

Treatment of the spurious detections around bright stars by the Gaia satellite.

Author: Jaume Jaume Bestard.

Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Jorge Torra

(Dated: January 19, 2015)

Abstract: We study the environment of bright stars detected and stored by the Gaia satellite. The main goal is to determine how the environment is related to the real stars. This study will allow us to select the best values to identify these detections in the algorithm.

I. THE GAIA MISSION

The satellite was launched on 19 December 2013 from French Guiana. The mission aims to measure the positions and proper motions of more than one billion stars and other objects with unprecedented accuracy, complemented with multi-band photometry and spectrometry. Currently, the Gaia satellite orbits around the L2 Lagrangian point, which is at about 1.5×10^6 km from the Earth, opposing the Sun.

The mission's result, the publication is scheduled for 2022, will be the most complete and accurate three-dimensional map of our Galaxy.

A. The Gaia Focal Plane

The satellite will be scanning the whole sky systematically during its lifetime. The scan rate, around its axis, is 60 arcseconds every second, that is, every 6 hours the Gaia satellite does a complete revolution.

The astrometric measurements are obtained from the combination of two astrometric telescopes, FOV0 and FOV1 (with an angular separation of 106.5 called *basic angle*), converging into a unique astrometric focal plane. The sections of the focal plane, whose structure is shown in FIG. 1, are as follows:

- The Sky Mapper (SM) is in charge of detecting the objects to be measured and determining the way it should be done in the focal plane. The first column of the SM is for telescope FOV0 and the second for telescope FOV1. This 14 CCDs are read out completely.
- The Astrometric Field (AF) is not read out entirely due to the massive data rates and read-out noise. Instead, only small *windows* around the target star are transported across the CCDs in TDI (Time Delay Integration, needed time to transport the charge from one pixel to next one, $1 \text{ TDI} = 1 \text{ pixel} = 982.8 \mu\text{s}$) mode and then these windows are processed. A PSF or LSF fitting to the observed windows gives the input for the initial astrometric data. The AF1 (first column) is used to confirm the detected SM image.

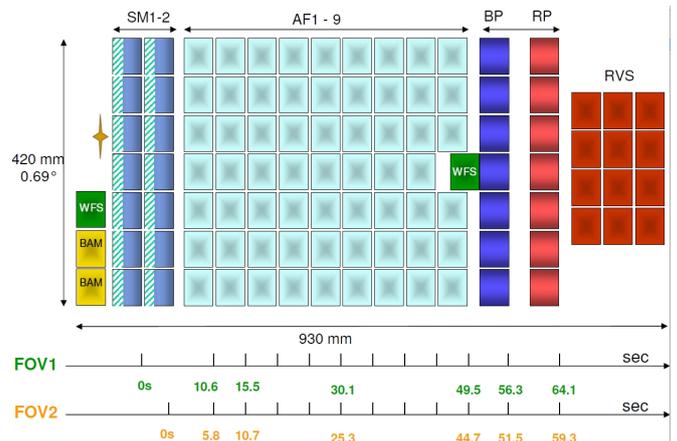


FIG. 1: The Gaia focal plane layout. We will consider the AC (across-scan) and AL (along-scan) directions as vertical and horizontal, respectively. The two bottom lines give the time for targets to reach the different parts of the field from SM1 or SM2.

- The Broad Band Photometer (BBP) is in charge of measuring the flux of every object in 2 different wavelength bands. The photometer consists of two low-resolution prisms dispersing all the light that comes through the FOVs. The first disperser operates in the range 330–680 nm (called RP, Red Photometer) and the second one in the range 640–1050 nm (BP, Blue Photometer).
- The Radial Velocity Spectrometre (RVS) provides radial velocities from spectrophotometer covering the range around CaII line. The RVS operates in windowed mode as the other Gaia instruments.
- The BAM (Basic Angle Monitoring) and WFS (WaveFront Sensors) monitor the behaviour of the optic of Gaia.

II. THE SPURIOUS DETECTIONS

Due to optical diffraction the bright sources create spikes in their near environment that the on-board detection system considers as new sources which take up

a lot of memory in the data base. Stars brighter than at least magnitude $G = 15$ produce these spurious detections. The brightest stars, with $G = 5$ mag, can cause easily hundreds or even thousands of false detections and fainter stars just a few. An example is shown in FIG. 2. We need to eliminate all the spurious detections from the reduction chain.

Our aim is to characterize the distribution of false detections. In the following figures we will use TDI and pixels to refer AL and AC coordinates, respectively. We also will classify them in FOV and CCD rows, where each of the seven rows has 2000 pixels in AC direction.

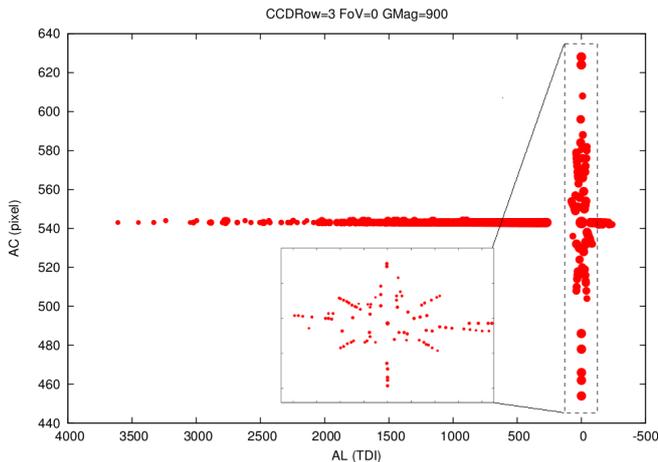


FIG. 2: Typical distribution of spurious detections caused by a bright star. In this case the star has $G = 6.92$. The point's size varies according to the magnitude of the detections. To see the distribution in AC better, a zoom is showed where a few diagonal distributions and only one vertical line can be appreciated.

We can see an important concentration of detections in the AL direction, about 70%. The distribution is symmetric in AC direction and asymmetric in AL due to charge transfer inefficiency (CTI). For this reason, we must treat separately the AC and the AL distributions, and then after and before the bright star.

Once the problem is understood, we have to find certain parameters to define the principal pattern. In our study we have analysed the following parameters:

- Length of every wing, both AC and AL.
- Magnitude and geometric position of the first detections of every wing.
- Links between consecutive spurious detections.

Each parameter is studied as a function of the principal source's magnitude.

To do that we have to develop an algorithm for the identification of spurious detection. It works with general values to determine these parameters for each source. To improve the results we examine them to obtain the best values for an accurate solution.

III. ANALYSIS AND STATISTICS OF THE PARAMETERS

In the previous section we described some parameters that we use in our analysis. In this section we discuss the values for each of these parameters.

Only two figures for one of the couples FOV/CCD of every parameter are shown. We use them as an example to explain how they can serve to improve our algorithm.

A. First detections

We need to know how the first detections are and behave because they are the sole link between the environment and the *parent* source (the one originating the spurious detections). For this reason, we have studied the magnitudes and positions of first detections exhaustively.

One rule has been imposed for this parameter: *the first detection in any direction after a parent will be handled as the first spurious detection.*

Firstly, we analyse their positions and magnitudes in the horizontal (AL) wings.

In most cases the horizontal wings are inside the range *parent position* ± 3 pixels. This depends on apparent velocity of the observed stars over the focal plane. In our algorithm we imposed the ± 3 as a constant, that means, the first detection has to be found inside this range to be considered the first detection. Their horizontal position with respect to the parent and GMag (the magnitude of the stars is encoded on board as $G = 21 - \frac{GMag}{21}$) are our variable parameters in the algorithm.

Initially, the maximum distance where the first detection can be found is 1000 TDIs and 500 TDIs after and before the parent, respectively. The range of magnitudes that restricts the first detections is from 80 to 350 GMag.

In FIG. 3, a regular pattern described by these two parameters can be observed. When the GMag of the parent grows, the position and GMag of the first detection does too. The bottom points of the second figure are out of the typical range due to the initial values in the algorithm.

Looking to FIG. 3, we can consider several regions for these parameters in function of the parent magnitude. Discriminating different values could avoid situations like the ones in the bottom of FIG. 3 and obtain better results.

Next, let us analyse the vertical directions likewise in FIG. 4

The most brightest stars do not present vertical wings in our thickness condition, *parent position* ± 3 TDIs, due to the maximum number of windows in the same AC column.

Once analysed the first detections, different regions of the parent magnitude to impose better values can be defined. TABLE I is only an example of tabulated ranges.

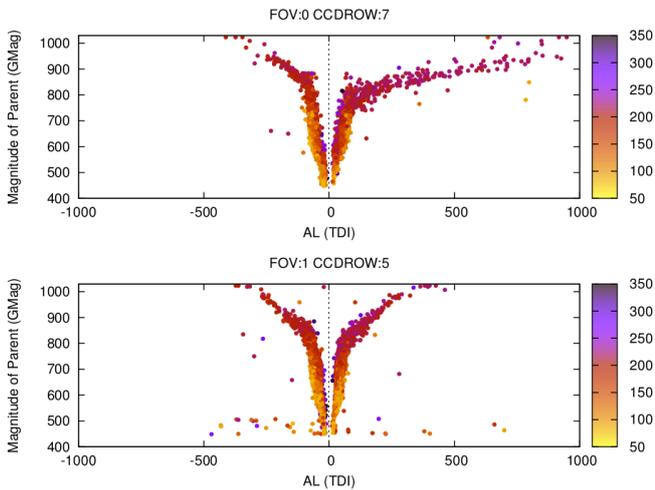


FIG. 3: Magnitude and position for the first detection in the horizontal direction. The color points indicate the magnitude of the first detections in reference to the colourbox.

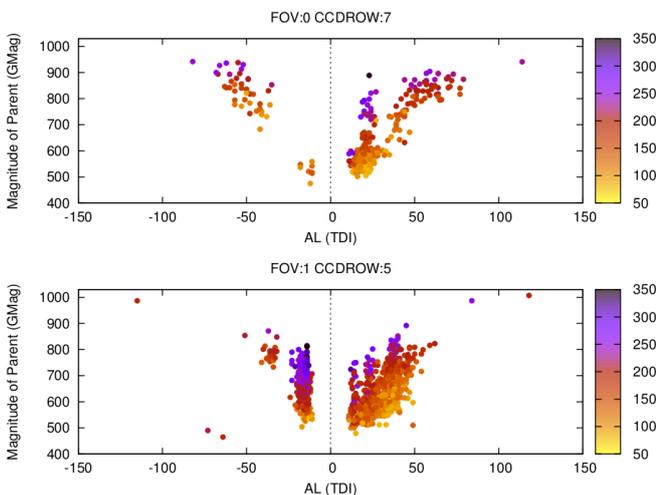


FIG. 4: Magnitude and position for the first detection in the vertical wing. In this direction, we can see that spurious detections are not created if the GMag of the parent does not reach a minimum value, about 500 GMag. The defined maximum distance to consider the first AC detection in the algorithm is 200 pixels.

	Horizontal	Vertical
Range of magnitudes (GMag)	448 – 700	448 – 650
	701 – 1023	651 – 900
	-	901 – 1023

TABLE I: In any case these or any other ranges are still definitive. The team has to decide the best way to define them and what values are better. The minimum GMag considered in our algorithm is 448.

B. Distance from the parent

We analyse the different cases and the used method to handle these length parameters.

In FIG. 5 the behaviour of the horizontal lengths is shown with respect to the magnitude of their parents. The position of the first detection is much better determined than the lengths, because the search of the first detections is direct (first detection after, before, above and below the parent), however, to find the respective lengths depends on all the other spurious detections.

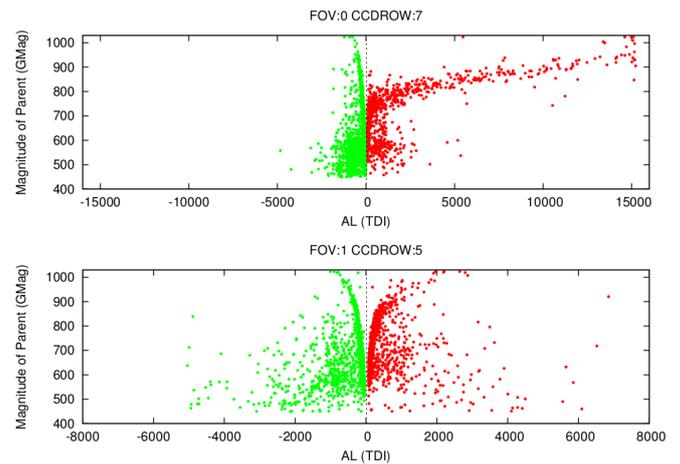


FIG. 5: This parameter shows better the difference between the two focal planes. In FOV:0 CCD:7 some parents are registered with wings about 15000 TDIs and in the other case they have just lengths about 7000 TDIs.

The same procedure is valid for the vertical wings, shown in FIG. 6

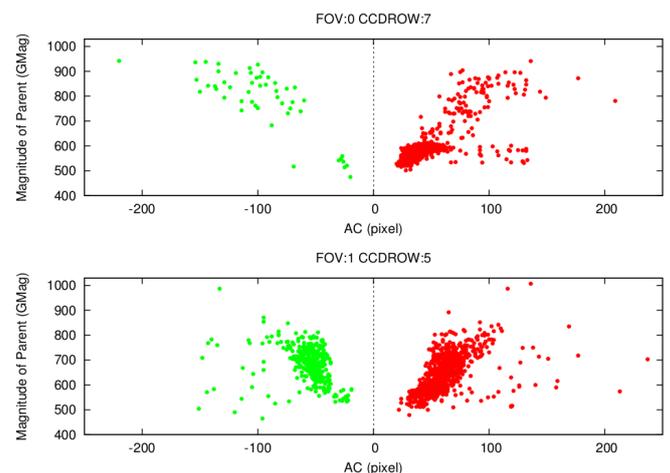


FIG. 6: The position of the last detection found in the AC wings. Initially, in our algorithm a maximum AC distance of 500 pixels with respect to the parent is imposed to find spurious detections.

Next, we explain how the algorithm finds these lengths.

Once the first detection is found, the algorithm starts to check for more detections in the same direction. The following detections must satisfy some criteria:

1. Its GMag has to be inside the range with respect to the previous spurious detection:

$$-64(128) < \text{GMag}_i - \text{GMag}_{i-1} < 64(128)$$

(depending on we are considering an horizontal or a vertical direction, respectively).

2. A maximum gap of 1000 TDIs (horizontal) or 100 (vertical) pixels between two consecutive detections is allowed:

$$AL_i - AL_{i-1} < 1000 \text{ TDIs}$$

$$AC_i - AC_{i-1} < 100 \text{ pixels}$$

3. It has to be inside the thickness condition of both directions:

$$|AC_i(AL_i) - \text{parent position}| < 3 \text{ pixels (TDIs)}$$

These parameters are also adjustable to improve the results. They are studied in the Section III C.

If one spurious detection does not satisfy the previous criteria, the algorithm stops and takes the position of the previous one which satisfies these criteria as the last spurious detection in this direction.

C. Links between consecutive spurious detections

The length of every wing depends on the links between spurious detections: their GMags and positions.

In Section III B, the criteria to identify detections as spurious is already mentioned. These parameters are adjustable and we need to study them to obtain better results for our statistics.

It is important to know that these parameters are defined to avoid irregularities in the principal pattern. Here we will explain the different irregularities that have been detected in our study.

1. Distribution of magnitude

Usually, the magnitude of the spurious detections follows a typical distribution but there are cases where this distribution is partially lost. These cases are shown in FIG. 7.

2. Geometric position of the spurious detections

On the other hand, it is possible to find large position gaps between two consecutive spurious detections, as we show in FIG. 8. For this reason, the second criteria cited in Section III B has to be applied in these cases.

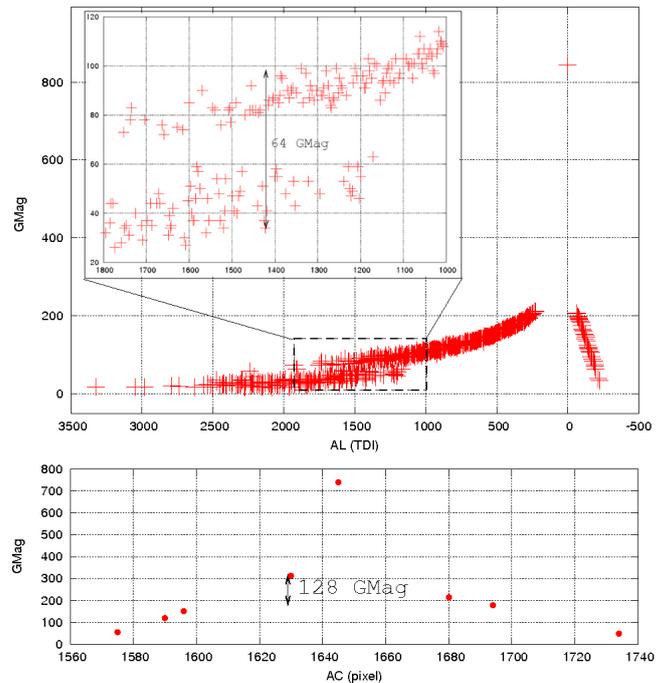


FIG. 7: This figure shows odd cases in horizontal (upper square) and vertical (lower square) directions. Without the first criteria in Section III B the algorithm might not detect the whole wing. The figures do not correspond to the same source.

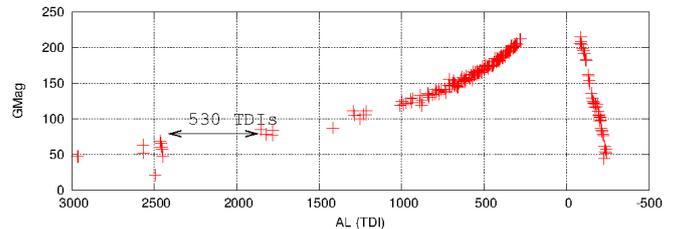


FIG. 8: The distribution of magnitude suggests that the last detections are also spurious, but there is a position gap between them. In this case, it is a gap of 530 TDIs for a parent of 912 GMag. For this reason, 1000 TDIs is the maximum value for the horizontal wings and 100 pixels for the vertical ones.

IV. RESULTS AND CONCLUSIONS

We have analysed the general trends of the spurious detections in Gaia. As a result the best values to eliminate the spurious detections are found.

Currently, our algorithm runs in the Initial Data Treatment (IDT). For efficiency, it works within the Cleaning Cross Matching, which establishes the link between the intermediate data objects and the catalogue of sources. Our algorithm writes the detections that it considers spurious to a blacklist and prevents the creation of new

sources in the working catalogue of IDT and MDB (Main Data Base).

The next tables still not definitive, show main results of this analysis. The general values in the tables are the initial values in our algorithm and the tabulated values are possible results of our study.

First detections before (after) the parent		
Ranges of GMag	General (TDI)	Tabulation (TDI)
448 – 700	500 (1000)	200 (200)
701 – 800		300 (500)
801 – 1023		500 (1000)

TABLE II: Tabulated values for the first detection in the horizontal direction. The numbers in brackets are the distances of the first detections after the parent.

First detections above and below the parent		
Ranges of GMag	General (pixel)	Tabulation (pixel)
448 – 700	200	75
701 – 900		100
901 – 1023		150

TABLE III: Tabulated values for the first detection in the vertical direction.

Wing's length before (after) the parent		
Ranges of GMag	General (TDI)	Tabulation (TDI)
448 – 600	5000 (20000)	300 (300)
601 – 800		500 (1000)
801 – 1023		1500 (16000)

TABLE IV: Tabulated values for the lengths in the horizontal direction. The numbers in brackets are the lengths after the parent.

Wing's length in the vertical direction		
Ranges of GMag	General (pixel)	Tabulation (pixel)
448 – 600	500	80
601 – 700		100
701 – 1023		200

TABLE V: Tabulated values for the lengths in the vertical direction.

We have also analysed a number of cases with different parameters, boundary conditions, and so on, but for this paper we have only included the most common elements of the spurious detections.

Acknowledgments

I express my thanks to Dr. Jordi Torra, Dr. C. Fabricius and the rest of the team of ICCUB-IEEC for their support and guidance during the development of this project.

In addition, I would like to thank my family for their support and comprehension during this investigation and all this time.

-
- [1] J. Portell i de Mora, *Payload Data Handling, Telemetry and Data Compression Systems for Gaia*, (PhD. Thesis, Barcelona, June 2005).
- [2] ESA, *Gaia Data Processing & Analysis Consortium*, GAIA-CD-SP-DPAC-FM-030-2 (April 2007).
- [3] [CF-030] C. Fabricius, *Spurious detections: A*

- proposal for their treatment in IDT and IDU*, GAIA-C3-TN-UB-CF-030-01, (October 2014).
- [4] J. Portell i de Mora, *Detection Classifier for CU3-IDT*, GAIA-C3-TN-UB-JP-069-D, (October 2014).