

Detection of Faint Stars Near Gaia Objects

Shahin Jafarzadeh

Lund Observatory
Lund University



2009-EXA34

Degree project of 60 higher education credits (for a degree of Master)
April 2009

Lund Observatory
Box 43
SE-221 00 Lund
Sweden

Detection of Faint Stars Near Gaia Objects

A dissertation presented

by

Shahin Jafarzadeh

to

Lund Observatory (Institutionen för astronomi)

in partial fulfillment of the requirements

for the degree of

Master of Science

in the subject of

Astronomy and Astrophysics

Thesis supervisor:

Professor Lennart Lindegren

Thesis examiners:

Dr. Ingemar Lundström

Dr. Jennifer Simmerer

Lund University

Lund, Sweden

April 2009

©2009 - Shahin Jafarzadeh

All rights reserved.

Abstract

The upcoming ESA space astrometry mission Gaia, planned to be launched in early 2012, has been designed to scan the whole sky (flux limited, corresponding to magnitude 20 in the Gaia passband) during a five-year observational phase. It will measure the astrometric parameters (positions, parallaxes and proper motions) of more than one billion stars, complemented by multi-epoch spectrometric and radial-velocity observations. The astrometric accuracy ranges from 10 μ arcsec for bright objects to 200 μ arcsec for the faintest. The science goals cover a wide range of problems in modern astronomy, with a focus on Galactic structure, kinematics and evolution, and stellar astrophysics.

To achieve the science goals, it is important that possible disturbing effects on the astrometric measurements are well understood. This is necessary both in order to design the instrument so that these effects can be minimized, and to design appropriate algorithms of data analysis for detecting and correcting remaining effects. For example, the measurements of a target star may be disturbed by faint stars, apparently close to the target and therefore not normally detected, because they fall below the flux limit or are hidden behind the main target's flux. Because of the high accuracy aimed for, even very faint stars could produce significant errors if they remain undetected and therefore not taken into account in the data analysis.

In this work I investigate the detectability of faint stars by means of computer simulations of a two-dimensional image reconstruction technique. Using the actual scanning law proposed for the Gaia, one-dimensional (1D) images are first simulated for the whole five-year mission and for several random positions on the sky. Then, the two-dimensional (2D) images are produced by stacking the 1D images. The quality of the produced 2D images is very important for the successful detection of very faint stars. Because the pixel geometry of the images is rather peculiar, several well-known stacking algorithms (e.g., shift-and-add, sub-sampling and drizzle methods) are examined. Based on these techniques, an improved stacking algorithm is proposed.

The resulting 2D images are strongly affected by artifacts from the original pixel geometry and the non-uniform distribution of scanning directions, and must be corrected by deconvolution. The CLEAN method, which is a well-known deconvolution technique in radio astronomy, presents itself as the best solution for detecting point sources in the simulated Gaia images. Using this iterative method, the positions as well as the G magnitudes of the point-source objects down to a computed threshold level are precisely obtained.

Using Monte Carlo simulations, it is investigated how the detection probability depends on the apparent (angular) separation of the two stars. This probability is a function of the signal-to-noise ratio, which depends on the object's magnitude and the number of scans across the object according to the Gaia scanning law. Minimum and maximum separation limits are therefore computed for different conditions of magnitude difference and number of scans for some random positions on the sky. I also investigate how the detection probability depends on the positional angle of the faint star with respect to the target, and the precision of the resulting position and magnitude of the faint star.

Finally, an easy-to-use graphical user interface (GUI) called Gaia FSDtool is described. This is a standalone executable MATLAB code written by the author, which allows the user to choose the input parameters and immediately see the simulated images, stacking and detecting processes, as well as the final results on the detected objects. All of the output of this GUI is computed by the same MATLAB functions that were used for the investigations described in this work.

Contents

Title Page	i
Abstract	iii
Table of Contents	v
Acknowledgments	vii
1 Introduction	1
2 The Gaia Mission	4
2.1 Astrometry with Gaia	5
2.2 The Gaia scanning law	6
2.3 Layout and structure of the CCDs	7
2.4 The photometric G band	9
2.5 Accuracy of Gaia observations	10
3 Gaia Imaging Simulations	11
3.1 The Line Spread Function	11
3.2 Sources of noise	12
3.2.1 Readout noise	12
3.2.2 Dark current	13
3.2.3 Background electrons	13
3.2.4 Photon noise	13
3.3 Simulation of 1D images	13
3.4 Image stacking: Producing two-dimensional images	15
3.4.1 Shift-and-add technique	15
3.4.2 Sub-sampling	16
3.4.3 Drizzle method	18
3.4.4 Drizzle method without shrinking	18
3.4.5 A new geometrical stacking method for increased quality	19
4 Detection of Faint Stars	22
4.1 Image de-biasing	23
4.2 Noise level and detection threshold	23
4.3 The CLEAN method	25
4.4 Gaussian convolution: Producing the clean map	26
4.5 Computing the positions	27

4.6	Computing the magnitudes	27
4.7	Disturbing stars detection	27
5	Statistical Studies and Results	29
5.1	Optimizing the CLEAN parameters	29
5.1.1	Selection of the loop gain	29
5.1.2	Selection of the size of the Gaussian clean beam	30
5.2	Detection limit for a single star	30
5.3	Probability of true detection	33
5.3.1	Variation of detection probability with separation	34
5.3.2	Variation of detection probability with magnitude difference	37
5.3.3	Variation of detection probability with the number of transits	37
5.3.4	Minimum separation for a true detection	37
5.3.5	Maximum magnitude difference for a true detection	37
5.4	Precision of the parameters of the disturbing star	40
6	Gaia FSDtool: an easy-to-use graphical user interface (GUI)	41
6.1	Developing the GUI in MATLAB	41
6.2	Implementation of Gaia FSDtool	42
6.3	Inputs options	42
6.3.1	Data selection	43
6.3.2	Initial characterizations	44
6.4	Output options	45
6.5	Running the Gaia FSDtool	46
6.6	Code availability and requirements	46
7	Discussion and Conclusions	48
	Bibliography	50

Acknowledgments

First and foremost, I am deeply indebted to my supervisor, Professor Lennart Lindegren, for proposing this research project, his great guidance and supports during the research as well as enormous helps to make it possible for me to prepare this thesis as a successful work. Not only his enthusiasm, inspiration, encouragement and great efforts to explain things clearly and simply at every step of this particular research project; but also his way of facing any kind of scientific problem in general, have been invaluable resources to me. He is truly an admirable scientist who I tried to learn the research attitude from; and it has really been a great honor for me to work with him.

I also wish to express my appreciation to Dr David Hobbs and Mr Berry Holl for providing essential inputs as well as valuable hints to this project and also Dr David Hobbs for providing me the access to the Gaia super computer at Lund Observatory, which was useful for some parts of the extensive statistical studies of this work.

In addition, my special thanks go to Dr Sofia Feltzing and all master students at Lund Observatory for their valuable feed-backs on my previous presentations. My office-mate, Nils Håkansson, as well as my best friends, Alexander Hansson, Berry Holl, Johan Sandel and Shahab Moghadas who all made my lunch and free times enjoyable.

Furthermore, I am grateful to all the staff members of Lund Observatory for their friendly environment and providing all the supports I have needed at this institute, and in particular, I would like to thank Professor Dainis Dravins for all his supports and encouragements as well as Dr Ingemar Lundström, the study director of Lund Observatory, who has always been kind and supportive since the first day I started my study at Lund University.

Finally, I have to say how grateful I am with my family: my uncles in Denmark, my lovely sister, Shadi, and lastly but most importantly, my parents who have always supported and encouraged me to pursue my interests.

“Tusen tack alla”

Chapter 1

Introduction

The role of astronomical observations

Observation plays an essential role in all natural sciences, and especially in astronomy. Electromagnetic radiation (visible light, X-rays, UV, ...) or particle radiation (neutrinos, cosmic rays) received from the distant objects are gathered by telescopes or antennae, analyzed and detected by various instruments (cameras, photometers, spectrometers, ...). The raw data obtained in this way are later processed to extract the meaningful scientific data which can be used to learn new facts about the observed targets.

One of the basic difficulties in this raw data processing is considering and if possible removing all the disturbing effects that influence the original data. Otherwise, in the absence of a good knowledge about all those disturbing effects, the data may lead to an inaccurate or even erroneous results. Therefore, inasmuch as one of the most important disturbances arise from the instrumental effects, all astronomical instruments including telescopes and detectors have been continually improved during the last centuries in all respects such as size, technology and quality. These improvements have mainly been done to increase the accuracy of the measurements and to reach a larger number of distant and faint objects. However, in addition to the power and sophistication of the instruments, there are several important and unavoidable factors which may affect the obtained data.

Recently, many advances in optics and telescope technology, computer control, digital imaging and data processing have improved the performance of astronomical instruments by reducing some of these disturbances. The history of science shows well how the constant development of instruments and technologies, aiming for ever higher sensitivity and accuracy, has expanded our knowledge about the universe.

Using space telescopes further eliminates some important disturbances such as atmospheric seeing and refraction, and instrument flexure due to the gravity. Space can also provide a more stable environment than is possible from the ground. These factors are especially important for high-accuracy astrometry, which explains why space observations are indispensable also in the visual region of the electromagnetic spectrum. Space astrometry was pioneered by the Hipparcos satellite, launched by the European Space Agency (ESA) in August 1989 and operational until August 1993. Its catalogue (ESA 1997; van Leeuwen 2008) provides astrometric data to about 1 milliarcsec accuracy for about 100,000 stars, and is widely used by the astronomical community in many fields of stellar and Galactic

research. In this thesis I will focus on Gaia, the upcoming successor of Hipparcos, which will observe 10,000 times as many stars to much higher accuracy. Understanding the disturbing effects in the Gaia observations is even more challenging than it was for Hipparcos. Here, the perturbations caused by faint, normally undetected stars will be considered.

Imaging with Gaia

Gaia is a major mission in the European Space Agency's science programme, with a planned launch in early 2012. The satellite will measure the parallaxes, positions, proper motions and photometric properties of more than 1 billion celestial objects with a very high astrometric accuracy. The typical parallax accuracy ranges from about $10 \mu\text{arcsec}$ for stars brighter than magnitude 12, to about $200 \mu\text{arcsec}$ for the faintest stars (magnitude 20). Practically all stars in the magnitude range 6 to 20 will be measured.

The very high accuracy of the Gaia measurements makes it necessary to understand very well all possible disturbances. As a potentially important case, the astrometric measurements of a given object may be disturbed by much fainter stars, apparently very close to the main target of interest and not normally detected, because they are hidden behind the main target's flux. Therefore, the flux profile of the main target is shifted due to the hidden fainter stars, causing some additional uncertainty in the positional measurements. By detecting the faint stars, and correcting the measurements for their fluxes, more accurate results can be achieved on the intended targets both in terms of its position (for astrometry) and intensity (for photometry). In addition, the detected faint objects could be very interesting by themselves, e.g., as faint components of binary stars, or features of the gravitational lensing (Einstein cross or ring) around quasars.

In this work, I investigate the detectability of such faint stars close to the main Gaia targets. The detection is complicated by the fact that Gaia is not optimized for imaging, but for one-dimensional position measurements in the scanning direction defined by the instantaneous spin axis of the satellite. The raw CCD data essentially represent one-dimensional (1D) images. Before the faint stars can be detected, it is necessary to reconstruct the two-dimensional (2D) images by means of some fairly sophisticated image processing technique (stacking and deconvolution).

Structure of the thesis

This master thesis consists naturally of four parts which are relatively independent:

1. An overview of Gaia mission (Chapter 2),
2. Simulation of 2D images from Gaia data (Chapter 3),
3. Detection of faint stars in the 2D images, including statistical studies and results (Chapters 4 and 5), and
4. A graphical user interface (GUI) for the codes used in Chapters 3, 4 and 5 (Chapter 6).

The first part gives some general background information about the Gaia satellite, including the instrumental characterizations which are later used for simulating the observations.

The second and third parts can be read independently and reflect two different projects where the result of the earlier part is used as the input data for the later one. Part two presents the different methods which were examined to find the best approach for simulating the 2D images with the best possible resolution. Part three forms the main body of the thesis and introduces the application of the CLEAN method for detecting the much fainter stars around the main targets of the Gaia observations. In addition, the results of different statistical studies using the CLEAN method are presented in Chapter 5.

Finally, the fourth part presents a user-friendly graphical interface for the image simulation and detection process.

Chapter 2

The Gaia Mission

Gaia is a European space mission which was originally proposed to the European Space Agency (ESA) by Lindegren et al. in 1993. This space satellite is planned to be sent in early 2012 into a controlled Lissajous orbit around the L2 Lagrangian point of the Earth–Sun system. This orbit will have several advantages, e.g., a thermo-mechanically stable environment as well as observation conditions uninterrupted by the Earth, Sun and Moon. Gaia will scan the whole sky (flux limited) for several times during a 5 year operational lifetime (this might be extended to 6 years) with the expected time for having the final catalogue around 2021. The limiting magnitude is $G = 20$ in the unfiltered Gaia passband (G).

Gaia is scientifically planned to do (i) very high-accuracy astrometric measurements (i.e., measuring positions, parallaxes and proper motions) of more than 1 billion stars, (ii) very low-resolution spectrophotometric measurements in the wavelength range of 330–1050 nm, and (iii) radial-velocity measurements of the brighter objects ($G < 16.5$ mag). Also, all point-like (unresolved) extragalactic and solar-system objects exceeding the flux limit will be observed. Therefore, the scientific case for Gaia covers an extremely wide range of topics in astrophysics, cosmology and solar-system research.

Gaia’s astrometric observations are based on the same principles as the Hipparcos satellite but with several great improvements, including (i) the much better resolution and photon collection power due to the much larger optics; (ii) the much higher quantum efficiency and wider spectral range of the CCDs, compared with the photomultipliers used in Hipparcos; and (iii) the parallel observations of tens of thousands objects in different CCDs. With all these improvements and several other changes, Gaia will observe some 10^4 times as many objects as Hipparcos with a 10–100 times higher accuracy.

The two observing fields will continuously move on the sky and cause the images to move across the telescope’s focal plane. Each point on the sky will be scanned several dozens of times and with different scanning angles during the five years’ mission. The number of scans (field-of-view transits) varies from around 40 to 230 times (80 times on average). Each field crossing essentially represents an one-dimensional scan across the object.

In the focal plane of the telescope (the effective focal length is equal to 35 m) there is a very large assembly of CCDs for the different purposes of detection, astrometric, photometric and spectrometric measurements. Using the detecting CCDs (Sky Mappers), the

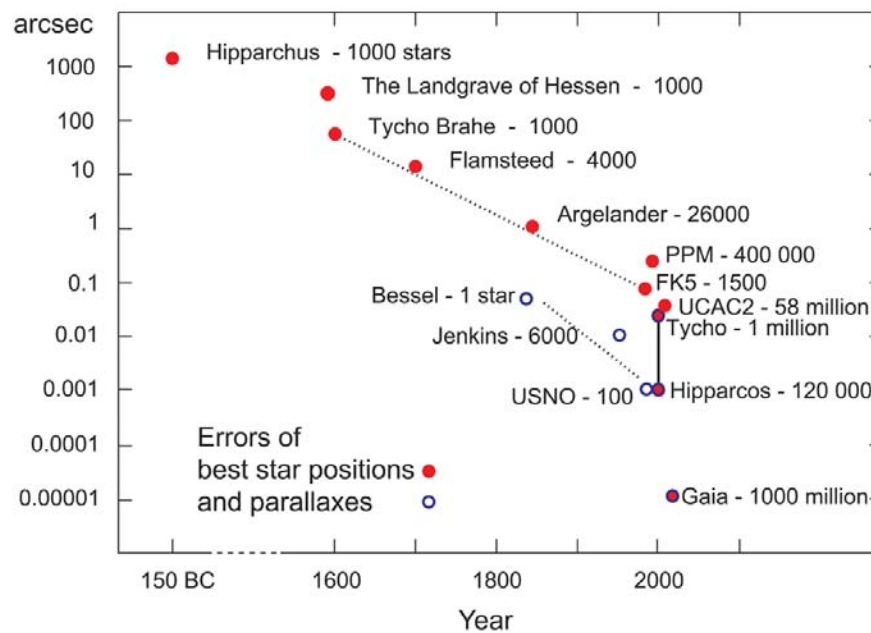


FIGURE 2.1: The accuracy comparison of the astrometric measurements from Hipparchus to Gaia satellite. (Figure courtesy: ESA)

onboard processing system detects all point-like objects down to the Gaia limiting magnitude $G = 20$ mag. On each field-of-view transit, the object will pass across a number of CCDs in the focal plane and the resulting photoelectron counts will be transmitted to the ground. The astrometric observations are done with 62 CCDs named AF (Astro Field).

2.1 Astrometry with Gaia

Measuring the positions of the point-source celestial objects and hence their distances and proper motions is the main objective of the astrometric measurements. This information gives essential knowledge about the kinematic and dynamics of the Galaxy and hence the origin and evolution of the Milky Way can be understood. Astrometry is one of the oldest investigating fields in astronomy, and the number of involved objects as well as the accuracy of the measurements have increased all the time since Hipparchus compiled the first astrometric list of about 1000 stars in the second century BC. The astrometric surveys and their accuracies have especially been much improved over the last few decades thanks to the Hipparcos mission (1989–1993) and ground-based surveys interpolating the Hipparcos catalogue to much fainter stars. Gaia will be the next major step in this process. Figure 2.1 illustrates the evolution of the accuracy astrometric measurements during the last 2200 years.

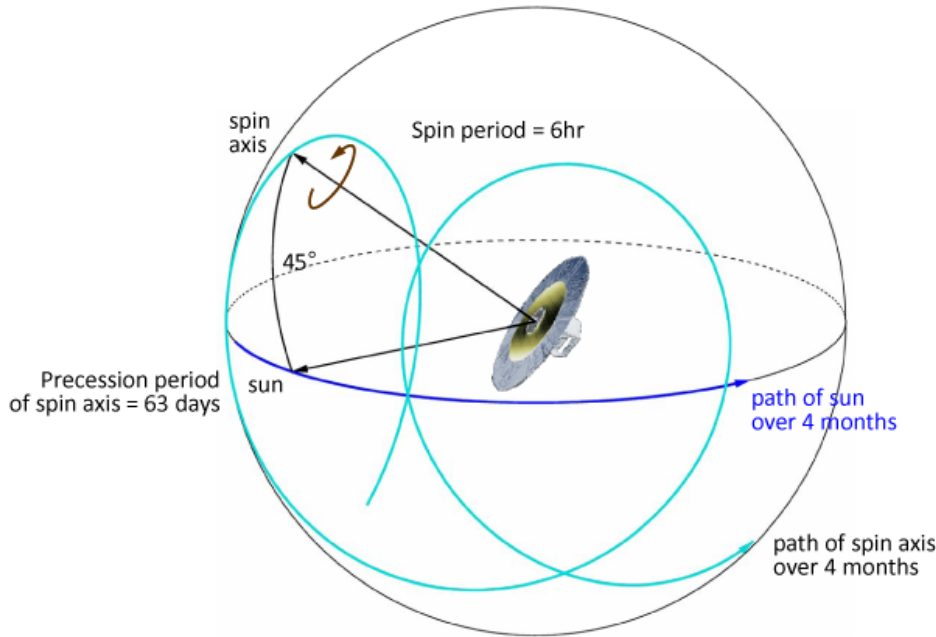


FIGURE 2.2: This schematic graph shows how the Gaia satellite will scan the whole sky during the 5-year mission according to a pre-defined scanning law. The satellite rotates around its spin axis with a period of 6 hours while the spin axis precesses around the solar direction with a fixed angle of 45° and a precessional period of 63 days. Therefore, the satellite also moves on the sky with the Sun which completes the ecliptic orbit in one year. (Figure courtesy: L. Lindegren)

2.2 The Gaia scanning law

The Gaia satellite will continuously scan the whole sky during the mission lifetime of 5 to 6 years and therefore the satellite orientation at any time directs the two observing telescopes to two specific regions on the sky (the preceding and following field of view). The design of this ‘scanning law’ is based on the previous Hipparcos mission and consists of three different rotational motions, as illustrated in Fig. 2.2.

As shown in the figure, the angle between the satellite’s spin axis and the Sun is constant and equal to 45° and the spin period is 6 hours. The combination of the three motions will cover the whole sky in six months and over a few years results in a rather smooth sky coverage with a geometry that allows us to separate the astrometric parameters for every observed object. The scanning will be done according to a precisely defined scanning law calculated by Lindegren (1998) and modified in 2002 because of some changes in the payload design. Therefore, the pointing of the telescopes at any time during the mission can be precisely predicted using a FORTRAN code provided by Lindegren (2001) based on the adopted scanning law. In this case, for any position in the sky, it is possible to compute the total number of transits during the 5 years along all the transit properties, e.g., the position angle of each scan across a given position on the sky.

The colour-coded sky map in Fig. 2.3 shows the number of field-of-view transits as a

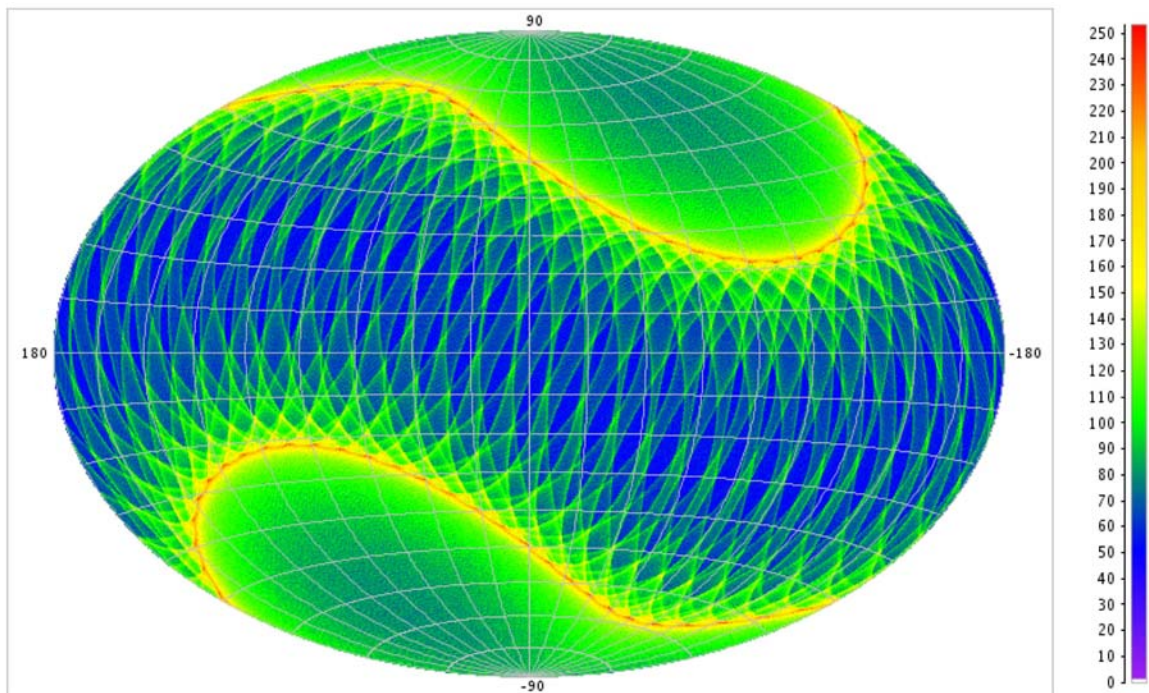


FIGURE 2.3: Distribution of the predicted number of field-of-view transits for the whole sky in ICRS (barycentric, equatorial) coordinates according to the pre-defined scanning law. This prediction was done by B. Holl using the AGISLab simulation software.

function of equatorial coordinates, according to a simulation by Holl (2008). As seen in this map, the number of transits varies approximately between 40 and 240, with a mean value of 80. The number is non-uniformly distributed in the sky, with the lowest numbers in the ecliptic region and the highest number at ecliptic latitudes $\pm 45^\circ$ (this is due to the Sun–spin axis angle of 45°). In addition, the transit angles (i.e., the position angles of the scans) also have a non-uniform distribution for any specific coordinates.

2.3 Layout and structure of the CCDs

Gaia scans the sky with its two different fields of view (separated by 106.5°) according to a pre-defined scanning law which was described in Sect. 2.2. During this scanning, the satellite is rotating all the time about an axis which is perpendicular to the observing directions. This causes star images to move across the field of view, so that it takes about 4.4 s to cross one CCD. The CCDs are operated in the TDI (time, delay and integration) mode, which means that the charge image is transported along the pixels at the same rate as the optical image. The integration time is therefore fixed to the time it takes the images to cross the CCD, i.e., 4.4 s. Then the produced photoelectrons in the CCD pixels are completely transported to the read-out register of the CCD when the star leaves the CCD field. In its very large CCD assembly for different observing purposes, Gaia has 14 sky mapper (SM) CCDs, 62 main astrometric fields (AF), 7 blue photometer (BP), 7 red photometer (RP), 12

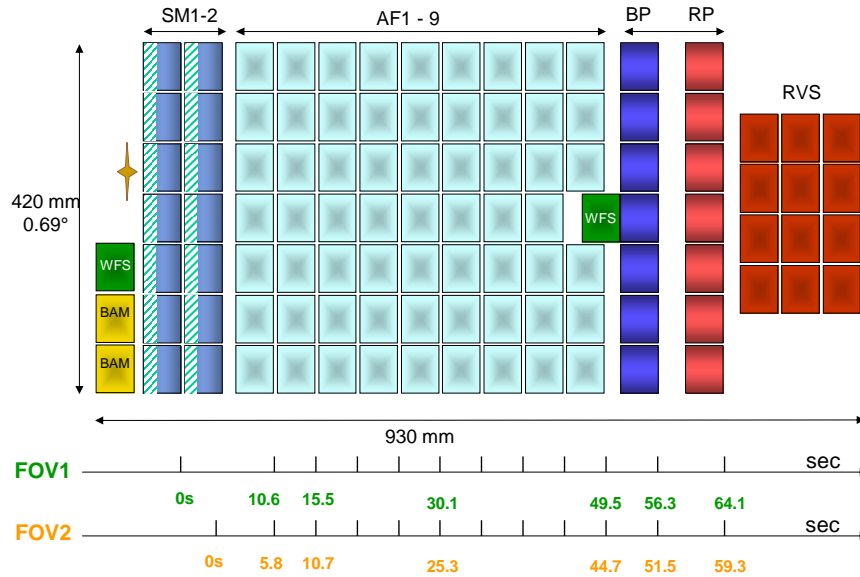


FIGURE 2.4: Layout of CCDs in the focal plane of the Gaia satellite. Stars from two different viewing directions pass all CCDs (i.e. the coloured rectangles) from left to right on the time scales shown at the bottom. (Figure courtesy: ESA – Gaia DPAC)

radial velocity spectrograph (RVS) CCDs, as well as 2 CCDs for the basic angle monitoring (BAM) system and 2 wave-front sensors (WFS) for initial mirror alignment and focusing (Fig. 2.4). The actual size of each CCD is 45 mm (along scan, AL) by 59 mm (across scan, AC). The pixel size is 10 μm (AL) by 30 μm (AC).

Among all these CCDs, only the AF CCDs are considered in the current work. These CCDs are not completely read out because of the data rate and read-out noise. Only a small window around the main target is considered. Before the stars enter the field of the AF CCDs, the SM CCDs are completely read out, the on-board software detects the images and, using a numerical centroiding process, the centres of the detected images (down to the Gaia magnitude $G = 20$ mag) are precisely computed. This information is then used to predict the times when the star images leave each AF CCD, and the pixels in a window centred on each image is read out. The size of the window depends on the magnitude of the star and also on which CCD is considered (AF1–9). For this study it is assumed that the window consists of 12 along-scan (AL) and 8 across-scan (AC) pixels, see Fig. 2.5. Moreover, the all generated photoelectrons in the 8 AC pixels are combined at the CCD readout in order to improve the data rate and readout noise. The result, at the end, is a one-dimensional image of 12 AL samples.

As shown in Fig. 2.5, each window has a size of 120 by 240 μm in the AL and AC directions, or 707.196 by 1414.384 mas (milli-arcseconds) on the sky. This means that each μm in the images approximately corresponds to 5.89 mas on the sky.

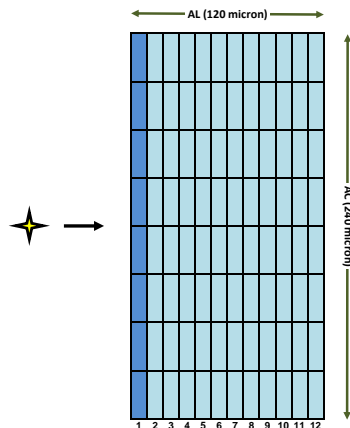


FIGURE 2.5: A schematic map of the CCD pixels read out in one window. The numbers at the bottom show the number of each sample. The photoelectrons in each such column (e.g., column 1 shown in darker colour) are binned together at readout producing a single value representing the integrated intensity in that column.

2.4 The photometric G band

The light collected in the AF CCDs is not filtered and hence, when used for broad-band photometry, will give the brightness or the corresponding magnitude of the objects in the white-light photometric ‘ G ’ band. According to the definition of this band, based on the silver coating of the six mirrors of the Gaia telescope as well as the quantum efficiency of the AF CCDs, the G magnitude of an object is defined as

$$G = -2.5 \log_{10} \left(\frac{s}{s_{A0V}} \right), \quad (2.1)$$

where s is the number of photons received from an object with the spectral photon distribution $N(\lambda)$ (expressed, for example, in $\text{ph s}^{-1} \text{m}^{-2} \text{nm}^{-1}$), and s_{A0V} is the corresponding value for an A0V star with visual magnitude $V = 0$. Knowing the dimensions of entrance pupils in the AL and AC directions (D and H , respectively), the mirror reflectivity for the silver coating $T_{Ag}(\lambda)$, the integration time ($\tau \simeq 4.4$ s) as well as the quantum efficiency $Q_{AF}(\lambda)$ of an AF CCD, s can be written as:

$$s = DH\tau \int_0^{\infty} N(\lambda) T_{Ag}^6(\lambda) Q_{AF}(\lambda) d\lambda. \quad (2.2)$$

The normalized response of the G band in different wavelengths along the normalized transmission curves for the standard (Bessell) U (ultraviolet), B (blue), V (visual), R (red) and I (infrared) filters are presented in the left panel of Fig. 2.6. As can be seen from this diagram, the mean wavelengths for the U , B , V , R , I and G bands are approximately 365 nm, 436 nm, 545 nm, 641 nm, 798 nm and 600 nm, with full widths at half maximum of about 58 nm, 90 nm, 82 nm, 150 nm, 155 nm and 430 nm, respectively. The plotted $UBVRI$ response curves are based on the standard filters response data published by Bessell (1990) and the G response curve is calculated as

$$G(\lambda) = T_{Ag}^6(\lambda) Q_{AF}(\lambda) \quad (2.3)$$

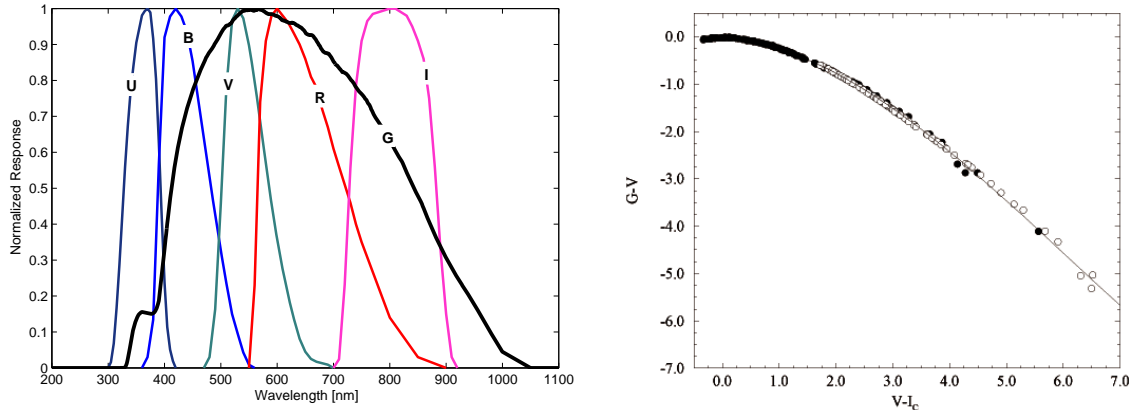


FIGURE 2.6: Left: normalized response curves of the G band along with the standard (Bessell) $UBVRI$ filters. Right: $G - V$ versus $V - I$ colour-colour diagram.

using data from the Gaia Parameter Database (Lindegren, private comm.).

In addition, the right panel in Fig. 2.6 shows the relation between the G -band and the standard Johnson V -band as a colour-colour diagram. Using this diagram, the G magnitude of any star with known V and I magnitudes can be calculated.

2.5 Accuracy of Gaia observations

As already mentioned, Gaia is designed for very high-accuracy positional measurements. This is achieved by using large and high-quality instrument components such as the primary mirrors, high quantum efficiency CCDs, a large mosaic of CCDs, as well as the optimized observing law which results in many scans of each position on the sky at different epochs and in different position angles.

By considering all these instrumental parameters as well the all possible error sources (e.g., wavefront errors and readout noise), the accuracy of the astrometric measurements can be predicted. The predicted accuracies shown in Table 2.1 were presented by Lindegren et al. (2008) based on the Gaia Accuracy Analysis Tool (de Bruijne, 2005) and using the latest data of the instruments and the mission. The accuracies were computed for single solar-type (G0V) stars without extinction.

TABLE 2.1: Predicted accuracies of the astrometric measurements for single G0V stars, with no extinction (from Lindegren et al., 2008)

V magnitude	6-13	14	15	16	17	18	19	20
parallax [μas]	8	13	21	34	55	90	155	275
proper motion [$\mu\text{as yr}^{-1}$]	5	7	11	18	30	50	80	145
position at mid-epoch ($\cong 2015$) [μas]	6	10	16	25	40	70	115	205

Chapter 3

Gaia Imaging Simulations

This chapter describes the simulation of the two-dimensional Gaia images used in the further analysis. In fact, it is a very necessary step at this time to be able to make realistic image simulations of the future Gaia observations, in order to develop and test reliable analysis algorithms in a good time before the launch of the mission. The processing procedures can then be tested and improved in several development cycles before the real observations arrive. To cover this need, simulation code should be written by considering all factors which contribute to the image production.

The code developed for this thesis project is a step in this direction. Among all the possible astronomical objects, only stars are simulated in the current work, for two main reasons: (i) they can normally be considered as point sources and hence the simulations and then the reduction algorithms can be done without the additional complexities of extended objects, and (ii) they will make the main contribution to the Gaia object database. In addition, among all the astrometric, photometric and spectroscopic measurements of the Gaia satellite, I will only focus on the images produced in the astrometric fields (AF). The astrometric CCDs will produce several one dimensional (1D) images for any point in the sky according to a Line Spread Function (LSF) that will be determined as part of the instrument calibration. Then, the two dimensional (2D) images can be produced using the obtained raw data. In the following sections, I will show the complete chain of simulating 1D images of Gaia observations by considering the appropriate CCD characteristics and possible noises as well as the scanning law. Furthermore, the different methods and algorithms tested, for producing the corresponding 2D images, are discussed.

3.1 The Line Spread Function

Line Spread Function (hereafter LSF) for 1D images is the marginal distribution of the corresponding Point Spread Function (PSF) for 2D images. It shows how the light of a point source is distributed over the pixels of a CCD. This function, which is a convolution of the optical diffraction image with the instrumental responses (pixel size etc.) was originally modelled and numerically computed under the realistic assumptions of the Gaia instruments by Lindegren (1998a). In these computations, it is considered that the LSF is a function of the object's spectral distribution folded with the Gaia response (G band) and of the

position in the field of view, as well as being different in the two observing directions.

The real *LSF* is thus complicated and depending on many factors, but for the present work it will be approximated by a Gaussian profile as:

$$L(u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u^2}{2\sigma^2}}, \quad (3.1)$$

where u is the along-scan displacement of the pixel from the image centre. The definition of the image centre is in general a matter of convention, but for a symmetric LSF it is natural to put it at the point of symmetry, as for the function above. The standard deviation σ is approximately equal to one pixel (or 10 μm , or 58.9 mas) for the AF observations.

3.2 Sources of noise

Like any other astronomical observations, the Gaia data will also be affected by different natural and artificial noise sources. The sky background and cosmic rays are two examples of the background noises which come from natural sources. In addition, the quantum nature of light naturally causes the signal noise (i.e., photon noise) on the observed data. On the other hand, artificial noises mainly come from instrumental (e.g., CCD) effects among which can be noted the dark current (the electrons generated in the CCDs even in the absence of any signal) and the readout noise (the uncertainty of the CCDs' reading-out process). Among all different possibilities of noise sources, the effects of photon noise, background electrons and read-out noise are more considerable in Gaia observations, and so they will be considered in the image simulations. The effects of cosmic rays are not negligible, but only a small fraction of the observations are affected and these are relatively easy to filter out.

3.2.1 Readout noise

Readout noise is an uncertainty in the final output value when the number of electrons in each pixel (here, the sum of several pixels) is read at the serial register of the CCD. It is mainly produced by a combination of three random processes: (i) spurious electrons generated in the CCD's electronics, (ii) that the conversion from the analog signal to a digital value is not perfectly repeatable, and (iii) quantization noise in the analogue-to-digital (A/D) conversion. The first two processes are strongly dependent on the readout frequency, viz., the longer time is available for each A/D conversion, the smaller is the readout noise. This is the main reason why only small windows are read around each detected object, rather than all the pixels.

The readout noise causes some additional fluctuations of the output pixel values, which adds a noise component that is independent of the size of the signal and can be modelled as a Gaussian variable. For the reading of the AF CCDs, the typical readout noise has a standard deviation (RMS value) corresponding to about 5 electrons. This noise is equivalent to the photon noise from a uniform background of 25 electrons per sample. As the charges are summed on-board the satellite and reach the readout register of the CCD when the star(s) leave the CCD, the total noise is estimated together with the background noise at a certain value.

3.2.2 Dark current

The thermally generated electrons in CCD pixels, even in absent of any signal, is called dark current noise which contributes to the measured signal. The value of the dark current noise for the Gaia CCDs is very small because the CCDs are cooled to about 160 K. The dark current is therefore neglected for this study.

3.2.3 Background electrons

The background photoelectrons are mainly produced by (i) the zodiacal light and (ii) galactic and extragalactic backgrounds (e.g., diffuse nebulae, planetary nebulae, HII regions). The current total predicted background, containing the read-out noise and dark current, is estimated as 100 photoelectrons per sample (i.e., for the sum of eight across-scan pixels). It starts to become important for objects fainter than $G = 16$ mag.

3.2.4 Photon noise

The photon noise from the source and sky counts is normally assumed to have a Poisson distribution. It is the result of the statistical nature of the photon detection process, and therefore depends only on the detection probability, i.e., intensity of the light. The photon noise for the Gaia observation is simulated by generating a random Poissonian variable from the expected number of electrons counted in each sample.

A star of magnitude $G = 15$ generates about 75,000 photoelectrons during the 4.4 s it takes to move over the CCD.

3.3 Simulation of 1D images

According to the Gaia CCDs structure and in term of the Line Spread Function defined in Sect. 3.1, and by considering the different noise sources, the one dimensional (1D) images are simulated. Each 1D image consists of 12 samples, denoted N_1, N_2, \dots, N_{12} . For a single star the expected number of electrons in the k th sample is expressed as follows:

$$E[N_k] = \alpha L(k - \kappa) + \beta, \quad (3.2)$$

where α is the intensity ('amplitude') of the star in electrons, calculated according to the object's G magnitude as

$$\alpha = 75000 \times 10^{-0.4(G-15)}. \quad (3.3)$$

and β represents the additive background and readout noise, expressed as the equivalent number of electrons in each sample.

κ is a parameter specifying the exact location of the image centre with respect to the 12 samples. For example, $\kappa = 6.5$ for an image that is exactly centred in the window, i.e., between samples 6 and 7.

The counts are then computed by adding the random noise (P_k) based on Poissonian statistics:

$$N_k = \alpha L(k - \kappa) + \beta + P_k, \quad (3.4)$$

where

$$\mathbb{E} \left[P_k^2 \right] = \alpha L(k - \kappa) + \beta. \quad (3.5)$$

All the simulations in this work are implemented in MATLAB, the matrix-oriented programming language, and by considering the expressions and initial values mentioned in this and previous sections. In all the calculations, the photoelectron counts N_k are represented as real numbers (in double precision), which means that fractional electrons are often considered.

In the following, differential angular coordinates (x, y) define the position of the star(s) relative to some reference position on the sky. The $+x$ axis points to North and the $+y$ axis to East. Each 1D image corresponds to a scan across the object in position angle θ , where $\theta = 0$ for a scan along the x axis and $\theta = 90^\circ$ for a scan along the y axis.

The windows (1D images) are *approximately* centred on the reference point $(x, y) = (0, 0)$, but they cannot be *exactly* centred because the phasing of the CCD pixels in each scan is independent of the choice of reference point. In practice the phasing of the pixels will be random, which is simulated by generating the positional parameter κ in (3.4) randomly between the two central samples, (i.e., between 6 and 7), for a star at the reference point. In the case that the star is not located at the reference point, or when there is more than one star in the window, the star positions are added to the along-scan position parameter as shown below:

$$\delta = \frac{x_i \cos \theta_j + y_i \sin \theta_j}{10}, \quad (3.6)$$

and

$$c = \kappa_j + \delta \quad (3.7)$$

where x_i and y_i are the map coordinates of the i th star, θ_j is the position angle of the j th scan, and the number 10 is the size (in μm) of each sample in the along scan direction. c is the new along-scan position parameter which should be used instead of κ in Eq. (3.4) in this case. The map coordinates (x, y) will be further used in Sect. 3.4.2. As an example of one simulated 1D image, Fig. 3.1 shows the case of an object at equatorial coordinates R.A. = 20° and Dec. = 75° when three stars are considered in the field of view with the following information:

- $G = 16$ mag, located at the center
- $G = 16.4$ mag, located at $45 \mu\text{m}$ from the center to the left
- $G = 16.6$ mag and located at $35 \mu\text{m}$ from the center to the right

This image is illustrated to scale and the window position $\kappa = 6.38$ was considered in the simulation.

It is obvious even from this simple example that locating the objects and finding their accurate positions in the 1D images will be difficult and sometimes impossible, even though some of the disturbing stars can cause a shift in positions of the brighter star. Figure 3.2 shows three different cases of simulated 1D images along with their photoelectron distributions and their corresponding LSFs.

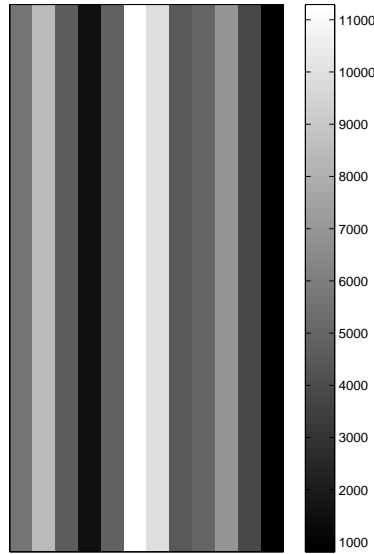


FIGURE 3.1: An example of a simulated 1D Gaia image. This image, shown to scale, was simulated for three stars in the window, with $G = 16, 16.4$ and 16.6 mag and located at $0, 265.05$ and 206.15 mas from the center.

3.4 Image stacking: Producing two-dimensional images

Two-dimensional images can generally give more scientific information and thus it is desirable to produce them from the 1D images directly given by the Gaia observations. During the five-year mission, many 1D images of the same point of the sky and with different position angles will be obtained. These images can later be stacked to make a 2D image, using the precise satellite pointing information derived as part of the data processing. The general method for this image stacking process is the superposition of all the 1D images onto a fixed and well defined map. But the problem is that the geometry of the images is quite peculiar, with very elongated samples of different sizes, in different orientations and different sub-pixel phases, which could increase the uncertainty and reduce the resolution of the final image. Therefore, in the following sections the different possible techniques which I tested in this work are explained; and at the end, a new enhanced image stacking method is proposed for analyzing the Gaia images.

3.4.1 Shift-and-add technique

The simplest and most common method of image stacking is the shift-and add technique, which is often used to obtain a high quality images from several short exposures. Therefore, it was the first method I tested for the present purpose. In this application, a square map with 300×300 square cells of $1 \times 1 \mu\text{m}^2$ was created and then all the 1D images were added to this map. This adding process involves determining the differential shifts of each image and then shifting all images back to a common centre. The most important point is keeping all the images in their original position angles. Therefore the sample value (number of photoelectrons) of each pixel of each 1D image is added to the corresponding map cells

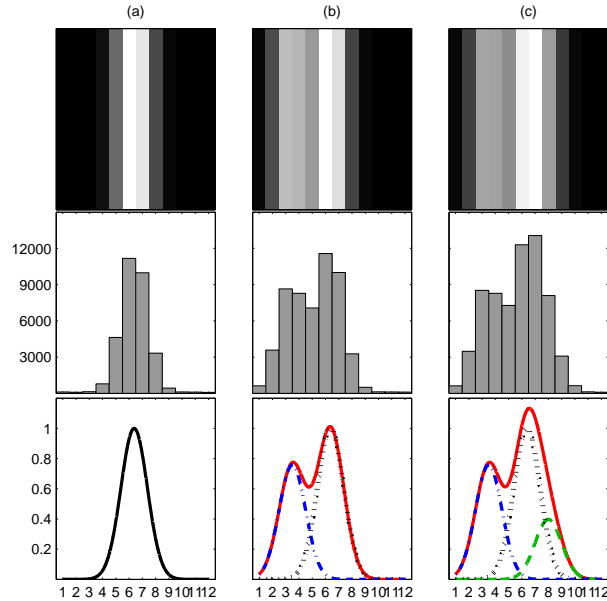


FIGURE 3.2: Three examples of simulated one-dimensional images (top panels) along with their distributions of the sample values (middle panels) and the corresponding line spread functions (bottom panels). In these examples, the different cases are: (a) one star with $G = 16$ located at the center, (b) two stars with $G = 16$ and 16.3 mag at the center and $30 \mu\text{m}$ away from the center towards the left, and (c) three stars consisting of the stars in (b) plus another star with $G = 17$ at $15 \mu\text{m}$ away from the center towards the right.

which are much smaller than the pixel sizes. Although this method achieved a 2D image, the image quality obtained was not satisfactory at all, because of the pixel geometry and inadequate process of adding the electrons to the map cells.

3.4.2 Sub-sampling

The second stacking method investigated is slightly different but expected to be much more efficient than the previously described shift-and-add technique. In this method, each sample in a 1D image was considered as a mosaic of square sub-samples (one μm on each side). Then the number of photoelectrons in the sample was uniformly divided among the sub-samples. In this way, each sample with the size of $10 \times 240 \mu\text{m}$ consisted of 2400 sub-samples, each of with an electron count equal to $N_k/2400$. After, the mosaic of sub-samples should be translated and then rotated in order to stack it on the map defined in the previous section. It should be mentioned here that the sub-sample size was chosen to be the same as cell size of the map in order to give a sufficient overlap.

This algorithm can be summarized as below:

1. A flux map of 301×301 cells with a cell size of $1 \times 1 \mu\text{m}^2$ is created for the stacking process. For further calculations, this flux map is considered as a square matrix with indices i, j counted from an origin at the left-top corner. The Cartesian coordinate system (x, y) has its origin at the centre of the map, with the $+x$ and $+y$ axes in the directions of decreasing i and j , respectively.

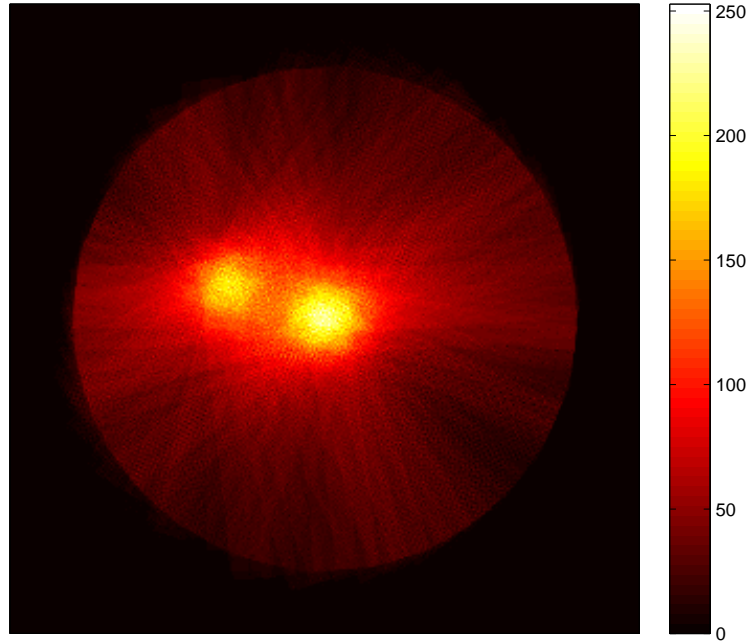


FIGURE 3.3: An example of the 2D image produced using the sub-sampling technique. This image was simulated for two stars: $G = 16$ mag at image center, and $G = 16.3$ mag at $50 \mu\text{m}$ from the center in position angle 70° .

2. Each one-dimensional image is placed on the map in a way that the centre of the image is located exactly on the map's centre and considering the correct position angle.
3. The algorithm goes through each sample and divides it into 2400 sub-samples.
4. The position of the centre of each $1 \times 1 \mu\text{m}^2$ sub-sample is found in a Cartesian coordinate system (d, e) having its axes parallel to the along-scan and across-scan directions, respectively.
5. The position found in the (d, e) coordinate system is converted to the (x, y) coordinate system by taking into account the phase (κ_j) and position angle (θ_j) of the scan.
6. The matrix indices (i, j) are calculated from the (x, y) by shifting the origin and truncating to integers.
7. On the other hand, the photoelectron count in each sub-sample is computed by dividing the number of electrons in the corresponding sample (i.e., as determined in 1D image simulating) by 2400.
8. The calculated sub-sample value in previous step is added to the corresponding array element (i, j) of the map computed in Step 6.
9. The above procedures are repeated for all sub-samples in a sample, for all samples in a 1D image, and for all 1D images of the target object.

Although the resolution obtained in these images (e.g. Fig. 3.3) is satisfactory, the sub-sampling produces a lot of extra noise due to the uneven mapping from the rotated and shifted 1D images to the common map. This is unacceptable in view of the high accuracy of Gaia. Therefore, more sophisticated algorithms had to be investigated.

3.4.3 Drizzle method

The ‘Drizzle’ method, which is also formally known as the Variable-Pixel Linear Reconstruction, is a technique for linear reconstruction of an image from undersampled and dithered data, originally developed by Fruchter & Hook (2002) for the combined Hubble Deep Field images. The Drizzle method is conceptually similar to sub-sampling but with the great advantage that it calculates a weighted average of the (sub-)pixels contributing to the re-sampled image, which greatly reduces the noise introduced by the resampling. The weight factors are calculated from the amount of overlapping between the input and output pixels.

In addition, as normally applied to astronomical images, each pixel from the original images is shrunk (e.g., by a factor 2) and placed at the centre of initial pixel, before it is mapped onto the finer output grid. Shrinking the input pixels reduces the amount of smearing introduced by the resampling.

Because the Drizzle method is demonstrated to work very well in astronomical applications, it was interesting to try and apply this approach for the Gaia image stacking. After shrinking the pixels, sub-sampling and adding the 1D image to the map, it was however noted that the empty spaces between the shrunk pixels cause a non-smooth image on the output map. It was understood that this effect is produced by the much larger ratio of the pixel size and the output map’s cell size than what was originally considered for the Hubble images. Therefore, in this original form the Drizzle method does not seem to be adequate for the present application.

3.4.4 Drizzle method without shrinking

To improve the performance of the Drizzle method, I ignored the pixel shrinking used in its original approach and only applied its computation of overlapping areas for weighting the sub-samples. In this way, the procedure is exactly as explained in Sect. 3.4.2 but instead of adding the whole sub-sample value (i.e., the calculated number of photoelectrons in each sub-sample) to the corresponding output cell according to the position of the sub-sample center, the overlapping areas were computed and the photoelectrons distributed according to the areas. The following steps show how this distribution was done:

1. The first six steps are the same as for the sub-sampling algorithm, explained in Sect. 3.4.2.
2. The matrix indices (i, j) are calculated from the (x, y) by shifting the origin and truncating to integers (this step is also the same as for the sub-sampling).
3. By considering that the truncated (integer) part of i and j represent the *mean* position of the sub-sample on the output map, the decimal part shows how the *actual* position is displaced from the centre of the output cell. If, for simplicity, we only consider an

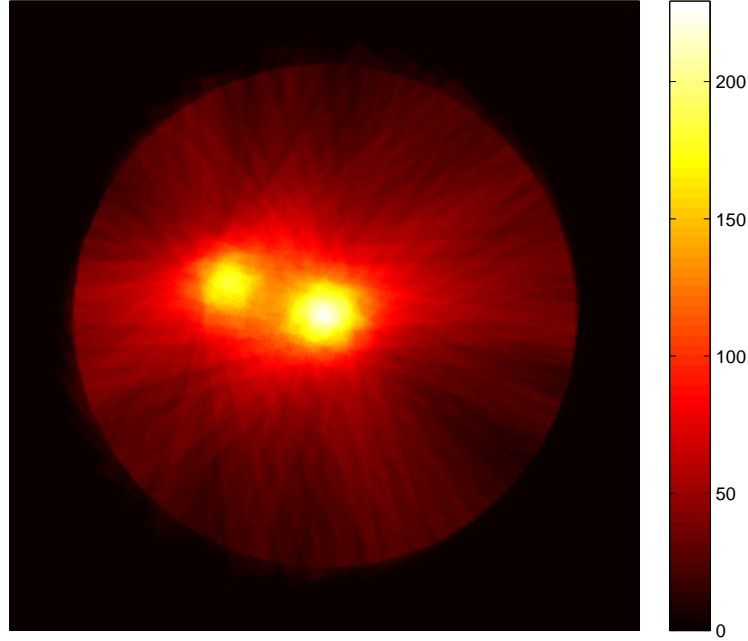


FIGURE 3.4: An example of the 2D image produced using the Drizzle method without shrinking. This image was simulated for two stars: $G = 16$ mag at image center, and $G = 16.3$ mag at $50 \mu\text{m}$ from the center in position angle 70° .

image which is aligned with the output map (i.e., position angle 0° , 90° , or 270°), the decimal parts of the real-valued array indices easily give the overlapping area fractions for the output cells at (i,j) , $(i+1,j)$, $(i,j+1)$ and $(i+1,j+1)$.

4. According to the overlapping area fractions, the value of each sub-sample (i.e., number of photoelectrons in the sub-sample) is distributed among the cells on the output map.
5. Again, the whole procedures explained in the previous steps have to be repeated for all sub-samples in a sample, then for all samples in a 1D image, and then for all the 1D images of the target object.

As it is clearly evident from the 2D image obtained with this modified Drizzle method (Fig. 3.4), the resolution and resampling noise suppression is much better than the previous tested approaches. Nonetheless, an even better method is investigated in the following.

3.4.5 A new geometrical stacking method for increased quality

With the ideas from all the methods of image stacking explained in the previous sections, and in order to improve on those techniques, different geometrical changes were examined and luckily a new approach was found. In this method, unlike the previously discussed approaches, a rectangle is defined around each 1D sample, and the entire output map is then searched for sub-samples belonging to a certain sample (i.e., with a specific index k). Thus, the whole rectangular region of the sample is recognized in the output map, and the

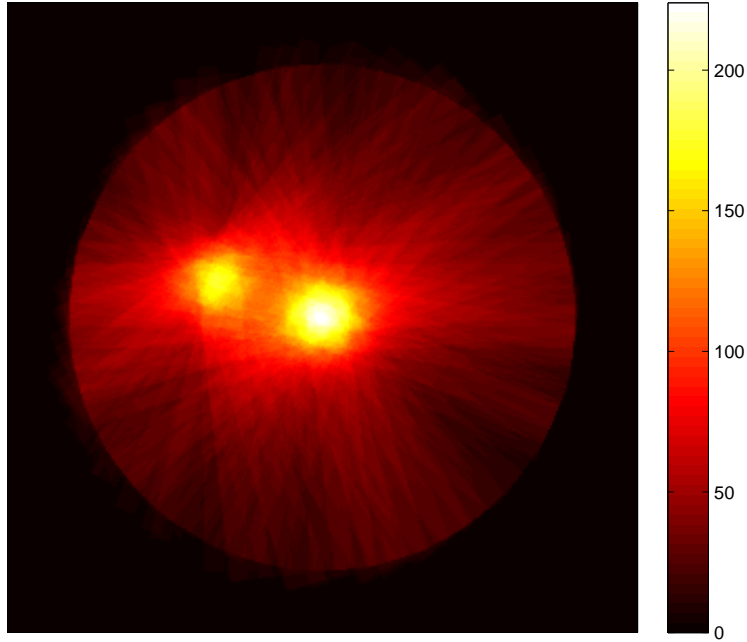


FIGURE 3.5: An example of the 2D image produced using the proposed geometrical stacking method. This image was simulated for two stars: $G = 16$ mag at image center, and $G = 16.3$ mag at $50 \mu\text{m}$ from the center in position angle 70° .

corresponding sample value is uniformly added to that rectangular area. By applying this simple algorithm, a much higher resolution was obtained and its effect was considerable on the further analysis of the 2D images for detecting stars. In summary, the proposed method is as follows:

1. For each cell (i, j) of the output map, the corresponding coordinates (x, y) and then (d, e) are computed.
2. According to the (d, e) coordinates of the 1D image, the sample index (k) is determined.
3. The sub-sample values are calculated for the k th sample by dividing the sample value by 2400 (i.e., the number of $1 \times 1 \mu\text{m}^2$ sub-samples in each sample).
4. All the cells in the output map having the same sample index (k) receives its corresponding sub-sample value calculated in the previous step.
5. All the above steps have to be repeated for all one-dimensional images of the target object.

This stacking process is proposed as the most successful test for producing the Gaia two-dimensional images. Figures 3.5 to 3.7 show some images produced with this algorithm.

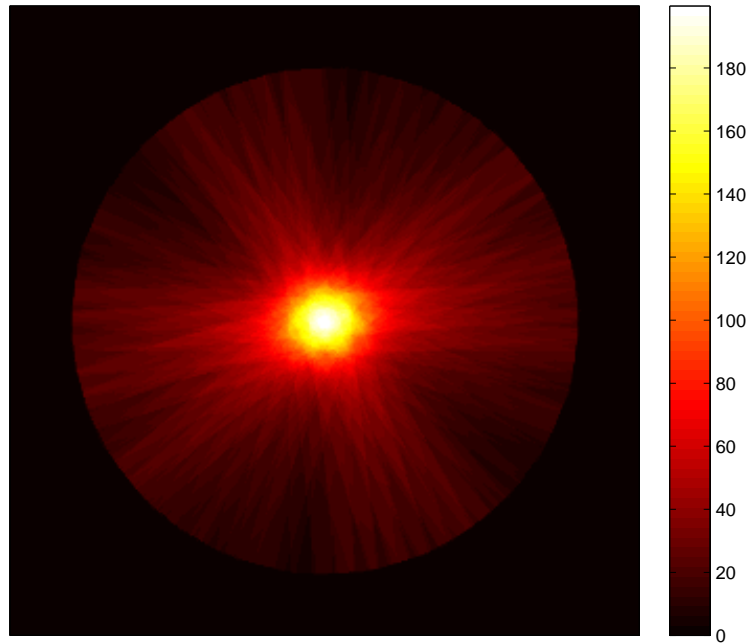


FIGURE 3.6: An example of the 2D image produced using the proposed geometrical stacking method. This image was simulated for one star with $G = 16$ mag located at the image center.

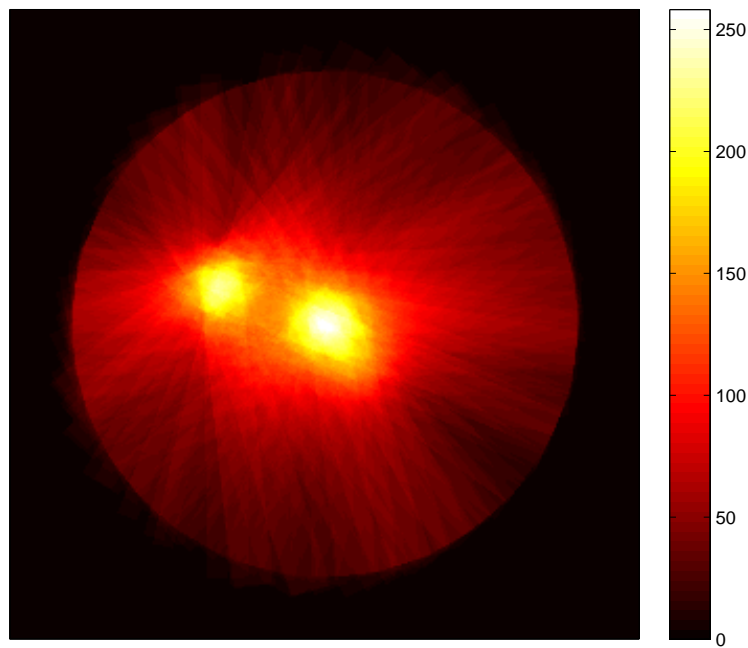


FIGURE 3.7: An example of the 2D image produced using the proposed geometrical stacking method. This image was simulated for three stars with $G = 16$ mag at image center, $G = 16$ mag at $55 \mu\text{m}$ from the center in position angle 70° and $G = 17$ mag at $25 \mu\text{m}$ from the center in position angle 220° .

Chapter 4

Detection of Faint Stars

Any detection of objects in astronomical images is based on some theoretical assumptions which allow us to discern between signal and noise. According to this theory, one can set up criteria for what should be regarded as a detection. Finding the most reliable detecting approach is an important issue because there will always be an uncertainty on the final decision. Whether the star (as a signal) is present or not, and also whether any star is detected or not, the outcome has some uncertainty and the reliability of the result depends both on the detection criterion and on the actual situation. In detection theory, the different possibilities are categorized in the four cases:

1. When the star is actually present on the image, and it is detected through the detection process: *true detection*
2. When the star is actually present on the image, but it is not detected through the detection process: *missed detection*
3. When the star is not actually present on the image, and it is not detected through the detection process: *correct non-detection*
4. When the star is not actually present on the image, though a star is detected through the detection process: *false alarm*.

In any given situation, one of the two hypothesis must be selected: (i) a star is detected, or (ii) no star is detected. No third option is allowed in a detection process. Therefore, the sum of the probabilities of detection and false alarm is always 100 percent, whether or not a star is actually present. Thus as a first step before applying the detection algorithm, the desired detection probability or alternatively a reasonable probability of false alarm should be decided. As a criterion for star detection at a certain position in the 2D Gaia image, it is reasonable to require that the number of photoelectrons in the corresponding cell of the output map must exceed a certain threshold. A procedure for finding the threshold level in the simulated images will be explained in Sect. 4.2. The criterion adopted there is that the probability of false alarm should only be 1 percent.

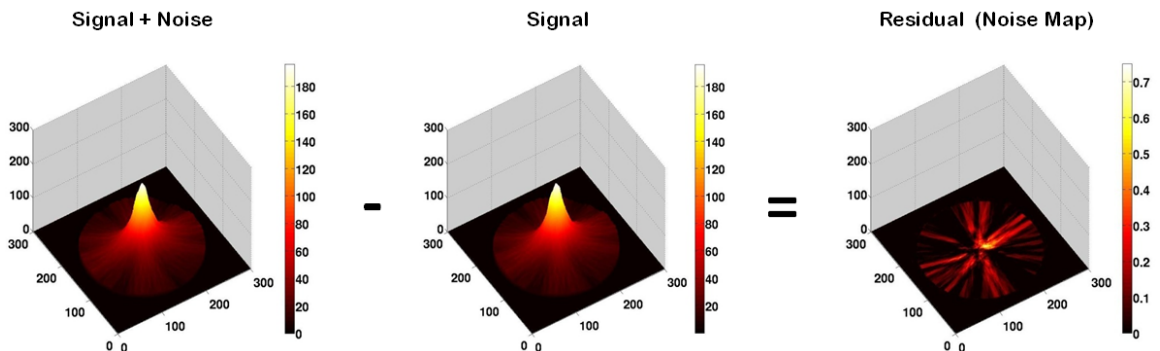


FIGURE 4.1: Finding the noise map by subtracting the signal map from the signal+noise map generated under otherwise identical conditions.

4.1 Image de-biasing

Before applying threshold detection, it is convenient to remove the constant background level from the images, which is referred to as image de-biasing. This step can only be done if we have a good understanding about the several different noise sources. For the real Gaia data processing, de-biasing is far from trivial, but for the simulated images it is easy. Background electrons, introduced in Sect. 3.2.3, added a constant amount of electrons to any one-dimensional image during the simulations. This amount, taken to be 100 electrons per sample independent of the signal, is subtracted from the 1D images before the stacking process. Applying the de-biasing to the 1D images was decided for simplicity because there would be no difference in the final results if it was instead applied to the 2D image.

4.2 Noise level and detection threshold

Similar to all detection algorithms, star detection in the Gaia 2D images needs to determine a threshold such that any value above the threshold can be considered as starlight. If the threshold level is low, then more stars might be detected, but there is also an increased number of false alarms. On the other hand, if the threshold level is high, then fewer stars might be detected, but with a smaller probability of false alarm. Therefore, the threshold should be set at a level which satisfies a certain probability of false alarm for any case under study.

Because the constant noise (i.e., background) was subtracted from the image in the image de-biasing process, and because the remaining noise in the Gaia images (i.e., with Poisson statistics) are not constant by their nature (e.g., photon noise depends on the star brightness), a fixed threshold level cannot be used. It is expected that the level should be related to the magnitude of the brightest star in the image. As in this work I am simulating the images myself, I can control the initial parameters of the image. For example, to find a suitable threshold level for a certain magnitude, I can simulate two sets of 1D images with exactly the same initial parameters (magnitudes, positions, background level, scanning geometry, ...) except that one of the sets has no Poisson noise. After de-biasing and stacking the 1D images, two 2D images exist for the same configuration: one with the

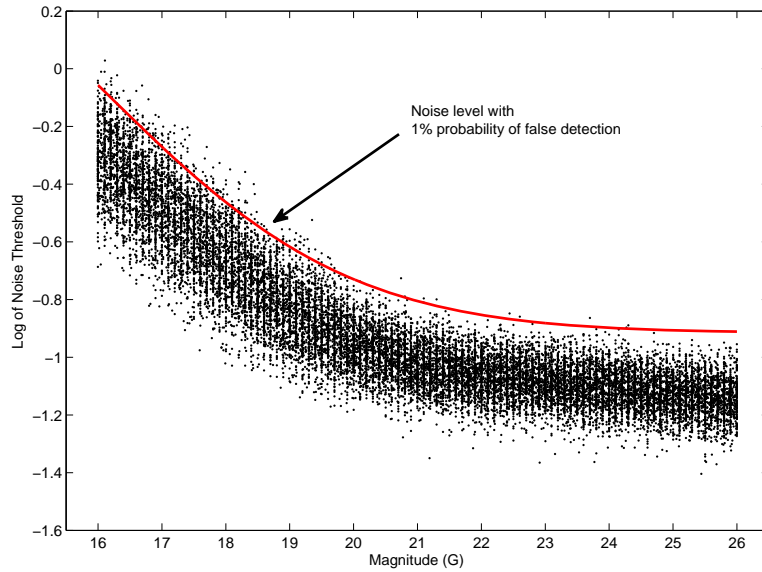


FIGURE 4.2: Points: peak values in the noise maps for different G magnitudes of the target star. Curve: the adopted threshold level for detection of faint stars with one percent probability of false alarm.

Poisson noise, and one without. Now, If I subtract the second one (only signal) from the first one (signal+noise), what is left will be the noise map. Figure 4.1 shows this procedure schematically. The noise map looks very strange because of the resampling and stacking of the originally uncorrelated noise in the 1D images.

In this case, the highest value in the noise map can be considered as a minimum threshold level for detecting faint stars in a given configuration with a bright target star at the centre of the map. Repeating this calculation for target stars of different magnitude should give a diagram of the minimum threshold level as function of magnitude. But, since the Poisson noise is random and different in each calculation, there is a scatter of peak values at each magnitude. Therefore, thousands of cases were calculated and the peak values plotted as a function of the G magnitude (Fig. 4.2). Then, setting the probability of false alarm to be one percent, a curve of threshold values was fitted such that only one percent of the points are above the threshold.

As seen from the scatter of points in Fig. 4.2, it is clear that the threshold curve has two regions: (i) the right part where the target star is actually too faint to be detected by the Gaia on-board detection system, and where the noise reaches a constant level mainly of background and readout noise; and (ii) the left part where the noise is proportional to the square root of the signal (slope = -0.2 dex/mag), as expected for Poisson noise. By fitting straight lines in these two regions, finding their intersection point and combining them, the mean relation between noise and magnitude was found. This curve was then shifted upwards until only one percent of the points were left above the curve, in order to meet the condition of one percent probability of false alarm, giving the threshold curve plotted in the figure.

It can be seen that the above procedure to determine the threshold curve is not ideal,

since there are more points above the curve for bright target stars than for faint ones. This could have been corrected by shifting the curve slightly to fainter magnitudes.

4.3 The CLEAN method

The CLEAN method, which was originally proposed by Högbom (1974), has become one of the most successful deconvolution techniques in radio astronomy. In this method, the true image is considered to be entirely composed of point sources. The dirty map is successively decomposed into a set of Dirac delta functions convolved with the Point Spread Function (PSF, or dirty beam) through the CLEAN algorithm. This is done by finding the brightest point in the dirty map, subtracting a scaled PSF centred on this point, and adding the corresponding delta function on a blank (transition) map. As more and more flux is removed from the dirty map, the corresponding flux is transferred to the transition map in the form of delta functions. In each PSF subtraction, only a fraction of the flux at that point is removed. This fraction is named loop gain and can have any value between zero and one. Optimization of the loop gain is considered in Sect. 5.1.1.

All these procedures are iteratively done until the largest brightness of the residual map, which is used as a new dirty map in each iteration, reaches a specific pre-defined intensity limit, i.e., in this case the detection threshold.

The CLEAN method in its original form fits well for images in radio astronomy but needs some small changes for application to the Gaia images. The main problem is that the dirty beam (PSF) is not the same in different points of the map, because of the limited size of the 1D images both along and across scan. After some experiments, a successful adaptation to the present problem was however found. In this approach, the dirty beam is computed exactly as a simulated dirty map, without background and noise, and with only one star at the position of the brightest point of the current dirty map, and with a scaled intensity according to the loop gain.

The step by step procedure of the CLEAN method, applied to the Gaia 2D images, is shown below:

1. An empty map (i.e., with all cell values equal to zero) is created, of the same size as the dirty map, i.e., with 301×301 cells. This is the transition map.
2. The 2D image, as a dirty map, is considered as the input and the algorithm finds the position and the value of the cell with the largest number of photoelectrons.
3. Because the position of the maximum can be between the cells and not exactly at the centre of a cell, a small algorithm is used to find the accurate (sub-cell) position of the maximum. This algorithm can be itemized as:
 - A 3×3 sub-matrix of the dirty map is considered around the point of interest, where the central point is the cell with the largest value.
 - Inside this sub-matrix, two quadratic polynomials are fitted to the values summed in the horizontal and vertical directions.
 - The maxima of the fitted polynomials are computed analytically and give the coordinates of the maximum in the horizontal and vertical directions.

- These coordinates are transferred to the original dirty map, as we know the position of the sub-matrix in the dirty map.
4. The accurate position determined in Step 3 and its intensity obtained in Step 2 is added to the transition map.
 5. In a parallel algorithm which is exactly like the algorithm used for simulating the 1D images and then stacking them to produce the 2D images, a dirty map is computed with only one star at the position determined in Step 3 and the magnitude corresponding to the number of electrons obtained in Step 2 multiplied by the loop gain (i.e., 0.5). This newly produced image corresponds to the shifted and scaled dirty beam, but rigorously takes into account the actual size of each 1D window.
 6. The (shifted and scaled) dirty beam is subtracted from the dirty map, cell by cell.
 7. The resulting residual map is considered as a new dirty map and the process is repeated from Step 2 until the residuals reach the threshold level determined in the previous section.

It should be noted here that the threshold level is chosen according to the largest cell value in the original dirty map, and therefore corresponds to the brightest star in the 2D image.

After all the above procedures, the resulting transition map contains several clusters of point sources, with each cluster located around the true position of the detected star. In fact, the overlapping flux profiles of different stars in the image cause small differences in the determined positions of each star. In the next step, the transition map is used to produce a clean map from which the accurate positions and magnitudes of the detected stars can be determined.

4.4 Gaussian convolution: Producing the clean map

In order to produce the clean map, the transition map obtained in the previous section is convolved with a 2D Gaussian function representing the ‘clean beam’. In this way the best estimate for the position of stars can be made from the small displacements in the positions which are caused by the overlapping flux profiles. The width of the clean beam should roughly correspond to the theoretical resolution of the instrument, or about 80 mas (FWHM). The adopted clean beam is a two-dimensional Gaussian function with standard deviation $\sigma = 4 \mu\text{m}$ (23.6 mas), which has a FWHM of $9.4 \mu\text{m}$ (55 mas). The investigation for the most acceptable width of this Gaussian function will be shown in Sect. 5.1.2. By adding the final residual image obtained at the end of CLEAN method (Sect. 4.3) to the convolved image, the detected stars are shown again in the whole image (clean map). Figure 4.3 shows an example which illustrates the whole CLEAN process from the dirty map (i.e., the 2D image produced as described in Sect. 3.4) to the clean map as the final result. In this example, three stars are assumed in the image simulation though two of them are almost overlapped and cannot normally be detected. From the CLEAN method iterations, only a few frames are shown in this figure.

4.5 Computing the positions

An algorithm similar to the one used for finding the accurate position of the brightest cell of the dirty map in Step 3 of the CLEAN method (Sect. 4.3) is used for determining the positions of the detected stars on the clean map. The only small difference is that the position of maximum cell value is searched in the clean map image instead, and then the small 3×3 matrix is defined, centred at that position. Then, after applying the previously explained method and finding the position of the brightest star, the second algorithm similar to the CLEAN method is applied. This time, a new map with only one star at the position of the brightest star found and with its brightness is produced and subtracted from the clean map. Now, the whole positional detection algorithm, is repeated for the remaining stars in the residual clean map. This iterative procedure is stopped when the maximum cell value is smaller than the corresponding threshold level, computed in Sect. 4.2. As can be seen, finding the position exactly follows the CLEAN method procedures but applied to the clean map instead of the dirty map, and using the clean beam (Gaussian) instead of the dirty beam.

4.6 Computing the magnitudes

The magnitudes of the detected stars can be calculated according to the known relation between the G magnitude and the number of photoelectrons mentioned in Sect. 3.3. By considering the procedures which result in the clean map, it can be understood that each point in the transition map contains a fraction of the total number of photoelectrons for the corresponded star. Therefore, by adding the values of points in the cluster corresponding to a star, the total number of electrons and hence the G magnitude of that star can be computed.

4.7 Disturbing stars detection

In the current work, the detection processes explained in this chapter are mainly concerned with objects fainter than $G = 16$ mag, and even objects which are fainter than the Gaia completeness limit ($G = 20$ mag) if they are still bright enough to be detected in the images. In fact, a main consideration for this work is that there can sometimes be very faint stars (even beyond the Gaia completeness limit) which are apparently close the main target and might not be detected in the skymapper CCDs because they hide behind the flux profile (i.e., PSF) of the brighter main target. These objects, which were described in Chapter 2 as one of the possible disturbing effects in the Gaia observations, can be detected in this way, which could result in more accurate, corrected astrometric measurements. Their detection also results in more accurate photometric measurements by allowing a better estimate of the background level. But there are some limitations of these detections which depend on the magnitude differences between the main target and the disturbing star(s) as well as their apparent separations. These limitations will be investigated in Chapter 5.

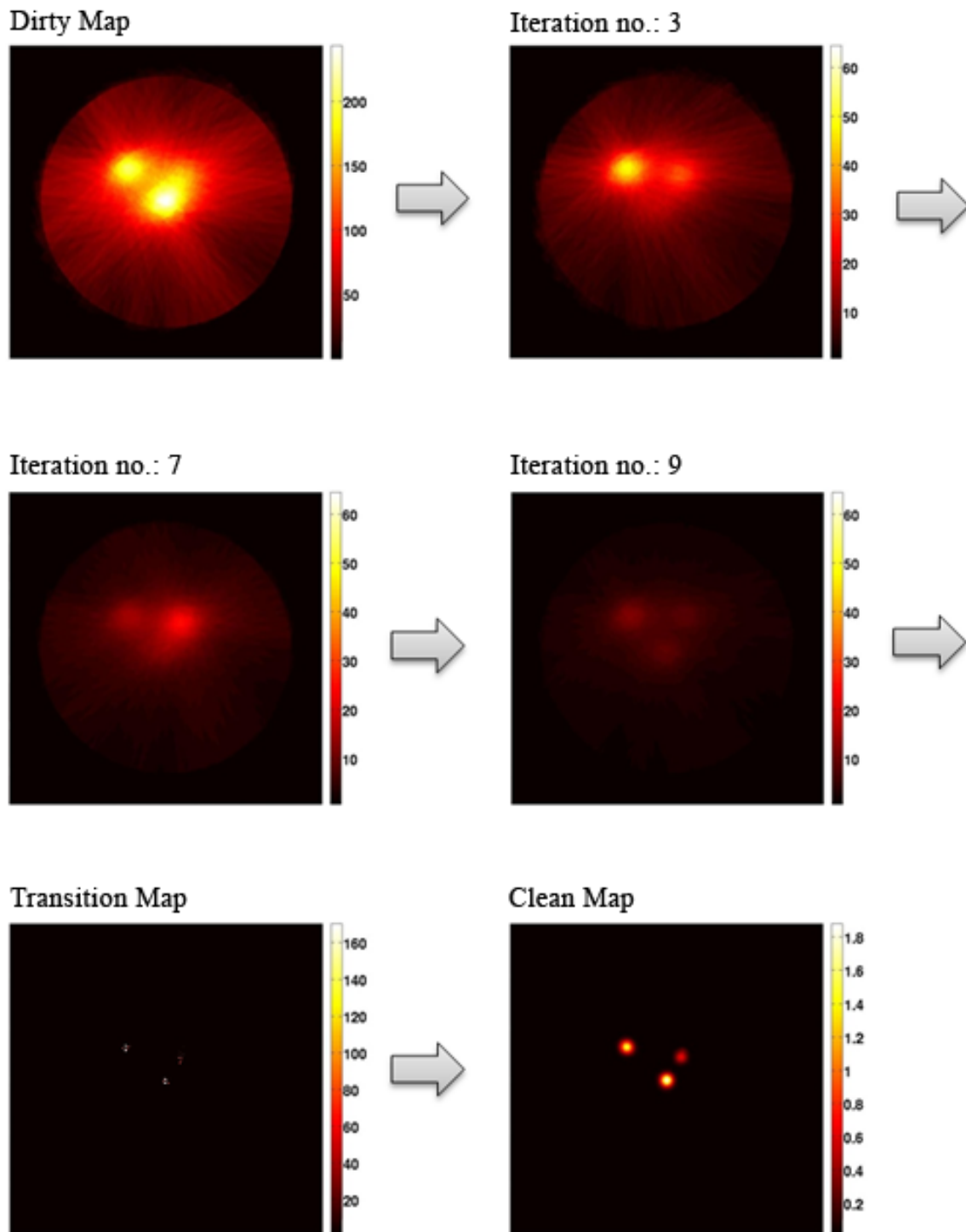


FIGURE 4.3: CLEAN process from the dirty map (top-left frame) to the clean map (bottom-right frame). The dirty map is the same as in Fig. 3.7, with three stars of magnitudes $G = 16, 16$ and 17 .

Chapter 5

Statistical Studies and Results

This chapter describes a comprehensive set of statistical studies to compute the probability of detection under different conditions. In addition, the variable parameters used in the CLEAN method are optimized by a set of dedicated statistical studies.

5.1 Optimizing the CLEAN parameters

In Sects. 4.3 and 4.4 two parameters used in CLEAN procedure were described: the loop gain and the standard deviation of the Gaussian convolution (clean beam). In the following subsections, the effects of varying these two parameters are shown for both positional and brightness measurements and then the optimum values are obtained.

5.1.1 Selection of the loop gain

Loop gain, as was explained in Sect. 4.3, is the fraction of the largest pixel value on the residual image subtracted in each iteration, and therefore can have any value between 0 and 1. By choosing this fraction equal to 1, the whole intensity at the maximum is subtracted at once in the CLEAN method, which can only give a good result in the simplest cases, e.g., if in reality there is only one star in the image or if the stars' profiles are not overlapping at all. On the other hand, although an infinitesimally small loop gain may seem to give a more safe result, Högbom (1974) showed that in practice there is little improvement when the loop gain is smaller than 0.5: a very small value for the loop gain only causes an unnecessarily large number of iterations in the CLEAN procedure.

Although a loop gain around 0.5 is therefore indicated, the variation of this parameter was studied in relation to different variables. Figures 5.2–5.6 show the final results as functions of the loop gain and for different possible values of the Gaussian beam width. In these plots are shown, from top to bottom, the differences of the input (true) and output (derived) parameters of the map coordinates x and y , the total error in position ($d = \sqrt{\Delta x^2 + \Delta y^2}$), and the intensity error. The positional errors are expressed in number of cells on the map, i.e., in the unit $1 \mu\text{m}$ (5.89 mas). The intensity error is expressed in electrons. All these variations are studied for one star with $G = 16$ mag located at the center of the image. The figures confirm that 0.5 is a good choice for the loop gain.

5.1.2 Selection of the size of the Gaussian clean beam

The chosen standard deviation σ of the Gaussian clean beam should be larger than the output map's cell size ($1 \mu\text{m}$) and smaller than the pixel size of the Gaia CCDs ($10 \mu\text{m}$ in along scan direction). It can be seen from Figs. 5.2–5.6 that the least variation on the stellar positions and magnitude differences happens around $\sigma = 4 \pm 1 \mu\text{m}$. Therefore, $\sigma = 4 \mu$ (23.6 mas) is decided as the best size for the Gaussian convolution.

5.2 Detection limit for a single star

To understand the brightness limit of true detection for Gaia, only one star is simulated on the image and the probability of true detection for that star is investigated. In this case, the detection is accepted only when only one star is detected i.e. when the only detected star has a brightness greater than the threshold level. It should be noted that the Gaia has a flux limit corresponding to the $G = 20 \text{ mag}$, meaning that fainter stars are not detected in the SM CCDs (Sect. 2.3) in the ordinary observations and hence not observed in the astrometric CCDs. Therefore, these fainter limits apply to stars which can be detected using the proposed detection technique during the data analysis phase. Høg et al. (2003) showed that even those fainter objects have expected astrometric measurement precisions of 2 and 5 mas for the limiting magnitudes of $G = 21$ and $G = 21.5 \text{ mag}$, respectively.

Figure 5.1 shows the variation of the detectability of one faint star as a function of the G magnitude for a simulated image of 41 scans (dashed line) and 103 scans (dashed line).

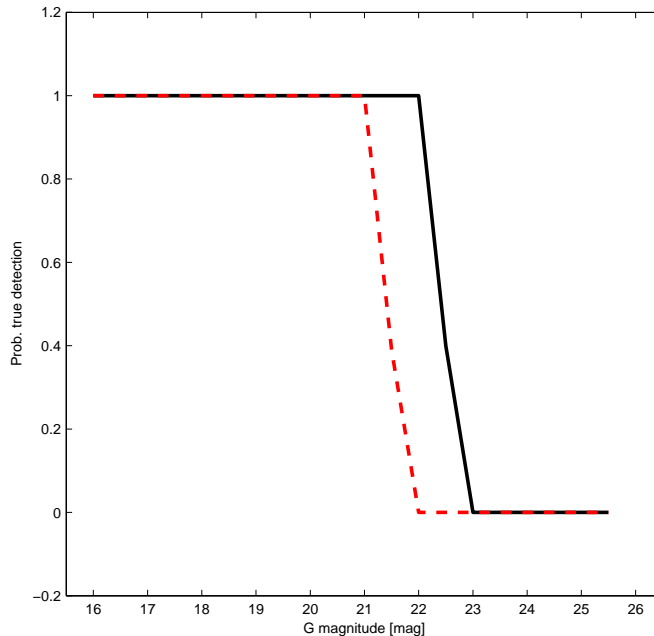


FIGURE 5.1: Brightness limit: the probability of detecting a (single) star as function of its G magnitude.

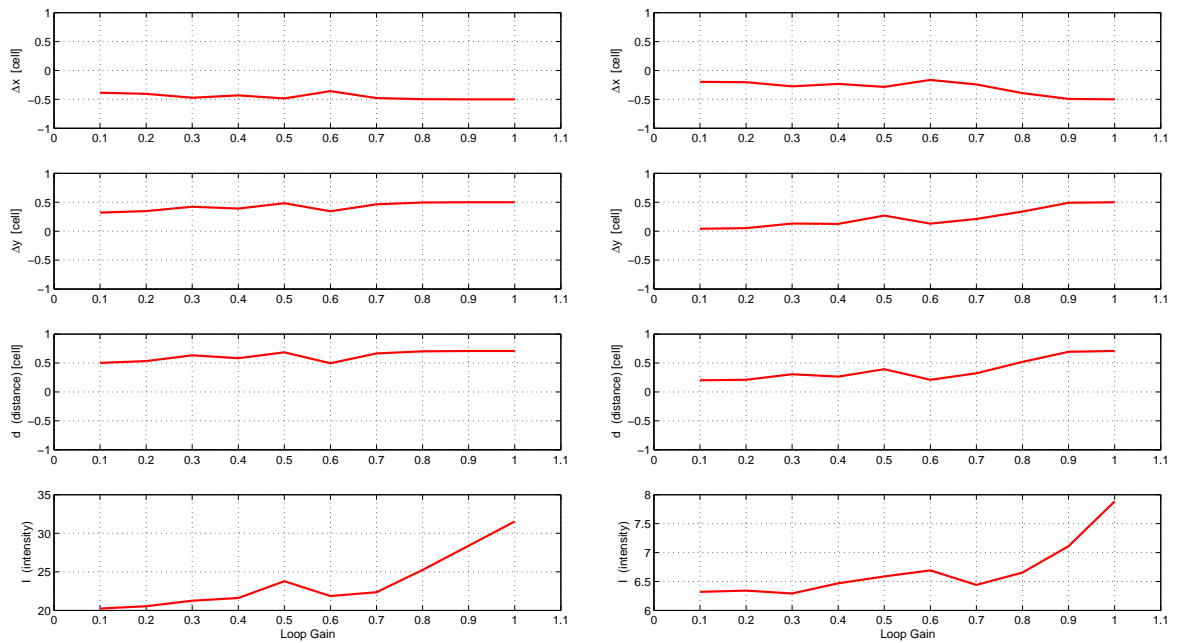


FIGURE 5.2: Loop gain optimization. The graphs show the derived positions and brightness vary with loop gain for a $G = 16$ mag star at the center, using a clean beam with $\sigma = 1 \mu\text{m}$ (left) and $2 \mu\text{m}$ (right).

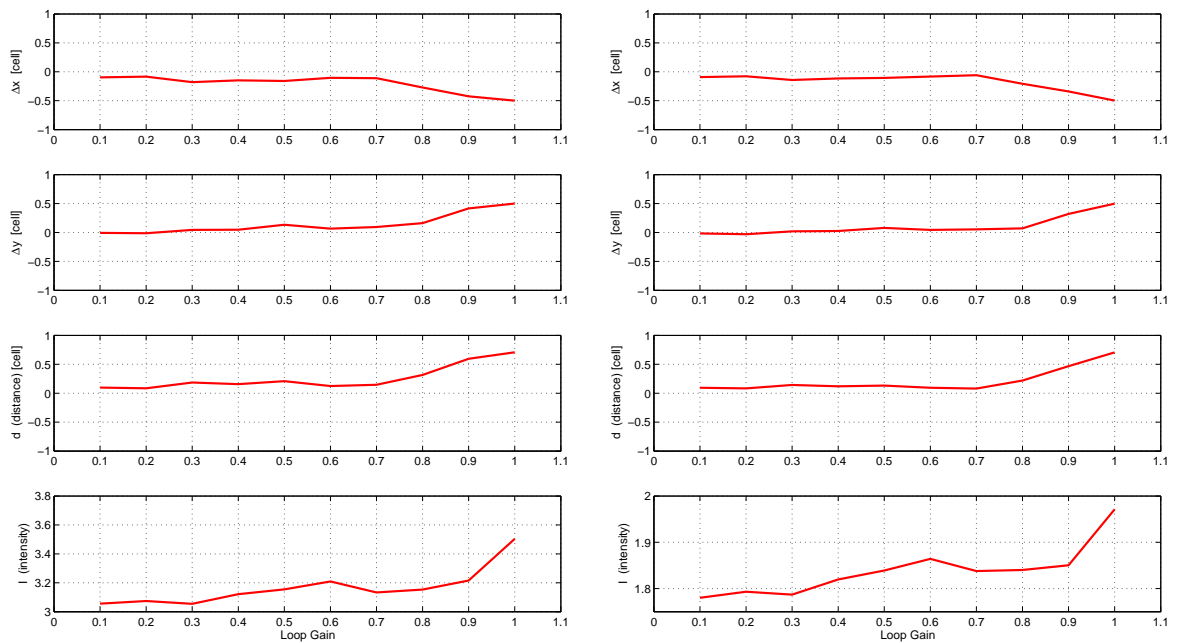


FIGURE 5.3: Loop gain optimization. The graphs show the derived positions and brightness vary with loop gain for a $G = 16$ mag star at the center, using a clean beam with $\sigma = 3 \mu\text{m}$ (left) and $4 \mu\text{m}$ (right).

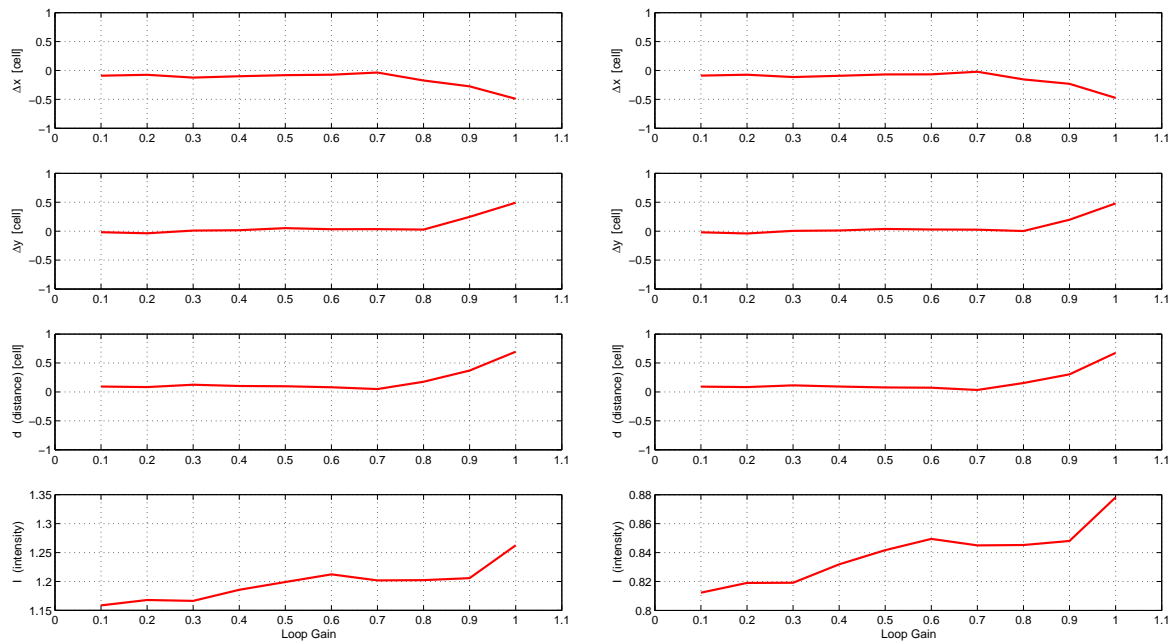


FIGURE 5.4: Loop gain optimization. The graphs show the derived positions and brightness vary with loop gain for a $G = 16$ mag star at the center, using a clean beam with $\sigma = 5 \mu\text{m}$ (left) and $6 \mu\text{m}$ (right).

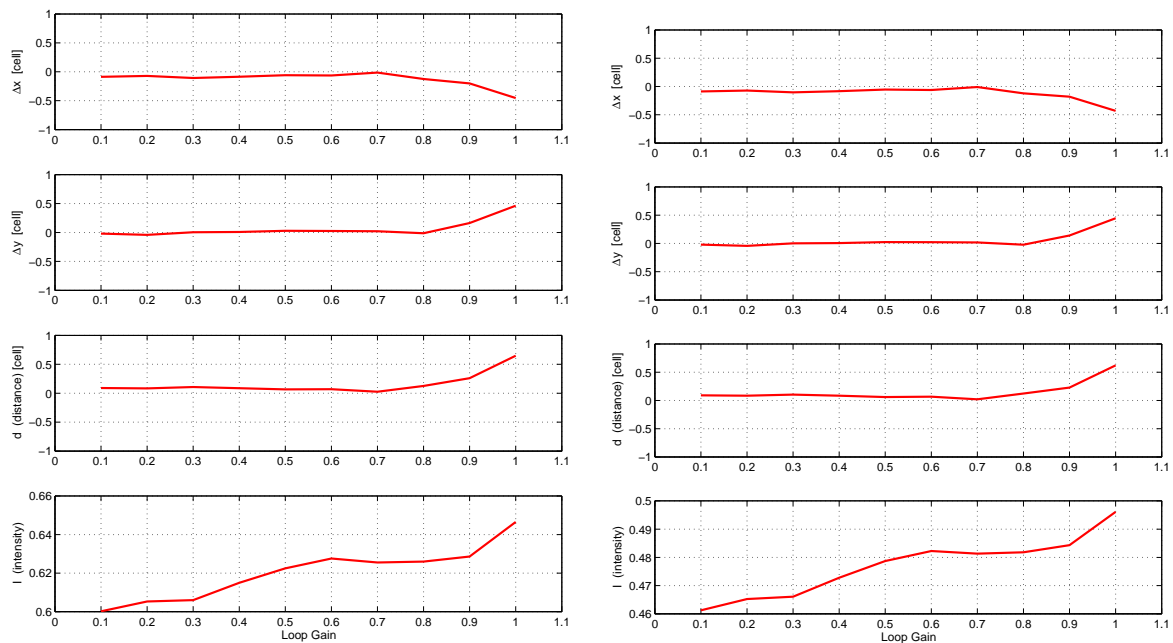


FIGURE 5.5: Loop gain optimization. The graphs show the derived positions and brightness vary with loop gain for a $G = 16$ mag star at the center, using a clean beam with $\sigma = 7 \mu\text{m}$ (left) and $8 \mu\text{m}$ (right).

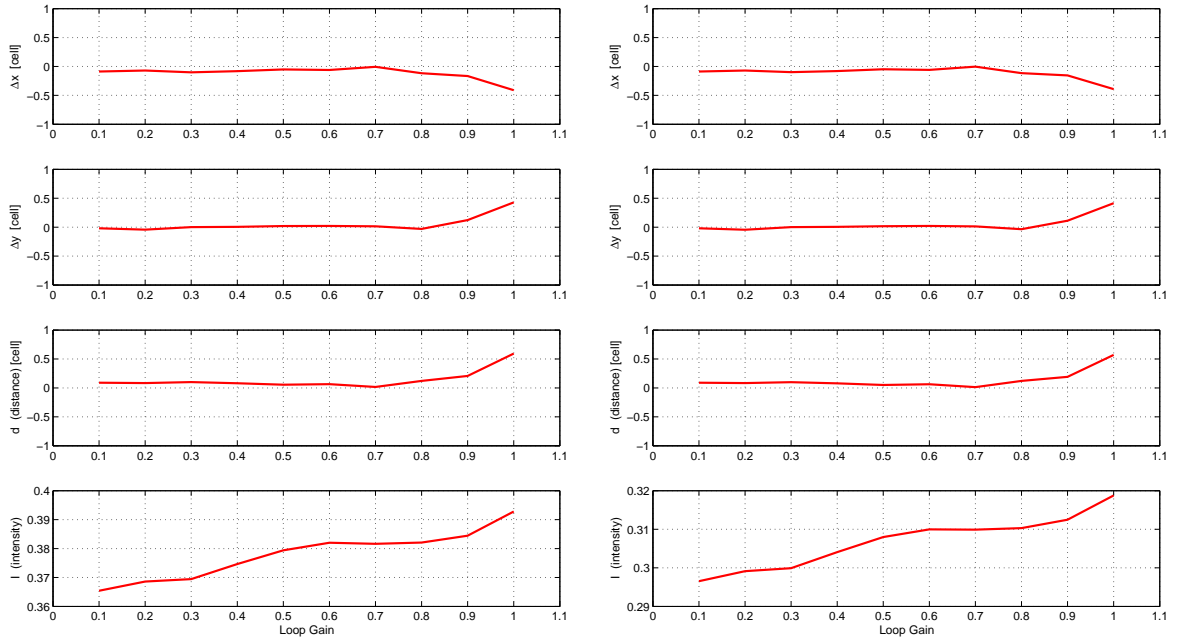


FIGURE 5.6: Loop gain optimization. The graphs show the derived positions and brightness vary with loop gain for a $G = 16$ mag star at the center, using a clean beam with $\sigma = 9 \mu\text{m}$ (left) and $10 \mu\text{m}$ (right).

5.3 Probability of true detection

To be able to compute the probability of true detection, a certain case is assumed, consisting of two stars with a given magnitude difference (ΔG_{in}) and apparent separation (ρ_{in}) and for a given sky coordinates. Then, the whole simulating and detecting processes are performed until the final results with the magnitudes (ΔG_{out}) and positions (ρ_{out}) of the detected star(s) are obtained. Following the detection theory, explained in the introductory part of Chapter 4, the unknown hypothesis is not allowed and therefore the detection is assessed in relation to a set of pre-defined criteria in terms of the two values, ‘one’ for true detection and ‘zero’ for non-true detection. The probability of true detection can then be computed by running the whole process several times and then finding the percentage of values ‘one’ obtained.

The pre-defined criteria are itemized as follows, based on the detection errors:

- The difference between the input magnitude difference of the two stars is less than 1 mag: $\delta(\Delta G) \equiv |\Delta G_{\text{out}} - \Delta G_{\text{in}}| < 1 \text{ mag}$
- The difference between the input and output separation of the two stars is less than the size of Gaussian clean beam: $\delta\rho \equiv |\rho_{\text{out}} - \rho_{\text{in}}| < 4 \mu\text{m}$

Therefore, the true detection is achieved when (i) exactly two stars are detected and (ii) their results meet both the above criteria. If the stars are detected, but the magnitude difference

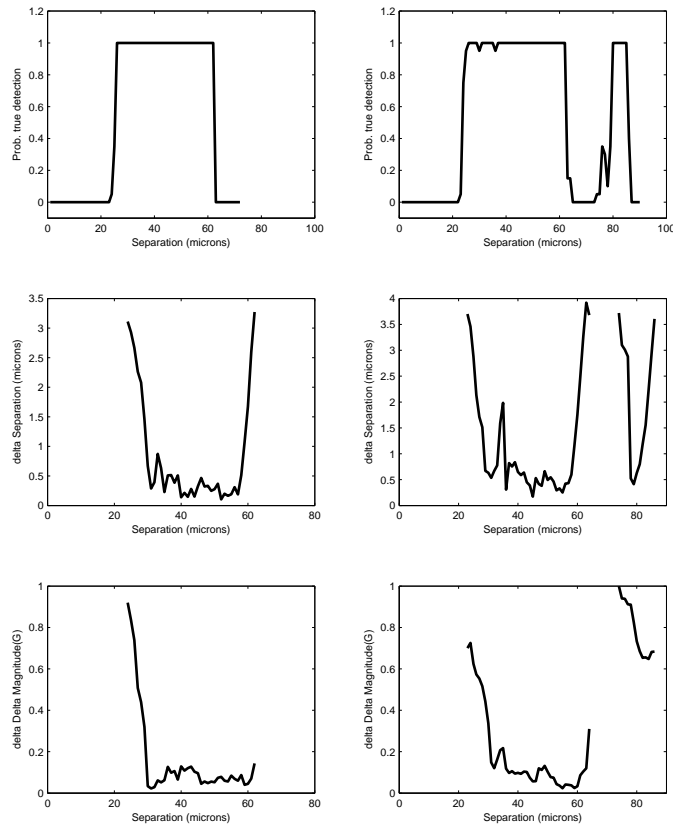


FIGURE 5.7: Probability of true detection (top panels), difference between calculated and real separation (middle panels) and difference between calculated and real magnitude difference (bottom panels), versus true separations, for the two cases of $G_1 = 16$ and $G_2 = 17$ (left panels) and $G_1 = 16$ and $G_2 = 19$ (right panels). All separations are in μm (with $1 \mu\text{m} = 5.89 \text{ mas}$).

and/or separation do not meet these criteria, it is considered a non-true detection. Also, it is considered a non-true detection if the number of detected stars is different from the real number of stars, even if some of the detected stars agree with the real stars.

5.3.1 Variation of detection probability with separation

To study the variation of the probability of true detection, I compute the probability for certain magnitudes and different separations and then plot the probabilities, the mean absolute errors (i.e., simulated minus obtained) for separation as well as magnitude difference versus separation in Fig. 5.7. In these examples, two different sets of magnitude difference are shown: (a) $G_1 = 16$ and $G_2 = 17$ and (b) $G_1 = 16$ and $G_2 = 19$. The upper limit considered for the separation is different in the two cases, viz., $72 \mu\text{m}$ for (a) and (b), respectively.

From Fig. 5.7 it can be understood that very close pairs cannot be detected as two stars, as is needed for a true detection. It is also noted that separation of larger than $60 \mu\text{m}$

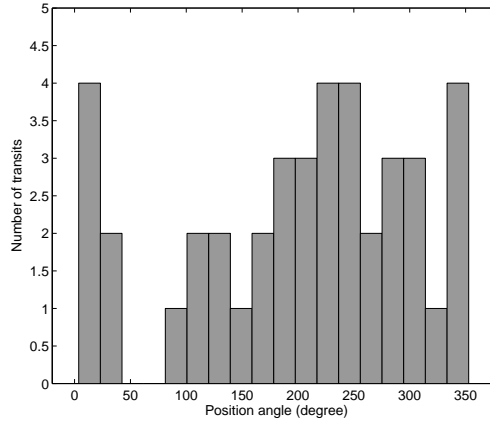


FIGURE 5.8: Distribution of the position angles (θ) of the 41 scans over the objects in Fig. 5.9.

have some strange variations. In these examples, the brighter star is put in the centre and the other is moved away from the central star. On the other hand the size of each one dimensional image of Gaia is $120 \mu\text{m}$ in the along-scan direction and hence $60 \mu\text{m}$ from the centre to the edges. Therefore, the second star will be outside the window if the stars are separated in the along-scan direction. On the other hand, the size of each 1D image is $240 \mu\text{m}$ in across-scan direction and hence $120 \mu\text{m}$ from the centre to the edges in that direction. Depending on the position angle of the second star in relation to the along-scan direction, it can then be inside or outside the window if the separation is between 60 and $120 \mu\text{m}$. In addition, for producing the 2D image, several 1D images are stacked in different position angles, which are not uniformly distributed. Therefore, the detectability of a star with separation greater than $60 \mu\text{m}$ depends on the number of transits and their distribution in position angle relative to the star.

To further investigate the dependence on position angle, the probability of true detection is computed for an example of two stars with $G_1 = 16$ mag at center and $G_2 = 20$ mag at 18 different position angles in steps of 20° . Results are presented in Fig. 5.9. The images for this example were simulated for equatorial coordinates $\text{RA} = 20$ hr and $\text{Dec} = 75^\circ$ which gives 41 field-of-view transits according to the Gaia scanning law. The distribution of the transit angles is shown in Fig. 5.8.

By comparing the different panels in Fig. 5.9, no specific systematic can be seen and even no significant correlation between the plots with different position angles and the transit angle distribution in Fig. 5.8, though some differences are evident. Therefore, to avoid the probable effect of the transit angles distribution on the probability calculation, the detectability should only be investigated within $60 \mu\text{m}$ from the image center, though the separation could be larger if one of the stars was not located at the center.

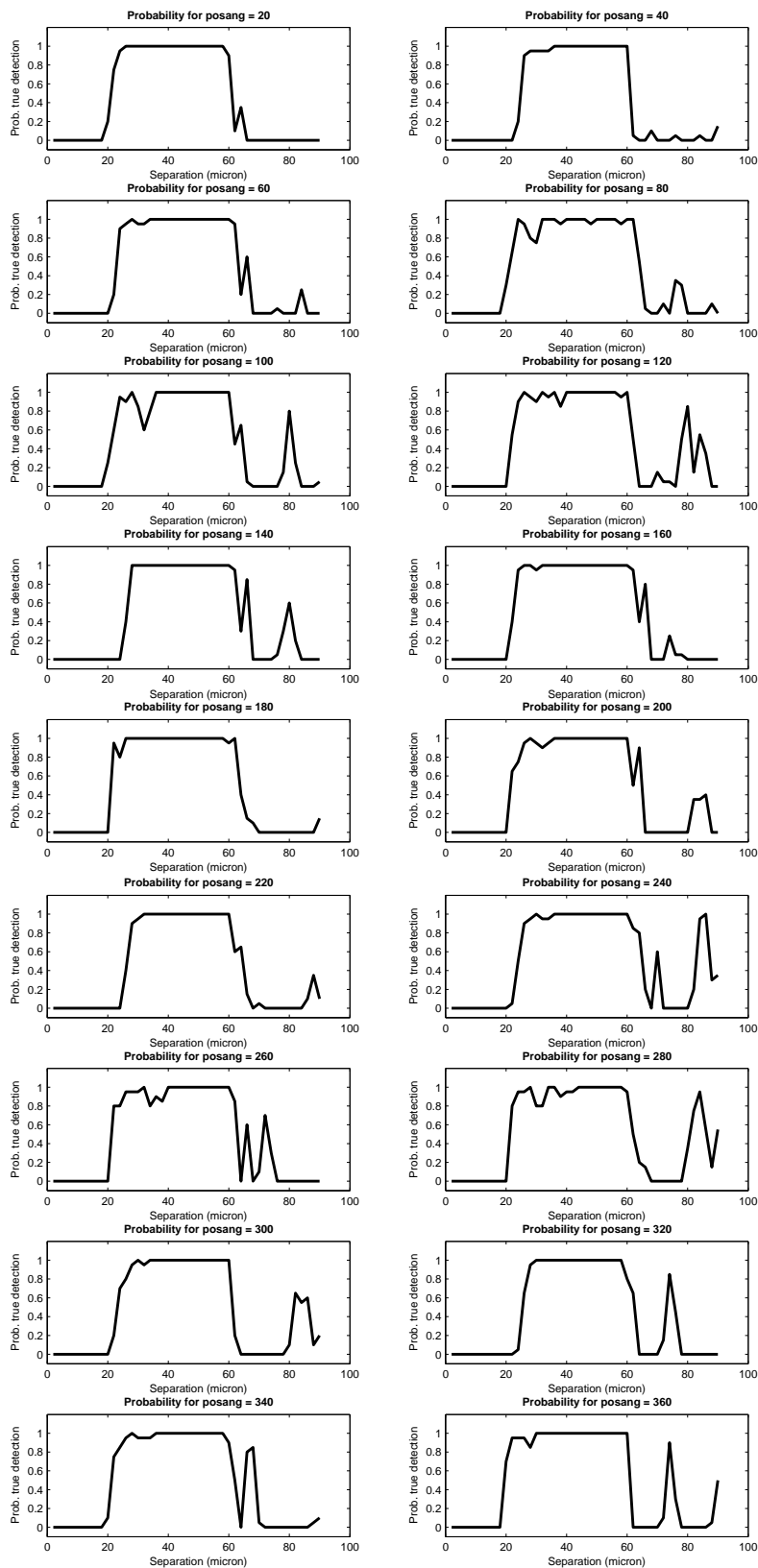


FIGURE 5.9: Probability of true detection for different position angles of the second star ($G_2 = 20$) while the first star ($G_1 = 16$) is located at center.

5.3.2 Variation of detection probability with magnitude difference

As discussed above, the detection probability may be affected by the non-uniform distribution of the transit angles. For computing representative mean values of the true detection probability, 16 different position angles were therefore considered with the first star at the center of the image and the second star on a circle around it with a certain radius (separation).

Results for different sets of G magnitudes and for magnitude differences of 2 and 4 mag are plotted in Fig. 5.10. These probabilities are again computed for simulated images of the sky position $RA = 20$ hr and $Dec = 75^\circ$ with 41 transits.

5.3.3 Variation of detection probability with the number of transits

One thing is for sure, by increasing the number of transits which are stacked to produce a 2D image, the signal-to-noise ratio (SNR) increases and so it is expected that the probability of the true detection is increased. As shown in Fig. 2.3, the number of transits varies a lot over the sky, and the number 41 used above is smaller than average. Thus higher detection probability can be expected in other parts of the sky.

To investigate this hypothesis, the experiments reported in Fig. 5.10 (for 41 scans) were repeated for different positions in the sky, which will have different number of scans during the Gaia observations. results are shown in Figs. 5.11 (103 scans), 5.12 (157 scans) and 5.13 (204 scans). As is clear from these figures, the probability of true detection is indeed increased at least when the disturbing star is fainter than $G = 20$ mag.

5.3.4 Minimum separation for a true detection

The probability plots for different magnitudes, magnitude differences as well as different number of scans, shown in previous sections, can be used to define more precisely the minimum possible separation for a true detection, i.e., when both stars are correctly identified. From all the plots presented, it is concluded that pairs closer than approximately $20 \mu\text{m}$ in the image, or equivalently 120 mas angular distance on the the sky, cannot be detected as a pair with the proposed algorithm. This minimum separation clearly seems not to have any relation with the number of scans (or consequently SNR) or the magnitude difference. Stars with larger separation can be detected with different probabilities depending on these factors.

5.3.5 Maximum magnitude difference for a true detection

The probability plots also give an indication of the maximum magnitude difference for a true detection of a pair. This clearly depends both on the magnitude of the brighter star and the number of scans. Under favourable conditions it can be at least 4 magnitudes.

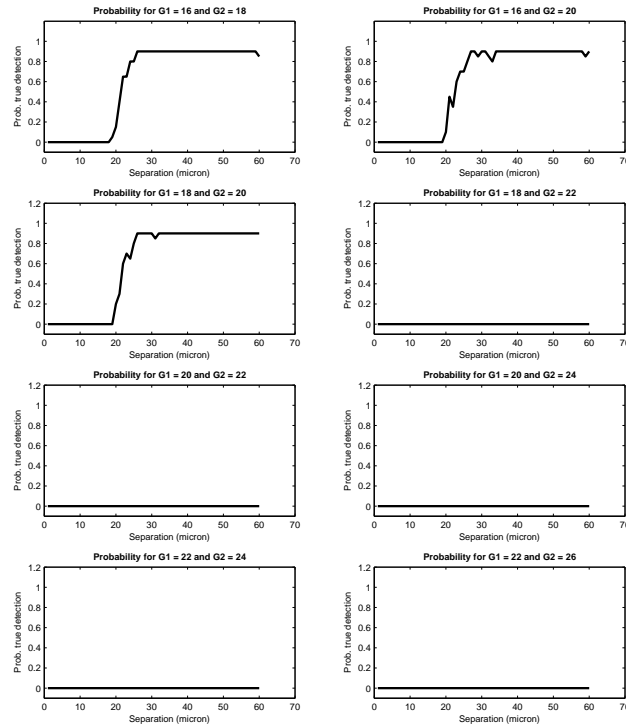


FIGURE 5.10: Mean probability of true detection versus separation, for different magnitudes, and for a position in the sky with 41 scans.

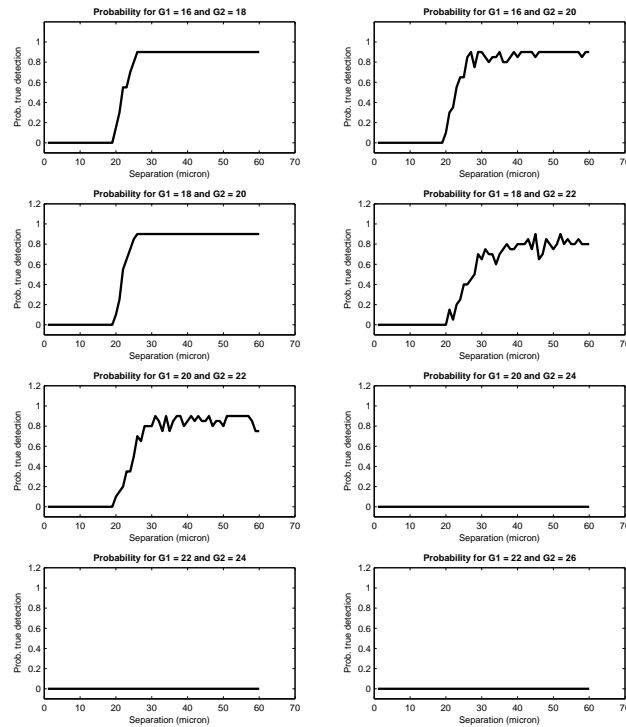


FIGURE 5.11: Mean probability of true detection versus separation, for different magnitudes, and for a position in the sky with 103 scans.

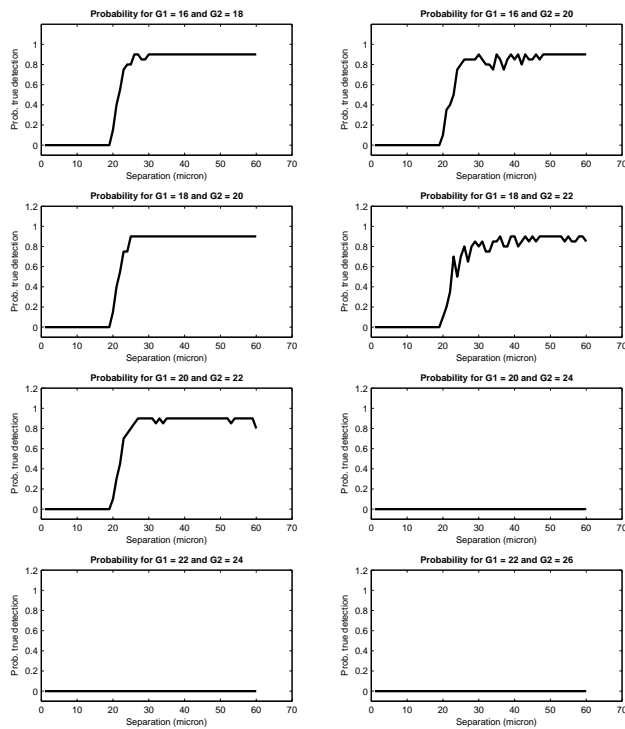


FIGURE 5.12: Mean probability of true detection versus separation, for different magnitudes, and for a position in the sky with 157 scans.

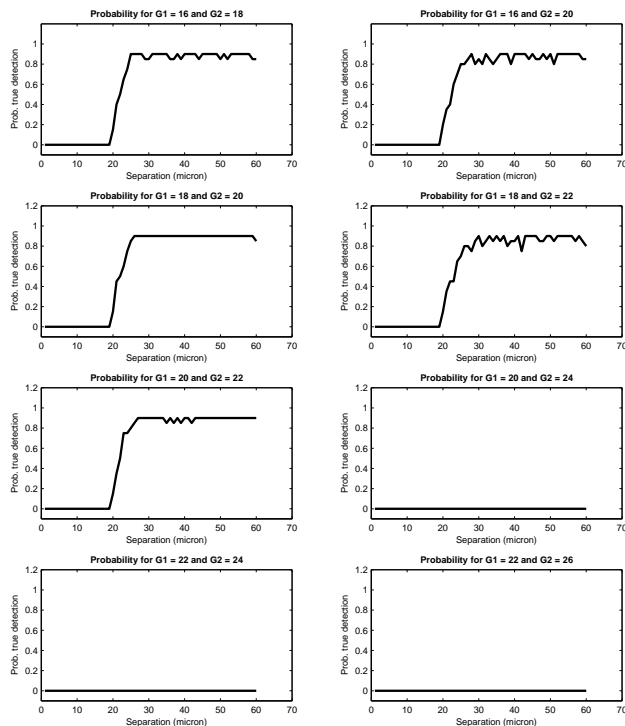


FIGURE 5.13: Mean probability of true detection versus separation, for different magnitudes, and for a position in the sky with 204 scans.

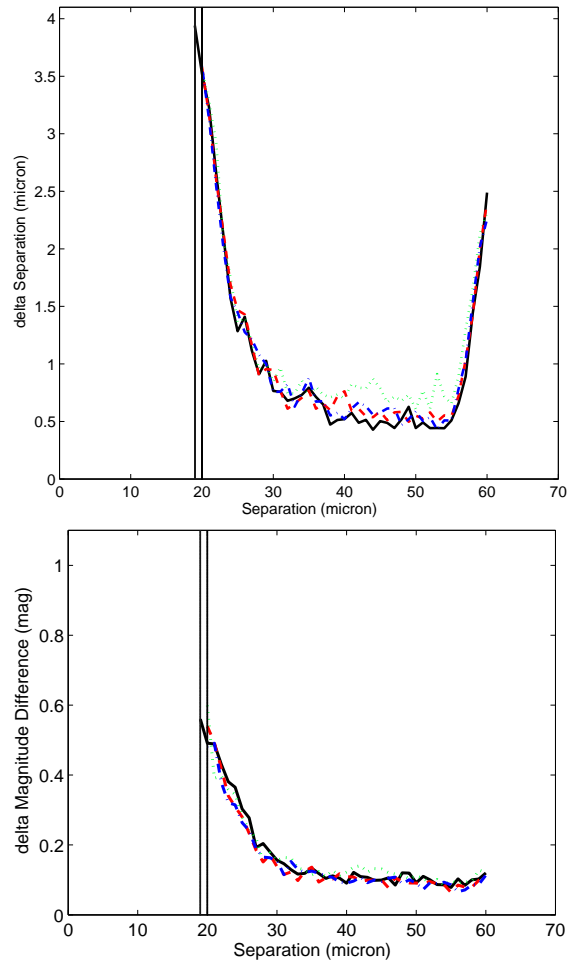


FIGURE 5.14: Mean absolute errors in the separation (top) and magnitude difference (bottom) versus the separation of the two stars. The different curves are for the cases of 41 scans (black, solid line), 103 scans (green, dotted line), 157 scans (red, dashed line) and 204 scans (blue, dashed-dotted line).

5.4 Precision of the parameters of the disturbing star

In order to correct the measurements of a target star for the disturbing effects of a fainter star, it is not sufficient to detect the disturbing star, its position and magnitude difference with respect to the target star must also be determined to a certain accuracy.

To investigate the accuracies achieved with the proposed algorithm, the mean absolute errors in separation and magnitude difference (i.e., the mean values of the $\delta\rho$ and $\delta(\Delta G)$ defined in Sect. 5.3) are plotted in Fig. 5.14. As expected, the errors decrease from a large value close to the separation limit (19–20 μm) to smaller values for well-separated pairs (35–55 μm), and then increases again for separations close to 60 μ , when the second star is outside the window in some scans. Surprisingly, the errors for the well-separated pairs do not depend much on the number of scans.

Chapter 6

Gaia FSDtool: an easy-to-use graphical user interface (GUI)

Gaia FSDtool (Faint Star Detecting tool) is a graphical user-friendly interface which was developed based on several functions written in MATLAB, the matrix-oriented programming language. This program is run with or without the commercial MATLAB program and provides an interactive and convenient tool for simulating the Gaia images as well as the detection process, based on the methods described in previous chapters.

The graphical interface is designed to facilitate user interaction with the tool. With the user-provided input parameters for any arbitrary point-source in the sky, the program simulates the 1D Gaia images, produces the very high quality 2D image, detects the stars on the image and outputs the positions and magnitudes of the detected objects. In addition, Gaia FSDtool provides different kinds of images and information in output which the user can choose and manipulate, e.g., for printing or recording on physical discs.

6.1 Developing the GUI in MATLAB

MATLAB is a matrix-oriented programming language which provides a very concisely implementation of many complex mathematical calculations without more complications like memory allocation. Therefore, it is a convenient coding environment for any data analysis and visualization routines. MATLAB was used for coding all the procedures of this work and also for coding a graphical user interface (GUI) to visualize the simulations and analysis procedures.

In general, a GUI is a graphical display in one or several windows which contains controls and the called components that gives the user this possibility to control different tasks. Therefore, the user does not have to type a command or make a script to achieve the tasks and even does not need to understand the details of the calculations. Instead, the user can only select the data from a pop-up window, or alternatively generate it by entering some parameters and then choose more specific characterizations which are needed for the intended simulation and analysis. Then, by clicking some buttons, the necessary functions for the actions are invoked. In MATLAB, the GUIs are coded as a general event-driven program and each ‘control’ has an unique coded routine which calls back to MATLAB to

ask for a specific action. In fact, each ‘callback’ is triggered by a user action like clicking a mouse button, selecting a menu item, typing a value, or pressing the enter keyboard button. Then, the GUI responds to the action and shows the next pre-defined step.

6.2 Implementation of Gaia FSDtool

The Gaia FSDtool program consists of a computational core based on all simulation and detection algorithms, which is controlled by the GUI. The first step of implementing the GUI was designing the windows’ faces according to our need. Therefore, all variable parameters which can be modified by user should be considered while the graphical interface of the windows should be also be thought about.

In this implementation, although the initial FSDtool parameters are considered to be set by user as inputs when the program is started, the default values of the major parameters according to the current decisions for the Gaia satellite instruments are given for more simplicity.

Gaia FSDtool consists of five different windows (GUIs), which ask the user different types of inputs in different windows and give the final results on the last GUI. Thus, at any step the user can go back and modify the parameters as often as needed. The GUI can easily be closed by a ‘Cancel’ button at any step. The program features are briefly described in the following sections.

6.3 Inputs options

Figures 6.1 to 6.6 in the following sections show different windows of the Gaia FSDtool’s GUI from the starting point to the final level, respectively. Figure 6.1 illustrates the first window that the user is faced with, when the code is started. This window gives some basic facts about the GUI as well as the Gaia mission, and asks the user to click on the ‘Next’ button to continue the program.

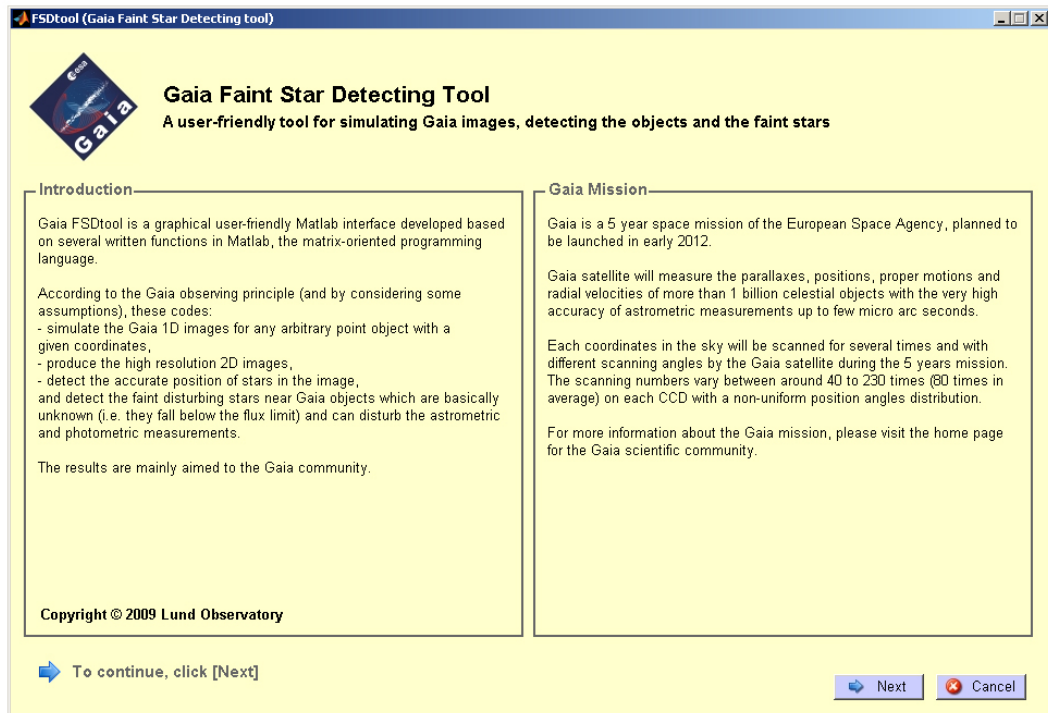


FIGURE 6.1: The first GUI of the Gaia FSDtool.

Gaia FSDtool is provided as an executable program for both Windows and Linux operating systems which means that it can be run even if the target machine does not have the MATLAB program installed. In addition, the GUI can be called from the MATLAB command window by changing the directory to the source codes folder and then typing the word 'FSDtool' in the command window. Although different input parameters of this program are explained in details in previous chapters, the different GUI inputs options are summarized in the following.

6.3.1 Data selection

After the brief introduction of the first window, the user reaches the second window where a few initial parameters should be selected as inputs. As it is shown in Fig. 6.2, the coordinates of an arbitrary object in the sky can be entered in one of the coordinate systems of equatorial, ecliptic or galactic. Then, pressing the 'Generate Data Set' button causes a data set to be generated according to the scanning law of Gaia proposed by Lindegren (see Sect. 2.2). This set contains the number of transits for the chosen coordinates together with information about each transit, e.g., the transit angle (position angle θ of the transit).

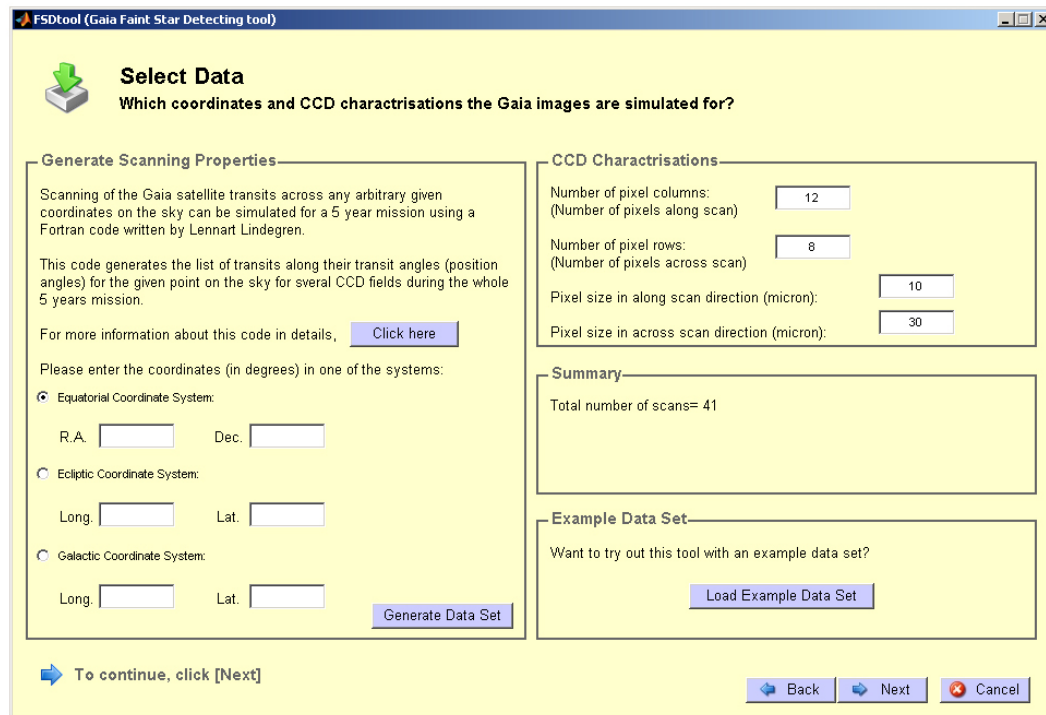


FIGURE 6.2: Data selection GUI. The second display of the Gaia FSDtool.

On the right side of the GUI, the basic sample characterizations of the CCD should be entered. These basic parameters and their default values are:

1. Number of pixel columns, which is the number of pixels in along-scan direction: 12
2. Number of pixel rows, that shows the number of pixels in across-scan direction: 8
3. Pixel size in along-scan direction: $10 \mu\text{m}$
4. Pixel size in across-scan direction: $30 \mu\text{m}$

The second box on the right side shows the number of scans according to the generated data set or the previously computed sample data set which can be loaded by clicking on the ‘Load Example Data Set’ button. Now, the ‘Next’ button is activated and will re-direct the user to the next window upon clicking this ‘Next’ button.

6.3.2 Initial characterizations

After adjusting the observing and CCD parameters, the user can decide about the initial stellar properties (i.e., G magnitudes and the positions) which are going to be simulated. According to the user interest, one or many stars can be considered in the image simulation, with different G magnitudes. The positions of the stars should also be entered in map coordinates (in μm) with any number between -60 and 60 (the size of each 1D window in along scan direction). After choosing the stellar parameters, as shown in Fig. 6.3, the

signal characterization can be adjusted. This consists of the number of background electrons (Sect. 3.2.3) as well as the standard deviation σ for the Line Spread Function of the Gaia CCDs (Sect. 3.1). As is shown in the figure, the current default values are 100 electrons and $10 \mu\text{m}$ (the size of 1 pixel in along scan direction), respectively.

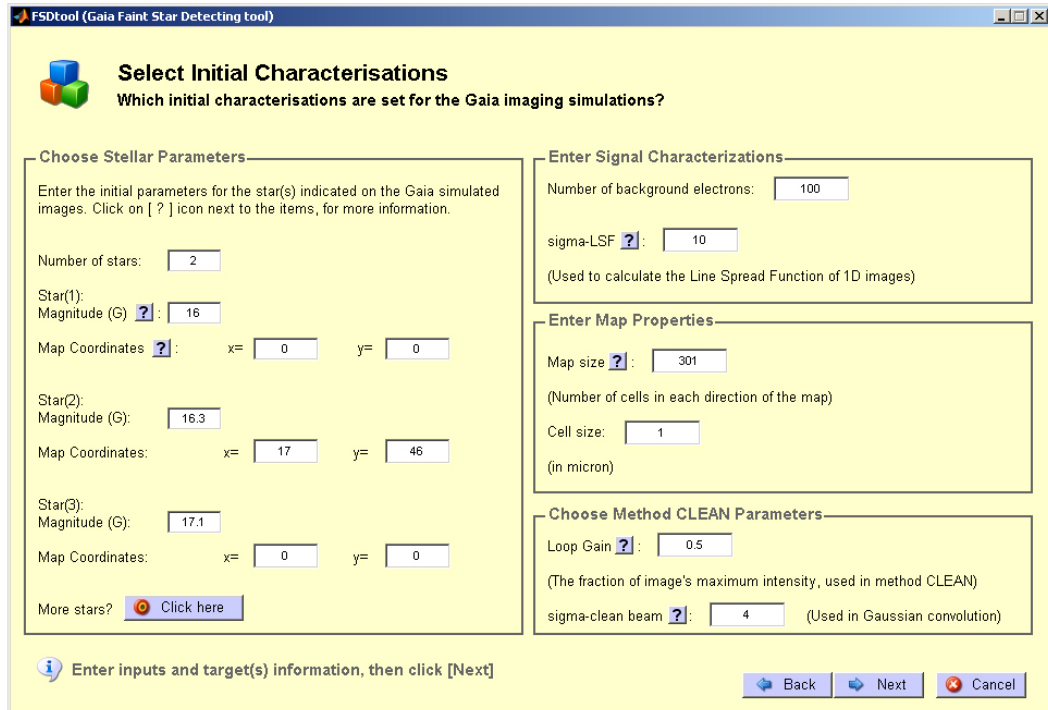


FIGURE 6.3: Initial characterization GUI. The third GUI of the Gaia FSDtool, in which the stellar parameters as well as all adjustable properties of the image simulations and the detecting procedures can be entered.

The second and third panels on the right side of this GUI enable the user to decide about the size of the whole map and also each cell of the map, used for producing the 2D image. According to the stacking process explained in Sect. 3.4.5, the values 301 and $1 \mu\text{m}$ are the default sizes for the each side of the map and each cell, respectively. In addition, the loop gain and the standard deviation σ used in CLEAN technique can be entered. The optimized values according to the statistical study in Sect. 5.1 are 0.5 and $4 \mu\text{m}$, respectively.

6.4 Output options

After setting the all initial parameters, the output options can be selected. Figure 6.4 illustrates this GUI with different possible options. For simulating the real observations, all the possible sources of noise including the background electrons and the Poisson noise (explained in Sect. 3.2) should be considered, though for investigating the effect of these noises on the final simulations and analysis, one can decide not to consider one or both types of noise.

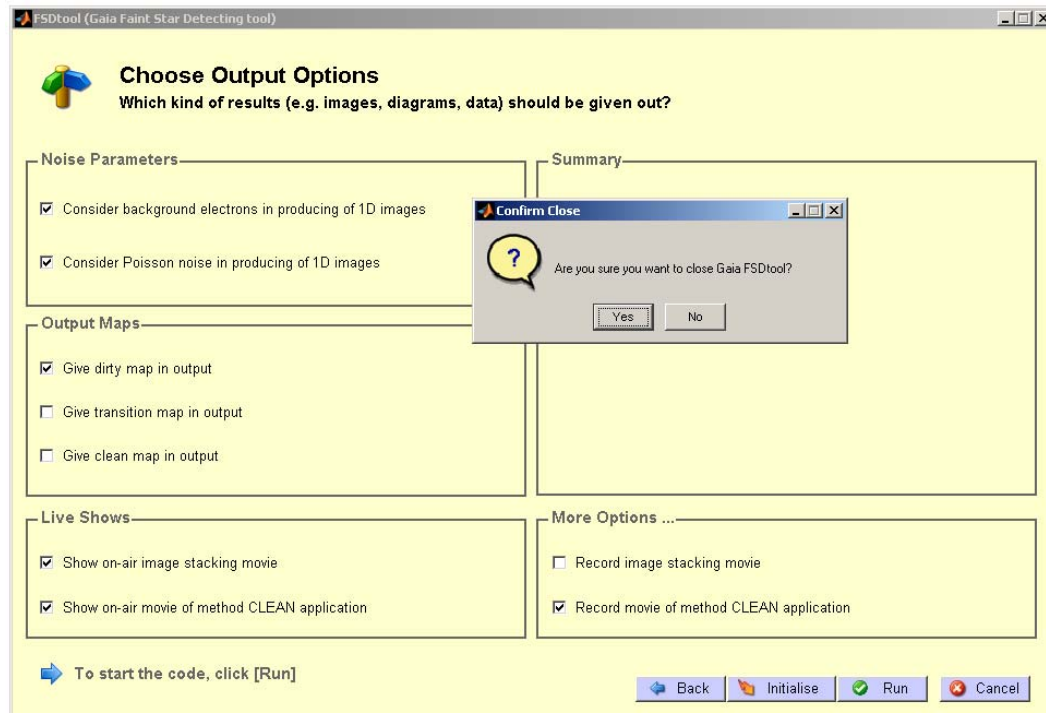


FIGURE 6.4: Output options GUI. The fourth GUI of the Gaia FSDtool, which gives the user the possibility of choosing the options of interest for the output results.

Furthermore, different images like the dirty map, transition map, and clean map can be selected for the output results, and the image stacking and the CLEAN procedure can be chosen to be shown during the process.

6.5 Running the Gaia FSDtool

According to all selected parameters and options, this window displays the result. As an example of the running program, Fig. 6.5 shows the image stacking of 41 scans for a specific position while 38 of the 1D images have already been stacked. Finally, after the detection process and finding the positions and magnitudes of the detected stars, the GUI looks like Fig. 6.6 for an example of one $G = 15$ mag star located at the center of the image.

6.6 Code availability and requirements

The Gaia FSDtool program will be publicly available on the world wide web in the near future. This program was written and tested in MATLAB R2008a and R2008b and should be completely platform independent, though it was only tested under Windows XP, Windows Vista and Linux CentOS 5. As it is an executable standalone program, it can be run without MATLAB installed, but it needs java installed on the target machine. A minimum screen resolution of 1024×768 is preferred for the best display.

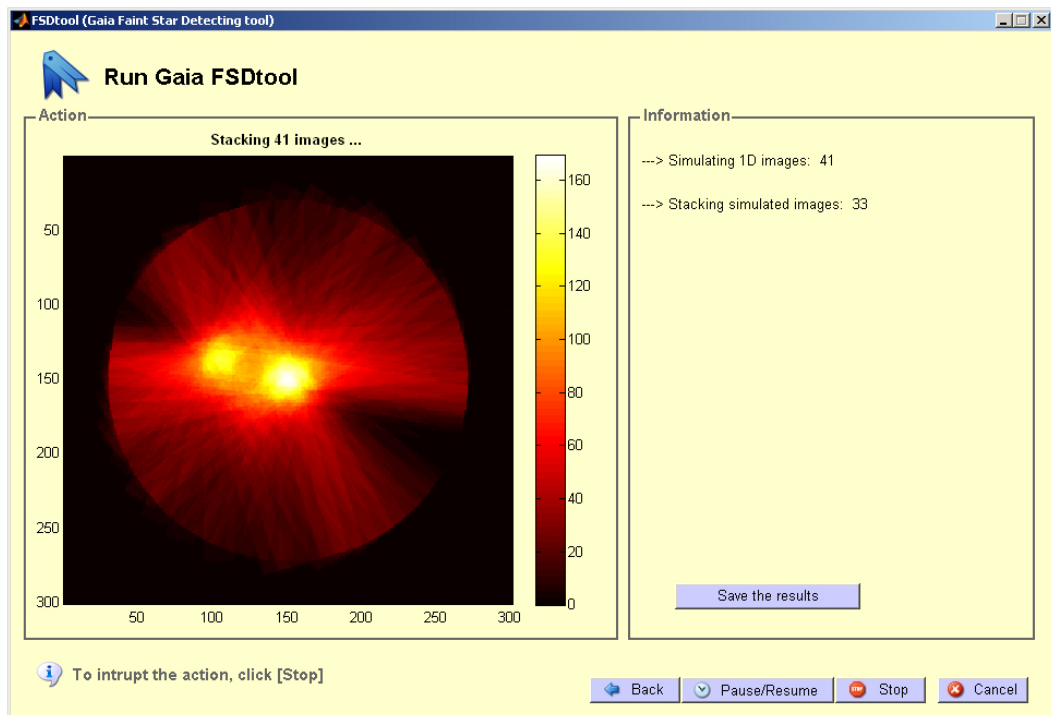


FIGURE 6.5: Final display of the Gaia FSDtool, which here displays the stacking process as an example.

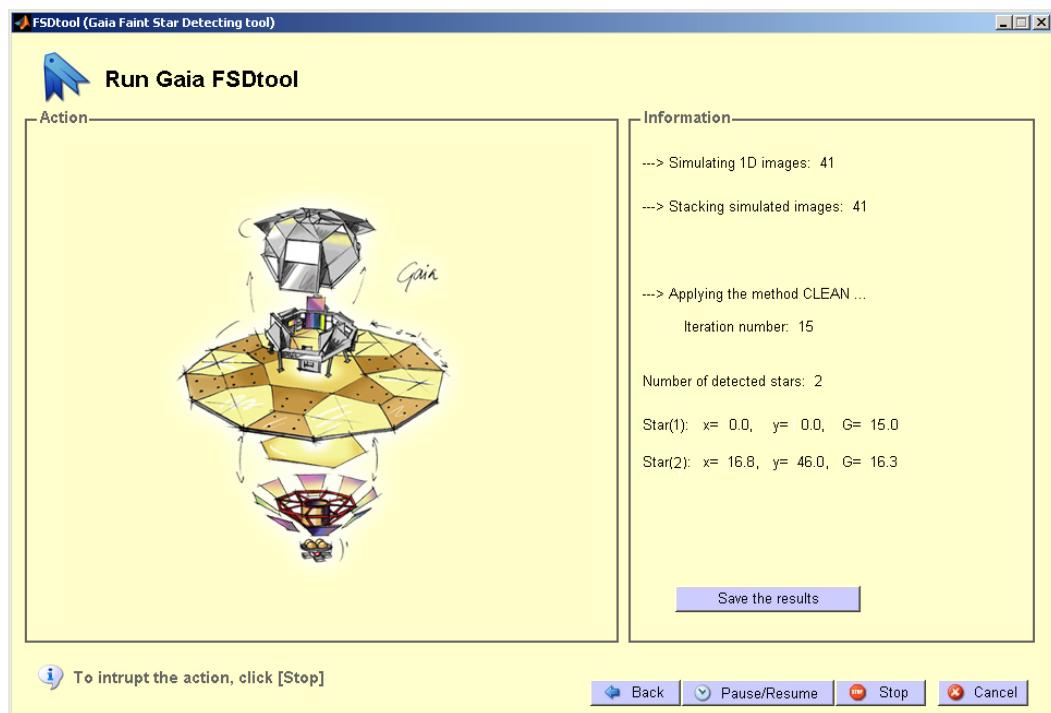


FIGURE 6.6: Final results. The positions and magnitudes of the detected stars are shown in the right panel.

Chapter 7

Discussion and Conclusions

Simulation of the one-dimensional (1D) images from Gaia observations was done according to the known scanning law, CCD characterization and with consideration of the possible noise sources. Each 1D image is, in fact, a $120 \times 240 \mu\text{m}^2$ image with 12 samples of $10 \times 240 \mu\text{m}^2$ size. The number of photoelectrons recorded per sample (assumed to be uniformly distributed in the sample) depends only on the number, positions and magnitudes of the stars in the field of each image. This means that the 1D image simulation does not depend on the observing direction on the sky. On the other hand, according to the Gaia scanning law (Sect. 2.2), the number of scanned 1D images for each specific coordinates on the sky is totally different and therefore different number of scans with different scanning angles were simulated for a few random positions on the sky.

From those simulated 1D images for a particular point on the sky, a two-dimensional (2D) image was produced using different stacking techniques. It was concluded that the bad quality of the produced 2D image can strongly affect the final results of detection and therefore, several well-known approaches were adapted for the simulated images and then examined as candidates methods. The shift-and-add, sub-sampling and drizzle methods were tested and found to be unsatisfactory, although the drizzle method worked much better than the other two. Therefore, the drizzle technique was selected and modified from the original form aimed at extended objects to a form more oriented towards point-source objects. Although this last method seemed to work perfectly, some uncertainty on the final results lead to the discovery of some very tiny features left on the produced images which were causing those uncertainties. Therefore, with the ideas from all those methods of image stacking and by examining different geometrical changes, a new method was found. This later proposed method appeared to result the highest quality 2D image without those extra tiny features.

For detecting stars in the resulting 2D maps, a well-known deconvolution method in radio astronomy, the CLEAN algorithm, was modified for the current purpose. This method is an iterative procedure which uses the known shape of the dirty beam (i.e., the effective PSF produced by the image stacking) for differentiating between the real positions of stars and artifacts produced by the dirty beam. By applying this method on the simulated 2D images, the number of detected stars and information about their positions and magnitudes could be derived according to some pre-defined criteria. A threshold level for detection was

calculated by Monte Carlo experiments for a given probability of false alarm (1 percent was used).

A statistical study was then made of the probability of detection in the case when the true image contains two stars of different magnitudes and separations. The detection was considered to be true only if it gave the correct number of stars (in this case two stars), and their relative positions and magnitudes were correct within certain limits. The main conclusions from this study are:

- The minimum separation for a true detection is about $20 \mu\text{m}$ on the CCD, equal to two along-scan pixels or about 120 mas on the sky. This roughly corresponds to the theoretical resolution of the Gaia telescopes.
- The maximum separation for reliable detection is about $60 \mu\text{m}$ or 350 mas. This is determined by the along-scan size of the 1D windows.
- The maximum detectable magnitude difference is at least 4 mag.
- Disturbing stars at least two magnitudes below the completeness limit of Gaia can be detected close to the target star.
- In the studied cases, the positions and magnitudes of the detected faint disturbing stars are recovered with a precision of about $0.5 \mu\text{m}$ (3 mas) and 0.1 mag, respectively.

Thus image reconstruction with the algorithms described here could be used to detect faint disturbing stars around the Gaia targets. The positions and magnitudes of the disturbing stars determined by the CLEAN method are probably not accurate enough to be directly used for correcting the astrometric and photometric results of the target star, but they can be used as starting values for a least-squares determination of all the parameters involved. The main difficulty with this least-squares solution is to decide on the relevant model (e.g., number of stars) and find a good starting approximation. The image reconstruction technique can provide both.

Finally, a graphical user interface (GUI), Gaia FSDtool, written in MATLAB but which can be run as a standalone program, allows the user to perform simulations and image analysis according to the above-mentioned procedures in an interactive and convenient manner. For example, it can be used for examining the effect of different input parameters on the final results and to investigate more complicated situations involving an arbitrary number of stars.

Bibliography

- [1] Bessell M. S., 1990, *PASP*, 102, 1181
- [2] De Angeli F. et al., 2006, Proposal for the Gaia Data Processing: GAIA-CD-SP-DPAC-FM-030-1
- [3] de Bruijne J. H. J., 2005, Proceedings of the Gaia Symposium "The Three-Dimensional Universe with Gaia" (ESA SP-576), 35
- [4] ESA-RSSD Gaia homepage at <http://www.rssd.esa.int/GAIA>
- [5] Fabricius C., Lindegren L., 2002, *Ap & SS*, 280, 51
- [6] Fruchter A. S., Hook R. N., 2002, *PASP*, 114, 144
- [7] Gaia Parameter Database, Private Communication
- [8] Høg E., Arenou F., Hjorth P., Jørgensen U.G., Mignard, F., Wolff, S., 2003, GAIA-CUO-118
- [9] Högbom J. A., 1974, *A & AS*, 15, 417
- [10] Hobbs D., Lindegren L., Holl B., Lammers U., O'Mullane W., 2008, Proceedings of the International Astronomical Union, Volume 248, 268
- [11] Jordan S., *Astronomische Nachrichten*, 2008, 329, 875
- [12] Kovalevsky J., 2003, *A & A*, 404, 743
- [13] Lindegren L., Perryman M.A.C., Bastian U., Dainty J.C., Høg E., van Leeuwen F., Kovalevsky J., Labeyrie A., Mignard F., Noordam J.E., Le Poole R.S., Thejll P. and Vakili F., 1993, The first proposal for Gaia as an ESA Cornerstone Mission
- [14] Lindegren L., 1998, Point Spread Function for GAIA including aberrations: SAG-LL-025
- [15] Lindegren L., 1998, Simulation of GAIA Scanning of Arbitrary Directions: SAG-LL-026
- [16] Lindegren L., 2005, Proceedings of the Gaia Symposium "The Three-Dimensional Universe with Gaia" (ESA SP-576), 29

-
- [17] Lindegren L., de Bruijne J. H. J., 2005, ASP Conference Series, Vol. 338, *25*
- [18] Lindegren L.; Babusiaux C., Bailer-Jones C., Bastian U., Brown A. G. A., Cropper M., Høg E., Jordi C., Katz D., van Leeuwen F., Luri X., Mignard F., de Bruijne J. H. J., Prusti T., 2008, Proceedings of the International Astronomical Union, Volume 248, *217*
- [19] Mignard F., 2005, ASP Conference Series, Vol. 338, *15*
- [20] Mignard F., 2008, SF2A-2008: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, *23*
- [21] Nurmi P., 2005, Ap & SS, 298, *559*
- [22] Quist C. F., Lindegren L., 1999, A & AS, 138, *327*
- [23] Quist C. F., Lindegren L., Soederhjelm S., 1997, Proceedings of the ESA Symposium Hipparcos, *257*
- [24] Ribak E., 1986, J. Opt. Soc. Am. A, 3, *2069*
- [25] Schwarz U. J., 1978, A & A, 65, *345*