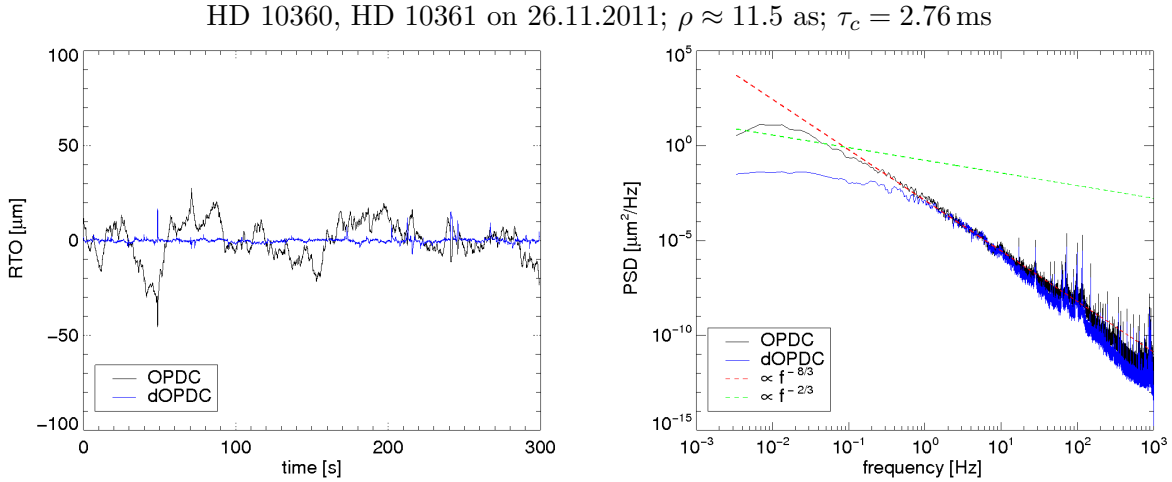


Figure 4.17. – *Left:* Linearly detrended and centered real-time offsets as recorded by both fringe tracking controllers (OPDC on the pimary and dOPDC on the secondary source) during a tracking sequence on a star pair, separated by ≈ 11.5 as on the sky and with magnitudes in the K-band around 3.5. *Right:* Temporal power spectra of the same signals, overplotted with the two major asymptotes of the Kolmogorov infinite outer-scale model. Model and data correspond well for mid frequencies, where the power law exponent is fitted to -2.89 on the OPDC data. A damp in the power of the atmospheric turbulences that are corrected for by the differential delay lines is evident at frequencies below ~ 1 Hz in the dOPDC spectrum.

During the 300 s sequence full tracking was achieved on both targets, when the controller state IDLE state is counted in as well, which for the main tracker contributes to about 2.6 % of the time. The RTO signals over time reveal different behavior of the differential OPD controller, compared to the offsets on the main target. High amplitude but low-frequency disturbances are measured and already applied to the main delay, hence to both targets, so that they do not reappear within the RTO recordings on the dOPDC.
This clear manifestation of the phase reference benefit becomes even more obvious within the temporal power spectra, which have been derived from both RTO sequences after linear detrending and centering around zero.
While for frequencies above ~ 1 Hz the spectra of both trackers are fairly comparable, though a slight suppression in amplitude seems apparent at the dOPDC for frequencies greater than ~ 10 Hz, an evident divergence between the two curves is found at the low-frequency part of the spectra. In this temporal regime the atmospheric piston is compensated for by the main delay tracker and results in the observed and expected damp in the power of the dOPDC signal.
A fit to the power law within the range of 1–50 Hz to the main delay data yields an exponent of around -2.89, which is a little bit steeper than reported values from PRIMA fringe tracking by [Sah09] or [Sah11], but still can be regarded as a proper affirmation of the theoretical prediction, with a deviation of less than 10 % from the Kolmogorov value of $-8/3$.

The flattening of the spectrum at lower frequencies is also apparent, with a fitted exponent of about -0.7 for frequencies below 1 Hz, but a clear distinction between the two regimes cannot be made, as the region of transition appears rather broad.

Common to the spectra of both controllers are noise peaks at several well-defined frequencies above 1 Hz and a clear bump in the range between ~ 60 Hz and ~ 100 Hz. These peaks, located above the closed loop fringe tracker bandwidth of about ~ 15 Hz ([DiL08]), can result from any high-frequency vibrations on any optical elements along the path of the travelling light beams through the system. They are regularly encountered within the tracking data and have been previously reported and partially discussed in [Sah09], where at least some of them could have been attributed to originate from differential injection of the ABCD signals into the fibers, which can result in erroneous phase and delay estimates, which, misinterpreted by the tracking controller as actual fringe motions, are then falsely corrected for.

Furthermore, the predicted steep decay in the spectrum above the discussed transient f_{t2} is not confirmed within this tracking sequence, although it should be noted that an obvious change in the power law exponent towards increased values in magnitude is visible for frequencies beyond ~ 150 Hz. This feature is encountered also within spectra of other data sets and might be interesting to be investigated further.

However, in summary it can be stated that the advantages of phase referenced fringe tracking, achieved by PRIMA's dual-feed mode implementation, are clearly evident in the flatness of the power spectrum of the dOPDC real-time offsets beneath $f \approx 1$ Hz and that the secondary fringe tracker mostly compensates for the remaining low amplitude perturbations at high frequencies.

4.3.2. Baseline Calibration

PRIMA has been designed to measure differential delays regarding fringe envelope centers of two simultaneously observed sources, from which the desired astrometric quantities, the star separations on the sky, can be principally derived, when the basic relation Eq. (4.21) is applied. To achieve this, the baseline, so far roughly defined as the vector connecting the two telescopes, needs to be known beforehand.
As will be outlined in the following two sections, in reality and in particular when regarding PRIMA's dual-feed character, the baseline concept becomes complex and the determination of this astrometrically essential quantity is anything but a trivial task.
In ideal situations the static entity, in reference to the observatory, which together with the coordinates of the observed source determines the measured delays (Eq. (4.18)) that need to be applied to compensate the external OPD, is called the wide-angle baseline and can be defined as the joining vector of the fixed telescopes' pivot points, positions of the intersections of the azimuth and altitude axes, as these determine the effective path lengths of the stellar beams in the system.
Still, when it comes to real instruments in uncontrolled environments, the positions of the telescopes' pivot points will not remain locally fixed, but can be smeared out due to optical

misalignments or any slight material deformations. This applies also to the the wide-angle baseline, which consequently becomes unstable and subject to time dependent variabilities.
In addition to that, conceptional difficulties arise due to the implementation of the laser metrology system, which does not stringently cover the entire actually applied delays, but only the path lengths between the laser beams' locations of injection (extraction) and their endpoints. Although in the case of PRIMA the metrology covers the point of beam combination, its coverage of the beam train is limited by the retroreflectors within the star separators, which do not coincide with the pivot points. This generally results in offsets between the effective baseline and that one, which is detected by the system and can be calibrated.
Furthermore, as the metrology reaches only up to the STSs M9 mirrors, a longitudinal offset is present due to the unmonitored path length between the metrology endpoint and the instrument's entrance pupils at the M2s.
In order to incorporate these uncertainties, the delay equation can be extended by error terms, as introduced in [Sah13]:

$$\underbrace{\mathrm{DL2}-\mathrm{DL1}}_{d}+\epsilon_{\mathrm{tr}}+\epsilon=\hat{\boldsymbol{s}}\cdot(\underbrace{\boldsymbol{B}_{\mathrm{W}}-\boldsymbol{\mu}_1+\boldsymbol{\mu}_2}_{\boldsymbol{B}_{\mathrm{eff}}})+C \quad . \tag{4.81}$$

Here the effective instrument baseline $\boldsymbol{B}_{\mathrm{eff}}=\boldsymbol{T}_2+\boldsymbol{\mu}_2-(\boldsymbol{T}_1+\boldsymbol{\mu}_1)$ is composed of the ideal wide-angle baseline $\boldsymbol{B}_{\mathrm{W}}=\boldsymbol{T}_2-\boldsymbol{T}_1$ and the terms $\boldsymbol{\mu}_{1,2}$, which describe the metrology endpoint offsets from the telescopes' pivot points. The left hand side of the equation specifies the effectively applied delay that needs to be adjusted to compensate for the external geometric delay and the instrument's constant term C. Besides the measured delay d, it accounts for contributions from any tracking errors ϵ_{tr} and path discrepancies ϵ due to not monitored stellar beam path segments.

The instrument's effective baseline can be obtained from delay measurements during observations of stellar sources with known coordinates, as widely and uniformly as possible spread across the accessible celestial sphere.
For this reason, during the night of November, 21th/22th 2011 several stars have been observed by PRIMA in both modes (NORMAL/SWAPPED).
The stars have been chosen to be single stars from the FK6 catalog, and the catalog's *long-term prediction mode* (LTP) has been incorporated into the given star positions and proper motions ([Wie02]). The parallax values have been extracted from the new reduction of the Hipparcos data ([vLee07]). The nine stars that have been observed during this night and which have been used eventually to calibrate the baseline are listed in Tab. 4.6.

The observations have been performed with the two telescopes being mounted on the VLTI stations G2 and J2 (see Fig. 4.18), which results in an approximate baseline lenght of ≈ 91.2 meters. As PRIMA is by principle constructed to perform measurements on two stars at the same time, but for this calibration only single objects have been observed, the system's set-up had to be adapted. This was achieved by introducing a fake secondary source at a distance of 8 microarcseconds in right ascension to the actual star. In this manner, the telescopes were pointing to the position between those two objects, and the stellar beams were picked up as usual by the field selectors.

Table 4.6. – Stars and properties that have been used for the baseline calibration in Nov. 2011. The coordinates have been derived from the long-term prediction mode of the FK6 catalog and are given here in the form as have been provided to IPHASE (equinox and epoch J2000.0). The parallax entries are taken from re-reduced Hipparcos data ([vLee07]).

	RA	DEC	μ_α	μ_δ	parallax	radial vel.
	[h : m : s]	[d : m : s]	[s/yr]	[as/yr]	[as]	[km/s]
HIP 14146	03 02 23.49941	-23 37 28.0799	-0.01073429	-0.0552	0.0368	-9.8
HIP 22449	04 49 50.41257	+06 57 40.6112	+0.03123425	+0.01217	0.12394	+24.3
HIP 36795	07 34 03.18071	-22 17 45.8509	-0.0028749429	+0.04759	0.03953	+61.4
HIP 34088	07 04 06.53089	+20 34 13.0768	-0.0004970226	-0.00014	0.00237	+6.7
HIP 24659	05 17 29.08546	-34 53 42.7336	+0.0072145027	-0.33495	0.03082	+21.1
HIP 50799	10 22 19.58617	-41 38 59.8705	-0.0023500477	+0.06020	0.016	+20.9
HIP 20384	04 21 53.32834	-63 23 11.0090	+0.01285945	+0.17426	0.00848	+45.0
HIP 50954	10 24 23.70642	-74 01 53.7970	-0.0037246118	-0.02687	0.06164	-4.8
HIP 56647	11 36 56.93254	-00 49 25.5012	+0.0001873527	+0.04339	0.01797	+1.0

During the NORMAL mode observations the fringes have been tracked by FSUB and, after swapping the input channels, the fringes have been locked in SWAPPED mode by FSUA. As this field rotation is performed by the de-rotator, its changing position can also influence the effective baseline. In particular an internal longitudinal offset Δ_{AB} is expected, which, when solving the baseline for the individual modes separately, is covered by the corresponding constant term, but which needs to be accounted for when the baseline solution is obtained from the combined data of both modes.

For this baseline calibration the main delay measurements, obtained from the main delay line metrology (not PRIMET), have been used, and the system has been assumed to be stable over the night so that a time dependency was considered to be solely introduced by effects from longitudinal dispersion.

The delay data has been reduced by the ADRS pipeline up to "Level 2", after which the delays, averaged onto a 1-second time grid, have been corrected for dispersion manually by the approach described by Eq. (4.16) so that the final equations to be solved for by the calibration routine becomes:

$$n_{\mathrm{corr}} \frac{n_{\mathrm{s,G}}(T,P,H,\kappa_{\mathrm{s}})}{n_{\mathrm{met}}(T,P,H,\kappa_{\mathrm{met}})} \cdot d = \hat{\boldsymbol{s}} \cdot \boldsymbol{B}_{\mathrm{eff}} + C \pm \Delta_{\mathrm{AB}} \quad . \tag{4.82}$$

In order to retrieve the refractive indices for the stellar bandpass (2.25 μm) and respectively the metrolology wavelength ($\lambda_{\mathrm{vac}} = 632.991354\,\mathrm{nm}$), the models, as provided by [Mat06] and [Bir1993], have been applied, wherefore the temperature and the air pressure in the tunnels have been extracted from the individual observation raw file headers, and an average site humidity of 12.5 % has been assumed.

The resulting group refractive indices have been determined to lie between $n_{\mathrm{s,G}} - 1 = 1.995 \times 10^{-4}$ and $n_{\mathrm{s,G}} - 1 = 1.999 \times 10^{-4}$, while the refractive indices regarding the metrology fluctuated

between $n_{\text{met}} - 1 = 2.016 \times 10^{-4}$ and $n_{\text{met}} - 1 = 2.020 \times 10^{-4}$.
An additional correction term ($n_{\text{corr}} = \lambda_{\text{vac}}/\lambda_{\text{mod}}$) has been applied, as, according to [Sah13], the controller that converts the metrology measurement to the delays does not use the vacuum wavelength but a modified wavelength $\lambda_{\text{mod}} = 632.863\,\text{nm}$, which seemingly has been implemented as an average refraction correction.

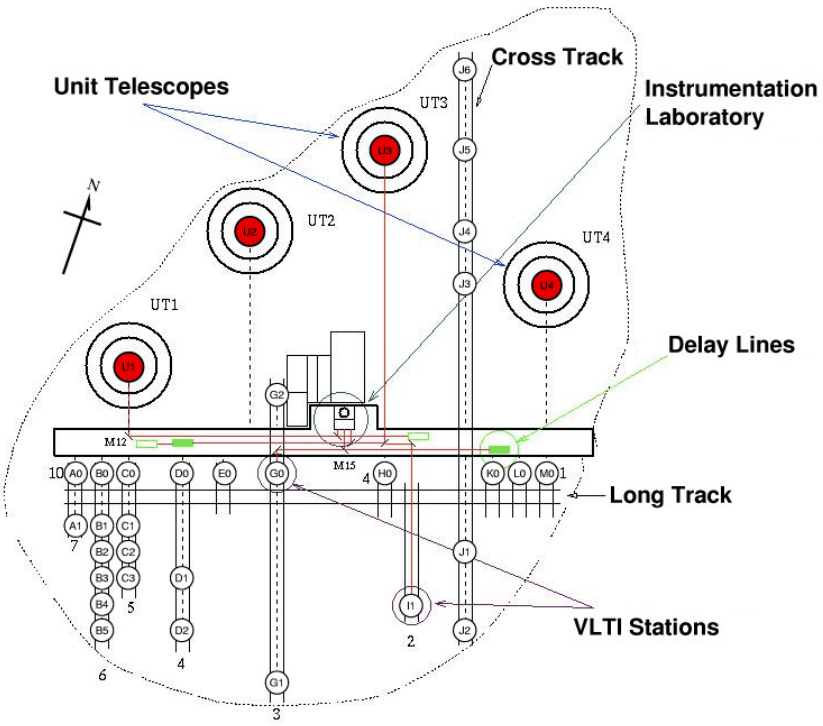


Figure 4.18. – The VLTI station map depicts the site platform with the UT positions (UT1, ..., UT4) and the stations (A0, ..., M0), which the ATs can be positioned onto. Due to those flexible system configurations, various interferometric baselines, distinguished by length and orientation, become accessible. (Image from [ESO03])

In order to solve the delay equation, the corrected delays and the star catalog (Tab. 4.6) have then been fed manually into IPHASE. In the case of only one mode data (NORMAL or SWAPPED), a four parameter fit has been performed to determine the global constant term C and the three baseline components: B_x, the component in west direction on the ground, B_y, the component along the south direction and B_z, the difference in altitude above ground.
The fifth parameter Δ_{AB}, also assumed constant over the night, has been applied only, when combined data has been considered. According to Eq. (4.82), this offset is applied with opposite sign to the measured delays of the two concurring modes. While it is arbitrary and a pure question of definition, in this analysis here the positive sign has been attributed to the NORMAL observations and respectively the negative to the SWAPPED.
Besides the delay data to the given times (UTC) of the measurements and the star properties, the IPHASE software needs also to be provided with the observing site's geographic coordinates.

As the station positions of the two AT stations can be retrieved from [ESO03] with respect to a well-defined coordinate center of the VLTI array with longitude $\varphi_{\mathrm{VLTI}} = -70.40498688°$ (east positive) and latitude $\lambda_{\mathrm{VLTI}} = -24.62743941°$, the precise position of the midpoint between the two telescopes can be determined.
With the coordinates E = 38.063 m (offset in east direction) and N = −12.289 m (offset in north direction) of the G2 station and E = 114.46 m, N = −62.151 m of the AT station J2, the geographic position of the midpoint (E = 76.2615 m, N = 37.22 m) has been determined to $\varphi_{\mathrm{G2-J2}} = -70.404233°$ and $\lambda_{\mathrm{G2-J2}} = -24.627774°$, wherefore a site elevation of 2681 m has been adopted.[7]
Furthermore, if provided with the necessary information, IPHASE is capable of correcting for earth orientation at the time of observation. For this reason, parameters on polar motion and UT1-UTC have been applied after being retrieved from the [IERS13] long term earth orientation data.

Table 4.7. – Results of the baseline calibration, where the fit to the delay equation has been performed within IPHASE. The best fit values, their formal errors and the residual RMS are listed for the four parameter models of the individual datasets of observations in NORMAL and SWAPPED mode, as well as the results of the five parameter fit to the data combined from those two set-ups. The number of data points (1-second averages of the main delay readings) is 1,413 during NORMAL and 1,790 during SWAPPED observations.

	$\mathrm{NORMAL_B}$		$\mathrm{SWAPPED_A}$		combined	
	9 stars; N = 1413		8 stars; N = 1790		9 stars; N = 3203	
	value [μm]	σ [μm]	value [μm]	σ [μm]	value [μm]	σ [μm]
B_x	76387563.9	4.8	76387584.2	5.0	76387577.3	3.5
B_y	-49878178.2	2.5	-49878172.2	3.1	-49878174.2	2.1
B_z	-16679.1	9.8	-16676.2	10.9	-16681.5	7.5
C	-60920915.0	8.7	-60920731.2	9.1	-60920827.3	6.4
Δ_{AB}	-	-	-	-	-94.7	0.9
RMS [μm]	38.1		54.1		47.8	

The results of the final IPHASE baseline calibration are summarized in Tab. 4.7.
The first noticeable outcome is that the fits to the individual observational mode data differ mainly in the determination of the constant zero term, while the three baseline components agree down to the maximum difference of 20.3 μm in the x-component ($(\Delta B_x, \Delta B_x, \Delta B_z, \Delta C) = (20.3, 6.0, 2.9, 183.8)\,\mu$m). This was expected due to the difference in the de-rotator position between the two set-ups and is successfully modelled by the fit to the combined data by $\Delta_{\mathrm{AB}} = -94.7\,\mu$m, which is approximately half the constant term difference ΔC.
The RMS of the combined data fit residuals is about $\approx 48\,\mu$m (lower for NORMAL data only, but higher for the SWAPPED fit) and the formal errors of the fit parameters appear rather

[7] The coordinate center coordinates and the site elevation are retrieved from header keywords of the fits files.

small, with a relative error of about 4.5×10^{-8} for the x- and y-components of the baseline, and approximately 4.3×10^{-4} for the z-component, which naturally is due to its magnitude being several orders smaller than the absolute values of the two other components.
Furthermore, it should be noted that the B_z component is generally difficult to determine as it correlates strongly with the constant term C.
Nevertheless, taking a deeper look into the residuals and their distributions against several parameters seems to reveal hidden systematics within the delay measurements and questions the confidence on the accuracy of the calibration. Although no clear dependency of the residuals regarding the azimuth angle of the observation or the elevation can be derived from the graphs in Fig. 4.20, some structure seems to be indicated with varying applied delay and becomes rather obvious, when studying the residuals against the time of observation during the span of about nine hours, during which the frames for this calibration have been acquired.
As one can see, some systematic change in either the instrument's environment or varying influences from the atmosphere excited a clear trend in the delay measurements, mirrored by the monotonic behavior of the residuals against time with a drift of about 170 μm between the first observation of the star HIP 14146 and the last in the night HIP 56647. Applying a stochastic error of this order of magnitude onto the delay measurements, naturally increases the formal errors of the fit parameters by a factor of $\sim$3.5.
Also, as indicated by the accelerated increase over time, either those systematics became increasingly effective during the last hour of the observation sequence due to any hidden reasons, or the time scale of the variation of their underlying mechanisms changed. However, this has a clear impact on the distribution of the residuals, resulting in a deviation from a natural gaussian shape (see Fig. 4.19).

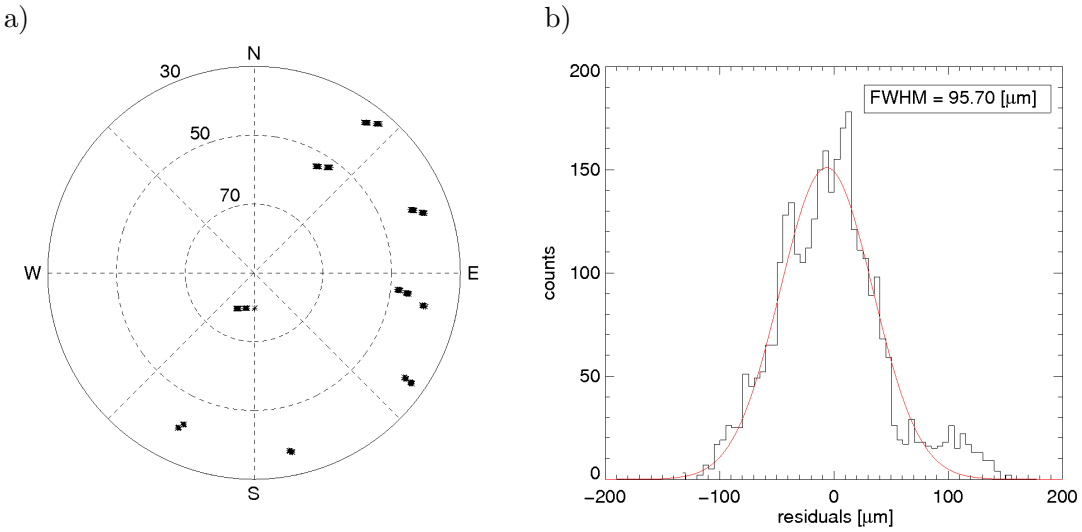


Figure 4.19. – a): Star distribution in azimuth angle and elevation. As one can see the northwest part of the sky is empty, which is mainly due to observing restrictions from the used telescope stations and the applicable delays. b): Distribution of the residuals of the fit to the combined data of both observational modes. The bin size is 5 μm and the FWHM was determined by a gaussian fit.

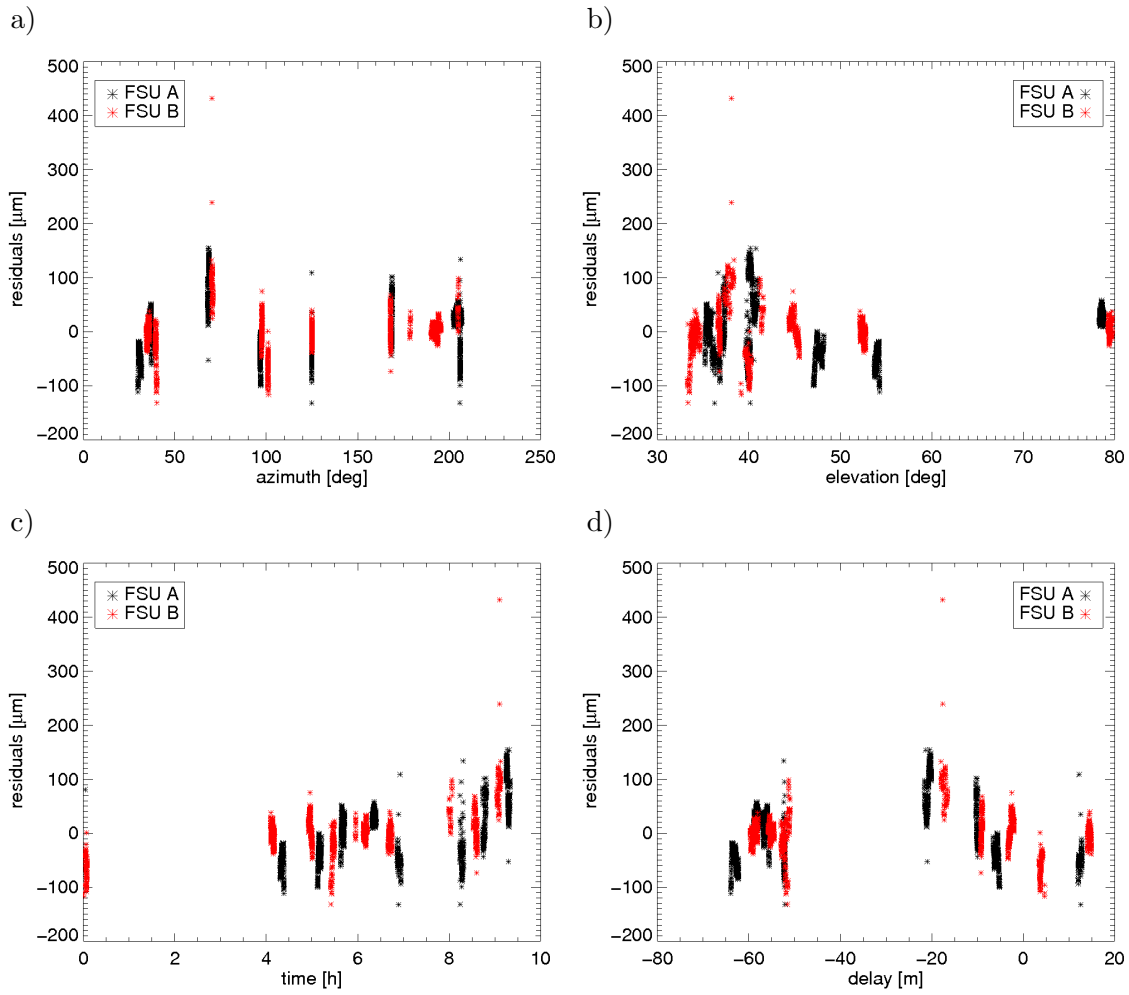


Figure 4.20. – Residuals of the combined baseline fit against the azimuth angle and elevation of the observed calibrator stars, the time of observation, in reference to the first data set, and against the applied delay.

Although these features are alarming and must be investigated further to definitely pinpoint the sources of the variations, the analysis also suggests that the system's precision allows to calibrate the instrument's wide-angle baseline with an accuracy, as is demanded by the science goals of the PRIMA program.
One should remember that the relative precision of the baseline directly scales the uncertainty of the obtained astrometric measurements of star separations and that a relative accuracy of 10^{-6} ($100\,\mu$m on a 100 m baseline) should suffice to obtain astrometric uncertainties below $50\,\mu$as for absolute star separations on the sky up to 30 as, which is the limit on target selection for the exoplanet search program (see chapter 5).

The results of the baseline calibration, which have been presented here and have been further used for the astrometric analysis in Sec. 4.3.3, qualitatively resemble those acquired by [Sah13] on the same dataset, although the absolute values of the baseline components differ

up to the maximum of $\approx 860\,\mu$m on the less accurate z-component and $\approx 570\,\mu$m on the y-component. These deviations have been verified to mainly originate from a difference between the applied geographic coordinates of the observation site and the polar motion correction.
An agreement of the results to about $100\,\mu$m has been obtained when those differences have been left aside for the modelling.

4.3.3. Astrometry from Commissioning Data

During the PRIMA commisioning runs several bright binary targets have been observed in dual-feed mode in order to ensure the instrument's functionality and to test the modelling of the recorded data with respect to the desired observable, namely the astrometric separation vector between the two sources on the sky.
Within this section the derivation of the star separation from the raw PRIMA tracking (and scanning) data will be outlined and the results from the performed analysis on different observations will be presented regarding the question towards the instrument's precision and accuracy.

4.3.3.1. Data Reduction

For the purpose of this analysis, the raw fits files, as recorded by the instrument during the observations, have been processed. All the essential information can be obtained from three different file table categories. The measured differential delay is retrieved from the fits file extension METROLOGY_DATA, sampled at 4 kHz, the group delay measurements of the two FSUs are given within the IMAGING_DATA_FSUA (resp. IMAGING_DATA_FSUB) tables, which are recorded at 1 kHz, and the OPD controllers' states are acquired from the corresponding file extensions OPDC/DOPDC, with the data sampled at 4 kHz.
Before the astrometric fit is performed, some basic data reduction is performed to the raw files, at one hand in order to ease the data handling, as the high sampling entails large file sizes, and on the other hand also to correct for instrumental and conceptional effects.
Consequently, the following procedure is applied to all dual-feed raw data files, which have been recorded while simultaneously fringe tracking on two targets:

1. The raw differential delay data depends on the PRIMET laser metrology, which detects each passing fringe. Those are internally stored in fringe counters with a size of 32 bits. Therefore, the absolute count is limited, and a counter overflow occurs, when the maximum value is reached. Whenever this happens, it results in two jumps of the raw Δd measurements, the first by $-\Delta_{\mathrm{FC}}$, followed by $+\Delta_{\mathrm{FC}}$ shortly after the first occurrence, where $\Delta_{\mathrm{FC}} \approx 486.783\,\mu$m ([Sah13]).
   Such metrology counter overflows, illustrated in Fig. 4.21, are stored in corresponding keywords within the file headers and need to be corrected for by shifting the data in-between by Δ_{FC}.
   In those cases, where the fringe counter oveflows occur at the beginning or the end of

a data frame, only one of the jumps is recorded and the differential metrology data is adapted appropriately.

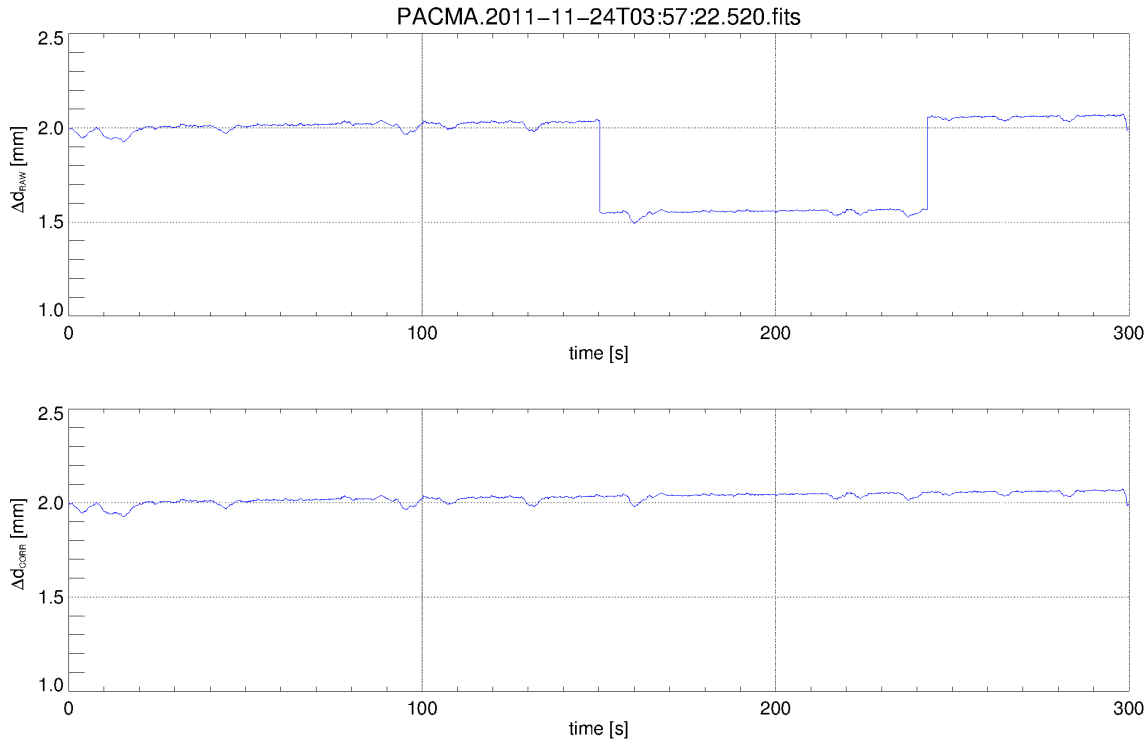


Figure 4.21. – Raw PRIMET Δd recordings (*top*) and the corresponding recordings corrected for the occurring jumps by the amount of $\Delta_{\mathrm{FC}} \approx 486.783\,\mu$m due to fringe counter overflows (*botttom*).

2. The OPD and dOPD controller states are investigated regarding the tracking state, and all metrology data that has been recorded, while one of the controllers was not fringe tracking (state 7), is discarded.

3. The group delay data of both fringe sensor units is smoothed due to its noisy nature and interpolated onto the time grid of the metrology readings in order to apply the following group delay correction onto the differential delays:

$$\Delta d_{\mathrm{corr}} = \Delta d - \mathrm{GD_B} + \mathrm{GD_A} \quad . \tag{4.83}$$

4. Although at this point of the reduction one is in principle left with only differential delay data, while, according to the path delay controllers, both stars were tracked, experience has shown that, especially at the points of transitions between different controller states, thus at the beginning and the end of individual tracking sequences, the differential metrology data can be quite unstable and shows considerable variations.
Sigma clipping is applied in order to get rid of these recordings. For this reason a 2nd

order polynomial fit is performed and all data points, for which the residuals happen to be greater than three times the residual standard deviation, are being rejected.

5. The remaining delay data is averaged onto a 1-second binned time grid, which after the reduction procedure is finished, and the data is considered clean and ready for further analysis.

It should be noted that no dispersion correction is applied, as the differential delay is introduced by the differential delay lines in vacuum conditions, and any differential dispersive effects from the light paths outside of the DDLs are neglected.
Also, any tracking errors that do not reveal themselves within the group delay measurements might still be persistent if not discarded by the sigma clipping method.

4.3.3.2. Astrometric Fit and Separation Vector

As implied by Eq. (3.47), the differential delay measurements can be used to fit for the separation vector $\Delta\boldsymbol{s} = \hat{\boldsymbol{s}}_2 - \hat{\boldsymbol{s}}_1$ between the two observed sources on the sky. The vector is defined either by its components $\Delta s_\alpha = \Delta\alpha^* = (\alpha_2 - \alpha_2)\cos\delta$ along right ascension and $\Delta s_\delta = \Delta\delta = \delta_2 - \delta_1$ in direction of declination in the tangential plane at the objects' midpoint coordinates, or equivalently by the magnitude of the separation Δs and θ_s, the position angle (PA) (see Fig. 4.22).

The model, which is applied to the measured differential delay data, reads as

$$\begin{aligned} \Delta d(t) &= \pm B_p \cdot \Delta s \cdot \cos(\theta_{BL}(t) - \theta_s) + Z_0 \\ &= \pm B_p \cdot [\underbrace{\Delta s \cdot \cos\theta_s}_{\Delta s_\delta = \Delta\delta} \cos\theta_{BL}(t) + \underbrace{\Delta s \cdot \sin\theta_s}_{\Delta s_\alpha = \Delta\alpha^*} \sin\theta_{BL}(t)] + Z_0 \quad , \end{aligned} \tag{4.84}$$

with the positive sign applied to the data points recorded in NORMAL mode and the negative in SWAPPED.
Therefore, at least one observation frame from each of the two modes needs to be supplied to disentangle the measured delay along the baseline orientation from the instrumental contributions and to solve for the PRIMET zero term Z_0.
As the time dependence of the measurements due to earth rotation is accounted for by the changing baseline angle $\theta_{BL}(t)$, the observatory bound wide-angle baseline, which was solved for in the previous section (4.3.2), is projected onto the sky by applying Eq. (3.44). The midpoint coordinates, at which the tangential plane is defined, are recovered from the two target coordinates provided within the file headers.
For the entire analysis presented here the wide-angle baseline solution, with its components solved for by the combined data set (given in Tab. 4.7), has been used, and the error terms, involving any metrology endpoint offsets from the telescopes' pivot points and uncertainties due to unmonitored light path segments (see Eq. (4.81)) have been neglected.

From here on, any data set, composed of at least one observation frame from each mode, during which the PRIMET metrology was not reset, and no substantial glitches occurred so

that the zero term can be assumed stable and constant and can be solved for, will be denoted as a single epoch observation. In order to distinguish the various data sets, unique run identifiers will be attributed to these observations.

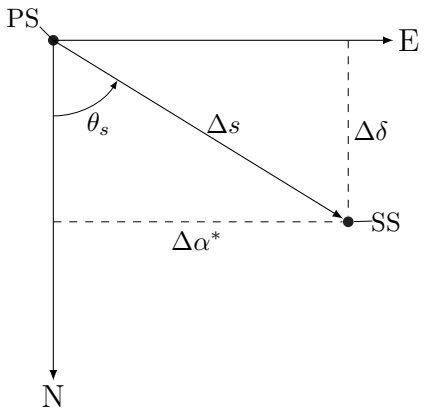


Figure 4.22. – The separation vector $\Delta\boldsymbol{s}$ defines the relative position of the two stellar targets on the sky. It can be expressed by either its orthogonal components ($\Delta\alpha^*$, $\Delta\delta$) along the directions of the right ascension and declination respectively or by its absolute value and the position angle θ_s, which gives the angular offset of the secondary source (SS) to the primary (PS), counted from the celestial north pole over the east direction.

The astrometric multiple linear regression fit to the data was performed by a standard weighted least squares routine written in *IDL* (Interactive Data Language) and is based on standard procedures summarized in [Pre1992].
The corresponding weightings are obtained frame-wise from the standard deviations σ_d of the residuals of a 2nd order polynomial fit to the reduced and averaged differential delays, which then gives their dispersion in delay space.

The applied astrometric fit solves for the constant zero term and the remaining two astrometric parameters, namely the separations along the directions of right ascension and declination, from which the desired observables are computed by the standard formulae $\Delta s = \sqrt{(\Delta s_\alpha)^2 + (\Delta s_\delta)^2}$ and $\theta_s = \tan^{-1} \frac{\Delta s_\alpha}{\Delta s_\delta}$.

An example of the averaged delay data of such an observation run is plotted in Fig. 4.23, where the original but reduced NORMAL and SWAPPED measurements from an observation over nearly 3 hours of the binary HD 202730, separated by ≈ 7 as, is illustrated, as well as the differential delays after having been corrected for the zero term.

For the estimation of the 1-sigma errors of the three fit parameters ($\delta\Delta s_\alpha$, $\delta\Delta s_\delta$, δZ_0), the frame dependent statistical delay dispersions σ_d, which as mentioned above are exploited for the file dependent weightings, are scaled linearly by a factor c_d to the point, where the reduced chi-squared χ^2_{red} of the fit becomes 1. As a matter of course, these adapted errors $\sigma_c = \sigma_d \cdot c_d = \sqrt{\sigma_d^2 + \sigma_{\mathrm{sys}}^2}$ no longer describe the pure statistical uncertainty in the delay mea-

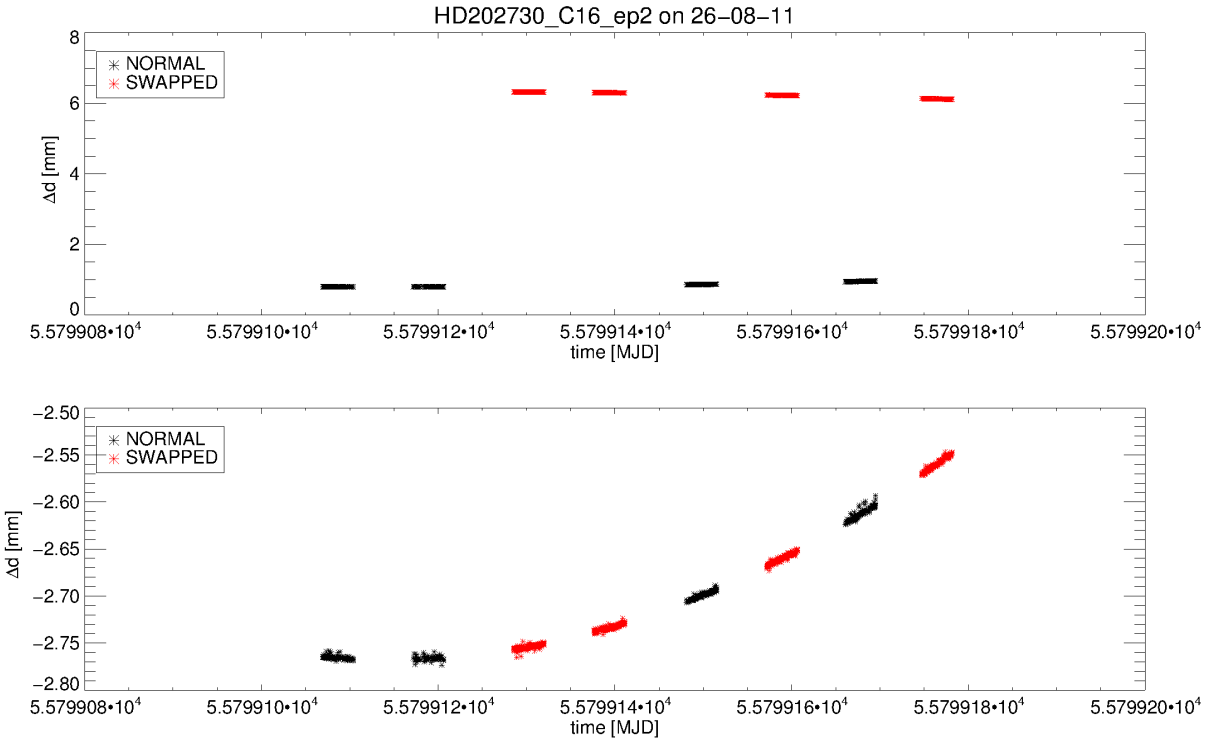


Figure 4.23. – *Top:* Reduced and averaged PRIMET differential delay data with, according to Eq. (4.84), the sign inverted measurements in SWAPPED mode and an overall offset due to the metrology zero term. *Bottom:* The corrected delays, where the constant term has been subtracted and the SWAPPED data has been multiplied by -1.

surement itself, but contain a systematic contribution σ_{sys}, which potentially displaces the measurement from the applied model.
From those uncertainties the errors of the observables are derived by the means of standard gaussian error propagation, with the entities regarded being independent:

$$\delta\Delta s = \frac{\sqrt{(\Delta s_\alpha)^2(\delta\Delta s_\alpha)^2 + (\Delta s_\delta)^2(\delta\Delta s_\delta)^2}}{\Delta s} \tag{4.85}$$

$$\delta\theta_s = \frac{\sqrt{(\Delta s_\delta)^2(\delta\Delta s_\alpha)^2 + (\Delta s_\alpha)^2(\delta\Delta s_\delta)^2}}{\Delta s^2} \quad . \tag{4.86}$$

In order to verify this approach of the error estimation, an alternative bootstrap method has been implemented and conducted for each processed observation run. For this purpose the astrometric fit is performed N times (2,000 times in the course of this analysis), but with re-sampled data sets.
While the sample size remains the same for all of those sets, where the number of values in the different modes is maintained, the varying samples are randomly drawn from the original data collection with the elements being returned after each individual draw.
Consequently, this method provides N fit solutions, from which mean values for Δs_α and Δs_δ

are derived, and the errors are estimated from their corresponding standard deviations.
The separation and position angle with errors are then again determined by the same formulae and methods, as described above for the simple astrometric fit.
Furthermore, a principal component analysis of the N fit parameter distribution in the sky plane gives the orientations of the highest and respectively lowest obtained precision.
Inherently, the orientation with the lowest dispersion is determined by the distribution of the projected baseline angles during the observation and can therefore be considered the average baseline orientation $\langle \theta_{BL} \rangle$ for the particular run.

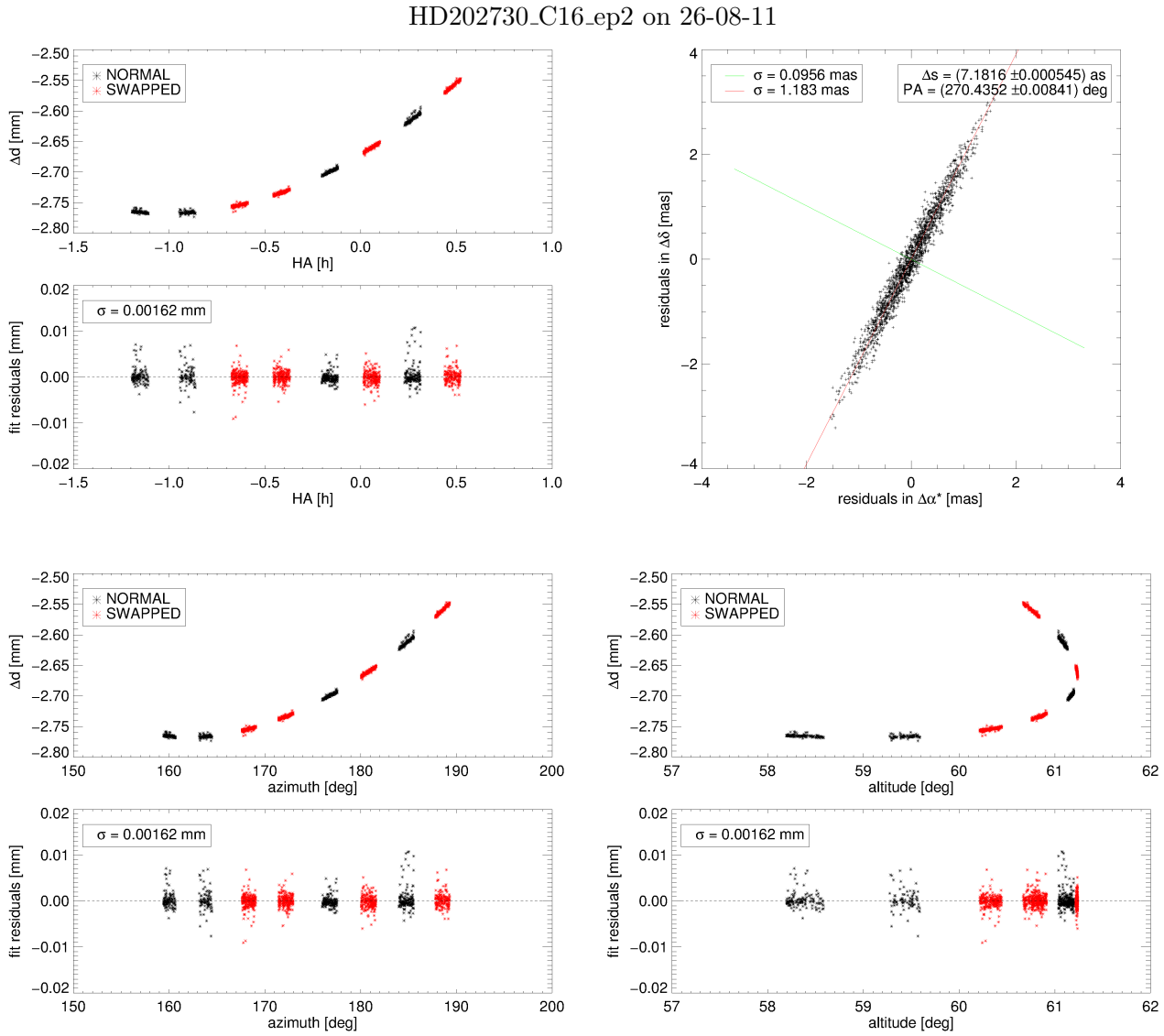


Figure 4.24. – Corrected differential delays Δd and residuals of the astrometric fit to an observation sequence of the binary HD 202730 on 26-08-11. Also depicted is the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

Figure 4.24 shows the results of this entire astrometric modelling to the same observation run of the system HD 202730 as from which the differential delays, illustrated in Fig. 4.23, have been extracted. It depicts the differential delays, corrected for the zero term, as well as the fit residuals in delay space against the target system's hour angle, the azimuth and the altitude during the observations. The residuals in SWAPPED mode are sign inverted for this illustration.
The astrometric fit to the 1,387 data points from altogether 8 frames (4 in NORMAL, 4 in SWAPPED) yields as results a separation of $\Delta_s = (7181.58 \pm 0.53)\,\mathrm{mas}$ and a position angle of $(270.4354 \pm 0.0082)^\circ$.
The residual structure is flat and does not show any significant systematics, which is why the overall residual standard deviation, at about $\approx 1.6\,\mu$m, is governed by the delay dispersions from the individual files, varying within the range between $\approx 1.0\,\mu$m and $\approx 2.5\,\mu$m.
Also shown is the bootstrap fit parameter distribution, together with the results from the resampling analysis. The precision along the average baseline angle $\langle\theta_{BL}\rangle = 117.09^\circ$ was determined to $\sigma_\parallel = 96\,\mu$as, whereas the dispersion along its orthogonal, and consequently least precise direction, is more than 10 times higher and quantified by $\sigma_\perp = 1.18\,$mas.
The uncertainties in separation $\delta\Delta s = 0.54\,$mas and position angle $\delta\theta_s = 0.0082^\circ$, derived by bootstrapping, agree very well with the corresponding ones from the simple fit.

However, although in this case at first inspection the model seems to describe the data quite precisely, it will become obvious in the following that this is not the usual case and that the indicated precision of the measurement does not reflect the instrument's accuracy.

4.3.3.3. Results and Implications

Within this section some results from astrometric fits to PRIMA dual-feed data, obtained between Jan. and Nov. 2011, is being presented and discussed. During this period of time, several visual binary targets have been observed as part of the system's characterization.
Table 4.8 lists the main properties of the systems, which have been obtained from *The Washington Visual Double Star Catalog* ([Mas01]), while all results from analyzed single epoch observation sequences are compiled in Tab. 4.9.
The individual sequences are distinguished by an unique run identifier, and the corresponding plots to the fits, if not presented here, can be found in App. D.
Lists of the individual fits files, belonging to the analyzed sequences are given in App. C.

It should also be mentioned that, although only astrometric results to tracking data is discussed here, the fringe scanning observations can also be used to fit for the desired star separations, as long as the fringe tracker on one of the targets provides the phase reference and the position of the fringe envelopes can be acquired from the measurements of the second FSU.
For the sake of completeness, several results, obtained by this approach, are also listed in Tab. 4.9. The procedure of data reduction was similar to the one described in 4.3.3.1, with the difference that the single data points were obtained by a gaussian fit to the SNR within

Table 4.8. – Properties of the observed binary systems between commissionings 14 and 17. The entries are taken from the WDS catalog ([Mas01]), and the separation and position angle values are referred to the catalog specific *date of the last satisfactory observation*, called the observation date here.

system	WDS	Obs. date [yr]	θ_s [deg]	Δs [as]	$\Delta\mu^*_\alpha$ [mas/yr]	$\Delta\mu_\delta$ [mas/yr]
HD100286	11323-2916	2007	210	9.4	3	-4
HD10268	01397-3728	1998	277	19.8	0	-14
HD10360	01398-5612	2009	188	11.6	-116	-204
HD108248J	12266-6306 (AB)	2012	111	3.6	-8	7
HD131977	14575-2125 (AB)	2011	305	25.6	-93	156
HD156274	17191-4638 (AB)	2009	256	10.2	-8	0
HD18622	02583-4018	2009	91	8.4	51	-22
HD202730	21199-5327	2010	268	7.3	-13	-9
HD66598	08031-3228	2000	135	35.7	13	-7

each scan over the fringe package in delay space. Also, the correction for the group delay was performed only with reference to the tracking FSU.

Precision versus Accuracy

The fit to the tracking sequence HD202730_C16_ep2 (see Fig. 4.24) of the rather short separation target HD 202730 appears to be quite representative for the PRIMA data obtained from this system. An observation sequence three days later (HD202730_C16_ep3) is shown in Fig. 4.25.

Both of these runs are comparable in qualitative results, as the residuals appear flat and of similar magnitude. The uncertainties regarding the derived separations of both fits are of the same order of about ≈ 0.5 mas.
This appears promising at the first inspection, but a deviation of around ≈ 10 mas of the absolute separation values arises suspicion.
Inspecting the results from all analyzed sequences of this system's observations, one sees that the scatter in separation is even bigger (compare also with Fig. 4.26), although it must be said that the fit of the run HD202730_C16_ep1 seems to be strongly affected by some systematics from the last three observation frames (see Fig. D.18 on p. 226).

Still, this dispersion, which is by around one or even two orders of magnitude higher than the implicated precision of the measurements, is alarming and cannot be explained by the stars' proper motions and an actual change of separation, but must be the result of unrecognized systematics.

Table 4.9. – Results from the single epoch analysis. For each processed observation sequence (epoch) the number of valid 1-second averaged data points and the astrometric observables with errors from the astrometric fit, defined by Eq. (4.84), are listed. Also tabulated are the corresponding 1-sigma uncertainties from the bootstrap approach and the dispersions $\sigma_{\parallel}$ and $\sigma_{\perp}$ along the directions of the axes of maximum and minimum precision from the principal component analysis. The entries for the fits to pure tracking data (*above*) and to scanning data (*below*) are separated by the dashed line.

sequence ID	date	N	Astrometric fit		Bootstrap				
			Δs	θ_s	$\delta\Delta s$	$\delta\theta_s$	$\langle\theta_{BL}\rangle$	$\sigma_{\parallel}$	$\sigma_{\perp}$
	[dd-mm-yy]		[mas]	[deg]	[mas]	[deg]	[deg]	[mas]	[mas]
HD100286_C14_ep2	31-01-11	941	9369.6 ± 2.8	210.1057 ±0.0134	2.5	0.0120	121.64	0.080	3.13
HD10268_C16_ep1	29-08-11	939	19486.91 ±0.72	278.7361 ±0.0034	0.72	0.0034	120.81	0.125	1.37
HD10360_C16_ep1	26-08-11	1167	11429.34 ±0.45	188.1107 ±0.0020	0.45	0.0020	130.98	0.063	0.60
HD10360_C17_ep1	20-11-11	366	11398.3 ±13.5	187.9759 ±0.0398	15.0	0.0448	119.86	0.071	17.41
HD10360_C17_ep2	21-11-11	1225	11433.79 ±0.22	188.0723 ±0.0018	0.22	0.0014	144.06	0.099	0.35
HD10360_C17_ep4	24-11-11	3545	11423.18 ±0.12	188.0270 ±0.0008	0.14	0.0009	144.51	0.034	0.23
HD10360_C17_ep5	25-11-11	3065	11408.98 ±0.21	187.9906 ±0.0008	0.22	0.0008	127.03	0.031	0.27
HD10360_C17_ep6	26-11-11	3376	11420.48 ±0.37	188.0264 ±0.0015	0.36	0.0014	128.12	0.039	0.46
HD10360_C17_ep7	26-11-11	5285	11432.48 ±0.06	188.0723 ±0.0006	0.05	0.0005	156.99	0.023	0.12
HD108248J_C14_ep1	27-01-11	341	3897.4 ±10.8	113.9064 ±0.2130	9.9	0.1947	122.40	0.062	16.51
HD108248J_C14_ep2	29-01-11	309	3866.1 ±19.4	116.0821 ±0.3557	19.8	0.3638	125.09	0.573	31.58
HD108248J_C14_ep3	01-02-11	359	3917.3 ± 2.5	112.7284 ±0.0465	3.0	0.0570	123.30	0.648	4.85
HD131977_C14_ep1	02-02-11	249	25789.0 ±61.9	310.3967 ±0.1534	65.1	0.1613	113.65	0.239	97.54
HD156274_C15_ep3	20-07-11	2978	10019.8 ± 1.3	257.2951 ±0.0132	1.2	0.0127	116.30	0.179	2.52
HD156274_C15_ep4	20-07-11	402	10078.2 ±19.8	257.2105 ±0.1113	19.2	0.1075	135.42	0.280	26.93
HD18622_C15_ep2	21-07-11	167	8303.4 ± 7.0	90.5292 ±0.2242	7.3	0.2328	102.02	1.095	34.50
HD202730_C15_ep1	20-07-11	603	7187.8 ± 2.3	270.1519 ±0.0382	2.6	0.0459	114.43	0.380	6.32
HD202730_C16_ep1	24-08-11	581	7142.9 ± 1.6	270.7321 ±0.0135	1.5	0.0126	133.76	0.437	2.14
HD202730_C16_ep2	26-08-11	1387	7181.58 ±0.53	270.4354 ±0.0082	0.54	0.0084	117.09	0.096	1.18
HD202730_C16_ep3	29-08-11	1088	7192.27 ±0.45	270.3241 ±0.0048	0.46	0.0049	126.83	0.061	0.77
HD66598_C14_ep1	27-01-11	189	35969.4 ±49.9	136.2926 ±0.0783	50.8	0.0797	125.28	0.134	71.30
HD66598_C14_ep2	27-01-11	325	35883.7 ±28.2	135.2607 ±0.0451	27.3	0.0436	139.00	0.122	38.59
HD66598_C14_ep3	28-01-11	928	35885.4 ± 1.8	135.5776 ±0.0029	1.5	0.0024	127.34	0.129	2.10
HD66598_C14_ep4	29-01-11	438	35865.7 ± 3.3	135.5377 ±0.0052	3.5	0.0056	123.53	0.207	4.95
HD66598_C14_ep6v2	30-01-11	2140	35891.6 ± 2.1	135.5810 ±0.0034	2.3	0.0037	121.14	0.452	3.26
HD66598_C14_ep9	02-02-11	750	36113.5 ± 3.8	133.6556 ±0.0063	4.0	0.0065	109.32	0.389	5.70
HD66598_C14_ep10	02-02-11	1087	35851.3 ± 2.0	135.4276 ±0.0032	2.1	0.0033	118.73	0.151	2.90
HD66598_C14_ep11	02-02-11	680	35933.7 ± 8.6	135.8939 ±0.0136	6.5	0.0103	132.96	2.090	8.89
HD66598_C14_ep12	04-02-11	1468	35914.61 ±0.82	135.6074 ±0.0013	0.70	0.0011	117.00	0.089	0.98
HD66598_C17_ep1	24-11-11	3975	35841.3 ± 1.2	135.2511 ±0.0019	1.4	0.0022	111.74	0.106	1.96
HD66598_C17_ep2	25-11-11	6007	35826.28 ±0.51	135.1908 ±0.0008	0.51	0.0008	113.09	0.083	0.73
HD66598_C17_ep3	26-11-11	1769	35873.0 ± 4.1	135.4654 ±0.0065	3.9	0.0061	118.90	0.303	5.45
HD10360_C17_s_ep1v2	20-11-11	67	11477.4 ±90.8	184.2621 ±0.3071	85.2	0.2272	117.87	0.618	96.61
HD10360_C17_s_ep2	21-11-11	780	11438.5 ±10.5	188.1159 ±0.0513	10.3	0.0667	142.58	0.837	16.82
HD10360_C17_s_ep3	24-11-11	84	11425.0 ± 4.9	188.0478 ±0.0200	4.4	0.0224	135.34	1.424	6.11
HD202730_C16_s_ep1	24-08-11	187	7138.5 ±11.5	270.7808 ±0.0768	2.9	0.0242	132.75	1.892	3.71
HD66598_C17_s_ep1	26-11-11	1062	35889.8 ± 9.0	135.5513 ±0.0143	8.4	0.0132	120.48	0.238	11.76

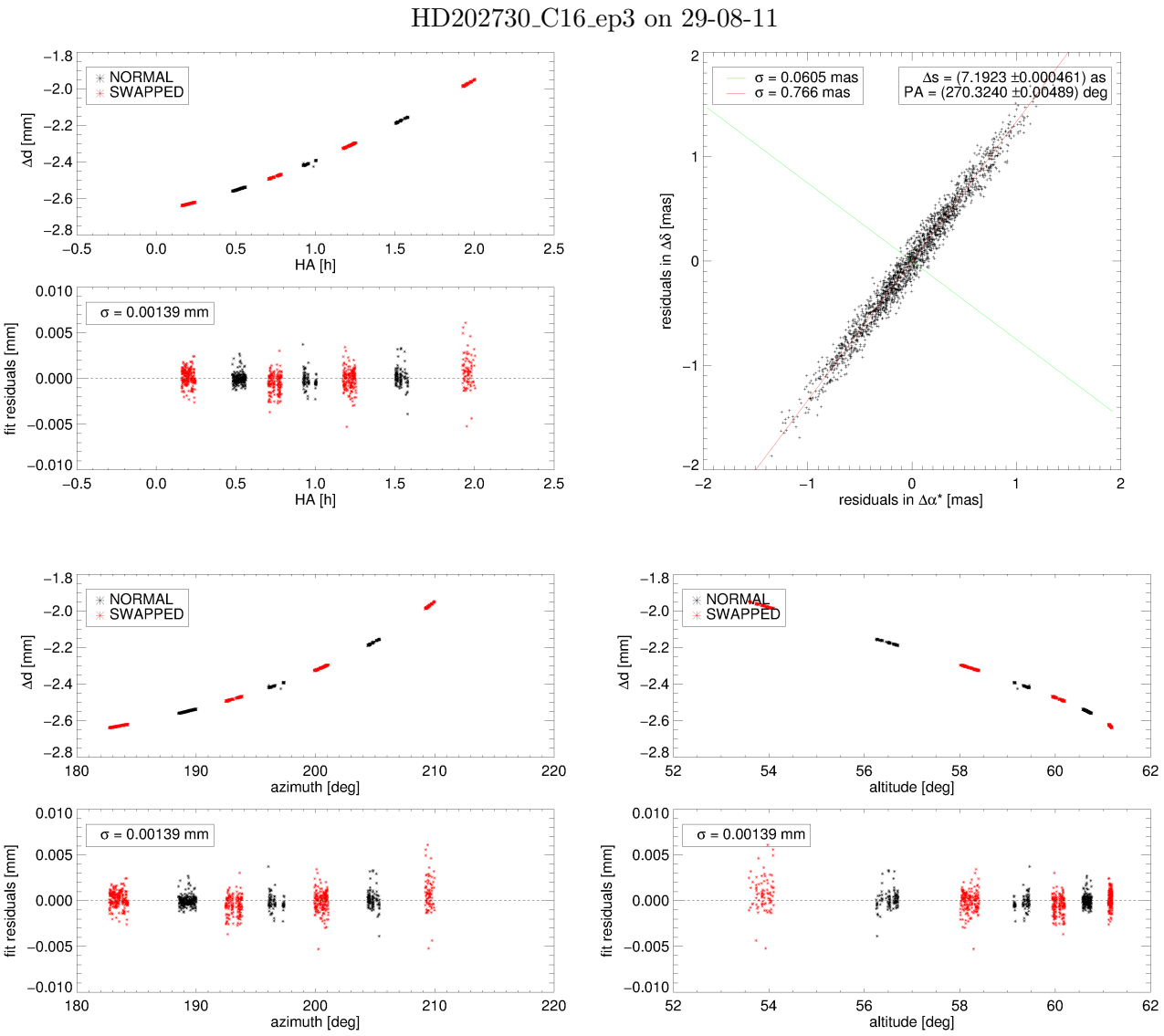


Figure 4.25. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD202730_C16_ep3 on 29-08-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

Some of those are already implied, when observation sequences of some wider separated targets are examined. Figures 4.27–4.29, belonging to runs HD10360_C17_ep5 ($\Delta s \approx 11.5$ as), HD10268_C16_ep1 ($\Delta s \approx 20$ as) and HD66598_C17_ep2 ($\Delta s \approx 36$ as), show all significant residual structures, which seem to become more and more apparent with increasing star separations. In particular for the systems HD 10268 and HD 66598 the structure, with respect to the hour angle of observation and the azimuth angle, are clear and even appear to be smoothly shaped and repeating.
Consequently, the standard deviations of the residuals exceed the dispersion in delay measurements of the individual files. For HD10268_C16_ep1 they are determined to about $\approx 1.9\,\mu$m

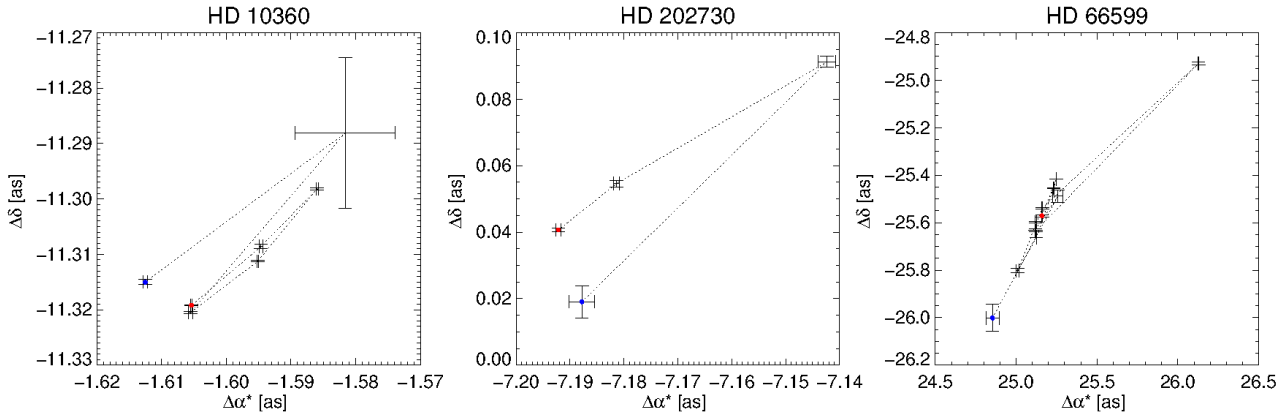


Figure 4.26. – Results in terms of star separations in right ascension and declination from several single epoch observations, where the timeline is indicated by the dotted line and the earliest observation sequence is marked with a blue dot and the latest by a red one. The scatter within the sky plane, together with the non-overlapping error bars, clearly depict a discrepancy between the instrument's alleged precision and its actual achieved accuracy.

and for HD66598_C17_ep2 to $\sigma \approx 2.8\,\mu$m, which, with the applied baseline, translate to about ≈ 4.5 mas and respectively to ≈ 7.5 mas on the sky.

The implication of the systematics being correlated with the azimuth or the hour angle of observation is even strengthened, when observations of several epochs are combined as one astrometric measurement.
This is accomplished by expanding the model given by Eq. (4.84) to a multi epoch fit with $N_{\rm ep} + 2$ parameters, where $N_{\rm ep}$ is the number of epochs and as a matter of course the number of constant terms that need to be derived and corrected for:

$$\Delta d = \pm B_p \cdot (\Delta s_\delta \cos\theta_{BL} + \Delta s_\alpha \sin\theta_{BL}) + \sum_{i=1}^{N_{\rm ep}} U_i Z_{0,i} \quad . \tag{4.87}$$

As before, the sign swapping is performed with respect to the mode the data is taken with, and the unit step function

$$U_i = \begin{cases} 1 & \text{for data within epoch}\,, i \\ 0 & \text{for data not within epoch}\,, i \end{cases} \tag{4.88}$$

determines, which of the individual zero terms is appplied to which measurement during the fit.

The combination of HD 66598 observations over three nights in Nov. 2011, with a resulting time span of $\Delta T = 2.14$ d between the first and last data point, is shown in Fig. 4.30 and convincingly illustrates the reproducibility of the residual structures.

The clearly shaped dependence on the hour angle, with a residual standard deviation of more than $5\,\mu$m and about ≈ 12.4 mas on the sky, already suggests that the source of these sys-

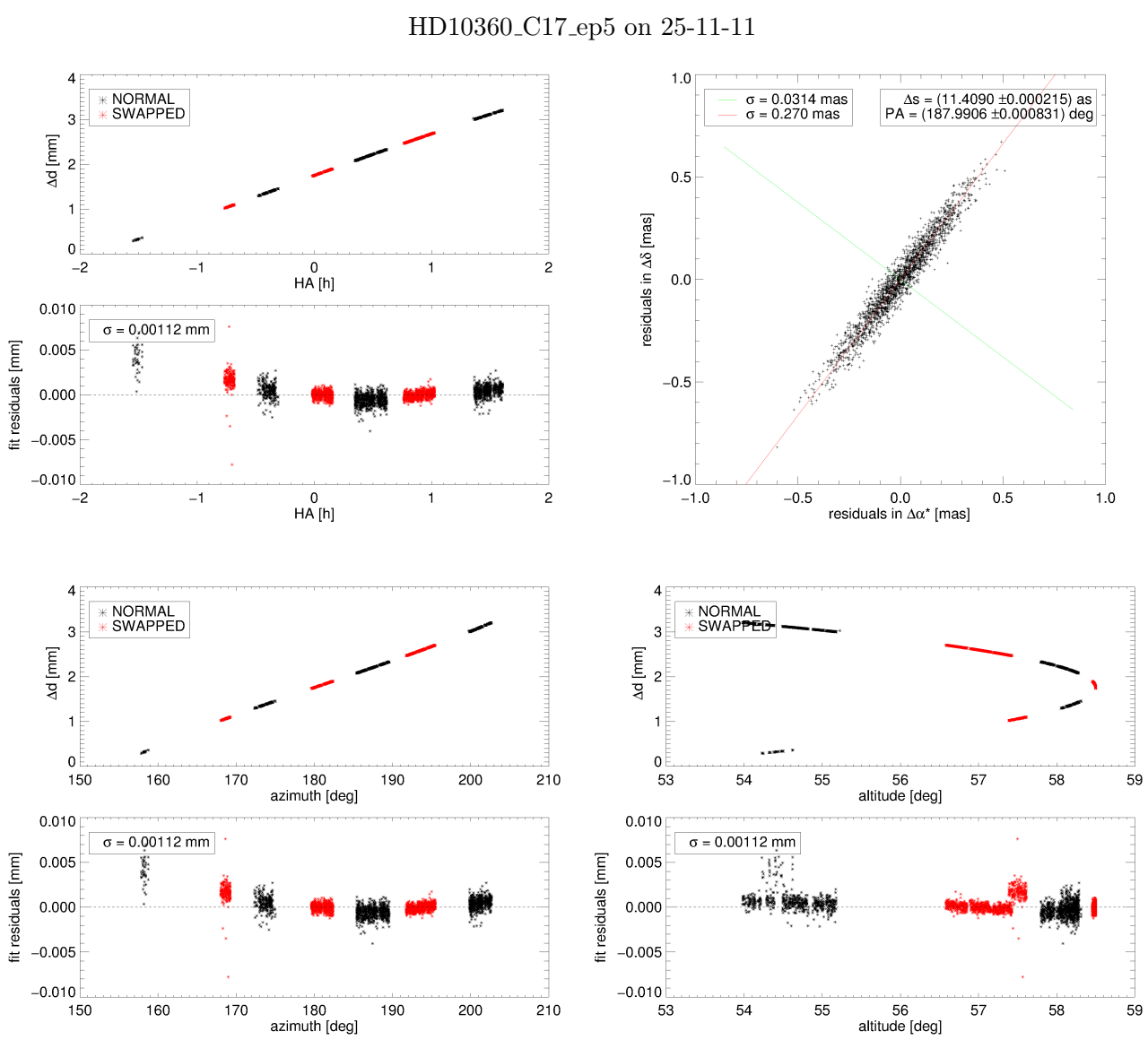


Figure 4.27. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_ep5 on 25-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis. Systematics, correlated with azimuth and hour angle are visible.

tematics might be correlated to the system's field rotation procedure that is performed by the de-rotator devices within the star separators of the telescopes.

In fact, the reason for this behavior is discussed in [Sah13]. It appears that misalignments of the instrument's optics, which have been detected in form of occurrences of obscurances of the stellar pupils, have been the major source for these large field rotation dependent systematics. After a following instrument intervention in March 2012, during which the system, in particular the ATs, has undergone some realignment, the situation improved, and the magnitude of the systematics was reduced by a factor of 5 in residual standard deviation.

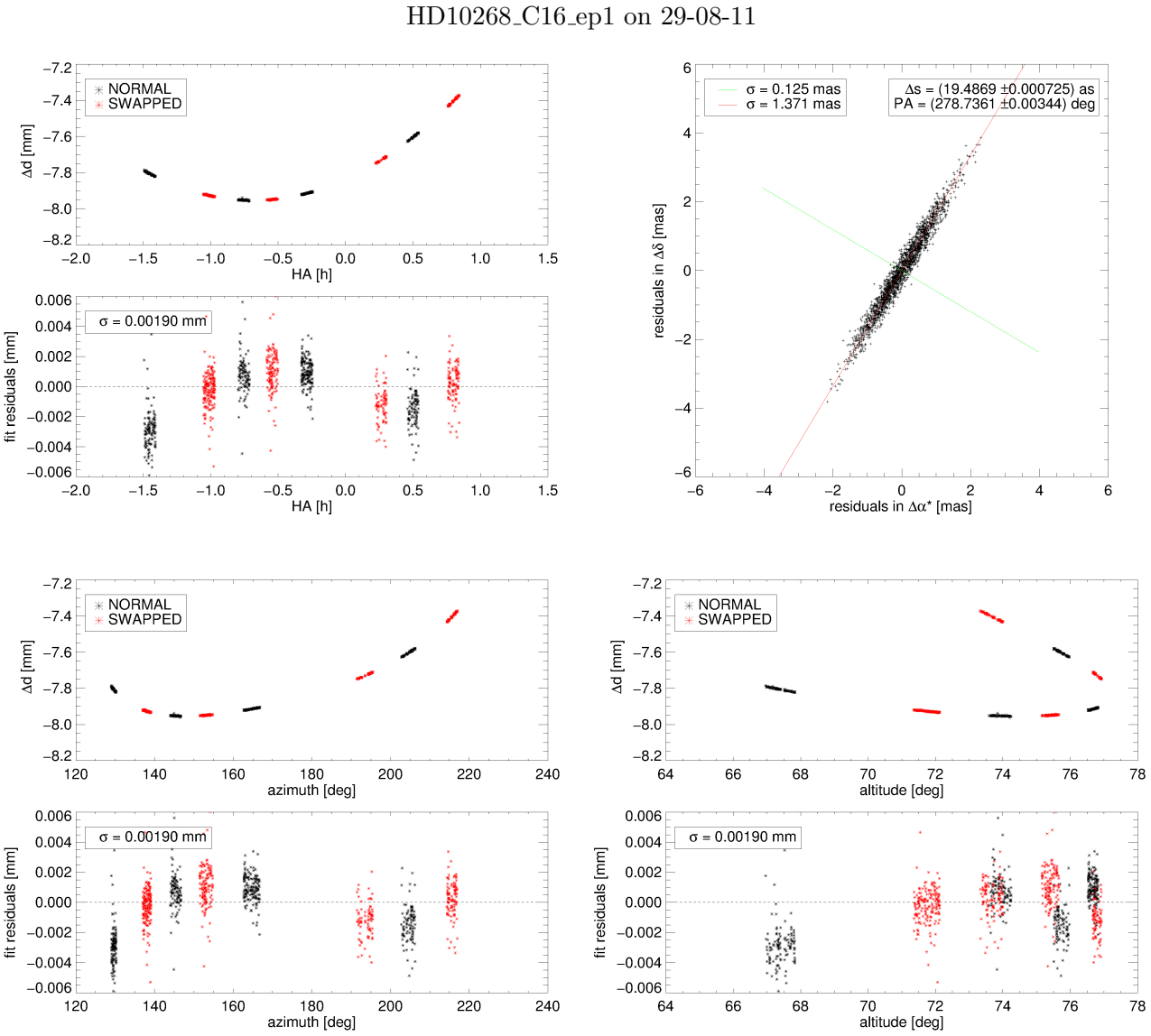


Figure 4.28. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10268_C16_ep1 on 29-08-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis. Well shaped systematics, correlated with azimuth and hour angle are visible.

All in all, it can be said that the instrument's precision, which is suggested by the 1-σ uncertainties of the fits and from the bootstrap analysis, as well as by the dispersions along the axis of highest precision during an observation (average baseline orientation), is promising with regard to the goal of having an astrometric instrument at the level of tens of microarcseconds. The best precision achieved, according to this analysis presented here, was obtained from an observation sequence on the system HD 10360 in Nov. 2011 (HD10360_C17_ep7), with an error of $\delta\Delta s = 60\,\mu$as and even $\sigma_{||} = 23\,\mu$as along the best axis.

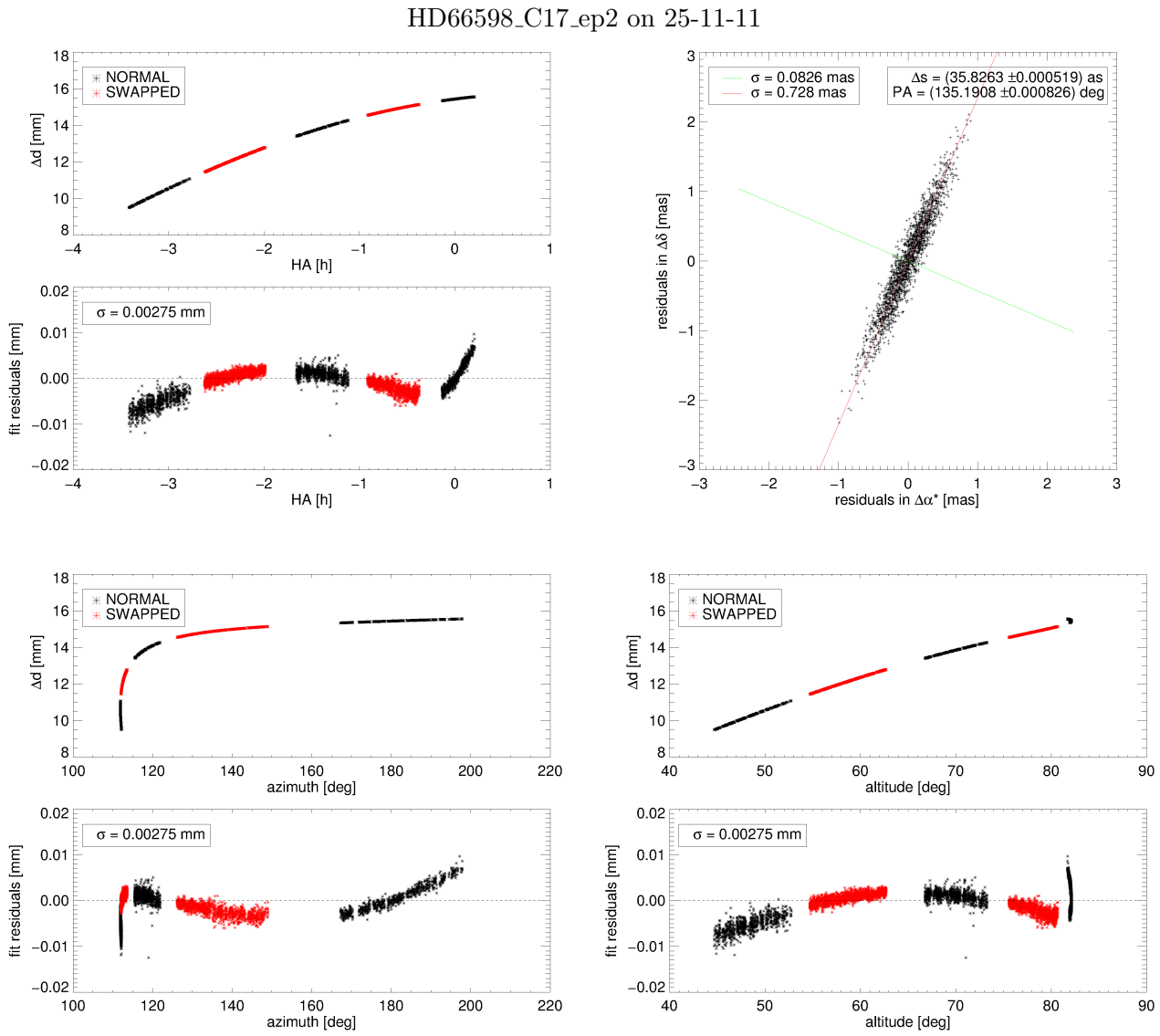


Figure 4.29. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C17_ep2 on 25-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis. Well shaped systematics, correlated with azimuth and hour angle are visible.

Still, these values explore only the level of precision, at which the differential delay measurements can be performed and which are valid only on short time scales.
If it comes to the accuracy of the instrument, the situation is different, as several observations of the same targets feature large scatters up to tens of milliarcseconds, which are too high to be explainable by the intrinsically changing system separations over time.
Although the realignment of the system seems to have decreased the amplitude of the systematics ([Sah13]), the science requirements are not fulfilled.
Several sources of remaining errors and possible improvements can be thought of. Besides the

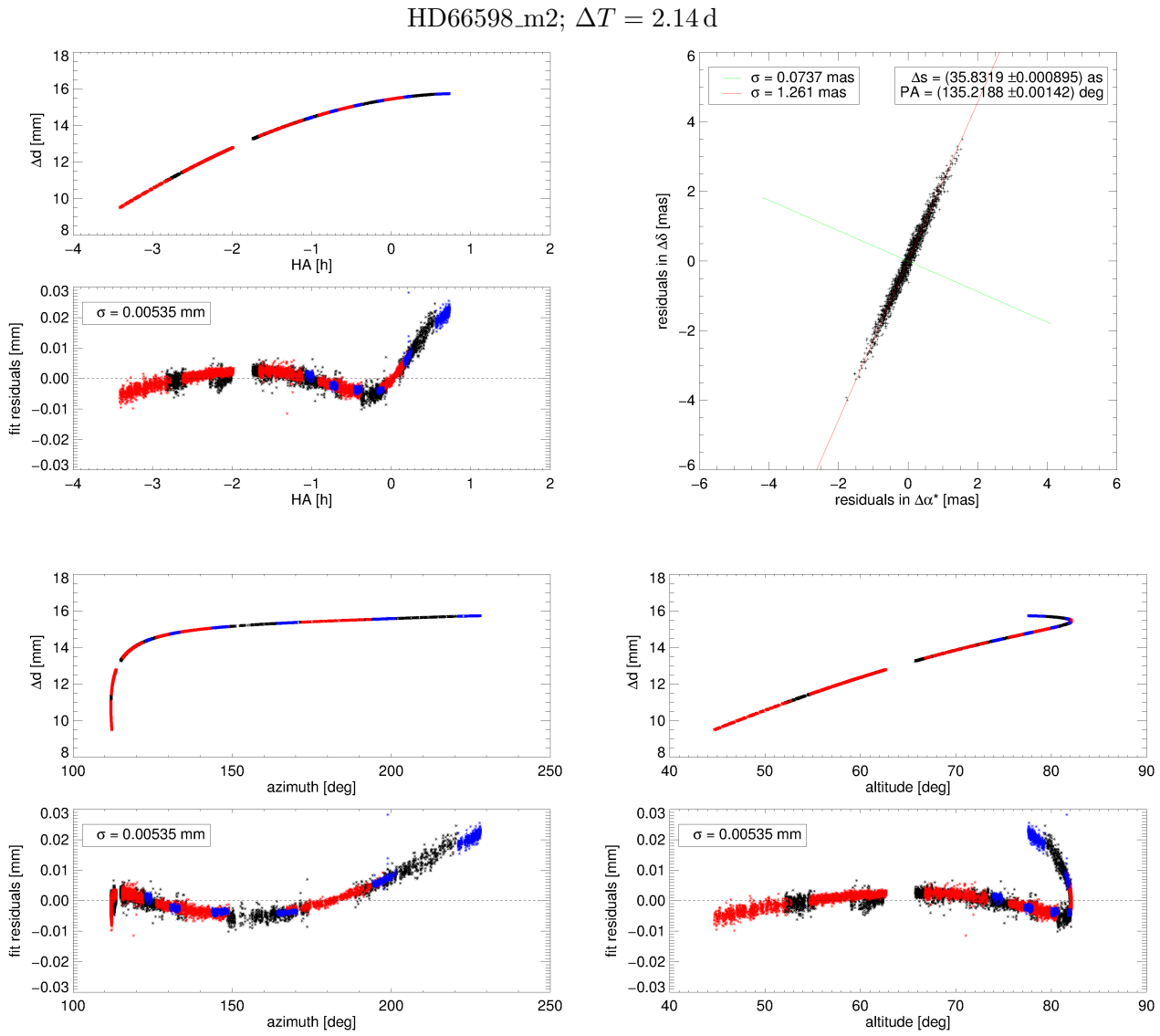


Figure 4.30. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit (see Eq. (4.87)) to the observation sequence HD66598_m2, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis. The three different epochs are color-coded, and as before a well shaped correlation with azimuth and hour angle is clearly visible.

fact that the astrometric analysis itself could be improved in terms of data reduction, as well as by modelling the systematics, if their structure can be correlated to some defined parameters, it must be taken into account that for this analysis the calibrated baseline was assumed stable and constant over the time span of all processed observation epochs.
This clearly is not the case, because any motions of the telescopes and their components, as well as any system's realignments, actions that are usually carried out between the commissioning runs, will in general introduce systematics of unknown magnitude, and therefore the system in

principle would need to be recalibrated.
Beyond that, one of the key factors, neglected up to this point, is presumed to be highly responsible for the remaining systematics and therefore for the lack of accuracy.
Recalling equation (4.81), and in particular the error terms introduced there, one sees that for the modelling of the differential delay one ends up with:

$$d_2 + \epsilon_2 - (d_1 + \epsilon_1) = \Delta d + \Delta\epsilon = \Delta\boldsymbol{s} \cdot (\underbrace{\boldsymbol{B}_\mathrm{W} - \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2}_{\boldsymbol{B}_\mathrm{NAB}}) + \underbrace{C_2 - C_1}_{\Delta C} \quad , \tag{4.89}$$

where the tracking errors have been neglected, and the metrology endpoints for the two individual interferometers, one tracking the primary at one FSU and the second the secondary target on the other, are assumed to be identical at the two telescopes.

The baseline, effective for the astrometric measurement, in this context also called *narrow-angle baseline* (NAB), is affected by the offsets of the metrology endpoints from the telescopes' pivot points, while the measured differential delay is influenced by any time dependent differential path discrepancies due to the unmonitored beam path above the metrology endpoint at M9 up to the entrance pupils at the M2s.
After the accomplished realignment of the instrument, the latter source of error is suspected to be the main remaining uncertainty.
In order to achieve the aspired accuracy in the astrometric measurements on the level of 10-50 μas, these systematics need to be further examined, quantified and calibrated.
Furthermore, an expansion of the PRIMA metrology beyond its actual endpoints is under consideration.

Table 4.10. – Results from the multi epoch analysis. For each processed multi epoch observation, for which the included individual epochs are specified in App. C, the overall time span of observation, the number of valid 1-second averaged data points and the astrometric observables with errors from the astrometric fit, defined by Eq. (4.87), are listed. Also tabulated are the corresponding 1-sigma uncertainties from the bootstrap approach and the dispersions $\sigma_\parallel$ and $\sigma_\perp$, along the directions of the axes of maximum and minimum precision from the principal component analysis.

sequence ID	ΔT	N	Astrometric fit		Bootstrap				
			Δs	θ_s	$\delta\Delta s$	$\delta\theta_s$	$\langle\theta_{BL}\rangle$	$\sigma_\parallel$	$\sigma_\perp$
	[days]		[mas]	[deg]	[mas]	[deg]	[deg]	[mas]	[mas]
HD10360_m1	6.19	16862	11429.049 ±0.047	188.04785 ±0.00028	0.053	0.00034	143.79	0.020	0.084
HD10360_m2	91.92	18029	11429.393 ±0.057	188.05189 ±0.00034	0.055	0.00042	148.87	0.023	0.10
HD108248J_m1	5.03	1009	3909.76 ± 1.94	112.7736 ± 0.0387	2.73	0.0544	122.42	0.063	4.61
HD108248J_m2	5.03	700	3909.78 ± 2.10	112.7734 ± 0.0418	2.92	0.0583	122.39	0.063	4.93
HD156274_m1	0.15	3380	10000.88 ± 0.94	257.4250 ± 0.0092	0.85	0.0084	117.96	0.171	1.69
HD202730_m1	5.05	3056	7183.71 ± 0.30	270.4123 ± 0.0035	0.30	0.0035	123.56	0.055	0.53
HD202730_m2	40.02	3659	7181.74 ± 0.33	270.4313 ± 0.0038	0.31	0.0037	123.34	0.058	0.56
HD66598_m1v2	8.06	7255	35886.82 ± 1.10	135.5540 ± 0.0017	1.54	0.0024	123.16	0.272	2.15
HD66598_m2	2.14	11751	35831.93 ± 0.72	135.2189 ± 0.0011	0.90	0.0014	113.66	0.074	1.26
HD66598_m3v2	303.20	19004	35867.78 ± 0.77	135.3960 ± 0.0012	1.37	0.0022	117.70	0.133	1.93

Chapter 5

ESPRI: Exoplanet Search with PRIMA

As pointed out broadly within the precedent chapters, the design and actual implementation of the instrument PRIMA was mainly motivated by the prospect of a dual-feed interferometer with the capability of performing high-precision astrometry on the level of tens of microarcseconds, by which extrasolar planets could be detected.
With this application in mind, ESO was supported by an international consortium, consisting of working groups from the Observatoire de Genève, the Max Planck Institute for Astronomy and the Landessternwarte Königstuhl (Heidelberg), in the development of the key components, both on the hardware and software side, for which in turn the consortium has been awarded guaranteed observation time (GTO) for the years after, once the instrument is successfully commissioned.
Based on this perspective, the consortium members have mutually found a program for the astrometric search of extrasolar planets, ESPRI (Exoplanet Search with PRIMA), which has been planned and prepared during the recent years.

As radial velocity surveys and transit based search programs have proven to be highly effective and successful within this branch of research, ESPRI, based on the astrometric method, should be regarded as an complementary approach to the key questions in this field.
The potential of the applied technique lies in filling up the gaps in the exoplanetary detection space, which result from restrictions involved in the other methods. The orbits, determined by the transit method, are conceptionally limited to usually high inclinations, and the Doppler spectroscopy of radial velocity measurements relies on sufficiently narrow and well-defined lines in the stellar spectrum, as well as on the star's stability, which significantly limits the type of stars, whose signals can be examined at all.
Considering the capability of exploring these, otherwise hardly accessible, objects and possible exoplanetary systems and the advantage of determining the detected systems' inclinations, as the true orbits can be deduced from the two dimensional measurements of the stellar reflex motion, some certain key questions and aspects have been focused on during the planning and design phase of the program.
Based on those, it has been initially decided to concentrate the efforts and observation time on three groups of possible science targets ([Ref06]):

1. **RV stars**:
   Stars, around which companions have been already found by the radial velocity method, are not only suitable for the system's ultimate science verification, but the information about the systems can be significantly increased by follow-up astrometry, as the ambiguity between the companion's mass and inclination can be solved and the true mass can be

derived from the time resolved observations.
Furthermore, since, in contrast to RV detections, the astrometric method is more sensitive to wider and consequently longer orbits, so far undetected additional companions could not only be found, but the multiple systems could be investigated regarding their degree of coplanarity.

2. **Nearby stars**:
   As the astrometric signature of planet harboring stars, in terms of the measured angular amplitude on the sky, is reciprocally proportional to the distance between the examined system and the observer, the most nearby stars, with distances below ~ 15 pc, are the most promising target candidates and therefore serve best for this kind of a survey.
   With given astrometric accuracies at the level of tens to hundred microarcseconds, Saturn like planets at reasonable orbit extents around these kind of stars can be detected (see Fig. 2.21), which is why a blind search throughout a sufficiently large sample has been foreseen.

3. **Young stars**:
   In particular with respect to the question of planet formation and the evolution of planetary systems it is substantially interesting to investigate the occurrences of detected companions around stars of various stages of evolution. In this context the astrometric method might prove itself superior to the RV technique, by which young stars are usually omitted due to their activity that can spoil the Doppler signal.
   In order to study the early stages of planetary systems, young stars, with ages in the range of 5–300 Myr and distances up to $\sim$140 pc, have been considered.
   However, depending on its amplitude and time scales, stellar activity could also affect the astrometric measurements, if a displacement of the photocenter in the observing K-band is introduced. Although the time scales of the possibly most problematic effects, such as non-radial pulsations or astrometric modulations, rotationally induced by asymmetric distributions of starspots, can be considered distinctly shorter than the periods of the companions' orbits, their signatures, which in unfavorable cases and nearby stars could reach up to about 10 μas ([Set07]), would contribute to some astrometric noise if accuracies at this level are achieved in the long term.
   Therefore, the stellar activity of the young science target stars needs to be studied independently in order to minimize the number of false positive detections.

5.1. Preparatory Observations and Target Selection

After defining the types of stars that would be observed during the systematic astrometric exoplanet search with PRIMA, in order to approach the key questions, such as the general investigation of the planetary mass distribution, the analysis of formation processes of multiple planetary systems and also the study of the systems' evolution depending on the stellar age, the next task, fundamental to the success of the survey, was to preselect suitable science targets.
Due to the specific method of observation and search conception, the objects in question need to fulfill several requirements in order to meet with the technical constraints imposed by the

instrument.
A substantial set of aspects must have been considered during this phase of preparation. Besides the elemental restriction of the targets' position on the sky, which results from the observing site's location at Paranal, limiting the observing angles of declinations to the range between around $+27°$ and $-83°$ over the year, and in addition to the fact that the angular extent of the star, depending on the combination between its diameter and distance, must not be resolved, which would cause a drop in the visibility, the objects' brightness is the most critical property regarding the effectiveness of fringe tracking.
Since at this point of time the brightness limits of the system could not have been predicted with certainty, the limiting magnitudes for the possible primary science targets have been set to the initial science goal of 13 mag in K-band and for the secondary reference stars, at which due to the benefit from the phase referencing longer coherence times would be achieved, to the most optimistic limit of $K = 16$ mag.

The second and most critical factor, with regard to the search technique, is the availability and distribution of adequate reference objects in the vicinity of the target science star.
Assuming that the reference star, to which the separation on the sky is being determined, is bright enough for fringe tracking, according to Eq. (3.1), its angular distance to the primary source is linearly correlated with the theoretical astrometric error. Therefore, one is interested in secondary stars within the isoplanatic patch around the primary targets, and for the search of suitable references a limit of 30 arcseconds has been agreed on, while separations below 10 as are pursued in order to achieve the highest possible accuracies.
However, at the other side of the range the sources' angular distance must not fall below a critical limit, at which cross-talk between the input channels might occur, as the star separators fail to properly isolate the light from the distinct sources, which must be fed into the system on different light paths. In order to avoid this, objects with apparent separations below 2 as have been decided to be refused.
The number of the possible reference stars that could be used for the differential astrometric measurements and their distribution within the field are also matters of considerable importance.
Although in the typical case, where the science targets are situated relatively near to the observer, especially when compared to the references that are supposed to be bright enough but distant background stars, any astrometric signal, which in fact is distance depending, would most probably be resulting from relative motions of the primary sources, distinct measurements against multiple references are beneficial, as these would allow to validate the occurrences and eliminate false detections due to motions of the secondary objects.
Also, as the separation measurements are one-dimensional along the orientation of the projected baseline, well distributed secondary objects, in terms of position angles with reference to the primary target, are helpful to reduce the uncertainty in the determination of the motion directions on the sky and improve the real orbit determinations, when the measurements are combined, as it results in less redundancy regarding the relationship between the position angles and the direction of the time dependent projected baselines.
Apart from these main criteria for the search of suitable targets, the observed stellar objects need to be characterized for calibration reasons, since for example color differences could induce

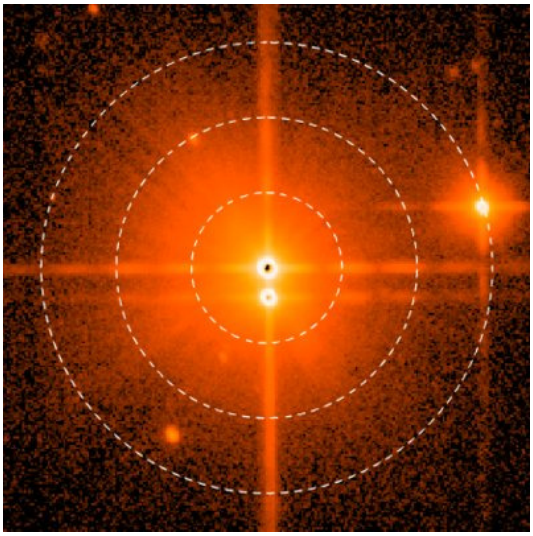
differential dispersion effects, and in the sense of astrometric stability, which could be affected by stellar activity or any photocenter shifting mechanisms.

The available standard catalogs, as the 2MASS Point Source Catalog ([Skr06]) and others, have been studied with the aim of defining a first and preliminary preselection of suitable science targets, based on the main selection criteria regarding the observability, the brightness and the availability of reference targets within the preferred angular range.
Unfortunately, those archives have turned out as not satisfactorily adequate for this purpose due to several reasons. At the one hand, their sensitivity in K-band is usually considerably below the $K = 16\,\mathrm{mag}$ limit for the secondary stars, and on the other hand it has been found that saturation of bright objects often results in bright halos around them so that reference stars within the preferable zone of 10 as could not be found.
Mainly for these reasons a preparatory observation program has been planned and carried out by the ESPRI team, whose main purpose was to observe possible science targets and to look for suitable reference targets in their surroundings.

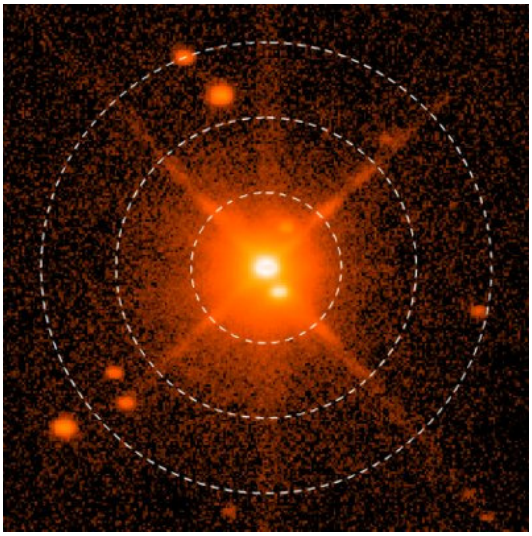
The program was started in 2004 and lasted for several years, during which photometric observations in J- and K-band of nearly 1,000 stars have been carried out with *SOFI*, ESO's infrared imaging camera on the 3.6 m telescope NTT at La Silla in Chile and with *Omega-Cass*, the MPIA's infrared camera installed at the 3.5 m telescope on Calar Alto, Spain.
The results of this survey have been used to set up the database *ASTRID* (AStrometic Target & Reference stars Interactive Database), in which information about the possible target stars and, if detected, also about their associated reference stars are stored.
Depending on the instrument's final performance and the resulting criteria, mainly in terms of brightness and availability of reference stars, the suitable science targets can be then recovered.
In order to obtain the required data from the observations, the obtained images have been processed by the standard means of astronomical data reduction, during which they have been corrected for typical sensitivity variations over the detector pixels (flat fielding, bad pixels, etc.).

A jitter imaging technique (see [Dev1999] for principles) has been applied to enhance the sensitivity of the observations. It is effective in separating any sky background variations from the real astronomic signals, which is achieved by taking multiple exposures of the same source, with the telescope pointing varying from exposure to exposure by some defined small offsets. From those initial images sky background variations are estimated and filtered so that the resulting recombined products' photometric sensitivity is significantly improved due to the subtraction of the background signal.

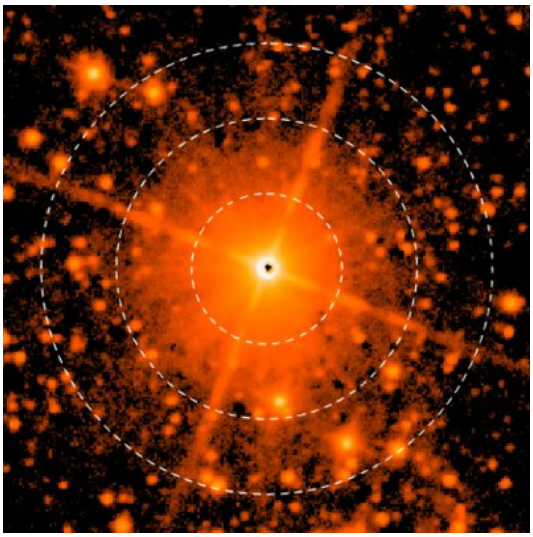
The final reduced images, of which four examples are presented in Fig. 5.1, have then been analyzed for any occurrences of suitable secondary stars.
This has been achieved by the usage of the *Source-Extractor* ([Ber1996]) software, which is designed to automatically detect stellar sources within astronomical images. It recovers the amount of flux that is associated with the corresponding detected objects and determines the positions of their photometric barycenters.



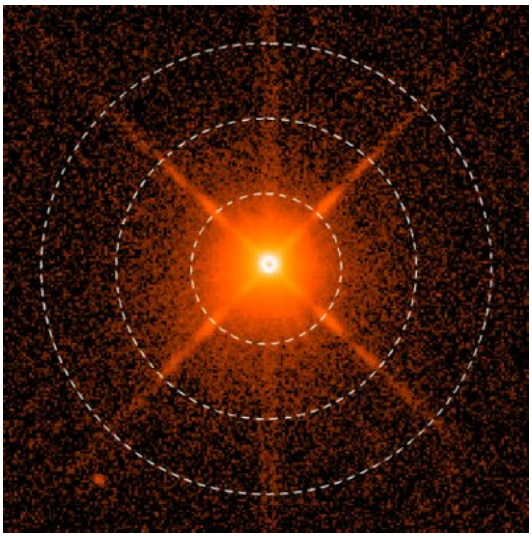
(a) Target young star candidate ($K \approx 5$ mag) with potential reference stars at separations 3.4 as ($K \approx 8.3$ mag), 25.5 as ($K \approx 12$ mag) and 29.7 as ($K \approx 7.9$ mag).



(b) Target young star candidate ($K \approx 6.5$ mag) with two potential reference stars at separations below 10 as, namely at 3.2 as ($K \approx 9.2$ mag) and at 6.1 as ($K \approx 13.8$ mag). Three more references with $K \approx 13.5$ mag, $K \approx 15.9$ mag and $K \approx 15.7$ mag are situated at separations of around 25 as.



(c) Nearby star candidate ($K \approx 5$ mag) with several potential reference sources, the brightest and closest with $K \approx 10.7$ mag at 17.7 as. However, the field is pretty crowded, which might disturb fringe tracking at the dim secondaries.



(d) Observed radial velocity target ($K \approx 6.4$ mag) with no suitable, bright enough reference stars.

Figure 5.1. – Four reduced sample SOFI K-band images as obtained from the preparatory observations. The science targets are centered, while the dashed circles indicate angular distances at 10, 20 and respectively 30 as. (The images have been created by R. Launhardt for the database ASTRID.)

The resulting fluxes have been then converted into K- and J-band magnitudes by a calibration onto the scale of the 2MASS catalog, wherefore for each night and filter several stars have been compared to their catalog's brightness entries to recover the scale zero point.
Any findings of potential secondary stars with $K < 16$ mag and angular distances, with reference to their primary stars, below 30 as have been stored for further consideration into the database.

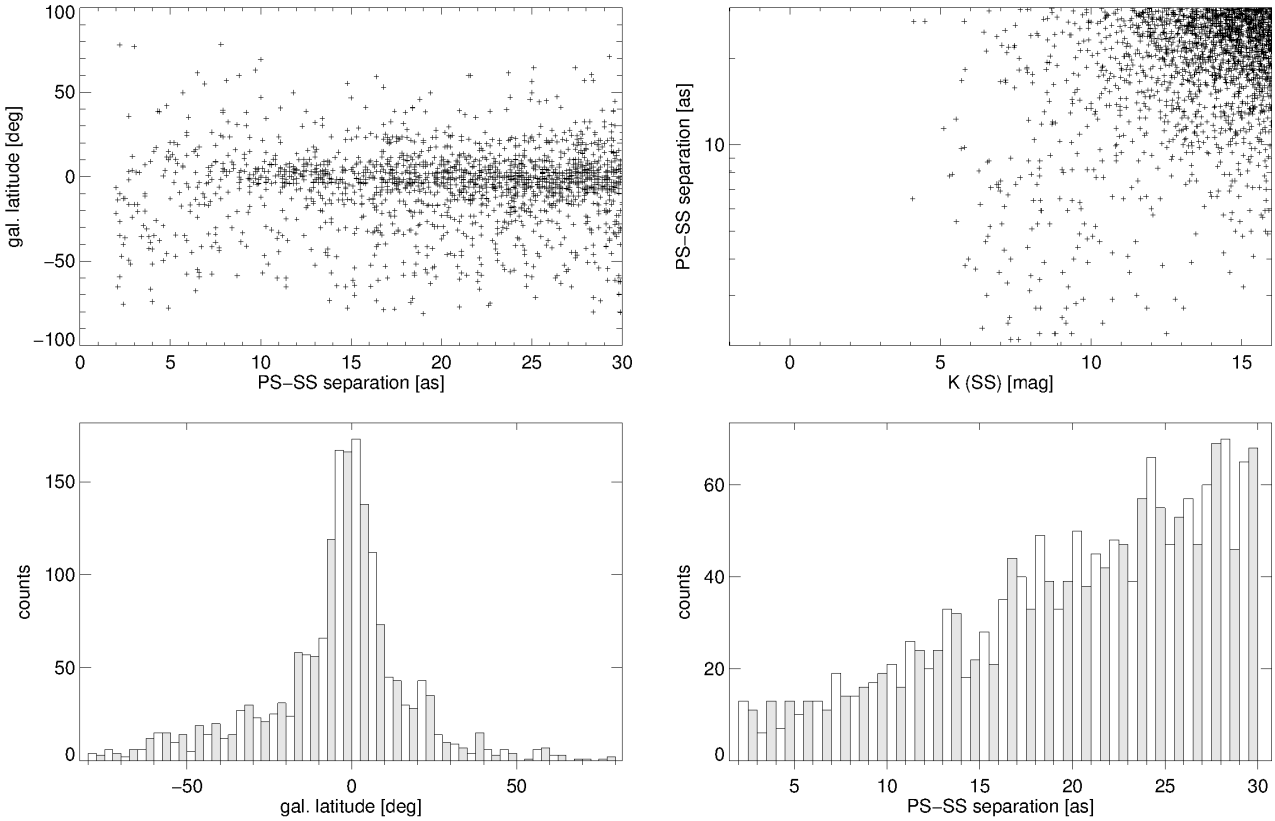


Figure 5.2. – 1,877 potential reference stars around 572 science target candidates, which have been detected by the analysis of the preparatory observations, are plotted here against their positions on the sky with reference to the galactic plane (*left*) and their distributions, regarding the angular distances to the primary sources and their K-band brightness (*right*).

Based on this gathered information from the preliminary photometry and astrometry, possible science targets can be selected from the pool of target candidates.
Fig. 5.2 summarizes the findings from the preparatory observations, where 572 primary targets, with at least one reference star candidate, have been taken into account. All in all, 1,877 potential secondaries are being considered. As expected, one can see that the probability to find suitable secondary stars is highly increased towards low galactic latitudes, hence towards the galactic plane, where the density of observable stars is naturally higher.
However, some of the observed fields are considerably crowded, which could lead to disturbing

noise signals and impose problems during fringe tracking on some references (see Fig. 5.1 (c)).

Looking at the distribution of the potentially suitable secondary targets, one can see that their occurrences accumulate with increasing separations and decreasing brightness. Ultimately, the decision, which of these can be used for the ESPRI survey, will be highly affected by PRIMA's performance, particularly considering the brightness limit for fringe tracking of both, the primary and the secondary sources.

5.2. Program Prospects

Based on the analysis of the preparatory observations and also spectroscopic characterizations in several cases, the ESPRI team is sufficiently endowed with a collection of many suitable target stars at hand, and is in principle prepared to begin the science program observations, whenever the instrument will be ready for science operation.
Unfortunately, the delays due to unexpected problems during the at present incomplete progress from the system installation at the VLTI infrastructure to a fully commissioned instrument in terms of astrometry, but also efficiency, involve not solely certain uncertainties regarding the question, what will be achievable compared to the initially defined goals, but also entail rather economic considerations, regarding the planned survey's expected outcome, in particular with reference to concurring projects as for example ESA's space mission *Gaia*.

Facing the already apparent but initially unexpected limitations in terms of brightness limits for fringe tracking with PRIMA, it is highly important to have estimations on the number of science targets, which can be observed in the end and to be flexible at adapting the survey to any instrument based restrictions.
The gathered information in ASTRID is essential in this context. In order to get a feeling, in which cases the program might encounter a lack of appropriate targets, the 572 primarily considered target star candidates (see Sec. 5.1) have been downselected by applying more stringent constraints of qualification.
Figures 5.3–5.5 illustrate different distributions of 255 primary stars candidates that remain under consideration after employing stricter constraints. The fringe tracking limits for this subsample have been relocated to $K = 8\,\mathrm{mag}$ for primary and respectively to $K = 14\,\mathrm{mag}$ for secondary targets. Furthermore, besides limiting the range of preferred target positions on the sky to allowed objects' declinations between $+20°$ and $-70°$, star group dependent distance limits have been refined to $d < 100$ pc for the young stars and $d < 15$ pc for the stars falling into the nearby star category.
In addition to that, known spectroscopic binaries and targets within crowded fields have been also removed, although it should be noted that the minimum angle between an object, on which fringe tracking is being performed, and its nearest neighbor, before the effect of wavefronts overlapping becomes noticeably disturbing, is also a question of brightness difference and is still to be investigated for PRIMA.

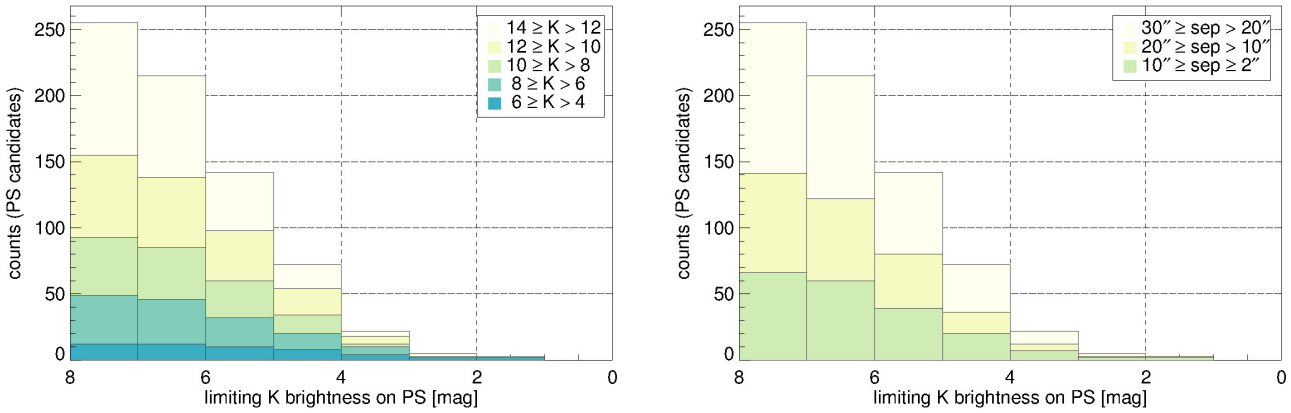


Figure 5.3. – Cumulative histograms of the 255 target star candidates (see text for subsample restrictions) as a function of the fringe tracking brightness limit on the primary sources. *Left:* The cumulative numbers are subdivided into color coded subclasses of different brightness ranges of the brightest secondary source linked to the corresponding primary target. *Right:* The cumulative numbers are subdivided into color coded subclasses of different angular separation ranges to the brightest secondary source linked to the corresponding primary target.

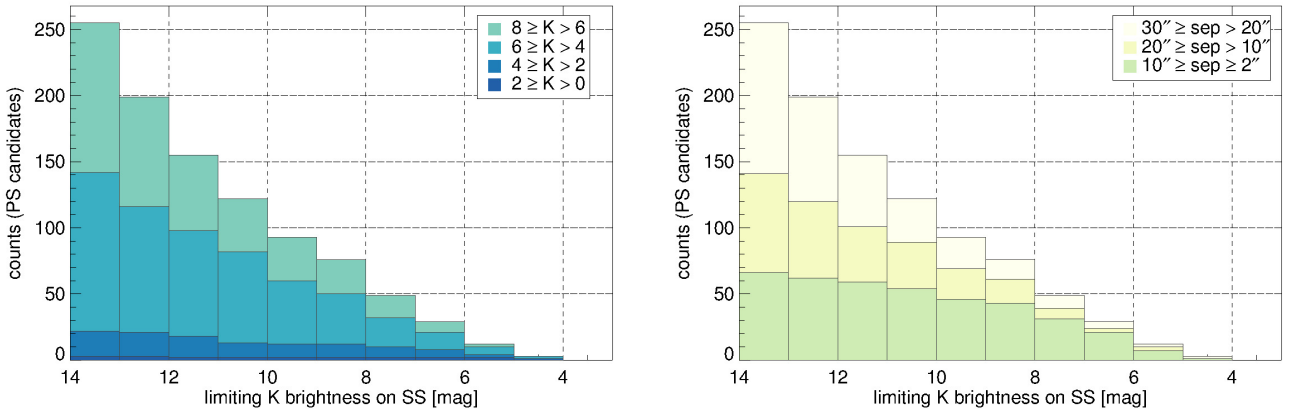


Figure 5.4. – Cumulative histograms of the 255 target star candidates (see text for subsample restrictions) as a function of the fringe tracking brightness limit on the secondary sources. *Left:* The cumulative numbers are subdivided into color coded subclasses of different brightness ranges of the primary targets. *Right:* The cumulative numbers are subdivided into color coded subclasses of different angular brightness ranges to the brightest secondary source linked to the corresponding primary target.

Again, the importance of the final fringe tracking limits, depending on the brightness of the primary, as well as of the secondary sources, are stressed out by the evaluated histograms. Assuming that the system will be able to track primary stars at a sufficient signal to noise ratio down to $K_{\mathrm{PS,lim}} = 8\,\mathrm{mag}$, the number of the possible science target candidates is mainly

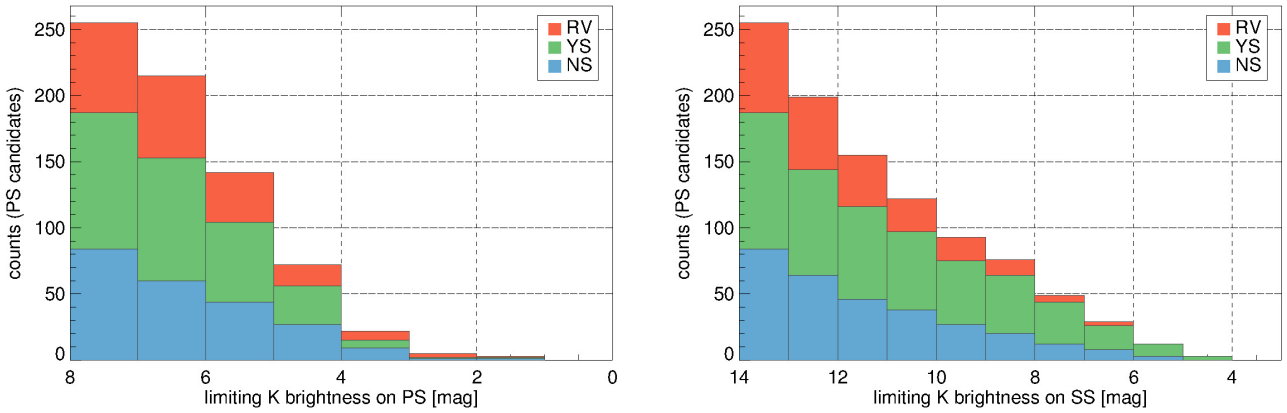


Figure 5.5. – Cumulative histograms of the 255 target star candidates (see text for subsample restrictions) as a function of the fringe tracking brightness limit on the primary sources (*left*) and the secondary stars respectively (*right*). The cumulative numbers are subdivided into ESPRI's three different star group categories (RV=Radial Velocity stars, YS=Young stars, NS=Nearby stars).

determined by the secondary targets' brightness limitations.
While at a limit of $K_{\mathrm{SS,lim}} = 12$ mag one would be still left with around 160 targets, only around 100 targets would remain observable, if the limit would be found to be decreased by two magnitudes down to $K_{\mathrm{SS,lim}} = 10$ mag.
The situation becomes more complicated, when the separations to the secondary targets, which affect the achievable astrometric accuracies, are taken into account. Among the discussed subsample, less than 70 out of the 255 target candidates offer potential secondaries within the mostly favorized range of separations below 10 as, and 141 feature their brightest references at angular distances up to 20 as.
In this context it should also be considered that some of the candidate primary-secondary pairs, in particular those with both components being relatively bright, might still be physically bound systems. Although this number should not be too high, a closer look at the systems might disqualify them for science operations.

The cumulative distributions as a function of the fringe tracking limits, displayed in Fig. 5.5, give an additional overview of the numbers with respect to the star categories involved in the ESPRI program.
In particular, the number of suitable target RV stars candidates is strongly affected by the fringe tracking limit of the secondary stars. While 68 out of the 255 stars are radial velocity stars with known companions, which makes them quite valuable targets for the program, as the study of those guarantees positive results, provided that the expected astrometric signature lies within the instrument's level of accuracy, so that they can be used as save sources for science verification, not more than 18 of them could be observed with $K_{\mathrm{SS,lim}} = 10$ mag.
The other 187 among the 255 stars subsample are subjects to the blind exoplanet survey. Their number is particularly important, as the number of positive results, in the sense that an as-

trometric signal due to an unseen companion is obtained, solely depends on the statistical probability, determined by the occurrence rate of detectable planets around them.
As these numbers are only vaguely known, an estimation of the occurrence frequency of giant planets with masses above $0.47\,M_{Jup}$ around solar type stars at orbits up to $\sim 5\,AU$ has been derived by [Nae05] to be about $7.3 \pm 1.5\,\%$, it is probably reasonable to state that less than $10\,\%$ of the observed stars will exhibit a detectable signature.
These numbers naturally depend on the ultimate detection limits of PRIMA, which at this moment can only hardly be estimated, supporting the importance of having a flexible and strong in number star candidate collection.

The temporal delay of the ESPRI program also becomes an issue, as concurring programs are carried out. In particular ESA's space mission Gaia, which according to [ESA13] is expected to be launched in October 2013, will acquire data that will also allow astrometry in the microarcsecond regime. This entails a huge overlap not only on ESPRI's science goals but also specifically on the individual science objects themselves.
After its launch and after Gaia has reached its final destination, a Lissajous orbit around the second Lagrange point (L2), the spacecraft will begin its 5-years lasting science operations, during which the whole sky will be observed in a scanning mode. At this time, its astrometric instrument, operating in a broad white-light Gaia (G) band, defined by the wavelength coverage of around 330–1050 nm, will be able to measure the relative separations of any objects within the field of view within the brightness range between $G = 6\,mag$ and $G = 20\,mag$.
It is expected that by this method about 70 sets of measurements will be gathered on each observable star and that a final accuracy down to $25\,\mu as$ can be achieved for all sources brighter than 15 mag, and even less than 10 mag for objects at $\lesssim 12\,mag$. This astrometric survey will thus be able to sample about 150,000 FGK stars up to a distance of around 200 pc and optimistic estimations implicate that about 2,000 exoplanets in single systems can be expected to be found, with distance depending detection limitations on the companions' mass.

In the context of this perspective and with ESPRI's given strict time budget, it is being discussed to expand the sample of possible science targets, in particular the candidate list for the blind survey, with a degraded emphasis on the sources' distances, but more with the objective to find targets complementary to those observed by Gaia. The most promising approach in this sense is to examine stars, which might be too bright for Gaia due to its saturation limit around $G = 6\,mag$. As the so far gathered ESPRI sample is presumed to be rather complete for such bright stars at spectral types between G0 and M2 ([Lau12]), objects at the tails of this range are being considered, and the available catalogs are searched for possible target candidates in the spectral range of A0–F5, but not earlier, as the occurrence rate of giant planets around more massive stars is decreasing significantly. On the opposite end of the spectrum, M dwarfs with M1.5–M5 are being investigated.

The next step towards PRIMA's scientific operation and the beginning of the ESPRI program is the understanding and minimization, or a sufficient modelling of the instrument's systematics, which at present manifest themselves within the astrometric measurements (see Sec. 4.3.3).
As the astrometric method is increasingly sensitive to wide exoplanetary orbits, results from

the blind search cannot be expected before one or two years after the start of the observations. However, several classes of objects could serve as early science objects or also provide means for the instrument's science verification. Besides low motion visual binaries, which naturally offer a way to probe the measurements' stability and reproducibility, radial velocity stars with high anticipated astrometric amplitudes could be the first objects, at which the reflex motions could be detected and verified. Additionally, stellar hierarchical triple systems, consisting of two visual components on long orbits, bright enough for fringe tracking, and where one of these in addition is a short-period spectroscopic binary, could be used for the first and quickest orbit determinations.

Once it is verified that astrometry can be performed at the 50 μas level, and beyond this the operation of the system becomes sufficiently reliable and stable in terms of repeatability, but also regarding the fringe tracking sensitivity so that the brightness limitations remain sustainable, the ESPRI survey can be initiated.

Chapter 6

Summary

This thesis has been designed to provide a framework for the presentation of the development and implementation of the dual-feed interferometer PRIMA, one of the next generation VLTI instruments at the Cerro Paranal in Chile.
After the system's installation and commissioning it will allow simultaneous fringe tracking on two separate sources, by which highly precise differerential astrometric measurements at the level of tens of microarcseconds will become realizable.
The instrument's set-up and its characteristics are mainly motivated by the search for and exploration of extrasolar planetary systems.
This fairly young field in observational astronomy is crucial to our understanding of the formation and evolution of planetary systems and the physical mechanisms behind it, as the findings of more and more systems and the characterization of their properties serve as the ultimate testing environments for theoretical elaborations.

Since the beginning of this new scientific era, when exoplanet detection became possible and reliable during the 1990s, hundreds of extrasolar systems have been discovered and examined. However, the nature of the systems in question is rather challenging in the context of direct observation. The immense contrast in brightness between the host stars and their companions, as well as the relation between the distances to the observed systems and the corresponding orbits' spatial extents require high instrumental sensitivity and extraordinary observing strategies.
For these reasons, mostly indirect techniques have been developed and have proven to be particularly successful in this field.
Doppler spectroscopy, by which the stars' radial velocity profile is searched for any periodic signals due to invisible companions, was the detection method of first choice for a very long time. Meanwhile, photometric surveys became sufficiently sensitive to detect dips in recorded lightcurves, which are induced by occultations of the stellar disk by companions on orbits with high orbital inclinations. Nowadays, this transit method became also very powerful, and in particular the space based survey Kepler alone has already identified several hundreds of exoplanet candidates.
Still, although the techniques are continuously improved, the indirect approaches suffer from conceptional limitations, as the methods' sensitivity is dependent on certain system parameters and configurations, so that exoplanet surveys are usually subjects to biases and can be considered as complete only within their characteristic ranges of detectability.
This is why it is desirable to have different complementary strategies to overcome those difficulties.
The astrometric detection method could serve as such one. Although, just as the RV technique,

it is based on the detection of the stars' reflex motions, those are derived not from velocity measurements along the direction of the line of sight, but from precise determinations of the stellar positions over time on the sky. Consequently and contrarily, it is more sensitive to companions on wide and long orbits and additionally is not affected that much by the orbital inclination. While only a minimum mass of the companion can be deduced from RV measurements, by the means of astrometry the real orientation of the orbit with reference to the observer and also the true planetary mass can be deduced.
On the other hand, since in reality the projection of the reflex motion on the sky is measured, the detectability regimes are strictly limited by the combination of the individual instrument's precision and the observed system's absolute distance.

Aiming at reaching the capability for the detection of astrometric signals from Saturn-like planets around nearby stars, reliable positioning needs to be performed at accuracies below 100 μas. This can be achieved by differential measurements, where the science target's position over time is compared to that of one or several reference sources. PRIMA has been primarily designed with this objective in mind and allows simultaneous fringe tracking on two separate sources. The astrometric quantity, the sources' angular displacement on the sky, is then derived from the measured difference in delay that needs to be applied in order to compensate the external path differences to the two apertures and to detect the individual fringes of each of the targets.
In addition to that, the simultaneous observation provides the means to extend the coherence time on one of the sources, as the other can be used as a phase reference so that sensed fringe motions, which are introduced by atmospheric turbulences, can be compensated for both at the same time.
Consequently, the brightness limitations on the secondary source are loosened and dimmer targets can be observed. In order to achieve this, it must be ensured that the light from the two stars is affected by the atmosphere in a sufficiently equal manner, which is the case if they are not separated by angular distances greater than the isoplanatic angle.
In theory, the instrument, which is operated in the infrared K-band, should be capable to achieve the aspired accuracy level, when interferometric baselines of the order of 100 m are used and stars with separations below 30 as are observed.

In order to perform these kind of measurements, PRIMA has been equipped with components, which are particularly necessary for double star interferometry. Together, the star separators, the differential delay lines, the laser metrology system and the fringe sensor units provide the possibility to select two particular sources from the field, to feed them into the system, and to monitor and control the applied differential delay in principle at precisions below 5 nm and at frequencies above 200 Hz, which allows stable fringe tracking on both sources.
The real-time fringe phase measurements are achieved within the fringe sensor units by the means of spatial phase modulation, where the fringes are sampled at four different positions in phase space, with preferably equal phase differences of 90°, and the group delay, which is the trigger for the tracking algorithm, is determined from five spectral channels.
The data reduction, which as final products delivers the star separation vectors on the sky, will be usually performed by an automated data reduction software, which when provided with

the necessary calibration sequences, corrects for instrumental, environmental and observational error sources. This procedure includes corrections to detector recordings due to unequal pixel sensitivities and relative phase errors, the determination of the metrology zero point, the correction for dispersion effects and measurement corrections due to the sky background, the sources' spectrum, the earth orientation, and also differential astrometric corrections, such as proper motions, light time delay, parallax, relativistic light deflection and aberration.

In the course of this work, the instrument's performance has been investigated, based on the observations that have been gathered during the system's commissioning runs.
It has been shown that stable fringe tracking is achieved for sufficient bright sources and various conditions, although the performance seems to decrease at coherence times below 3 ms. Also, in agreement with the theory, the delay variations power spectra resemble the Kolmogorov model throughout the moderate range between 1 and 50 Hz and the damping of high amplitude but low-frequency fringe motions (below $\sim$1 Hz) due to atmospheric disturbances on the secondary sources, which is the imminent effect of the phase referencing method, is apparent within the recorded data.
However, although the recorded tracking sequences in good conditions imply stable fringe tracking with mean phase variations below 1/4 of the wavelength of observation, a more detailed study reveals phase jumps, which are not noticed by the tracking controllers and which can be interpreted as undetected glitches, which after the fringe is not tracked at its center any more. The reason for those occurrences is supposed to be a nonlinear behavior of the group delay estimation, which might be resulting from differential photometric effects between the two source input channels and from non-nominal phase offsets between the four ABCD signals.
Their separation in phase is achieved by a polarization dependent encoding set-up, and naturally any polarization effects along the beam train affect its functionality.
As a matter of fact, overall general phase offset deviations up to $\pm 20^\circ$ and dispersion effects over the different spectral channels have been measured during the past observations and need to be studied in more detail so that in future for science operation they could be calibrated or even properly modelled and corrected for.

Apart from these technical issues, the analysis of the commissioning data, with regard to the astrometric performance, has uncovered grave uncertainties within the determination of the star separations.
Main delay measurements of observations of nine different stars during one single night have been used to solve for the instrument's wide-angle baseline, which was then applied to several single and multi epoch observations in order to perform the corresponding astrometric multiple linear regression fits, which deliver the amount and the orientation of the separations of the two observed stars on the sky.
This investigation has shown that the corresponding fits' 1-σ uncertainties, and likewise bootstrap analyses regarding the determination of the separations, imply achievable measurement precisions down to $\sim 60\,\mu$as, and even lower for the separation along the projected baseline orientation.
Unfortunately, this seems to apply only to measurements on short time scales, whereas observations of the same target pairs at different points in time show a large and only partially

comprehensible scatter of up to even tens of milliarcseconds, which cannot be explained by the observed systems' temporal evolution but obviously originates from systematics during the measurements and which limits the instrument's ultimate accuracy.
Although the situation seems to have improved after a recent system intervention, in the course of which some alignment issues have been removed, the systematic errors have not disappeared entirely. The until now persisting residuals are believed on one hand to result from unstable offsets of the metrology endpoints from the telescopes' pivot points, which affect the length and stability of the effective interferometric baseline, and on the other to be a consequence of the fact that the metrology endpoints are situated at the level of the star separators' M9 mirrors and that the sections of the beam paths between those and the instrument's entrance pupils at the M2 mirrors remain unmonitored.
For the sake of reaching a final and reliable astrometric accuracy between 10 and 50 μas, the remaining errors need to be understood and modelled or even eliminated, as would be the case if the metrology paths could be expanded, which is under consideration by ESO and the ESPRI consortium.

The decisions regarding further instrument improvements are expected before the beginning of 2014 and will be ultimately decisive for the future of the planned exoplanet survey ESPRI.
Since the program's acceptance, it has been prepared so that science operations can start whenever the instrument would be successfully commissioned and open to the scientific community.
Preparatory observations of about 1,000 target star candidates of different types, as stars with known companions that have been found by the RV method, young stars and generally nearby stars, which might be suitable for a blind search, have been performed during the recent years to search for adequate reference targets.
From the analysis of those observations, where emphasis has been laid on the stars' brightness and separation to the possible references, a database of well qualified targets has been compiled in order to facilitate the ultimate target selection and to be flexible regarding the limitations that will be imposed by the instrument's performance.
Besides the astrometric accuracy, the most critical factor in this context is the brightness limit for both, the primary and secondary stars, at which stable fringe tracking can be performed.
It is difficult to foresee, when PRIMA will be ready, so that the ESPRI program might commence, but it is clear already that in consequence of the delays of the instrument's final commissioning, due to all encountered unexpected difficulties, it will have to compete against concurring surveys as Gaia, ESA's space mission, which will be launched probably before the end of this year and which is also designed to perform astrometry at the microarcsecond level.
As a response to that, and in order to maintain a justifiable status within the community, the ESPRI team is currently expanding its set of possible science candidates to stars that cannot be observed by Gaia so that complementarity of the two projects is ensured.
With its final target collection at hand, ESPRI will hopefully be able to begin with early science observations next year, which simultaneously could provide an appropriate opportunity for the instrument's science verification.

Appendix A

The Three-Body Problem

Following the discussion of the topic as it can be found in [dePat01], the principles of the approach to the three-body problem are summarized here.
While the motion of two gravitationally interacting bodies is completely integrable, and analytical solutions to the problem are easily derived, the treatment of three massive objects becomes complicated enough so that in practice numerical integrations to the system are usually inevitable.
However, certain approximations can be adopted in order to simplify the general problem and to derive some principle results. If the mass of one of the bodies is assumed negligible in comparison to the other two, then this particular body can be considered as a test particle that does not affect the others, and the problem is simplified to the *restricted three-body problem*. In the case when additionally the relative orbit of the two massive objects is approximated by a circle, the system is simplified further to the so called *circular restricted three-body problem*, which will be considered here in the following.

Lagrangian Points and Co-Orbital Motion

Treating the circular restricted three-body problem in a non-inertial, synodic frame with its origin at the center of mass of the two massive bodies m_1, m_2, located at the fixed positions x_1, x_2, and which is rotating at the same rate as the two bodies' orbital frequency about its z-axis, perpendicular to the orbital plane, the two-body system becomes stationary, and the motion of the third (test) body can be described by the *Jacobi's constant of motion*:

$$C_J = \left(x^2 + y^2\right) + \left(\frac{2m_1}{|\boldsymbol{r} - \boldsymbol{r}_1|} + \frac{2m_2}{|\boldsymbol{r} - \boldsymbol{r}_2|}\right) - v^2 \quad , \tag{A.1}$$

where v denotes the test particle's velocity with reference to the rotating frame, and $|\boldsymbol{r} - \boldsymbol{r}_i|$ give its distance to the corresponding two massive bodies.
With units chosen so that the gravitational constant, the sum of the masses $m_1 + m_2$, the distance $|\boldsymbol{r}_2 - \boldsymbol{r}_1|$ and eventually the rotating frame's angular frequency equal one, the Jacobi constant equals twice the magnitude of the test body's energy (per unit mass), to which its kinetic energy and the effective potential contribute.
The particle's velocity for fixed values of C_J is now determined by its position within the orbital plane. Consequently, as $v^2 \geq 0$, *zero-velocity curves*, which are determined for different values of the Jacobi's constant by setting the velocity to zero, pose spatial boundaries to the test bodies' motion.
Several of such zero-velocity curves are illustrated in Fig. A.1, where the illustrated enclosed

areas restrict the orbits of the bodies with the corresponding C_J.

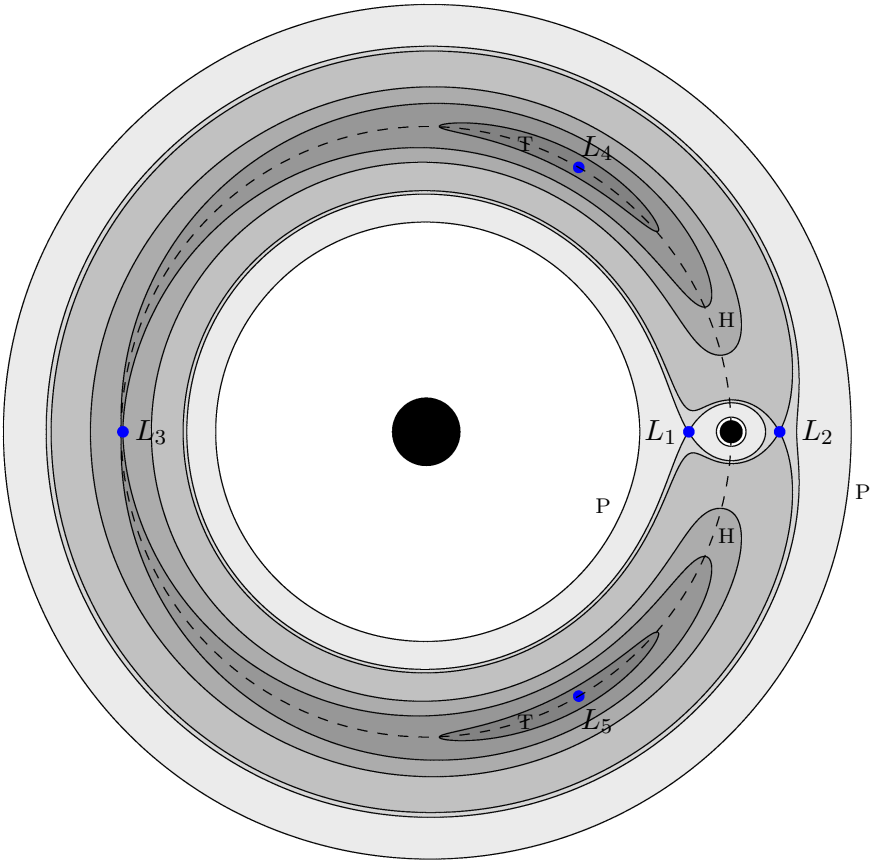


Figure A.1. – The graphic illustrates several zero-velocity curves for varying values of the Jacobi constant C_J (see Eq. (A.1)) within the frame co-rotating with the two bodies m_1 and m_2, which are depicted by the two filled black circles. The mass ratio between the primary body and the secondary body is chosen to be $m_1/m_2 = 100$. The locations of the five Lagrangian points are indicated by blue dots and the shadiness of the enclosed regions increases with decreasing C_J. The zero-velocity curves that are boundaries to the test particles' motion with the corresponding C_J and that define the shape of the passing, horseshoe and tadpole orbits are denoted by the letters P, H and T respectively.

However, several types of orbits are associated with the boundaries, as their shape is imposed onto the corresponding particles' paths of motion with reference to the co-rotational frame. Approximately circular *passing* orbits are only allowed above a critical value of C_J.
At lower values particles can move along the zero-velocity curve linked to the so called *horseshoe* orbits. Those occur when particles, with the adequate C_J values move at the near side to the system's center of mass, where its distance to the primary body is lower than that between the primary and secondary. The angular frequency of its orbit then exceeds that of the two-mass

system, and it catches up with the secondary body. At some point, when approaching the secondary, it is accelerated into a higher orbit, where its angular frequency drops below that of the reference two-body system. Consequently, it falls behind again and appears to move backwards until it is caught up by the secondary, at which point it is pulled into a lower orbit, and the cycle begins again.
Bodies with even lower values of C_J do not perform the whole horseshoe cycle, and their motion along the corresponding zero-velocity curves is referred to as *tadpole* or *Trojan* orbits.

At five points within the system, the net force onto the test particles becomes zero. Three of these *Lagrangian* points (L_1, L_2 and L_3) are saddle points of the effective potential and therefore unstable, while the triangular points L_4 and L_5 mark stable extrema.

The Hill Sphere

The region around the smaller secondary object within a two-body system, where the force onto a test particle is dominated by the secondary's gravitation, is usually called the *Hill sphere*, and its radius is approximated by ([dePat01])

$$R_H = \left(\frac{m_2}{3(m_1 + m_2)}\right)^{1/3} a \quad , \tag{A.2}$$

where a denotes the radius of the secondary's circular orbit about the system's primary mass m_1. Consequently, stable (*planetocentric*) orbits around the secondary body are confined to the region within the Hill sphere, while stable (*heliocentric*) orbits around the primary mass are only possible if the corresponding bodies' trajectories do not intersect with each other's Hill spheres.
Naturally, when entering the Hill spheres of more massive bodies, smaller objects can become gravitationally bound to those and stay in temporary orbits around them.

Appendix B

Transport of Energy in Electromagnetic Fields

Due to mathematical convenience electromagnetic fields are often described by complex functions. Although only the real parts are of physical significance, this handy approach is justified as long as mathematical operations on them do not mix the real and imaginary parts, which could alter the result, when the imaginary parts are ignored in the end. This needs to be considered particularly, when handling with cross or dot products of complex vectors.

The Poynting vector $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$, which represents the energy flow density of electromagnetic waves, is one example, where the operation must be treated with care, when the fields are given as complex quantities.
Since optical detectors usually are not able to resolve the temporal oscillation of the energy flow, it is interesting to consider the time average of a time-dependent signal $< f(t) >$, which can be defined by:

$$< f(t) > = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\,\mathrm{d}t \quad . \tag{B.1}$$

Now let us consider two complex vector fields with a periodic time-dependency:

$$\begin{aligned} \boldsymbol{a}(\boldsymbol{r},\boldsymbol{t}) &= \hat{\boldsymbol{a}}_0(\boldsymbol{r})\, e^{-i\omega t} \quad , \\ \boldsymbol{b}(\boldsymbol{r},\boldsymbol{t}) &= \hat{\boldsymbol{b}}_0(\boldsymbol{r})\, e^{-i\omega t} \quad , \end{aligned} \tag{B.2}$$

and their complex conjugates $\boldsymbol{a}^*$, $\boldsymbol{b}^*$.
Since one is interested in the entity $< \boldsymbol{a} \times \boldsymbol{b} >$, but only the real parts are considered of significance, the term

$$< (\mathrm{Re}\,\boldsymbol{a}) \times (\mathrm{Re}\,\boldsymbol{b}) > = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} (\mathrm{Re}\,\boldsymbol{a}) \times (\mathrm{Re}\,\boldsymbol{b})\,\mathrm{d}t \quad , \tag{B.3}$$

needs to be evaluated.
Using the relation

$$(\mathrm{Re}\,\boldsymbol{a}) \times (\mathrm{Re}\,\boldsymbol{b}) = \frac{1}{2}\,(\boldsymbol{a}+\boldsymbol{a}^*) \times \frac{1}{2}\,(\boldsymbol{b}+\boldsymbol{b}^*) = \frac{1}{4}\left[(\boldsymbol{a}\times\boldsymbol{b}) + (\boldsymbol{a}\times\boldsymbol{b}^*) + (\boldsymbol{a}^*\times\boldsymbol{b}) + (\boldsymbol{a}^*\times\boldsymbol{b}^*)\right], \tag{B.4}$$

one sees that four summands contribute to the result:

$$< (\mathrm{Re}\,\boldsymbol{a}) \times (\mathrm{Re}\,\boldsymbol{b}) > = \frac{1}{4} < \boldsymbol{a}\times\boldsymbol{b} > + < \boldsymbol{a}\times\boldsymbol{b}^* > + < \boldsymbol{a}^*\times\boldsymbol{b} > + < \boldsymbol{a}^*\times\boldsymbol{b}^* > . \tag{B.5}$$

Evaluating the integrals of the time average, one sees that

$$\begin{aligned} < \boldsymbol{a} \times \boldsymbol{b} > &= \lim_{T\to\infty} \frac{1}{2T}\, \hat{\boldsymbol{a}}_0 \times \hat{\boldsymbol{b}}_0 \int_{-T}^{T} e^{-2i\omega t}\, \mathrm{d}t \\ &= \frac{\hat{\boldsymbol{a}}_0 \times \hat{\boldsymbol{b}}_0}{2T}\, \frac{e^{-2i\omega t}}{-2i\omega}\Bigg|_{-T}^{T} \\ &= \frac{\hat{\boldsymbol{a}}_0 \times \hat{\boldsymbol{b}}_0}{2\omega T} \left(\frac{e^{2i\omega T} - e^{-2i\omega T}}{2i} \right) \\ &= \frac{\hat{\boldsymbol{a}}_0 \times \hat{\boldsymbol{b}}_0}{2\omega T} \sin(2\omega T) = 0 \quad , \end{aligned} \tag{B.6}$$

for $\lim T \to \infty$.
In analogy to this, it becomes clear that

$$< \boldsymbol{a}^* \times \boldsymbol{b}^* > = 0 \quad , \tag{B.7}$$

whereas non-vanishing contributions are left over from:

$$< \boldsymbol{a} \times \boldsymbol{b}^* > = \hat{\boldsymbol{a}}_0 \times \hat{\boldsymbol{b}}_0^* \quad , \tag{B.8}$$

$$< \boldsymbol{a}^* \times \boldsymbol{b} > = \hat{\boldsymbol{a}}_0^* \times \hat{\boldsymbol{b}}_0 \quad . \tag{B.9}$$

Given these results,

$$\begin{aligned} < (\mathrm{Re}\, \boldsymbol{a}) \times (\mathrm{Re}\, \boldsymbol{b}) > &= \frac{1}{4} \left[\left(\hat{\boldsymbol{a}}_0 \times \hat{\boldsymbol{b}}_0^* \right) + \left(\hat{\boldsymbol{a}}_0^* \times \hat{\boldsymbol{b}}_0 \right) \right] \\ &= \frac{1}{2}\, \mathrm{Re} \left(\hat{\boldsymbol{a}}_0 \times \hat{\boldsymbol{b}}_0^* \right) = \frac{1}{2}\, \mathrm{Re} \left(\hat{\boldsymbol{a}}_0^* \times \hat{\boldsymbol{b}}_0 \right) \quad , \end{aligned} \tag{B.10}$$

can be concluded.

Now, if the electric and magnetic fields are given by complex vector fields as in Eq. (B.2), then this result can be directly applied on the time average of the Poynting vector, yielding:

$$< \boldsymbol{S} > = \frac{1}{2}\, \mathrm{Re} \left(\hat{\boldsymbol{E}}_0 \times \hat{\boldsymbol{H}}_0^* \right) \quad . \tag{B.11}$$

In the particular case of plane waves, where $\boldsymbol{E} \perp \boldsymbol{H}$, it follow that:

$$\begin{aligned} < \boldsymbol{S} > &= \frac{c\epsilon_0}{2}\, \mathrm{Re} \left(\hat{E}_0 \hat{E}_0^* \right) \\ &= \frac{c\epsilon_0}{2} \hat{E}_0^2 \quad . \end{aligned} \tag{B.12}$$

This or similar derivations can be found in various textbooks on electrodynamics, such as [Nol04].

Appendix C

PRIMA Commissioning Files

In the following, the names of the PRIMA commissioning files that have been used for the analyses throughout the thesis are listed for reasons of reference.

In Sec. 4.1.3.2:

PACMA.2011-11-19T19:08:18.928.fits

In Sec. 4.3.1, Fig. 4.13:

PACMA.2010-11-06T08:11:11.172.fits
PACMA.2011-08-29T07:17:46.611.fits

Fig. 4.12:

PACMA.2010-11-08T04:21:06.365.fits
PACMA.2010-11-06T08:11:11.172.fits

In Sec. 4.3.1.1, Fig. 4.14:

PACMA.2011-11-21T04:46:50.719.fits

In Sec. 4.3.1.2, Fig. 4.17:

PACMA.2011-11-26T02:50:20.981.fits

In Sec. 4.3.2:

PACMA.2011-11-21T23:57:33.219.fits
PACMA.2011-11-22T04:03:21.985.fits
PACMA.2011-11-22T04:15:49.645.fits
PACMA.2011-11-22T04:53:08.325.fits
PACMA.2011-11-22T05:04:54.935.fits
PACMA.2011-11-22T05:22:07.585.fits
PACMA.2011-11-22T05:35:25.235.fits
PACMA.2011-11-22T05:54:46.625.fits
PACMA.2011-11-22T06:04:56.545.fits
PACMA.2011-11-22T06:16:38.336.fits
PACMA.2011-11-22T06:37:42.935.fits
PACMA.2011-11-22T06:49:34.145.fits
PACMA.2011-11-22T07:56:47.164.fits
PACMA.2011-11-22T08:11:53.994.fits
PACMA.2011-11-22T08:29:37.784.fits
PACMA.2011-11-22T08:41:30.394.fits
PACMA.2011-11-22T09:00:18.824.fits
PACMA.2011-11-22T09:12:22.614.fits

Single epoch astrometric fits in Sec. 4.3.3 and App. C:

HD100286_C14_ep2:

PACMA.2011-01-31T05:52:28.093.fits
PACMA.2011-01-31T06:09:29.912.fits
PACMA.2011-01-31T06:59:53.249.fits
PACMA.2011-01-31T07:22:36.278.fits

HD10268_C16_ep1:

PACMA.2011-08-29T06:23:14.554.fits
PACMA.2011-08-29T06:49:35.271.fits
PACMA.2011-08-29T07:04:56.110.fits
PACMA.2011-08-29T07:17:46.611.fits
PACMA.2011-08-29T07:33:08.649.fits
PACMA.2011-08-29T08:05:46.077.fits
PACMA.2011-08-29T08:20:06.037.fits
PACMA.2011-08-29T08:38:10.595.fits

HD10360_C16_ep1:

PACMA.2011-08-26T08:38:24.206.fits
PACMA.2011-08-26T09:00:03.636.fits
PACMA.2011-08-26T09:13:05.206.fits
PACMA.2011-08-26T09:26:12.035.fits
PACMA.2011-08-26T09:41:13.865.fits
PACMA.2011-08-26T09:53:40.495.fits

HD10360_C17_ep1:

PACMA.2011-11-20T02:08:58.057.fits
PACMA.2011-11-20T02:44:16.887.fits

HD10360_C17_ep2:

PACMA.2011-11-21T04:22:34.939.fits
PACMA.2011-11-21T04:51:51.429.fits
PACMA.2011-11-21T05:16:39.280.fits
PACMA.2011-11-21T05:36:18.050.fits
PACMA.2011-11-21T05:41:58.771.fits
PACMA.2011-11-21T06:10:48.071.fits
PACMA.2011-11-21T06:29:32.982.fits
PACMA.2011-11-21T06:35:31.881.fits
PACMA.2011-11-21T06:51:26.441.fits

HD10360_C17_ep4:

PACMA.2011-11-24T01:39:11.418.fits
PACMA.2011-11-24T02:03:54.017.fits
PACMA.2011-11-24T02:12:37.327.fits
PACMA.2011-11-24T02:19:10.477.fits
PACMA.2011-11-24T02:24:48.967.fits
PACMA.2011-11-24T02:42:53.058.fits
PACMA.2011-11-24T02:48:34.337.fits
PACMA.2011-11-24T02:54:05.839.fits
PACMA.2011-11-24T02:59:32.128.fits
PACMA.2011-11-24T03:05:03.629.fits
PACMA.2011-11-24T03:10:31.348.fits
PACMA.2011-11-24T03:28:26.979.fits
PACMA.2011-11-24T03:34:00.089.fits
PACMA.2011-11-24T03:44:56.310.fits
PACMA.2011-11-24T03:51:54.420.fits
PACMA.2011-11-24T03:57:22.520.fits
PACMA.2011-11-24T04:02:51.010.fits
PACMA.2011-11-24T04:08:24.511.fits
PACMA.2011-11-24T04:13:52.231.fits
PACMA.2011-11-24T04:19:50.731.fits
PACMA.2011-11-24T04:26:12.721.fits
PACMA.2011-11-24T04:52:35.741.fits
PACMA.2011-11-24T05:03:32.362.fits
PACMA.2011-11-24T05:09:00.472.fits
PACMA.2011-11-24T05:14:38.962.fits

HD10360_C17_ep5:

PACMA.2011-11-25T00:34:09.786.fits
PACMA.2011-11-25T01:20:59.677.fits
PACMA.2011-11-25T01:38:06.926.fits
PACMA.2011-11-25T01:43:38.237.fits
PACMA.2011-11-25T02:05:44.977.fits
PACMA.2011-11-25T02:11:27.697.fits
PACMA.2011-11-25T02:27:31.098.fits
PACMA.2011-11-25T02:33:02.408.fits
PACMA.2011-11-25T02:38:43.708.fits
PACMA.2011-11-25T02:52:27.528.fits
PACMA.2011-11-25T02:57:56.038.fits
PACMA.2011-11-25T03:03:34.529.fits
PACMA.2011-11-25T03:27:16.099.fits
PACMA.2011-11-25T03:32:44.609.fits
PACMA.2011-11-25T03:38:12.729.fits

HD10360_C17_ep6:

PACMA.2011-11-26T02:00:16.849.fits
PACMA.2011-11-26T02:05:45.339.fits
PACMA.2011-11-26T02:11:26.449.fits
PACMA.2011-11-26T02:14:54.749.fits
PACMA.2011-11-26T02:21:28.250.fits
PACMA.2011-11-26T02:27:01.349.fits
PACMA.2011-11-26T02:32:37.480.fits
PACMA.2011-11-26T02:38:41.380.fits
PACMA.2011-11-26T02:50:20.981.fits
PACMA.2011-11-26T02:55:47.240.fits
PACMA.2011-11-26T03:01:15.741.fits
PACMA.2011-11-26T03:06:48.861.fits
PACMA.2011-11-26T03:12:57.151.fits
PACMA.2011-11-26T03:18:28.642.fits
PACMA.2011-11-26T03:24:09.511.fits
PACMA.2011-11-26T03:30:22.812.fits

HD10360_C17_ep7:

PACMA.2011-11-26T03:54:40.432.fits
PACMA.2011-11-26T04:00:06.721.fits
PACMA.2011-11-26T04:05:35.021.fits
PACMA.2011-11-26T04:11:48.502.fits
PACMA.2011-11-26T04:26:12.292.fits
PACMA.2011-11-26T04:31:40.402.fits
PACMA.2011-11-26T04:37:08.512.fits
PACMA.2011-11-26T04:42:46.623.fits
PACMA.2011-11-26T04:54:18.053.fits
PACMA.2011-11-26T04:59:53.973.fits
PACMA.2011-11-26T05:05:27.093.fits
PACMA.2011-11-26T05:12:05.403.fits
PACMA.2011-11-26T05:28:49.494.fits
PACMA.2011-11-26T05:34:17.613.fits
PACMA.2011-11-26T05:39:45.714.fits
PACMA.2011-11-26T05:45:18.834.fits
PACMA.2011-11-26T05:56:40.254.fits
PACMA.2011-11-26T06:02:14.165.fits
PACMA.2011-11-26T06:07:40.075.fits
PACMA.2011-11-26T06:13:23.555.fits
PACMA.2011-11-26T06:14:12.244.fits
PACMA.2011-11-26T06:25:52.895.fits
PACMA.2011-11-26T06:31:19.185.fits
PACMA.2011-11-26T06:36:50.706.fits
PACMA.2011-11-26T06:42:28.425.fits

HD108248J_C14_ep1:

PACMA.2011-01-27T08:52:50.486.fits
PACMA.2011-01-27T09:20:48.604.fits

HD108248J_C14_ep2:

PACMA.2011-01-29T08:49:00.462.fits
PACMA.2011-01-29T09:09:07.131.fits
PACMA.2011-01-29T09:26:42.500.fits

HD108248J_C14_ep3:

PACMA.2011-02-01T08:30:46.552.fits
PACMA.2011-02-01T08:52:09.720.fits
PACMA.2011-02-01T09:01:12.390.fits
PACMA.2011-02-01T09:17:57.059.fits
PACMA.2011-02-01T09:23:42.729.fits
PACMA.2011-02-01T09:36:22.029.fits

HD131977_C14_ep1:

PACMA.2011-02-02T08:22:06.969.fits
PACMA.2011-02-02T08:45:09.778.fits

HD156274_C15_ep3:

PACMA.2011-07-20T00:22:00.348.fits
PACMA.2011-07-20T00:27:46.208.fits
PACMA.2011-07-20T00:33:53.929.fits
PACMA.2011-07-20T00:44:06.668.fits
PACMA.2011-07-20T00:49:49.558.fits
PACMA.2011-07-20T00:58:12.668.fits
PACMA.2011-07-20T01:05:52.158.fits
PACMA.2011-07-20T01:14:07.278.fits
PACMA.2011-07-20T01:22:07.777.fits
PACMA.2011-07-20T01:30:21.108.fits
PACMA.2011-07-20T01:49:52.228.fits

PACMA.2011-07-20T01:55:50.718.fits
PACMA.2011-07-20T02:04:08.447.fits
PACMA.2011-07-20T02:09:51.558.fits
PACMA.2011-07-20T02:19:50.067.fits
PACMA.2011-07-20T02:28:09.757.fits
PACMA.2011-07-20T02:33:48.258.fits
PACMA.2011-07-20T02:48:00.468.fits

HD156274_C15_ep4:

PACMA.2011-07-20T03:12:39.027.fits
PACMA.2011-07-20T03:19:07.137.fits
PACMA.2011-07-20T03:38:49.786.fits
PACMA.2011-07-20T03:47:12.907.fits

HD18622_C15_ep2:

PACMA.2011-07-21T08:54:02.203.fits
PACMA.2011-07-21T09:08:04.803.fits
PACMA.2011-07-21T09:26:17.832.fits

HD202730_C15_ep1:

PACMA.2011-07-20T05:05:22.883.fits
PACMA.2011-07-20T05:11:13.984.fits
PACMA.2011-07-20T05:37:08.163.fits
PACMA.2011-07-20T05:54:12.723.fits
PACMA.2011-07-20T06:02:30.463.fits
PACMA.2011-07-20T06:35:55.773.fits
PACMA.2011-07-20T07:36:48.261.fits

HD202730_C16_ep1:

PACMA.2011-08-24T04:21:12.062.fits
PACMA.2011-08-24T04:35:13.883.fits
PACMA.2011-08-24T04:56:12.712.fits
PACMA.2011-08-24T05:09:34.362.fits
PACMA.2011-08-24T05:27:51.241.fits
PACMA.2011-08-24T05:39:41.391.fits
PACMA.2011-08-24T05:53:27.210.fits
PACMA.2011-08-24T06:08:19.891.fits
PACMA.2011-08-24T06:22:25.680.fits
PACMA.2011-08-24T06:28:20.780.fits

HD202730_C16_ep2:

PACMA.2011-08-26T02:33:59.188.fits
PACMA.2011-08-26T02:48:43.028.fits
PACMA.2011-08-26T03:05:07.627.fits
PACMA.2011-08-26T03:18:13.148.fits
PACMA.2011-08-26T03:33:14.768.fits
PACMA.2011-08-26T03:46:21.598.fits
PACMA.2011-08-26T03:59:08.207.fits
PACMA.2011-08-26T04:11:37.006.fits

HD202730_C16_ep3:

PACMA.2011-08-29T03:43:08.355.fits
PACMA.2011-08-29T04:02:10.713.fits
PACMA.2011-08-29T04:15:46.003.fits
PACMA.2011-08-29T04:28:45.242.fits
PACMA.2011-08-29T04:43:44.821.fits
PACMA.2011-08-29T05:03:24.909.fits
PACMA.2011-08-29T05:28:37.778.fits

HD66598_C14_ep1:

PACMA.2011-01-27T04:24:08.843.fits
PACMA.2011-01-27T04:51:46.620.fits

HD66598_C14_ep2:

PACMA.2011-01-27T05:59:48.956.fits
PACMA.2011-01-27T06:15:44.126.fits

HD66598_C14_ep3:

PACMA.2011-01-28T04:25:50.326.fits
PACMA.2011-01-28T04:57:15.634.fits
PACMA.2011-01-28T05:13:45.903.fits
PACMA.2011-01-28T05:32:31.172.fits
PACMA.2011-01-28T05:52:41.881.fits

HD66598_C14_ep4:

PACMA.2011-01-29T03:50:57.575.fits
PACMA.2011-01-29T04:08:33.384.fits
PACMA.2011-01-29T04:54:12.140.fits

HD66598_C14_ep6v2:

PACMA.2011-01-30T02:37:54.252.fits
PACMA.2011-01-30T02:58:00.371.fits
PACMA.2011-01-30T03:16:52.719.fits
PACMA.2011-01-30T03:38:00.089.fits
PACMA.2011-01-30T03:58:00.548.fits
PACMA.2011-01-30T04:17:47.718.fits
PACMA.2011-01-30T04:36:36.667.fits
PACMA.2011-01-30T04:56:39.156.fits
PACMA.2011-01-30T05:16:06.725.fits

HD66598_C14_ep9:

PACMA.2011-02-02T01:29:41.271.fits
PACMA.2011-02-02T01:59:19.800.fits
PACMA.2011-02-02T02:09:14.469.fits
PACMA.2011-02-02T02:17:44.138.fits

HD66598_C14_ep10:

PACMA.2011-02-02T02:58:03.897.fits
PACMA.2011-02-02T03:08:24.566.fits
PACMA.2011-02-02T03:23:53.435.fits
PACMA.2011-02-02T03:33:54.105.fits
PACMA.2011-02-02T03:44:14.775.fits

HD66598_C14_ep11:

PACMA.2011-02-02T04:37:15.531.fits
PACMA.2011-02-02T04:59:40.090.fits
PACMA.2011-02-02T05:27:10.618.fits
PACMA.2011-02-02T06:06:55.676.fits

HD66598_C14_ep12:

PACMA.2011-02-04T02:34:37.029.fits
PACMA.2011-02-04T02:54:55.848.fits
PACMA.2011-02-04T03:14:31.387.fits
PACMA.2011-02-04T03:34:19.926.fits

PACMA.2011-02-04T04:35:26.013.fits
PACMA.2011-02-04T05:38:57.190.fits
PACMA.2011-02-04T05:44:34.860.fits
PACMA.2011-02-04T05:50:08.530.fits
PACMA.2011-02-04T06:01:06.869.fits

HD66598_C17_ep1:

PACMA.2011-11-24T05:43:34.102.fits
PACMA.2011-11-24T05:49:05.612.fits
PACMA.2011-11-24T05:54:56.882.fits
PACMA.2011-11-24T06:16:32.632.fits
PACMA.2011-11-24T06:22:00.732.fits
PACMA.2011-11-24T06:27:41.853.fits
PACMA.2011-11-24T06:48:49.462.fits
PACMA.2011-11-24T06:54:20.763.fits
PACMA.2011-11-24T06:59:47.052.fits
PACMA.2011-11-24T07:05:15.162.fits
PACMA.2011-11-24T07:10:43.473.fits
PACMA.2011-11-24T07:16:49.593.fits
PACMA.2011-11-24T07:22:15.493.fits
PACMA.2011-11-24T07:27:43.604.fits
PACMA.2011-11-24T07:33:12.103.fits
PACMA.2011-11-24T07:38:40.214.fits
PACMA.2011-11-24T07:50:31.004.fits
PACMA.2011-11-24T08:10:12.494.fits
PACMA.2011-11-24T08:15:43.995.fits
PACMA.2011-11-24T08:21:10.284.fits
PACMA.2011-11-24T08:26:41.765.fits
PACMA.2011-11-24T08:37:50.295.fits
PACMA.2011-11-24T08:43:16.585.fits
PACMA.2011-11-24T08:48:44.895.fits
PACMA.2011-11-24T08:54:23.375.fits
PACMA.2011-11-24T08:59:51.876.fits
PACMA.2011-11-24T09:05:23.387.fits

HD66598_C17_ep2

PACMA.2011-11-25T05:04:28.618.fits
PACMA.2011-11-25T05:09:56.928.fits
PACMA.2011-11-25T05:15:25.048.fits
PACMA.2011-11-25T05:21:26.549.fits
PACMA.2011-11-25T05:26:52.469.fits
PACMA.2011-11-25T05:32:20.589.fits
PACMA.2011-11-25T05:38:11.710.fits
PACMA.2011-11-25T05:52:07.939.fits
PACMA.2011-11-25T05:57:39.450.fits
PACMA.2011-11-25T06:03:05.330.fits
PACMA.2011-11-25T06:08:33.449.fits
PACMA.2011-11-25T06:14:01.570.fits
PACMA.2011-11-25T06:19:29.290.fits
PACMA.2011-11-25T06:25:12.620.fits
PACMA.2011-11-25T06:49:44.800.fits
PACMA.2011-11-25T06:55:11.100.fits
PACMA.2011-11-25T07:00:39.210.fits
PACMA.2011-11-25T07:06:07.720.fits
PACMA.2011-11-25T07:11:38.441.fits
PACMA.2011-11-25T07:17:17.311.fits
PACMA.2011-11-25T07:34:19.160.fits
PACMA.2011-11-25T07:39:47.661.fits
PACMA.2011-11-25T07:45:15.781.fits
PACMA.2011-11-25T07:50:47.282.fits
PACMA.2011-11-25T07:56:13.531.fits
PACMA.2011-11-25T08:01:51.641.fits
PACMA.2011-11-25T08:21:00.532.fits
PACMA.2011-11-25T08:26:28.651.fits
PACMA.2011-11-25T08:31:56.962.fits
PACMA.2011-11-25T08:37:27.682.fits

HD66598_C17_ep3:

PACMA.2011-11-26T07:21:31.766.fits
PACMA.2011-11-26T07:39:27.846.fits
PACMA.2011-11-26T07:57:34.796.fits
PACMA.2011-11-26T08:14:57.127.fits
PACMA.2011-11-26T08:35:00.437.fits
PACMA.2011-11-26T08:58:36.917.fits
PACMA.2011-11-26T09:04:05.037.fits

HD10360_C17_s_ep1v2:

PACMA.2011-11-20T01:55:20.187.fits
PACMA.2011-11-20T01:58:02.877.fits
PACMA.2011-11-20T02:38:16.177.fits

HD10360_C17_s_ep2:

PACMA.2011-11-21T04:16:15.659.fits
PACMA.2011-11-21T04:46:50.719.fits
PACMA.2011-11-21T05:10:37.799.fits
PACMA.2011-11-21T05:31:26.370.fits
PACMA.2011-11-21T05:59:28.110.fits
PACMA.2011-11-21T06:23:58.111.fits

HD10360_C17_s_ep3:

PACMA.2011-11-24T01:57:43.297.fits
PACMA.2011-11-24T02:37:55.768.fits
PACMA.2011-11-24T03:16:39.068.fits
PACMA.2011-11-24T04:41:45.361.fits
PACMA.2011-11-24T01:33:01.927.fits
PACMA.2011-11-24T03:16:39.068.fits

HD202730_C16_s_ep1:

PACMA.2011-08-24T04:30:59.252.fits
PACMA.2011-08-24T05:05:09.721.fits
PACMA.2011-08-24T05:35:22.121.fits

HD66598_C17_s_ep1:

PACMA.2011-11-26T07:34:28.345.fits
PACMA.2011-11-26T07:52:44.076.fits
PACMA.2011-11-26T08:10:08.416.fits
PACMA.2011-11-26T08:53:21.197.fits

Epoch identifiers of the individual epochs included within the multi epoch astrometric fits in Sec. 4.3.3 and App. C:

HD10360_m1:

HD10360_C17_ep1
HD10360_C17_ep2
HD10360_C17_ep4
HD10360_C17_ep5
HD10360_C17_ep6
HD10360_C17_ep7

HD10360_m2:

HD10360_C16_ep1
HD10360_C17_ep1
HD10360_C17_ep2
HD10360_C17_ep4
HD10360_C17_ep5
HD10360_C17_ep6
HD10360_C17_ep7

HD108248J_m1:

HD108248J_C14_ep1
HD108248J_C14_ep2
HD108248J_C14_ep3

HD108248J_m2:

HD108248J_C14_ep1
HD108248J_C14_ep3

HD156274_m1:

HD156274_C15_ep3
HD156274_C15_ep4

HD202730_m1:

HD202730_C16_ep1
HD202730_C16_ep2
HD202730_C16_ep3

HD202730_m2:

HD202730_C15_ep1
HD202730_C16_ep1
HD202730_C16_ep2
HD202730_C16_ep3

HD66598_m1v2:

HD66598_C14_ep1
HD66598_C14_ep2
HD66598_C14_ep3
HD66598_C14_ep4
HD66598_C14_ep6v2
HD66598_C14_ep10
HD66598_C14_ep11
HD66598_C14_ep12

HD66598_m2:

HD66598_C17_ep1
HD66598_C17_ep2
HD66598_C17_ep3

HD66598_m3v2:

HD66598_C14_ep1
HD66598_C14_ep2
HD66598_C14_ep3
HD66598_C14_ep4
HD66598_C14_ep6v2
HD66598_C14_ep10
HD66598_C14_ep11
HD66598_C14_ep12
HD66598_C17_ep1
HD66598_C17_ep2
HD66598_C17_ep3

Appendix D

Plots to the Astrometric Fits

For the sake of commpleteness, the plots to all performed astrometric fits of single and multi epoch sequences are included here.

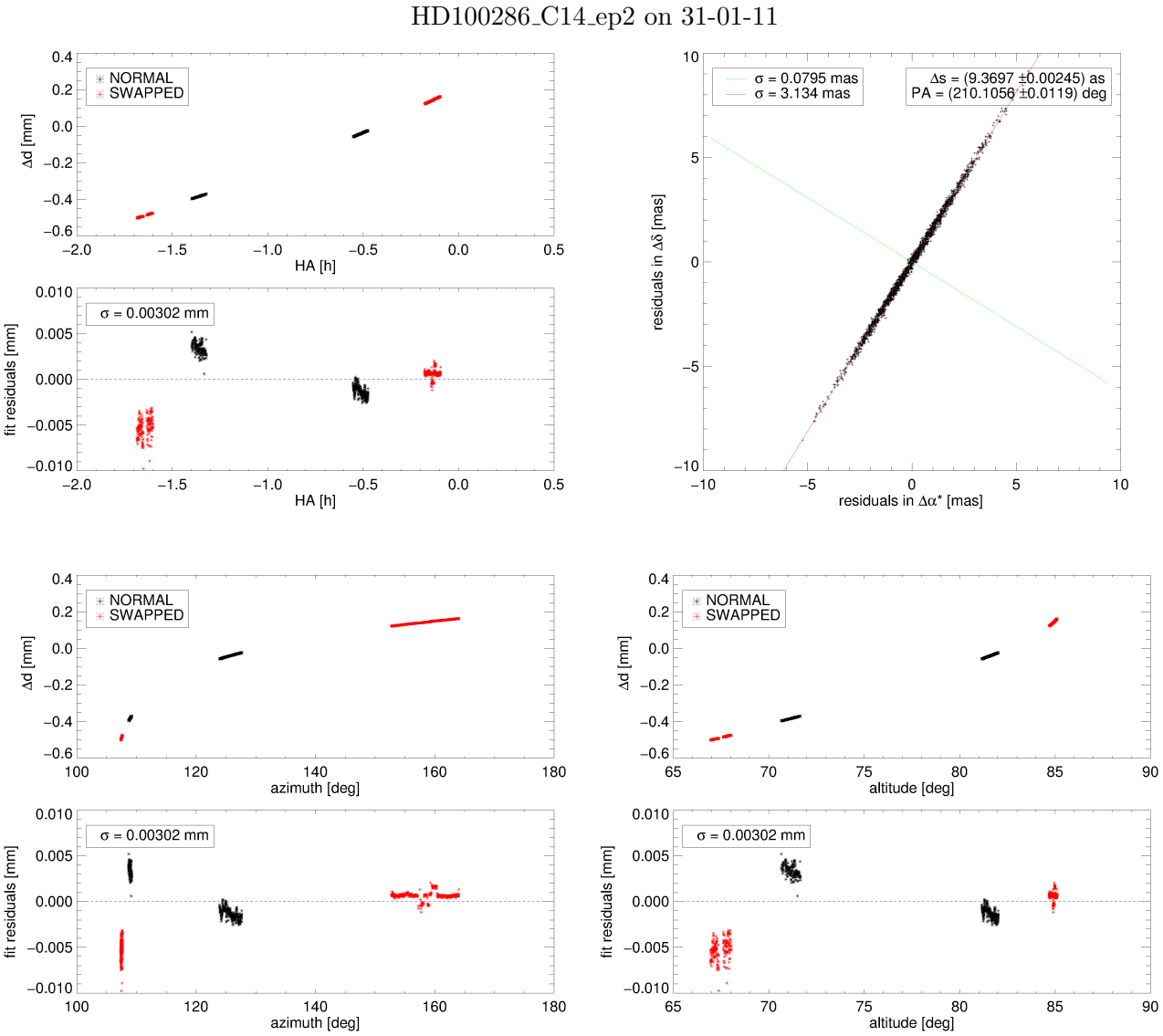


Figure D.1. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD100286_C14_ep2 on 31-01-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

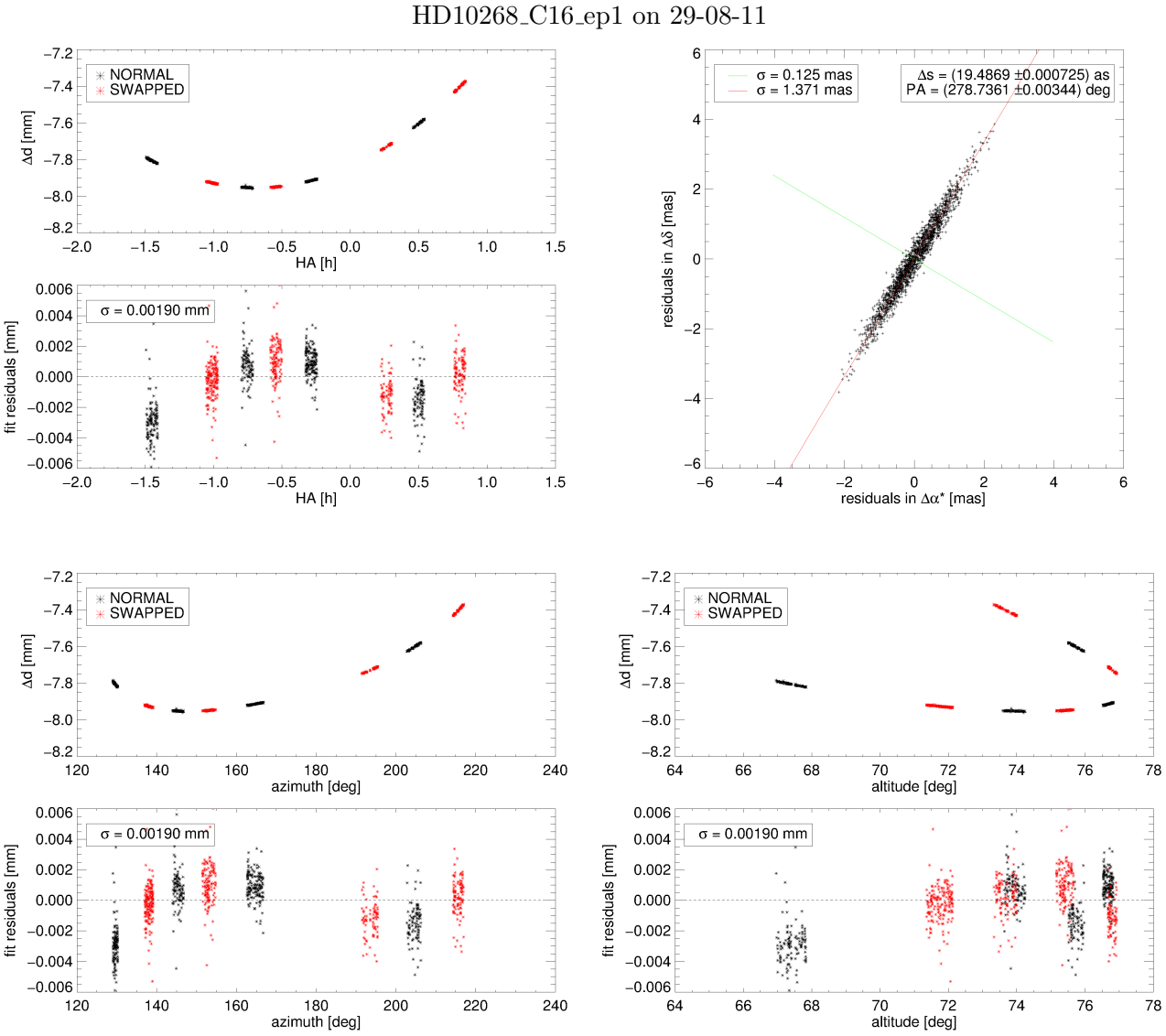


Figure D.2. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10268_C16_ep1 on 29-08-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

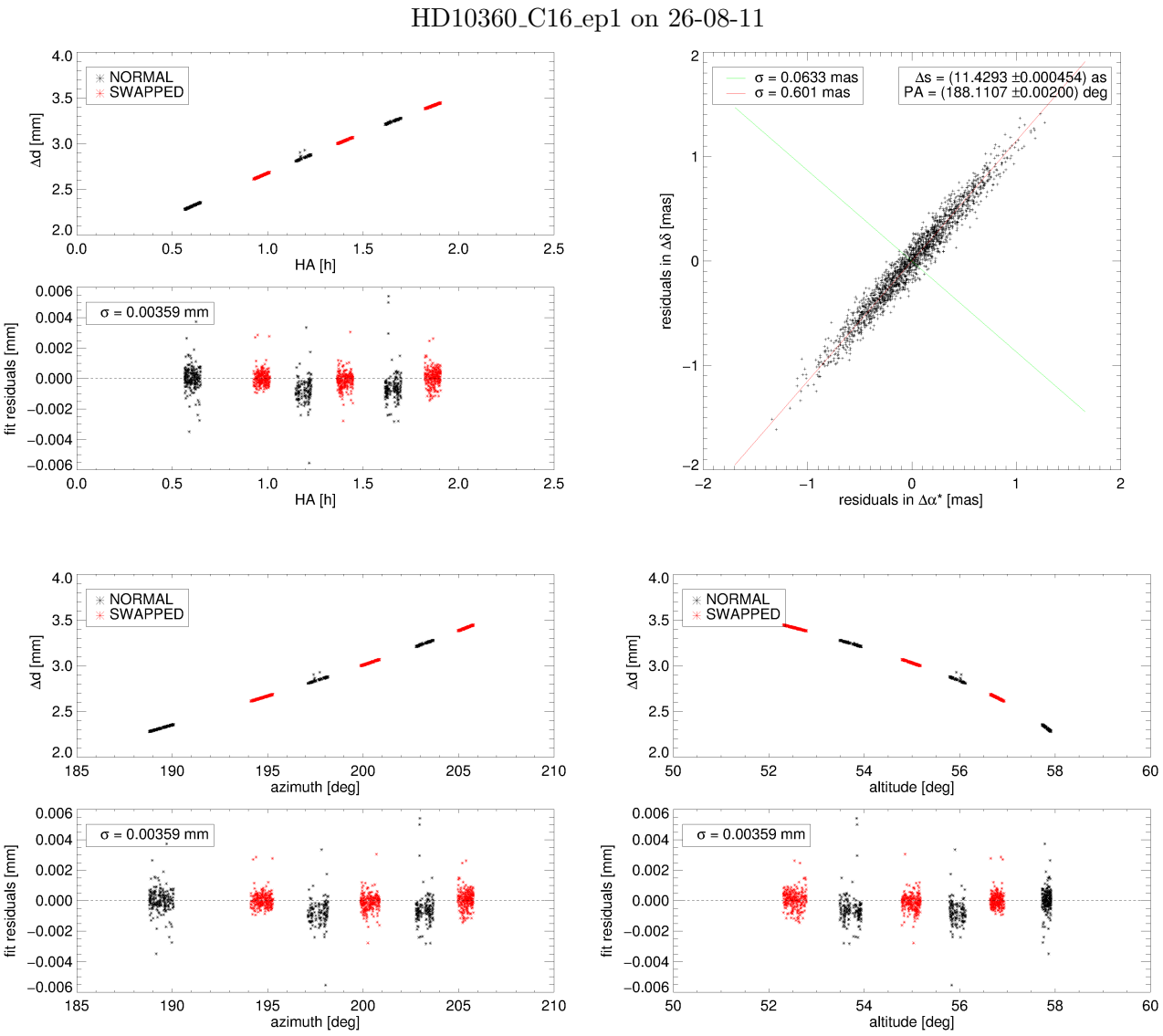


Figure D.3. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C16_ep1 on 26-08-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

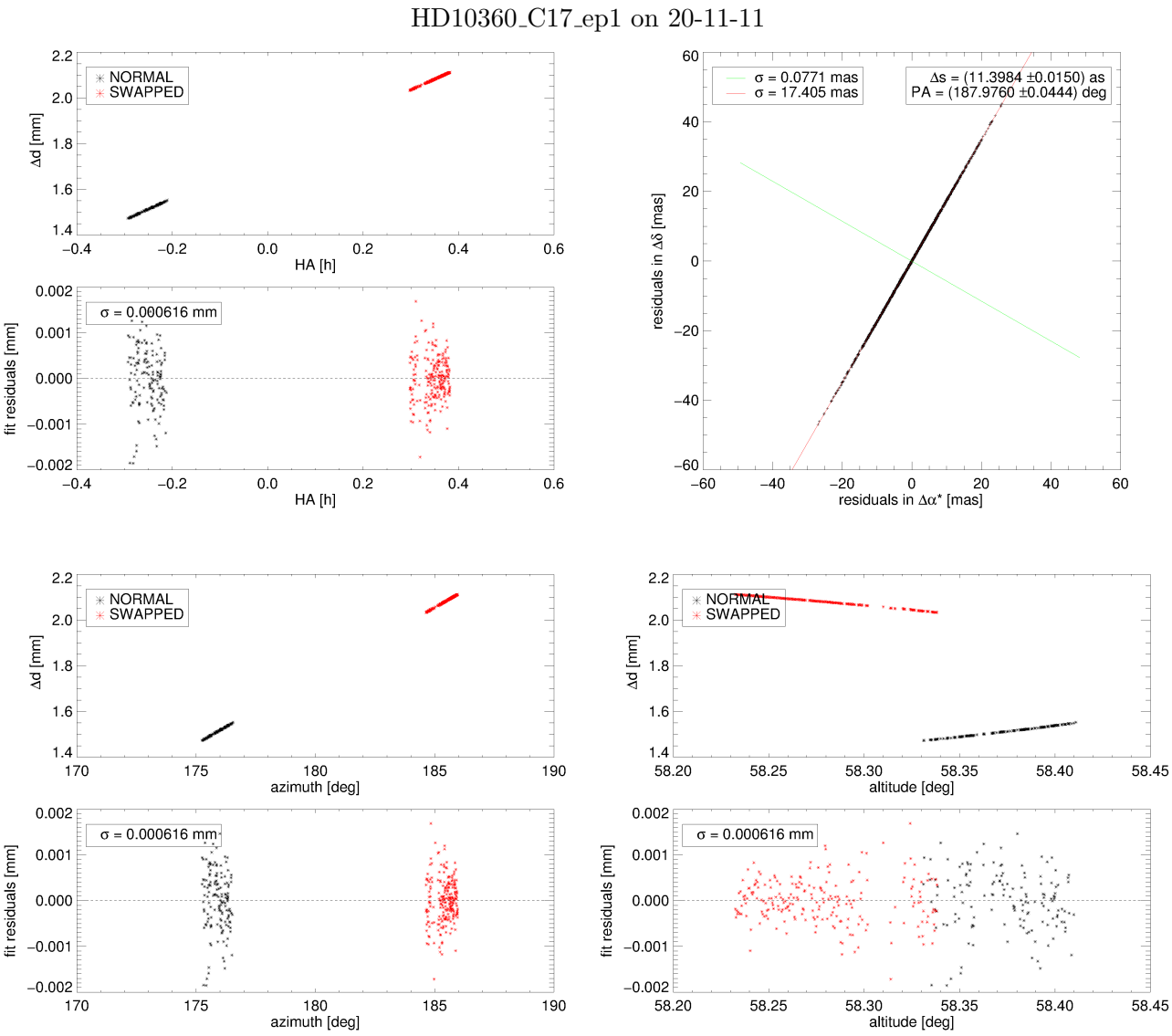


Figure D.4. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_ep1 on 20-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

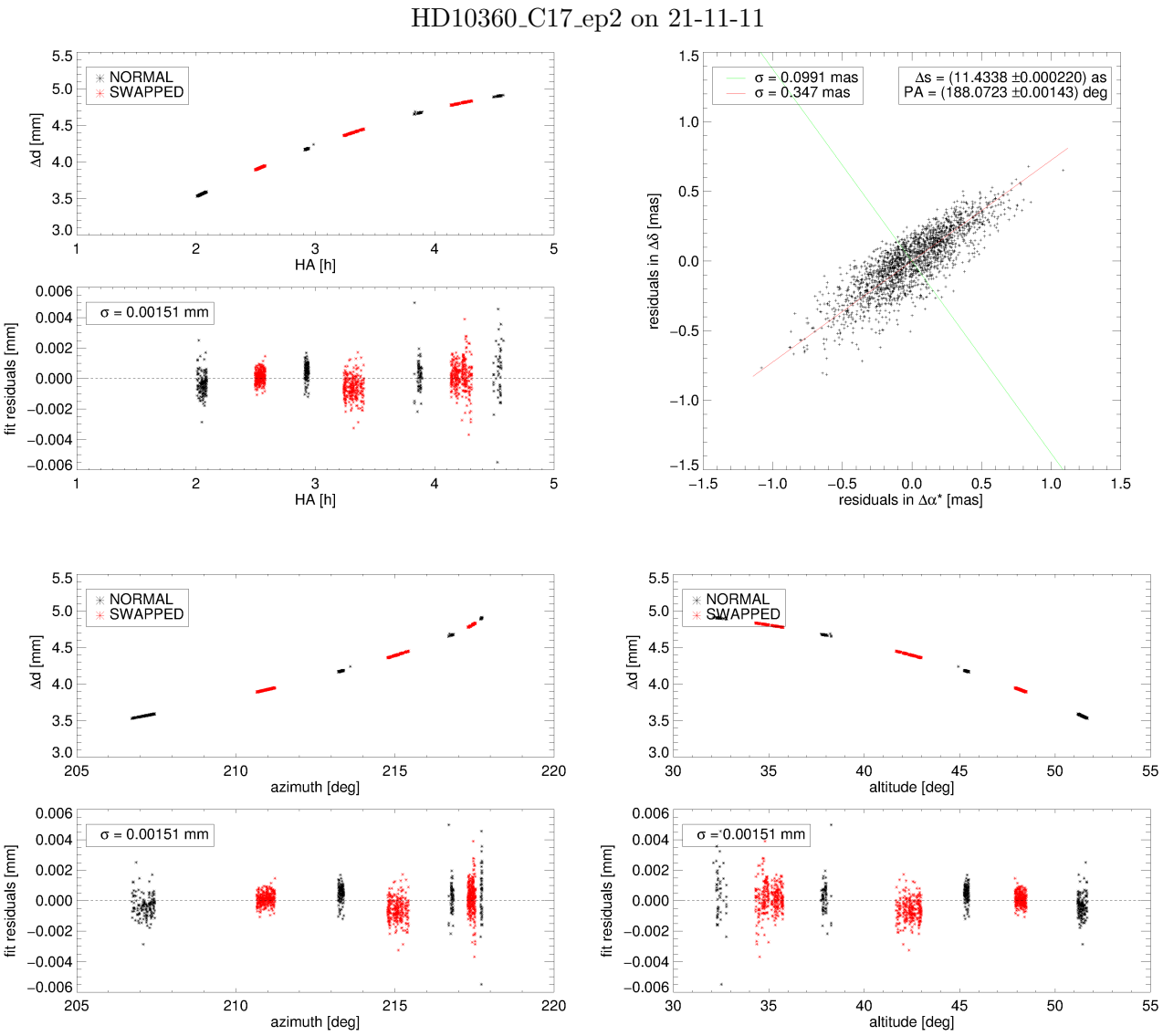


Figure D.5. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_ep2 on 21-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

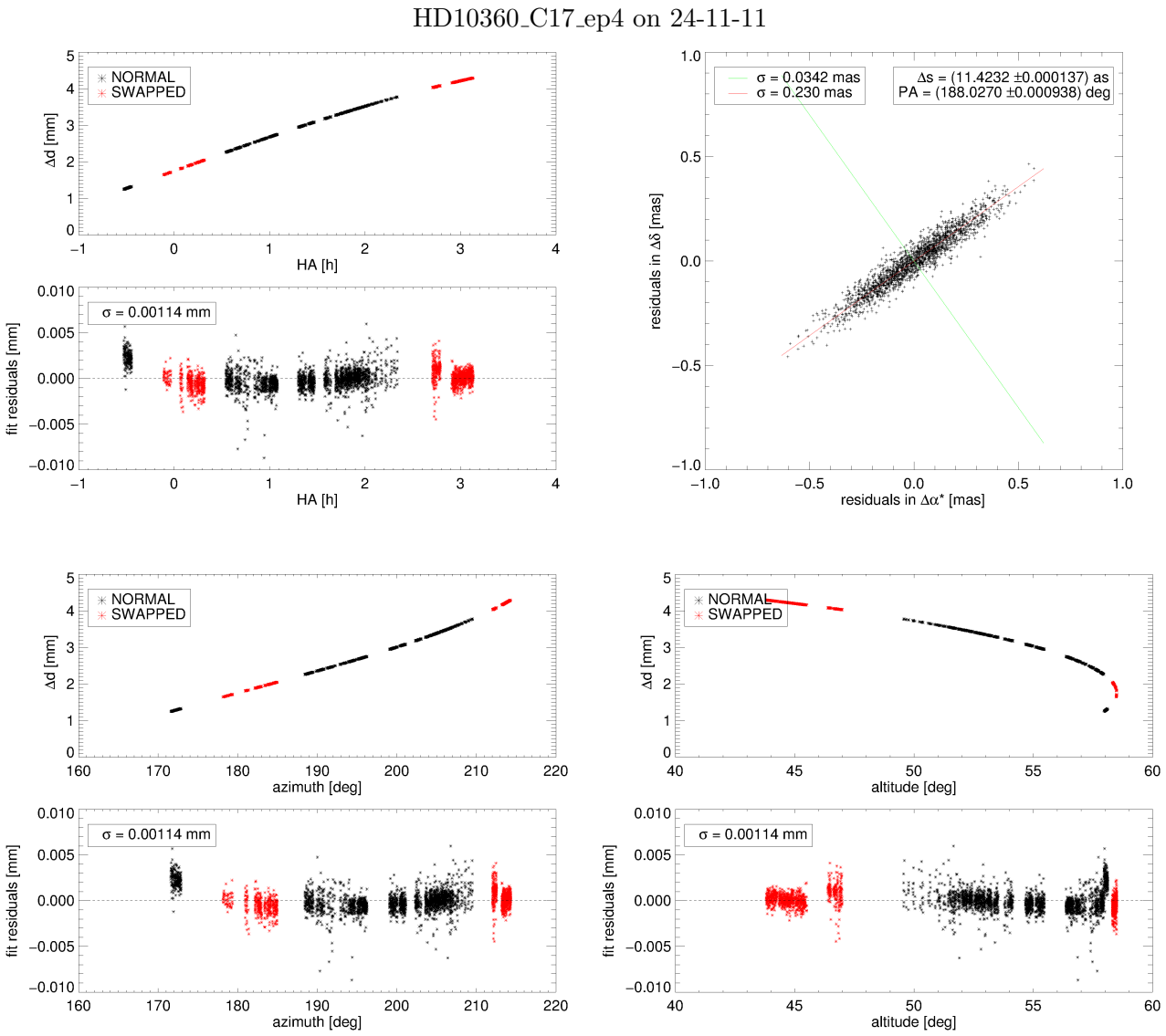


Figure D.6. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_ep4 on 24-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

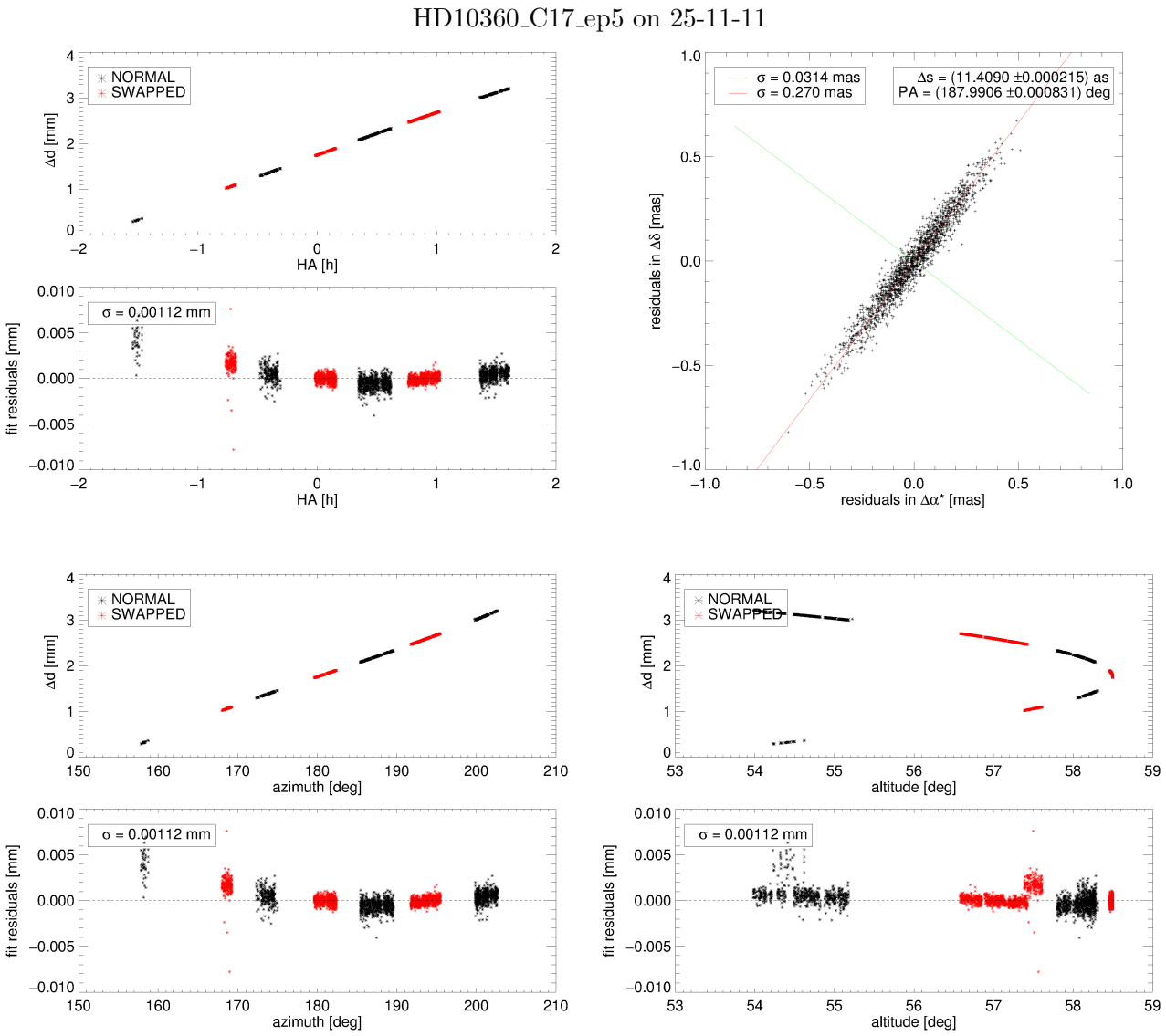


Figure D.7. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_ep5 on 25-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

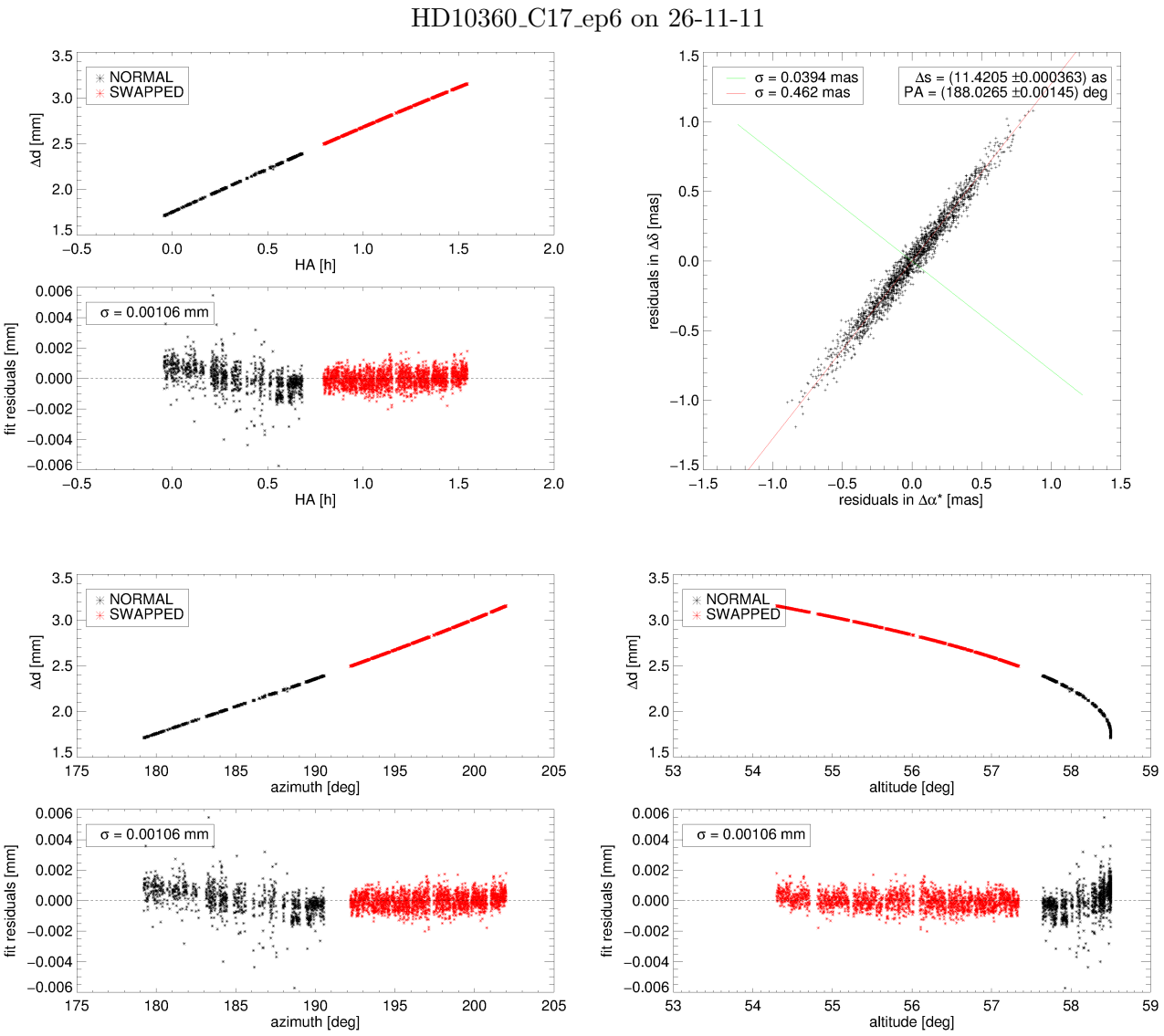


Figure D.8. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_ep6 on 26-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

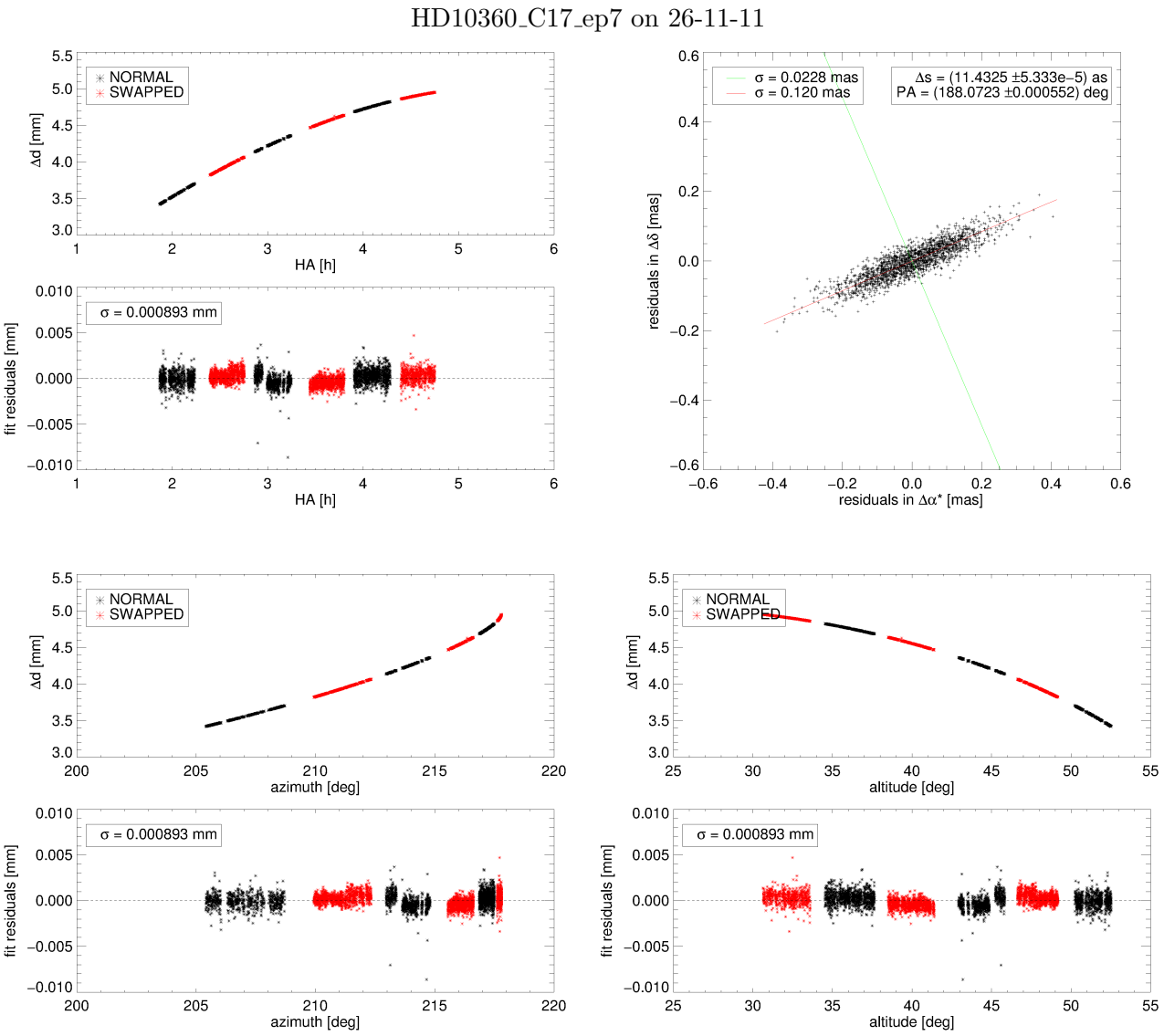


Figure D.9. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_ep7 on 26-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

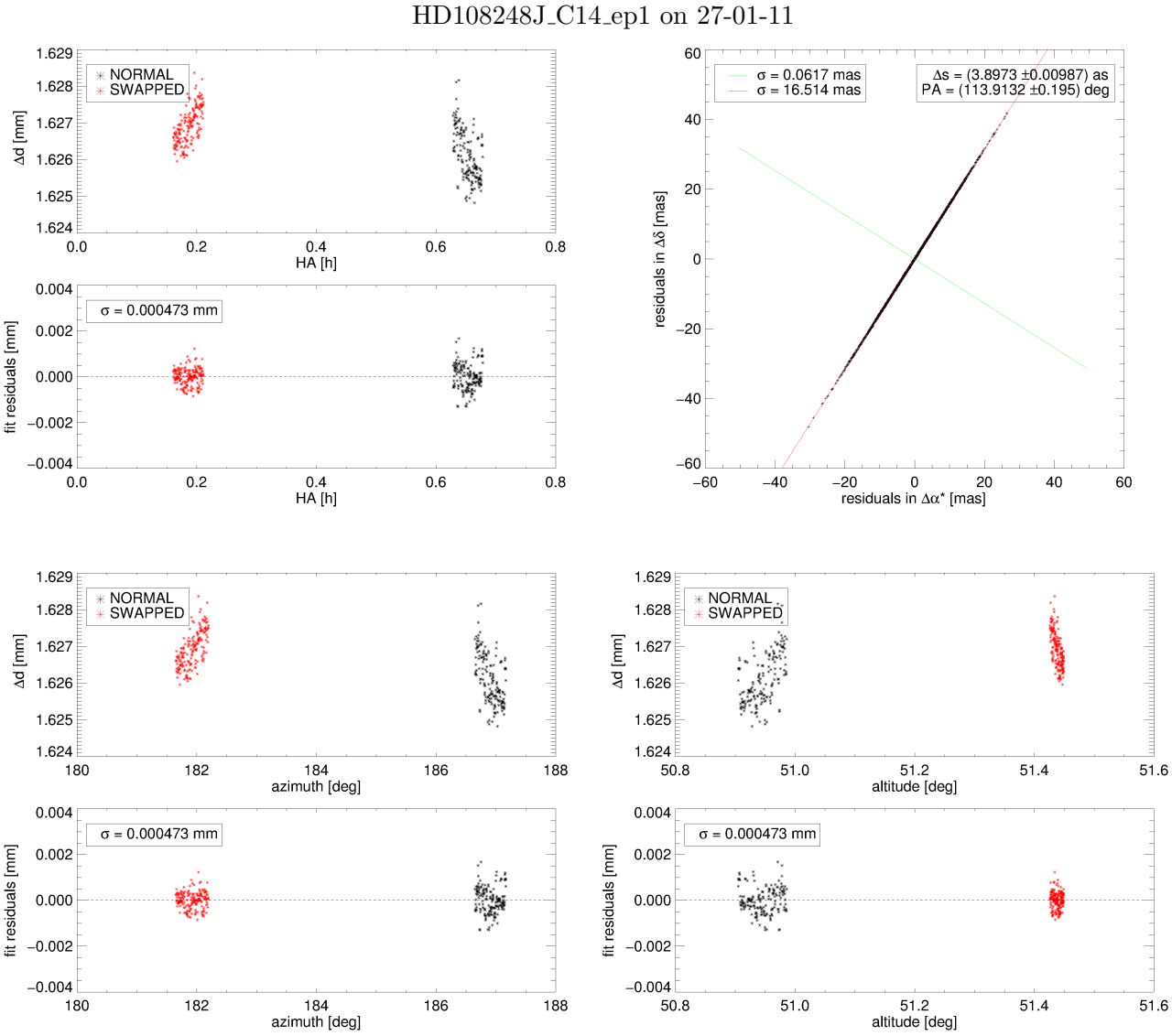


Figure D.10. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD108248J_C14_ep1 on 27-01-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

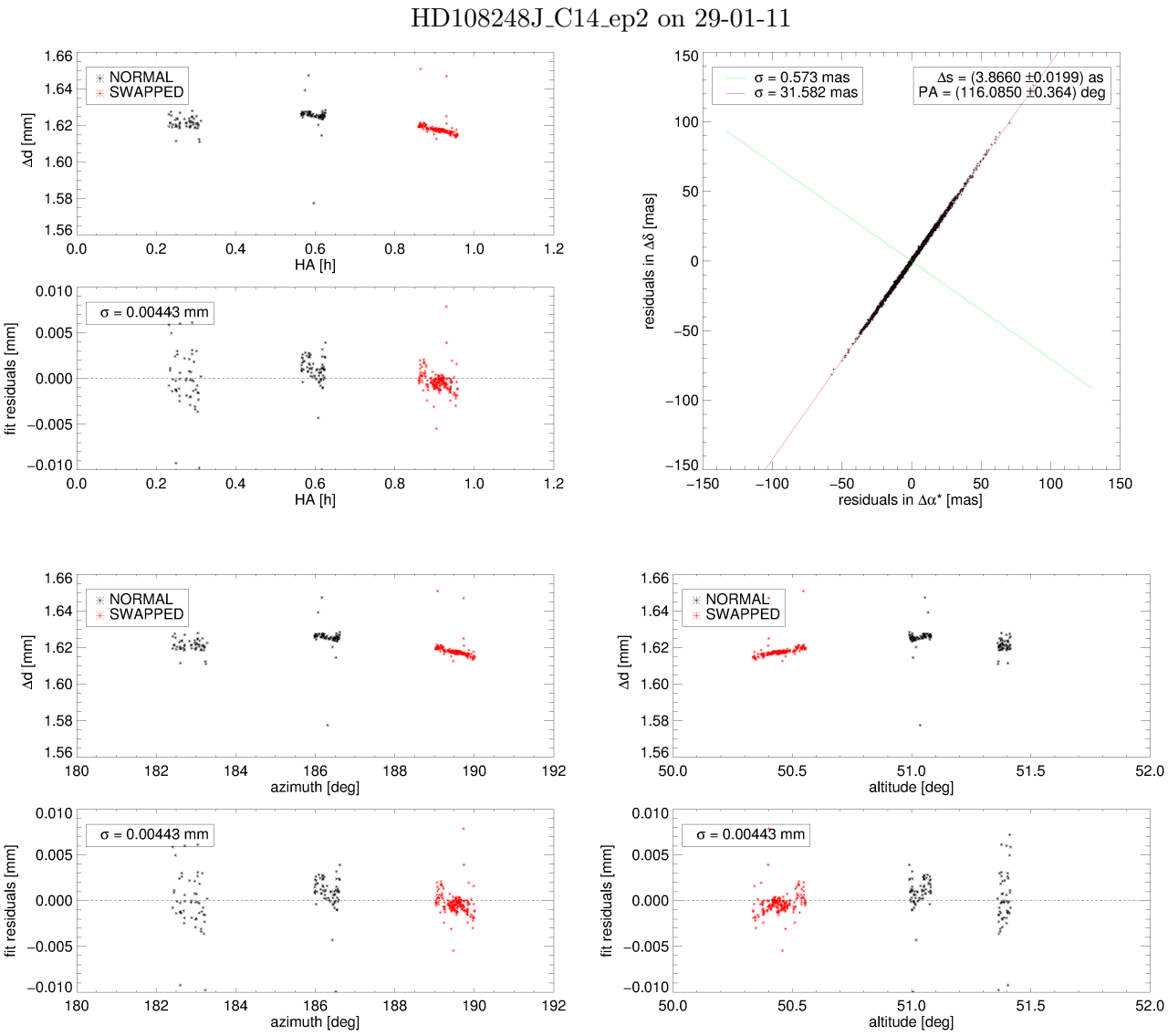


Figure D.11. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD108248J_C14_ep2 on 29-01-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

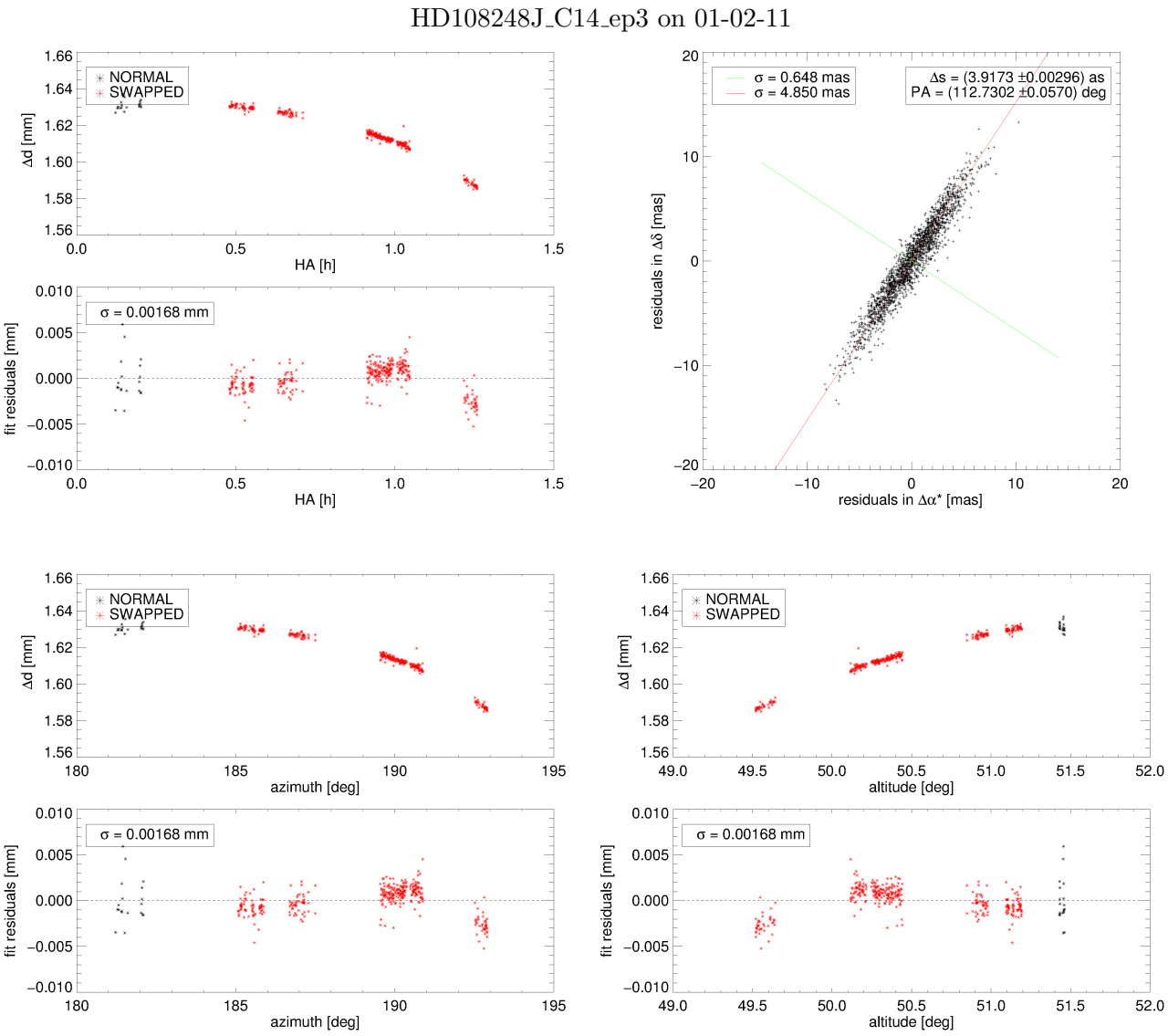


Figure D.12. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD108248J_C14_ep3 on 01-02-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

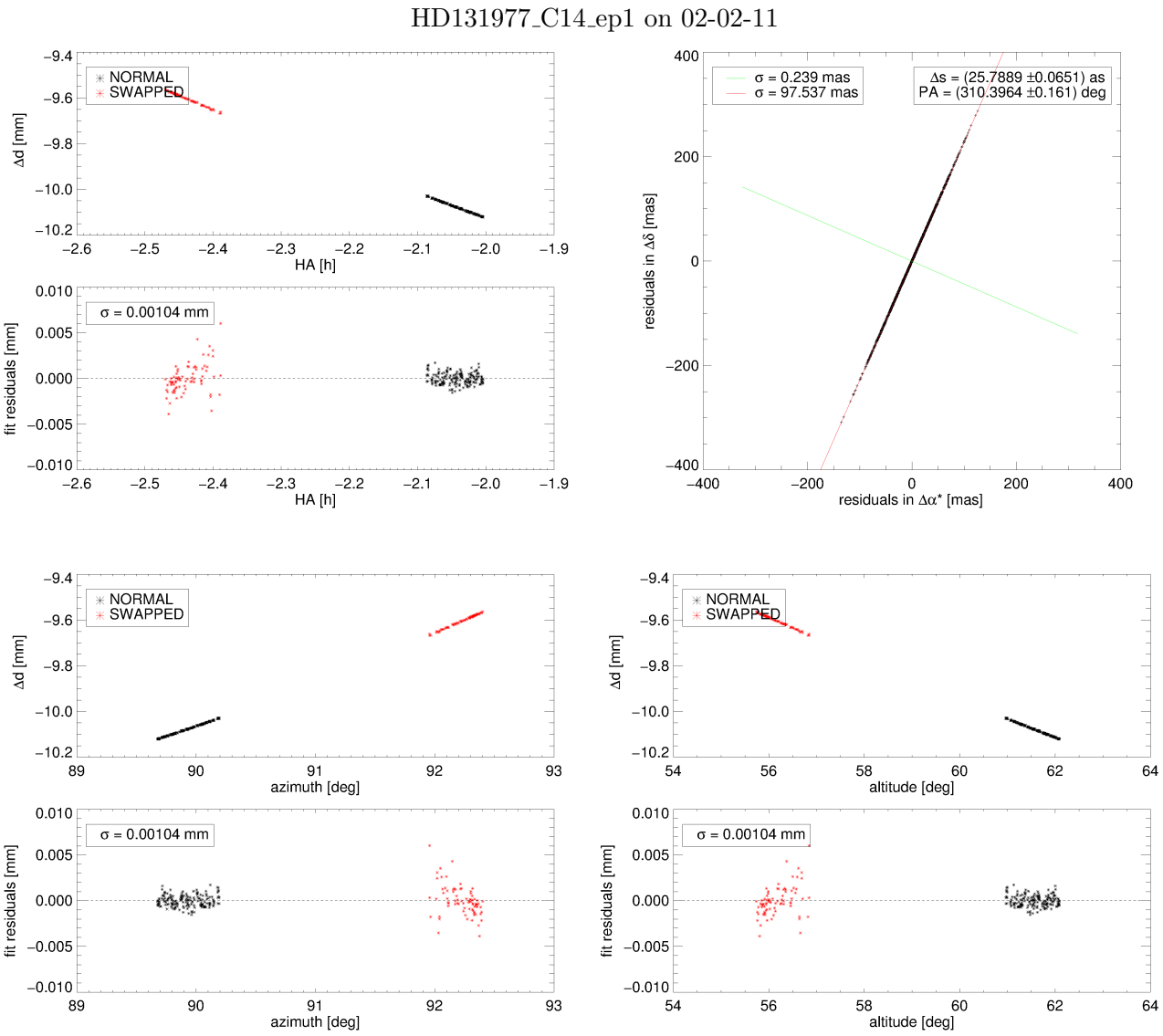


Figure D.13. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD131977_C14_ep1 on 02-02-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

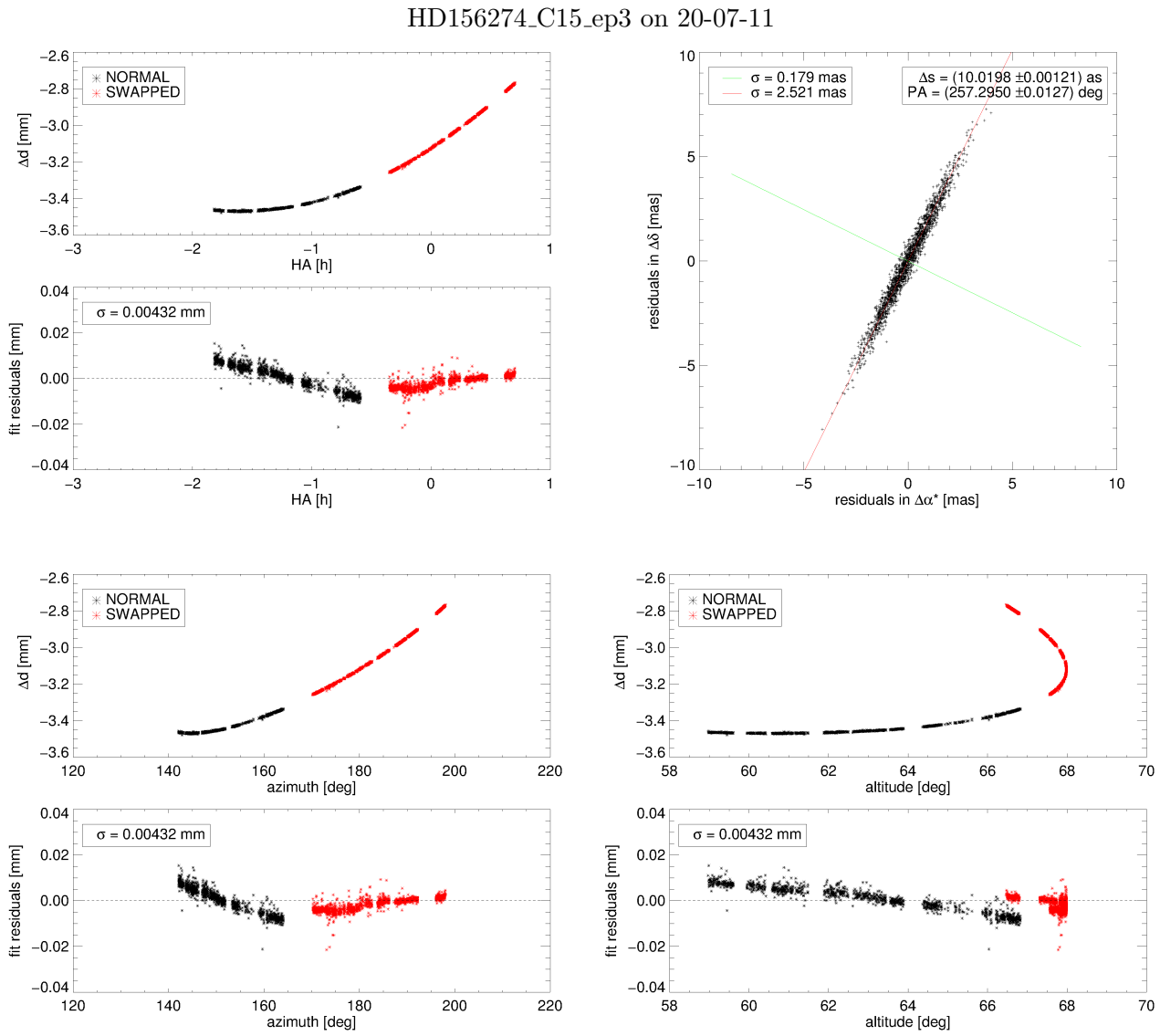


Figure D.14. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD156274_C15_ep3 on 20-07-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

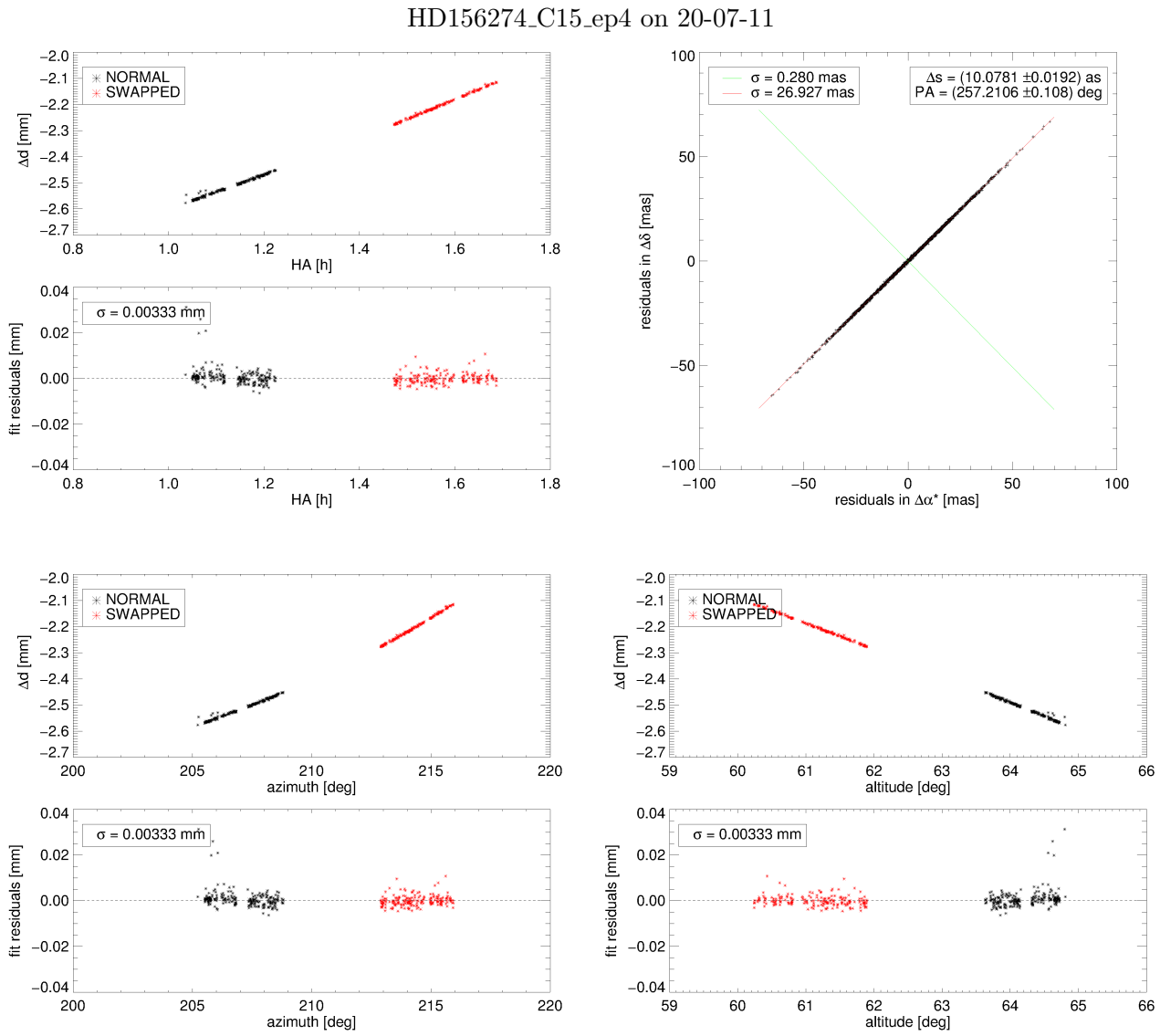


Figure D.15. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD156274_C15_ep4 on 20-07-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

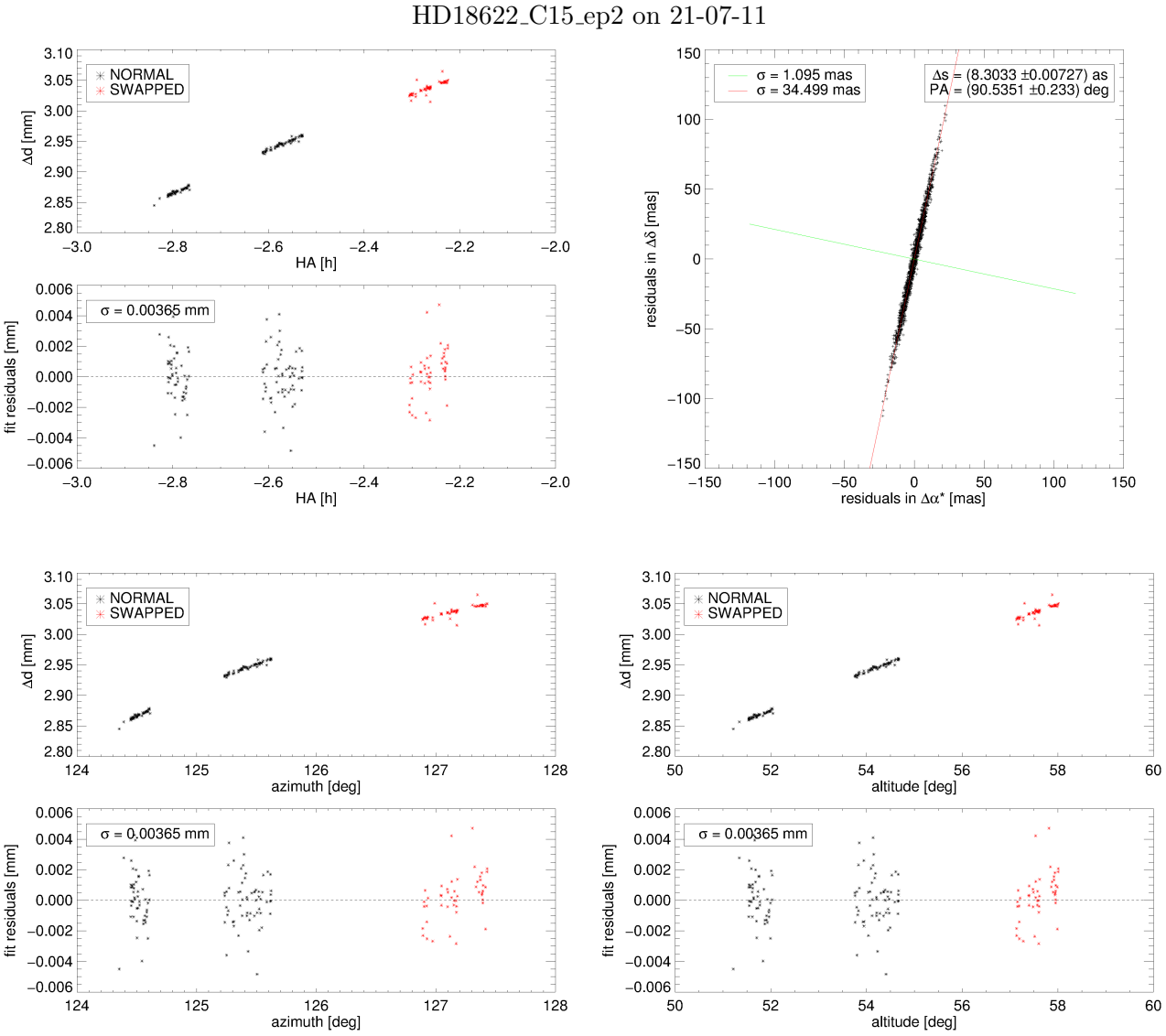


Figure D.16. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD18622_C15_ep2 on 21-07-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

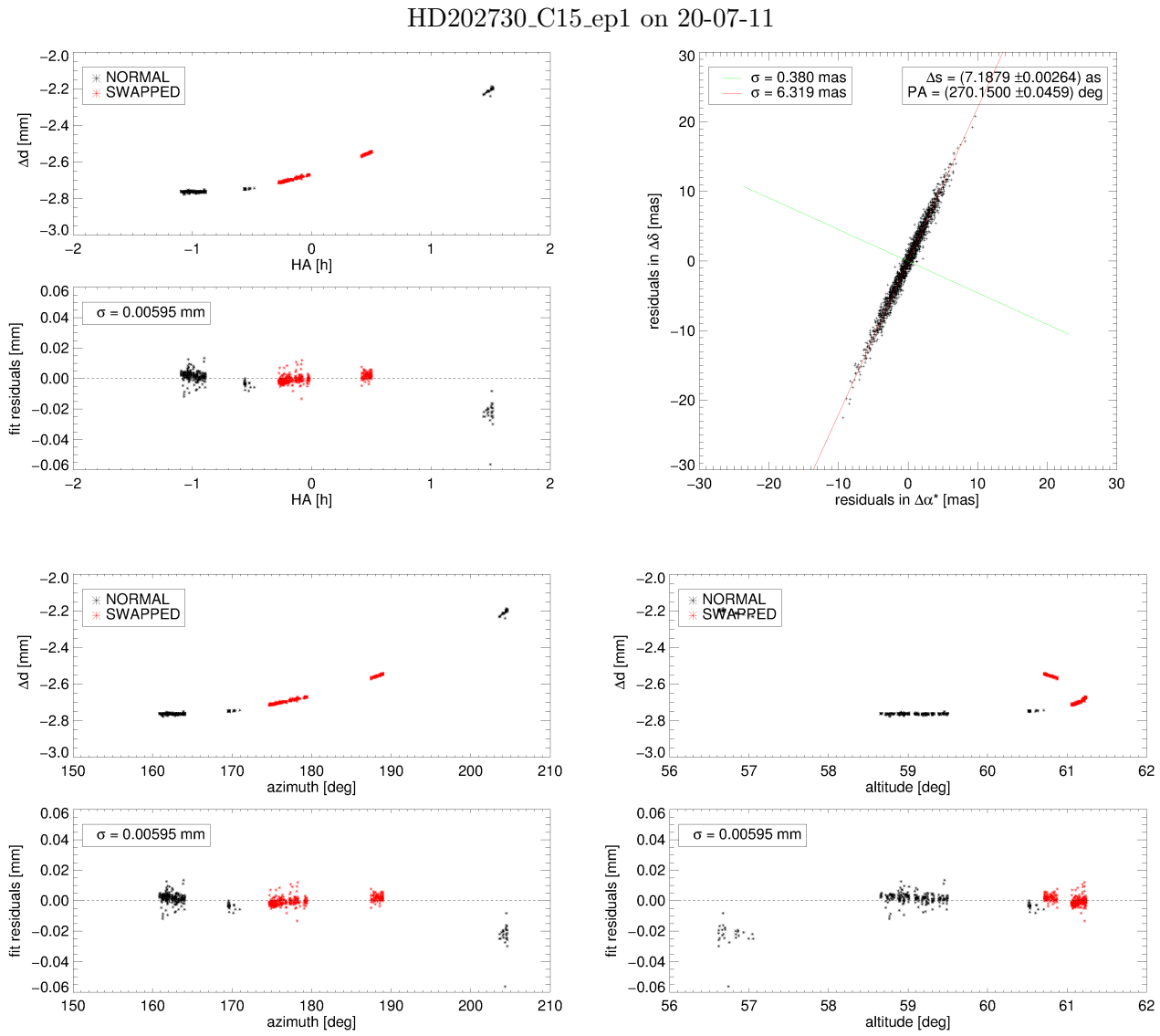


Figure D.17. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD202730_C15_ep1 on 20-07-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

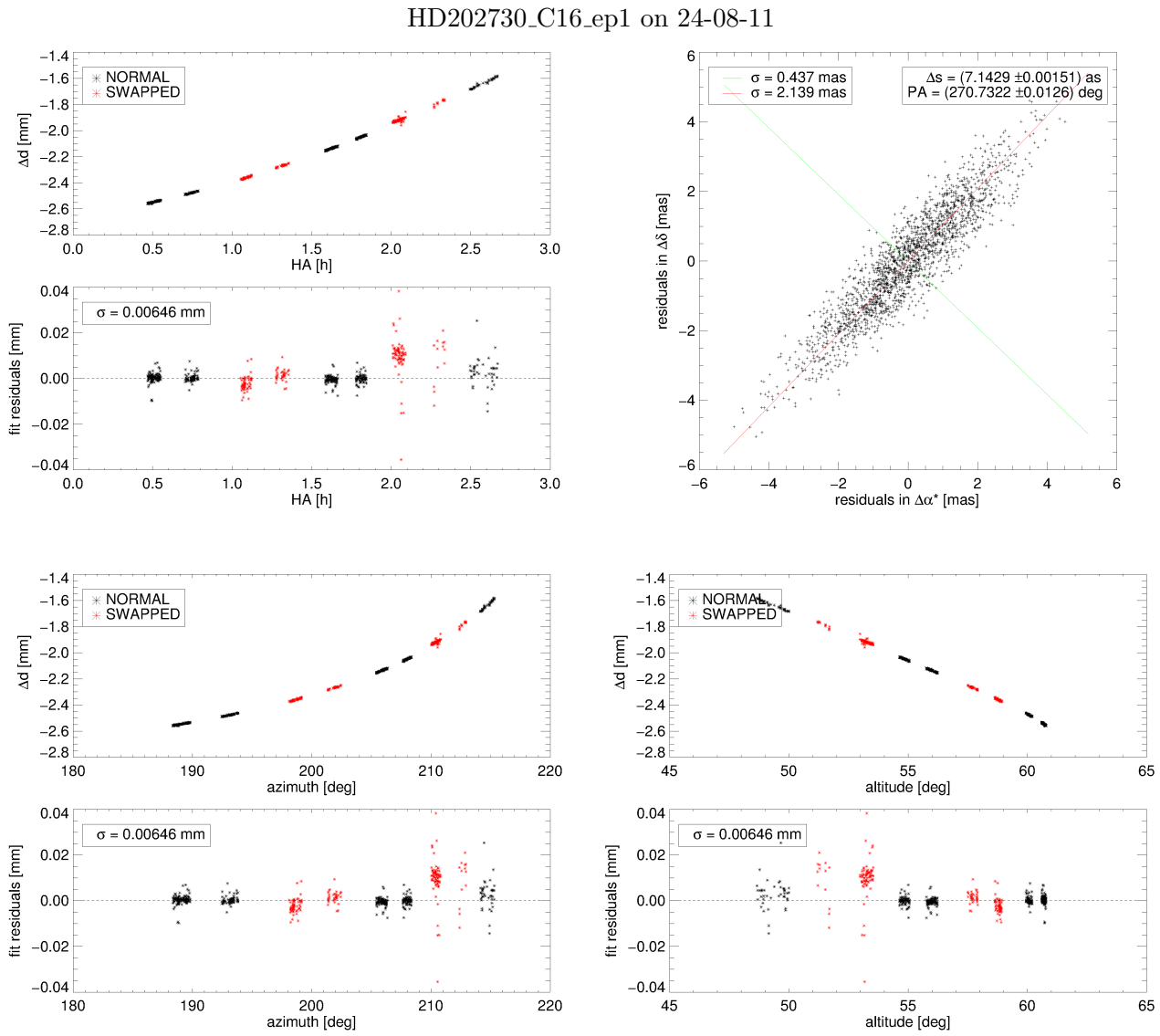


Figure D.18. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD202730_C16_ep1 on 24-08-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

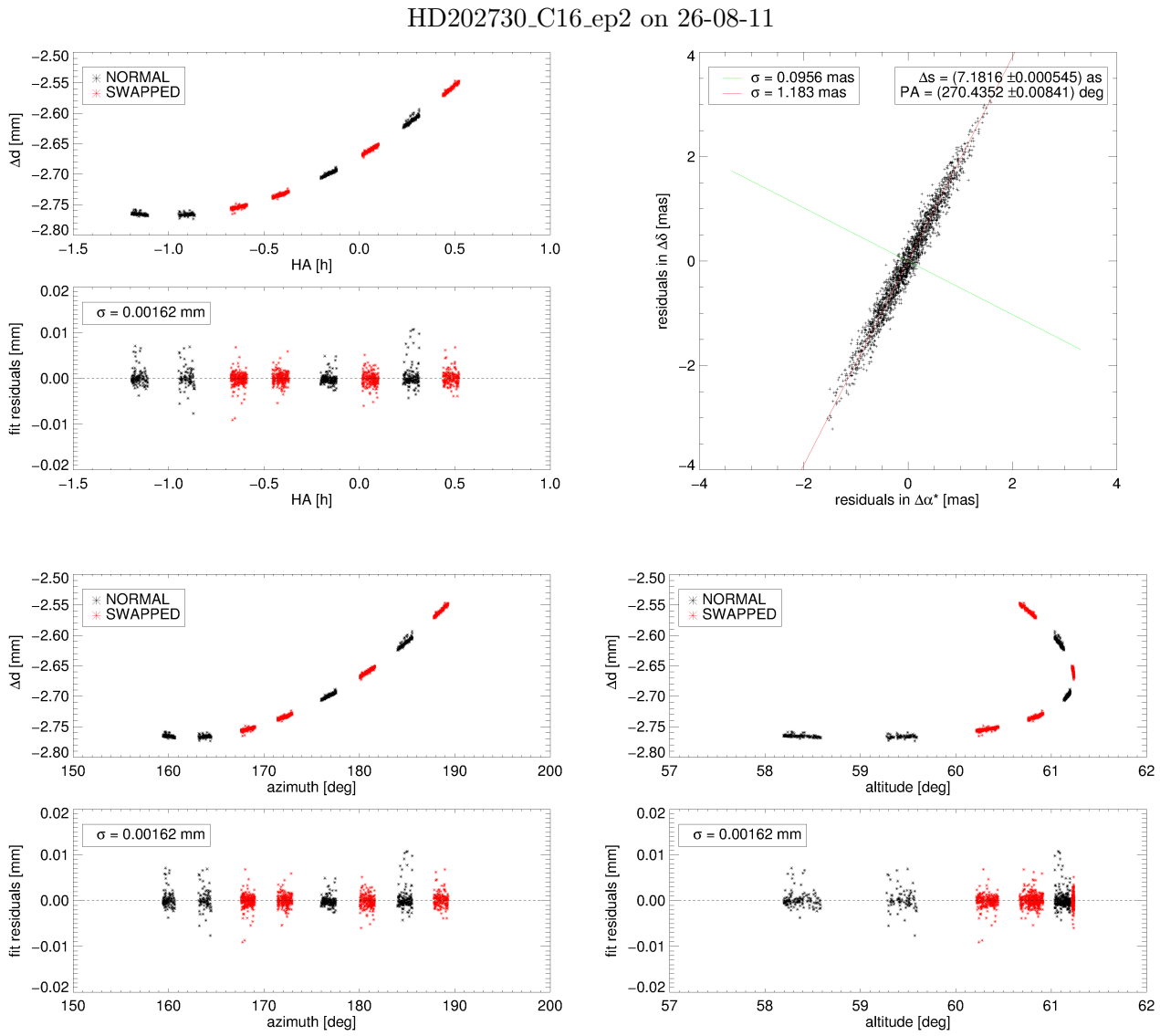


Figure D.19. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD202730_C16_ep2 on 26-08-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

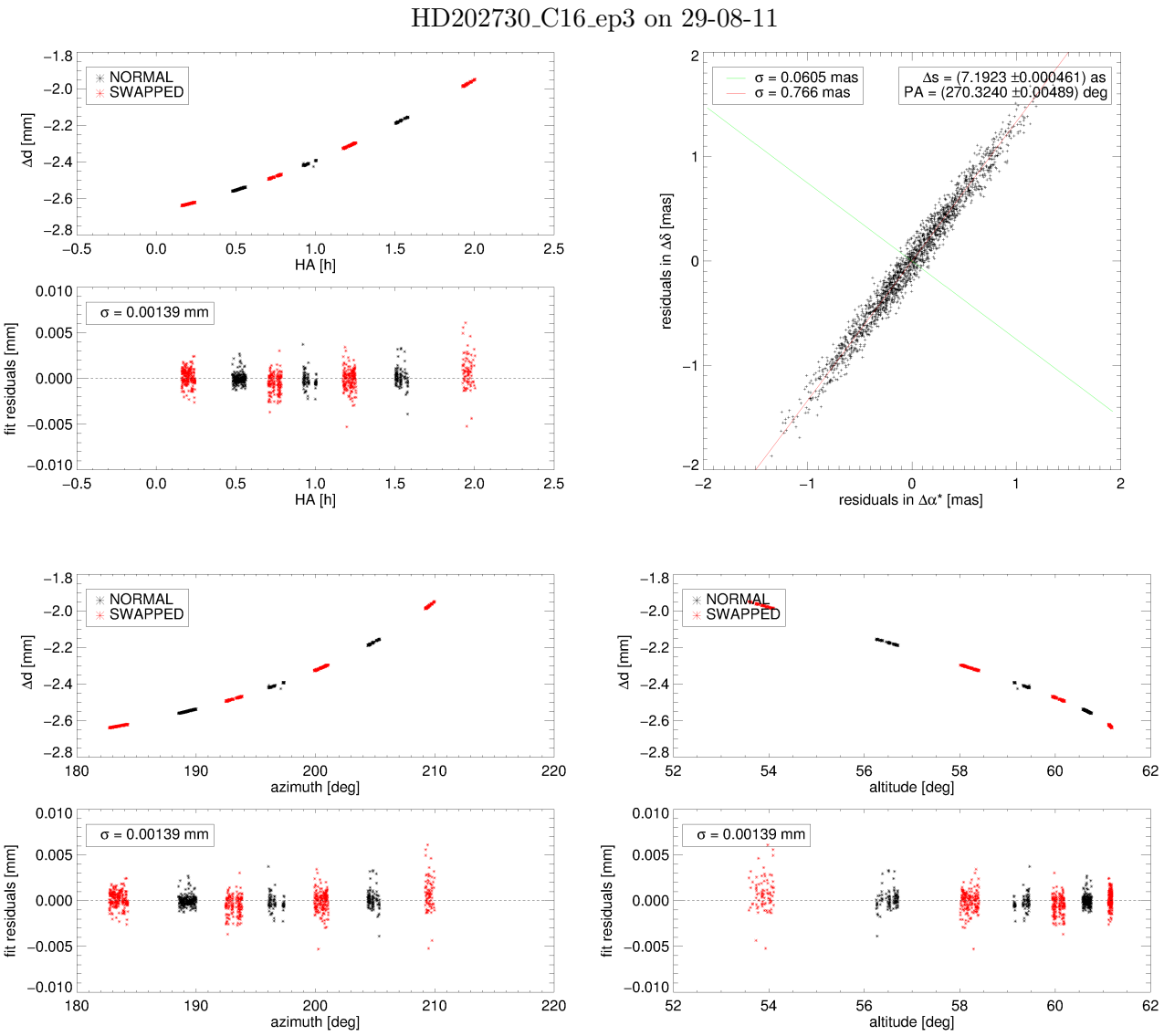


Figure D.20. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD202730_C16_ep3 on 29-08-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

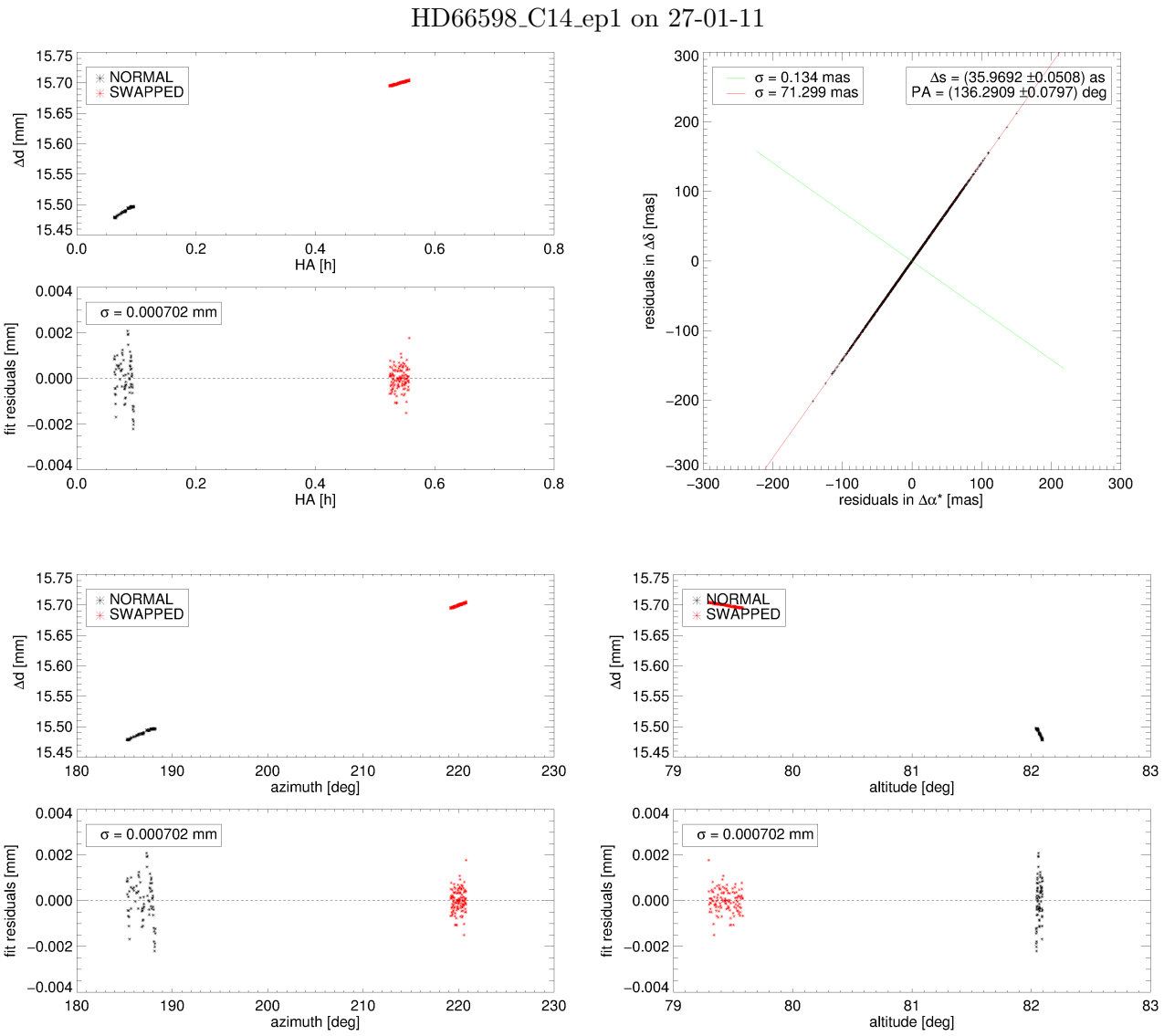


Figure D.21. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep1 on 27-01-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

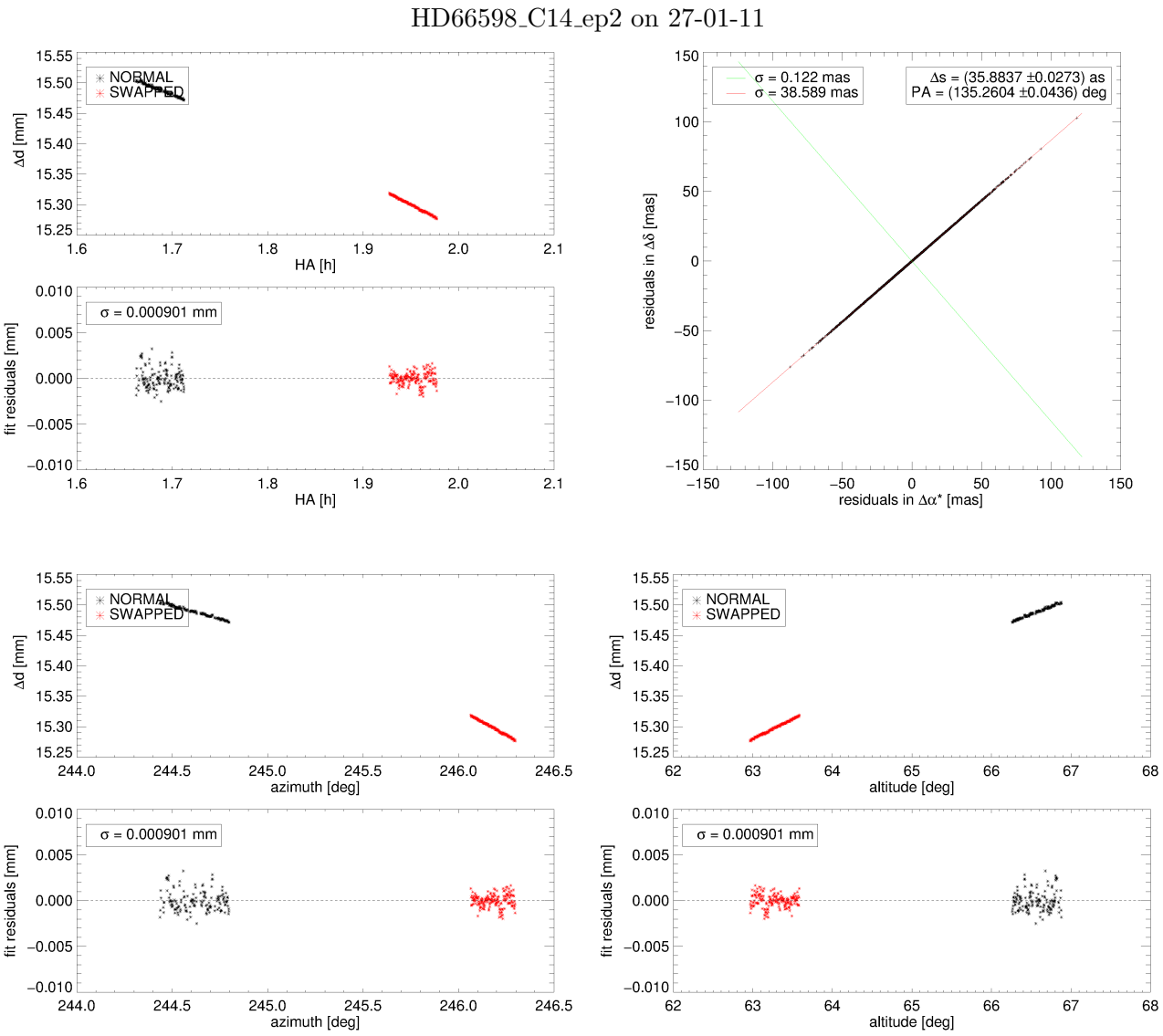


Figure D.22. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep2 on 27-01-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

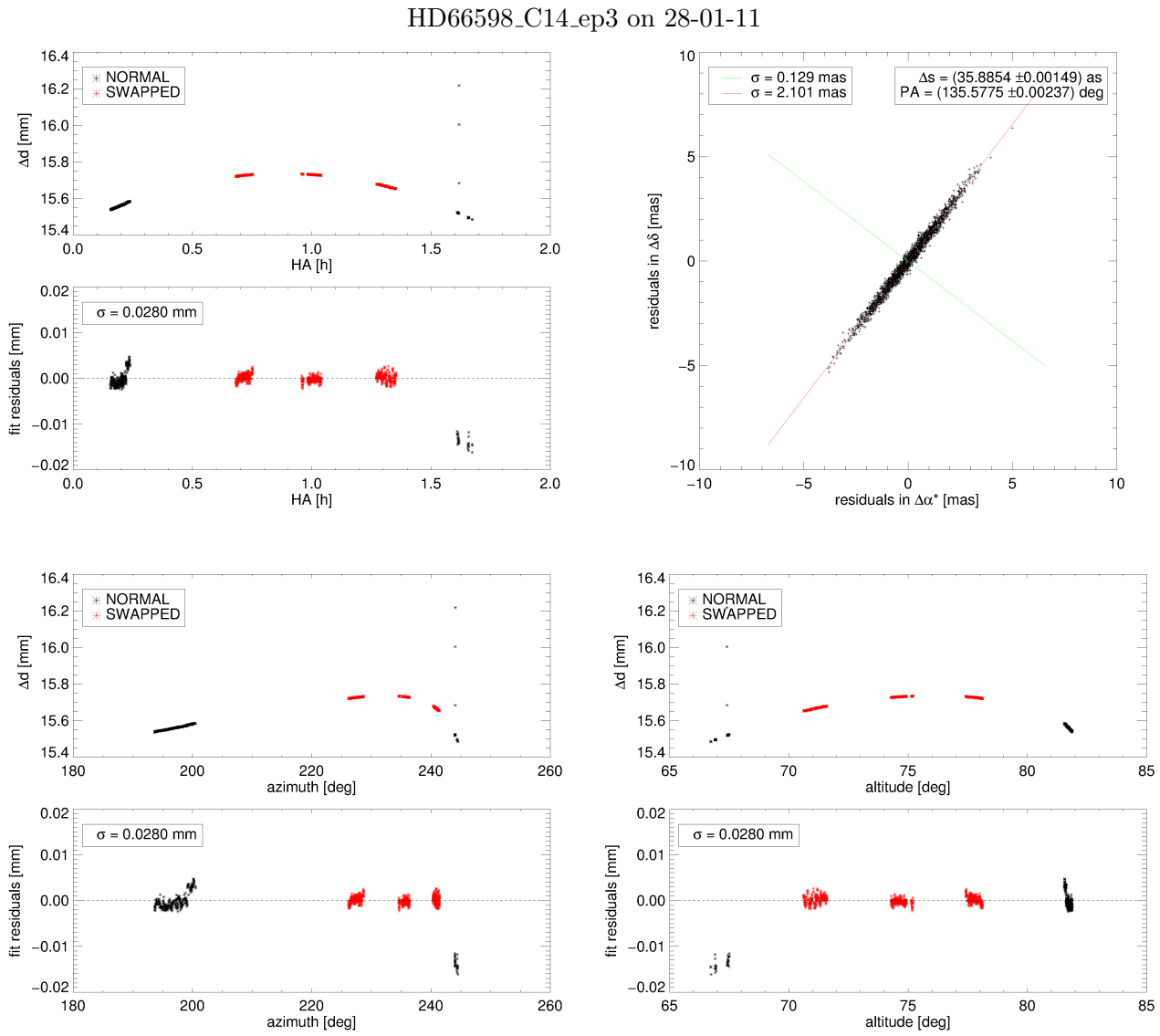


Figure D.23. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep3 on 28-01-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

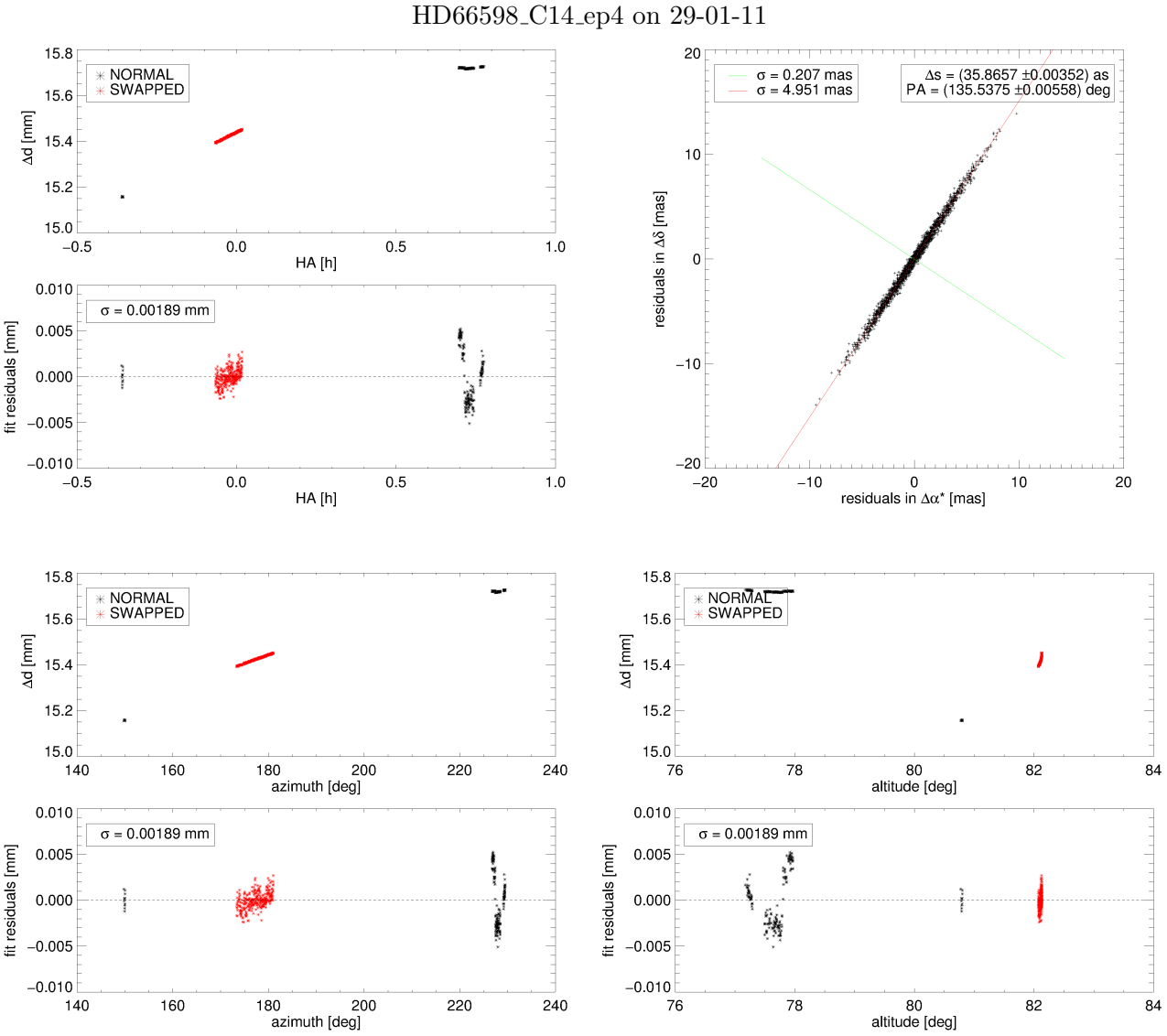


Figure D.24. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep4 on 29-01-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

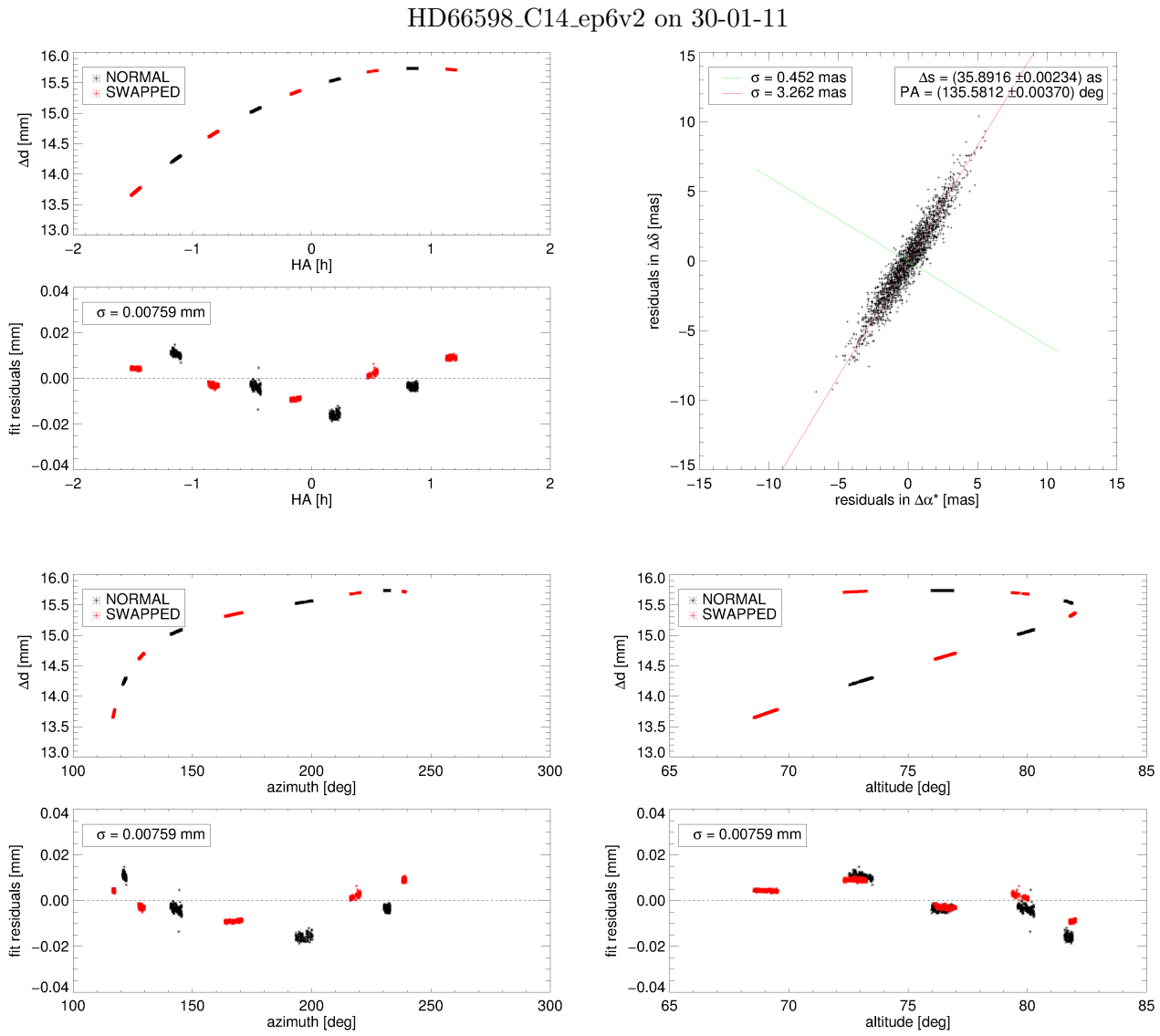


Figure D.25. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep6v2 on 30-01-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

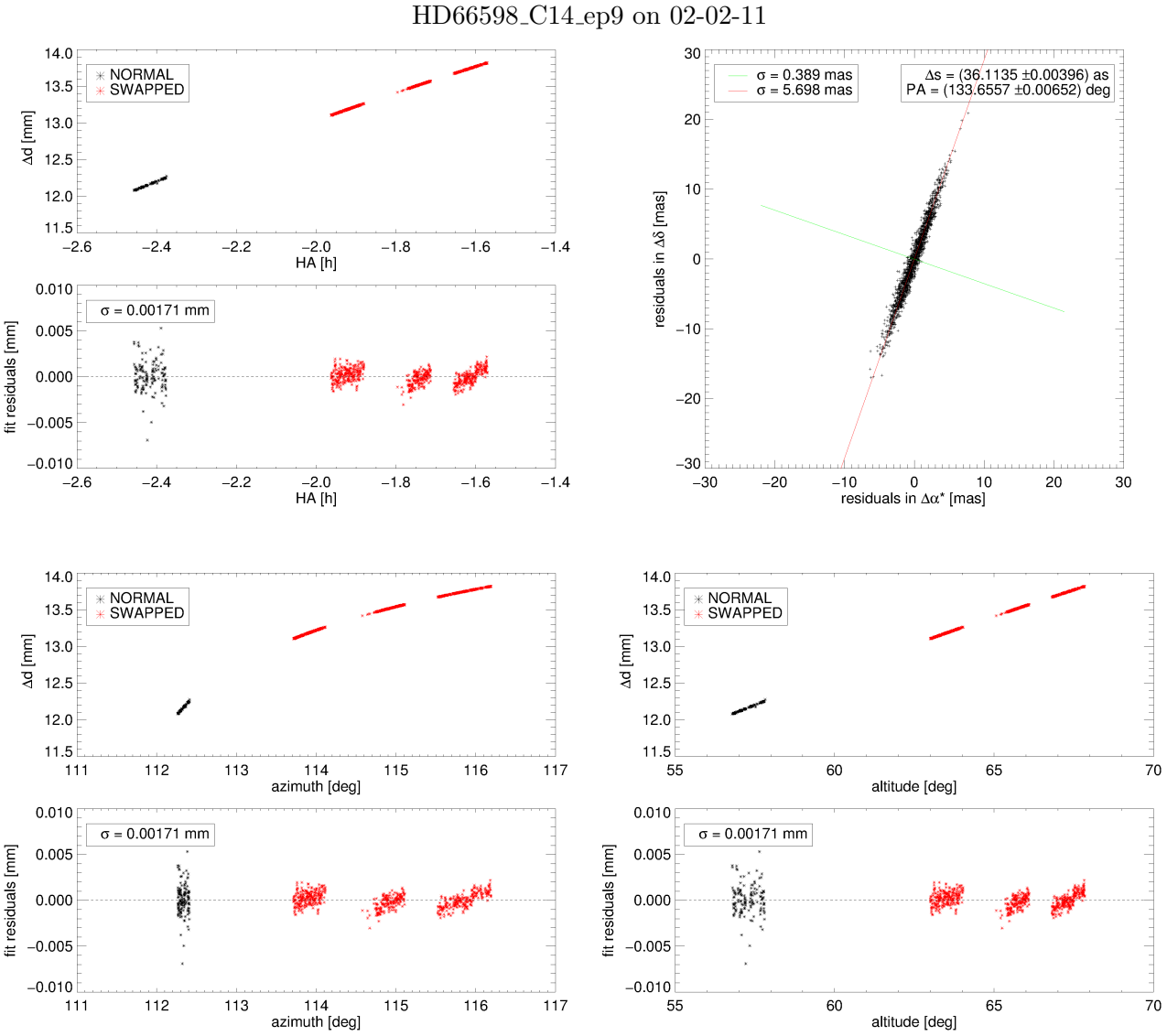


Figure D.26. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep9 on 02-02-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

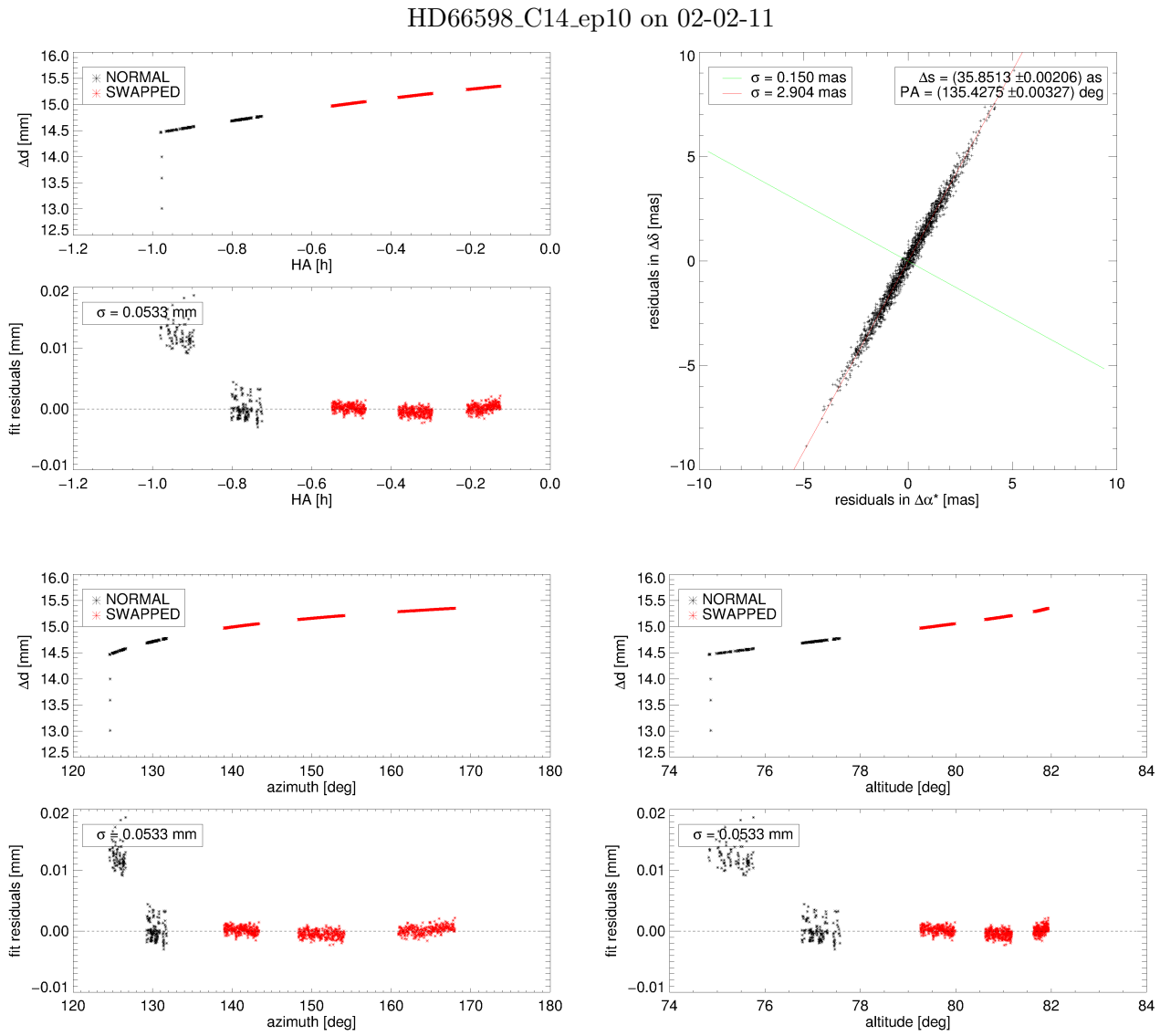


Figure D.27. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep10 on 02-02-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

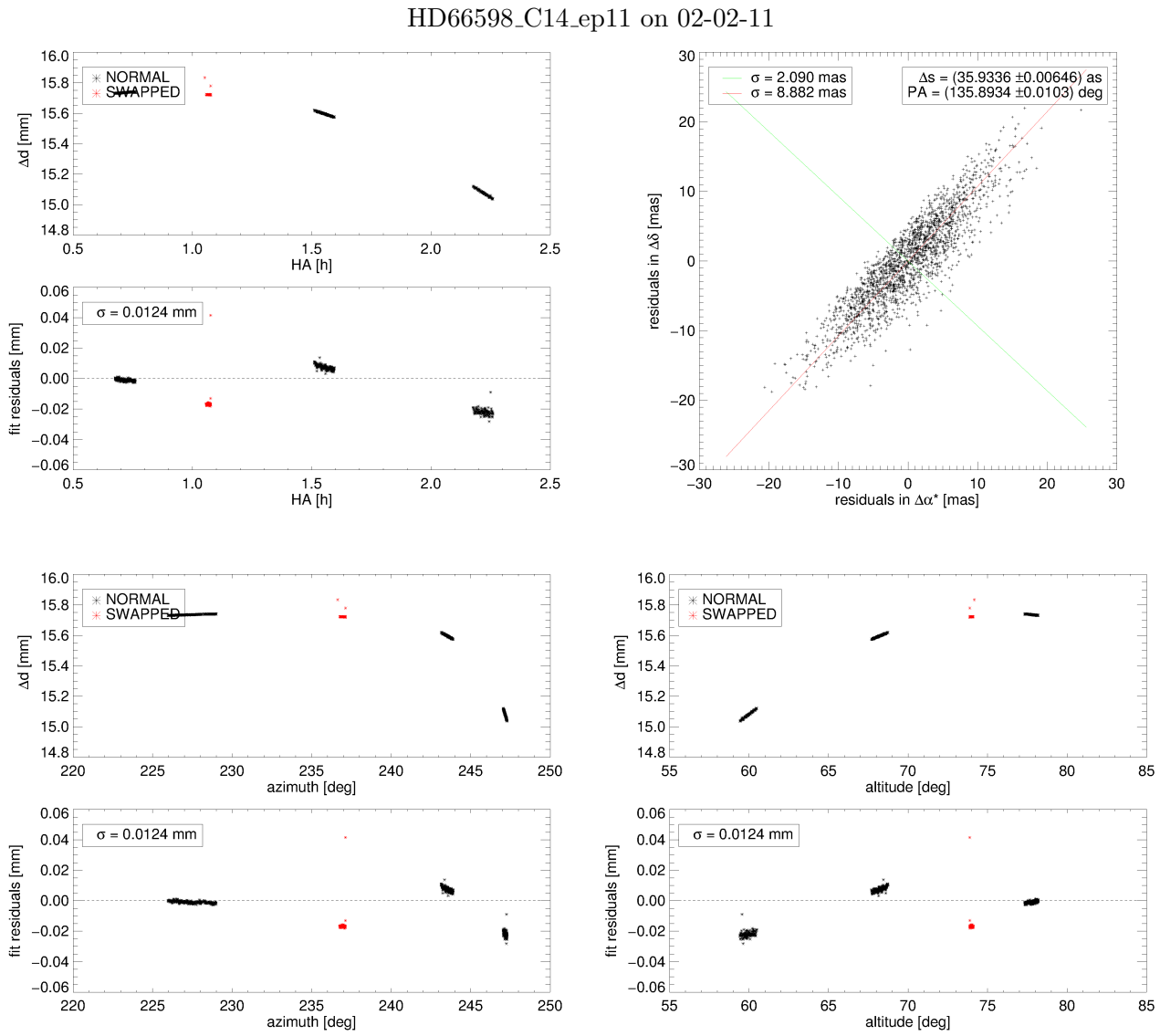


Figure D.28. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep11 on 02-02-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

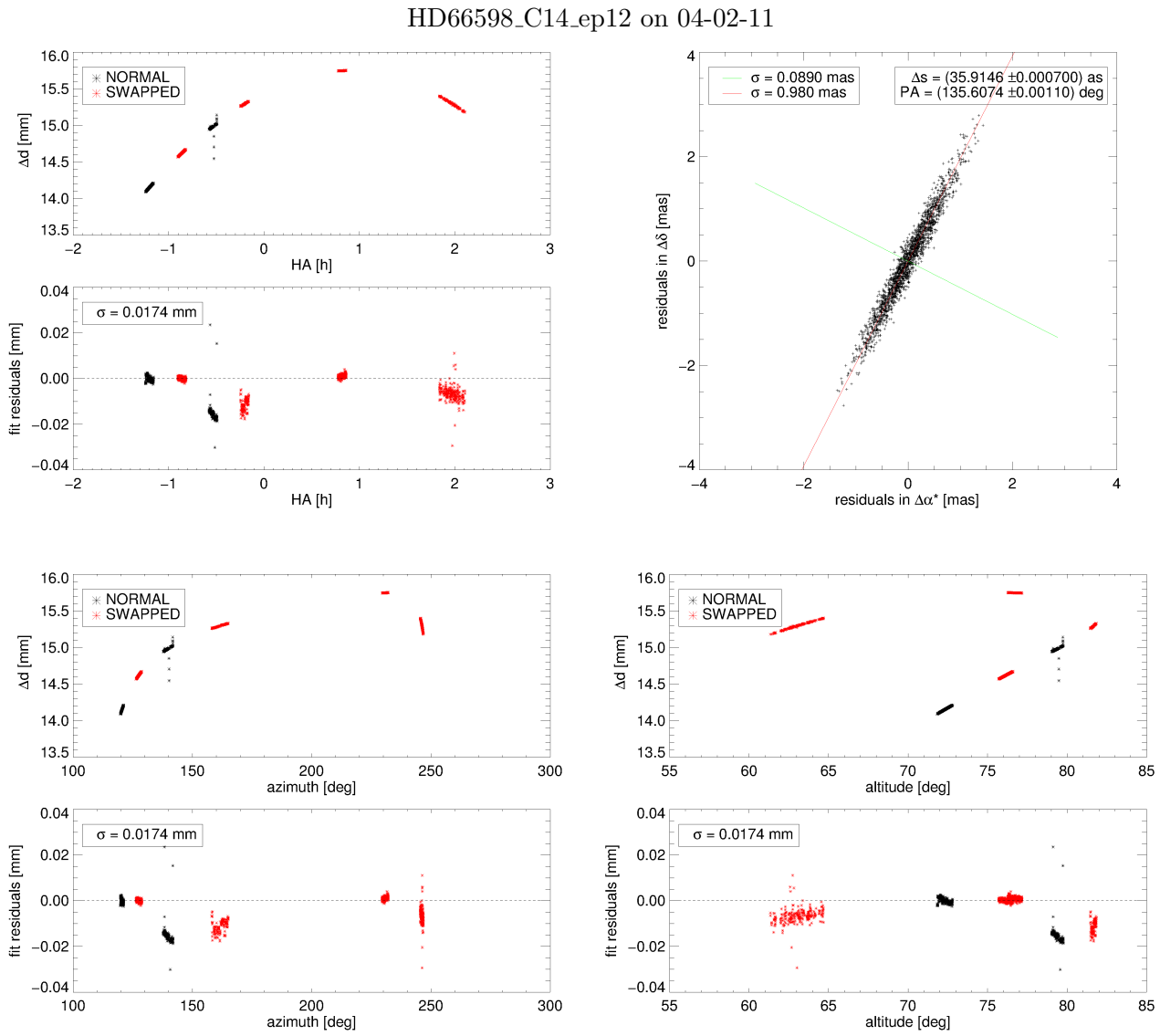


Figure D.29. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C14_ep12 on 04-02-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

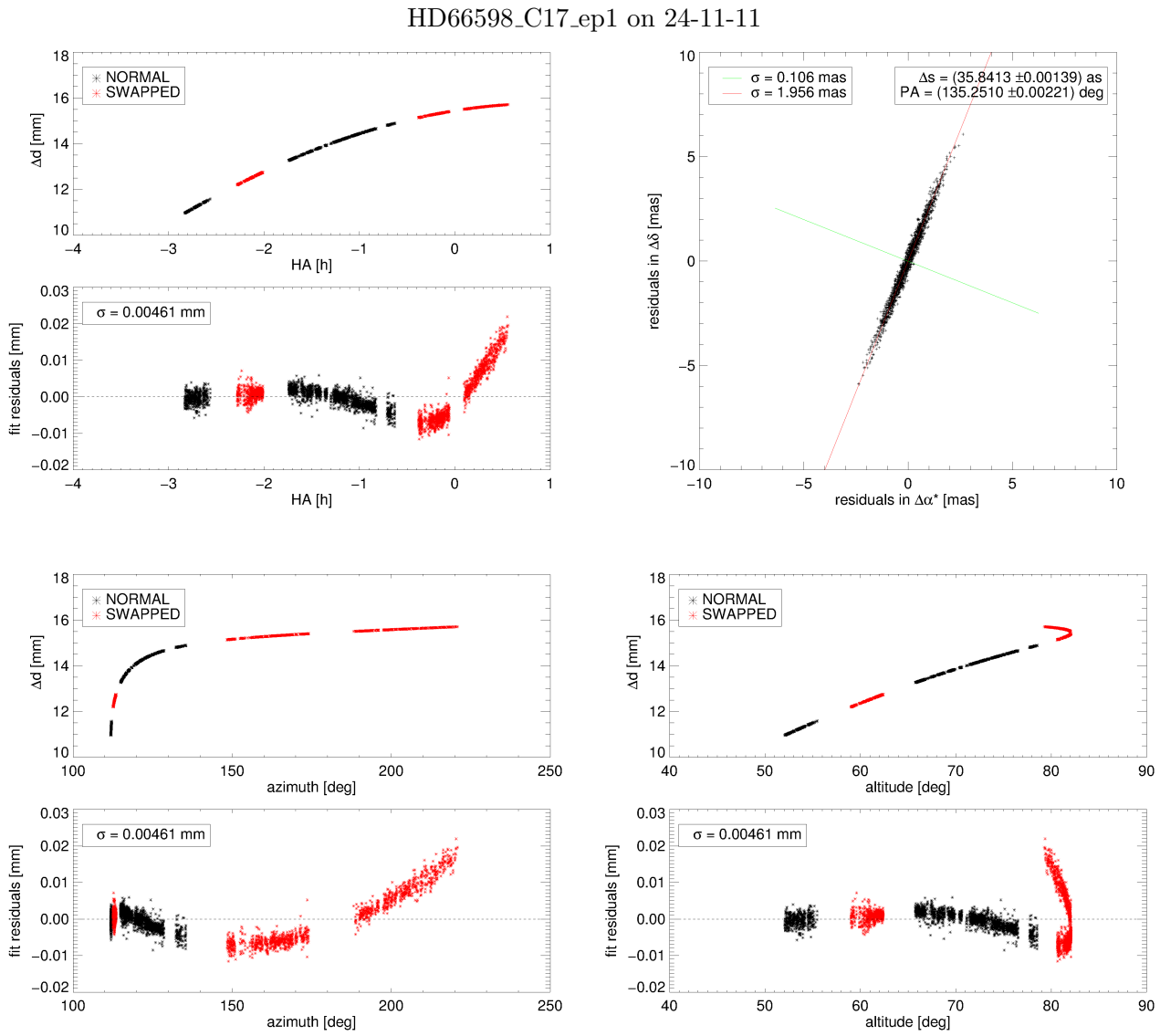


Figure D.30. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C17_ep1 on 24-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

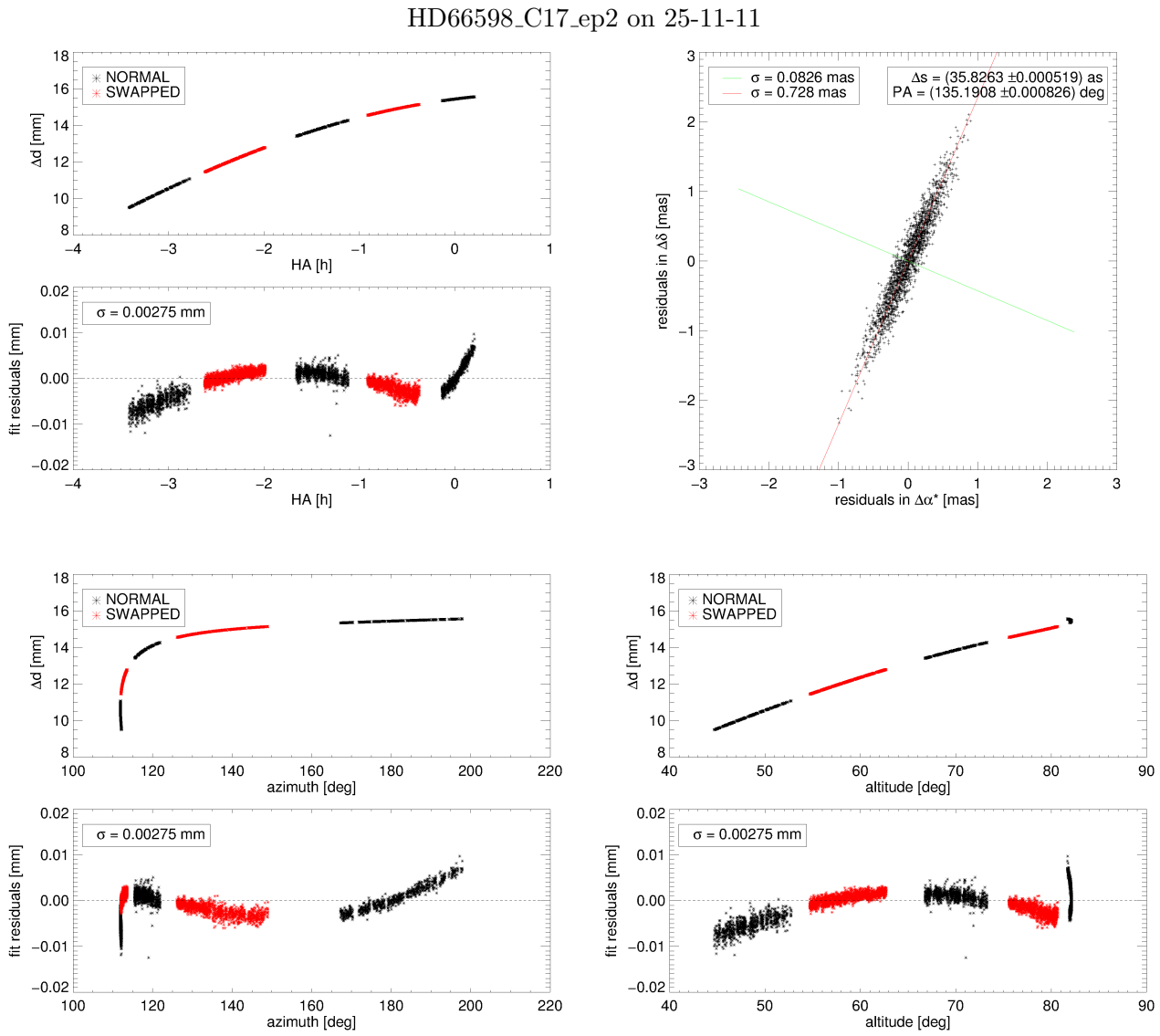


Figure D.31. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C17_ep2 on 25-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

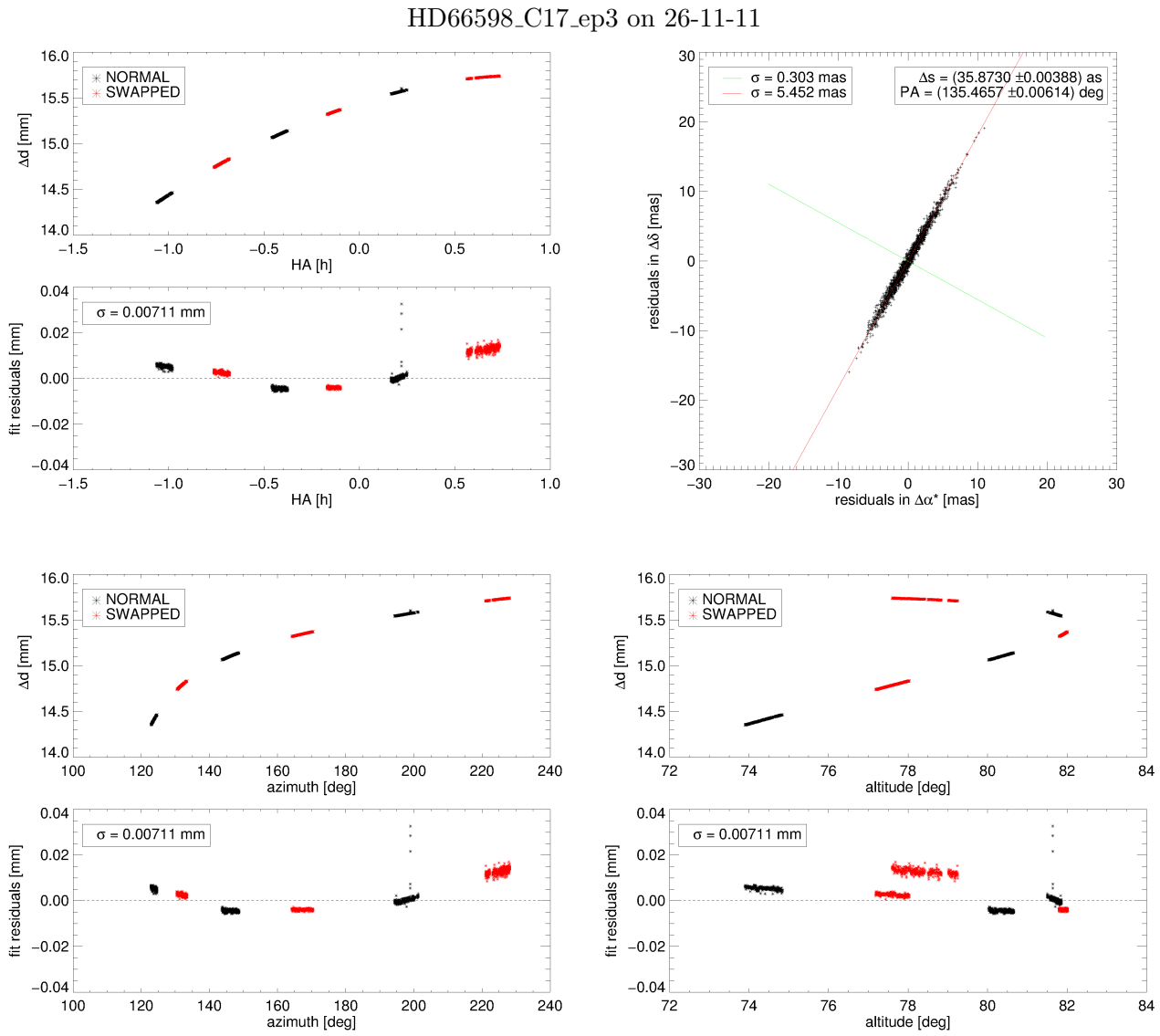


Figure D.32. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C17_ep3 on 26-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

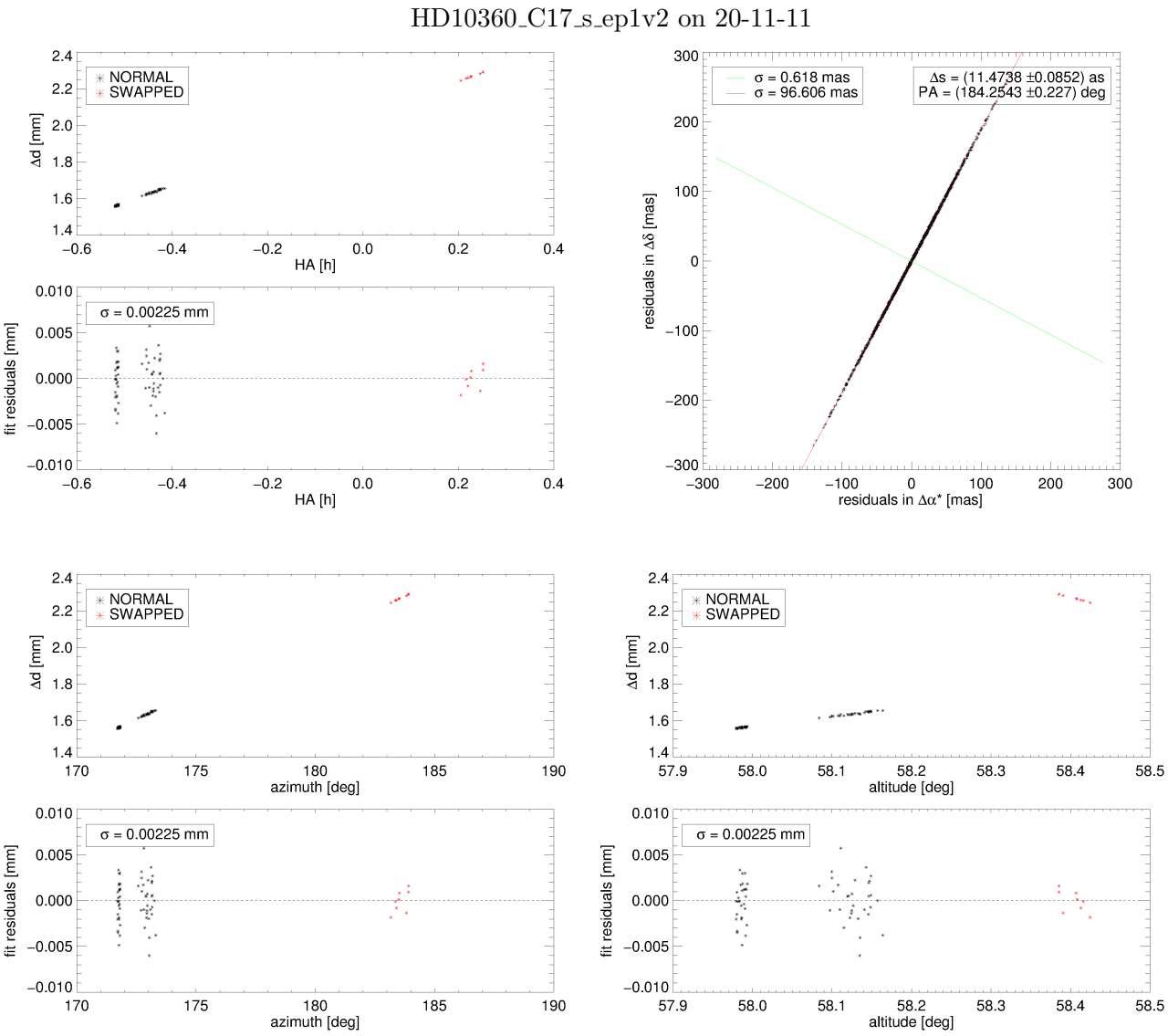


Figure D.33. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_s_ep1v2 on 20-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

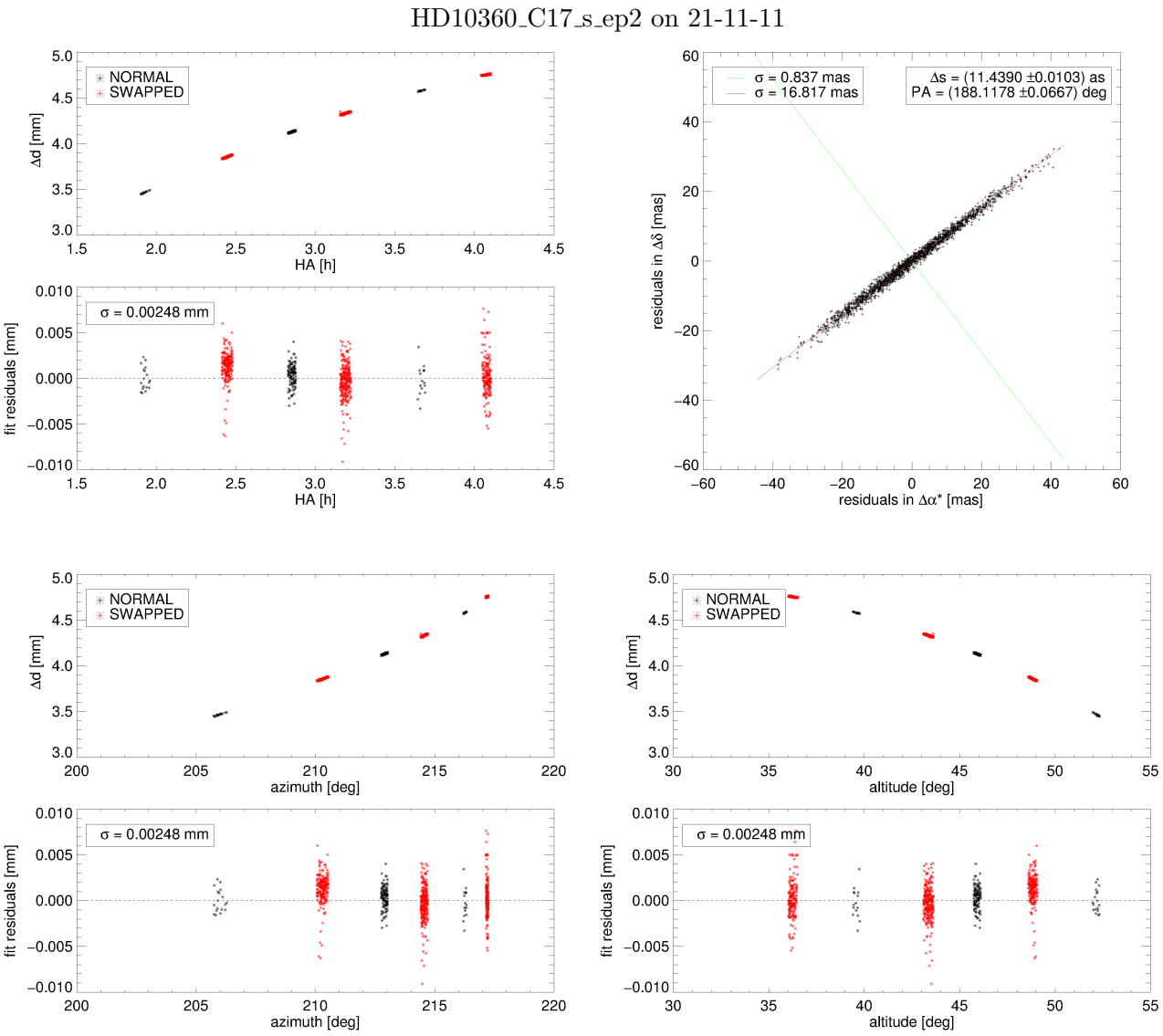


Figure D.34. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_s_ep2 on 21-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

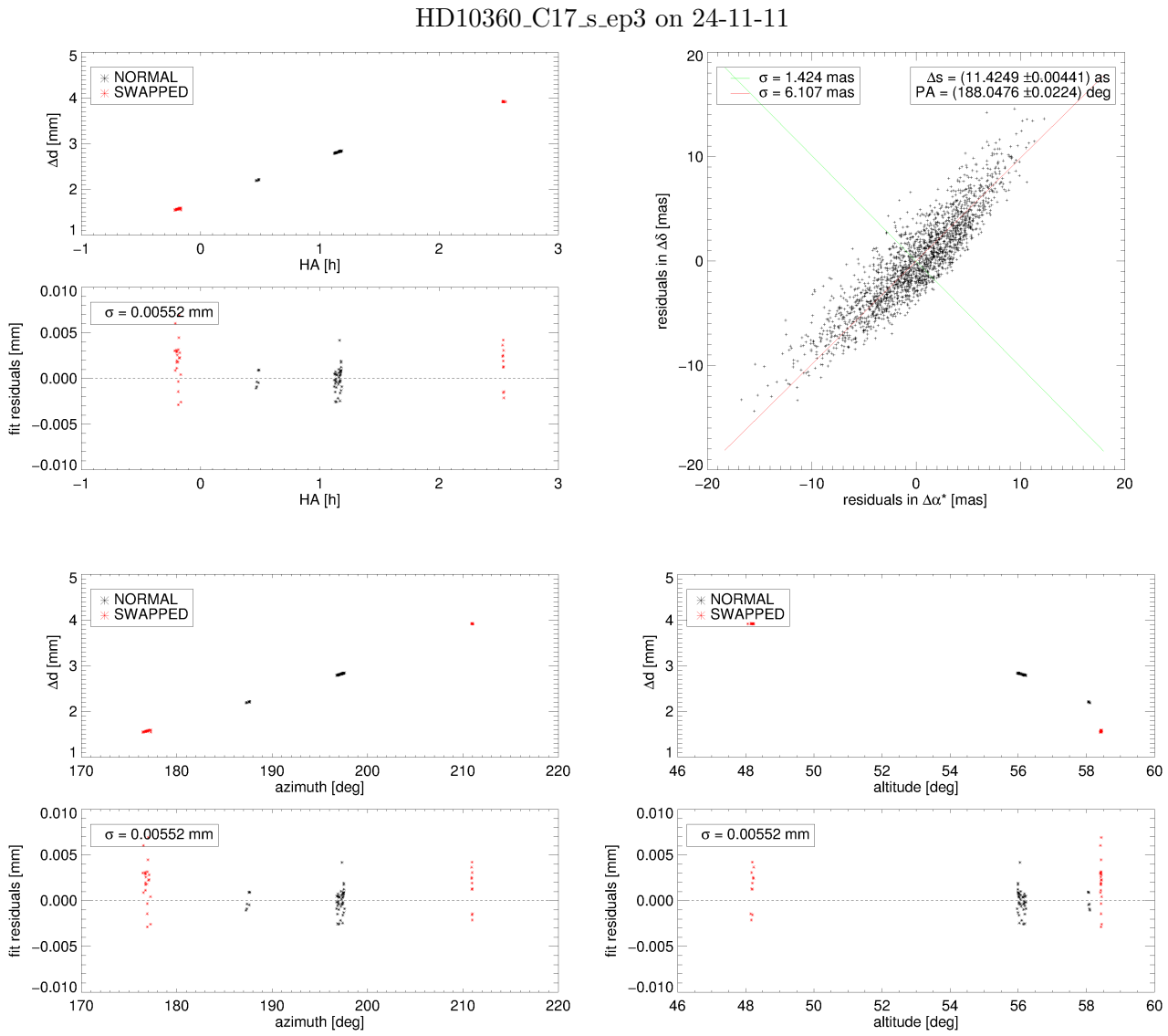


Figure D.35. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD10360_C17_s_ep3 on 24-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

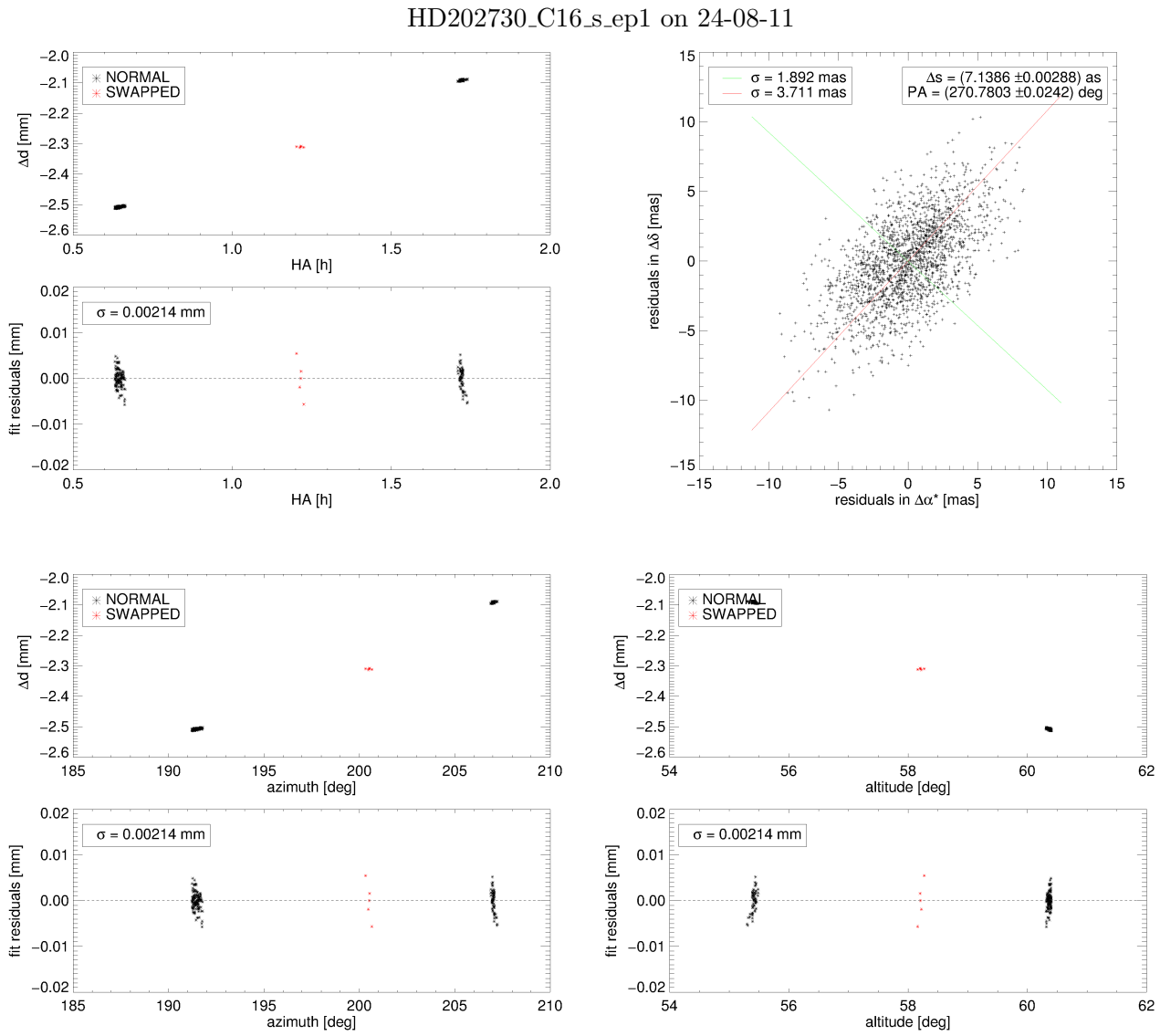


Figure D.36. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD202730_C16_s_ep1 on 24-08-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

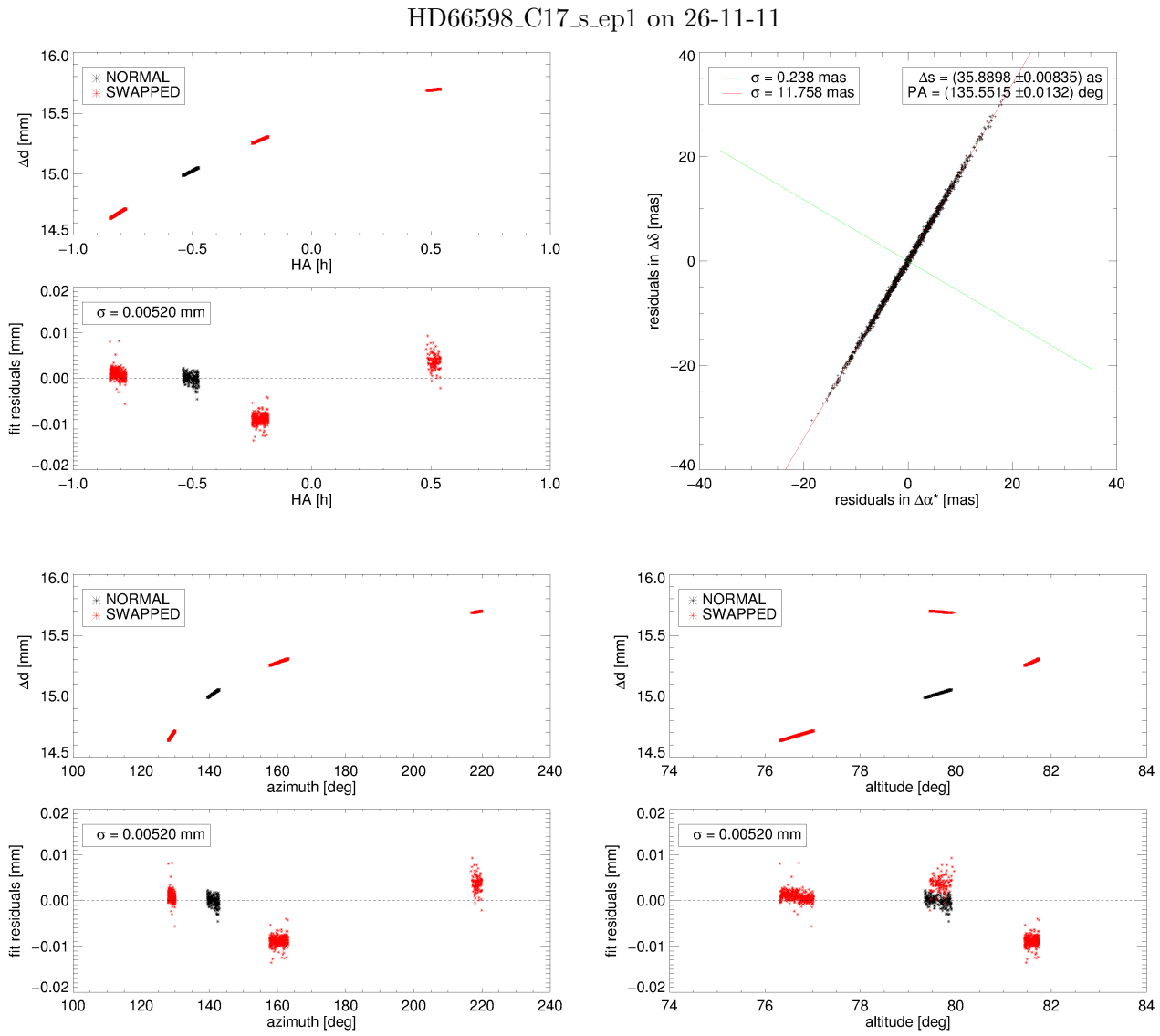


Figure D.37. – Corrected differential delays Δd and residuals of the astrometric fit to the observation sequence HD66598_C17_s_ep1 on 26-11-11, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

Multi Epoch Fits:

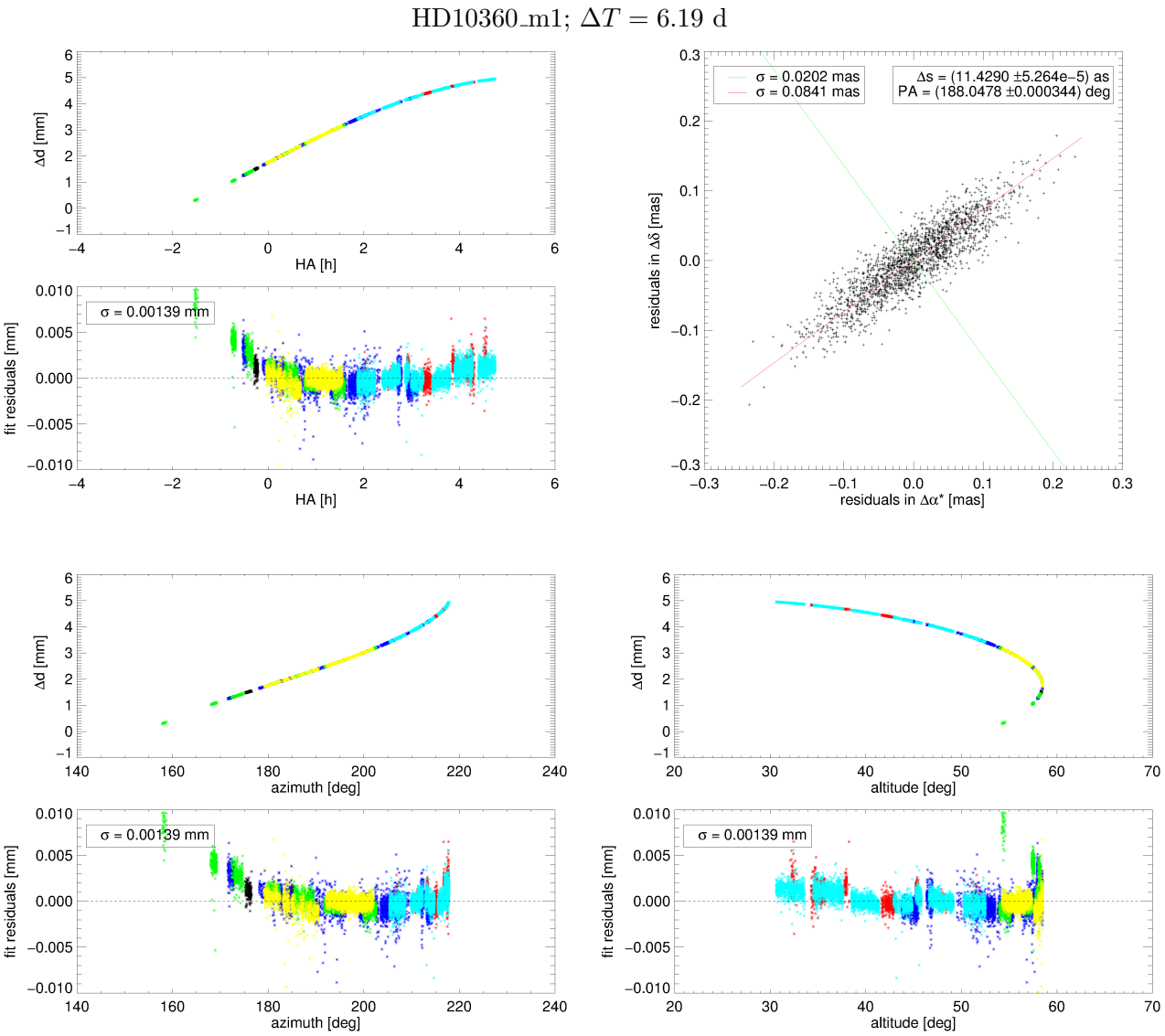


Figure D.38. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD10360_m1, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

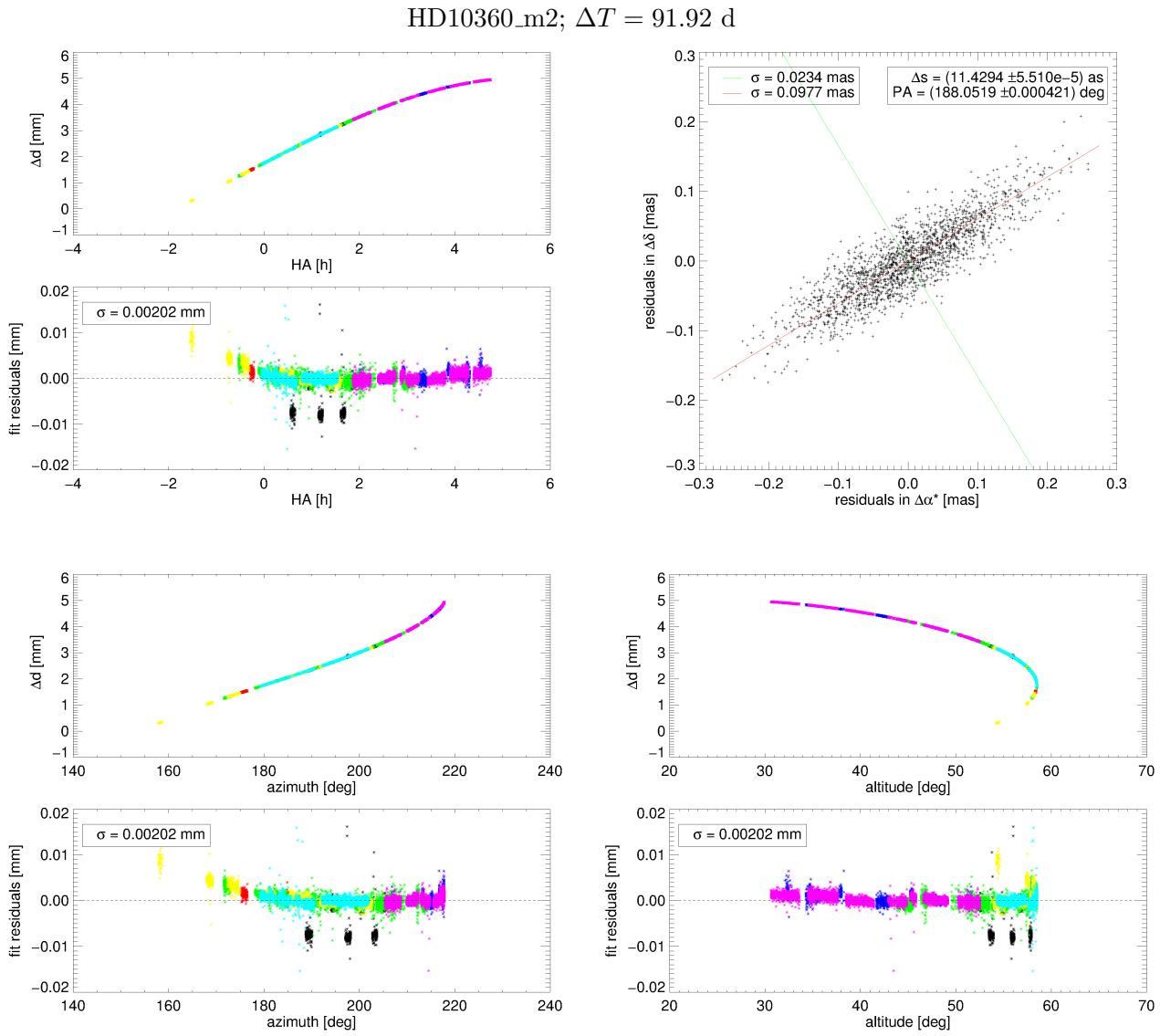


Figure D.39. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD10360_m2, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

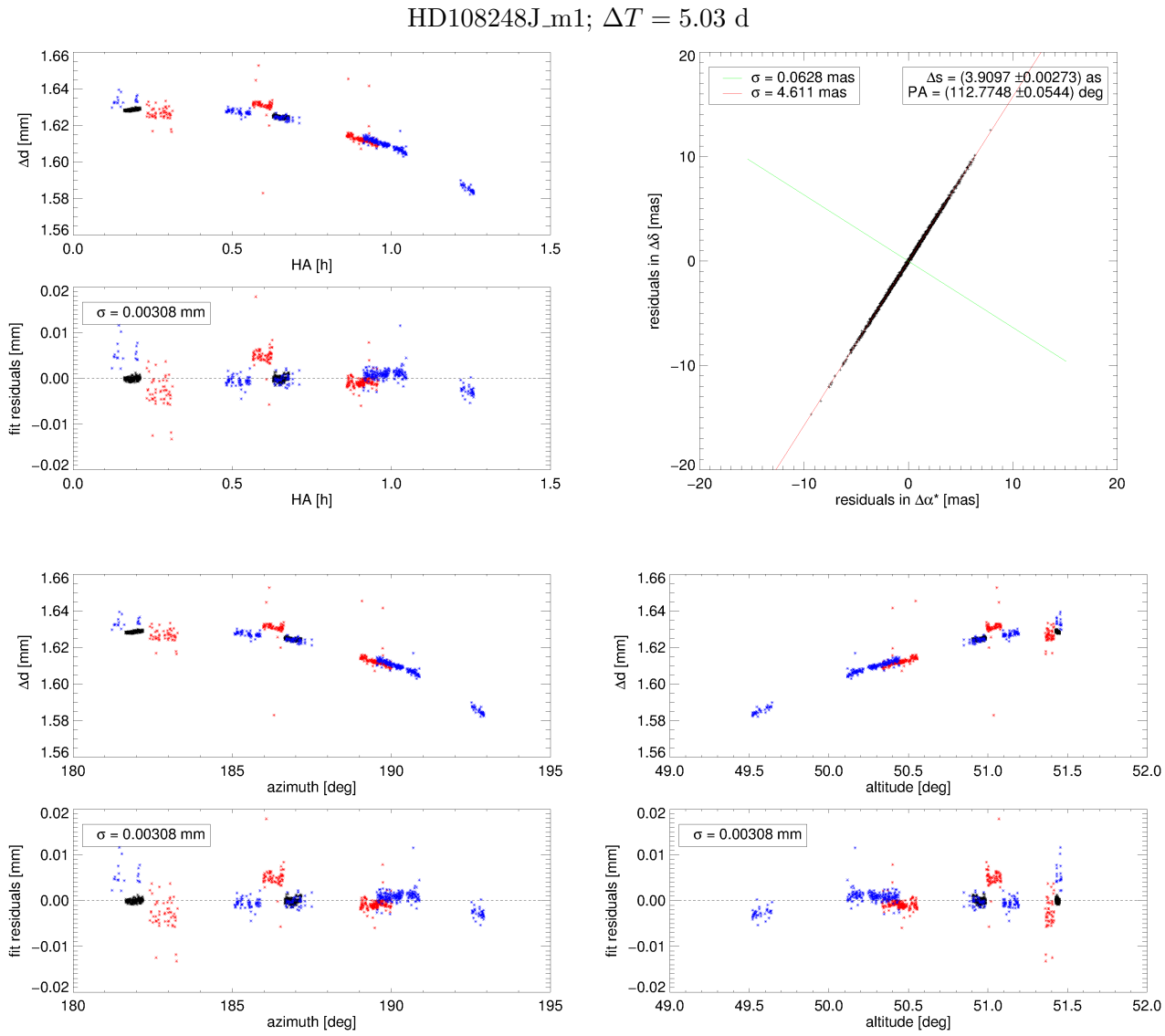


Figure D.40. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD108248J_m1, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

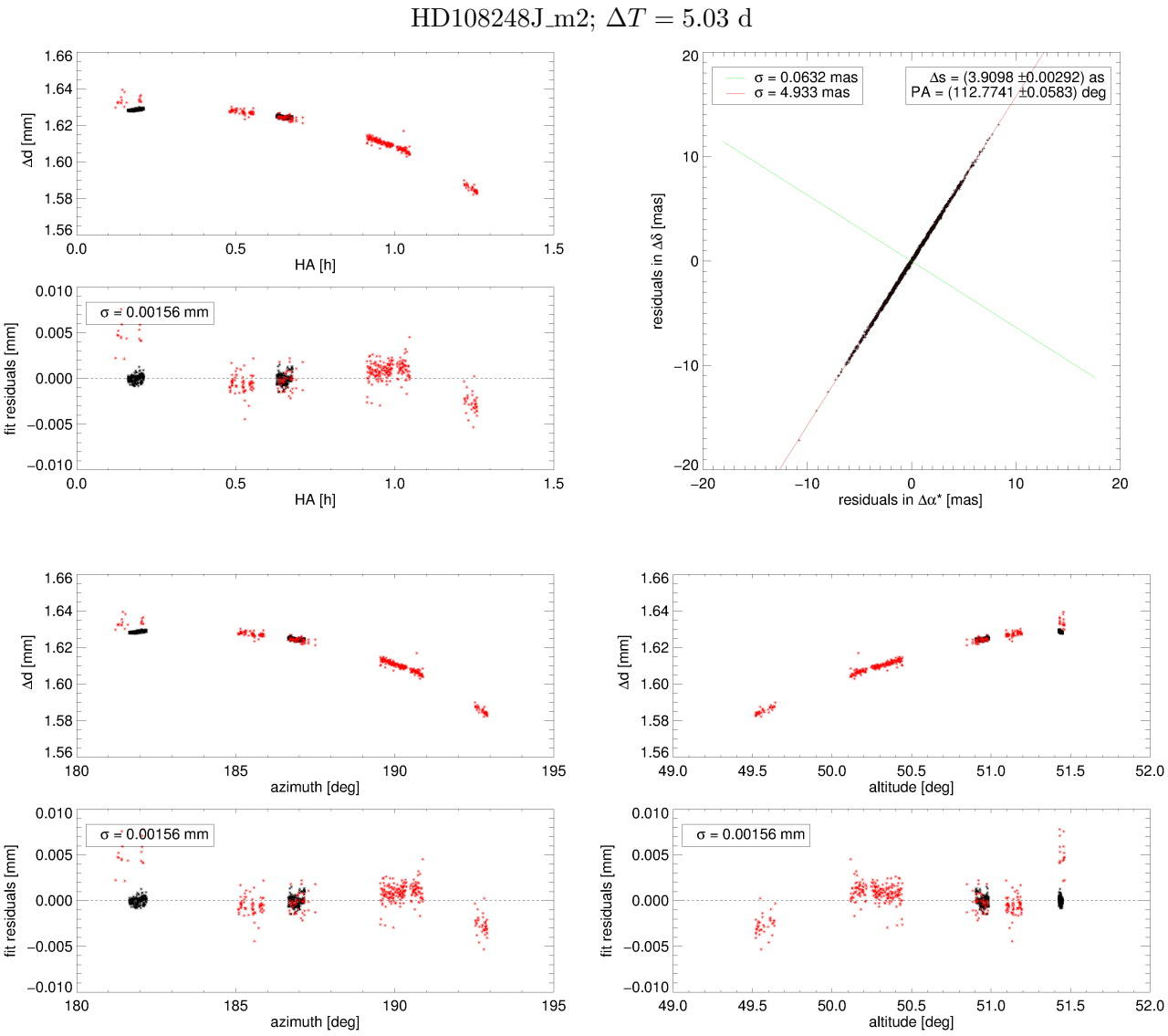


Figure D.41. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD108248J_m2, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

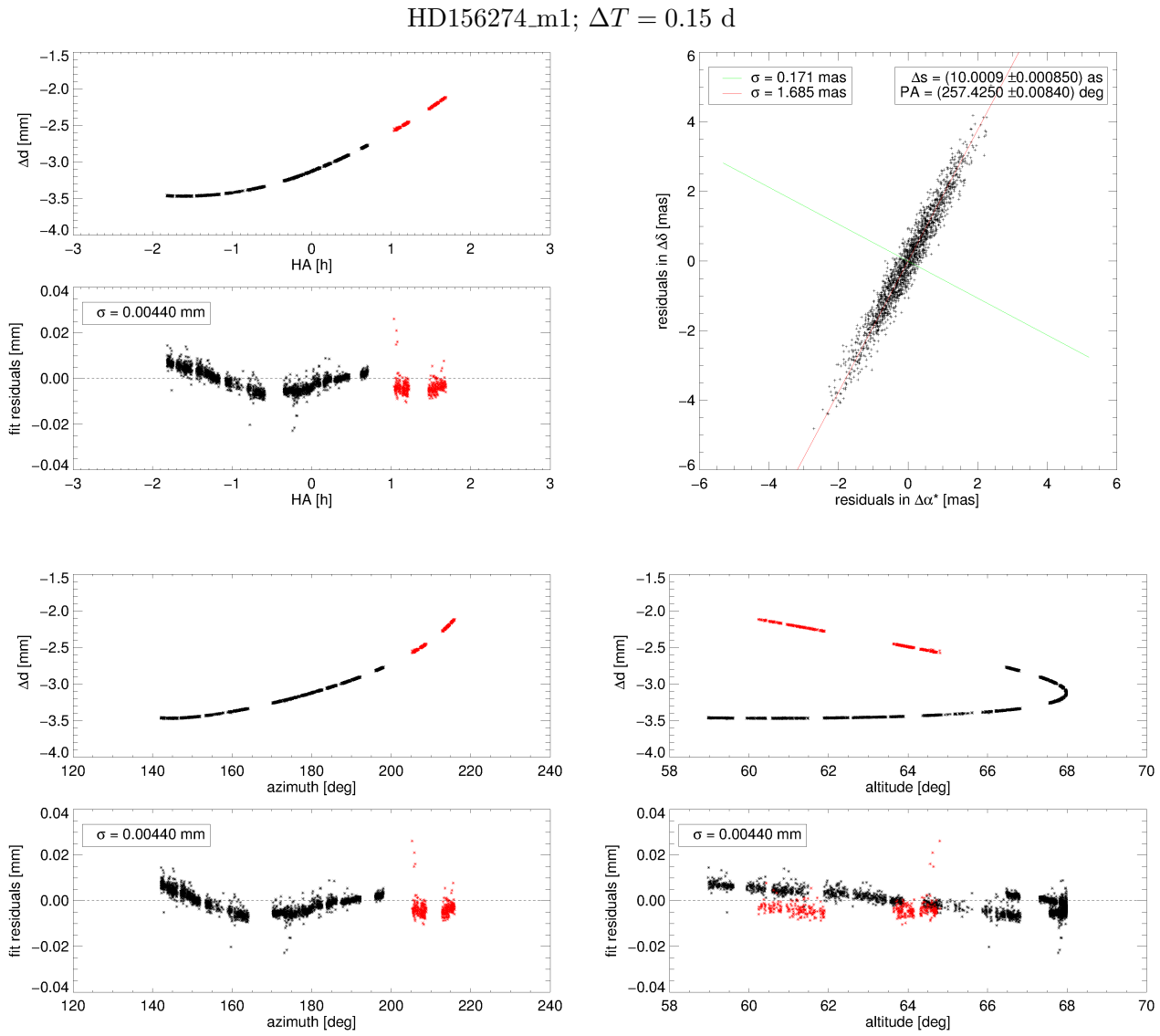


Figure D.42. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD156274_m1, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

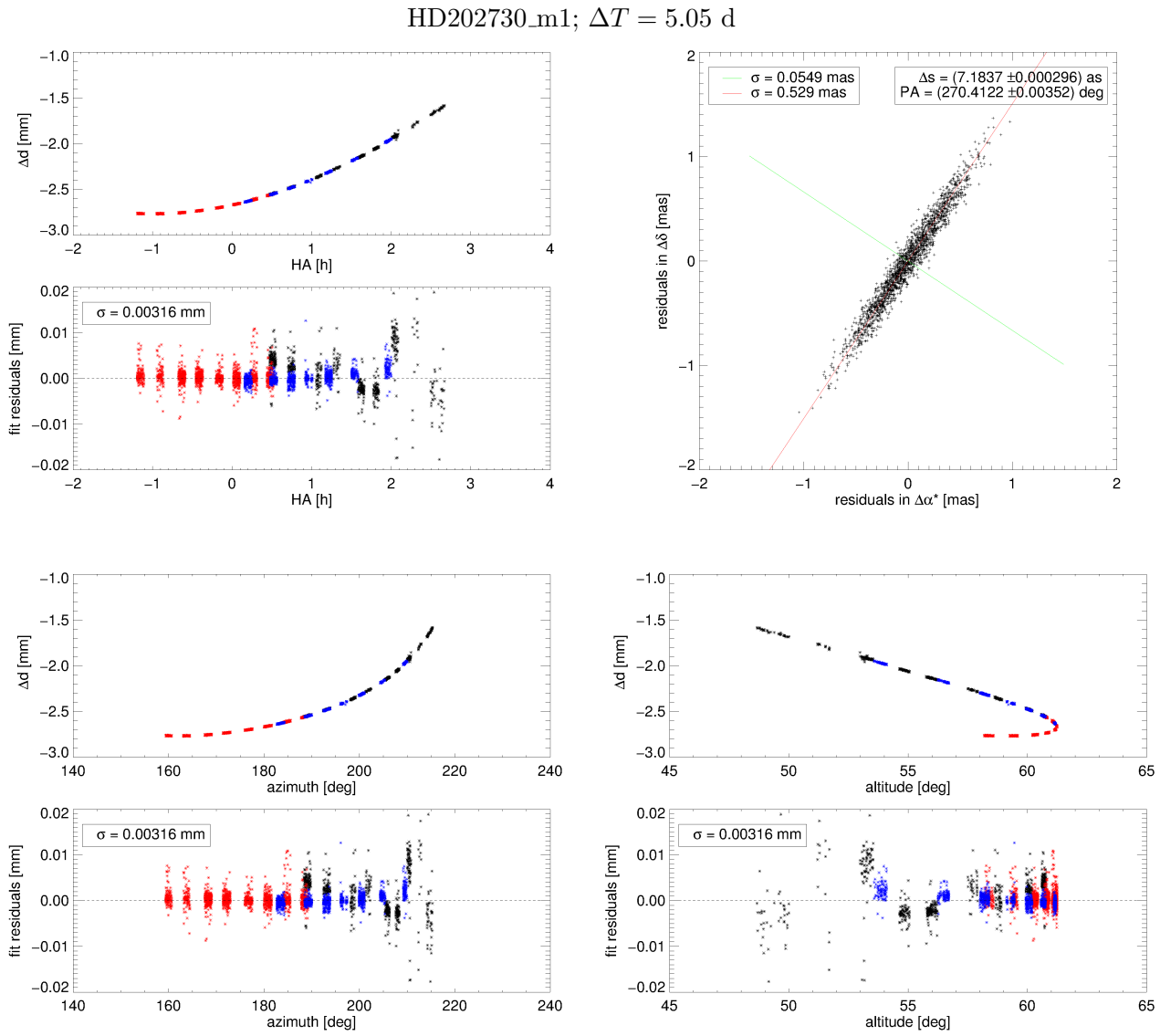


Figure D.43. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD202730_m1, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

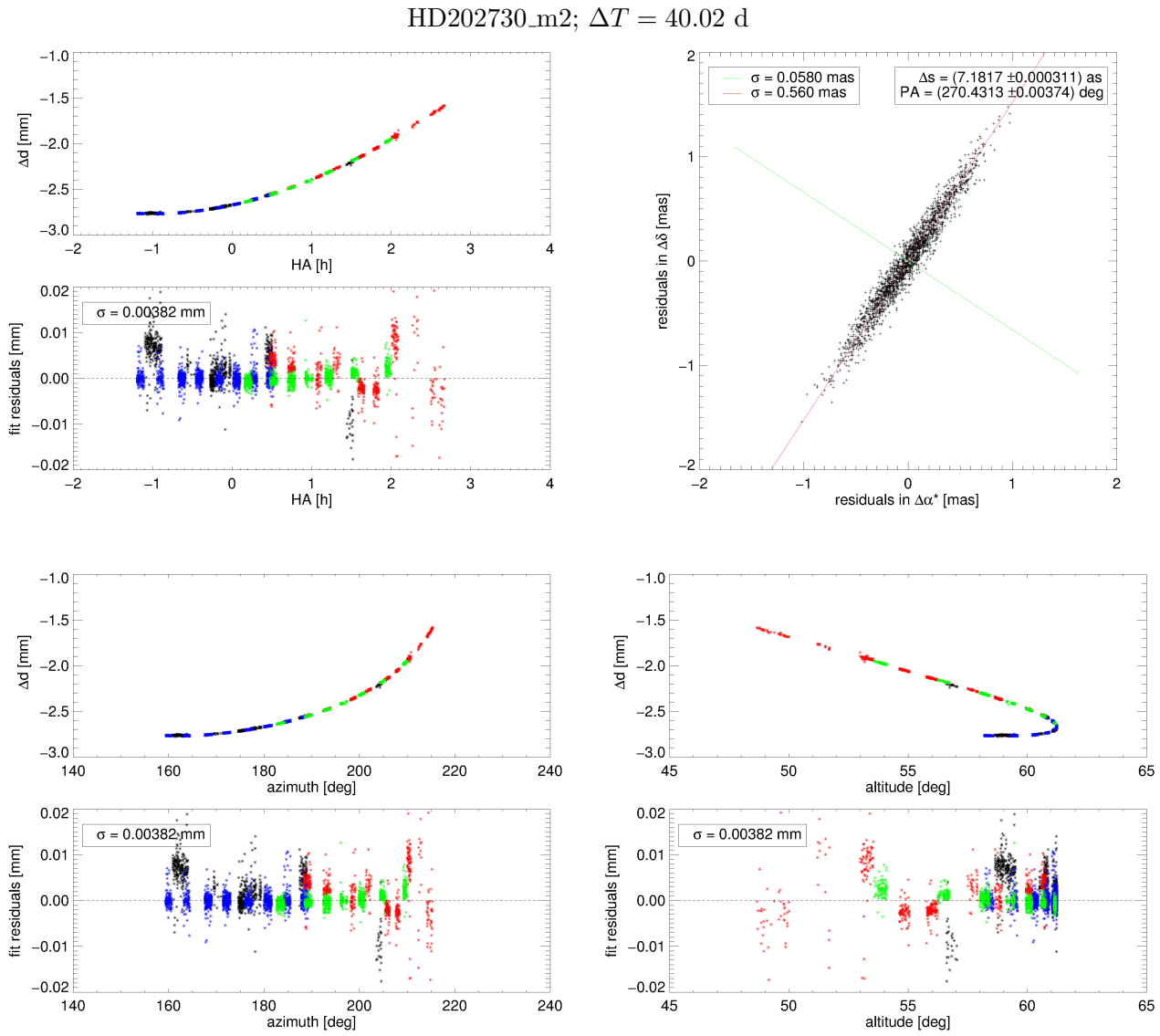


Figure D.44. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD202730_m2, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

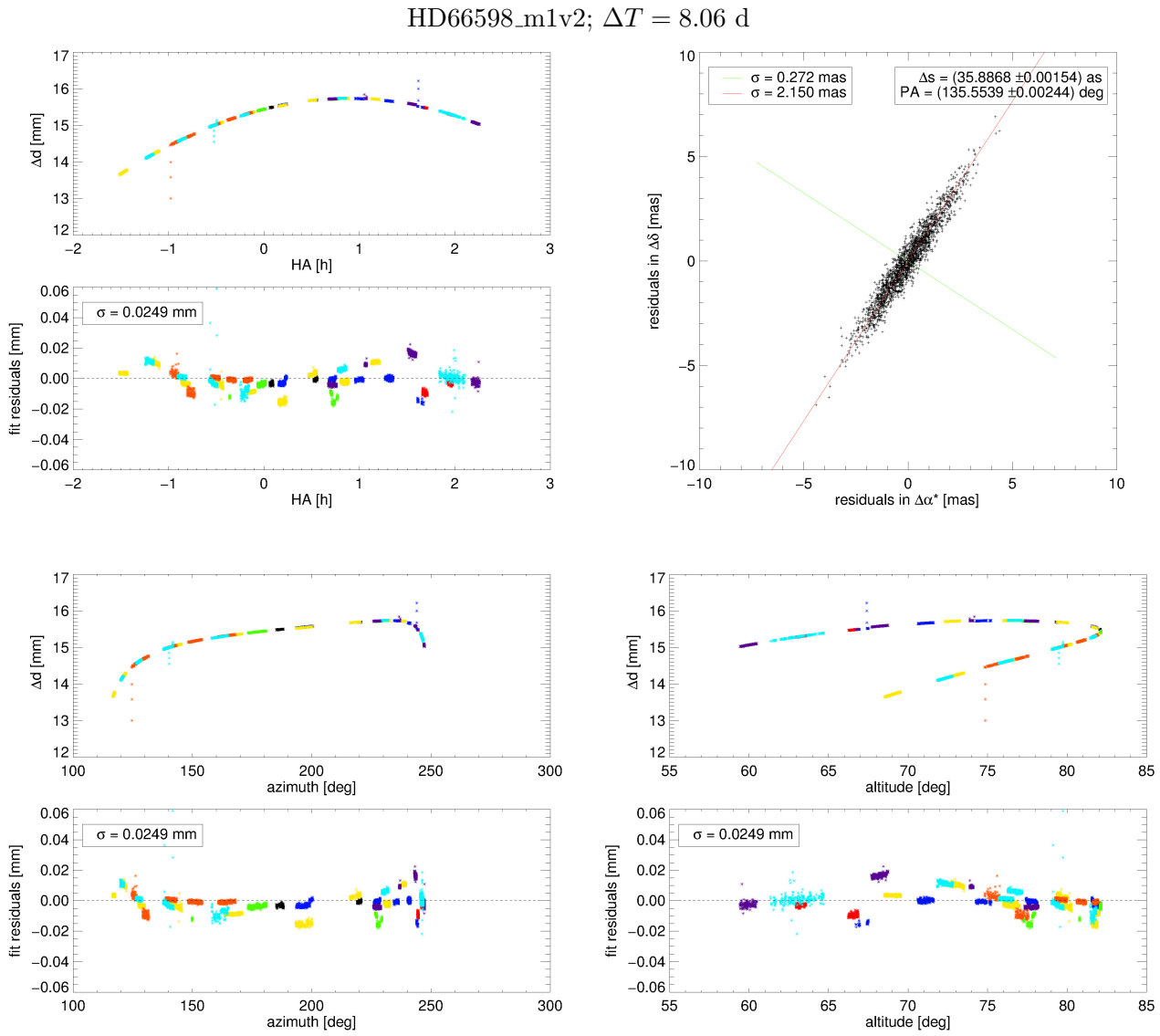


Figure D.45. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD66598_m1v2, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

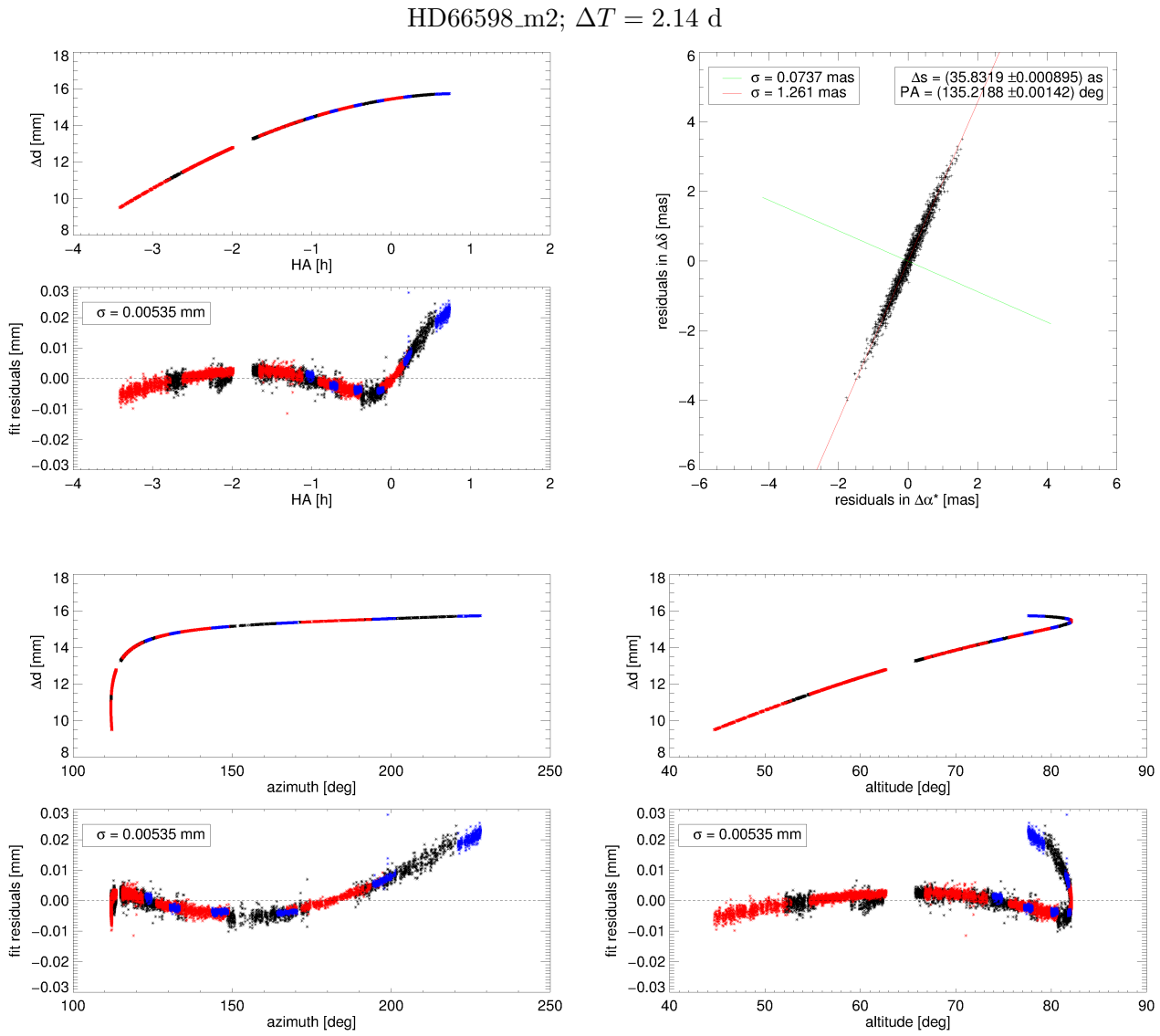


Figure D.46. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD66598_m2, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

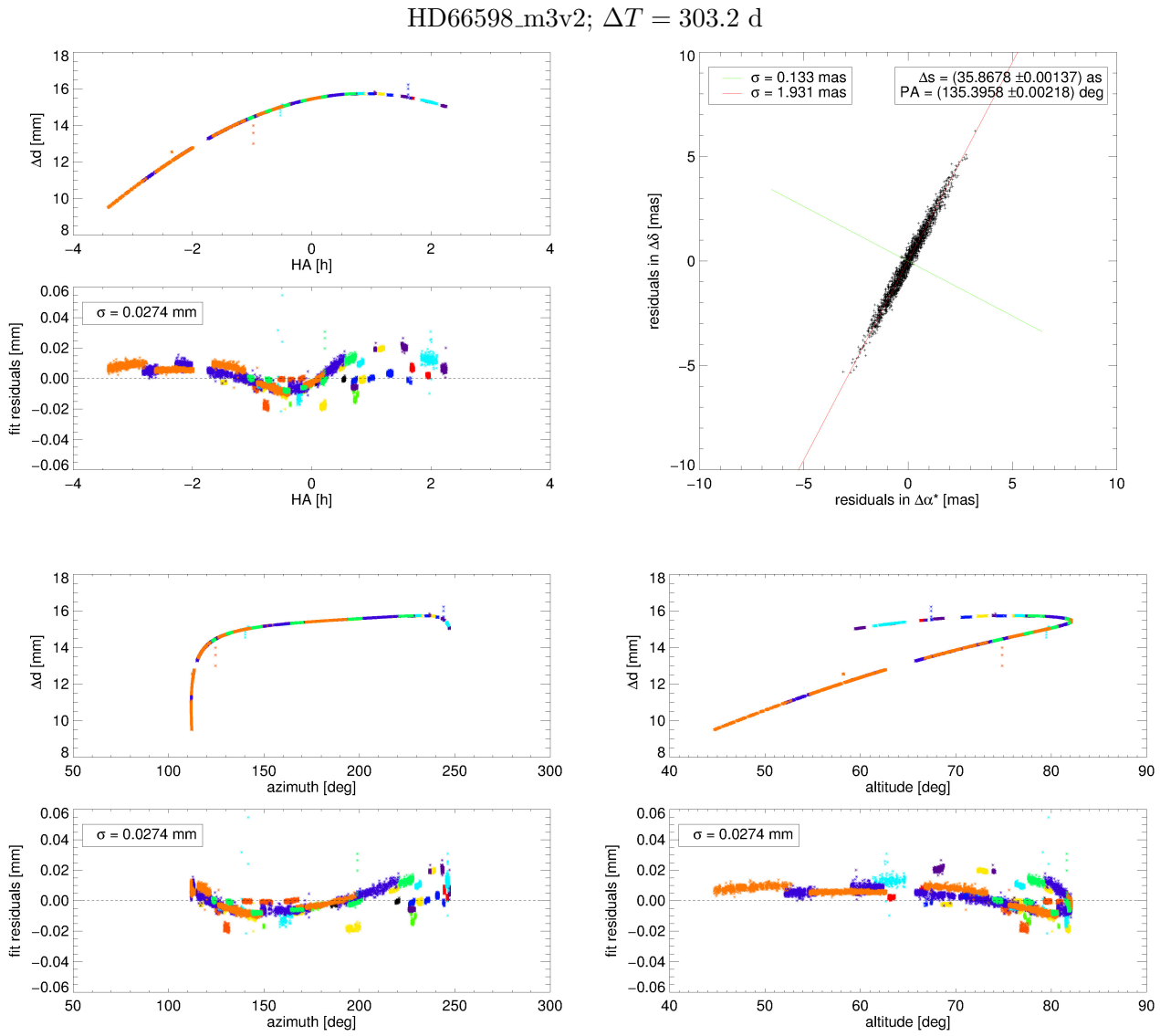


Figure D.47. – Corrected differential delays Δd and residuals of the multi epoch astrometric fit to the observation sequence HD66598_m3v2, as well as the distribution of the 2,000 fits to the resampled data sets from the bootstrap analysis.

Bibliography

[Abu06] Abuter, R. et al. 2006, *VLTI-PRIMA fringe tracking testbed*, Advances in Stellar Interferometry. Edited by Monnier, John D.; Schöller, Markus; Danchi, William C.. Proceedings of the SPIE, Volume 6268, pp. 62682Y

[Afs05] Afshordi, N.; Mukhopadhyay, B.; Narayan, R. 2005, *Bypass to Turbulence in Hydrodynamic Accretion: Lagrangian Analysis of Energy Growth*, The Astrophysical Journal, Volume 629, Issue 1, pp. 373-382

[And1982] Anders, E.; Ebihara, M. 1982, *Solar-system abundances of the elements*, Geochimica et Cosmochimica Acta, vol. 46, Nov. 1982, p. 2363-2380

[App1989] Appenzeller, I.; Mundt, R. 1989, *T Tauri stars*, Astronomy and Astrophysics Review (ISSN 0935-4956), vol. 1, Nov. 1989, p. 291-334

[Arm05] Armitage, P. J.; Rice, W. K. M. 2005, *Planetary migration*, eprint arXiv:astro-ph/0507492

[Arm07] Armitage, P. J. 2007, *Lecture notes on the formation and early evolution of planetary systems*, eprint arXiv:astro-ph/0701485

[Bac1996] Bachiller, R. 1996, *Bipolar Molecular Outflows From Young Stars and Protostars*, Annu. Rev. Astron. Astrophys. 1996, 34:111-154

[Bai10] Bailes, M. 2010, *The art of precision pulsar timing*, Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis, Proceedings of the International Astronomical Union, IAU Symposium, Volume 261, p. 212-217

[Bak05] Baker, J., Bizarro, M. et al. 2005, *Early planetesimal melting from an age of* 4.5662Gyr *for differentiated meteorites*, Nature 436: 1127-1131

[Bal1991] Balbus, S. A.; Hawley, J. F. 1991, *A powerful local shear instability in weakly magnetized disks. I - Linear analysis. II - Nonlinear evolution*, Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 376, July 20, 1991, p. 214-233

[Bal1998] Balbus, S. A.; Hawley, J. F. 1998, *Instability, turbulence, and enhanced transport in accretion disks*, Rev. Mod. Phys. 70, 1-53

[Bal06] Balbus, S. A.; Hawley, J. F. 2006, *An Exact, Three-dimensional, Time-dependent Wave Solution in Local Keplerian Flow*, The Astrophysical Journal, Volume 652, Issue 2, pp. 1020-1027

[Bal1983] Bally, J.; Lada, C. J. 1983, *The high-velocity molecular flows near young stellar objects*, Astrophysical Journal, Part 1, vol. 265, Feb. 15, 1983, p. 824-847

[Bas1983] Bastian, U. et al. 2005, *The definition of T Tauri and Herbig Ae/Be stars*, Astronomy and Astrophysics (ISSN 0004-6361), vol. 126, no. 2, Oct. 1983, p. 438, 439

[Bea06] Beaulieu, J.-P. et al. 2006, *Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing*, Nature, Volume 439, Issue 7075, pp. 437-440

[Ben10] Beanedict, G. F. et al. 2010, *The Mass of HD 38529c from Hubble Space Telescope Astrometry and High-precision Radial Velocities*, The Astronomical Journal, Volume 139, Issue 5, pp. 1844-1856

[Ben1999] Benz, W.; Asphaug, E. 1999, *Catastrophic Disruptions Revisited*, Icarus, Volume 142, Issue Icarus, pp. 5-20

[Ben11] Benz, W. 2011, *Planet Formation*, Lecture at IMPRS Summer School: Characterizing Exoplanets - From Formation to Atmospheres, Heidelberg 2011

[Ber1996] Bertin, E.; Arnouts, S. 1996, *SExtractor: Software for source extraction.*, Astronomy and Astrophysics Supplement, v.117, p.393-404

[Bir1993] Birch, K. P.; Downs, M. J. 1993, *An Updated Edlén Equation for the Refractive Index of Air*, Metrologia 1993, 30, 155-162

[Blö1998] Blöcker, T.; Holweger, H. et al. 1998, *Lithium Depletion in the Sun: A Study of Mixing Based on Hydrodynamical Simulations*, Space Science Reviews, v. 85, Issue 1/2, p. 105-112

[Bor12] Borucki, W. J. et al. 2012, *Kepler-22b: A 2.4 Earth-radius Planet in the Habitable Zone of a Sun-like Star*, The Astrophysical Journal, Volume 745, Issue 2, article id. 120

[Bra08] Brauer, F.; Dullemond, C. P.; Henning, Th. et al. 2008, *Coagulation, fragmentation and radial motion of solid particles in protoplanetary disks*, Astronomy and Astrophysics, Volume 480, Issue 3, 2008, pp.859-877

[Bro2000] Brown, T. M. et al. 2000, *Seeking the Atmospheric Transmission Spectrum of HD209458b*, American Astronomical Society, 197th AAS Meeting, #11.05; Bulletin of the American Astronomical Society, Vol. 32, p.1417

[Bus95] Buscher, A. F. et al. 1995, *Interferometric seeing measurements on Mt. Wilson: power spectra and outer scales*, Applied Optics LP, vol. 34, Issue 6, p.1081

[Car06] Carballido, A. et al. 2006, *Mid-plane sedimentation of large solid bodies in turbulent protoplanetary discs*, Monthly Notices of the Royal Astronomical Society, Volume 373, Issue 4, pp. 1633-1640

[Cha1939] Chandrasekhar, S. 1939, *An introduction to the study of stellar structure*, Univ. of Chicago Press, Chicago, 1939

[Cha04] Chauvin, G. et al. 2004, *A giant planet candidate near a young brown dwarf. Direct VLT/NACO observations using IR wavefront sensing*, Astronomy and Astrophysics, v.425, p.L29-L32

[Cla09] Clarke, C. J. 2009, *Pseudo-viscous modelling of self-gravitating discs and the formation of low mass ratio binaries*, Monthly Notices of the Royal Astronomical Society, Volume 396, Issue 2, pp. 1066-1074

[Col1987] Colavita, M. M.; Shao, M.; Staelin, D. H. 1987, *Atmospheric phase measurements with the Mark III stellar interferometer*, Applied Optics (ISSN 0003-6935), vol. 26, Oct. 1, 1987, p. 4106-4112.

[Col09] Colavita, M. M. 2009, *Adverse effects in dual-feed interferometry*, New Astronomy Reviews, Volume 53, Issue 11-12, p. 344-352

[Dav1995] Davis J. et al. 1995, *Atmospheric path variations for baselines up to 80m measured with the Sydney University Stellar Interferometer*, Monthly Notices of the Royal Astronomical Society, Volume 273, Issue 3, pp. L53-L58

[Daw12] Dawson, R. I.; Murray-Clay, R. 2012, *Neptune's wild days: constraints from the eccentricity distribution of the classical Kuiper Belt*, eprint arXiv:1202.6060

[dePat01] de Pater, I.; Lissauer, J. J. 2001, *Planetary Sciences*, Cambridge University Press

[Dee1998] Deeg, H. 1998, *Photometric Detection of Extrasolar Planets by the Transit-Method*, Brown dwarfs and extrasolar planets, Proceedings of a Workshop held in Puerto de la Cruz, Tenerife, Spain, 17-21 March 1997, ASP Conference Series #134, edited by Rafael Rebolo; Eduardo L. Martin; Maria Rosa Zapatero Osorio, p. 216

[Del08] Delplancke, F. 2008, *The PRIMA facility phase-referenced imaging and micro-arcsecond astrometry*, New Astronomy Reviews, Volume 52, Issue 2-5, p. 199-207

[Des07] Desch, S. J. 2007, *Mass Distribution and Planet Formation in the Solar Nebula*, The Astrophysical Journal, vol. 671, Issue 1, pp. 878-893

[Dev1999] Devillard, N. 1999, *Infrared Jitter Imaging Data Reduction: Algorithms and Implementation*, Astronomical Data Analysis Software and Systems VIII, ASP Conference Series, Vol. 172. Ed. David M. Mehringer, Raymond L. Plante, and Douglas A. Roberts. ISBN: 1-886733-94-5 (1999), p. 333.

[DiF03] Di Folco, E. et al. 2003, *Atmospheric and internal turbulence measured on the Very Large Telescope Interferometer with VINCI*, Interferometry for Optical Astronomy II. Edited by Wesley A. Traub . Proceedings of the SPIE, Volume 4838, pp. 1115-1126

[DiL08] Di Lieto, N. et al. 2008, *Adaptive vibration cancellation on large telescopes for stellar interferometry*, Optical and Infrared Interferometry. Edited by Schöller, Markus et al., Proc. of SPIE, Volume 7013, pp. 70130H

[Doy08] Doyle, L. R. 2008, *Overview of extrasolar planet detection methods*, Extrasolar Planets: XVI Canary Islands Winter School of Astrophysics. Edited by Hans Deeg, Juan Antonio Belmonte and Antonio Aparicio. ISBN-13 978-0-521-86808-2 (HB). Published by Cambridge University Press, Cambridge, UK, 2008, p.1

[Dull05] Dullemond, C. P.; Dominik, C. 2005, *Dust coagulation in protoplanetary disks: A rapid depletion of small grains*, Astronomy and Astrophysics, Volume 434, Issue 3, May II 2005, pp.971-986

[Dur07] Durisen, R. H. et al. 2005, *Gravitational Instabilities in Gaseous Protoplanetary Disks and Implications for Giant Planet Formation*, Protostars and Planets V, B. Reipurth, D. Jewitt, and K. Keil (eds.), University of Arizona Press, Tucson, 951 pp., 2007., p.607-622

[Eli08] Elias, N. et al. 2008, *Astrometric Data-Reduction Software Design Report*, ESO Document VLT-TRE-AOS-15752-0003, Issue 3.0

[ESA1997] ESA, 1997, *The Hipparcos and Tycho Catalogues*, ESA SP-1200

[ESO09] Haguenauer, P.; Kaufer, A. 2009, *VLT Observatory requirements for visitor instruments at the VLTI*, Doc. Nr.: VLT-ICD-ESO-15000-4824, Issue: 1

[Fai12] Faigler, S. et al. 2012, *Seven New Binaries Discovered in the Kepler Light Curves through the BEER Method Confirmed by Radial-velocity Observations*, The Astrophysical Journal, Volume 746, Issue 2, article id. 185

[Fei07] Feigelson, E. et al. 2007, *X-Ray Properties of Young Stars and Stellar Clusters*, Protostars and Planets V, B. Reipurth, D. Jewitt, and K. Keil (eds.), University of Arizona Press, Tucson, 951 pp., 2007., p.313-328

[For03] Ford, E. B.; Rasio, F. A.; Yu, K. 2003, *Dynamical Instabilities in Extrasolar Planetary Systems*, Scientific Frontiers in Research on Extrasolar Planets, ASP Conference Series, Vol 294, Edited by Drake Deming and Sara Seager. (San Francisco: ASP) ISBN: 1-58381-141-9, 2003, pp. 181-188

[Gam01] Gammie, C. F. 2001, *Nonlinear Outcome of Gravitational Instability in Cooling, Gaseous Disks*, The Astrophysical Journal, Volume 553, Issue 1, pp. 174-183

[Gat1973] Gatewood, G.; Eichhorn, H. 1973, *An unsuccessful search for a planetary companion of Barnard's star BD* +43561, Astronomical Journal, Vol. 78, p. 769 - 776

[Gla1993] Gladman, B. 1993, *Dynamics of systems of two close planets*, Icarus, vol. 106, p. 247

[Gla00] Glassgold, A. E.; Feigelson, E. D.; Montmerle, T. 2000, *Effects of Energetic Radiation in Young Stellar Objects*, Protostars and Planets IV (Book - Tucson: University of Arizona Press; eds Mannings, V., Boss, A.P., Russell, S. S.), p. 429

[Gli11] Glindemann, A. 2011, *Principles of Stellar Interferometry*, Astronomy and Astrophysics Library,

ISBN 978-3-642-15027-2. Springer-Verlag Berlin Heidelberg, 2011

[Gol1979] Goldreich, P.; Tremaine, S. 1979, *The excitation of density waves at the Lindblad and corotation resonances by an external potential*, Astrophysical Journal, Part 1, vol. 233, no. 3, Nov. 1, 1979, p. 857-871

[Gott05] Gott, J. R. III et al. 2005, *A Map of the Universe*, The Astrophysical Journal, Volume 624, Issue 2, pp. 463-484

[Gra1987] Grant, E. 1987, *Celestial Orbs in the Latin Middle Ages*, Isis, Vol. 78, No. 2 (Jun., 1987), pp. 152-173

[Gui05] Guillot, T. 2005, *The Interiors of Giant Planets: Models and Outstanding Questions*, Annual Review of Earth and Planetary Sciences, vol. 33, p.493-530

[Gul1998] Gullbring, E. et al. 1998, *Disk Accretion Rates for T Tauri Stars*, Astrophysical Journal v.492, p.323

[Gütt10] Güttler, C. et al. 2010, *The outcome of protoplanetary dust growth: pebbles, boulders, or planetesimals?. I. Mapping the zoo of laboratory collision experiments*, Astronomy and Astrophysics, Volume 513, id.A56

[Har1998] Hartmann, L. et al. 1998, *Accretion and the Evolutiion of T Tauri Disks*, Astrophysical Journal v.495, p.385

[Hay1981a] Hayashi, C. 1981, *Formation of the planets*, Fundamental problems in the theory of stellar evolution; Proceedings of the Symposium, Kyoto, Japan, July 22-25, 1980. (A82-34012 16-90) Dordrecht, D. Reidel Publishing Co., 1981, p. 113-126

[Hay1981b] Hayashi, C. 1981, *Structure of the Solar Nebula, Growth and Decay of Magnetic Fields and Effects of Magnetic and Turbulent Viscosities on the Nebula*, Progress of Theoretical Physics Supplement, No. 70, pp. 35-53

[Hei1988] Heintz, W. D. 1988, *The binary star 70 Ophiuchi revisited*, Journal of the Royal Astronomical Society of Canada (ISSN 0035-872X), vol. 82, June 1988, p. 140-145

[Her1973] Hershey, J. L. 1973, *Astrometric analysis of the field of AC* +656955 *from plates taken with the Sproul 24-inch refractor*, Astronomical Journal, Vol. 78, p. 421 - 425

[Huy1698] Huygens, C. 1698, *The Celestial Worlds Discovered*, London: F. Cass. 1968

[Iko2000] Ikoma, M.; Nakazawa, K.; Emori, H. 2000, *Formation of Giant Planets: Dependences on Core Accretion Rate and Grain Opacity*, The Astrophysical Journal, Volume 537, Issue 2, pp. 1013-1025

[Isr09] Israelian, G.; Delgado, M. et al. 2009, *Enhanced lithium depletion in Sun-like stars with orbiting planets*, Nature, Volume 462, Issue 7270, pp. 189-191

[Jac1855] Jacob, W. S. 1855, *On certain Anomalies presented by the Binary Star 70 Ophiuchi*, Monthly Notices of the Royal Astronomical Society, Vol. 15, p.228

[Jea1902] Jeans, J. H. 1902, *The Stability of a Spherical Nebula*, Philosophical Transactions of the Royal Society of London. Series A 199: 1-53

[Ji2006] Ji, H. et al. 2006, *Hydrodynamic turbulence cannot transport angular momentum effectively in astrophysical disks*, Nature, Volume 444, Issue 7117, pp. 343-346

[Joh05] Johansen, A.; Klahr, H. 2005, *Dust Diffusion in Protoplanetary Disks by Magnetorotational Turbulence*, The Astrophysical Journal, Volume 634, Issue 2, pp. 1353-1371

[Joh09] Johansen, A. et al. 2009, *Particle Clumping and Planetesimal Formation Depend Strongly on Metallicity*, The Astrophysical Journal Letters, Volume 704, Issue 2, pp. L75-L79

[Joh11] Johansen, A. 2011, *Protoplanetary discs: Setting the conditions for planet formation*, Lecture at

IMPRS Summer School: Characterizing Exoplanets - From Formation to Atmospheres, Heidelberg 2011

[Joy1945] Joy, A. H. 1945, *T Tauri Variable Stars*, Astrophysical Journal, vol. 102, p.168

[Jur08] Jurić, M.; Tremaine, S. 2008, *Dynamical Origin of Extrasolar Planet Eccentricity Distribution*, The Astrophysical Journal, Volume 686, Issue 1, pp. 603-620

[Kal08] Kalas, P. et al. 2008, *Optical Images of an Exosolar Planet 25 Light-Years from Earth*, Science, Volume 322, Issue 5906, pp. 1345- (2008)

[Kam10] Kaminski, A. 2010, *PRIMA - Phase Offsets between Quadratures ABCD*, Internal PRIMA-DAWG report, 02.10.2010

[Ken08] Kennedy, G. M.; Kenyon, S. J. 2008, *Planet Formation around Stars of Various Masses: The Snow Line and the Frequency of Giant Planets*, The Astrophysical Journal, 673:502-512

[Kli03] Klioner, S. A. 2003, *A Practical Relativistic Model for Microarcsecond Astrometry in Space*, The Astronomical Journal, Volume 125, Issue 3, pp. 1580-1597

[Kok1998] Kokubo, E.; Ida, S. 1998, *Oligarchic Growth of Protoplanets*, Icarus, Volume 131, Issue 1, pp. 171-178

[Kon03] Konacki, M.; Wolszczan, A. 2003, *Masses and Orbital Inclinations of Planets in the PSR B1257+12 System*, The Astrophysical Journal, Volume 591, Issue 2, pp. L147-L150

[Kov04] Kovalevsky, J.; Seidelmann, P. K. 2004, *Fundamentals of Astrometry*, Fundamentals of Astrometry, by Jean Kovalevsky and P. Kenneth Seidelmann, pp. 420. ISBN 0521642167. Cambridge, UK: Cambridge University Press, July 2004

[Kür1999] Kürster, M. et al. 1999 *Precise radial velocities of Proxima Centauri. Strong constraints on a substellar companion*, Astronomy and Astrophysics, v.344, p.L5-L8

[Lan03] Lane, B. F. 2003, *High-precision infra-red stellar interferometry*, Thesis (PhD). CALIFORNIA INSTITUTE OF TECHNOLOGY, Source DAI-B 64/02, p. 763

[Lau12] Launhardt, R. et al. 2012, *ESPRI science goals and targets - revisited*, , ESPRI document Nr.: VLT-PLA-AOS-15759-0006, Issue: 0.1, December 20th, 2012

[Lee2000] Lee, M. H. 2000, *On the Validity of the Coagulation Equation and the Nature of Runaway Growth*, Icarus, Volume 143, pp. 74-86

[Lev1996] Lévêque, S. et al. 1996, *Longitudinal Dispersion Compensation for the Very Large Telescope Interferometer*, Astrophysics and Space Science, Volume 239, Issue 2, pp.305-314

[Lev03] Lévêque, S. et al. 2003, *Toward nanometer accuracy laser metrology for phase-referenced interferometry with the VLTI*, Interferometry for Optical Astronomy II. Edited by Wesley A. Traub . Proceedings of the SPIE, Volume 4838, pp. 983-994

[Lin1980] Lindegren, L. 1980, *Atmospheric limitations of narrow-field optical astrometry*, Astronomy and Astrophysics, vol. 89, no. 1-2, Sept. 1980, p. 41-47

[Lis1993] Lissauer, J. J. 1993, *Planet formation*, Annual review of astronomy and astrophysics. Vol. 31 (A94-12726 02-90), p. 129-174

[Liz1987] Lizano, S.; Shu, F. H. 1995, *Heating of molecular cloud cores*, IAU, Reunion Regional Latinoamericana de Astronomia, 5th, Merida, Mexico, Oct. 6-10, 1986; Revista Mexicana de Astronomia y Astrofisica (ISSN 0185-1101), vol. 14, May 1987, p. 587-594

[Lyn1974] Lynden-Bell, D.; Pringle, J. E. 1974, *The evolution of viscous discs and the origin of the nebular variables*, Monthly Notices of the Royal Astronomical Society, Vol. Vol. 168, p. 603 - 637

[Lyr10] Lyra, W. et al. 2010, *Orbital Migration of Low-mass Planets in Evolutionary Radiative Models:*

Avoiding Catastrophic Infall, The Astrophysical Journal Letters, Volume 715, Issue 2, pp. L68-L73

[May1995] Mayor, M.; Queloz, D. 1995, *A Jupiter-mass companion to a solar-type star*, Nature 378: 355-359

[May11] Mayor, M. et al. 2011, *The HARPS search for southern extra-solar planets XXXIV. Occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets*, eprint arXiv:1109.2497

[Mar08] Marois, Ch. et al. 2008, *Direct Imaging of Multiple Planets Orbiting the Star HR 8799*, Science, Volume 322, Issue 5906, pp. 1348- (2008)

[Mas01] Mason, B. D. et al. 2013, *The Washington Visual Double Star Catalog (Mason+ 2001-2013)*, VizieR On-line Data Catalog: B/wds. Originally published in: The Astronomical Journal, Volume 122, Issue 6, pp. 3466-3471

[Mat06] Mathar, R. J. 2006, *Refractive Index of Humid Air in the Infrared: Model Fits*, J. Opt. A: Pure Appl. Opt. 9 (5) (2007) 470-476

[Maz10] Mazeh, T.; Faigler, S. 2010, *Detection of the ellipsoidal and the relativistic beaming effects in the CoRoT-3 lightcurve*, Astronomy and Astrophysics, Volume 521, id.L59

[Mil08] Militzer, B.; Hubbard, W. B. et al. 2008, *A Massive Core in Jupiter Predicted from First-Principles Simulations*, The Astrophysical Journal, 688:45-48

[Mich1921] Michelson, A. A.; Pease, F. G. 1921, *Measurement of the diameter of alpha Orionis with the interferometer.*, Astrophysical Journal, 53, 249-259

[Mor05] Morbidelli, A. 2005, *Origin and Dynamical Evolution of Comets and their Reservoirs*, eprint arXiv:astro-ph/0512256

[Mor09] Mordasini, Ch. et al. 2009, *Extrasolar planet population synthesis. II. Statistical comparison with observations*, Astronomy and Astrophysics, Volume 501, Issue 3, 2009, pp.1161-1184

[Mor11] Mordasini, Ch. 2011, *Application of recent results on the orbital migration of low mass planets: convergence zones*, The Astrophysics of Planetary Systems: Formation, Structure, and Dynamical Evolution, Proceedings of the International Astronomical Union, IAU Symposium, Volume 276, p. 72-75

[Mos10] Moshovitz, N. et al. 2010, *Formation of Jupiter using opacities based on detailed grain physics*, Icarus, Volume 209, Issue 2, p. 616-624

[Mül10] Müller, A. et al. 2010, *First results using PRIMA FSU as a fringe tracker for MIDI*, Optical and Infrared Interferometry II. Edited by Danchi, William C. et al. Proceedings of the SPIE, Volume 7734, pp. 773420-773420-15

[Mut11] Muterspaugh, M. W. et al. 2011, *The Planet in the HR 7162 Binary System Discovered by PHASES Astrometry*, American Astronomical Society, AAS Meeting #217, #429.09; Bulletin of the American Astronomical Society, Vol. 43, 2011

[Mun1983] Mundt, R.; Fried, J. W. 1983, *Jets from young stars*, Astrophysical Journal, Part 2 - Letters to the Editor (ISSN 0004-637X), vol. 274, Nov. 15, 1983, p. L83-L86

[Nae05] Naef, D. et al. 2005, *The ELODIE planet search: synthetic view of the survey and its global detection threshold*, Proceedings of the 13th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun, held 5-9 July, 2004 in Hamburg, Germany. Edited by F. Favata, G.A.J. Hussain, and B. Battrick. ESA SP-560, European Space Agency, 2005., p.833

[Nij08] Nijenhuis, J. et al. 2008, *Simultaneous observation of two stars using the PRIMA Star Separator*, Optical and Infrared Interferometry. Edited by Schöller, Markus et al., Proc. of SPIE, Volume 7013, 70133

[Nol04] Nolting, W. 2004, *Grundkurs Theoretische Physik 3*, 7. Aufl., ISBN: 3-540-20509-8, Springer

[Opp09] Oppenheimer, B. R.; Hinkley, S. 2009, *High-Contrast Observations in Optical and Infrared Astronomy*, Annual Review of Astronomy & Astrophysics, vol. 47, Issue 1, pp.253-289

[Par06] Paardekooper, S.-J.; Mellema, G. 2006, *Halting type I planet migration in non-isothermal disks*, Astronomy and Astrophysics, Volume 459, Issue 1, November III 2006, pp.L17-L20

[Pep08] Pepe, F. et al. 2008, *The ESPRI project: Differential Delay Lines for PRIMA*, Optical and Infrared Interferometry. Edited by Schöller, Markus et al., Proc. of SPIE, Volume 7013, 70130

[Pre1992] Press, W. H. et al. 1992, *Numerical Recipes in C. The Art of Scientific Computing*, Second Edition, Cambridge Univ. Press, 1992

[Pri1981] Pringle, J. E. 1981, *Accretion discs in astrophysics*, Annual review of astronomy and astrophysics. Volume 19. (A82-11551 02-90) Palo Alto, CA, Annual Reviews, Inc., 1981, p. 137-162

[Pud07] Pudritz, R. E. 2007, *Disk Winds, Jets, and Outflows: Theoretical and Computational Foundations*, Protostars and Planets V, B. Reipurth, D. Jewitt, and K. Keil (eds.), University of Arizona Press, Tucson, 951 pp., 2007., p.277-294

[Qui99] Quirrenbach, A. 1999, *Observing Through the Turbulent Atmosphere*, Course Notes from the 1999 michelson Summer School, August 15-19, 1999, JPL Publication 00-009 07/00

[Ref06] Reffert, S. et al. 2006, *The PRIMA astrometric planet search: goals and prospects*, Advances in Stellar Interferometry. Edited by Monnier, John D.; Schöller, Markus; Danchi, William C.. Proceedings of the SPIE, Volume 6268, pp. 626846

[Ref11] Reffert, S.; Quirrenbach, A. 2011, *Mass constraints on substellar companion candidates from the re-reduced Hipparcos intermediate astrometric data: nine confirmed planets and two confirmed brown dwarfs*, Astronomy & Astrophysics, Volume 527, id.A140

[Sah09] Sahlmann, J. et al. 2009, *The PRIMA fringe sensor unit*, Astronomy and Astrophysics, Volume 507, Issue 3, 2009, pp.1739-1757

[Sah10] Sahlmann, J. et al. 2010, *First results from fringe tracking with the PRIMA fringe sensor unit*, Optical and Infrared Interferometry II. Edited by Danchi, William C.; Delplancke, Françoise; Rajagopal, Jayadev K. Proceedings of the SPIE, Volume 7734, pp. 773422-773422-12

[Sah11] Sahlmann, J. et al. 2011, *PRIMA astrometry data modelling and analysis of auxiliary data (Comm14)*, ESO Document VLT-TRE-ESO-15770-5337, Issue 1.0

[Sah13] Sahlmann, J. et al. 2013, *The ESPRI project: astrometric exoplanet search with PRIMA I. Instrument description and performance of first light observations*, Astronomy & Astrophysics, Volume 551, id.A52, 31 pp

[Sal11] Salmeron, R. 2011, *Formation of stars and planets: the role of magnetic fields*, Astrophysics and Space Science, Volume 335, Issue 2, pp.353-371

[Sch11] Schneider, J. et al. 2011, *Defining and cataloging exoplanets: the exoplanet.eu database*, Astronomy and Astrophysics, Vol. 532, id.A79

[Schu06] Schuler, N. 2006, *Frequency-comb stabilized laser sources for absolute distance metrology at the Very Large Telescope Interferometer*, PhD thesis, Université Louis Pasteur (2006)

[Set07] Setiawan, J. et al. 2007, *PRIMA Astrometry Target and Reference Selection Users Guide*, ESPRI document Nr.: VLT-MAN-AOS-15759-0003, Issue: 2, March 09th, 2007

[Sha1973] Shakura, N. I.; Sunyaev, R. A. 1973, *Black holes in binary systems. Observational appearance.*, Astronomy and Astrophysics, Vol. 24, p. 337 - 355

[Sha1977] Shao, M.; Staelin, D. H. 1977, *Long-baseline optical interferometer for astrometry*, Optical Society of America, Journal, vol. 67, Jan. 1977, p. 81-86

[Sha1992] Shao, M.; Colavita, M. M. 1992, *Potential of long-baseline infrared interferometry for narrow-angle astrometry*, Astronomy and Astrophysics, vol. 262, no. 1, p. 353-358.

[Shu1977] Shu, F. H. 1977, *Self-similar collapse of isothermal spheres and star formation*, Astrophysical Journal, Part 1, vol. 214, June 1, 1977, p. 488-497

[Shu1987] Shu, F. H.; Lizano, S.; Adams, F. C. 1987, *Star formation in molecular cloud cores*, Star forming regions; Proceedings of the Symposium, Tokyo, Japan, Nov. 11-15, 1985 (A87-45601 20-90). Dordrecht, D. Reidel Publishing Co., 1987, p.417-433

[Skr06] Skrutskie, M. F. et al. 2006, *The Two Micron All Sky Survey (2MASS)*, The Astronomical Journal, Volume 131, Issue 2, pp. 1163-1183

[Sta1988] Stahler, S. W. 1988, *Deuterium and the stellar birthline*, Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 332, Sept. 15, 1988, p. 804-825

[Sum11] Sumi, T. et al. 2011, *Unbound or Distant Planetary Mass Population Detected by Gravitational Microlensing*, Nature, Volume 473, Issue 7347, p. 349-352

[Tak1996] Takeuchi, T.; Miyama, S. M.; Lin, D. N. C.. 1996, *Gap Formation in Protoplanetary Disks*, Astrophysical Journal v.460, p.832

[Tan02] Tanaka, H.; Takeuchi, T.; Ward, W. R. 2002, *Three-Dimensional Interaction between a Planet and an Isothermal Gaseous Disk. I. Corotation and Lindblad Torques and Planet Migration*, The Astrophysical Journal, Volume 565, Issue 2, pp. 1257-1274

[Tang1990] Tango, W. J. 1990, *Dispersion in stellar interferometry*, Applied Optics, vol. 29, Feb. 1, 1990, p. 516-521

[Ted02] Tedesco, E. F.; Desert, F.-X. 2002, *The Infrared Space Observatory Deep Asteroid Search*, The Astrophysical Journal, 123:2070-2082

[Ter1984] Terebey, S.; Shu, F. H.; Cassen, P. 1984, *The collapse of the cores of slowly rotating isothermal clouds*, Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 286, Nov. 15, 1984, p. 529-551

[Too1964] Toomre, A. 1964, *On the gravitational stability of a disk of stars*, Astrophysical Journal, vol. 139, p. 1217-1238

[Tra1999] Traub, W. A. 1999, *Beam Combination and Fringe Measurement*, Course Notes from the 1999 michelson Summer School, August 15-19, 1999, JPL Publication 00-009 07/00

[Tubb08] Tubbs, R. et al. 2008, *Astrometric Error Budget*, ESO Document VLT-TRE-AOS-15753-0001, Issue 2.1

[Udr11] Udry, S. 2011, *Detections and characterisation of extra-solar planets*, Lecture at IMPRS Summer School: Characterizing Exoplanets - From Formation to Atmospheres, Heidelberg 2011

[Ume1981] Umebayashi, T.; Nakano, T. 1981, *Fluxes of Energetic Particles and the Ionization Rate in Very Dense Interstellar Clouds*, Publications of the Astronomical Society of Japan, Vol. 33, P. 617, 1981

[Uns1999] Unsöld, A.; Baschek, B. 1999, *Der Neue Kosmos*, 6. Aufl., Springer

[vBelle08] van Belle, G. T. et al. 2008, *The VLTI PRIMA Facility*, The ESO Messenger, 134:6-11

[vdKam63] van de Kamp, P. 1963, *Astrometric Study of Barnard's Star*, Astronomical Journal, Vol. 68, p. 295

[vLee07] van Leeuwen, F. 2007, *Hipparcos, the New Reduction of the Raw Data*, Astrophysics and Space Science Library, Vol. 350, by F. van Leeuwen

[Ver04] Veras, D.; Armitage, P. J. 2004, *Outward migration of extrasolar planets to large orbital radii*, Monthly Notices of the Royal Astronomical Society, Volume 347, Issue 2, pp. 613-624

[Wall08] Wallace, P. T. 2008, *Interferometer Optical-Path-Length Analysis System*, Version 7.0, September 2008

[Wam1997] Wambsganss, J. 1997, *Discovering Galactic planets by gravitational microlensing: magnification patterns and light curves*, Monthly Notices of the Royal Astronomical Society, Volume 284, Issue 1, pp. 172-188

[War1997] Ward, W. R. 1997, *Protoplanet Migration by Nebula Tides*, Icarus, Volume 126, Issue 2, pp. 261-281

[Wei1977] Weidenschilling, S. J. 1977, *The distribution of mass in the planetary system and solar nebula*, Astrophysics and Space Science, vol. 51, no. 1, Sept. 1977, p. 153-158

[Wei1977b] Weidenschilling, S. J. 1977, *Aerodynamics of solid bodies in the solar nebula*, Monthly Notices of the Royal Astronomical Society, vol. 180, July 1977, p. 57-70

[Wei09] Weigert, A.; Wendker, H. J.; Wisotzki, L. 2009, *Astronomie und Astrophysik*, 5. Aufl., Wiley-VCH Verlag

[Wie02] Wielen, R. et al. 2002, *Sixth Catalogue of Fundamental Stars (FK6) (Wielen+ 2000)*, VizieR On-line Data Catalog: I/264

[Wol1992] Wolszczan, A.; Frail, D. A. 1992, *A planetary system around the millisecond pulsar PSR1257+12*, Nature 355: 145-147

[Wya1975] Wyant, J. C. 1975, *Use of an ac heterodyne lateral shear interferometer with real-time wavefront correction systems*, Applied Optics, vol. 14, Nov. 1975, p. 2622-2626

[You05] Youdin, A. N.; Goodman, J. 2005, *Streaming Instabilities in Protoplanetary Disks*, The Astrophysical Journal, Volume 620, Issue 1, pp. 459-469

[Zap2000] Zapatero Osorio, M. R. et al. 2000, *Discovery of Young, Isolated Planetary Mass Objects in the σ Orionis Star Cluster*, Science, Volume 290, Issue 5489, pp. 103-107

Web Ressources:

[ESA13] ESA, 2013 *Astrophysics Missions: Gaia*, http://www.rssd.esa.int/index.php?project=GAIA&page=index

[ESO03] ESO, *VLTI Station Position Technical Data*, http://www.eso.org/observing/etc/doc/vlti/baseline/vltistations.html

[Hub1999] Hubblesite.org News Release Archive 1999, *Vast Stellar Disks Set Stage for Planet Birth in New Hubble Images*, News Release Number: STScI-1999-05, http://hubblesite.org/newscenter/archive/releases/1999/05/

[IERS13] IERS, *Long term earth orientation data, EOP 08 C04 (IAU2000A)*, http://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html

[MPIA10] MPIA Heidelberg/R. Launhardt, *PRIMA hardware development at MPIA*, http://www.mpia.de/PRIMA-DDL/hardware.php

[NASA08] NASA/JPL-Caltech/R.Hurt (SSC/Caltech), http://www.spitzer.caltech.edu/images/1925-ssc2008-10b-A-Roadmap-to-the-Milky-Way-Annotated

[NStED] NASA/IPAC/NExScI Star and Exoplanet Database, http://nsted.ipac.caltech.edu/index.html

Further Acknowledgements:

- This work is based on observations made with the European Southern Observatory telescopes, obtained from the ESO/ST-ECF Science Archive Facility.
- This research has made use of the Washington Double Star Catalog maintained at the U.S. Naval Observatory.
- This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.
- This research has made use of the SIMBAD database, operated at the CDS, France.

Acknowledgment

Most of all, I would like to thank my supervisor Andreas Quirrenbach for the unique opportunity to participate in this project and for his assistance over the past years.

Furthermore, I would like to thank Joachim Krautter for agreeing on acting as the second thesis referee and additionally Cornelis Dullemond and Ingeborg Levin for accepting to be jury members at the defense.

Special thanks go to my co-supervisors Sabine Reffert and Ralf Launhardt, who provided support whenever and wherever needed.

I would like to acknowledge ESO in general for providing the commissioning data the thesis is based on and in particular the ESO department in Santiago, Chile for offering me a six months studentship, during which I had the chance to participate in PRIMA's commissioning runs and encountered a fruitful and motivational working environment.

I am also grateful to all the members of the ESPRI consortium and the DAWG for the overall encouraging collaboration and also for their trust, which allowed me to conduct ESPRI observations at the La Silla facilities. In particular, I would like to thank Johannes Sahlmann for willingly sharing his expertise on the subject on several occasions.

In general, I thank the entire LSW exoplanet group for the educational time I have spent here, while I would like to single out Ingo Stilz and Rainer Köhler, who have been particularly supporting and who helped me a lot regarding anything concerning the data reduction pipeline and interferometry.

I owe my gratitude to my wife for her infinite believe in me and for her understanding during the whole process of working on the thesis.

At last I would like my parents for their incredible support in any respect and for teaching me that anything is achievable.