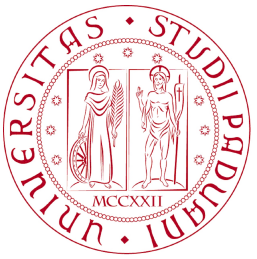

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Astrophysics and Cosmology

Final dissertation

GAIA SPECTROSCOPY AND APPLICATION TO ASTEROID FAMILIES

Thesis supervisor

Prof. Monica Lazzarin

Thesis co-supervisor

Dr. Paolo Tanga

Candidate

Roberto Balossi

Academic Year 2022/23

ABSTRACT

Context. The third data release by the Gaia mission of the European Space Agency (DR3) contains 60 518 asteroid reflectance spectra over the range $374\ nm - 1034\ nm$, representing a huge improvement in spectroscopic data compared to previous surveys available in the literature. Gaia spectra are however affected by systematic deviations in the near infrared when compared to ground-based and template spectra.

Aims. I explored possible correlations of the deviations with several parameters of the observed objects, such as the magnitude and the spectral type, in order to assess the internal consistency of Gaia spectra. In addition, I explored different approaches able to mitigate the deviations and classify Gaia spectra into a consistent taxonomy scheme. Finally, I determined whether if asteroid families could be identified using Gaia spectra alone.

Methods. In order to determine if the deviations were correlating with the magnitude, I quantified them for groups of asteroids of the same spectral type and different luminosity. To explore the correlation with the taxonomy class, I quantified the deviations for groups of objects of different spectral types. I then developed a procedure able to mitigate the deviations, which I tested with the *classy* Python tool. Finally, I tested a color-based taxonomy, which by converting Gaia reflectance spectra in SDSS colors retrieves the spectral type of a specific asteroid given its position on a color-color diagram.
With this taxonomy system I classified Gaia observations in specific regions of the Main Belt in order to identify the Tirela and Watsonia families. By converting the observations in diameters, I then fitted the slopes of the v-shapes in order to constrain the ages of the families.

Results. I found that the systematic deviations in the near infrared correlate with the spectral type, being maximum for S- complex asteroids, intermediate for the C- complex and almost null for M- and P- types.
Classifying Gaia observations with the color-based taxonomy, I identified the Tirela and the Watsonia families, whose memberships resulted to be in very good agreement with the literature. I then constrained the age of the Tirela family, which resulted to be $670 \pm 570\ Myr$ for the inner side and $688 \pm 377\ Myr$ for the outer side. The same analysis for the Watsonia family turned out to be less precise due to the limited number of objects observed by Gaia and resulted in an age of the order of 1 Gyr.

Conclusions. Gaia spectra have been found to be internally coherent, such that asteroids of the same spectral type present similar spectra. In addition, I proved that asteroid families can be recognized using Gaia spectra alone. In particular, the Tirela and the Watsonia families were correctly identified. I then constrained their ages, showing one of the many possibilities in which Gaia spectra can be exploited.

Contents

Abstract

Contents

List of Figures

1 Introduction **1**

2 Asteroids and Asteroid Families **4**
2.1 Asteroids 4
2.1.1 Introduction 4
2.1.2 Dynamical evolution of asteroids orbits 5
2.1.3 Dynamical classification 9
2.1.4 Asteroid spectroscopy 12
2.1.5 Taxonomic classification 13
2.2 Asteroid Families 20
2.2.1 Introduction 20
2.2.2 Identification of asteroid families 21
2.2.3 Age determination: the v-shape method 25

3 The Gaia Mission **31**
3.1 Introduction 31
3.2 The spacecraft 33
3.3 Solar System objects in Gaia Data Release 3 38
3.4 Spectrophotometry of asteroids in Gaia DR3 40

4 Modelling of Gaia spectra **44**
4.1 Introduction 44
4.2 Pre-processing of the reflectance spectra 47
4.3 Correlation with the spectral type 51
4.4 Correction procedure 54
4.5 Classification with *classy* 58
4.6 Color based taxonomy 63

5 Analysis of the Tirela family **67**
5.1 Introduction 67
5.2 Identification of Tirela family members 68
5.3 V-Shape of the Tirela Family 75
5.4 Age Determination of the Tirela Family 77

6 Analysis of the Watsonia family **88**
6.1 Introduction 88
6.2 Identification of Watsonia family members 89
6.3 A possible v-shape for the Watsonia family? 95

7 Conclusion **101**

Bibliography **105**

List of Figures

2.1 *Left panel:* Diurnal Yarkovsky effect for a prograde rotator with spin axis perpendicular to the orbital plane moving around the Sun on a perfect circular orbit. *Right panel:* Seasonal Yarkovsky effect for an object in circular orbit around the Sun with spin axis lying in the orbital plane. Credit: Bottke et al. (2006). 7

2.2 Evolution of the taxonomic schemes from Tholen (left) to Bus-DeMeo (centre) and Mahlke (right). The arrows show the evolution of each spectral type. Credit: Mahlke et al. (2022) 19

2.3 Tirela, Watsonia and some of the largest families plotted against the background of Main Belt asteroids in the $a - \sin(i)$ plane. 21

2.4 $a - 1/D$ plot for the Massalia family. Blue crosses are family members, blue stars are resonant objects and black points are background objects. The thick black lines are the best fit lines for the slopes of the v-shape. Outliers are marked with a blue $\oplus$ symbol, while the red $\oplus$ marks the position of (20) Massalia. Credit: AstDys. 27

2.5 Asteroid families ages. See text for better description. (Credit: AstDys). 29

3.1 Illustration of the focal-plane assembly with a human hand for scale comparison. (Credit: Gaia Collaboration (2016)). 34

3.2 The procedure to compute mean reflectance spectra illustrated for the case of (1459) Magnya. A more complete description of the figure is found in the text. Credit: Gaia Collaboration (2022a). 41

3.3 Mean reflectance spectrum of (4) Vesta. Black symbols have RSF=0, yellow symbols have RSF=1 and red symbols have RSF=2. Plot obtained using the *classy* Python package. 42

4.1 Gaia spectra of the S- type (39) Laetitia, in blue, and the L- type (729) Watsonia, in orange. Despite being of two different spectral types, the two spectra appear to be very similar due to the deviations in the NIR. 46

4.2 Example spectra of (3) Juno on the left and (31) Euphrosyne on the right. Gaia data (in black) are found to be systematically redder than the ground based SMASS spectra (in light blue). 47

4.3 Left panel: Gaia observations of S- type asteroids between 15 and 17 *mag* compared to Mahlke template spectrum. Right panel: The processed mean spectrum compared to the Mahlke template. 48

4.4 Relative difference from the Mahlke template spectrum for S- type asteroids subdivided in the four luminosity classes. Data at the same wavelength have been shifted by 5 *nm* along the x-axis for a better visualization. The black dotted line identifies the zero level. 50
4.5 Same as Figure 4.4, but for C- type asteroids. 51
4.6 Same as Figure 4.4, but for P- type asteroids. 52
4.7 Relative difference from the Mahlke template spectrum for asteroids of different spectral types. Data at the same wavelength have been shifted by 2 *nm* along the x-axis for a better visualization. The black dotted line identifies the zero level. 53
4.8 Example spectra of (3) Juno (left) and (31) Euphrosyne (right) with highlighted the maximum of the reflectance (red star) and the slope in the red (black line). The two parameters are used in the correction procedure. 55
4.9 Slope - maximum reflectance diagram of asteroids observed by Gaia with $SNR >$ 100 and classified by Mahlke (black letters). The seven regions used to define the corrections are reported on top of the diagram in different colors with the relative names. 56
4.10 Corrected Gaia spectra (in black) compared to the original Gaia spectra (in red), to SMASS spectra (light blue, when present) and to Mahlke template spectra (black dotted line) for numbered asteroids between 1 and 34. 59
4.11 Corrected Gaia spectra (in black) compared to the original Gaia spectra (in red), to SMASS spectra (light blue, when present) and to Mahlke template spectra (black dotted line) for numbered asteroids between 35 and 75. 60
4.12 *Left panel*: the confusion matrix between Mahlke asteroids and uncorrected Gaia spectra classified by *classy*. *Right panel*: same as the left panel, but with corrected Gaia spectra. 62
4.13 The SDSS filters response curves without the atmospheric transmission. The filters are, from left to right, u, g, r, i and z. Gaia reflectance spectra cover the interval between 3740 Å (in the u filter) and 10340 Å (in the z filter). 64
4.14 Color distribution for all the Gaia observations (black dots) and the asteroids in sample of the outer Main Belt centered on the Tirela Family (red stars). The yellow dashed lines separate the three major taxonomic groups. 65

5.1 The number of asteroids per magnitude for the complete sample (blue), the Gaia sample (yellow), the Milani classification (red, left panel) and the Nesvorny classification (red, right panel). 69
5.2 Distribution in the $a_p - 1/H$ space for the objects of the complete sample (in blue), the Gaia sample (in yellow), the Milani classification (in red, left panel) and the Nesvornny classification (in red, right panel). 70
5.3 *Left panel*: Albedo distribution for the objects of the complete sample (in blue), of the Gaia sample (in yellow) and of the Milani classification (in red). *Right panel*: Spatial distribution in the $a - e$ plane for the objects of the complete sample divided according to the albedo. 71

5.4 Number of objects per spectral type for Gaia observations classified by the color-based taxonomy. 72
5.5 Spatial distribution in the $a - e$ space for Gaia observations in the Tirela region with relative taxonomic types. 73
5.6 Normalized distribution of the asteroids in the Tirela region. L- types are shown in red, S- types in green and C- types in blue. 74
5.7 *Left panel*: Distribution in the $a-1/D$ plane for family members determined from Gaia observations (blue points), resonant objects (red crosses) and interlopers (black circles). *Right panel*: Distribution in the $a-1/D$ plane for family members included in the Milani classification (in blue) and not included (in red). 76
5.8 Illustration of the procedure to fit the v-shape of the Tirela family for N varying between 8 and 22. The blue points are the family members, the red dashed lines indicate the division in bins, the black circles are the points used to fit the sides and the black lines are the fitted lines. 79
5.9 Same as in Figure 5.8, but with the small cluster of seven objects in the outer side included in the fit. 80
5.10 The slopes retrieved from the fit for $8 < N < 22$ for the case without the small cluster (left panel) and with the cluster included (right panel). The blue indicates the inner side and the orange the outer side. 82
5.11 The v-shape of the Tirela family. In blue the family members, in black the lines fitting the sides of the v-shape. 83

6.1 *Left panel:* Absolute magnitude distribution for asteroids in the complete sample (in blue), in the Gaia sample (in yellow) and in Milani classification (in red). *Right panel:* Distribution in the $a - 1/H$ plane of asteroids in the complete sample (yellow circles), in the Gaia sample (in blue), in Milani classification (in red) and in common between Gaia observations and Milani classification (in black). 91
6.2 Albedo distribution for the objects of the complete sample (left panel), of the Gaia sample (middle panel) and of the Milani classification (right panel). 92
6.3 *Left panel:* Spectral types of the asteroids observed by Gaia in the Watsonia region. The red indicates objects with $\rho_V > 0.08$, the blue objects with $\rho_V < 0.08$ and the grey objects without albedo. *Right panel:* same as the left panel, but with the orange indicating objects included in the Milani classification and the green objects not included. 93
6.4 Spectral types of the asteroids observed by Gaia in the Watsonia region, but without considering the second more probable class for the objects initially classified as U- types. The red indicates objects with $\rho_V > 0.08$, the blue objects with $\rho_V < 0.08$, the grey objects without albedo. 94
6.5 *Left panel:* Distribution in the $a - 1/D$ plane for family members determined from Gaia observations (blue points) and interlopers (black crosses). *Right panel:* Distribution in the $a - 1/D$ plane for family members included in the Milani classification (in blue) and not included (in red). Each object is marked with the letter corresponding to its spectral type. 96

6.6 The family members (black points) in the $a - 1/D$ plane with overplotted the v-shapes corresponding to 0.5 Gyr (in blue), to 1.0 Gyr (in red) and to 1.5 Gyr (in green). 99

Chapter 1

Introduction

Gaia is an extraordinary mission of the European Space Agency (ESA) launched in 2013 whose ambition is to chart a three-dimensional map of the Milky Way by observing one billion of its constituent stars. However, Gaia has been built in order to detect every point source of the sky, among which unresolved galaxies, quasars and Solar System objects. Astrometry and low-resolution spectro-photometry data are obtained for all sources, while high resolution spectra are obtained for the brightest stars in order to determine their radial velocities.

Gaia has now reached its third data release (Gaia DR3) which contains astrometry and broad band photometry for more than 1.8 billion objects collected during 34 months of satellite operations, from August 5, 2014 to May 28, 2017. In addition, Gaia DR3 contains over 3 million astrometry measurements for more than 150 000 Solar System objects and 60 518 asteroid reflectance spectra in the range $374 - 1034\ nm$. Gaia DR3 represents a substantial improvement compared to previous surveys: at the moment it was released, only 7600 asteroid spectra were in fact available in the literature.

Gaia asteroid reflectance spectra are however affected by systematic deviations in the near-ultraviolet (NUV) and in the near-infrared (NIR) when compared to ground-based and template spectra. Deviations in the NUV have been found to be related to the mean solar analogue spectrum used in the data reduction (Tinaut-Ruano et al., 2023), while, to date, deviations in the NIR have not yet an explanation. However, understanding how the deviations behave is essential before attempt to conduct any scientific research based on the spectra. In particular, it is necessary to assess if Gaia spectra are internally coherent, such that asteroids of the same spectral type present similar spectra. This in fact would allow to compare them with literature and to quantify the deviations. In this work I therefore explored possible correlations of the deviations with several parameters,

such as the magnitudes and the spectral types of the observed asteroids, in order to asses the internal consistency of the spectra.

Asteroid families are groups of asteroids orbiting around the Sun on very close trajectories and usually sharing their physical properties, which have been formed after past asteroid collisions. For this reason, they are perfect natural laboratories to investigate the collisional and the dynamical evolution of the Main Belt.
Asteroid families are usually identified in the space of proper elements (semi-major axis a, eccentricity e and inclination i) as dense clusters of asteroids over a more diffused background. However, the large dataset of asteroid reflectance spectra published in Gaia DR3 provides the opportunity to explore asteroid families from a spectroscopic point of view. In this work, I therefore tried to assess the feasibility of recognizing asteroid families using Gaia data alone. In particular, I focused on the Tirela and the Watsonia families, which both consist of L- type asteroids.
The Tirela family represents the perfect benchmark to test Gaia spectra. It is in fact a large family located in the outer Main Belt, where the background asteroids belong to the C- complex. The difference between the spectral types of the family members and the background should in principle allow to identify the family using spectroscopic observations alone.
The Watsonia family instead is smaller than the Tirela family and located in the middle of the Main Belt, where the difference between the spectral types of the family members and the background is less marked. The Watsonia family thus represents a more complicate case than the Tirela family.
After having identified the families from Gaia spectra, I then constrained their ages implementing the v-shape method proposed by Milani et al. (2014) and Spoto et al. (2015). By using NEOWISE albedos, I converted the observations in diameters and then I fitted the v-shape in the $a - 1/D$ plane with a procedure I developed. By combining the slopes retrieved from the fit and a Yarkovsky calibration, I then constrained the ages of the families.

This thesis is subdivided in seven chapters, including the present chapter (Chapter 1) of introduction to the research topic. This chapter describes the objectives of the study and a brief thesis outline. Chapter 2 provides a review of past literature related to asteroids and asteroid families. Chapter 3 describes the Gaia mission, the DR3 and the asteroid reflectance spectra published in it. Chapter 4 describes the attempts to find a dependency of the deviations in the NIR with some asteroids parameters. A correction procedure mitigating the deviations is proposed, while at the end of the chapter a taxonomy based

on color indexes computed from Gaia spectra is described. Chapter 5 describes how the Tirela family has been identified using Gaia spectra alone and how its age has been constrained with the v-shape method. Chapter 6 reports a similar analysis conducted for the Watsonia family. Chapter 7 summarizes the conclusions and the research contributions of the thesis.

Chapter 2

Asteroids and Asteroid Families

2.1 Asteroids

2.1.1 Introduction

4.56 billion years ago the Solar System consisted of a rotating disk of gas and dust, known as protoplanetary disk, surrounding the Sun. Planets formed from the disk, while the left-over material condensed in small bodies, progenitors of the current asteroid population. Physical evidences of these formation phases are not preserved on the Earth nor on other planets because of the geological processes that occurred on them, such as erosion, metamorphism and remelting. Some asteroids instead are very small bodies that have changed very little since their formation. They are the oldest material in the Solar System and they provide a window to probe its early evolutionary phases. For example, important information on the chemical mixture of the protoplanetary disk and on the conditions from which the Earth formed can be retrieved from the composition and structure of meteorites, which are fragments of asteroids that survived the passage through the atmosphere and struck the surface of the Earth (Pfalzner et al. (2015), Libourel et al. (2017)).

Asteroids played a significant role in the history of the Earth. They are in fact believed to have delivered to our planet the building blocks of life, such as carbon-based molecules, volatile materials and, especially, water (Daly and Schultz (2018), Rivkin et al. (2002)). Since asteroids can collide with the Earth, they have significantly modified the planet's biosphere during the past. Statistically, there is an inverse relation between the size of the

object and the frequency of impact events (Donovan et al. (2017) and references therein). The Earth is in fact bombarded on a daily basis by tons of interplanetary material, which is usually so small that is destroyed by the atmosphere creating meteors. Statistically, every century an asteroid of the order of few tens of meters in diameter is expected to collide with the Earth. Two recent examples that have been observed are Chelyabinsk in 2013 and Tunguska in 1908, for which both objects had a diameter of about 20 meters. Collisions with objects larger than few kilometers are extremely rare and happen approximately once every tens of million years. These events cause global disasters which may bring to the extinction of almost all life forms on the planet, as happened 65 million years ago with the dinosaurs.
Despite the probability of impacting the Earth and causing serious damages is remote, asteroids need to be continuously monitored to prevent such events. This is a further reason to study in detail orbital and physical properties of these small bodies.

2.1.2 Dynamical evolution of asteroids orbits

Since the discovery of (1) Ceres by Giuseppe Piazzi in 1801, more and more asteroids have been discovered. In particular, the last decades have seen an exponential increase in the number of discovered asteroids thanks to automatized ground-based surveys. To date, June 2023, almost 1,300,000 asteroids are known.
Most asteroids are located in the Main Belt, between the orbit of Mars and Jupiter, and in the Kuiper Belt, beyond Neptune (Section 2.1.3). Their distribution in these regions is not uniform, but it is shaped by orbital resonances with planets.

Resonances

Orbital resonances are defined as any system of two or more objects orbiting the same primary whose orbital mean motions are in a ratio of small whole numbers (Peale, 1976). They are divided in *mean motion*, *secular* and *diffusive resonances*.
Mean motion resonances occur every time that the ratio between the orbital period of an asteroid and of a planet is close to a ratio of small integers. In the Main Belt this usually occurs with Jupiter, while in the outer Solar System they are related to Neptune (Tiscareno and Malhotra, 2008). The Main Belt is crossed by numerous mean motion resonances which create dips in the distribution of semi-major axes known as Kirkwood gaps.

Secular resonances are instead related to the commensurability of the frequencies of variation of the eccentricity and of the inclination of orbits. In the Main Belt the major secular resonances are ν_5, ν_6 and ν_{16} with Jupiter and Saturn (Morbidelli and Henrard, 1991).
In addition to mean motion and secular resonances, the Main Belt is crossed by numerous diffusive resonances. These are high order resonances with Jupiter, three body resonances with Jupiter and another planet and mean resonances with planets other than Jupiter. Despite being very narrow, these resonances affect the orbital parameters of the asteroids located inside them by changing their values of eccentricity and orbital inclination, resulting in a slow chaotic evolution (Nesvorný et al., 2002).
Orbital resonances play an important role in the dynamical evolution of asteroids orbits. An object in a resonant position experiences close encounters with a major planet more often than an object located in another position. The gravitational interaction with the planet results in a chaotic variation in the orbital parameters of the asteroid, thus providing a potential escape route to the inner Solar System (Granvik et al., 2017). Resonant bodies can therefore be transported to Mars-, Earth-, and Venus- crossing orbits, thus replenishing the population of Near Earth Asteroids (NEAs), or can instead fall into the Sun.

Yarkovsky Effect

Orbital resonances are not the only mechanism able to modify the proper elements of an asteroid. The Yarkovsky effect, discovered by Ivan Osipovich Yarkovsky around 1900 (Beekman, 2005), is a non gravitational effect that causes a small force on a rotating object leading to long-term variations in its orbit. This force is caused by the anisotropic emission of thermal photons, which carry momentum, and it is most efficient for meteoroids and small asteroids with diameters between 10 cm and 10 km (Bottke et al., 2006). For larger objects, the force due to the photons is too small to cause significant variations of the orbit in the long term.
The Yarkosvsky effect was first measured for the 500 m diameter asteroid (6489) Golevka, for which a drift of 15 km in its semi-major axis was observed between 1991 and 2003 compared to its predicted position (Chesley et al., 2003). Precise Yarkovsky measurements are now available for few NEAs, such as (101955) Bennu or (99942) Apophis, thanks to radar observations, astrometric data or in site space missions, while to date no Yarkovsky measurements is available for any Main Belt object (Vigna et al., 2018).
Two components of the Yarkovsky effect can be recognized: the *diurnal* and the *seasonal components*.

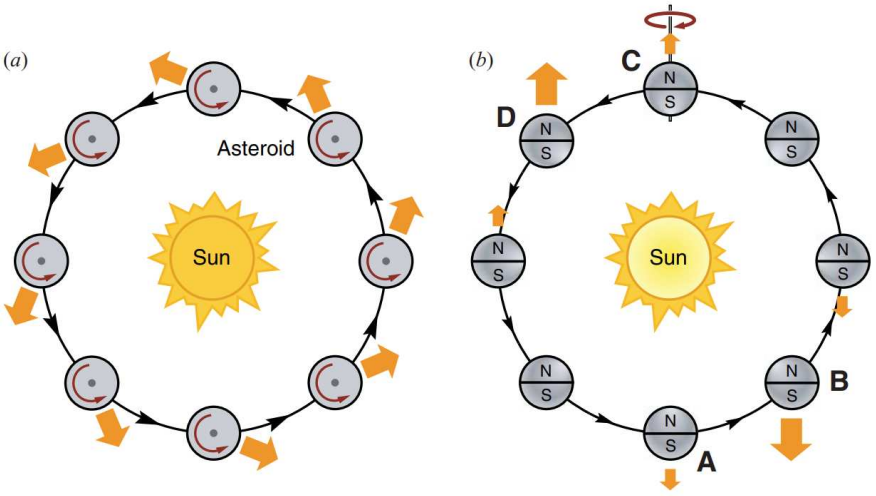

FIGURE 2.1: *Left panel:* Diurnal Yarkovsky effect for a prograde rotator with spin axis perpendicular to the orbital plane moving around the Sun on a perfect circular orbit. *Right panel:* Seasonal Yarkovsky effect for an object in circular orbit around the Sun with spin axis lying in the orbital plane.
Credit: Bottke et al. (2006).

The idea behind the diurnal component is that on a rotating body illuminated by the Sun the surface is warmed during the day, while it cools at night. The thermal inertia of the surface however causes a delay so that the hottest part of the body is its afternoon side rather than its subsolar point. The disalignment between the directions of absorption and re-emission of radiation yields to a net force along the direction of motion of the orbit. The sign of this force depends on the sense of rotation. If the object is a prograde rotator the force is directed along the motion of the orbit and the semi-major axis constantly increases. If instead the object is a retrograde rotator the force is directed on the opposite side of the motion of the orbit and the semi-major axis decreases.
The magnitude of the diurnal Yarkovsky effect depends on a lot of factors, among which the distance from the Sun, the rotation speed, the inclination of the rotation axis and the surface properties. All together these factors make difficult to estimate the magnitude of the force for a specific object. Intuitively, asteroids with virtual infinite or zero rotation periods experience no diurnal Yarkovsky force. This component is usually dominant for objects larger than about 100 meters and up to a few tens kilometers.
The left panel of Figure 2.1 shows the Yarkovsky diurnal component acting on a prograde rotator with spin axis perpendicular to the orbital plane moving around the Sun on a perfect circular orbit. In this case the force causes a constant outside drift of the semi-major axis.

The existence of the seasonal Yarkovsky component was found while searching for the cause of the secular decay of the orbit of the LAGEOS satellite (Rubincam, 1990).
The seasonal Yarkovsky effect is illustrated in the right panel of Figure 2.1 for an object in circular orbit around the Sun with spin axis lying in the orbital plane. It is the component of the force lying along the spin axis that gives rise to the seasonal effect. For example, in figure, when the object is in position A the northern hemisphere is warmed by the Sun. However, because of the thermal inertia, the maximum irradiation from the same hemisphere happens when the body is in B. The same happens for the southern hemisphere in C and D. The seasonal Yarkovsky effect thus causes a constant orbital decay which, unlike the diurnal component, is independent from the sense of rotation of the object. The effect also changes the other proper elements of the orbit, since for example it tends to circularize small eccentricities orbits.
The magnitude of the effect depends on the distance from the Sun and on the inclination of the spin axis with respect to the orbit. The effect vanishes for objects with rotation axis normal to the orbital plane. The seasonal component of the Yarkovsky effect is dominant for objects smaller than 100 meters.

YORP Effect

The reflection and re-emission of sunlight from the surface in addition to be responsible for the Yarkovsky effect also produces a net thermal torque on asteroids with irregular shapes. This torque integrated in time changes the rotation state of the asteroid, spinning up or down its rotation rate and modifying the inclination of the spin axis. This second-order variation of the Yarkovsky effect is known as the Yarkovsky-O'Keefe-Radzievskii-Paddack effect, or YORP for short (Rubincam (2000), Bottke et al. (2006)).
Depending on the physical properties of the object, in the long term the YORP effect can make the rotation rate and the orbital inclination changing chaotically or regularly. The first case is that of *tumbling asteroids* (Pravec et al., 2005), while the second is related to the concept of the *YORP cycle*. In certain geometric configurations, a fast-spinning asteroid experiences a gradual increase in its obliquity as well. When this becomes too large, the thermal torque changes sign and the asteroid begins to spin down. At a certain point, the torque changes again and the asteroids spins up again, in a continue cycle. If the spin becomes high enough, sudden reshaping of the object is also possible.
In addition to controlling the long term evolution of the rotation state of an object, the YORP effect also regulates the magnitude and the direction of the Yarkovsky effect, which depends on the rotation parameters.

2.1.3 Dynamical classification

Asteroids are divided into dynamical groups according to their proper elements a, e and $\sin(i)$, where a is the semi-major axis, e the eccentricity and $\sin(i)$ the inclination of the orbit, and their aphelia $Q = a(1 + e)$ and perihelia $q = a(1 - e)$.
In the following part of this section, the main dynamical subdivisions are listed, in order of crescent distance from the Sun.

Near Earth Asteroids (NEAs) are the closest asteroids to the Sun. In terms of orbital elements, they present a perihelion distance less than 1.3 *AU*. NEAs are divided into four subgroups (Atira, Aten, Apollo and Amor) according to their perihelia, aphelia and semi-major axes.
Atiras are the innermost asteroids, whose orbits are contained entirely within the orbit of the Earth. They present $a < 1.0$ *AU* and $Q < 0.983$ *AU* and are the least populous subgroup of NEAs. *Atens* are Earth crossing asteroids with semi-major axes smaller than the Earth's one ($a < 1.0$ *AU* and $Q > 0.983$ *AU*). *Apollos* are other Earth-crossing asteroids, but with semi-major axes larger than the Earth's one ($a > 1.0$ *AU*, $q < 1.017$ *AU*). Finally, *Amors* are Earth approaching asteroids that orbit between the Earth and Mars. NEAs may also be classified as *Potentially Hazardous Asteroids* (PHAs) according to parameters that measure the potential threat of an object during close encounters with our planet. In particular, all asteroids with an Earth Minimum Orbit Intersection Distance (MOID) of less than 0.05 *AU* and an absolute magnitude H less than 22 (which approximately translates to 140 m in diameter) are considered PHAs.

Mars Crossing Asteroids (MCAs) move on trajectories crossing Mars orbit and present perihelia included between Mars perihelion and aphelion. This group contains the few Martian trojans discovered until now. These are objects sharing their orbit with Mars, preceding (or following) it by 60° in the Lagrangian point L_4 (or L_5).
MCAs are more numerous than NEAs and they account for about the 2% of all the asteroids discovered. Their population is continuously resupplied by the diffusive resonances crossing the Main Belt (Section 2.1.2), for example via mean-motion resonances with Mars or the ν_6 secular resonance.

Main Belt Asteroids (MBAs) are objects located inside a torus shaped regions between the orbits of Mars and Jupiter, known as *Main Belt*. This region is estimated to contain between 1.1 and 1.9 million asteroids larger than 1 km and millions of smaller ones. The distribution of asteroids in the Main Belt is shaped by orbital resonances. As a result,

MBAs are divided in a large number of groups, sub-groups and families, some of which are illustrated in the following.
Hungaria asteroids are characterized by semi-major axes between 1.78 AU and 2.00 AU, eccentricities smaller than 0.20 and inclinations between 16° and 34°. They are named after (434) Hungaria.
Inner Main Belt (IMBs) asteroids are characterized by semi-major axes between 2.06 AU and 2.50 AU, boundaries defined by the 4:1 and the 3:1 mean motion resonances with Jupiter. Different subgroups are found within this region, such as the Phocaea and the Vesta families.
Middle Main Belt (MMBs) asteroids have semi-major axes between 2.50 AU and 2.82 AU, defined by the 3:1 and the 5:2 mean motion resonances with Jupiter. The Eunomia family is located inside this region.
Outer Main Belt (OMBs) asteroids are located between the 5:2 and the 2:1 mean motion resonances, with semi-major axes between 2.82 AU and 3.28 AU. Within this region the Koronis, Eos and Themis families, among others, are identified.
Cybele asteroids are in 7:4 orbital resonance with Jupiter and they are characterized by semi-major axes between 3.28 AU and 3.70 AU, eccentricities less than 0.3 and inclinations less than 25°. This group consists of about 2000 members and few collisional families.
Hilda asteroids are the last outpost of the Main Belt. They are in a 3:2 resonance with Jupiter and they present semi-major axes between 3.7 AU and 4.2 AU, eccentricities greater than 0.07 and inclinations less than 20°. They move on elliptical orbits that bring them close to Jupiter's orbit in correspondence of the Lagrangian points L_3, L_4 and L_5.

Jupiter Trojans are asteroids sharing their orbits with Jupiter in a 1:1 resonance and placed 60° ahead and behind it. Jupiter Trojans are distributed in two elongated regions around the two Lagrangian points L_4 and L_5. To date, June 2023, there are more than 10 thousands Jupiter Trojans discovered.
The term "trojans" is generally applied to every small body that is co-orbital with a planet. In addition to Jupiter and Mars, trojan asteroids are known to exist around Neptune, Uranus and the Earth, even if with much smaller populations than Jupiter Trojans.

Centaurs are planetesimals located between the orbits of Jupiter and Neptune, even if there is no clear definition of this group from a dynamical point of view. Centaurs usually move on unstable orbits due to interactions with the giant planets. They are probably transitory objects that have been scattered out of the Kuiper Belt and that will be soon ejected from the Solar System or become Jupiter family comets. Despite their lifetime is

just few millions years, it is estimated that there could be up to few millions Centaurs in the Solar System.
No Centaur has yet been photographed up close, therefore their physical properties are still quite unknown. They typically exhibit characteristics of both asteroids and comets.

Trans Neptunian Objects (TNOs) include every object beyond Neptune's orbit ($a > 30\ AU$). TNOs include a large variety of objects, subdivided in the following smaller dynamical groups.
Kuiper Belt Objects (KBOs) are located between 30 AU and 50 AU, despite this outer boundary is not well defined. KBOs present low inclinations and small eccentricities. The Kuiper Belt shares some similarities with the Main Belt, but it is far more extended and massive, and it includes most of the known dwarf planets, such as Pluto, Haumea, Orcus, Makemake and Quaoar.
KBOs are in turn divided in two subgroups: classical KBOs (Cubewanos) and resonant KBOs. Cubewanos, which take their name from the first object of this kind discovered, 1992 QB1, move along orbits which are not controlled by resonances with planets. Resonant KBOs, as the name suggest, are objects in mean motion resonance with Neptune (1:2, 2:3, etc). Among resonant KBOs a population of Plutinos is recognized, taking their name from Pluto. They are in 3:2 resonance with Neptune and present high inclinations and semi-major axes around 39.4 AU.
Scattered Disk Objects (SDOs) are characterized by wide orbits with semi-major axes ranging between 50 AU and 675 AU, perihelia greater than 30 AU, eccentricities up to 0.8 and high inclinations. It is believed that these objects are KBOs scattered on distant and very elliptical orbits after a past close encounter with Neptune.

Oort Cloud Objects (OCOs) are located within the Oort Cloud, a spherical cloud of comets lying outside the heliosphere at distances ranging about from 10000 AU to 100000 AU. It is estimated that the Oort Cloud contains millions of cometary nuclei, stable since at those distances the Sun radiation is very weak. The Oort Cloud is loosely bound to the Sun and it is easily perturbed by the gravitational pull of nearby stars or the Milky Way. Perturbed cometary nuclei may start moving towards the inner Solar System, thus resupplying the population of comets.

2.1.4 Asteroid spectroscopy

The incident solar flux arriving on an asteroid surface is split into two components, the reflected and the absorbed part, whose ratio depends upon the value of the albedo. The absorbed radiation warms up the surface of the object, which thus emits as a black body. The reflected sunlight instead is slightly modified by the material covering the surface, which adds to the spectrum some features coming from electronic and vibrational transitions. From Earth both the modified reflected sunlight and the black body emission are observed. The first dominates from the UV to the NIR (0.35 - 2.5 μm), while at longer wavelengths the contribution of the black body emission becomes detectable (Burbine, 2016).
Reflectance spectroscopy in the visible and in the near infrared analyzes the spectrum of the reflected sunlight and it is the preferred technique to investigate asteroids surfaces composition, which in turn provides information about the conditions and the processes present during the early phases of evolution of the Solar System. When conducting this type of analysis it is very important to disentangle between the spectral features related to the formation conditions and the ones due to the subsequent evolution of the asteroid. There are three main processes able to alter the primordial surface of an object: collisional evolution (craters or impacts with other asteroids), space weathering (due to the irradiation of cosmic rays and solar wind and collisions with micrometeorites) and the evolution of the surface regolith structural and thermal properties (Gaia Collaboration (2022a) and references therein).
Asteroids are composed by a mixture of minerals, the most common of which are olivine, pyroxene, iron-nickel metal, spinel, feldspar and hydrated phyllosilicates and organic compounds. Each one of them produces a unique reflectance spectrum with specific diagnostic features (Reddy et al., 2015). By analyzing these features and determining which wavelengths are absorbed and how strongly, it is possible to establish which minerals are present on the surface and in which quantity. From some high quality spectra it is also possible to determine the average composition of the object and the relative abundances. In addition, the wider the observed wavelength interval, the more information about the chemical and compositional properties of the asteroid are retrieved.

Asteroid spectra have been obtained since the late 1960s. One of the first example is reported in McCord et al. (1970), who published the spectral measurements for (4) Vesta in the $0.3-1.1$ μm interval. One of the first surveys containing several asteroid reflectance spectra was the Eight-Color Asteroid Survey (ECAS) (Zellner et al., 1985) in the 1980s.

In the last decades, the development of CCD spectrographs has made possible to obtain spectra of significantly fainter asteroids with a much higher spectral resolution than achievable with filter photometry. Up to now thousands of asteroid spectra have been collected as part of different surveys, such as the SMall Asteroid Spectroscopic Survey in the visible light (SMASS) phase I (Xu et al., 1995), phase II (Bus and Binzel, 2002b), and in the NIR (Burbine and Binzel, 2002); the MIT-Hawaii Near-Earth Object Spectroscopic Survey (MITHNEOS) (Binzel et al., 2019) and the Small Solar System Objects Spectroscopic Survey (S^3OS^2) (Lazzaro et al., 2004). Some surveys have been dedicated to specific groups of asteroids, such as the NEAs observed by SINEO (Lazzarin et al., 2005).
The majority of the surveys available in the literature have been carried out in the visible and in the near infrared, while in many cases the UV information downward 450 nm has been lost. The Gaia Data Release 3 (DR3) (Chapter 3) contains the reflectance spectra of 60 518 Solar System small bodies in the interval $374 - 1034$ nm. It represents a huge improvement in spectroscopic data compared to previous surveys.

2.1.5 Taxonomic classification

Taxonomy is the classification into categories of objects sharing similar characterizing properties. For asteroids, these properties are the spectral features and the albedos.
Since these properties are related to the mineralogy surface (Section 2.1.4), objects belonging to different taxonomic classes present different surface properties. A taxonomic scheme is thus a useful tool to investigate the compositional distribution of asteroids in the Main Belt, to search for link with meteorites and to probe the formation and the evolution of the Solar System.

The need for a classification scheme describing the different observable features of asteroid spectra became clear in the 1970s as soon as there were enough data to distinguish between meaningful groups. The first taxonomies, such as Chapman et al. (1971), were based on broad band colors obtained for few bright objects. They noticed a clear separation in two distinct types of objects, the silicaceous asteroids denoted with S and the carbonaceous ones denoted with C.
In the past decades higher resolution observations of fainter objects revealed the underlying continuum distribution between these two complexes and more taxonomic classes were defined. During the years numerous taxonomy systems have been proposed, the

most used the ones from Tholen (1984), Bus and Binzel (2002a), DeMeo et al. (2009) and the recently introduced Mahlke et al. (2022).

Tholen Taxonomy

Tholen built his classification scheme over more than 400 spectra in the $0.3-1.1\ \mu m$ range included in the Eight-Color Asteroid Survey (ECAS) and from estimates of geometric albedos computed from thermal radiometric surveys. Applying a Principal Component Analysis (PCA) on ECAS spectra eleven classes are identified (A, B, C, D, F, G, Q, R, S, T and V), with three of them (Q, R and V) with just one member. The E-, M- and P-classes present degenerate spectra and can be separated using the albedos. In case this information is not available, the M-, E- and P- objects are grouped into a generic X- class. The Tholen taxonomy presents some advantages that have never been fully replaced by the other taxonomies. The visual albedos, allowing to distinguish among degenerate spectra, have in fact been dropped by the following taxonomies, except the Mahlke system, while no other spectrographic survey until Gaia DR3 covered the region downward 450 nm, thus losing all the information in the UV region.

Bus & Binzel Taxonomy

The Bus & Binzel taxonomic system is based on a set of 1447 visible spectra included in the second phase of the Small Main-belt Asteroid Spectroscopic Survey (SMASSII). This taxonomy is built over the pre-existing taxonomy schemes and it defines three major groupings (the S-, C- and X- complexes) and thirteen classes (A, B, C, D, K, L, O, Q, R, S, T, V and X), some of them in common with Tholen and some newly introduced.
Asteroids showing intermediate spectral characteristics are reported with a double letter designation, where the first letter refers to the class that the asteroid spectrum resembles the most. The new intermediate classes are Cb, Cg, Ld, Sa, Sk, Sl, Sq and Sr. X- complex asteroids may be assigned to the Xc-, Xe- and Xk- subclasses depending on the strength of specific spectral features. Finally, inside the C- complex the Ch- and Cgh- classes are defined, indicating those objects showing a spectral feature at $0.7\ \mu m$ attributed to hydration.

Bus-DeMeo Taxonomy

The Bus-DeMeo taxonomic scheme is built over 371 asteroid spectra in the visible and the near IR range, from 0.45 to 2.45 μm. DeMeo defines 24 classes, almost all retained from the Bus & Binzel scheme, except for the Ld-, Sk- and Sl- types that are eliminated, while a new Sv- class is introduced. A mark 'w' at the end of the notation is added to indicate asteroids showing a larger slope compared to the other objects of the same spectral type.

In this system, several classes remain relatively rare. Only six objects fall into the A-class, while the eight objects belonging to the Q- class are all NEAs. The R- and O-classes both have one single member.

Mahlke Taxonomy

The Mahlke taxonomy scheme is built over 2983 spectra of 2125 asteroids observed in the visible and in the near infrared. The albedos are reintroduced in the classification.
Applying a probabilistic PCA (PPCA) in a lower-dimensional representation of the observations, Mahlke identifies 17 spectral classes, divided in the S-, C- and M- complexes. A new Z- class is introduced, while the spectral degeneracy of the X- complex is solved thanks to the albedos.
In some cases, a multiple letter designation referring to the presence of specific absorption features is assigned. The h- feature, exclusive to the C- complex, refers to an absorption feature centered at 0.7 μm associated to surface hydration. The e- and k- features can instead be assigned to every class and they respectively refer to a narrow absorption at 0.5 μm and to a broad absorption around 0.8 - 1.0 μm.
Unlike previous taxonomic schemes, in Mahlke system the assignment of an asteroid to a class is probabilistic, thus allowing to take into account the uncertainties of the classification and the degeneracy between classes.
In addition, an iteration algorithm allows to classify new observations, among which partial ones obtained just in the visible or in the near infrared. This last aspect is particularly well suited for classifying spectra from ongoing surveys, like Gaia.
Here follows a brief description of the spectral classes listed in Mahlke taxonomy. Most of these classes are in common with the other taxonomic systems.

C- COMPLEX

The C- complex is divided into the four classes B, C, Ch and P. Unlike previous taxonomies, the P- class has been moved away from the X- complex and included in the C- complex. The D- and Z- classes are referred to as endmembers because of some P-types showing intermediate spectra between the C- complex and these two classes. Any C- complex object presenting the h- feature is classified as Ch- type no matter what its original spectral type was.
C- complex asteroids present featureless spectra, except for the h- absorption when present. The diversity between classes resides in the albedos and in the different slopes and shapes of the spectra.
These objects are usually found in the outer Main Belt, in particular in the region beyond

the 3:1 mean motion resonance. They are carbonaceous and primitive objects.
The spectral classes of this complex are here briefly summarized:

- *B- types* generally present higher albedos than the other C- complex objects and they also tend to be brighter and bluer. Despite this, B- types do not really differentiate from C- types and form a continuous distribution with them. The archetype of this class is (2) Pallas and about 2% of the objects of this taxonomy are classified as B- types.

- *C- types* show featureless spectra with a neutral to small red slope and they may have an overall concave shape due to a broad absorption feature around 1.3 μm. The archetypes of this class are (1) Ceres and (10) Hygeia, while about 10% of the total sample is classified as C- type.

- *Ch- types* are all characterized by the hydration absorption feature at 0.7 μm. The rest of the spectrum is usually similar to C- types. The h- feature has been detected in about 5% of the asteroids of the sample.

- *P- types* present a red linear slope and they are spectrally degenerate with the M- types. The only way to differentiate between these two classes is by using the albedos, which are around 0.05 for P- types. P- types are found to be halfway between the C- complex and the endmember D- class. About 6% of the asteroids of the sample are classified as P- types.

- *D- types* are dark asteroids usually found in the outer Main Belt and among Jupiter trojans. They show strongly red slopes and featureless spectra, both in the visible and in the NIR. About 4% of the asteroids of the sample are classified as D- types.

- *Z- types* are extremely red objects with low albedos and red slopes larger than the D- types. Z- types present an irregular orbital distribution and they are scattered all over the Main Belt. Only the 1% of the sample is classified into the Z- class.

M- COMPLEX

The M- complex is the most heterogeneous among the three complexes and it comprises the K-, L- and M- classes and the endmember E- class. M- complex members do not in fact share similar mineralogical compositions and the only aspects they have in common are their albedos and their spectral appearance, characterized by absent or faint features at 0.9 μm and 1.9 μm.
The spectral classes of this complex are here briefly summarized:

- *K- types* are associated to the Eos family and are characterized by visual albedos around 0.10, a red slope in the visible, an absorption band at 1 μm and a neutral slope in the NIR, despite some objects belonging to this class, such as the archetype (221) Eos, present a blue NIR slope. When only NIR data are available, the K- class is degenerate with the B- class. The 2% of the sample is classified into the K- class.

- *L- types* are characterized by a wide absorption feature around 2 μm probably associated to calcium-aluminum inclusions. Besides this common feature, L- type spectra show a vast heterogeneity in slope and shape in the visible and in their albedos. This makes them degenerate with similar classes, like the M- and S- types. Slightly less than the 3% of the sample has been classified into the L- class.

- *M- types* form an heterogeneous group containing red asteroids presenting linear or even convex shapes. Featureless spectra are common among M- types, but most of them present silicate features at 0.9 μm and/or 1.9 μm. Two populations may be identified within the M- class, a chondritic one with archetype (21) Lutetia and a metallic one with archetype (16) Psyche, despite they cannot be separated by spectra alone. The 7% of the sample is classified as M- type.

- *E- types* show a steep slope in the visible and a flat one in the NIR, thus their spectra are degenerate with all the other classes of the M- complex. E- types are recognized from the albedo, which is generally above 0.5. A lot of E- types show the e- and k- features. About the 2% of the sample is classified as E- type.

S- COMPLEX

The S- complex includes the S- and Q- classes and the endmembers A-, R-, O- and V- classes. This complex is by far the largest among the three, probably due to an observational bias. S- types in fact dominate the Near Earth region and the inner Main Belt and they are characterized by high albedos, usually around 0.20. These aspects combined make them easier to detect than the dimmer and farther C- types. S- complex asteroids present continuous trend in slope and in silicate features, explained by variations in the mineral composition, especially of olivine and pyroxene.
The spectral classes of this complex are here briefly summarized:

- *S- types* show spectra particularly subject to variations in slope and band structure due to phase angle effects and space weathering. For this reason all the S- subclasses

introduced by the Bus-De Meo system have been removed by Mahlke and just the S- class is retained. 42% of the asteroids of the sample have been classified as S- types.

- *Q- types* are usually found in the Near-Earth region and compared to S- types they present higher albedos (around 0.30), a wider absorption band at 1 μm and a bluer slope over the whole spectral range. The archetype of this class is (1862) Apollo, while the 5% of the sample is classified as Q- type.

- *A- types* are characterized by a steep red slope and a deep absorption band due to olivine at 1.0 μm, while their albedo is between 0.20 and 0.30. Only 1.5% of the sample has been classified as A- type.

- *O- types* are extremely rare. This class was initially defined by Tholen for the single object (3628) Boznemcova, which presents a unique spectrum characterized by a wide round absorption feature at 1 μm. Only one other object, (7472) Kumakiri, has been classified as O- type in this system.

- *R- types* are characterized by 1 μm and 2 μm features that are deeper than the ones of S- types, while the width of the 1 μm absorption band is halfway between the V- and Q- types. The class, initially built by Tholen around the single object (349) Dembowska, contains 10 objects.

- *V- types*, taking their designation from (4) Vesta, are easily recognizable thanks to their very characteristic absorption features at 1 μm at 2 μm. The 7% of the sample is classified as V- type.

The evolution of the taxonomic schemes from Tholen to Bus-De Meo to Mahlke is reported in Figure 2.2. The arrows indicate the overall evolution of the spectral types. No analogues of the K- and L- types were defined by Tholen. The three dots indicate an unclear evolution between the three systems. This is the case for the X- complex as the albedo is not considered in the Bus-DeMeo system. The vertical lines indicate classes that have been suppressed from one taxonomy to the other. This is the case for the Bus-DeMeo Xn-class, whose characteristic features were grouped inside the k- feature by Mahlke, and the T- class, which instead has been completely removed from Mahlke system.

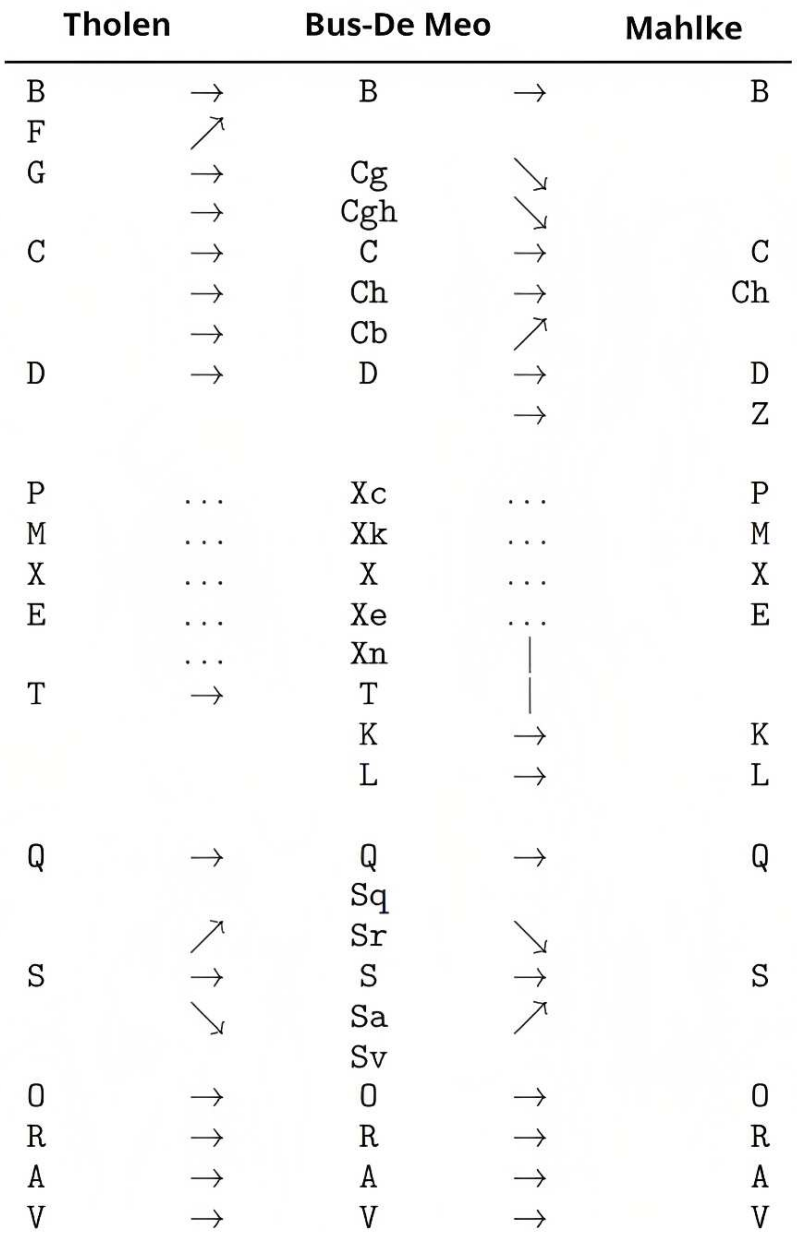

FIGURE 2.2: Evolution of the taxonomic schemes from Tholen (left) to Bus-DeMeo (centre) and Mahlke (right). The arrows show the evolution of each spectral type. Credit: Mahlke et al. (2022)

2.2 Asteroid Families

2.2.1 Introduction

Collisions among large asteroids are low probability events which played a fundamental role over the life of the Solar System. These collisional events produce groups of asteroids first identified by Hirayama (1918), commonly known as asteroid families, orbiting around the Sun on very close trajectories and usually sharing their physical properties. At present, about 30-35% of Main Belt asteroids are classified as family members. Some of the largest Main Belt families are shown in the $a - \sin(i)$ plane over the background asteroids in Figure 2.3. The Tirela (Chapter 5) and the Watsonia (Chapter 6) families are also reported there.
In most cases the parent body was shattered during the impact, but there are few exceptions of families resulted from a large cratering event that did not disrupt the parent body. The families of the first type are known as *fragmentation families*, while the families of the second type (for example Vesta) are known as *cratering families* and consist of a single large body and a swarm of asteroids of much smaller size. There is no net separation between the two classes and the definition to distinguish them is somehow arbitrary, despite there is a general convention stating that for disruption families the largest remnant has a size which is smaller than half the size of the parent body. Spoto et al. (2015) define as fragmentation families those groups for which the volume of the family without the largest member is more than 12% of the total volume, while for cratering families this percentage must be less than 12%.

Asteroid families are perfect natural laboratories to investigate the collisional and the dynamical evolution of the Main Belt.
Being created by large-scale collisions, families provide us the tools to test impact physics on planetary scales (Michel et al., 2015), while by determining their age it is possible to probe the first phases of life of the Solar System and its subsequent evolution (Spoto et al., 2015). Furthermore, the analysis of the present-day structures of families provides a way to understand the role that resonances and non gravitational forces such as the Yarkovsky effect played in the dynamical evolution of the Main Belt (Carruba et al., 2018).
Since families are usually created after the fragmentation of a large object, they provide a way to probe the interior of these primitive parents bodies, allowing to get important information about the internal composition and structure of asteroids (Masiero et al., 2015). In addition, collisions create thousands of fragments of different size, whose orbits

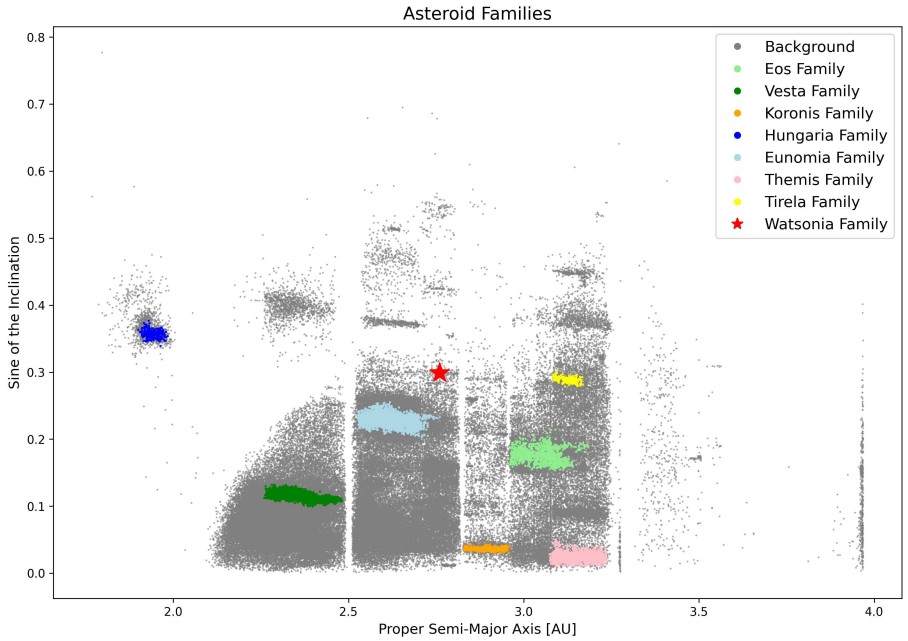

FIGURE 2.3: Tirela, Watsonia and some of the largest families plotted against the background of Main Belt asteroids in the $a - \sin(i)$ plane.

evolve due to a combination of further collisions, resonances and non gravitational effects. Some families are therefore believed to have played a very significant role in supplying the near Earth region with asteroids (Heck et al., 2017). Finally, since it is likely that some of the original bodies were water-bearing asteroids, families provide a way to estimate the actual content of water in the asteroids of the Main Belt and its delivery to the the near Earth region (Hsieh et al., 2018).

2.2.2 Identification of asteroid families

Families are dense clusters of asteroids that can be detected by density contrast with respect to a background in the phase space of appropriate parameters. In principle different datasets of parameters can be adopted to identify families, such as the proper elements a, e, $sin(i)$ or the physical observations.
Usually proper elements are preferred to physical properties since they contain much more information. Physical properties observations are in fact available for smaller catalogues than proper elements and they include many objects observed with low S/N and thus with poor accuracy. Physical properties can still result useful to solve degeneracies between families and/or interlopers. However, they can also be ambiguous as they are based on

the assumption of the homogeneity of the parent body.
Families identified from the space of proper elements are known as *dynamical families*, while *collisional families* are groups of asteroids that have been formed after a collision, being it fragmentation or cratering. Note that there is not a perfect one to one correspondence between collisional and dynamical families, since a dynamical family could include multiple collisional events or a single collisional family could be identified as more than one dynamical family due to resonance gaps inside the family.

The standard and most diffused approach to identify asteroid families in the space of proper elements is the hierarchical clustering method (HCM), first introduced by Zappalà et al. (1990). The HCM identifies families through a simple-linkage clustering procedure: it works in an agglomerative way such that at each step two clusters that contain the closest pair of elements (not yet belonging to the same cluster) are combined if their separation is less than a critical distance. The distance in the 3D proper elements space between two asteroids is defined by the following metric (Zappalà et al., 1990):

$$d = na_p\sqrt{\frac{5}{4}\left(\frac{\Delta a_p}{a_p}\right)^2 + 2(\Delta e_p)^2 + 2(\Delta \sin i_p)^2} \tag{2.1}$$

where d, usually expressed in m/s, corresponds to the relative velocity between the two objects. na_p is the heliocentric orbital velocity of an object on a circular orbit with semi-major axis $a_p = (a_{p,1} + a_{p,2})/2$, $\Delta a_p = a_{p,1} - a_{p,2}$, $\Delta e_p = e_{p,1} - e_{p,2}$ and $\Delta \sin i_p = \sin i_{p,1} - \sin i_{p,2}$, where the subscripts 1 and 2 refer to the two objects whose separation is being computed.
The simple-linkage is at the same time one of the main advantages and problems of the HCM. The method does not preclude a priori any shape of the family, since even irregular clusters with narrow bridging regions in the proper elements can be identify. This however is also related to the main problem of the HCM, known as *chaining effect*. Since only a single pair of objects is needed to merge two clusters, the method tends to incorporate together nearby groups, thus preventing to distinguish between families in some cases. This happens because the HCM is a statistical method that unavoidable associates interlopers to the family. Two nearby groups of asteroids not belonging to the same family could therefore be merged through the chaining effect due to two interlopers that are selected as the closest pair of objects in the two clusters.

The removal of interlopers is thus a procedure of great importance that was intensively explored in the past years. One solution would be to apply the HCM in a multidimensional

space, including the surface properties along with the proper elements. This was for example implemented by Parker et al. (2008), who used combinations of SDSS colors as fourth dimension, or by Carruba et al. (2013), who instead used the albedos. This allowed to greatly reduce the number of interlopers in families.
Another strategy could be to use the reflectance properties before applying the HCM: for example Masiero et al. (2013) divided the Main Belt in the two S- and C- complexes using WISE albedos and then ran the HCM separately over the two populations, thus avoiding interlopers of different spectral type.
These two approaches however present two major problems. The first, already discussed, is that reflectance properties are known with good accuracy only for a limited sample of asteroids. The second is that these methods implicit assume that every family is homogeneous in composition. However, heterogeneous families despite being rare exist as formed after the disruption of partially differentiated bodies. This approach thus would prevent these families from being discovered.
When dealing with reflectance properties as fourth dimension, a very helpful result based on statistical arguments was found by Migliorini et al. (1995), who obtained that with the HCM the expected percentage of interlopers should be around 10%. A value much higher than this would be a hint that the family is heterogeneous in composition or that two different families have been merged together.

Another drawback related to the HCM in addition to the chaining effect is that it becomes a very time consuming procedure when applied to a large number of asteroids. Since during the past years the available catalogues for the proper elements of asteroids have increased exponentially, this problem has became more and more significant. To improve the efficiency of the algorithm, improved versions of the classical HCM have been recently proposed by different authors.
One example is the multi step HCM proposed by Milani et al. (2014). Instead of running the HCM over all Main Belt asteroids, they proposed to apply it first only to large objects and then include the smaller ones in following steps. In order to reduce the computation time, the authors also proposed to divide the Main Belt in smaller regions such to deal with a smaller number of objects simultaneously. As first step, the HCM is applied to the objects brighter than a specific limit magnitude H_{lim} in a certain subgroup, so that the *family cores* are identified. In the following steps background asteroids are assigned to the cores if their distance from one core member is less than a specific critical distance. Some objects could result in belonging to more than one family and they are therefore

considered as part of the intersection. This method has been successfully applied by Milani et al. (2016) to a sample of more than 500 thousand asteroids, but with the continue growth of catalogues the multi step method will become more and more time consuming. In addition, intersection asteroids represent doubtful cases which require follow-up observations.
Some authors have instead tried to use machine learning (ML) based algorithm for the identification of asteroid families. Carruba et al. (2019) were the first to apply a ML-based HCM to a sub-population of high inclinations Main Belt asteroids. The ML approach identified all the significant groupings in that region, resulting faster and easier to use than the classical approach. However, the ML-based HCM showed some drawbacks, in particular related to its precision, since for more than half of the families the ratio between the number of the objects identified by this algorithm and by the classical approach was found to be less than 80%.
Both the ML-based HCM and the multi-step HCM were instead found to work pretty well in attaching new members to already existing families.
All these HCM based algorithms fail when trying to identify old families, generally older than 3 Gyr. Members of these families are in fact too dispersed in the space of the proper elements for the HCM to work, so that different approaches have been searched. The most efficient makes use of the v-shape (Section 2.2.3), whose idea is to test for possible v-shape number-densities amongst the background of unassociated asteroids. These methods, despite not being trivial to implement, have been successfully applied in some cases. For example Walsh et al. (2013) were the first to use this type of algorithm and they identified the Polana family, about 2 Gyr old, and the Eulalia family, about 0.8-1 Gyr old.
These algorithms however are extremely sensitive to many parameters and they robustness still need to be proven, since they can potentially generate false-positive results. A statistical study conducted by Deienno et al. (2021) showed that there is a probability up to 80% to detect 0.5 - 2.5 Gyr old families, while families older than 3 Gyr are extremely difficult to detect.

Of the various families classification systems that have been proposed in the literature during the years, only the two of them that will be used in the following scientific analysis are here briefly discussed. These two systems refers to the Asteroid Families Portal (AFP, Novaković et al. (2022)) and to the Asteroids Dynamic Site (AstDys, Milani et al. (2014)). The classification of AFP is based on a classical HCM, while AstDys uses a multi-step HCM, so, despite sharing some similarities, the two classification systems are somehow

different. For example, AFP identifies some families that are not present in AstDys or vice versa, while in some cases the largest member is different therefore the same family is reported with different names. For example, 1400 Tirela in AFP is named 1040 Klumpkea in AstDys. For families identified by both systems, the memberships are usually very different, since AstDys always tends to include less objects than AFP.
In addition to the family memberships, the two websites provide few tools that are extremely useful for the study of families. Both websites in fact include updated lists of the orbital and proper elements of known asteroids. AstDys in addition provides the v-shapes and the ages for few tens of families based on Spoto et al. (2015) and a graphic tool for the visualization of Main Belt asteroids in the space of the proper elements. AFP instead implements few algorithms that can be freely used. The website gives the possibility to apply the HCM algorithm described by Zappalà et al. (1990) to identify all the members of a family, the Backward Integration Method (BIM) implemented by Nesvorný et al. (2003) to estimate the age of a family and an automatic procedure to obtain a list of family interlopers based on the algorithm by Radović et al. (2017).

2.2.3 Age determination: the v-shape method

One of the main purposes of the identification of families is to try to constrain their ages, from which reconstruct their collisional and dynamical history and obtain insights on the first phases of the evolution of the Solar System. It must be kept in mind that it is possible to determine the age for collisional families, which not always coincide with the dynamical families identified from the proper elements.
Few methods to constrain ages exist, but here only the one described in Milani et al. (2014) and in Spoto et al. (2015) is reviewed, since a simplified procedure of this will be used in Chapter 5 to constrain the age of the Tirela family.
This method exploits the Yarkovsky effect, which generates secular perturbations in the proper elements of an asteroid inversely proportional to its diameter D.
When a family is created all the fragments are located around the same semi-major axis, with the only differences due to the initial ejection velocity field. The semi-major axes of the fragments then evolve in time due to the Yarkovsky effect, given by $da/dt = c\cos(\phi)/D$, where c is a calibration constant depending upon density and thermal properties and ϕ is the obliquity (angle of the spin axis with the normal to the orbit plane). Considering c as a constant inside the family, for asteroids of a given diameter their position at a certain time will depend only on their obliquity, while the drift in semi-major

axis decreases as the size of the asteroids increases. The family members thus form a sort of *v-shape* in the $a - 1/D$ plane, whose width increases with the age of the family. The largest body, from which the family usually takes its name, is approximately located at the vertex of the v. An example of v-shape for the Massalia family is shown in Figure 2.4, where the family members are marked as blue crosses.
The external objects corresponding to the prograde and retrograde spin states ($\phi = 0°$ and $\phi = 180°$) can be fitted by two straight lines, whose slopes can be converted in ages by mean of a Yarkovsky calibration.
This method works only if all the members of the family have the same age (therefore only for collisional families) and if the objects cover a sufficient interval in D. In addition, families are usually bounded in semi-major axis by strong resonances. Objects located inside them evolve into unstable orbits and should not be considered in the computation of the age. The cut in semi-major axis due to resonances is observed in most v-shapes, for example in Figure 2.4 Massalia is cut below 2.35 AU and over 2.47 AU. Finally, families usually presents large interlopers, which need to be removed to not spoil the computation of the slopes and thus of the age. Some evident outliers can already be removed from a visual inspection the v-shape, while others need a more exhaustive analysis.

Before constraining the age, the diameters of the family members must be determined. The following equation stemming from the definition of the albedo can be used:

$$D\ [km] = 1329\ [km]\ \frac{1}{\sqrt{\rho_V}}\ 10^{-H/5} \tag{2.2}$$

where ρ_V is the visual albedo and H is the absolute magnitude. The albedo is not known for all members, therefore usually a common value for all the family is assumed in the computation. This introduces the first error on the age of a family, since the positions of the objects in the $a - 1/D$ plane are not absolute, but present uncertainties that affect the computation of the slopes.
The region to be fitted to determine the slopes is then manually selected such to take into account the resonances that limit the family in the inner and outer semi-major axes. Next the $1/D$ axis is divided into a number of bins N, which are created such that they all roughly contain the same number of objects. The number of bins N depends on the number of objects belonging to the family. Since the limit in $1/D$ could be different for the inner and outer sides, the division in bins could not be the same for the two sides.
In the case of the inner side, for each bin the object with the minimum a is selected, while for the outer side, the object with the maximum a is selected. Those are the objects that

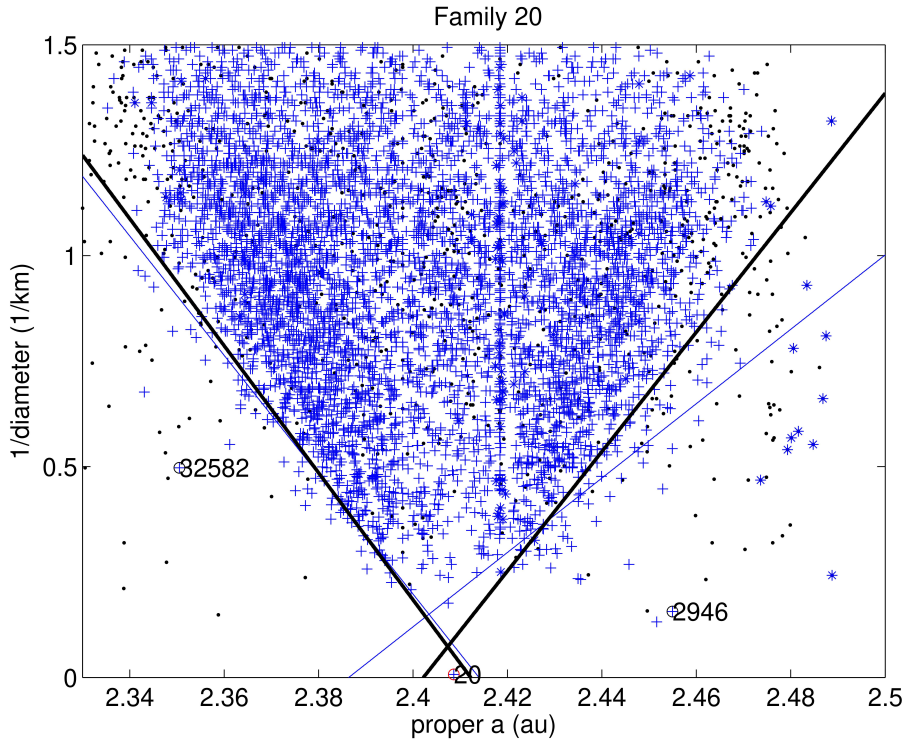

FIGURE 2.4: $a - 1/D$ plot for the Massalia family. Blue crosses are family members, blue stars are resonant objects and black points are background objects. The thick black lines are the best fit lines for the slopes of the v-shape. Outliers are marked with a blue $\oplus$ symbol, while the red $\oplus$ marks the position of (20) Massalia.
Credit: AstDys.

are used for the computation of the slopes through a linear regression. This procedure is iterative and includes the automatic outlier rejection scheme described in Carpino et al. (2003). At each iteration, the residuals of the observations, the expected covariance and their χ values are computed. An object is marked as an outlier if its χ^2 is larger than a threshold value χ^2_{rej}. The iteration ends when no more outliers are found and finally the slopes are computed from a linear regression of the remaining points. In Figure 2.4 two interlopers have been found by the procedure and have been marked by a $\oplus$ blue symbol. The light blue lines are the fitted lines before the removal of the interlopers, while the black thick lines are the final slopes that are used for the computation of the ages.
Since the slopes for the inner and outer sides are computed in two separate procedures, it could be possible to obtain very different values in the two cases. This could mean that the family is actually composed by two (or more) families of different ages. However, different values of the slopes could be also obtained if one of the two sides of the v-shape is poorly sampled or with some irregularities.

The age of the family can then be constrained from the inverse slope, which is the change $\Delta(a)$ accumulated over the family age by a family member with unit $1/D$. The age is in fact given by $\Delta(t) = \Delta(a)/(da/dt)$, where da/dt is the drift in semi-major axis due to the Yarkovsky effect for an object with obliquity 0° for the inner side and 180° for the outer side. da/dt is an unknown quantity that can be approximately calibrated. To date, no accurate measurements of the Yarkovsky effect for Main Belt asteroids exist, therefore Near Earth Asteroids are used at their place. Spoto et al. (2015) for example used data available for Bennu, currently the object for which both da/dt and the physical properties are known better. This procedure however introduces large uncertainties in the computation, with the final errors usually between 20% and 30%.
The drift in semi-major axis due to the Yarkovsky effect can be estimated from the scaling formula:

$$\frac{da}{dt} = \left(\frac{da}{dt}\right)_B \frac{\sqrt{a}_B\,(1-e_B^2)}{\sqrt{a}\,(1-e^2)}\,\frac{D_B}{D}\,\frac{\rho_B}{\rho}\,\frac{\cos(\phi)}{\cos(\phi_B)}\,\frac{1-A}{1-A_B} \tag{2.3}$$

where e is the eccentricity, ρ the density, A the albedo and $cos(\phi) = \pm 1$ depending if the inner or outer side is considered. The B subscripts refer to the Bennu parameters used for the calibration.
The diameter reported in Equation 2.3 is not the diameter of an actual asteroid, but the reference value corresponding to the inverse slope. The density corresponding to this diameter is computed by assuming the same porosity as Bennu and the same composition of the largest object of the same taxonomic type of the family.

Spoto et al. (2015) and Milani et al. (2016) extensively applied this method to Main Belt families and succeeded in constraining the ages for few tens of them. The results are reported in Figure 2.5, where the groupings on the horizontal axis correspond to fragmentation families, cratering families, young families and one-sided families. These last ones, be they fragmentation or cratering, are families in which just one the two sides is recognizable and therefore just one estimate of the age can be retrieved. On the y-axis, families are instead classified according to their ages: *primordial families* are older than 3.7 Gyr, *ancient families* have ages between 1.0 and 3.7 Gyr, *old families* between 0.1 and 1.0 Gyr and finally *young families* are younger than 0.1 Gyr. The results reported in this figure show the limitation of the HCM algorithm in detecting primordial families, since just possibly three of them have been identified in all the sample. Also young families are relatively rare, but for them there could be a bias in the v-shape method, since their dispersion in semi-major axis could be masked by the initial ejection velocity at disruption.

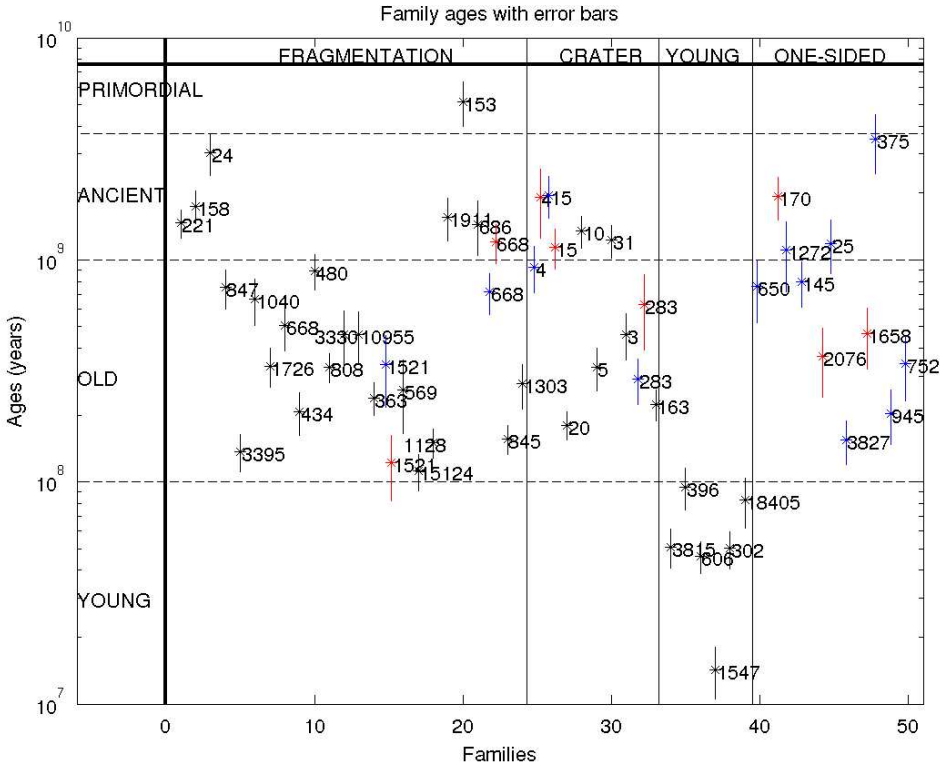

FIGURE 2.5: Asteroid families ages. See text for better description. (Credit: AstDys).

In Figure 2.5 there are three families (4 Vesta, 15 Eunomia and 1521 Seinajoki) for which the ages retrieved from the two sides are discordant and therefore both values have been reported in the plot. This suggest that these three families were created following at least two impacts. For Vesta this is known for a long time (it was first suggested by Farinella et al. (1996)) and it is confirmed by the presence of two large impact basins on its surface, whose ages are in agreement with the values reported in Figure 2.5.

Significant improvements in the precision with which family ages are known are expected in the future. To date there are several families with too few members to be statistically relevant for an age determination. The discovery of new asteroids may populate these families such that constraining their ages will become possible.
Planned future sky surveys will also provide accurate measurements of the physical properties of large catalogues of asteroids. This will help in removing the interlopers from the families and it will allow to have better estimates of the diameters of the objects, reducing one of the uncertainties affecting the ages.
Finally, regarding the calibration of the Yarkovsky effect, the future data releases of the

Gaia mission will contain high accurate astrometric observations for hundreds of thousands of Main Belt asteroids. As it was already proved by Gaia DR2 in 2018, the accuracy that Gaia observations can reach in determining the orbits of Main Belt asteroids is two orders of magnitude better than when it is done with ground based telescopes. Thanks to the future data releases, the Yarkovsky effect might be determined also for Main Belt asteroids. The final goal should be determining the Yarkovsky effect for at least one object of every family, or at least to have one measurement for every spectral type. This will allow to greatly reduce the uncertainties on family ages.

Chapter 3

The Gaia Mission

3.1 Introduction

Astrometry is the science of positional astronomy, which involves the precise measurements of the location and the movement of a celestial object, such a star, within the plane of the sky. The origin of astrometry can be dated back to the ancient Greece when Hipparcus around 190 BC thanks to stellar catalogues compiled by his predecessors discovered Earth's precession by studying the movement over time of the stars along the plane of the sky. In the last centuries, astrometry has led to a number of important discoveries, among which the stellar proper motion, the stellar aberration and nutation (Perryman, 2012).

In the last decades the accuracy of astrometric measurements has increased exponentially thanks to the improvement of technologies and to the possibility to observe from space, thus eliminating the effect of Earth's atmosphere. Astrometry in space was pioneered by the Hipparcos mission between 1989 and 1997 (van Leeuwen, 1997) and then followed by Gaia (Gaia Collaboration, 2016).

Designed and built by ESA, Gaia was launched in 2013 and it is expected to operate until 2025. The mission aims to unravel the structure, the dynamics and the evolution of the Milky Way through the observation of its stars. Since the mission observes every point-like source in the sky without being able to characterize its nature, Gaia will also detect comets, asteroids and quasars, among others, which altogether will form the largest and most precise space catalogue ever made. Astrometry and low-resolution spectrophotometry data are obtained for all sources, while for the brightest stars high resolution

spectra are acquired to determine radial velocities. Data, accompanied by several queries and visualization tools, are publicly available in the Gaia archive[1].

Gaia aims to investigate the structure, the dynamics and the evolution of the Milky Way through the study of the physical properties of its constituent stars and the analysis of the distribution and the kinematics of the luminous and dark masses. Gaia samples over one billion stars covering a large spatial volume which, despite representing just 1% of all the stars of the Galaxy, allows statistical work to be conducted. By combining ages determinations with physical properties measurements of individual stars, Gaia also probes the star formation histories of stellar populations of the Milky Way (Gaia Collaboration, 2016).
Gaia data covers most phases of stellar evolution, including pre-main sequence stars and peculiar objects, thus contributing to the study of star formation and evolution of single stars. Precise photometric measurements are instead bringing to the detection of numerous RR Lyrae and Cepheids. These data, in combination with parallaxes, allow to calibrate the period-luminosity relations to better accuracy, thus improving the quality of the cosmic-distance ladder. Astrometric measurements instead are used to probe oscillations in the positions of stars within the plane of the sky, which will bring to the detection of a huge number of exoplanets (Perryman et al., 2014).
Some of the stars detected by Gaia do not belong to the Milky Way, but to other galaxies of the Local Group, such as the Magellanic Clouds and M31, thus extending the study conducted to the Galaxy to other Local Group members. In addition Gaia also detects unresolved galaxies and quasars which can be used to conduct cosmological studies (Gaia Collaboration, 2016).
Although Gaia is designed to detect stars, it actually observes all point-like sources in the sky, among which Solar System objects (SSOs) (Tanga et al., 2022). Planets are usually too bright to be detected by Gaia, while numerous asteroids present a luminosity in the range of the spacecraft instrumentation. NEAs can sometimes escape to Gaia detection due to their high apparent motions, but for the other minor bodies, especially MBAs, Gaia provides a complete census of orbital parameters and taxonomy in a single and homogeneous photometric system.

[1] https://gea.esac.esa.int/archive/

3.2 The spacecraft

The concept of the Gaia mission is based upon the principle of scanning space astrometry (Lindegren and Bastian, 2010), which relies on a slowly spinning satellite measuring the crossing times of targets transiting in the focal plane.
Gaia is equipped with two fields of view (FoVs) separated by a large, constant angle (the basic angle) whose viewing directions are mapped into the same focal plane. The observations are converted in small scale separations between stars inside the same FoV and in large scale separations between stars in the different FoVs. In turn, because the parallax is proportional to the angle between the Sun and the star, the angular separations are converted in relative parallaxes. Once a calibration point is provided, the distances of the stars are retrieved (Gaia Collaboration, 2016).

The service and payload modules

From a mechanical point of view, the spacecraft consists of a service module (SVM) and a payload module (PLM).
The service module comprises all the mechanical, structural and thermal elements supporting the instruments and the spacecraft electronics. It includes the electrical harness, the solar-array panels, the propulsion systems, the sunshield assembly and the payload thermal tent. The SVM is needed to maintain the orbit and the attitude of the spacecraft, to handle the communications with the Earth and to protect the payload from the external environment, such as the solar radiation and the impact with micro-meteorites.
The payload module instead houses the two telescopes and the focal-plane assembly (FPA). In addition, the PLM includes the focal-plane computers VPUs, the detector electronics, an atomic master clock, the basic angle monitor BAM, the wave-front sensor WFS and the payload data-handling unit (PDHU).
Gaia is equipped with two identical, three-mirrors anastigmatic telescopes with apertures of 1.45 m x 0.50 m and focal lengths of 35 m. Both telescopes have a field of view of 1.8° x 0.7° degrees and they work close to the diffraction limit in the visible range (300 - 1050 nm). The two telescopes share a common focal plane and their lines of sight are separated by a fixed angle, called the "basic angle", equal to $\Gamma = 106.5°$. Only if the basic angle is fixed (or if its variations are known with high precision) the measurements obtained by the two telescopes can be combined to retrieve the angular separations between stars.
The telescopes and their relative viewing directions are usually referred to as FOV1 and FOV2 or as preceding and following.

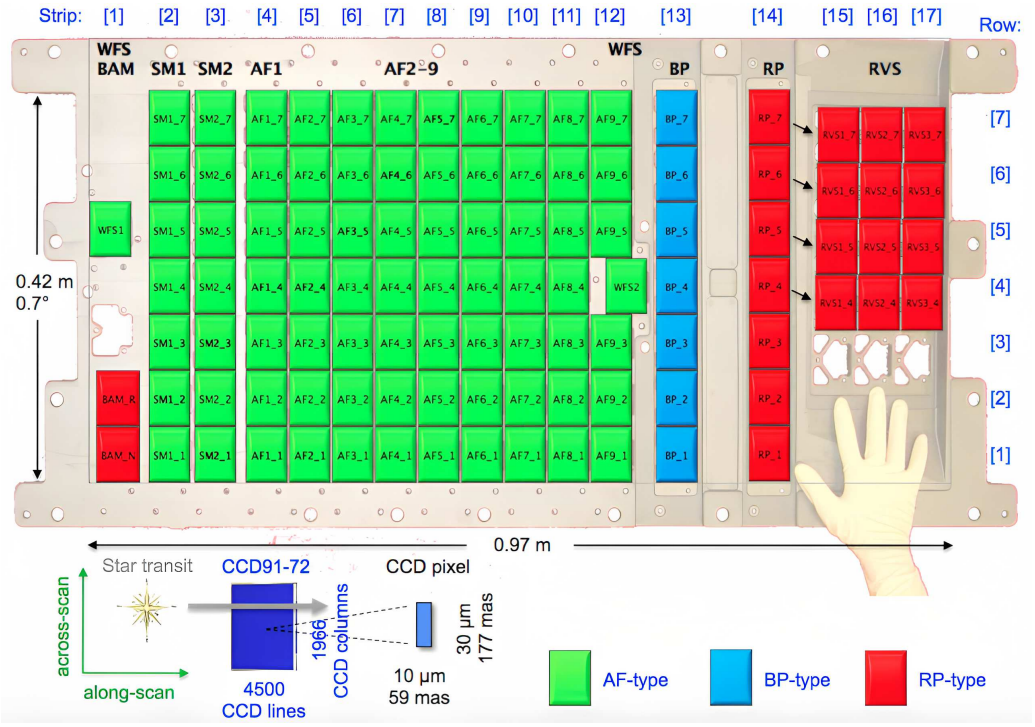

FIGURE 3.1: Illustration of the focal-plane assembly with a human hand for scale comparison. (Credit: Gaia Collaboration (2016)).

The focal-plane assembly

The focal-plane assembly, shared by both telescopes, consists of 106 detectors operating in time-delayed integration (TDI) mode. The focal plane houses five functionalities, two for internal use (metrology and sky mapping) and three with scientific functions (astrometry, spectrophotometry and spectroscopy).
A schematic illustration of the focal-plane assembly is reported in Figure 3.1. Because of the rotation on the spacecraft, the images of the observed sources slowly drift along the focal plane (from left to right in Figure 3.1) thus being detected by all the instruments. A brief description of the five functionalities implemented in the FPA follows here.

The *metrology* functionality includes two basic-angle monitors (indicated by BAM in the scheme) and two wave-front sensors (indicated by WFS).
As previously explained, the basic angle must remain ultra-stable or its variations must be known with extreme accuracy. The BAMs monitor for possible variations of the basic angle every 15 minutes with an accuracy of 1 μas by studying the interference patterns produced by fixed, artificial stars created by collimated laser beams (Fabricius et al., 2016).

The two wave-front sensors are instead needed to monitor the alignment and the optical quality of the telescopes. The alignment of the mirrors has in fact been performed on the ground, under the influence of gravitational forces, while Gaia is operating in zero-gravity conditions. By observing bright stars, the WFSs are able to derive information in order to correct the alignment and the focus of the telescopes while in orbit. This is done by mean of high-precision actuators moving the secondary mirrors of both telescopes.

Both telescopes are equipped with their own *sky mapper*, composed of seven detectors each and indicated with SM1 and SM2 in Figure 3.1. The SMs detect sources of interest in the field of views of the telescopes (point sources like stars and slightly extended objects such as asteroids and galaxies) and reject spurious observations such as cosmic rays or solar protons.
The identification of objects occurs thanks to detection algorithms implemented in the VPUs, which scan the images coming from each SM searching for fluxes emitted by celestial sources. A source is retained if its flux matches the criteria of the algorithms, otherwise it is discarded.
Once a source is confirmed, an estimate of its magnitude G (standing for the Gaia band) is determined from the SMs. Sources brighter than $G = 2\ mag$ are not properly observed by Gaia because they are saturated, while sources fainter than $G = 21\ mag$ usually escape detection. The typical precision on magnitude estimates is around 1 mag.
The objects that are detected with the SMs are then analyzed by the astrometry, spectrophotometry and spectroscopy instruments.

The "astrometric field" includes 62 CCDs completely devoted to *astrometry* (Gaia Collaboration, 2016), indicated by AF in Figure 3.1.
Objects detected by the SMs are confirmed again in the first column AF1 of the astrometric field. The sources are then followed along their slow drift along the focal plane and their luminosities and positions on the CCDs are accurately recorded at each moment. During the data processing, the positions are converted into angular separations and then in parallaxes. In addition, the photometry obtained with the AF instrument is suitable for variability studies. The astrometric field can handle a density up to one million objects per square degree. If this limit is overcome, only the brightest sources are detected.

As sources reach the end of the AF field, they are identified again by the spectrophotometer instrument, which includes 14 CCDs devoted to low resolution *spectrophotometry* in two different bandpasses (Riello et al., 2018). Spectra, dispersed along the scan direction, are generated by mean of a blue and red prisms, corresponding to the blue and red

photometers (BP and RP in Figure 3.1). The blue photometer covers the range 330-680 nm, while the red photometer the range 630-1050 nm, with an overlapping between 630 and 680 nm. The resolution in $\lambda/\Delta\lambda$ for the two photometers depends on the wavelength and ranges from 30 to 100 for the BP and from 70 to 100 for the RP.
The spectrophotometers are particularly useful to distinguish the nature of a point source (such as a star from a quasar) and for astrophysical characterization (such as metallicities and surface temperatures of stars, photometric redshifts of quasars and reflectance spectra of asteroids).
The RP and BP instruments can handle up to a density of 750 000 objects deg^{-2}.

The *spectroscopy* instrument consists of 12 CCDs arranged in three strips with four rows each, indicated with RVS in Figure 3.1. Medium resolution spectra with a resolving power of $R \approx 11700$ and dispersed along the scan direction are generated through an integral-field grating plate (Cropper et al., 2018).
The main aim of the RVS instrument is to measure the radial velocities of stars brighter than 16 mag through Doppler measurements of absorption lines in the spectral range 845-872 nm, covering the CaII triplet. Radial velocities are in fact required for kinematical and dynamical studies of the Galactic population. Additional information about rotational velocities, atmospheric parameters and individual element abundances are obtained for stars brighter than 11 mag.
The RVS instrument can handle up to a density of 35 000 objects deg^{-2}. Over this density, only the brightest sources are detected.
The wavelength range covered by the RSV is too small to be statistically relevant for asteroid studies.

The different colors in Figure 3.1 refer to the different types of CCDs used: the broad-band CCD (in green), the blue-enhanced CCD (in blue) and the red-enhanced CCD (in red). Each detector shares the same architecture (number of pixels, read out register, etc.) but they differ in the anti-reflection coating, thickness and resistivity. The detectors were built in order to optimize the performance of the observations they were designed to carry out. Broad-band detectors are used for the WFSs and for astrometry, blue CCDs are used for the BP spectrophotometer, red CCDs are used for the RP spectrophotometer, for the RVSs and for the BAMs.

Orbit and environment

Gaia lies at the second Lagrangian point L_2 of the Sun-Earth system (Gaia Collaboration, 2016). This point is located 1.5 million km away from the Earth, in the anti-solar

direction, and it corotates with the Earth completing a revolution around the Sun in one year. The spacecraft orbits around the L_2 along a trajectory with amplitudes of 120 000 km × 340 000 km x 180 000 km and with an orbital period of 180 days.
This orbit presents several advantages compared to an Earth-bound orbit, in particular because the Earth, the Moon and the Sun are always outside the fields of view of the telescopes. In addition, the environmental conditions are stable and allow to easily maintain the spacecraft. The temperature of the instrumentation is in fact found to be constant in time, with only small long-trends variations due to the variable distance from the Sun. The radiation level due to solar wind particles and cosmic rays and the rate of micrometeorites impact are constantly within the expected levels.

The scanning law

The scanning law describes the spacecraft pointing as a function of time. Gaia observes the sky through a uniform revolving scanning motion, which maximizes the uniformity of the sky coverage and optimizes the accuracy in the astrometric measurements (Gaia Collaboration, 2016).
The scanning motion can be divided in three independent components. The first is the six hour rotation of the spacecraft around its spin axis, the second is a 63-days precession of the spin axis around the solar direction and the third one is the annual orbit of the L_2-Earth system around the Sun.
The combination of the three motions enables a full sky coverage every six months and, on average, about 70 transits on the focal plane were estimated for each source during the five-years nominal mission (which however has been extended).
The slow but continuous precession of the spin axis around the solar direction causes a small drift in the across-scan direction of all sources transiting in the focal plane. The amplitude of the drift varies with time over the six hours spin period and it is corrected by the on-board instrumentation and calibrated in the ground processing.
During the time span in which the data published in the Gaia Data Release 3 (Section 3.3) have been obtained, two scanning laws have been employed: the *ecliptic-poles scanning law* (EPSL) and the *nominal scanning law* (NSL).
The ecliptic-poles scanning law was adopted during the first month of the mission in order to calibrate the instrumentation. During this period, the spin axis of the spacecraft stayed fixed in the ecliptic plane, resulting in the scanning of the ecliptic poles every six hours, thus providing an extremely good sampling of these regions. The EPSL had the advantage that there was no precession of the spin axis around the solar direction and therefore the small drift in the across-scan direction was null, making the calibration of the spacecraft

easier to perform.
The ecliptic poles scanning law was gradually abandoned until the nominal scanning law was adopted, which is planned to be used for the rest of the mission. In the NSL, the spin axis precesses with a period of 63 days around the solar direction with an opening angle of $\xi = 45°$. Because of this configuration, the pointing direction of Gaia varies its inclination with respect to the ecliptic between 90° and 45°. As a result of the intersection of the scanning law with the ecliptic plane, Solar Systems Objects (SSOs) are never observed at less then 45° elongation from the Sun and from the anti-solar direction. This has important consequences for the detection of asteroids (Section 3.3).

3.3 Solar System objects in Gaia Data Release 3

The Gaia mission has reached its third data release, Gaia DR3 (Gaia Collaboration, 2022b). Gaia DR3 contains a vast variety of observations collected during 34 months of satellite operations, from August 5, 2014 to May 28, 2017.
The release includes astrometry and broad band photometry for more than 1.8 billion objects. Low resolution BP/RP spectra and high resolution RVS spectra are also published. In addition, Gaia DR3 contains line broadening and chemical composition information derived from RVS spectra, variable stars classification and results on non-single stars, quasars and galaxies.
Finally, Gaia DR3 contains over 3 million astrometry measurements with magnitude $G < 20.9$ of more than 150 thousand Solar System objects and 60 518 low resolution asteroid reflectance spectra (Tanga et al., 2022). The release includes planetary satellites and small SSOs, the vast majority of which are asteroids with few transneptunian objects. Comets instead are not included in Gaia DR3.

Solar System objects exhibit apparent proper motions much larger than those of stars. While the different positions of a star, obtained from different transits, can be matched as they are very similar, this is not the case for moving objects such as the SSOs.
In Gaia DR3, a multi-step procedure (Tanga et al., 2022) has been implemented in order to link the SSOs detected by Gaia to the known population of numbered asteroids. Since this population is in constant evolution, a database of the known SSOs and their ephemerides is maintained and regularly updated in the processing.
In order to select the SSOs observed by Gaia, some filters are applied. Every transit contaminated by the presence of a star or another bright SSO is discarded a priori. In

the case of DR3, only SSOs which have been detected on at least eight transits over the 34 months of observations are retained. The remaining observations are then provided to a multi-step procedure, which basically match the approximate sources positions to the ephemerides of numbered SSOs.
The first criterion for matching a moving object detected by Gaia to a numbered asteroid relies on the accuracy of the orbit available in the catalogues. The ephemerides uncertainty is computed at the epoch of each transit and if it is larger than a limit value the transit is discarded. This step removes all SSO candidates observed by Gaia having uncertain orbits that could lead to spurious identifications.
The second criterion takes into account the relative predicted positions of numbered asteroids compared to the recorded transit positions of Gaia observations. The SSO candidate is matched to a numbered asteroid if the angular distance between prediction and observation is smaller than a given value.
A third criterion based on the difference between the apparent magnitudes of the observed transit and of the numbered SSO could in principle be introduced. However, apparent magnitudes of known SSOs are characterized by large uncertainties, usually of the order of 1 *mag*. The magnitude of a SSO at a given epoch depends in fact on the specific geometry of the observation. Furthermore, because of the intersection between the ecliptic and the scanning plane (Section 3.2), asteroids are never observed by Gaia at less than 45° elongation from the Sun and from the anti–solar direction. In addition to being observed at non negligible phase angles, an object may be detected by Gaia at multiple epochs in a huge variety of configurations, such as high and low proper motions and difference distances from the Sun. The apparent magnitudes computed by Gaia for the same object vary with the epoch, therefore this criterion based on magnitudes has not been implemented in Gaia DR3.
The rate of correct identifications obtained with this matching process was tested with a validation procedure and was found to be close to 100%.

Gaia DR3 also contains a small fraction of asteroids that have been found by their motion, but have not been matched to any known SSO. More and more unmatched observations will be linked to known objects thanks to the continuous growth of asteroid catalogues. For example, in 2017, 1320 transits resulted to not be linked to any known SSO. The same transits were then again analyzed in June 2021 by using an updated ephemerides database. This time, of the 1320 unmatched observations, 675 were matched to known asteroids.

3.4 Spectrophotometry of asteroids in Gaia DR3

Along with astrometric measurements, Gaia DR3 includes reflectance spectra in the 374 – 1034 nm wavelength range for 60 518 numbered asteroids (Gaia Collaboration, 2022a). The vast majority of them have been observed in the apparent magnitude range between 18 and 20, while the brightest appeared slightly brighter than magnitude 8 and the faintest presented a mean magnitude of 20.19.
The majority of the SSOs with a reflectance spectrum in Gaia DR3 are Main Belt asteroids, while 105 of them are NEAs, 729 Mars Crossers, 477 Jupiter trojans, 5 Centaurs and 7 TNOs. Spectra of satellites, unmatched asteroids and comets have not been included in Gaia DR3.

Each SSO has been observed by Gaia at multiple epochs and, for each epoch (corresponding to a transit on the focal plane), internally calibrated BP and RP spectra were extracted. Epoch spectra corresponding to the same object were then averaged together to produce a mean reflectance spectrum. Before doing that, a filtering procedure was applied.
Gaia observed 158 152 objects for a total of 3 218 410 transits. It was decided to keep a transit only if both RP and BP spectra were correctly computed. 529 694 epochs were therefore discarded because their BP or RP values were null. Other 196 970 transits were discarded because of some issues in the CCDs, such as hot columns or because of problems in the calibration. The filtering removed all transits for 23 SSOs. In addition, it was decided to discard all sources with less than three transits, which concerned 564 SSOs.
An epoch spectrum consists of a sample of 60 measurements corresponding to a set of pseudo-wavelengths and fluxes calibrated in the internal reference system of Gaia. Epoch spectra are not published in the DR3.
For a given epoch t, the reflectance spectrum $R(\lambda_i)_t$ at a given wavelength λ_i was determined by dividing the flux $f_t(\lambda_i)$ of the epoch spectrum with the solar analogue mean spectrum $F(\lambda)$ (Section 4.1):

$$R(\lambda_i)_t = \frac{1}{\xi_t} \frac{f_t(\lambda_i)}{F(\lambda)} \tag{3.1}$$

where ξ_t is a factor normalizing the spectrum to 1 at 550 nm and it is the same for the BP and the RP spectrophotometers. The uncertainties $\sigma_{R(\lambda_i)}$ on $R(\lambda_i)_t$ were computed by propagating the uncertainties on $f_t(\lambda_i)$ and $F(\lambda)$ assuming they are independent.
After the calibration, the RP and BP parts of a spectrum were expected to have very similar values in the range of wavelengths in common, between 630 nm and 680 nm.

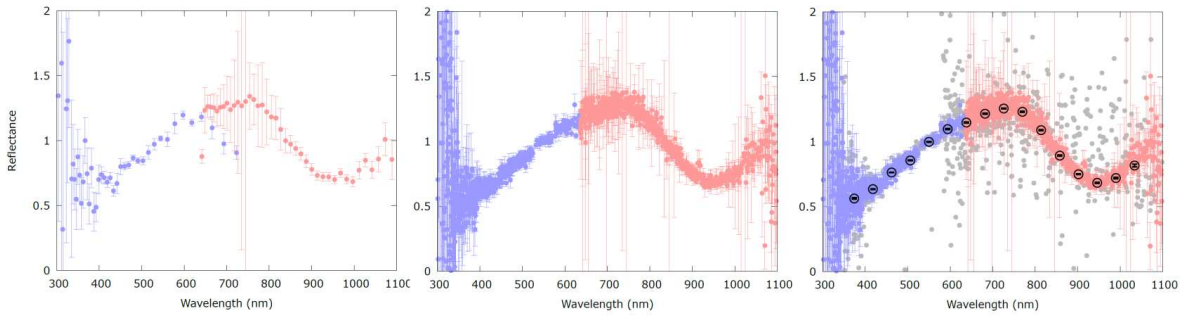

FIGURE 3.2: The procedure to compute mean reflectance spectra illustrated for the case of (1459) Magnya. A more complete description of the figure is found in the text. Credit: Gaia Collaboration (2022a).

In some problematic cases however the BP part had a lower flux than the RP part or vice versa. In addition, some of the brightest SSOs displayed high spectral frequency variability, while faint objects were found to be affected by large errors at $\lambda < 400$ nm and $\lambda > 1000$ nm.
In order to mitigate the contamination due to these problems, the reflectance spectra were processed and reduced to a set of 16 fixed bands with wavelengths λ_j every 44 nm in the range $374 - 1034$ nm. In order to create the mean reflectance spectrum, the median and the absolute median deviation (MAD) of all the values contained inside a band $[\lambda_j, \lambda_{j+1}]$ were computed. For each band, a σ-clipping procedure was applied to filter out all the values outside the range [median - 2.5 MAD, median + 2.5 MAD]. The procedure was repeated twice to remove the outliers. Finally, the remaining values inside each band were averaged together to obtain the final mean reflectance spectrum, which was then normalized at $\lambda = 550$ nm. The procedure is illustrated in Figure 3.2 for the asteroid (1459) Magnya. The left panel shows an example of epoch spectrum, with the overlapping between the BP and the RP spectra (colored in blue and red respectively). In the middle panel, all epoch reflectance spectra are plotted, with well evident the large errors affecting the data at $\lambda < 400$ nm and $\lambda > 1000$ nm. In the right panel, data filtered out by the σ-clipping procedure are shown in grey, while the final mean reflectance spectrum sampled in the 16 bands is over-plotted with black circles.
44 385 sources were removed after the filtering with this procedure. Mean reflectance spectra of 111 818 objects were computed using 1 980 726 transits.

Mean reflectance spectra associated to faint asteroids present a lower quality compared to spectra associated to bright sources. The average signal-to-noise ratio was therefore

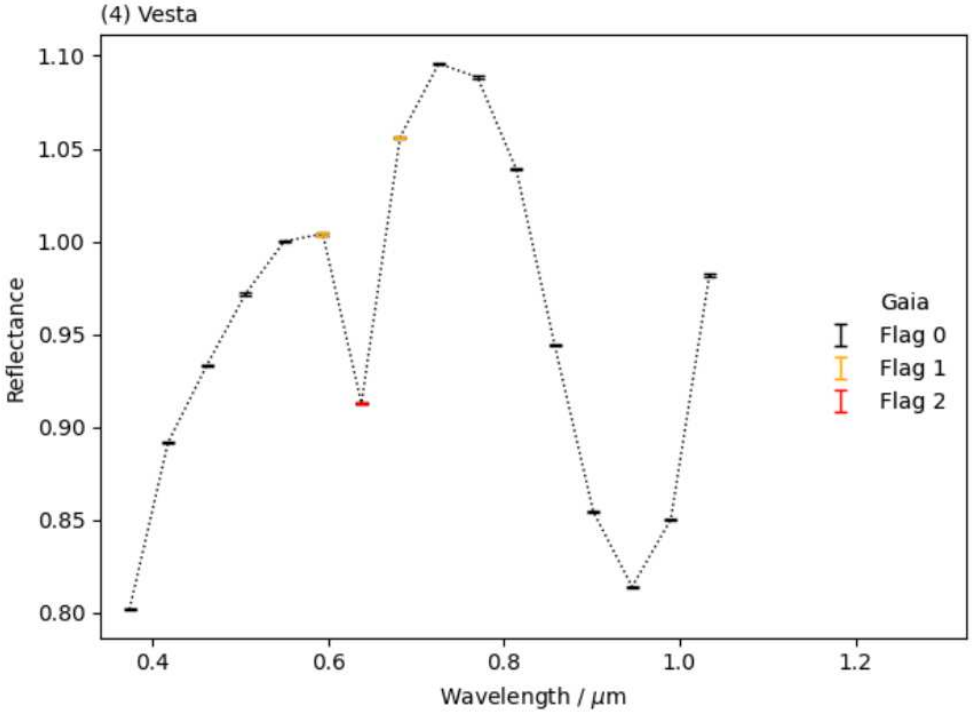

FIGURE 3.3: Mean reflectance spectrum of (4) Vesta. Black symbols have RSF=0, yellow symbols have RSF=1 and red symbols have RSF=2.
Plot obtained using the *classy* Python package.

introduced to assess the quality of a spectrum. It is defined as:

$$SNR = \frac{1}{12} \sum_{i=3}^{14} \frac{R(\lambda_i)}{\sigma_{R(\lambda_i)}} \tag{3.2}$$

Data points at 374, 418, 990 and 1034 nm where not considered in the computation since they are usually affected by large systematic errors.
The histogram distribution of the SNR showed a peak at $SNR = 13$. Objects below this peak are characterized by noisy spectra and magnitudes higher than 19, while objects at $SNR > 13$ are associated to more accurate spectra and bright magnitudes.
In Gaia DR3 it was decided to publish only the spectra characterized by $SNR > 13$, for a total of 60 518 asteroids. The remaining spectra will instead be released in Gaia DR4, expected for the late 2025.

$SNR > 13$ does not guarantee that spectra are scientifically exploitable, since they may still be affected by noise, variability or other imperfections. In order to assess the quality of the data points, a reflectance spectrum flag (RSF) system is introduced. A value equal

to 0, 1 or 2 is assigned to every wavelength depending if the reflectance point in that band is validated, suspected to be of poorer quality or compromised.
For a given spectrum, as a first step, RSF=0 is assigned to every wavelength. Then, RSF=2 is assigned to those points whose corresponding reflectances or uncertainties are not numbers (NaN), to those points whose reflectances are larger than 2.5 or smaller than 0.2 and to those points whose uncertainties are larger than 0.5. The points whose RSF remains zero are then fitted by mean of a smoothing natural spline function. Points with a distance from the smoothing spline larger than twice the corresponding uncertainty are marked with RSF=1. If the distance is larger than three times the corresponding uncertainty RSF=2 is assigned.
An example of spectrum with flagged data points is reported in Figure 3.3 for (4) Vesta. Black symbols with RSF=0 are good quality points, yellow symbols with RSF=1 are poor quality points, while the red point is clearly compromised and therefore flagged with RSF=2.

More information about the Gaia mission and the Data Release 3 can be found in the Gaia online documentation (*https://gea.esac.esa.int/archive/documentation/GDR3/index.html*).

Chapter 4

Modelling of Gaia spectra

4.1 Introduction

Gaia asteroid reflectance spectra published in the Data Release 3 are characterized by systematic deviations in the near infrared and in the near ultraviolet when compared to ground-based spectra. In particular, in the NUV Gaia spectra have been found to have lower reflectance than ground-based spectra and therefore they sistematically lie below them, while in the NIR they present higher reflectance, therefore sistematically lying above them.

The deviations in the blue have been found to be related to the mean solar analogue spectrum used in the calibration of Gaia asteroid reflectance spectra (Tinaut-Ruano et al., 2023).
As described in Section 2.1.4, from Earth we observe the sunlight reflected from the asteroid altered by the surface material. In order to study the asteroid spectral features, the asteroid spectrum needs to be divided by the solar spectrum. Since it is unpractical to observe the Sun with the same instrumentation used to detect asteroids, solar analogues (SAs) are used at its place, selected for their known similarities to the solar spectrum. If a solar analogue with a spectrum slightly different from that of the Sun is used in the data reduction, a systematic difference is introduced in the reflectance spectra. This is exactly what caused the deviations in the NUV in DR3 reflectance spectra, as reported by Tinaut-Ruano et al. (2023).
In order to reduce the observations, the Gaia team selected a list of SAs widely used in the literature and took an average of their Gaia spectra. However, the suitability of many

Wavelength [nm]	Correction Factor
374	1.07
418	1.05
462	1.02
506	1.01
550	1.00

TABLE 4.1: The multiplicative correction factors mitigating systematic deviations in the NUV proposed by Tinaut-Ruano et al. (2023).

of these SAs was assessed only in the visible, while only two of them, Hyades 64 and 16 Cyg B, were known to have almost indistinguishable NUV spectra from the Sun. Tinaut-Ruano et al. (2023) proved that the spectrum of the mean solar analogue used by the Gaia team differed significantly, up to 10%, from the spectrum of Hyades 64 for wavelengths below 550 nm. The division by this mean spectrum to obtain asteroid reflectance spectra thus introduced systematic deviations in the NUV.
Tinaut-Ruano et al. (2023) proposed empiric correction factors to be applied to all asteroid reflectance spectra to mitigate the deviations in the NUV. They computed the corrections by quantifying how much spectra of objects of a specific spectral type needed to be multiplied in order to match the corresponding template Bus-DeMeo spectra. The multiplicative factors, necessary for all wavelengths below 550 nm, are reported in Table 4.1.

To date, July 2023, the systematic deviations in the near infrared have not yet an explanation. However, understanding how they work is necessary before trying to do any science with Gaia data, since for example they may alter the spectrum of an object of a certain spectral type making it appearing as another spectral type. An example is shown in Figure 4.1 for the S- type (39) Laetitia, in blue, and the L- type (729) Watsonia, in orange. The systematic deviations modified the spectrum of (39) Laetitia, increasing its reflectance in the NIR and making it very similar to a L- type. Since a taxonomy classification based on Gaia spectra was a crucial aspect for the work I wanted to conduct with asteroid families, I therefore first explored possible correlations of the deviations in the NIR with many parameters, such as the magnitude and the spectral type of the observed objects.
Determining if the deviations correlate with one or more of the previously cited parameters is necessary for the correct interpretation of the spectra. In particular it is fundamental to assess if Gaia spectra are internally coherent, such that asteroids of the same spectral

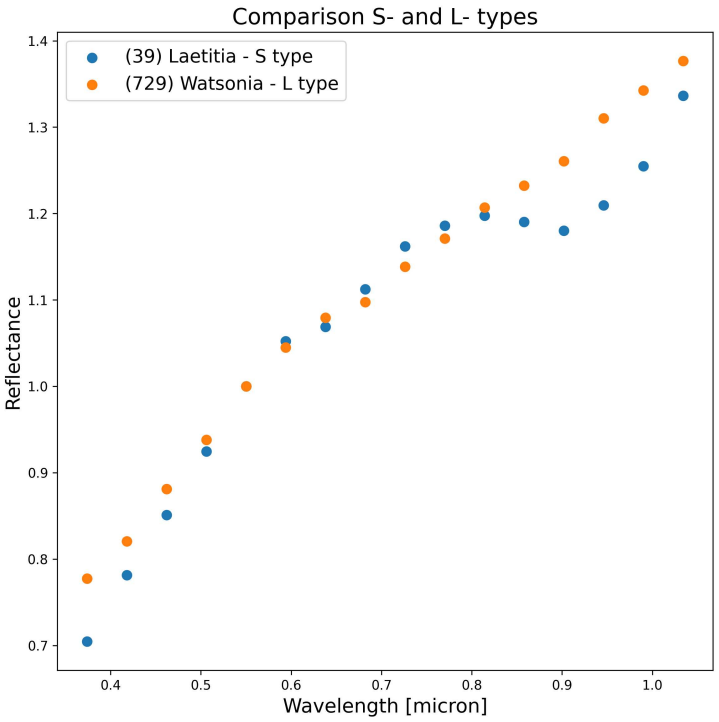

FIGURE 4.1: Gaia spectra of the S- type (39) Laetitia, in blue, and the L- type (729) Watsonia, in orange. Despite being of two different spectral types, the two spectra appear to be very similar due to the deviations in the NIR.

type have similar spectra, or if the deviations do not depend on the surface compositions, which would complicate their interpretation. If Gaia reflectance spectra were internally coherent it would be possible to quantify systematic deviations from ground-based spectra and taxonomic systems. The internal consistency of Gaia spectra could in principle allow to compare spectra with literature and to use the criteria proposed by taxonomic systems to classify Gaia observations, without however expecting a perfect correspondence between Gaia spectral types and literature.
Some examples of spectra affected by systematic deviations are reported in Figure 4.2 for the S- type (3) Juno, on the left, and the C- type (31) Euphrosyne, on the right. Gaia data are shown in black, while the ground-based SMASS spectra (Section 2.1.5) are reported in light blue. Gaia spectra are in good agreement with SMASS spectra in the NUV after they have been corrected with the multiplicative factor proposed by Tinaut-Ruano et al. (2023), while in the NIR they are systematically above the ground-based spectra. The entity of the deviations is different in the two cases.

In the following sections all my attempts to find a dependency of the deviations with several parameters are described. After a correlation with the spectral type was found,

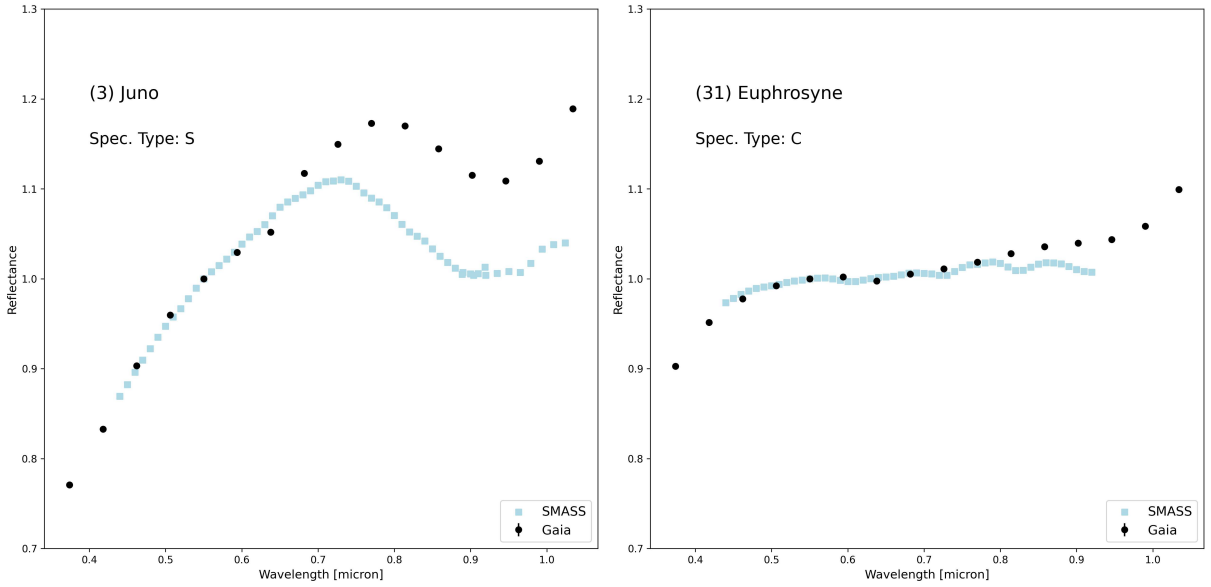

FIGURE 4.2: Example spectra of (3) Juno on the left and (31) Euphrosyne on the right. Gaia data (in black) are found to be systematically redder than the ground based SMASS spectra (in light blue).

an idea of a correction scheme analogue to Tinaut-Ruano et al. (2023) is proposed. Finally, a color taxonomy based on Gaia spectra that is currently being developed at the Observatoire de la Côte d'Azur by Alexey Sergeyev is described (A. Sergeyev, Personal Communication, 2023).

4.2 Pre-processing of the reflectance spectra

In order to quantify the systematic deviations in the NIR compared to the template spectra, I first considered asteroids of known spectral type. In the first instance I considered objects observed by Gaia and included in the Bus-DeMeo taxonomy, but I then moved to the Mahlke classification. The two systems are based on similar template spectra, but the Mahlke taxonomy includes more objects than Bus-DeMeo, thus improving the statistical significance of the analysis.

I divided Gaia observations in four luminosity classes: brighter than 15 *mag*, between 15 and 17 *mag*, between 17 and 19 *mag* and fainter than 19 *mag*. For a given spectral type, I computed the difference from the according template spectrum for each luminosity class, to search for a dependency of the deviations with the magnitude.

For each spectral type, I aimed to have few tens of objects for every luminosity class. I

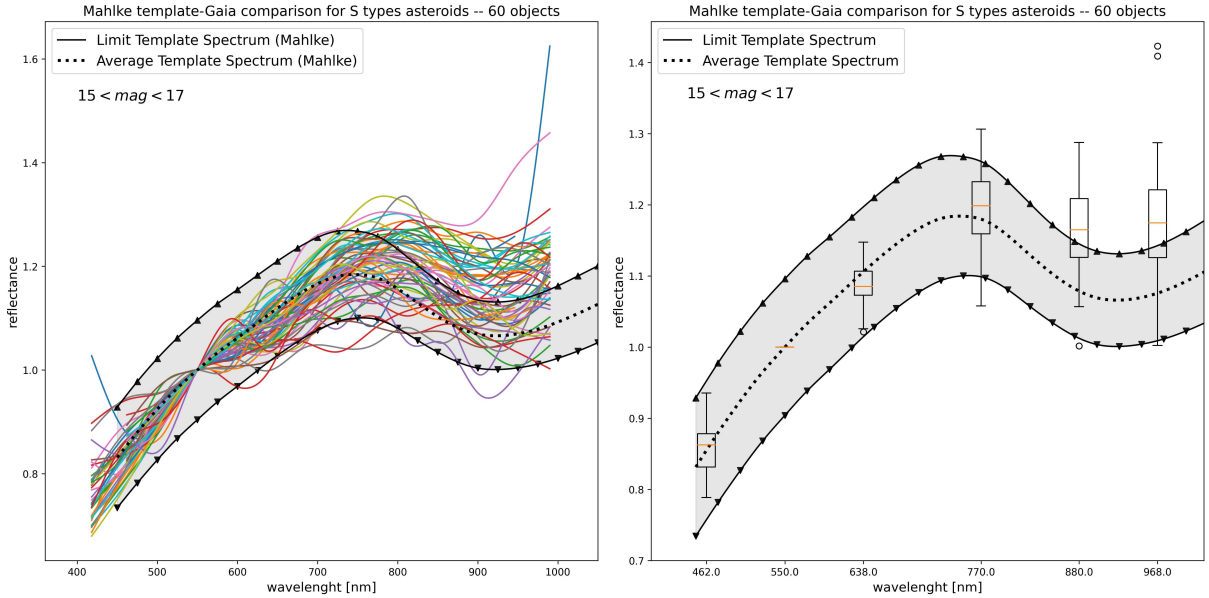

FIGURE 4.3: Left panel: Gaia observations of S- type asteroids between 15 and 17 *mag* compared to Mahlke template spectrum.
Right panel: The processed mean spectrum compared to the Mahlke template.

therefore had to limit my analysis to the S-, C- and P- types, since the other classes did not contain enough members to allow statistical studies to be conducted.
To make sure to not use misclassified objects or spectra presenting defects introduced during the calibration process, I checked by eye every Gaia spectrum used in the analysis. In addition, I did not consider Gaia spectra of NEAs because they are known to be affected by large errors due to their high proper motions.

For each spectral type and luminosity class, I combined together all Gaia observations in order to obtain an average spectrum from which quantify the deviation from the template spectrum. In order to do that, I first had to pre-process Gaia spectra. As already discussed in Section 3.4, data at 374 *nm* and at 1034 *nm* are characterized by large uncertainties, therefore I chose to remove these wavelengths from the computation. In addition, I discarded points flagged with RSF=1 or RSF=2. The remaining data points were interpolated by mean of a Python CubicSpline function.
The interpolated spectra were then compared to the template spectrum. An example is shown in the left panel of Figure 4.3 for S- types between 15 and 17 *mag*. Gaia observations are shown as multi-colors lines. The grey shaded area is the Mahlke template, which is included between the template spectrum (black dotted line) and its uncertainties (black continuous lines). Most Gaia spectra after 700 *nm* lied above the template spectrum.
In addition some spectra, in particular of faint objects, were affected by large oscillations,

therefore using them as they were would have spoiled the computation of the average spectrum. I therefore decided to reduce the number of data points from fourteen to six by creating a set of wavelengths at [462, 550, 638, 770, 880, 968] nm centered on specific spectral features. The wavelength at 462 nm was taken by averaging the data at 418, 462 and 506 nm and it describes the spectral behaviour in the NUV. The normalization at 550 nm was left unchanged. The wavelength at 638 nm was created by averaging the data at 594, 638 and 682 nm and the first derivative in this point can be interpreted as an indication of the spectral slope. The wavelength at 770 nm is obtained by averaging the data at 726, 770 and 814 nm and describes the maximum of the spectrum for S- complex objects. The wavelength at 880 nm was created by mediating the data at 858 and 902 nm and it samples the absorption band (when present). Finally, the wavelength at 968 nm was obtained by averaging the data at 946 and 990 nm and describes the spectral behaviour in the NIR.
I therefore reduced every Gaia spectrum to this set of six wavelengths and then I averaged them together in order to obtain a final mean spectrum for each spectral type and luminosity class. An example of mean spectrum sampled in the six wavelengths is reported in the right panel of Figure 4.3 for S- types between 15 and 17 mag. The grey shaded area represents the Mahlke template as in the left panel. The boxes indicate the reflectance values of the mean spectrum. The orange lines indicate the average value of the reflectance at each wavelength, while the rectangles indicate one quantile and the black lines two quantiles. The circles represent outliers deviating from the average values. The average spectrum follows the template at short wavelengths, while it deviates from it in the NIR. The entity of the deviations increases with the wavelength.

For every spectral and luminosity class, I then calculated the difference of the processed mean spectrum from the Mahlke template. The relative deviations were computed for each one of the six wavelengths λ_i as $Dev[\lambda_i] = (Gaia[\lambda_i] - Tem[\lambda_i])/Tem[\lambda_i]$, where $Gaia[\lambda_i]$ is the reflectance value of the Gaia mean spectrum at λ_i and $Tem[\lambda_i]$ is the reflectance value of the template spectrum at λ_i. Normalizing to the value of the template allows to compare results obtained from different spectral types.
The results obtained for S- type objects are reported in Figure 4.4, where the black dotted line identifies the zero level. The luminosity classes have been shifted by 5 nm along the x-axis for better clarity. Data are in good agreement with the template spectrum for wavelengths between 400 and 700 nm. At longer wavelengths the Gaia observations deviate from the expected behaviour. The deviation is zero at 550 nm because both spectra are normalized at this wavelength.

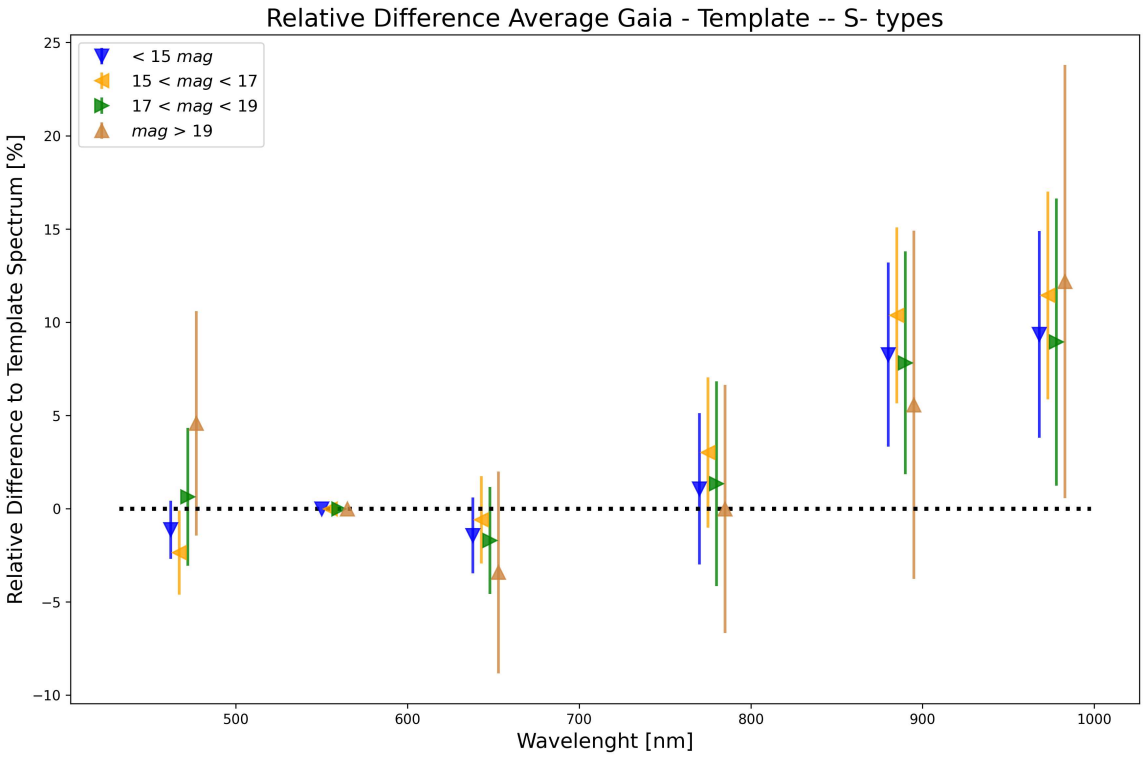

FIGURE 4.4: Relative difference from the Mahlke template spectrum for S- type asteroids subdivided in the four luminosity classes. Data at the same wavelength have been shifted by 5 *nm* along the x-axis for a better visualization. The black dotted line identifies the zero level.

Similar plots are reported in Figure 4.5 and Figure 4.6 for asteroids belonging to the C- and P- classes, respectively. For all the three cases, the data of objects fainter than 19 *mag* are affected by very large uncertainties. This confirms that, despite being observed by Gaia, faint asteroids present flawed spectra which cannot be analyzed singularly.
In all cases, no dependency of the deviations with the magnitude is observed. All classes show in fact approximately the same behaviour, with the deviations that are always consistent within the errorbars. However, there seem to be a dependency with the spectral type. In particular, I noticed that S- types deviate more from the template than C- types, than in turn deviate more than P- types. I therefore decided to explore this aspect in more detail.

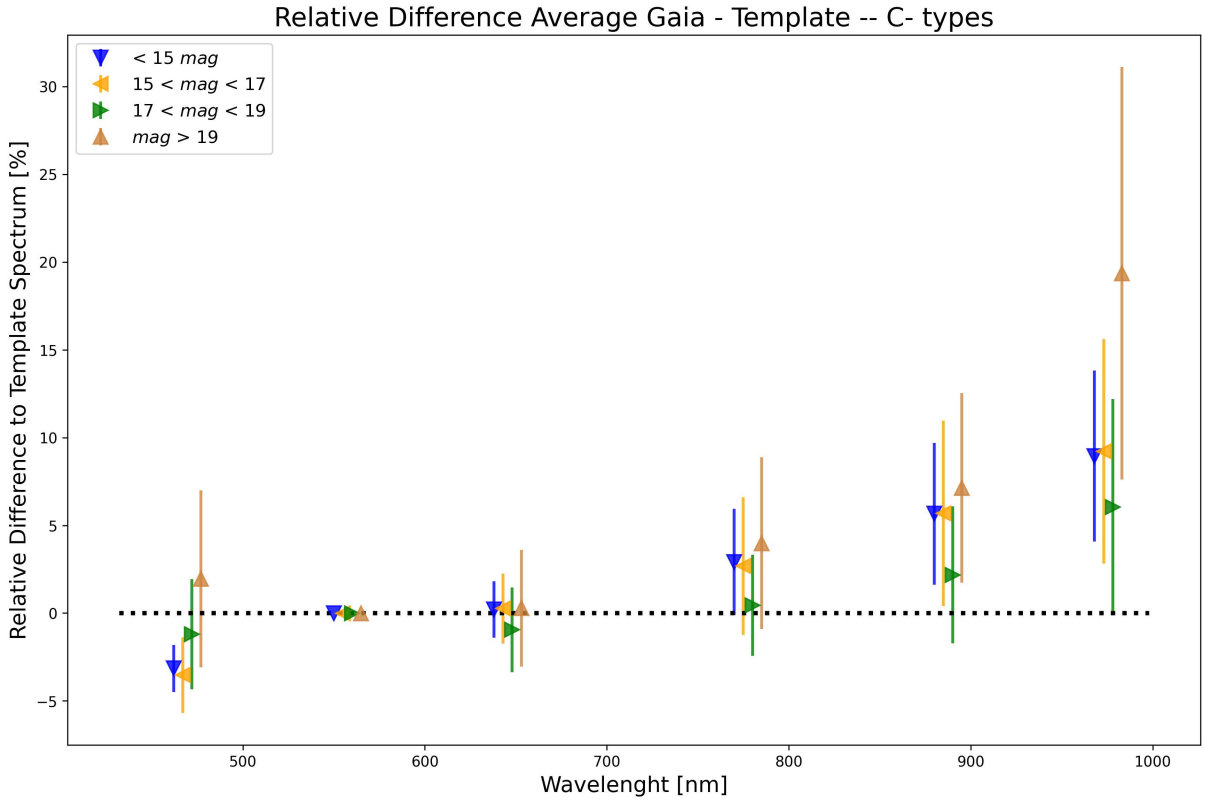

FIGURE 4.5: Same as Figure 4.4, but for C- type asteroids.

4.3 Correlation with the spectral type

In order to analyze the correlation of the deviations with the spectral type, I dropped the luminosity classes and grouped together asteroids belonging to the same taxonomy class. For each spectral type, I selected all objects in common between Gaia observations and Mahlke taxonomy and I retained only the classes containing a significant number of asteroids, at least 14. Removing the subdivision in luminosity allowed me to analyze more spectral types. For rich classes, such P, S, V, C and Ch, I did not consider all asteroids in common between Mahlke taxonomy and Gaia observations, but only a subsample homogeneous in magnitude.

To quantify the entity of the deviations for each spectral type I followed the procedure described in Section 4.2. For each spectral type, I first discarded NEAs, misclassified objects and flawed spectra. I then computed the mean reflectance spectrum for each spectral type sampled in the set of six wavelengths from which I calculated the deviations from the relative Mahlke template. The results are shown in Figure 4.7, with the wavelength on the x-axis and the relative difference from the template on the y-axis. The black dotted

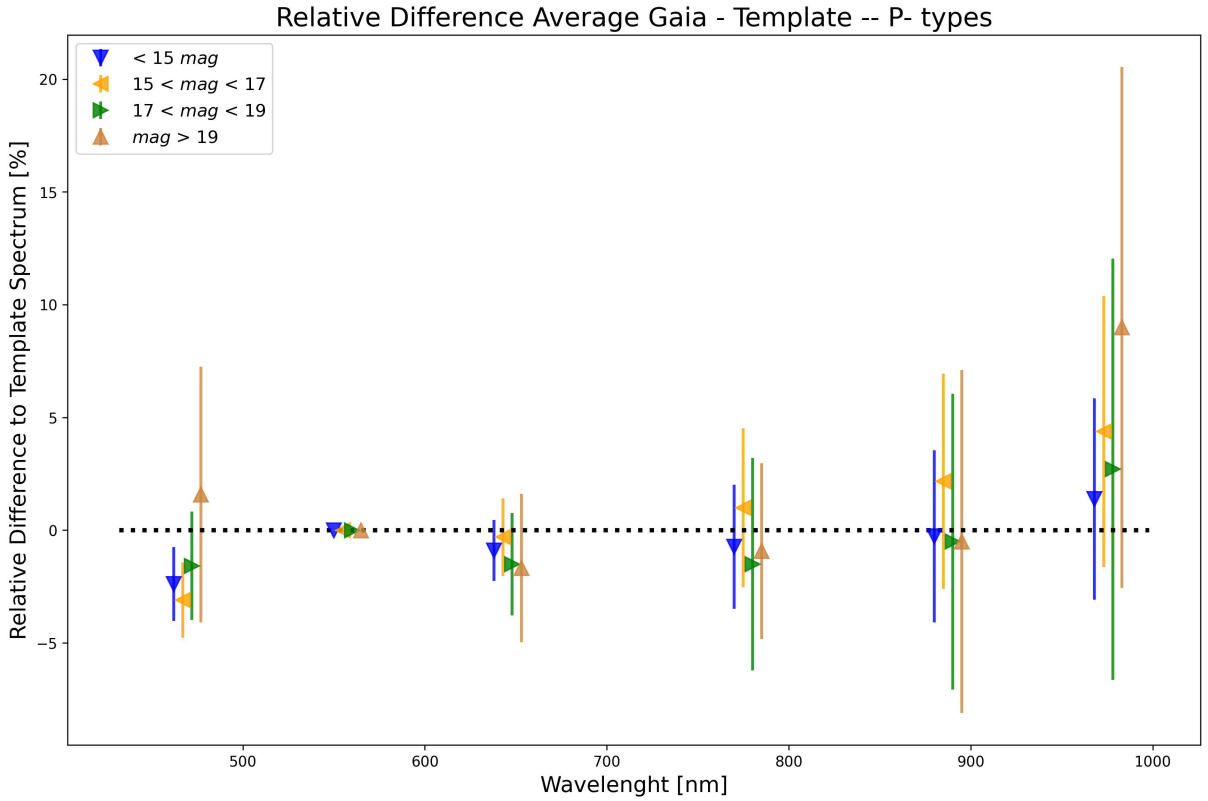

FIGURE 4.6: Same as Figure 4.4, but for P- type asteroids.

line identifies the zero level, while each spectral type is reported with a different color and marker. Each class has been shifted along the x-axis by 2 *nm* for a better visualization. Although measurements are characterized by large errors, a dependency of the deviations with the spectral type is clearly visible. In general, M- and P- objects are in good agreement with the template, C- and Ch- objects require a small correction, while S- types a larger one. A- types need a larger correction than S- types, while D- types an intermediate one between the S- and the C-complex. L- and K- types, despite having few objects, seem to require a correction similar to S- types.

Independently from the entity of the correction, almost all spectral types present the same behaviour, with the deviations increasing with the wavelength. The only exception is the V- class, which deviates more from the template around 900 *nm* than at the extreme. In order to confirm that this anomalous behaviour was intrinsic to Gaia observations and not due to the reduction to the six wavelengths, I explored more carefully the V- types. I therefore computed the deviations for the original Gaia spectra sampled in the sixteen wavelengths. In addition, to be sure that this behaviour was not due to the Mahlke taxonomy, I repeated the same procedure for V- types classified by Bus-DeMeo. For both

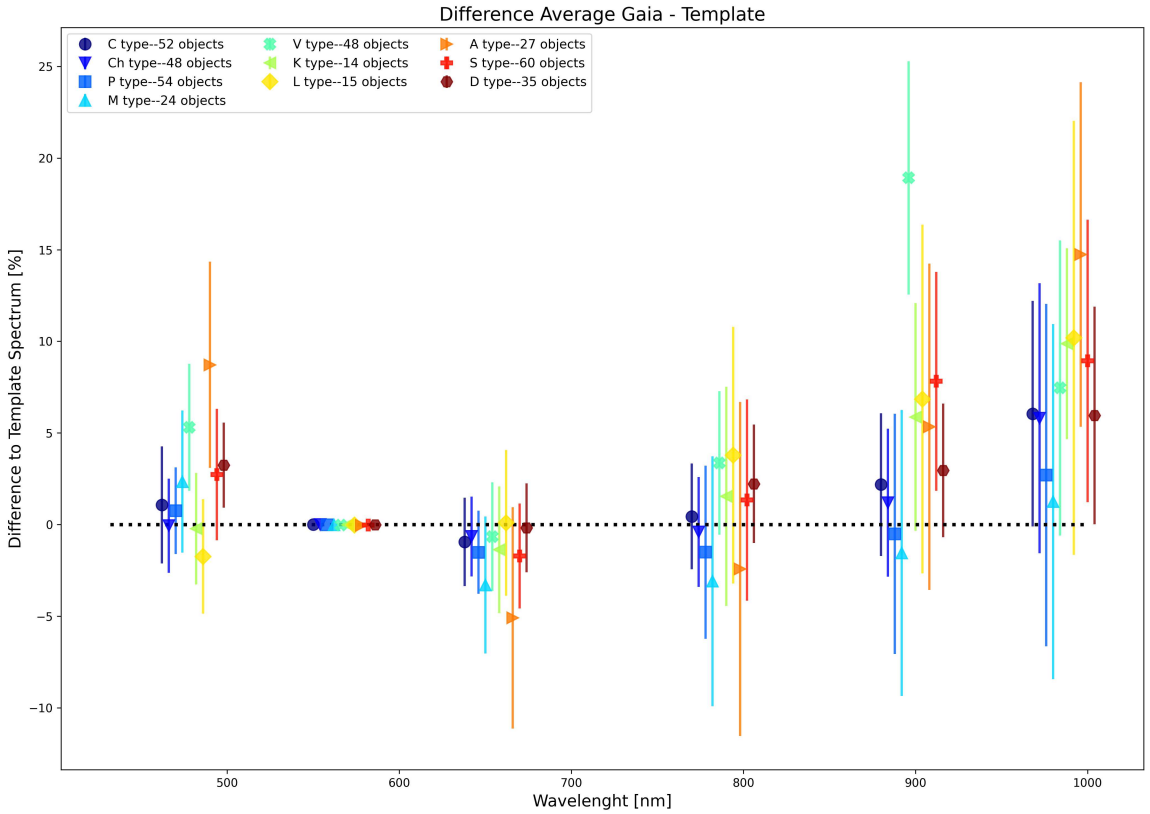

FIGURE 4.7: Relative difference from the Mahlke template spectrum for asteroids of different spectral types. Data at the same wavelength have been shifted by 2 *nm* along the x-axis for a better visualization. The black dotted line identifies the zero level.

taxonomic systems, V- types were found to be systematically above the template spectrum between 820 and 920 *nm*, while they were in better accordance at longer wavelengths. I therefore confirmed that the anomalous behaviour of V- types was real and intrinsic to the observations.

With this analysis I proved that Gaia spectra are internally coherent, since objects of the same spectral type present similar Gaia spectra, allowing to quantify the differences from ground-based and template spectra. However, since the entity of the deviations depends on the spectral type, and in turn on the surface composition, it is impossible to apply a single set of multiplicative factors to correct the observations in the NIR as it was done in the NUV by Tinaut-Ruano et al. (2023). Different correction factors are therefore required for each spectral type, or at least broad complex.

4.4 Correction procedure

The analysis described in Section 4.3 showed the impossibility to apply a unique correction in the NIR to all objects due to the correlation of the deviations with the spectral type. However, Gaia observed numerous objects of unknown spectral type, therefore the correction procedure must be independent from taxonomy.
In order to mitigate the deviations in the NIR, a method to correct the spectrum of an object must be developed. I explored two possibilities: in Sections 4.4 and 4.5 a procedure based on specific spectral features allowing to correct observations is described, while in Section 4.6 a taxonomy based on colors is built from Gaia observations.

In order to create a correction procedure, I wanted to find a combination of spectral features giving an indication of the spectral type of the considered object. I decided to focus on the slopes of specific regions of the spectrum. The slope has in fact the advantage of incorporating the differences between different parts of the spectrum, thus containing more information than a single band.
I tried different options until I found that a combination between the slope in the red and the maximum of the reflectance reproduced quite well a color-color diagram. In particular I computed the slope in the range between 800 nm and 900 nm, approximately corresponding to the olivine absorption band of S- types, while I selected the maximum of the reflectance in the interval between 730 nm and 830 nm.
I defined the two parameters using as reference S- types spectra interpolated by mean of a cubic spline function, as the ones shown in the left panel of Figure 4.3, but then I extended the procedure in order to work for every spectra, of known or unknown spectral type. I selected two wavelengths interval, the first centered on the maximum of S- types (from 730 nm to 830 nm) and the second centered on the minimum of the absorption band (from 870 nm to 970 nm). Inside the first interval I selected the maximum of the reflectance, which was taken as the first parameter, while inside the second interval I selected the minimum of the reflectance. The minimum and the maximum were then fitted by a straight line, whose slope was taken as the second parameter.
Example spectra are reported in Figure 4.8, where on the left is the S- type (3) Juno and on the right the C- type (31) Euphrosyne. The red star marks the maximum of the reflectance, whose y-value is taken as first parameter, while the slope of the black line is taken as second parameter. The right panel proves that the procedure to find the two parameters despite being built for S- types can be applied to every other spectrum.
The combination of the two parameters gives a good indication of the spectral type of the

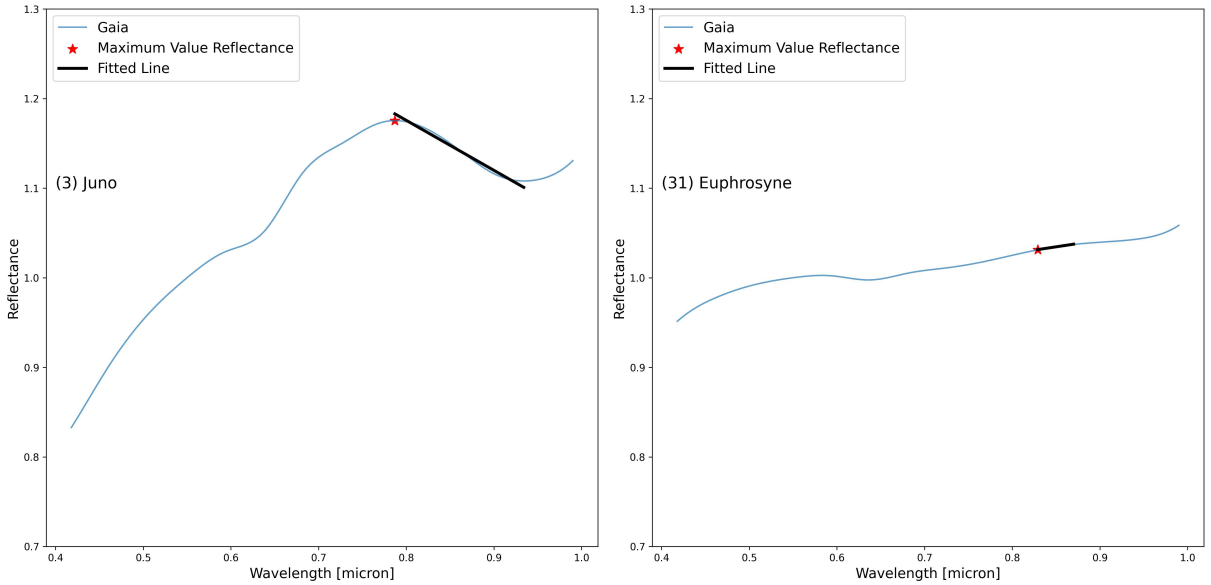

FIGURE 4.8: Example spectra of (3) Juno (left) and (31) Euphrosyne (right) with highlighted the maximum of the reflectance (red star) and the slope in the red (black line). The two parameters are used in the correction procedure.

considered object. From the slope in the red it is in fact possible to recognize V- types, which are characterized by very deep absorption band and therefore by large negative slopes, from A- and S- types, with intermediate negative slopes, and D-, M-, Z- and P- types, which instead present positive slopes. The maximum of the reflectance allows instead to separate the S- complex, centered at 1.20, from the C- complex, centered at 1.05.
Note that the slope in the blue could be used at the place of the maximum reflectance value given the correlation between these two quantities.

In order to develop a correction procedure, I first created a template slope-maximum reflectance diagram for Gaia spectra with $SNR > 100$ of asteroids included in Mahlke taxonomy, indicated in black with the correspondent taxonomy class letter in Figure 4.9. Asteroids arrange in this diagram forming two clusters, the S- complex on the center and the C- complex on the left. Some endmember classes are recognized: the A- class on the right of the S- complex, the V- class below the S- complex, the D- and Z- classes in the upper part of the diagram and the L- and K- classes between the D- types and the S- complex. The overall shape of the diagram in Figure 4.9 reminds very closely the one of a color-color diagram, such as the one reported in Parker et al. (2008) for the asteroids observed in the SDSS filters by the Sloan Digital Sky Survey Moving Object Catalog 4.

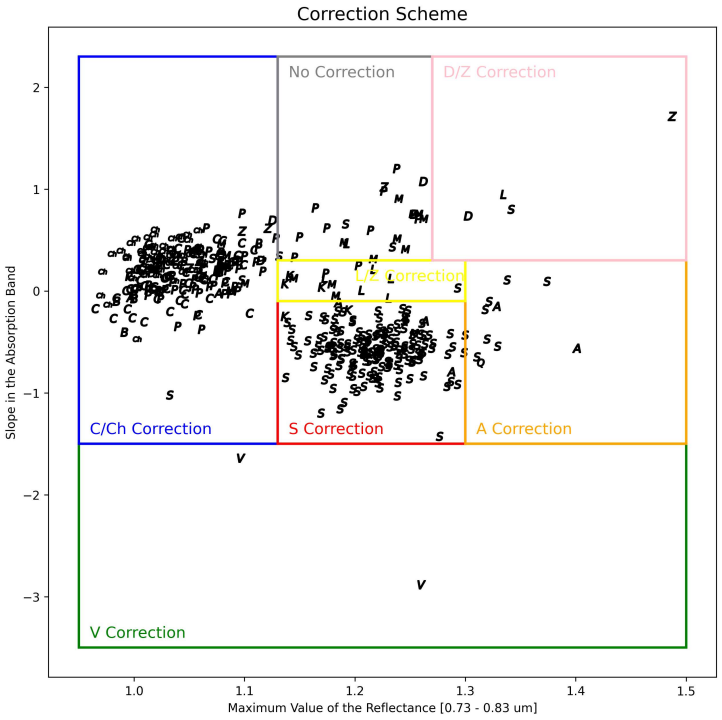

FIGURE 4.9: Slope - maximum reflectance diagram of asteroids observed by Gaia with $SNR > 100$ and classified by Mahlke (black letters). The seven regions used to define the corrections are reported on top of the diagram in different colors with the relative names.

As discussed in Section 4.3, asteroids of different spectral types require different corrections. I therefore empirically defined seven regions on top of the diagram, to each of which I assigned a name depending on the dominant spectral type inside them following Mahlke taxonomy. For example, the green area at the bottom is the V region and the red in the center is the S region. The boundaries in maximum reflectance and slope are reported in Table 4.2.

To each region I then assigned specific correction factors depending on the dominant spectral type within them. Figure 4.7 showed that L- and K- types needed similar corrections and therefore they have been united in the yellow area. Similarly, C- and Ch- types have been united in the blue area on the left. The correction factors relative to each region were computed by quantifying how much the average spectrum of the dominant spectral type needed to be lowered down to match the template spectrum. I found that correction factors were necessary for all wavelengths redder than 726 nm.

For the seven regions of the diagram, I defined six correction tables analogous to the one reported in Table 4.1 for the NUV. The grey area dominated by M- and P- types does not require any correction, since the Gaia spectra of these objects do not deviate from

	Upper Slope	Lower Slope	Upper Refl.	Lower Refl.
D/Z Region	2.3	0.3	1.50	1.27
S Region	-0.1	-1.5	1.30	1.13
A Region	0.3	-1.5	1.50	1.30
V Region	-1.5	-3.5	1.50	0.95
C/Ch Region	2.3	-1.5	1.13	0.95
L/K Region	0.3	-0.1	1.30	1.13
M/P Region	2.3	0.3	1.27	1.13

TABLE 4.2: Boundaries in slope and maximum reflectance value for the seven regions of the diagram in Figure 4.9.

WAVELENGTH (nm)	D/Z CORR.	S CORR.	A CORR.	V CORR.	C/Ch CORR.	L/K CORR.
726	0.98	0.99	1.00	0.99	0.99	0.98
770	0.98	0.97	0.98	0.94	0.98	0.96
814	0.98	0.94	0.94	0.87	0.97	0.95
858	0.97	0.92	0.93	0.83	0.96	0.94
902	0.97	0.92	0.92	0.85	0.95	0.94
946	0.95	0.91	0.90	0.91	0.94	0.93
990	0.92	0.90	0.88	0.98	0.93	0.90
1034	0.89	0.88	0.85	0.98	0.91	0.87

TABLE 4.3: Correction factors associated to the six regions defined in the slope - maximum reflectance value diagram.

the template spectra (Figure 4.7). The other six regions require a correction: the multiplicative factors are reported in Table 4.3.
Using this diagram it is possible to correct the spectrum of any object without knowing its taxonomy. After having derived the maximum reflectance and the slope with the previous procedure, one can see where the object falls in the diagram and apply the corresponding multiplicative factors. It is important to keep in mind that an object of known spectral type could fall in a region corresponding to another spectral type, such as for example the S- types in the A- region or the D- types in the M- and P- region. The object will be corrected according to the region it lies, regardless of its known spectral type.

In order to test the procedure I applied the correction to the first 60 numbered asteroids

in common between Gaia observations and Mahlke taxonomy (from (1) Ceres to (75) Eurydike). I selected the brightest asteroids since they present good Gaia spectra, while fainter objects are characterized by large errors and oscillations. In addition, for some of these objects SMASS spectra are also available. A comparison between the original Gaia spectra (in red), the corrected Gaia spectra (in black), the template spectrum of the associated Mahlke spectral type (the dotted black line) and the SMASS spectra (in light blue, when available) is reported in Figures 4.10 and 4.11. Gaia spectra have been interpolated by mean of a cubic spline function.
In general, the corrected Gaia spectra match better the template and SMASS spectra than the correspondent original spectra. There are some cases, such as (62) Erato, in which the corrected Gaia spectrum matches extremely well the SMASS spectrum, while they both deviate from the template spectrum. This is probably an indication that the object is misclassified by Mahlke taxonomy. In this system (62) Erato is in fact reported as a Ch- type, while in the literature it has mostly been classified as B- type, which is in good agreement with both Gaia and SMASS spectra. There are instead other cases, such as (2) Pallas, in which the corrected Gaia spectrum matches the template spectrum and not the SMASS spectrum. Finally, there are few cases, such as (10) Hygiea and (29) Amphitrite, in which both the original and corrected Gaia spectra are completely different from both the template and the SMASS spectra. This is very likely due to an intrinsic problem in Gaia reflectance spectra which cannot be solved with any correction.

4.5 Classification with *classy*

Classy[1] is a Python tool to compare asteroid reflectance spectra published in public repositories (SMASS, MITHNEOS, Gaia, etc.) and to classify spectra in the framework of the Mahlke asteroid taxonomy (Mahlke et al., 2022). *Classy* is just one of the several tools of this kind, however I chose it since it is recent and publicly available. In order to classify a spectrum, *classy* requires the wavelengths of observation and the corresponding reflectances. The uncertainties on the reflectances and the associated flags may also be provided. In addition, since Mahlke taxonomy makes use of the albedos to separate among spectral types, they should be provided to *classy* when available. However, Gaia observations are limited to the reflectance spectra and do not include albedos, therefore I did not consider them, thus preventing the identification of some spectral types. In addition,

[1] https://classy.readthedocs.io/en/latest/

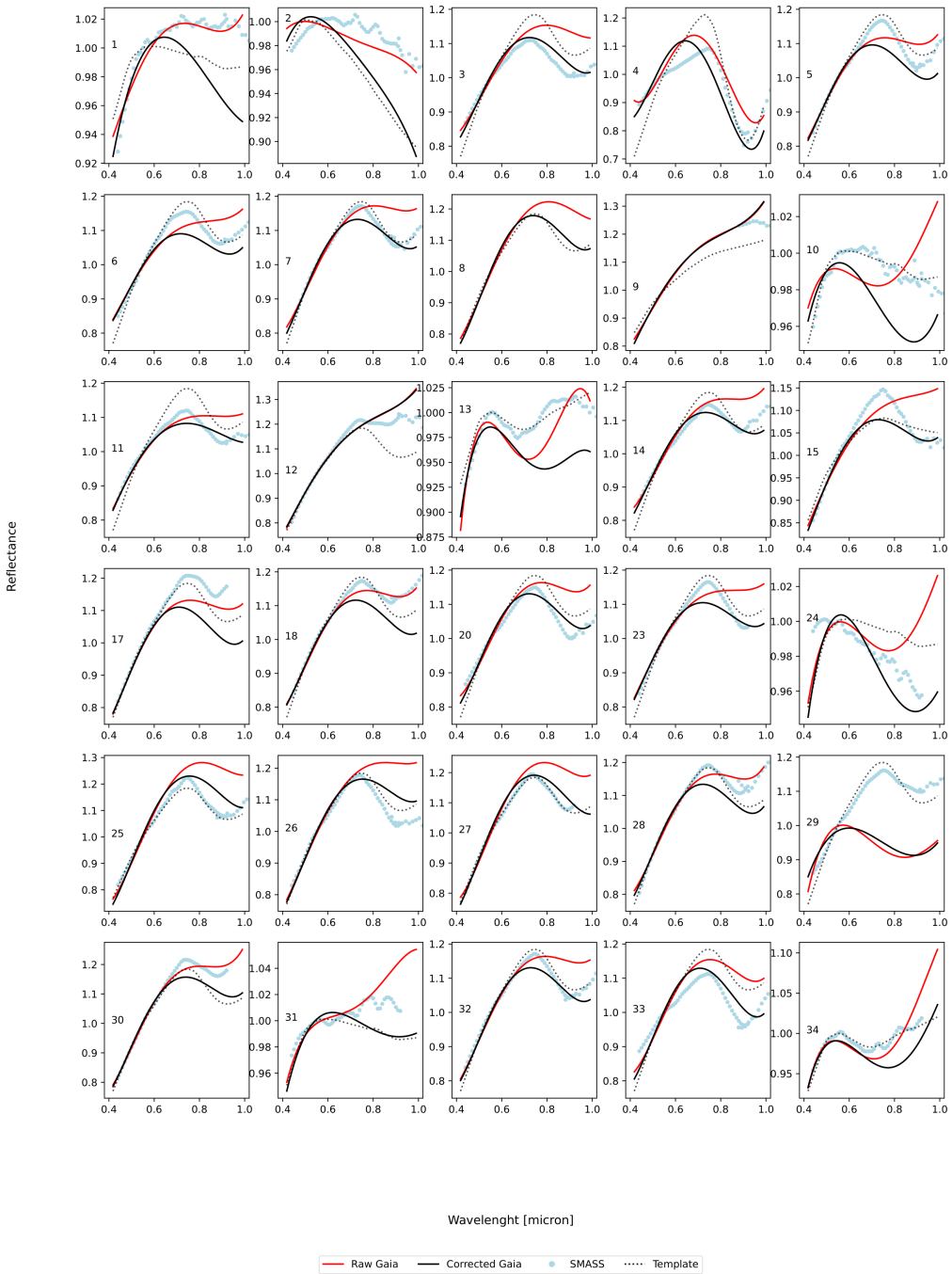

FIGURE 4.10: Corrected Gaia spectra (in black) compared to the original Gaia spectra (in red), to SMASS spectra (light blue, when present) and to Mahlke template spectra (black dotted line) for numbered asteroids between 1 and 34.

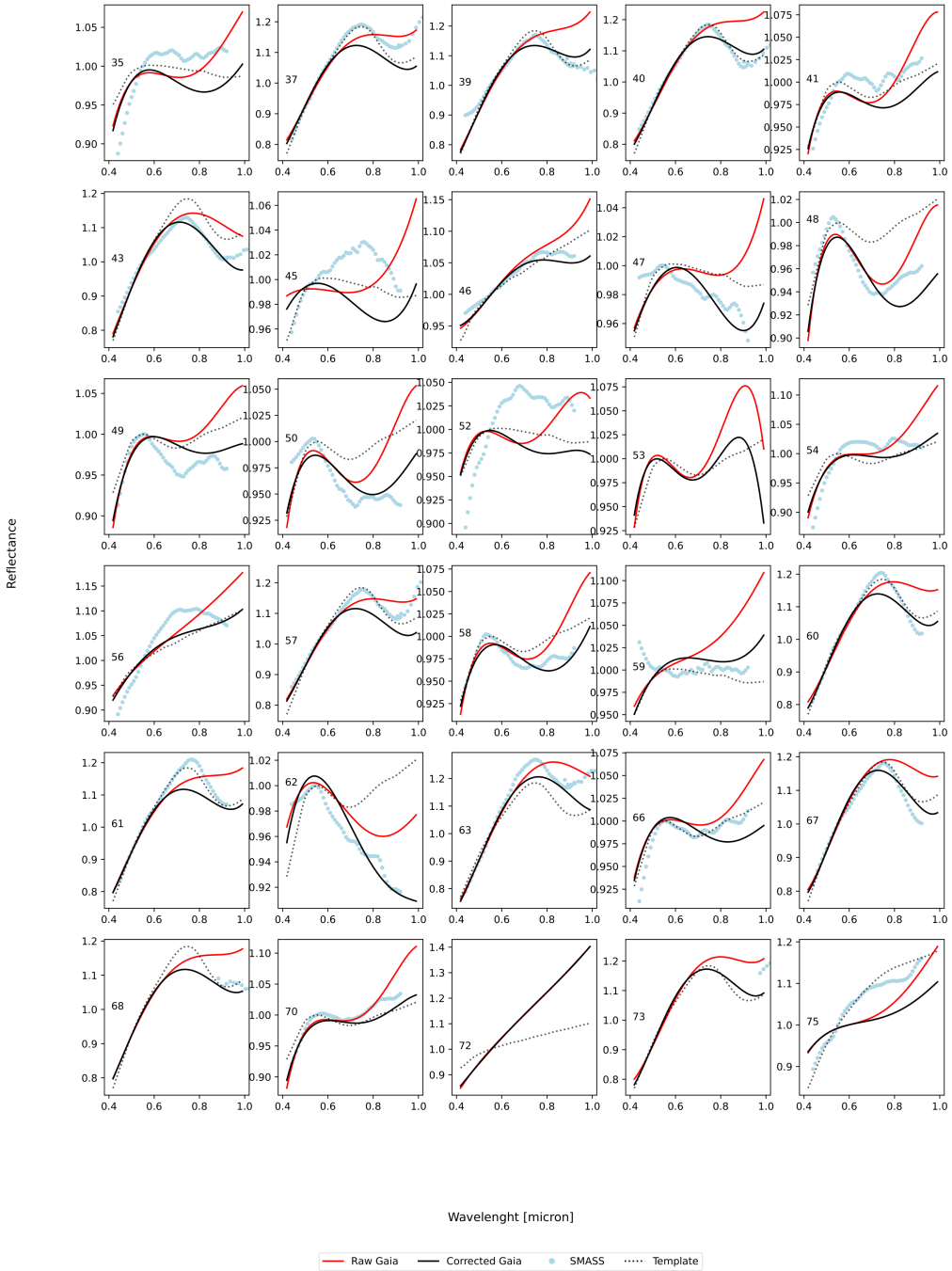

FIGURE 4.11: Corrected Gaia spectra (in black) compared to the original Gaia spectra (in red), to SMASS spectra (light blue, when present) and to Mahlke template spectra (black dotted line) for numbered asteroids between 35 and 75.

I was already expecting some degree of degeneracy because Mahlke defines spectral types using visible and infrared spectra, while Gaia data are limited to the visible range.

In order to test my correction procedure I classified with *classy* all the asteroids in common between Gaia observations and Mahlke taxonomy. Since in this taxonomy system class assignments are given probabilistically, I considered only the objects classified by Mahlke with a probability higher than 80%. In addition, the Ch- class does not have an associated class probability since it depends on the detection on the h- feature at 0.7 μm, therefore I merged it with the C- class.
I then classified Gaia spectra, corrected and non corrected, with *classy*. The results are reported in two confusion matrixes in Figure 4.12, where on the y-axis there are objects catalogued by Mahlke taxonomy and on the x-axis Gaia observations. The left panel shows the case of uncorrected Gaia spectra, the right panel the corrected spectra.
The first thing that can be noticed from both panels is that *classy* did not recognize some spectral types. Mahlke taxonomy makes in fact use of the albedos to distinguish the E- and P- classes, which have degenerate spectra with other taxonomic types. Since I did not provide the albedos, *classy* classified the majority of the E- types into the M- class (before the correction) and the K- class (after the correction). The majority of P- types are instead classified into the C- class, with which they present visible degenerate spectra. In addition, *classy* did not classify any spectra into the R- and O- classes, which however are very rare and particular types, containing very few objects.
In the left panel, despite a lot of objects have been correctly classified by *classy*, there is still a large confusion between some spectral types. Hundreds of asteroids classified as S- types in Mahlke taxonomy have been identified by *classy* of another spectral type, with a huge degeneracy with L- and X- types, while about one quarter of C- types is classified as S- types. There is then some expected degeneracy, for example the A- and V- types classified as S- types.
After applying the correction there is a general improvement in the performance of *classy*. More than two hundreds S- types have been removed from the X-, L- and M- classes and have been correctly classified into the S- class. There is also a huge improvement in the classification of V-, C- and L- types. However the correction worsens the D- class, whose objects are seen by *classy* as L- types. The Z- class, similar to the D- class, is an interesting case since *classy* is not able to recognize it in both cases.

The correction procedure has several limitations which limit the overall performance of *classy* and which can explain the remaining degeneracy among spectral types in the right

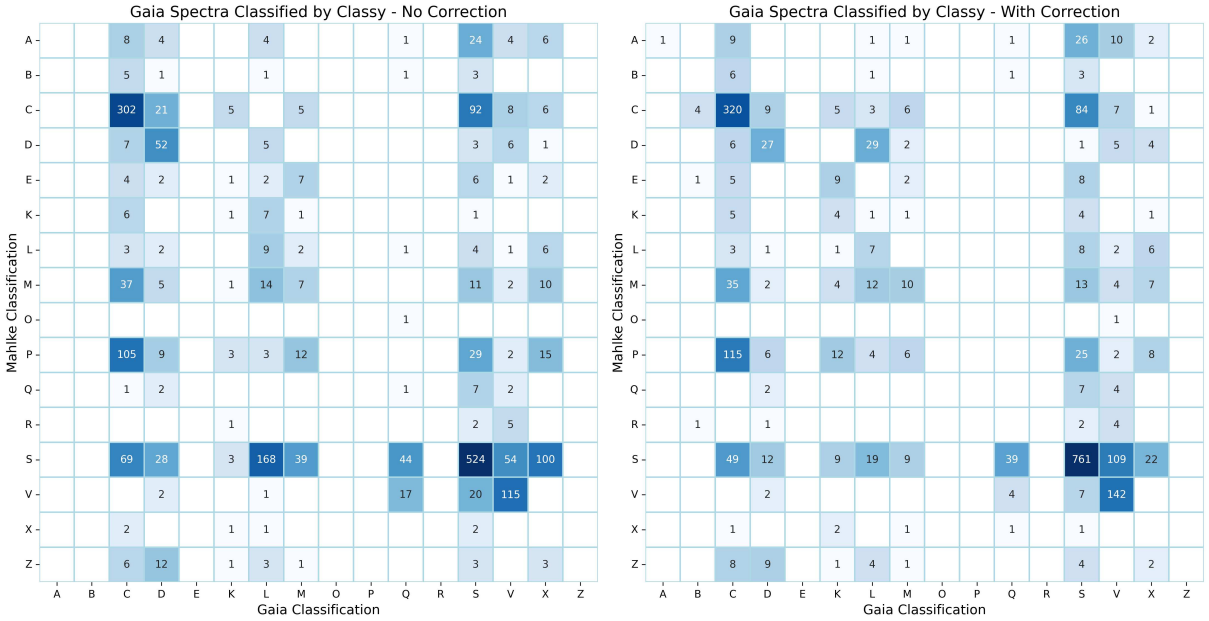

FIGURE 4.12: *Left panel*: the confusion matrix between Mahlke asteroids and uncorrected Gaia spectra classified by *classy*.
Right panel: same as the left panel, but with corrected Gaia spectra.

panel of Figure 4.12. Spectra of faint objects are in fact characterized by large uncertainties and oscillations, which can affect both the classification done by *classy* both the computation of the slope and the maximum of the reflectance, which in turn reflects on their positions on the diagram and on the multiplicative factors to apply in order to correct their spectra.
In addition, I empirically defined the boundaries of the regions of the diagram, but other better choices could be possible as well. This in particular may affect the correction of objects lying close to the boundaries. Finally, other parameters other than the slope and the maximum of the reflectance may be defined as well, allowing for a better correction procedure.
An even better solution would be to eliminate the regions and create a correction gradient inside the diagram, at each point of which would be associated specific corrections factors.

In order to test the performance of *classy* when classifying Gaia spectra of faint asteroids, I considered a sample of objects of known NEOWISE albedos selected in a region of the outer Main Belt centered on the Tirela family (Section 5). The sample contained low albedo background objects and high albedo family members, where low albedos have $\rho_V < 0.12$ and high albedos $\rho_V > 0.12$. Among this sample, I considered all objects having a spectrum in Gaia DR3 and I classified them with *classy*. Based on known taxonomy, I

was expecting a good correspondence between spectral types and albedos, with low albedo objects classified into the C- complex and high albedo family members into the L- class. However, *classy* classified most low albedo objects into the S- and L- classes and high albedos as C- or D- types. In addition, a large part of the sample was classified into the V- class, which is obviously wrong since objects of this spectral type are extremely uncommon in the outer Main Belt.
I therefore deduced that my procedure was not able to distinguish between family members and background and therefore it was not precise when dealing with faint objects.

Because of all the uncertainties affecting my correction procedure, I explored different approaches to Gaia spectra which did not require *classy*. I thus started working on a color-based taxonomy currently being developed by Alexey Sergeyev at the Observatoire de la Côte d'Azur (A. Sergeyev, Personal Communication, 2023). This taxonomy is described in the next section.

4.6 Color based taxonomy

The most used filters in astronomy are the SDSS filters introduced with the Sloan Digital Sky Survey (York et al., 2000). The five SDSS filters are ultraviolet (u, central wavelength 3543 Å), green (g, 4770 Å), red (r, 6231 Å), near infrared (i, 7625 Å) and infrared (z, 9134 Å) (Fukugita et al., 1996). The SDSS filters response curves without the atmospheric transmission are reported in Figure 4.13, where the u filter is in purple, g in blue, r in green, i in yellow and z in red. Gaia spectra cover an interval from 3740 Å in the u filter to 10340 Å in the z filter, spanning over the g, r and i filters.
In the ideal case, a filter works by blocking out light at all wavelengths except in a range of wavelengths where its transmission is 100%. In reality the filter response is not ideal and its sensitivity falls off slowly at shorter and longer wavelengths. In order to compute the brightness of a source as seen through a filter, the amount of light that it emits in the filter band must be considered. Colors are then obtained as the difference in brightness between two filters expressed as magnitudes. Example of SDSS colors are $g-r$, $r-i$ and $i-z$.
Colors have been used since the first taxonomies to distinguish between asteroids of different surface properties, in particular between the broad C- and S- complexes (Section 2.1.5). Color-color diagrams have been built for thousands of asteroids observed in the SDSS filters (Ivezić et al. (2001), Parker et al. (2008)), which confirmed the correlation

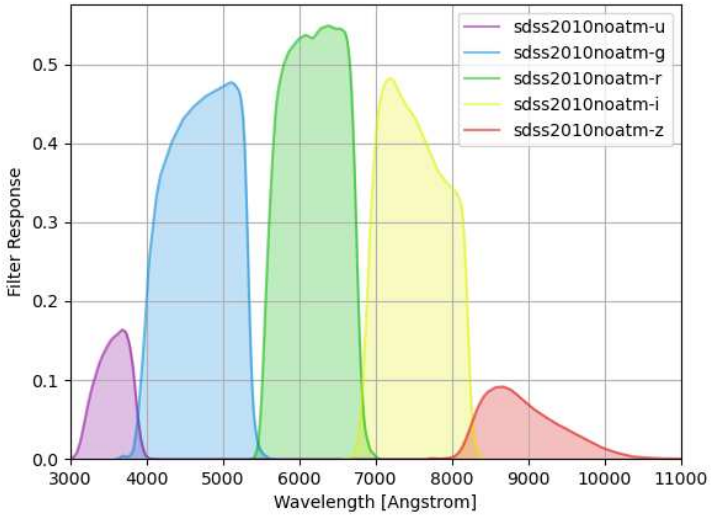

FIGURE 4.13: The SDSS filters response curves without the atmospheric transmission. The filters are, from left to right, u, g, r, i and z. Gaia reflectance spectra cover the interval between 3740 Å (in the u filter) and 10340 Å (in the z filter).

between spectral type and position on the diagram (Nesvorný et al. (2005), Ivezić et al. (2001)).
In order to explore asteroid families with Gaia data, I needed the spectral types. I therefore tested the taxonomy created by Alexey Sergeyev, who converted Gaia reflectance spectra in SDSS colors and built a color-color diagram giving an indication of the spectral type. (A. Sergeyev, Personal Communication, 2023).

Each Gaia spectrum has been integrated over the SDSS filters bandpasses in order to compute the associated magnitudes. Colors have then been computed as the difference between magnitudes. I then built a color-color diagram for all the over 60 thousand asteroid spectra obtained by Gaia, the black dots reported in Figure 4.14, where on the y-axis there is the color $i - z$ and on the x-axis the α^* parameter defined as $\alpha^* = 0.89(g - r) + 0.45(r - i) - 0.57$ by Ivezić et al. (2001). The yellow dashed lines, taken from Parker et al. (2008), divide the diagram into the three broad complexes C, S and V, indicated with the corresponding letter. The shape of the color-color diagram built from Gaia observations is overall very similar to the color-color diagram built from SDSS observations reported in Parker et al. (2008).
The diagram has then been divided into subregions, to each one of which a Bus-DeMeo spectral type was assigned, similarly to what was done by Sergeyev and Carry (2021) for the asteroids observed by the Sloan Digital Sky Survey and by Sergeyev et al. (2022)

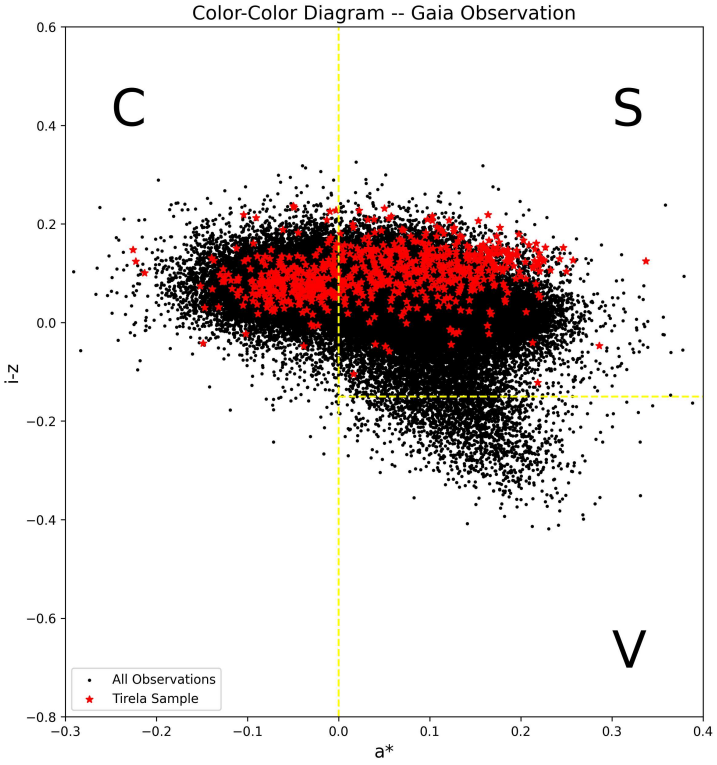

FIGURE 4.14: Color distribution for all the Gaia observations (black dots) and the asteroids in sample of the outer Main Belt centered on the Tirela Family (red stars). The yellow dashed lines separate the three major taxonomic groups.

for the objects detected from the SkyMapper Southern Survey. Each object observed by Gaia is therefore assigned to the taxonomy class corresponding region. For example, in Figure 4.14, V- types are on the bottom of the diagram and A- types on the right of the S- complex.

Gaia reflectance spectra are characterized by uncertainties, which must be considered in the computation of colors and that affect the positions of the objects on the diagram. This has two main implications. First, an asteroid close to a boundary could fall in more than one region thus receiving a double classification. Second, faint asteroids are usually characterized by large uncertainties that greatly affect their positions. These objects usually fall in more than one class even if they do not lie close to any boundary.
In order to take into account the uncertainties affecting Gaia spectra, the color-based taxonomy reports the spectral type of the considered object along with its probability to belong to that class. The second most probable spectral type and the associated probability are also included in the taxonomy. Doubtful cases, for example faint objects characterized by large uncertainties, are assigned to the U- class (unknown). Also for

these objects the second more probable class and the associated probability is reported. All Gaia observations are therefore assigned to a taxonomic class.

In a similar fashion to what I did with *classy*, I tested the color based taxonomy on the same subsample of asteroids of the outer Main Belt centered on the Tirela Family. In Section 4.5, I proved that *classy* was not able to correctly classify these objects, since most low albedos were assigned to the S- complex and high albedos to the C- complex, with a lot of unexpected V- types.
I therefore analyzed the spectral types resulting from the color based taxonomy of these objects, which are shown as red stars in Figure 4.14. The first aspect that can be noticed is that in this case there are no V- types, which already represents an improvement compared to *classy*. In addition, the sample was found to be split between the S- and the C- complexes, respectively the two red clusters on the right and the left of Figure 4.14. I then compared the spectral types with NEOWISE albedos and I found that almost all low albedo objects were classified as C- or D- types, while the high albedos felt in the S- and L- classes. The U- class contained objects of all albedos, therefore in these cases the second more probable class was selected. Also in this case I found a very good correspondence between spectral types and albedos. The few misclassifications could be imputed to flawed Gaia spectra.
This analysis thus proved the higher efficiency of the color based taxonomy over *classy*. I therefore decided to use this classification system to try to distinguish the Tirela (Chapter 5) and the Watsonia (Chapter 6) families using Gaia spectra alone.

Chapter 5

Analysis of the Tirela family

5.1 Introduction

Tirela is an asteroid family first identified by Nesvorný et al. (2005) and successively reidentified in the space of proper elements by different authors, such as Masiero et al. (2013), who divided the Main Belt in different albedo groups using albedos from WISE/-NEOWISE and then applied the HCM to these subpopulations. Milani et al. (2014) instead used a multi-steps HCM to identify dinamically associated groups of asteroids in the Main Belt (Section 2.2.2), from which they found different memberships to the family compared to previous works and renamed it the Klumpkea family.

The Tirela/Klumpkea family is located at the edge of the outer Main Belt, centered at the semi-major axis $a_p = 3.12\ AU$, and it possesses high eccentricity and inclination, $e_p = 0.20$ and $i_p = 16.8°$. The family is characterized by an high geometric albedo (0.2 - 0.3), whereas the background asteroids in the same region generally have low albedos and belong to the C- complex.
A determination of the age of the Tirela family is reported in Milani et al. (2016) who used the method based on the v-shape described in Spoto et al. (2015) (Section 2.2.3). Assuming a mean albedo of 0.204 ± 0.100 and a density $\rho\ (1km) = 2.275$, the authors obtained an age of $664 \pm 212\ Myr$ for the inner side and a very similar age of $661 \pm 244\ Myr$ for the outer side.

The family has been spectroscopically characterized for the first time by Mothé-Diniz and Nesvorný (2008), who obtained visible spectra for ten faint Tirela members. They found

that the family consisted of Ld- type asteroids (in the Bus-DeMeo classification scheme), while comparisons with minerals suggested similarities between the absorption observed in the spectrum of Tirela and the ones found in the mineral pigeonite.
High resolution NIR spectra for five family members are reported in Devogèle et al. (2018), who studied the family looking for possible correlations with the population of Barbarians. Tirela is in fact a particularly interesting family because it is an important repository of Barbarian asteroids.
Barbarians, named after (234) Barbara, are a rare class of asteroids showing anomalous polarimetric properties first discovered by Cellino et al. (2006). They present negative polarization at relatively large phase angles, while normal asteroids present positive polarization at the same phase angles. Despite being still unclear, the origin of this phenomenon is believed to be linked to large abundances in the asteroids surfaces of spinel, a mineral usually associated with the calcium-aluminium-rich inclusions (CAIs) on meteorites (Sunshine et al., 2008). Since CAIs are the oldest solid material in the Solar System (Connelly et al., 2012), Barbarian asteroids are a very interesting target for investigation.

Section 5.2 describes the procedure I implemented to distinguish between the Tirela family members and the background asteroids using Gaia spectra alone. In Section 5.3 the method I used to retrieve the v-shape is reported. Finally, an estimation of the age of the family is given in Section 5.4.

5.2 Identification of Tirela family members

In order to separate the family members from the background using Gaia data, I first considered all the asteroids contained inside a region in the 3D space of proper elements centered on the cluster of identified Tirela members with $3.00\ AU < a_p < 3.23\ AU$, $0.15 < e_p < 0.25$ and $0.27 < \sin(i_p) < 0.33$. This wide region contains numerous background asteroids and other families (or at least part of them) in addition to Tirela, such as the small Aegle family.
I retrieved the proper elements a_p, e and $\sin(i)$ and the absolute magnitudes H of the asteroids located within this region from the Asteroid Family Portal (AFP) (Section 2.2.2). I initially used the albedos retrieved from *rocks* (Mahlke et al., 2022), a Python client to explore and retrieve asteroid dynamical and physical parameters from The Virtual Observatory Solar system Open Database Network (SsODNet). However, these albedos are obtained with different methods which are sometimes very approximate and therefore

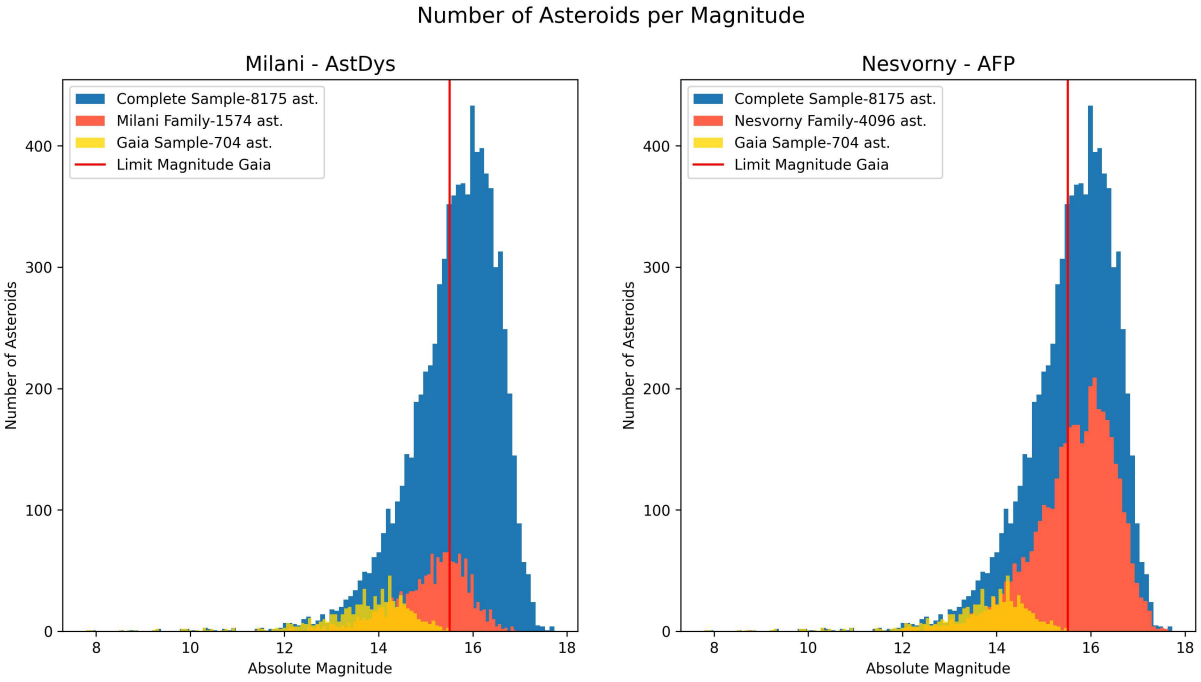

FIGURE 5.1: The number of asteroids per magnitude for the complete sample (blue), the Gaia sample (yellow), the Milani classification (red, left panel) and the Nesvorny classification (red, right panel).

this generate confusion when comparing them. I thus preferred to use WISE/NEOWISE albedos only. WISE is an all-sky thermal infrared survey that determined the albedos of more than 128 000 objects from their IR thermal flux (Masiero et al., 2011). The NEOWISE survey contains slightly less albedos than *rocks*, but all determined with the same method. For objects that have been observed more than once by the survey, I took as their albedos the weighted average of the measurements.

I found 8175 objects in total located within the selected region, 704 of which have a spectrum in the Gaia repository. Of the complete sample, only 41 objects had a spectra published in the literature before Gaia DR3, while 1887 objects had colors determined from photometric observations.

The huge difference between the total and the Gaia samples is due to an observational bias since Gaia is able to observe only the brightest objects. Figure 5.1 shows the distribution in absolute magnitude for the asteroids of the complete sample (in blue) and the ones observed by Gaia (in yellow). Gaia observations present a magnitude limit $H = 15.51$, marked by the vertical red line, while the faintest object in the complete sample has $H = 17.73$. The peak in the distribution of the magnitude of the objects of this region lies below Gaia limit. This affects the accuracy with which the age of the family can be

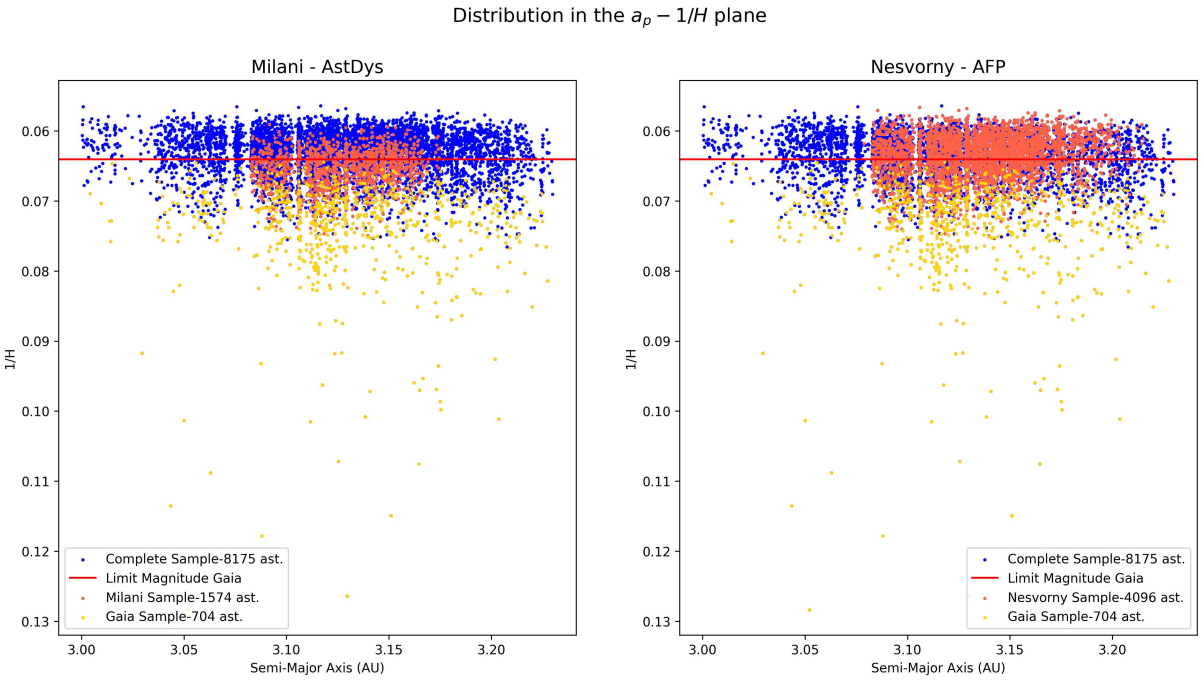

FIGURE 5.2: Distribution in the $a_p - 1/H$ space for the objects of the complete sample (in blue), the Gaia sample (in yellow), the Milani classification (in red, left panel) and the Nesvornny classification (in red, right panel).

determined (Section 5.4), since only the largest objects (corresponding to the bright tail of the distribution) are observed by Gaia.
As discussed in Section 2.2.2, two main classification systems for asteroid families exist: AFP (from now on Nesvorny) and AstDys (from now on Milani). The tendency of the Milani classification to include less objects than Nesvorny is here evident: the first lists 1574 objects as part of the family, while the latter 4096. The two systems are similar at bright magnitudes, but Nesvorny includes more objects at faint magnitudes. In both cases, the magnitude peak is fainter than the Gaia magnitude limit.

The observational bias in Gaia data is also evident in the distribution in the $a_p - 1/H$ plane reported in Figure 5.2, where the complete sample is showed in blue, the Gaia sample in yellow, the Milani classification in red in the left panel and the Nesvorny classification in red in the right panel. Red horizontal lines mark the Gaia magnitude limit.
Figure 5.2 highlights an additional difference between the Milani and the Nesvorny classification systems. The boundary in the inner semi-major axis for the family is the same for both classifications and it is defined by the resonance at 3.08 AU. Instead, in the outer boundary AstDys cuts the family at the resonance at 3.17 AU, while AFP includes objects reaching $a_p = 3.20$ AU. However, most of the family members listed by AFP in

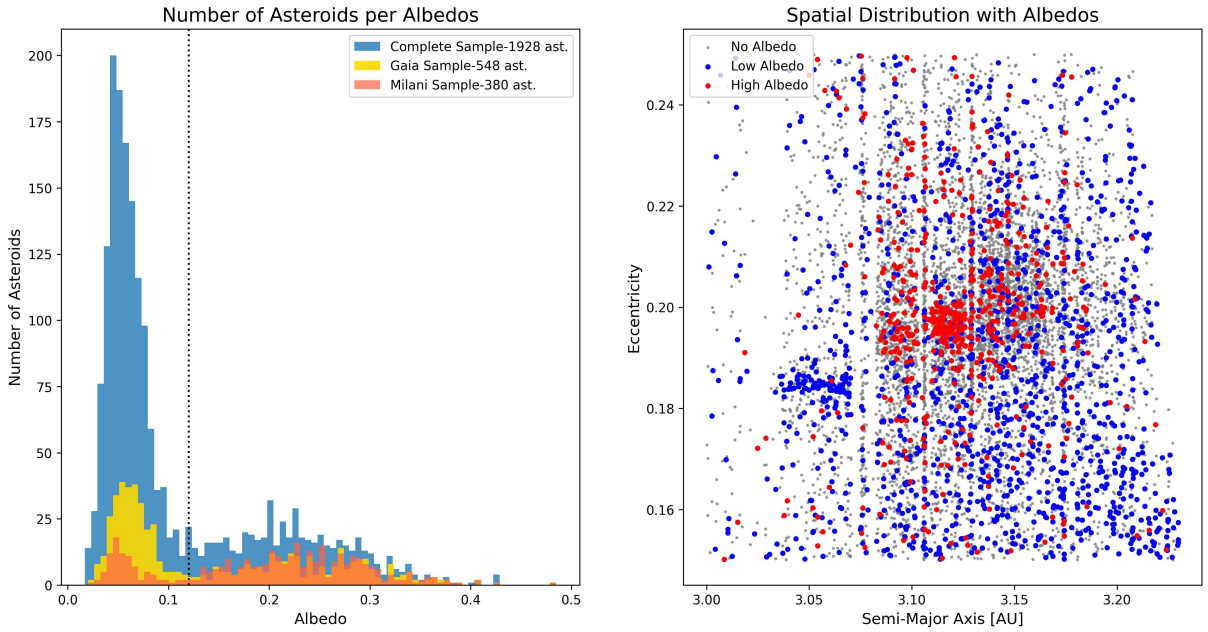

FIGURE 5.3: *Left panel*: Albedo distribution for the objects of the complete sample (in blue), of the Gaia sample (in yellow) and of the Milani classification (in red). *Right panel*: Spatial distribution in the $a - e$ plane for the objects of the complete sample divided according to the albedo.

this outer region possess low albedos and it is thus quite unlikely that they really belong to the family. The largest members of the Tirela family have in fact been observed to be characterized by high albedos and therefore, due to the homogeneity of the family, also smaller objects are expected to be like that. The family members in AstDys instead generally have high albedos, therefore in the following analysis only this classification system will be retained for comparison.

1928 objects out of the 8175 of the total sample and 548 out of the 704 of the Gaia observations possess a NEOWISE albedo. The distribution of the albedos is reported in the left panel of Figure 5.3, where the complete sample is showed in blue, the Gaia sample in yellow and the Milani classification in red. A bimodal distribution in the three samples is clearly visible, with a net separation around $\rho_V = 0.12$. The narrow peak at albedos lower than this limit corresponds to dark background asteroids, while the broad peak at higher albedos consists of Tirela family members plus few background asteroids belonging to the S- complex. It can also be noticed that the Milani classification includes few low albedos objects, which probably are interlopers included in the family due to the intrinsic limitations of the multi-step HCM used to identify the family.
The correlation between high albedos and family membership is confirmed by looking at

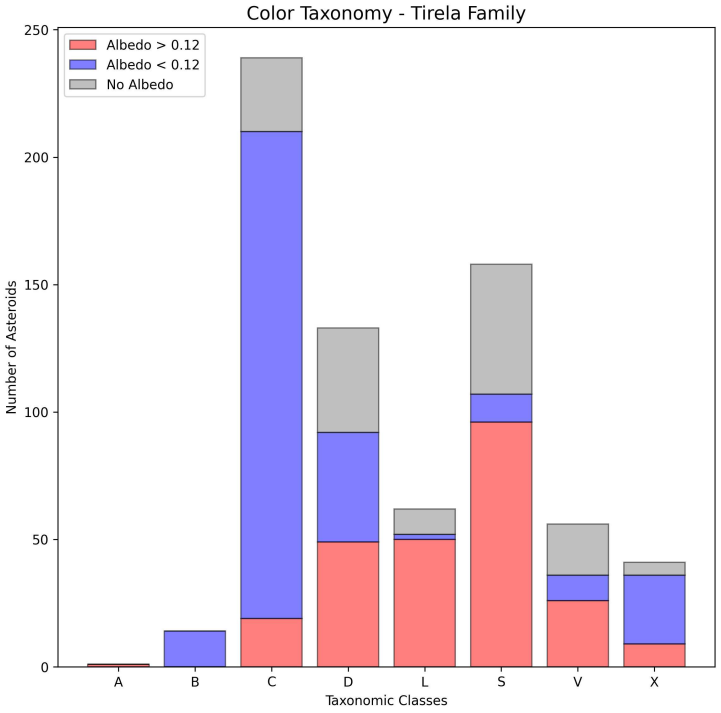

FIGURE 5.4: Number of objects per spectral type for Gaia observations classified by the color-based taxonomy.

the distribution in the proper elements space of the objects lying in this region. The right panel of Figure 5.3 shows the distribution of the asteroids of the complete sample in $a - e$ plane. Red points present albedos higher than the limit $\rho_V = 0.12$ which correspond to the broad peak in the left panel, while blue points are characterized by low albedos and correspond to the narrow peak. Small grey points are asteroids without albedo measurements. High albedo objects are concentrated in the family core, centered between 3.09 AU and 3.15 AU. Low albedo objects instead are scattered all around the space without a preferred distribution, except for a small cluster at 3.05 AU which is the Aegle family. This analysis showed that Tirela family members can be easily distinguished from the background by using the albedos. I wanted instead to address the question if it is possible to differentiate them by using Gaia data alone, without the help of any external information.

I wanted to classify Gaia observations in this region in order to separate L- type asteroids (or objects of similar classes) from the background of C- types. However, objects in this region are usually faint and the corresponding Gaia spectra are very noisy with large error bars. As discussed in Section 4.5, the Python package *classy* turned out to be not accurate enough in classifying faint spectra, therefore I decided to use the color taxonomy

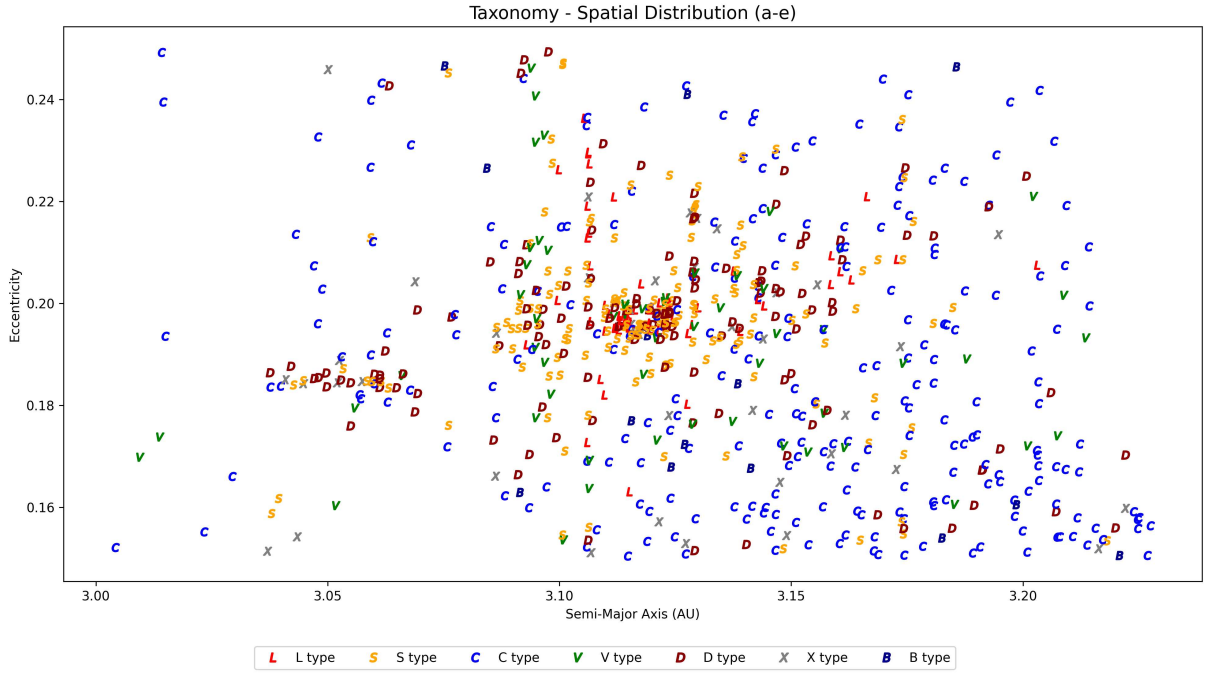

FIGURE 5.5: Spatial distribution in the $a - e$ space for Gaia observations in the Tirela region with relative taxonomic types.

at its place.

I computed the color indexes for all Gaia spectra in this region and I determined their taxonomic types from their positions on the color-color diagram, following the procedure described in Section 4.6. A large amount of objects turned out to belong to the U- class, therefore, in order to not discard them, I selected as their spectral type the second more probable class that resulted from the analysis.

The resulting distribution of the spectral types of the objects in this region is reported in Figure 5.4. To confirm the consistency of the color-based taxonomy system, I compared the spectral types with the albedos (when available). Almost all B- and C- types were found to possess albedos lower than $\rho_V = 0.12$, while almost all S- and L- types were found to be characterized by high albedos, as expected. In addition, I obtained a non negligible quantity of V- types, which should not be present in the outer Main Belt. Analyzing in more detail these objects I discovered that they all are very faint and with anomalous Gaia spectra characterized by wide oscillations within them. The color taxonomy interprets such oscillations as the deep absorption band typical of V- types and therefore classifies them as such. I therefore ignored these V- types.

The validity of the color taxonomy is confirmed by looking at the spatial distribution of the taxonomic types in the proper elements space, as shown for the $a - e$ plane in Figure

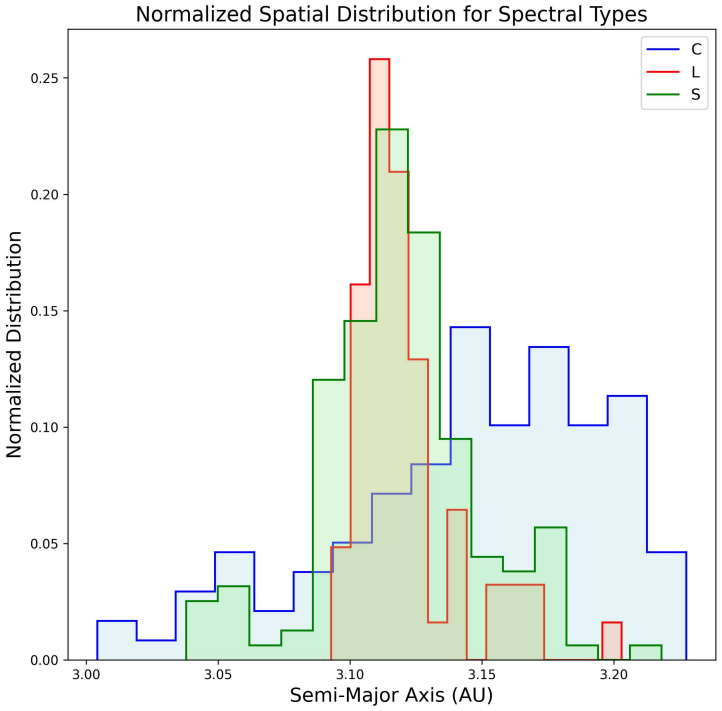

FIGURE 5.6: Normalized distribution of the asteroids in the Tirela region. L- types are shown in red, S- types in green and C- types in blue.

5.5. Except for few background S- types, all L- and S- types are concentrated in the center of the family, while C- types are more uniformly distributed over the whole domain. D-types concentrate in the Aegle family, while they are more diffuse for $a > 3.08\ AU$. This can be quantitatively proven by studying the normalized distribution in semi-major axis of each spectral type, from which I found that L- and S- types are the most concentrated in the center of the family at $a = 3.11\ AU$, while C- types have a more diffused distribution characterized by a broad peak centered at $a = 3.15\ AU$. The normalized distribution of the spectral types is reported in Figure 5.6, where only S-, L- and C- types are shown for better clarity.

As discussed in Section 5.1, the Tirela family in the literature is found to consist of L-type asteroids, while in my analysis I obtained that S- types are predominant in number in the center of the family compared to L- types. This result should not be seen as in contrast with previous works, since it is very likely related to the fact that Gaia spectra do not extend beyond the visible range. Due to the systematic deviations in the red end, L- and S- types present almost degenerate spectra in the visible and therefore the two spectral types can be easily confused one with the other. An example was already discussed in Figure 4.1 for the S- type (39) Laetitia and the L- type (729) Watsonia.

However, the relevant aspect that is important to stress here is that the color taxonomy is able to differentiate with high accuracy between family members (S- and L- types) and background (C- types).
I therefore considered as family members only the objects classified as L- and S- types. I discarded the B-, X- and C- types because they clearly are background, the V- types because they are very faint objects poorly classified, the A- types since they were too few to be statistically relevant and the D- types because their spatial distribution and their albedos were unclear.
By looking at the distribution of the remaining L- and S- types in Figure 5.5, I found that they concentrated between 3.08 AU and 3.18 AU, therefore I adopted these values as boundaries in the semi-major axis. S- types below the inner limit belong to the Aegle family, while the few S- and L- types outside the outer limit belong to the background.
Finally, the family is cut by few thin resonances. Resonance objects may in principle belong to the family, but their eccentricities have been pumped up (or down) by the resonances. In addition, their evolution in semi-major axis due to the Yarkovsky drift have been stopped because they have been trapped in those resonances. I therefore discarded all the resonant objects, no matter what their spectral types were. I identified three resonances: the first at 3.104 $AU < a <$ 3.108 AU, the second at 3.127 $AU < a <$ 3.130 AU and finally the third at 3.17 $AU < a <$ 3.18 AU. The remaining objects are the ones that I analyzed to constrain the age of the family, as explained in the following sections.

5.3 V-Shape of the Tirela Family

To constrain the age of the family, I developed a simplified version of the v-shape method implemented by Spoto et al. (2015) and summarized in Section 2.2.3.
In order to apply this method, the v-shape of the family in the $a-1/D$ plane must first be identified, which in turn requires the knowledge of the diameters of the family members. For the objects of known albedo, the diameter D is determined from Equation 2.2. For the objects of unknown ρ_V, I used an average albedo computed from the other family members. I found the average albedo of the family to be $\rho_V = 0.252 \pm 0.069$, compatible with the value reported in the literature ($\rho_V = 0.204 \pm 0.100$, Milani et al. (2016)).

The family members retrieved from Gaia observations are shown in blue in the $a - 1/D$ plane in the left panel of Figure 5.7. For better clarity, I also computed the diameters of

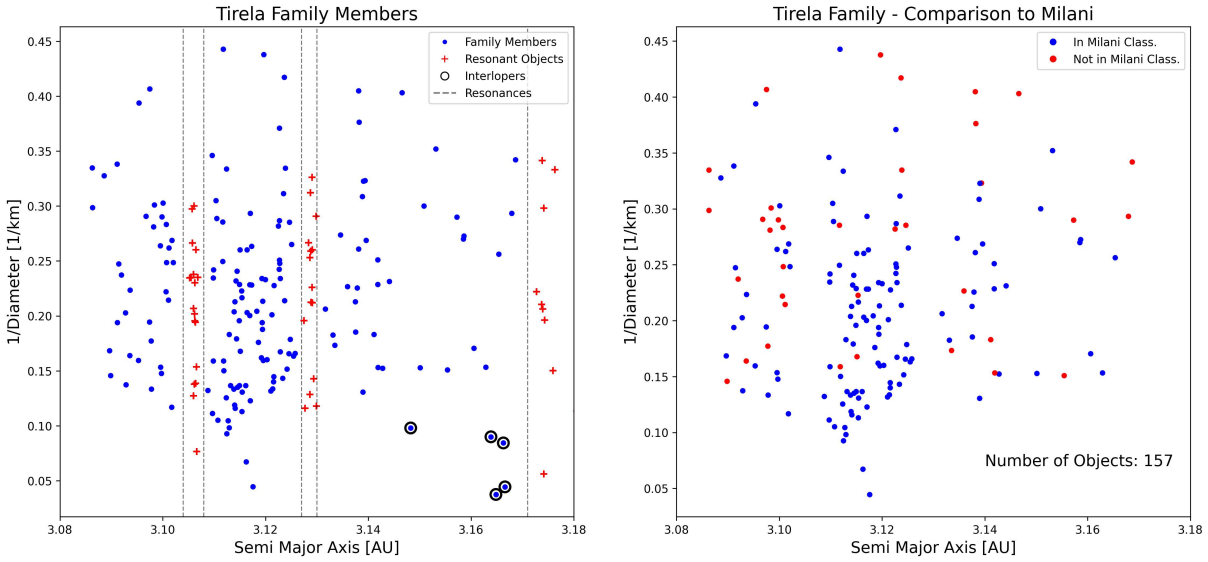

FIGURE 5.7: *Left panel*: Distribution in the $a - 1/D$ plane for family members determined from Gaia observations (blue points), resonant objects (red crosses) and interlopers (black circles).
Right panel: Distribution in the $a - 1/D$ plane for family members included in the Milani classification (in blue) and not included (in red).

the resonant objects, which are shown as red crosses. The positions of the resonances are marked by vertical black dashed lines.

The objects circled in black are evident interlopers that have been identified by visual inspection. They are in fact as big as Tirela and Klumpkea, the largest members of the family, but they are distant in semi-major axis. The Yarkovsky drift is not able to explain such a distance from the vertex of the v-shape and the only reasonable explanation is that they are background S- or L- types. This is also confirmed by the fact that these five objects are not included in the Milani classification. I therefore removed them from the family, which is left with 157 members determined from Gaia observations.

The family members are shown in the right panel of Figure 5.7, from which resonant objects and interlopers have been removed. In this plot, objects listed in the Milani classification are shown in blue, while asteroids identified by my analysis but not included in Milani are shown in red. There is a general good agreement between the family members retrieved from Gaia observations and the Milani classification, especially for large objects. A well defined v-shape can be recognized between 3.08 AU and 3.13 AU, while it becomes blurry for larger semi-major axes. This is due to the limited number of objects that Gaia observed in this region. In addition, there is a small group of seven objects aligned almost horizontally at $1/D = 0.15$ which deviate from the outer side of the v-shape. However,

five of these seven objects are listed as family members by Milani, therefore I decided not to exclude them, but to keep them for a more detailed analysis.
The objects reported in the right panel of Figure 5.7 are the family members that are analyzed in the following section to determine the slopes of the v-shape, from which the age of the family is constrained.

5.4 Age Determination of the Tirela Family

In order to fit the v-shapes and constrain the ages, Milani et al. (2014) and Spoto et al. (2015) developed an automatic procedure working for every type of family, cratering or fragmentation, old or young, rich or with few members or even one-sided families.
In my case I was dealing with only one family, which allowed me to develop a simplification of the method implemented in Milani et al. (2014) adapted to the Tirela family. This procedure is described in the following part of this Section. A comparison between my results and the ones reported in Milani et al. (2016) is also discussed.

Following Spoto et al. (2015), I first manually selected the region within which fit the v-shape. The boundaries in semi-major axis were taken to be between 3.08 AU and 3.18 AU (Section 5.2), same as in Milani et al. (2016).
The limit in $1/D$ could be different for the inner and outer sides and must be taken at the values at which the sides of the v-shape become vertical. Since in my case few objects belong to the family, the choice of the limit in $1/D$ greatly affects the performance of the fitting algorithm and thus the ages. On one side, choosing a low $1/D_{lim}$ would mitigate the uncertainties due to the small objects, however limiting the analysis to very few large objects making it not statistically relevant. On the other side, placing $1/D_{lim}$ at larger values would include more objects, but limiting the accuracy of the fit. I therefore took a compromise between these two aspects by selecting $1/D_{lim} = 0.22$ for both sides, keeping however in mind that a change in this parameter affects the results.
I then divided the y-axis in bins, such that each bin contained the same number of objects N. In order to create the bins, I started from the largest object and I moved up along the $1/D$ axis counting the number of points. Every time that the count reached N objects, I traced a bin. The procedure continued until reaching the limiting value $1/D_{lim}$. The last bin before this value was discarded if it contained less than N objects.
Spoto et al. (2015) implemented an automatic procedure that, given the number of asteroids inside the family, provided the best value of N for the subdivision in bins. In my work

I instead preferred to explore different values of N. I therefore repeated the subdivision in bins for N ranging between 8 and 22. The larger N is, the less the bins are and vice versa. The binning procedure is illustrated in Figure 5.8, where the bins boundaries are traced as red dashed lines and the blue points are the family members.
Inside each bin I then selected the two points characterized by the smallest and largest semi-major axes. These are the points that are going to be fitted to determine the slopes of the v-shape. They are marked as black circles in Figure 5.8. Apart from when N is very small, the two largest bodies (Tirela and Klumpkea) are never selected in the computation of the slopes. Being located at the vertex of the v-shape, they could be added to the points to fit. However, this would give more importance to the large objects respect to the small ones. The approach to follow therefore depends on the interpretation of the collisional event. In the case of cratering families, one considers that the impact energy was not large enough to displace the largest remnant in semi-major axis, relative to the original orbit. Since Tirela is a fragmentation family, neither Klumpkea or Tirela are added to the fit when they are not selected by the procedure.
The previously selected points are then fitted through a linear regression algorithm, which finds the best straight lines approximating the two sides of the v-shape, indicated in black in Figure 5.8.
Spoto et al. (2015) implemented the automatic rejection scheme of outliers described in Carpino et al. (2003). However, in my case having few objects to analyze I opted for an simplified version of this scheme. For each side, I computed the residuals res of the linear regression, which are given by the difference between the observed positions Y_{obs} of the points along the y-axis and their positions Y_{pred} predicted by the linear regression. I then computed the standard deviation STD_{res} of the residuals and I decided to mark a point as outlier if $res_i > 2\ STD_{res}$ in that point. When an outlier is found, the point is removed and the linear regression is repeated over the new set of points.
I decided to adopt a critical limit of 2 STD_{res} instead of the usual 3 STD_{res} because the STD_{res} was already very large since the fit was conducted on a small set of diffused data points. With this limit I found in fact only one outlier, marked as a black point in the upper left panel of Figure 5.9, while no outliers are present in Figure 5.8. The slopes of the black straight lines reported in the two Figures are analyzed to constrain the age of the family. On top of each panel it is reported the slope of the inner side S_{in}, the slope of the outer side S_{out} and their ratio R.

After having implemented the procedure, I had to decide which value of N between 8 and 22 was the best for fitting the v-shape. In addition, I also focused on which set of

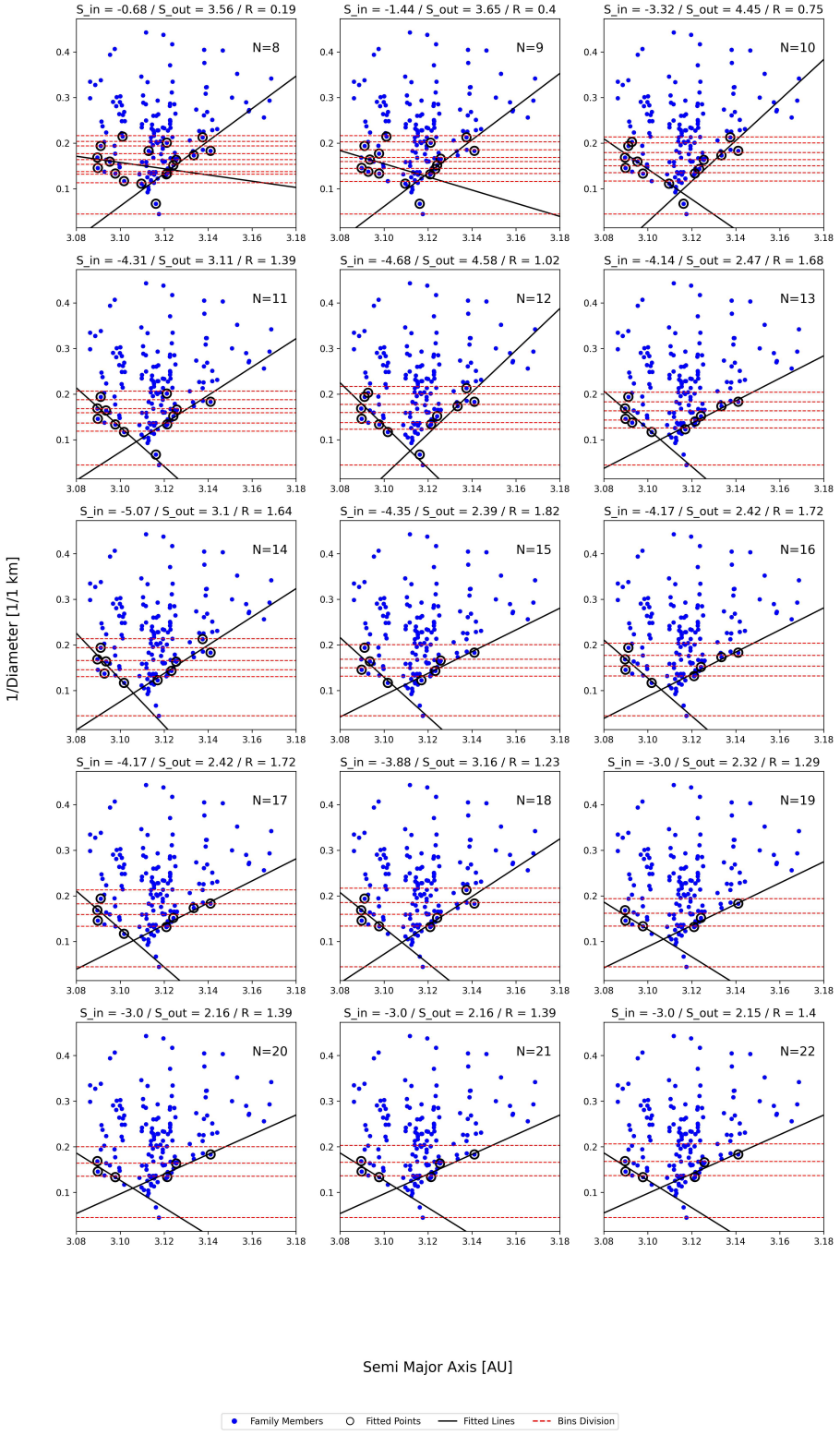

FIGURE 5.8: Illustration of the procedure to fit the v-shape of the Tirela family for N varying between 8 and 22. The blue points are the family members, the red dashed lines indicate the division in bins, the black circles are the points used to fit the sides and the black lines are the fitted lines.

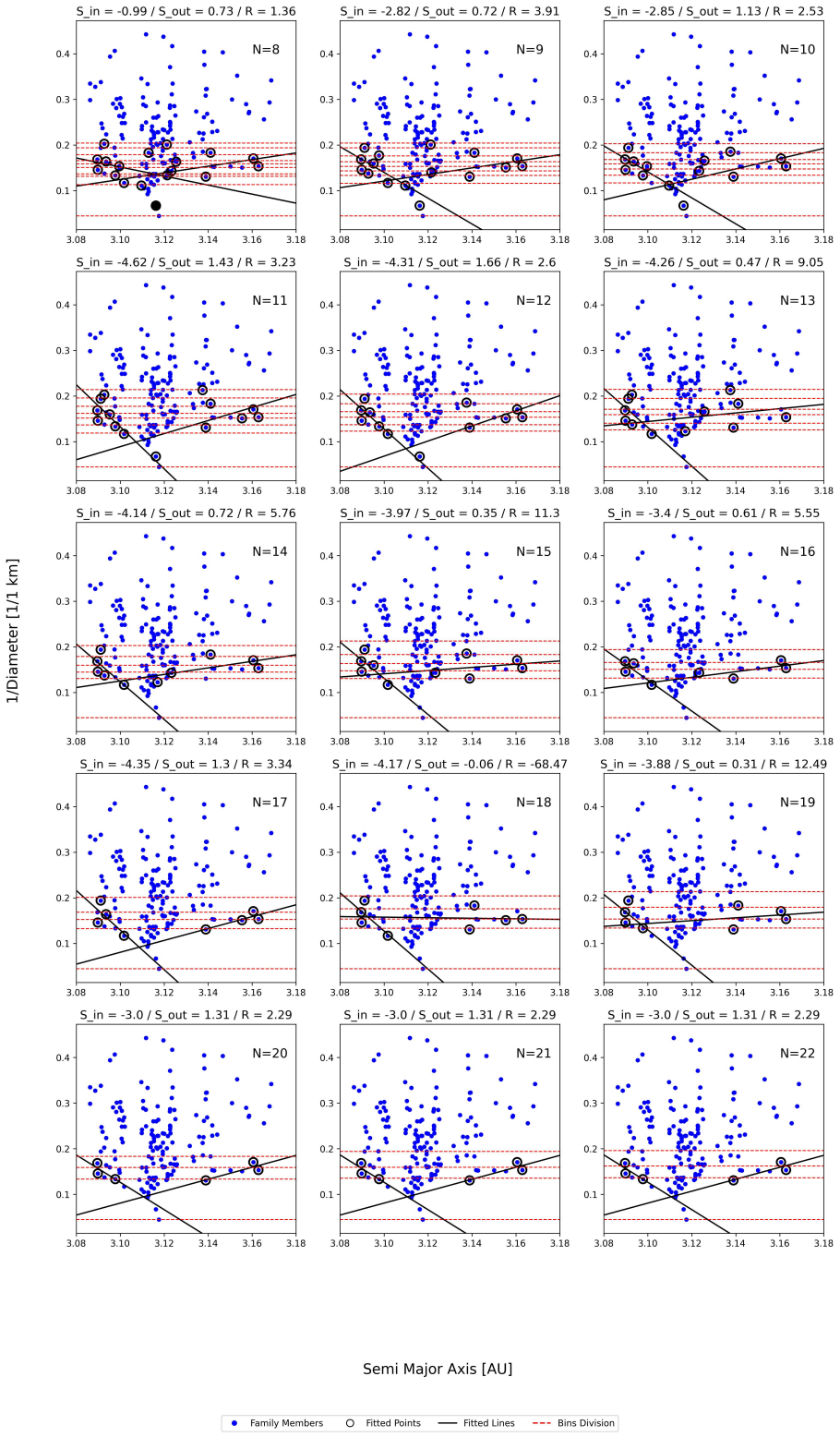

FIGURE 5.9: Same as in Figure 5.8, but with the small cluster of seven objects in the outer side included in the fit.

asteroids use in the analysis. As discussed at the end of Section 5.3, in Figure 5.7 there is a small cluster of seven asteroids deviating from the outer side of the v-shape which seem to be interlopers. However these objects are listed as part of the family by Milani. I therefore decided to compute the slopes of the v-shape for two cases: with this small cluster of asteroids removed (Figure 5.8) and with the cluster included (Figure 5.9).
In both cases it can be noticed that for $N < 11$ the fit is of poor quality since the bins are too narrow. On the other side, for $N > 18$ the bins are very large and only three points are used for the fit. The best value of N seems therefore to be between 11 and 18.
In Figure 5.9 the slope of the inner side for $11 < N < 18$ remains almost constant, with only small oscillations between $S_{in} = -4.62$ and $S_{in} = -3.40$ that can be explained by the different points used for the fit in each case. The outer side instead is quite dramatic: its slope varies in fact between $S_{out} = 1.66$ and $S_{out} = -0.06$, becoming negative in one case, which is obviously impossible. This is due to the small cluster, whose objects are selected by the procedure and fitted, spoiling the computation which is not anymore able to recognize the v-shape. Therefore I chose to discard these points for what concerns the computation of the age.
In Figure 5.8 the fit is conducted after having removed the small cluster. The slope of the inner side for $11 < N < 18$ varies between $S_{in} = -5.07$ and $S_{in} = -3.88$, while the behaviour of the outer side is more consistent than the previous case, with the slope varying between $S_{out} = 4.58$ and $S_{out} = 2.39$.
The results of the fit are summarized in Figure 5.10, where the x-axis reports N and the y-axis the slopes. The left panel summarizes Figure 5.8 and the right panel Figure 5.9. In both panels, the blue indicates the inner side and the orange the outer side. The errors on the inner slope are much larger than the errors on the outer slope. In the right panel, the inner and the outer slopes are incompatible except for $N > 20$, when only three points are used in the fit. In the left panel the slopes are instead in very good agreement except for $N < 9$, where the bins are too narrow. This clearly shows that the fit without the cluster is more reliable than the fit with the cluster included.
However, for some values of N of the left panel there is still some significant difference between the two slopes, with the ratio between them varying from $R = 1.02$ to $R = 1.82$. This could indicate that the inner and the outer sides correspond to different ages, which would suggest that the Tirela family is the result of more than one collisional events. However, this may be simply due to the limited number of points that are used in the fit, which affects the overall performance of the algorithm. Analysing the plots by eye it can in fact be noticed that the fitted line not always approximates correctly the outer side, but sometimes it is more inclined than it, thus lowering the value of its slope. This

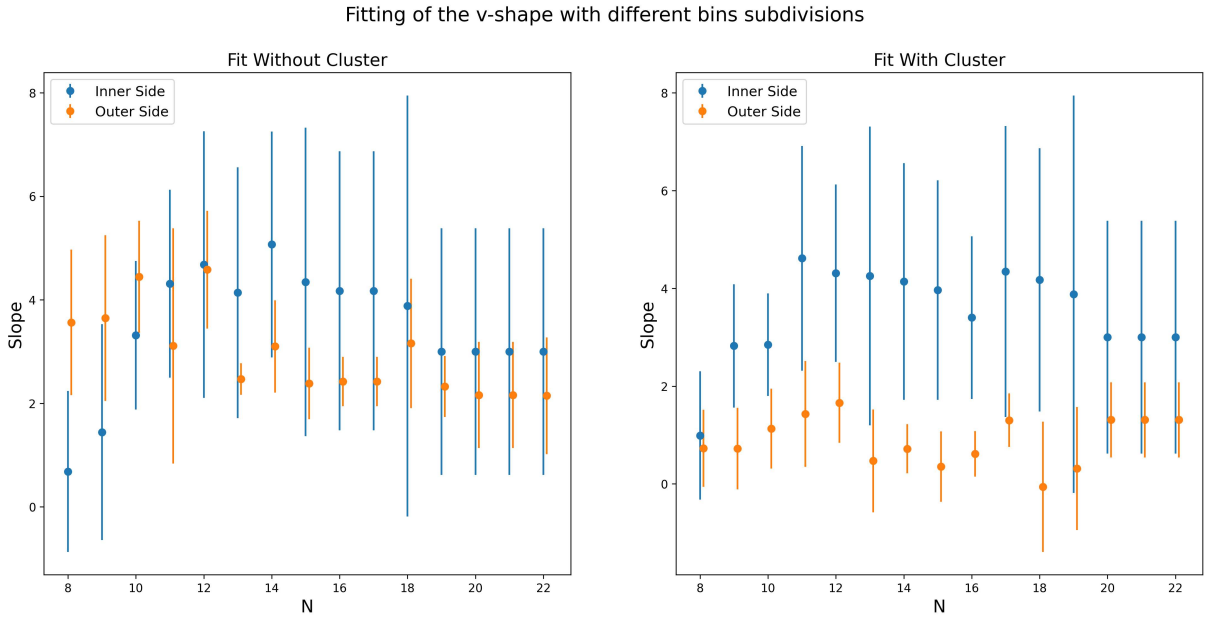

FIGURE 5.10: The slopes retrieved from the fit for $8 < N < 22$ for the case without the small cluster (left panel) and with the cluster included (right panel). The blue indicates the inner side and the orange the outer side.

would indicate that the inner and the outer sides correspond to a single age, which is in agreement with Milani et al. (2016). In addition, the authors also report that the outer slope is less well determined than the inner slope, imputing this to the $2/1J$ mean motion resonance, which is in agreement with what I found.
In order to constrain the age of the family I chose to focus on the case with $N = 12$ and without the small cluster on the outer side. The corresponding v-shape is showed in Figure 5.11, where the blue points are the family members and the black lines indicate the fitted v-shape.

The slope of the inner side of the v-shape is $S_{in} = -4.68 \pm 2.58$, while the slope of the outer side is $S_{out} = 4.58 \pm 1.14$. The associated errors have been retrieved by diagonalizing the covariance matrix computed during the linear regression.
In order to compare the results with Milani et al. (2016), I calculated the inverse slopes $1/S$ with the associated errors $\Delta(1/S)$ and the ratio $R = |S_{in}/S_{out}|$ with the associated error $\Delta(R) = R\ (\Delta S_{in}/S_{in} + \Delta S_{out}/S_{out})$. I obtained $1/S_{in} = -0.21 \pm 0.12$, $1/S_{out} = 0.22 \pm 0.05$ and $R = 1.02 \pm 0.81$.
The results are reported in Table 5.1, where they are compared to Milani et al. (2016). Their family consisted of 1950 members identified from the space of proper elements, while in my work I analyzed 150 objects identified from their spectra. Despite the two different

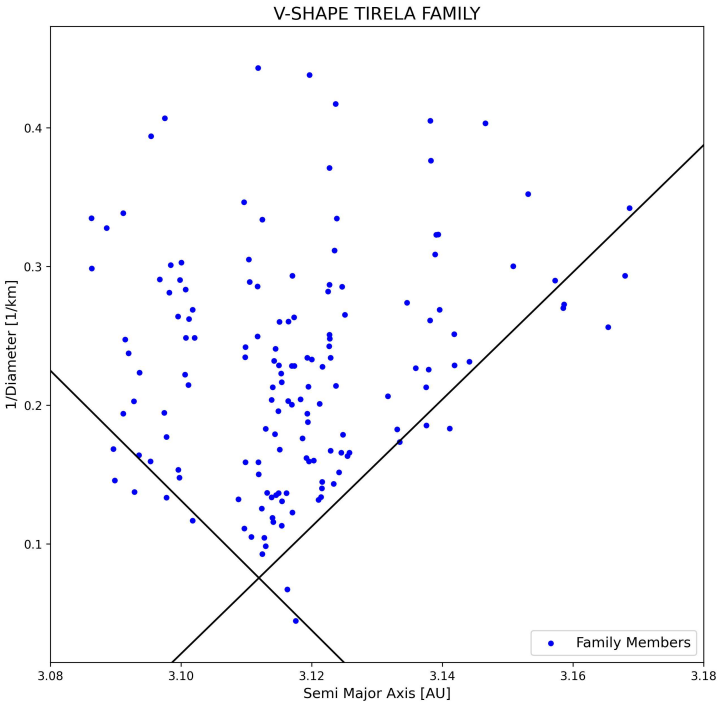

FIGURE 5.11: The v-shape of the Tirela family. In blue the family members, in black the lines fitting the sides of the v-shape.

	Milani et al. (2016)						This Work					
Side	Members	S	1/S	$\Delta(1/S)$	R	$\Delta(R)$	Members	S	1/S	$\Delta(1/S)$	R	$\Delta(R)$
IN	1950	-4.429	-0.226	0.025	0.99	0.24	150	-4.680	-0.214	0.118	1.02	0.81
OUT		4.494	0.223	0.048				4.583	0.218	0.054		

TABLE 5.1: The slopes of the v-shape of the Tirela family computed in this work and by Milani et al. (2016). The columns report the side, the number of objects belonging to the family, the slope S, the inverse slope $1/S$, the standard deviation of the inverse slope $\Delta(1/S)$, the ratio R and the standard deviation of the ratio $\Delta(R)$.

methods, my results are in good agreement with Milani et al. (2016).
My values are characterized by very large uncertainties. The v-shape in Figure 5.11 is in fact not very well defined, thus affecting the performance of the linear regression. Milani et al. (2016) identified instead more objects going to smaller diameters, thus reaching a better accuracy.

As explained in Section 2.2.3, in order to convert the slopes into ages it is necessary to determine a Yarkosvky calibration, which is the value of the Yarkovsky drift da/dt for a family member with unit diameter and spin axis obliquity equal to 0° for the inner side and to 180° for the outer side. Since the inverse slope is the change $\Delta(a)$ in semi-major

axis accumulated by a member with unit diameter over the age of the family, the age is simply $\Delta(t) = \Delta(a)/(da/dt)$.

Determining the Yarkovsky calibration da/dt requires the knowledge of the secular da/dt, the densities and the thermal properties of the family members (or at least of one of them). However, to date, these observations are not available for the Tirela family. In order to determine the age of the family, I used Equation 2.3 proposed by Milani et al. (2014), in which the Near Earth object Bennu is used for the calibration.

I noticed that in order to determine da/dt I would have used the same a, e, D and ρ used by Milani et al. (2016), with the only difference between their work and mine being the albedo. From NEOWISE albedos, I in fact obtained $\rho_{V,MW} = 0.252 \pm 0.069$ as the average albedo of the family, while Milani et al. (2016) adopted $\rho_{V,ML} = 0.204 \pm 0.100$, where the subscripts MW and ML refer respectively to "My Work" and "Milani". Equation 2.3 can therefore be simplified to:

$$\frac{da}{dt} = C\ (1 - \rho_V) \tag{5.1}$$

where C is a constant including the terms referring to Bennu, a, e, D and ρ. Dividing Equation 5.1 considered in the case of my work and of Milani and rearranging the terms I obtained:

$$(\frac{da}{dt})_{MW} = (\frac{da}{dt})_{ML}\ \frac{1 - \rho_{V,MW}}{1 - \rho_{V,ML}} \tag{5.2}$$

where $da/dt_{ML} = -3.40\ 10^{-4}$ AU/Myr for the inner side and $da/dt_{ML} = 3.37\ 10^{-4}$ AU/Myr for the outer side.

As stated by Milani et al. (2014) and Spoto et al. (2015), the calibration of the Yarkovsky drift is characterized by a relative uncertainty between 0.2 and 0.3. I therefore assumed the da/dt I found to be affected by an error of the 30%. From Equation 5.2 I obtained for the inner side $da/dt_{MW} = (-3.19 \pm 0.96)\ 10^{-4}$ AU/Myr and for the outer side $da/dt_{MW} = (3.17 \pm 0.95)\ 10^{-4}$ AU/Myr.

I then retrieved the age of the family as $Age = (1/S)/(da/dt)$, where $1/S$ is the inverse slope. I obtained an age of $670 \pm 570\ Myr$ for the inner side and an age of $688 \pm 377\ Myr$ for the outer side. The uncertainties have been computed by propagating the errors on the inverse slope and on the Yarkovsky calibration.

The results are reported in Table 5.2 where they are compared to Milani et al. (2016).

Side	Milani et al. (2016)			This Work		
	da/dt	Age	STD(Age)	da/dt	Age	STD(Age)
	10^{-4} AU/Myr	My	My	10^{-4} AU/Myr	My	My
IN	-3.40	664	212	-3.19	670	570
OUT	3.37	661	244	3.17	688	377

TABLE 5.2: Age estimation of the Tirela family obtained in Milani et al. (2016) and in this work. The columns report the side, the Yarkosvky calibration da/dt, the age and the standard deviation of the age.

The ages corresponding to the inner and the outer sides are compatible, thus indicating that the Tirela family is the result of a single collision. In addition, my results are in very good agreement with Milani et al. (2016).
Results reported in the literature may be discordant because of the different methods used to compute the ages. Because of this reason, I used as main comparison the paper by Milani et al. (2016), who used the procedure described in Spoto et al. (2015) on which my work is based. However, I compared my results with other authors that used different methods, such as Broz et al. (2012), who computed ages of Main Belt families by combining data in the (a, H) plane and N-body numerical simulations. They found the Tirela family to be younger than 1 Gyr, therefore compatible to my results.

Until now I focused on the case with $N = 12$ of Figure 5.8, which I selected because the two sides had similar slopes and because by eye the fit seemed to be the best compared to the other cases. However, the ages should be compatible between them and with literature independently of the value of N, apart from very low or high N. I therefore computed the ages for all the other cases with N varying between 11 and 18. The results are reported in Table 5.3, which proves that $N = 12$ is the best case among all the possible values of N. Despite not having the smallest uncertainties, this case is in fact the only one for which the inner and the outer ages are very similar, which is in agreement with the literature stating that the Tirela family is the result of a single collision. This thus justifies my previous decision to focus on this specific configuration.
For each N, the inner side has a very well defined shape, as confirmed by the ages, which are always very similar to the one determined for $N = 12$.
The outer side instead is less defined and this reflects on the age, which varies from 689 Gyr for $N = 12$ to 1324 Gyr for $N = 15$. The poor quality of the fit in the outer region is also proven by the uncertainties, which can be as large as 1 Gyr. As previously discussed, the older ages obtained for the outer side are likely due to the poor accuracy of the fitting

algorithm more than to a real difference in age between the two sides of the family.
The ages of the inner and outer sides are consistent with each other for every value of N. The large uncertainties affecting the measurements are due to the low numbers of objects found to belong to the family from Gaia spectra. It is however important to notice that all ages are in agreement with Milani et al. (2016) for every value of N, despite the two very different methods that are at the base of the identification of the family members.

From this analysis I proved that it is possible to recognize a family from the background using Gaia spectra alone, at least in a simple case in which the spectral type of the family is very different from the background. In addition, I managed to model the v-shape of the Tirela family and, by fitting its slopes, I constrained its age, which I found to be $670 \pm 570\ Myr$ for the inner side and $688 \pm 377\ Myr$ for the outer side, thus indicating a single collision scenario.
The contrast in albedo and composition between the family members and the background suggests that the parent body of the family did not originate in the present position, but that it likely came from a region of the inner Main Belt. Gaia spectra and my analysis seem to confirm this hypothesis.
The five objects circled in black in Figure 5.7 and the small cluster of seven objects of Figure 5.9 may indicate that the family may have experienced a more complex evolution than a single collision scenario. They in fact may not be interlopers, but large fragments originated from a previous collision, while smaller fragments would have been removed from the family after the interaction with resonances. This is a purely speculative scenario that cannot be proven with Gaia spectra alone and that could result from a Gaia observational bias, which was able to detect only the brightest objects of this region. More spectroscopic observations will be necessary to further explore the collisional and the dynamical evolution of the Tirela family.

N	Side	Age (Myr)	STD Age (Myr)
11	IN	726	523
	OUT	1015	1047
12	IN	670	570
	OUT	689	377
13	IN	756	669
	OUT	1279	542
14	IN	617	451
	OUT	1019	598
15	IN	720	710
	OUT	1324	780
16	IN	750	709
	OUT	1304	648
17	IN	750	709
	OUT	1304	648
18	IN	806	1086
	OUT	1000	695

TABLE 5.3: The ages and relative uncertainties determined from the v-shape of the Tirela family for N varying from 11 to 18 (where N is the number of objects within the same bin).

Chapter 6

Analysis of the Watsonia family

6.1 Introduction

Watsonia is an asteroid family first identified by Novaković et al. (2011), who applied the classical HCM to high inclination Main Belt asteroids.
Watsonia is located in the middle of the Main Belt, with 2.70 $AU < a_p < 2.83$ AU, and it is characterized by low to moderate orbital eccentricities, $0.10 < e_p < 0.15$, and relatively high inclinations, $16.5° < i_p < 18.0°$.
Watsonia is a very small family, with just 80 members listed in AstDys and 208 members listed in AFP. Tsirvoulis et al. (2014) tried to constrain the age of the family using the v-shape method, but because of the low number of family members they could not achieve a precise estimation and they report a rough age of the order of 1 Gyr.
As suggested by the same authors, the reason for which the Watsonia family is so depleted in members should be searched in its vicinity to the 5/2 mean motion resonance with Jupiter, which was proven to be very efficient in transporting asteroids to the Near Earth region. A significant number of fragments evolving in semi-major axis due to the Yarkovsky drift could have entered the resonance and thus removed from the family.

Watsonia, as Tirela, is an important repository of Barbarian asteroids (Section 5.1).
The parent member of the family, (729) Watsonia, was found to be a Barbarian by Gil-Hutton et al. (2014), while Cellino et al. (2014) observed nine objects belonging to the Watsonia region and discovered that seven of them showed a Barbarian behaviour, while the other two turned out to be interlopers.
Watsonia is known to be a L- type family, despite very few spectroscopic observations of

its members exist. A spectroscopic study of the family has been conducted by Devogèle et al. (2018), who reported the spectra of the three family members (729) Watsonia, (1372) Haremari and (3269) Vibert Douglas.

A still open problem about the Watsonia family concerns (980) Anacostia and (387) Aquitania. These are two fairly large bodies with diameters around 80/100 km that are not included in the family by any classification system because they possess higher eccentricities than the other family members. However, they share the same semi-major axes and inclinations and they present a lot of similarities in their physical properties. Most important, both (980) Anacostia and (387) Aquitania have been found to be Barbarians (Masiero and Cellino (2009), Gil-Hutton et al. (2008)).
Cellino et al. (2014) thus suggested a common origin for these two objects and the Watsonia family. They speculated that in a very early epoch a violent collision destroyed a big parent body which was the progenitor of all the Barbarians in this region. Most of the first generation fragments have been lost due to gravitational perturbations and the Yarkvosky effect, while the largest ones, (980) Anacostia and (387) Aquitania, due to their size have experienced almost zero Yarkovsky evolution and have remained in their original positions. Their relative high eccentricities may be due to perturbation effects with thin resonances crossing this region. In this picture, (729) Watsonia would be another first generation fragment which has then experienced a second collision that generated the present-day family. As stressed by the authors, this scenario is purely speculative and requires further analysis to be accepted as a plausible one.

Section 6.2 describes the procedure I implemented in order to distinguish the Watsonia family members from the background using Gaia data alone. In Section 6.3 a possible v-shape for the Watsonia family is discussed.

6.2 Identification of Watsonia family members

The Watsonia family shares some similarities with the Tirela family, since they both consist of L- type asteroids. However, the Watsonia family is much smaller than the Tirela family: AstDys lists only 80 asteroids in the former, while more than 1500 objects are included in the latter.
In addition, the two families are located in different parts of the Main Belt. The Tirela family lies in fact in the outer Main Belt, where the background asteroids belong to

the C- complex, which made easy to distinguish them from the L- type family members by using spectroscopic data alone. The Watsonia family instead is located in a high-inclination region poor of asteroids in the middle of the Main Belt, where the difference in spectral type between the background asteroids and the family members is less marked than in the Tirela case, thus reducing the efficacy to separate the family members from the background using Gaia data alone.
Keeping in mind that the Watsonia family would have been a more complicate case than the Tirela family, I however tried to reproduce the same analysis I conducted in Chapter 5.

First, I considered all asteroids contained inside a region centered on (729) Watsonia with $2.70\ AU < a_p < 2.83\ AU$, $0.105 < e_p < 0.150$ and $0.29 < \sin(i_p) < 0.31$. I retrieved their proper elements and absolute magnitudes from the Asteroid Family Portal, while for the albedos I considered NEOWISE data.
I found only 227 asteroids inside this region, 56 of which presented a spectrum in Gaia DR3. The distribution in absolute magnitude of these objects is reported in the left panel of Figure 6.1, where the complete sample is in blue and the Gaia sample in yellow. The red vertical line indicates the Gaia magnitude limit, $H = 15.93$, approximately corresponding to the peak of the distribution of the complete sample. In red it is reported the distribution of the 80 objects listed as part of the family by the Milani classification. Since the Nesvorny classification system includes in the family almost all the objects of the region (208 out of 227), I decided to not consider it for this analysis.
The right panel of Figure 6.1 shows instead the distribution in the $a - 1/H$ plane of the asteroids of the complete sample (yellow circles). The blue points are objects observed by Gaia and not included in Milani classification, while the red points are objects included in Milani and not observed by Gaia. Finally, black points are asteroids in common between Gaia observations and the Milani classification.
Objects observed by Gaia (blue and black points) are uniformly distributed in space, while the objects belonging to the family (red and black points) are almost all located at semi-major axes $a_p > 2.76\ AU$, which is also the position of the largest remnant (729) Watsonia. It therefore seems that, according to AstDys, the Watsonia family is one sided. This will be later verified by studying the v-shape of the family in the $a - 1/D$ plane.

The distribution of the albedo for the Watsonia family is reported in Figure 6.2, where the left panel shows the complete sample, the middle panel the Gaia sample and the right panel the Milani sample. A sort of bimodal distribution with limiting value at $\rho_V = 0.08$ can be recognized in the left and middle panels, even if less defined than in the case of the

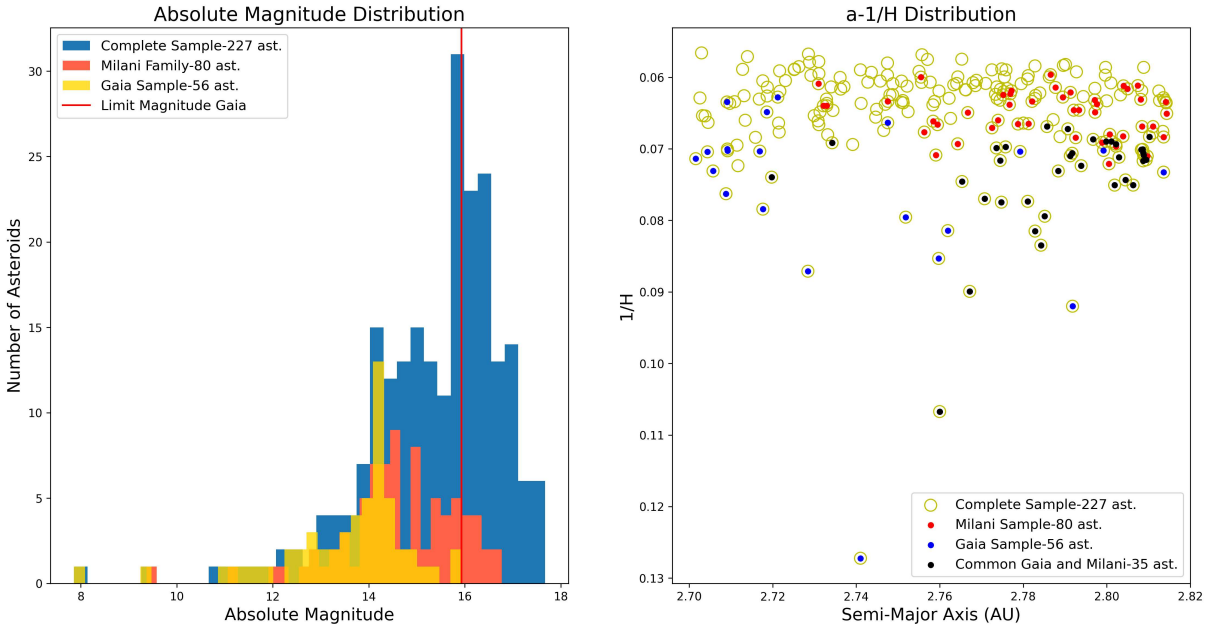

FIGURE 6.1: *Left panel:* Absolute magnitude distribution for asteroids in the complete sample (in blue), in the Gaia sample (in yellow) and in Milani classification (in red). *Right panel:* Distribution in the $a - 1/H$ plane of asteroids in the complete sample (yellow circles), in the Gaia sample (in blue), in Milani classification (in red) and in common between Gaia observations and Milani classification (in black).

Tirela family. The distribution of the total sample can in fact be subdivided in a small group of objects centered at $\rho_V = 0.05$ and in another bigger group centered at $\rho_V = 0.12$ and extending to very high albedos. The Gaia sample follows approximately the same behaviour of the complete sample, while in the distribution of the Milani classification high albedos ($\rho_V > 0.20$) and low albedos objects ($\rho_V < 0.10$) have been largely excluded from the family.
The albedo distribution of the objects in the Watsonia region thus confirmed the less marked difference between the spectral types of the family members and of the background.

In order to determine the spectral types of the objects observed by Gaia, I used the color-based taxonomy (Section 4.6). I first computed the color indexes for each object of the Gaia sample and I then retrieved their spectral types from their positions on the color-color diagram. For the objects assigned to the U- class, I selected as their spectral type the second more probable class.
The spectral types of the objects observed by Gaia in the Watsonia region are reported in the left panel of Figure 6.3, from where it seems that there is no predominance of any

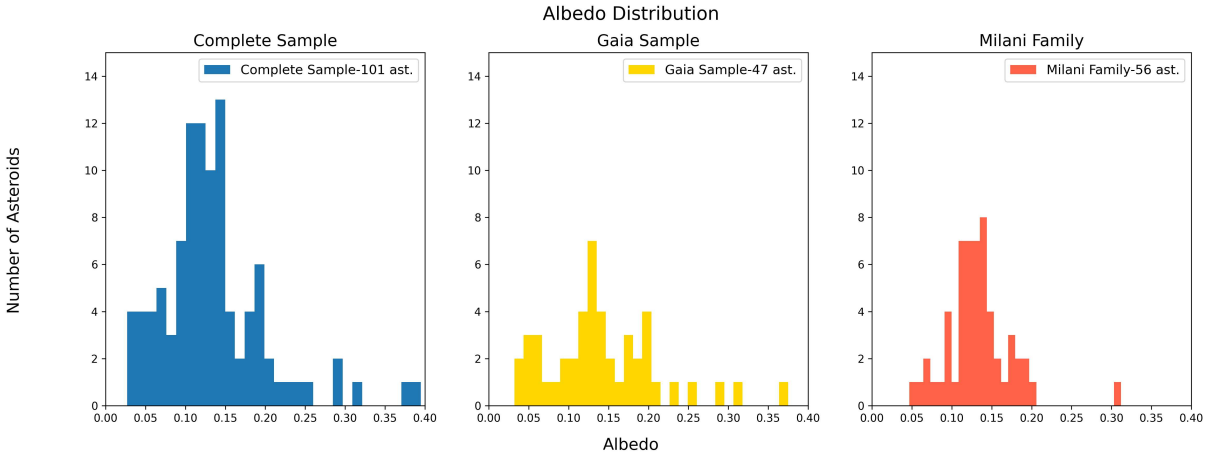

FIGURE 6.2: Albedo distribution for the objects of the complete sample (left panel), of the Gaia sample (middle panel) and of the Milani classification (right panel).

spectral type, with C-, D-, S- and X- types that are present in almost equal number, with then a small number of V- and L- types. I was not expecting to find any V- type in this region of the Main Belt, but, as already discussed for the case of Tirela, they are probably misclassified objects which present very flawed spectra. In addition, since the Watsonia family is reported to be a L- type family, I was expecting to find more objects classified as such. However, this may be due to their degeneracy with the S- types, as it was already observed for the Tirela family.

The spectral types in the left panel of Figure 6.3 have been divided according to the albedo, where in blue are asteroids with $\rho_V < 0.08$, in red with $\rho_V > 0.08$ and in grey objects without a NEOWISE albedo. The limiting value $\rho_V = 0.08$ has been selected from the bimodal distribution in Figure 6.2.

Despite the limited numbers of objects, I observed a general good agreement between the color-based taxonomy and the albedos. Only one object has been classified into the S-class despite its low albedo, while other five objects have been included in the C- class despite the high albedos.

The right panel of Figure 6.3 shows the spectral types retrieved from the Gaia spectra compared to the Milani classification. Objects observed by Gaia and included in the family by Milani are shown in orange, while objects not in common with Milani are shown in green.

I was expecting to find that the S- and L- types were included in the family, while the C-types were excluded. However, I observed that Milani included in the family objects of all spectral types, which is in contrast with the literature (Section 6.1) stating that the

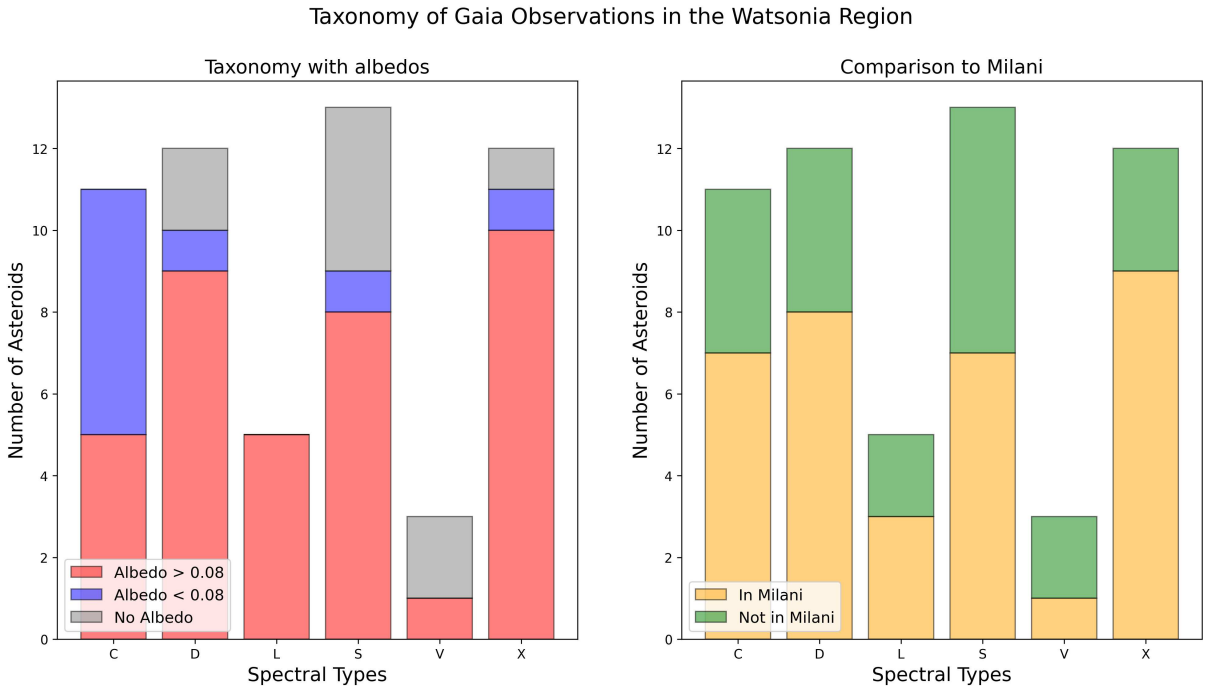

FIGURE 6.3: *Left panel:* Spectral types of the asteroids observed by Gaia in the Watsonia region. The red indicates objects with $\rho_V > 0.08$, the blue objects with $\rho_V < 0.08$ and the grey objects without albedo.
Right panel: same as the left panel, but with the orange indicating objects included in the Milani classification and the green objects not included.

Watsonia family consists of L- types.
In order to understand if this was due to a poor performance of the color-based taxonomy or to a real heterogeneous composition of the Watsonia family, I compared my results with the spectral types reported in the literature, which I retrieved from *rocks*. For each object, this Python package reports the taxonomy and the technique with which the spectral type has been determined. I considered only spectroscopic observations, discarding the objects whose taxonomy has been determined from photometry. Unfortunately, I found that of the 227 asteroids of the complete sample, only six presented a spectral type in the literature determined from spectroscopic observations. This therefore prevented me from giving a definitive answer to the previous question. However, on the other side, this analysis showed the great improvement in spectroscopic data due to the Gaia DR3 compared to literature. If in fact only six asteroids of the complete sample presented a spectral type determined from spectroscopic observations in the literature, I incremented the statistic to 56 objects using Gaia spectra and the color-based taxonomy.

Analyzing more in detail the spectral types retrieved from Gaia spectra I found that 35

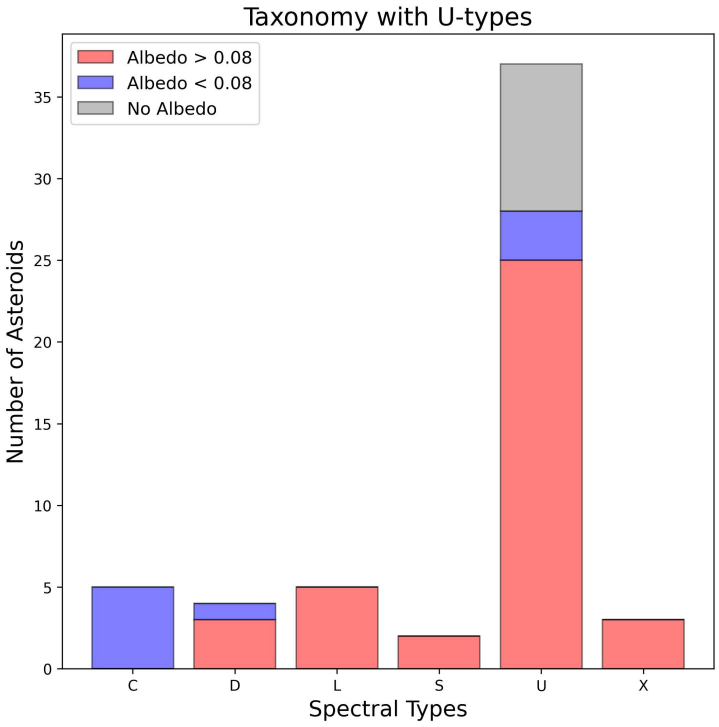

FIGURE 6.4: Spectral types of the asteroids observed by Gaia in the Watsonia region, but without considering the second more probable class for the objects initially classified as U- types. The red indicates objects with $\rho_V > 0.08$, the blue objects with $\rho_V < 0.08$, the grey objects without albedo.

objects, the 63% of the sample, were initially classified as U- types due to their bad quality. The distribution of the spectral types including the U- class is reported in Figure 6.4, where the blue indicates objects with $\rho_V < 0.08$, the red objects with $\rho_V > 0.08$ and the grey objects without albedo.

Of the 19 objects not included in the U- class, five are classified as C- types, all of which with low albedos. Other seven objects are instead classified as S- types or L- types, no one of which of low albedo. Almost all the objects classified as S- types in Figure 6.3 were initially classified as U- types. Same for the C- types of high albedos. Also the V- types have disappeared in Figure 6.4.

Despite the very limited sample, the correlation between albedos and spectral types determined with the color-based taxonomy has been proven to be valid also for the case of the Watsonia family. However, the majority of the asteroids observed by Gaia in this region were very faint and presented flawed spectra, therefore for most of them it resulted impossible to obtain a definitive spectral type. Despite this, Gaia DR3 still represents a huge improvement compared to literature, with 19 spectra of good quality compared to the 6 reported in the literature.

The limited number of objects observed by Gaia combined with the poor resolution of the spectra of most of them prevented me from conducting a detailed analysis on the Watsonia family such as the one reported in Chapter 5 for the Tirela family.
Due to the impossibility to retrieve the v-shape of the family using only the 19 spectra not classified into the U- class, I decided to maximize the number of objects by examining the spectral types of all the sample as reported in Figure 6.3. Among these, I then considered as part of the family the asteroids classified as L-, S-, D- and X- types. I excluded the three V- types, which I proved were characterized by flawed spectra, and the C- types, most of which presented low albedos.
I obtained that the Watsonia family, as seen by Gaia on the base of spectral properties, consisted of 42 members, which I analyzed in the following Section.

6.3 A possible v-shape for the Watsonia family?

In order to retrieve the v-shape of the family in the $a - 1/D$ plane, I first converted the absolute magnitudes H in diameters D. As in Section 5.3, I determined the diameters of the objects of known albedo from Equation 2.2, while for the remaining objects I used the average albedo of the family, which I found to be $\rho_V = 0.155 \pm 0.065$. For comparison, the albedo of (729) Watsonia is 0.124.
The distribution in the $a - 1/D$ plane is reported in the left panel of Figure 6.5, where the family members retrieved from Gaia observations are shown in blue. The Watsonia family is crossed by a resonance between 2.749 AU and 2.753 AU, indicated by the vertical black dashed lines, within which Gaia did not detect any object. The black X indicates (980) Anacostia, which, despite being located close to the vertex of the v-shape, is not considered as part of the family by any author because of its high eccentricity, as discussed in Section 6.1.
The right panel of Figure 6.5 reports the same objects of the left panel, except for (980) Anacostia, indicated by the corresponding spectral type letter retrieved from the color-based taxonomy. Blue objects are included in the Milani classification, while red objects are not part of it.
For each spectral type, there are some objects that are included in Milani and some objects that are not. If the Watsonia family really consists of only L- types as reported in the literature, this may be an indication that the color-based taxonomy is not accurate enough

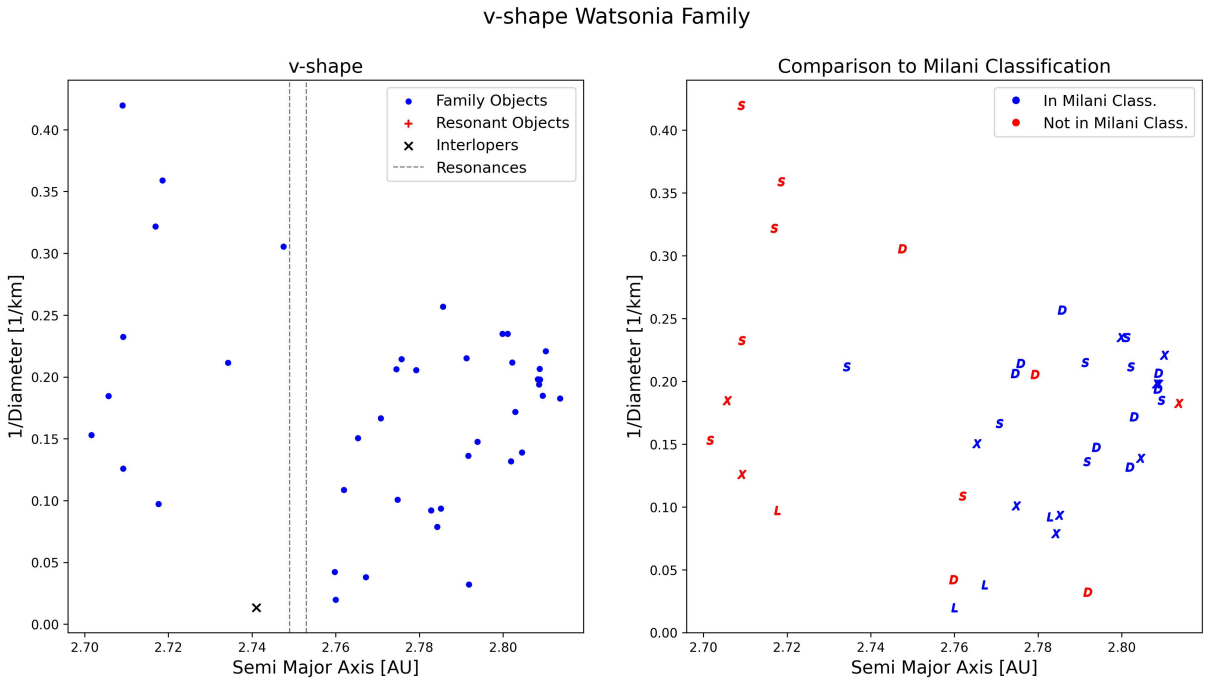

FIGURE 6.5: *Left panel:* Distribution in the $a - 1/D$ plane for family members determined from Gaia observations (blue points) and interlopers (black crosses). *Right panel:* Distribution in the $a - 1/D$ plane for family members included in the Milani classification (in blue) and not included (in red). Each object is marked with the letter corresponding to its spectral type.

to analyze this case or that the HCM fails to link the sparse distribution of fragments to the left of the resonance.

A sort of v-shape with its vertex on (729) Watsonia can be recognized in Figure 6.5, especially for $a > 2.76$ AU. The outer side, despite the limited number of objects, is in fact quite well defined. Instead, for $a < 2.76$ AU in the inner side there are only few objects randomly distributed in space and the v-shape cannot be recognized at all.
This is confirmed by the right panel of Figure 6.5. Milani in fact includes in the family only objects with $a > 2.76$ AU, while asteroids with smaller semi-major axes may therefore be defined as interlopers. However, these objects may also belong to the family and Milani does not include them because of limitations in the HCM method. The D- type between 2.78 AU and 2.80 AU which was the only object on the outer side deviating from the v-shape is instead excluded from the family.
The cluster of blue points included in the Milani classification resembles very closely a one-sided v-shape, with a very well defined outer boundary. This could have two explanations. The effect here observed may in fact be real, proving that the Watsonia family is actually

a one-sided family. This could be explained by assuming that, after the collision that created the family, part of the fragments experimented a positive drift in semi-major axis, thus forming the cluster of blue points that Gaia observed. Fragments that instead experimented a negative drift in semi-major axis could have been removed from the family after the interaction with the resonance highlighted in the left panel of Figure 6.5. The few objects in the left side could be fragments that jumped beyond the resonance immediately after the collision. This is however a purely speculative scenario that should require simulations of the dynamical evolution of the family to be proven.
The most likely explanation is that the one-sided v-shape is actually the result of an observational bias. Gaia in fact observed only 56 objects within the Watsonia region, 41 of which I classified as family members. Targeted observations may detect more asteroids in this region and thus reveal the inner side of v-shape. However, basing only on my work in which I analyzed Gaia data alone, it is impossible to give a definitive answer.

Despite having identified a v-shape for $a > 2.76\ AU$, I decided not to fit it in order to retrieve its slope and constrain the age of the family. As reported in Section 2.2.3, the age of a family is given by $Age = (1/S)/(da/dt)$, where $1/S$ is the inverse slope determined by fitting the v-shape and da/dt is the Yarkovsky calibration. However, both quantities resulted problematic to be determined in the case of the Watsonia family.
Due to the few objects detected by Gaia, the simplified procedure of the v-shape method I implemented would not work to fit the outer side. The subdivision in bins would in fact result in only two or three objects used for the linear regression, which therefore would be inaccurate and characterized by very large errors. In addition, from Gaia spectra it seems that the Watsonia family is a one sided-family, therefore only one age could be retrieved with the procedure. This would prevent me to compare the ages associated to the two sides and determine which subdivision in bins would be the best to implement the v-shape method (as I did in Section 5.4 for the Tirela family).
Fitting the v-shape therefore could result in an age that is not representative at all of the actual age of the family. More observations are necessary in order to populate the $a - 1/D$ plane with more objects in order to confirm if the Watsonia family is an actual one-sided family. This analysis should be repeated with the data that will be published in the Gaia DR4, which will contain more spectra than the DR3.
In addition, to date it is impossible to determine a valid Yarkovsky calibration for the Watsonia family. As reported in Equation 2.3, da/dt requires the knowledge of an albedo and a density representative for the family. While I determined an average albedo of the family by combining NEOWISE data ($\rho_V = 0.155 \pm 0.065$), an estimation of the density

is more complex to obtain. Since in the literature it is reported that the Watsonia family consists of L- types, the density of an L- type asteroid, even (729) Watsonia itself, could be used as a representative of all the family. However, in my work, I could not establish whether if the Watsonia family is homogeneous or heterogeneous in composition. From the color-based taxonomy I in fact obtained that the family consisted of a mixture of spectral types, with the L- class the least populated among them. As already shown in Figure 6.4, this is probably due to limitations in Gaia spectra and in the classificator. However, my results make impossible to use the Yarkovsky calibration proposed by Milani et al. (2014) and Spoto et al. (2015), which has been calibrated for homogeneous families. Very likely, an improvement in the color-based taxonomy and further spectroscopic observations at better resolution such as Gaia DR4 will prove that the objects observed by Gaia actually belong to the L- class, confirming that the Watsonia family is homogeneous in composition.

Since I could not implement the v-shape method for the Watsonia family in the same way I did for the Tirela family, I followed a different approach in order to constrain its age. In particular, I decided to overplot v-shapes corresponding to specific ages to see which one of them better adapted to the distribution of family members in the $a - 1/D$ plane.
To retrieve the slopes of the v-shape I rearranged the equation I used to constrain the age of the family from which I obtained:

$$S = (Age \cdot \frac{da}{dt})^{-1} \tag{6.1}$$

where S is the slope, Age the age of the family and da/dt the Yarkovsky calibration. However, from what it was discussed before, it resulted impossible to obtain a Yarkovsky calibration for the Watsonia family. I therefore decided to use at its place the da/dt determined for the Tirela family, reported in Table 5.2. The Tirela and Watsonia families both consist of L- types, therefore in first approximation their densities and albedos should be similar.
I then computed the slopes for the inner and outer sides of three v-shapes corresponding to 0.5, 1.0 and 1.5 Gyr and I selected (729) Watsonia as their vertex.
The resulting v-shapes are shown in Figure 6.6, where the blue dotted lines correspond to 0.5 Gyr, the red to 1.0 Gyr and the green to 1.5 Gyr. The v-shape that seems to better fit the family members is the one corresponding to 1.0 Gyr. This is in agreement with Tsirvoulis et al. (2014), who also found the age of the family to be of the order of 1 Gyr. My analysis is however extremely approximate. The Yarkovsky calibration determined for the Tirela family was in fact characterized by large uncertainties and using it for the

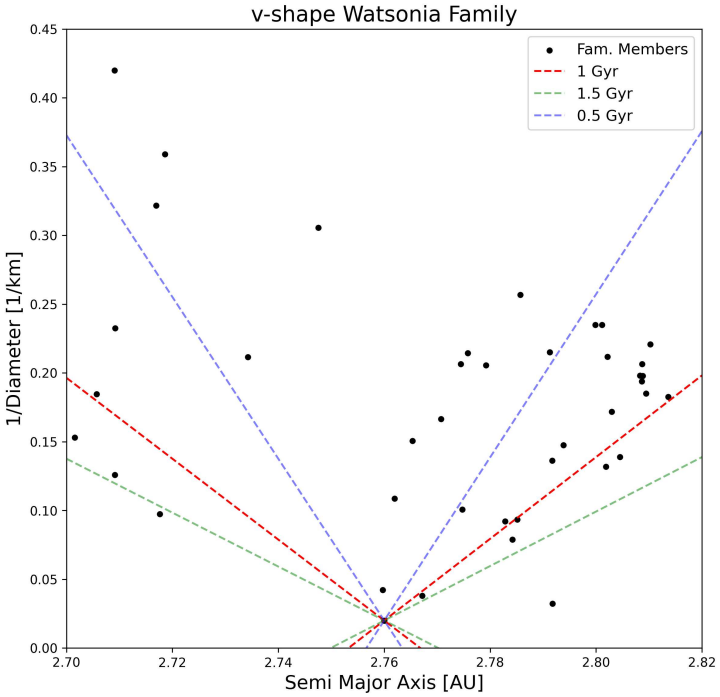

FIGURE 6.6: The family members (black points) in the $a - 1/D$ plane with overplotted the v-shapes corresponding to 0.5 Gyr (in blue), to 1.0 Gyr (in red) and to 1.5 Gyr (in green).

Watsonia family increased them. For example I know that its average albedo is quite different from that of the Tirela family.
This therefore affects the slopes of the v-shape, making difficult to constrain the age of the family. For example, an age of an 1.5 Gyr cannot be ruled out, while it seems quite unlikely that the family is younger than 0.5 Gyr.

The Watsonia family turned out to be a more intricate case than the Tirela family. The few objects in this region and the less marked difference between the spectral types of the family members and the background made more difficult to identify it using Gaia data alone.
Gaia observed 56 objects inside the Watsonia family, 19 of which had a good quality spectrum, and thus represents a huge improvement compared to literature, which instead reported only 6 spectroscopic observations of the same objects.
The intrinsic limitation of Gaia spectra and the color-based taxonomy made however complicated to distinguish the v-shape of the family, which appeared to be one-sided in my analysis. This prevented me from applying the v-shape method I implemented for the Tirela family, but I managed to constrain its age by overplotting v-shapes of specific ages

computed using the Yarkovsky calibration determined for the Tirela family. With this very approximate method I determined an age of the order of 1 Gyr.
Further spectroscopic observations, such as Gaia DR4, will be necessary in order to populate the $a - 1/D$ plane and confirm whether if the Watsonia family is really one-sided and heterogeneous in composition or if it is a result of an observational bias. Further observations will also be necessary to constrain with better accuracy the age of the family.

Chapter 7

Conclusion

The 60 518 asteroid reflectance spectra published in Gaia DR3 represent a huge improvement in spectroscopic data compared to previous surveys available in the literature, providing the opportunity to explore in deep asteroid families from a spectroscopic point of view.

Gaia spectra are however characterized by unexplained systematic deviations in the near infrared when compared to ground-based and template spectra, which prevent research work from being conducted. In the first part of this work I therefore explored possible correlations of these deviations with several parameters, such as the magnitudes and the spectral types of the observed objects. In particular, I wanted to prove that Gaia spectra were internally coherent, such that asteroids belonging to the same taxonomy class present similar spectra.

I first considered groups of asteroids of the same spectral type but of different magnitudes and I quantified their deviations from Mahlke template spectra, a recently published taxonomy scheme built over visible and near infrared spectra which reintroduced the albedo in the classification. I did not find any correlation of the entity of the deviations with the magnitude. I therefore dropped the luminosity classes and I computed the deviations from the templates for groups of asteroids of different taxonomy classes. I discovered that the entity of the deviations was correlating with the spectral type, being maximum for asteroids belonging to the S- complex, intermediate for C- complex asteroids and almost null for M- and P- types. I found that multiplicative correction factors were needed for every wavelength redder than 726 nm. I therefore proved that Gaia spectra are internally coherent, which allow to compare them with literature and to apply the criteria proposed by taxonomic systems to classify them.

However, since the deviations were correlating with the spectral type, I could not define an unique correction for all spectra. I therefore developed a procedure able to mitigate the deviations in the NIR for all Gaia spectra which did not depend on taxonomy. First, I defined two parameters differentiating among spectral types: the slope in the red, computed between 800 nm and 900 nm, and the maximum of the reflectance between 730 nm and 830 nm. I then built a template red slope - reflectance maximum diagram from bright objects of known taxonomy. On top of this diagram I defined seven regions, each one of them associated to specific correction factors. In order to correct the spectrum of an object, it is enough to compute the two parameters and apply the multiplicative factors associated to the region of the diagram within which it lies.
I then tested the procedure by correcting the spectra of the first 75 numbered asteroids observed by Gaia. Except for few cases in which Gaia spectra were irremediably flawed, I obtained a good agreement between the corrected Gaia spectra and the corresponding SMASS and Mahlke template spectra. I then used *classy*, a Python tool to classify asteroid reflectance spectra in the framework of Mahlke taxonomy, in order to classify all spectra in common between Gaia observations and Mahlke taxonomy. I found that the correction procedure improved the performance of *classy*. However, there still was a quite large number of misclassified Gaia spectra compared to the Mahlke spectral types, specifically concerning faint objects. Since the Tirela and the Watsonia families mostly consist of faint objects, I deduced that my correction procedure was not suited for the analysis I wanted to conduct on them.
I therefore looked for a new approach to classify Gaia spectra and I moved to a color-based taxonomy, which by converting Gaia reflectance spectra in SDSS colors retrieves the spectral type of a specific asteroid given its position on a color-color diagram. I tested this taxonomy by classifying a sample of faint asteroids of known albedos of the outer Main Belt centered on the Tirela family. I found a very good agreement between the spectral types retrieved from the color-based taxonomy and the albedos, with the Tirela family members of high albedos correctly classified as L- and S- types and the low albedo background asteroids classified as C- types. I therefore decided to adopt the color-based taxonomy in order to classify Gaia spectra and analyse the Tirela and the Watsonia families.

In the second part of this work I wanted to assess the feasibility of identifying asteroid families using Gaia spectra alone. In particular, I explored the Tirela and the Watsonia families, which both consist of L- type asteroids.
I first focused on the Tirela family, which being located in the outer Main Belt represented

the perfect benchmark to test Gaia data due to the very marked difference between the spectral types of the family members and the background asteroids.
I therefore considered all the asteroids contained in a region centered on Tirela, including family members and background. The NEOWISE albedos of these objects showed a bimodal distribution, with low albedos corresponding to the background and the high albedos to the family members. I then retrieved the spectral types of the asteroids observed by Gaia with the color-based taxonomy, from which I obtained a good agreement between S- complex family members, associated to high albedos, and C- complex background asteroids, associated to low albedos. I then considered as part of the family only the objects whose Gaia spectra were classified as L- or S- types. The family members I obtained are in very good agreement with the literature.
I then constrained the age of the family by implementing a simplified version of the v-shape method proposed by Spoto et al. (2015). I first converted the absolute magnitudes of the family members in diameters using NEOWISE albedos, when available, or an average albedo of the family when NEOWISE measurements were not available. After having identified the v-shape in the $a - 1/D$ plane, I fitted its inner and outer sides with a linear regression algorithm in order to determine its slopes. I then estimated a Yarkovsky calibration, which is the value of the Yarkovsky drift da/dt in semi-major axis for a family member of unit diameter, and I constrained the age of the family from $Age = (1/S)/(da/dt)$, where $1/S$ is the inverse of the slope. For the Tirela family, I obtained an age of $670 \pm 570\ Myr$ for the inner side and $688 \pm 377\ Myr$ for the outer side, which are consistent with the ages reported in the literature (Milani et al. (2016), Broz et al. (2012)).
I then moved to the study of the Watsonia family, which represented a more intricate case than the Tirela family because of its smaller size and its position in the middle of the Main Belt, where the difference between the spectral types of the family members and the background is less evident.
I repeated the same procedure implemented for the Tirela family. I first considered all the asteroids included in a region centered on the Watsonia family. However, Gaia observed only few tens of objects, most of which very faint and with flawed spectra. The correlation between spectral types and albedos seemed to be valid also in this case, despite the number of objects is too limited to be statistically relevant. In any case, this analysis proved the huge improvement in spectroscopic data of Gaia DR3 compared to previous surveys. Gaia in fact obtained 56 reflectance spectra for the asteroids in the Watsonia region, among which 19 of very good quality, while in the literature only 6 spectroscopic observations of the same objects are reported.

I then tried to retrieve the v-shape of the family in the $a - 1/D$ plane, but, because of the limited number of objects, my analysis resulted inconclusive. The Watsonia family appeared to be a one-sided family, meaning that only one side of the v-shape was recognizable, while the other side was very depleted in members. However, this could be due to a Gaia observational bias and not to a real physical effect.
The few objects detected by Gaia prevented me from applying the v-shape method in order to constrain the age of the family. Therefore I decided to compute v-shapes corresponding to different ages and to plot them over the family members in the $a - 1/D$ plane, from which I obtained that the Watsonia family has an age of the order of 1 Gyr. More observations will be however necessary to further explore the family, in particular to determine whether if it is actually one-sided and to constrain its age with better accuracy.

In this work I confirmed that Gaia spectra are internally coherent, such that asteroids of the same spectral type present similar spectra. This allowed me to study asteroid families, which I proved are possible to be identified using Gaia spectra alone. In particular, I succeeded in recognizing the Tirela and the Watsonia families, of which I constrained the ages, thus showing one of the many possibilities in which Gaia spectra can be exploited.
In this work I focused on two families only. A follow-up work could assess the feasibility of recognizing other types of families with different properties than Tirela and Watsonia. With Gaia spectra alone it is impossible to identify heterogeneous families or families for which the spectral type of the members is the same as the background. In these cases, Gaia data could be exploited a posteriori, in order to confirm the memberships or to identify interlopers.
Gaia DR4, that should be published at the end of 2025, will represent another huge improvement in spectroscopic data. It will in fact contain more spectra at higher S/N noise than the DR3. In addition, Gaia DR4 will detect fainter objects, thus improving the quality of the studies already conducted and increasing the statistics for small families such Watsonia.

References

Beekman, G., 2005. The nearly forgotten scientist ivan osipovich yarkovsky. Journal of the British Astronomical Association 115, 207–212.

Binzel, R., et al., 2019. Compositional distributions and evolutionary processes for the near-earth object population: Results from the mit-hawaii near-earth object spectroscopic survey (mithneos). Icarus 324, 41–76.

Bottke, W., et al., 2006. The yarkovsky and yorp effects: Implications for asteroid dynamics. Annual Review of Earth and Planetary Sciences 34, 157–191.

Broz, M., et al., 2012. Constraining the cometary flux through the asteroid belt during the late heavy bombardment. Astronomy & Astrophysics 551, 16 pp.

Burbine, T., 2016. Reflectance spectroscopy and asteroid taxonomy. Asteroids: Astronomical and Geological Bodies, Cambridge University Press , 133 – 184.

Burbine, T., Binzel, R., 2002. Small main-belt asteroid spectroscopic survey in the near-infrared. Icarus 159, 468–499.

Bus, S., Binzel, R., 2002a. Phase ii of the small main-belt asteroid spectroscopic survey: A feature-based taxonomy. Icarus 158, 146–177.

Bus, S., Binzel, R., 2002b. Phase ii of the small main-belt asteroid spectroscopic survey: The observations. Icarus 158, 106–145.

Carpino, M., et al., 2003. Error statistics of asteroid optical astrometric observations. Icarus 166, 248–270.

Carruba, V., et al., 2013. A multidomain approach to asteroid families' identification. Monthly Notices of the Royal Astronomical Society 433, 2075–2096.

Carruba, V., et al., 2018. Asteroid families interacting with secular resonances. Planetary and Space Science 157, 72–81.

Carruba, V., et al., 2019. Machine-learning identification of asteroid groups. Monthly Notices of the Royal Astronomical Society 488, 1377–1386.

Cellino, A., et al., 2006. The strange polarimetric behavior of asteroid (234) barbara. Icarus 180, 565–567.

Cellino, A., et al., 2014. A successful search for hidden barbarians in the watsonia asteroid family. Mon. Not. R. Astron. Soc. 439, 75–79.

Chapman, C., et al., 1971. A review of spectrophotometric studies of asteroids. Physical Studies of Minor Planets 267, 51–65.

Chesley, S., et al., 2003. Direct detection of the yarkovsky effect by radar ranging to asteroid 6489 golevka. Science 302, 1739–1742.

Connelly, J., et al., 2012. The absolute chronology and thermal processing of solids in the solar protoplanetary disk. Science 338, 651–655.

Cropper, M., et al., 2018. Gaia data release 2. gaia radial velocity spectrometer. Astronomy & Astrophysics 616, 19 pp.

Daly, T., Schultz, P., 2018. The delivery of water by impacts from planetary accretion to present. Science Advances 4.

Deienno, R., et al., 2021. Efficiency characterization of the v-shape asteroid family detection method. Icarus 357.

DeMeo, F., et al., 2009. An extension of the bus asteroid taxonomy into the near-infrared. Icarus 595, 160–180.

Devogèle, M., et al., 2018. New polarimetric and spectroscopic evidence of anomalous enrichment in spinel-bearing calcium-aluminium-rich inclusions among l-type asteroids. Icarus 304, 31–57.

Donovan, M., et al., 2017. A probabilistic asteroid impact risk model: assessment of sub-300 m impacts. Icarus 289, 106–119.

Fabricius, C., et al., 2016. Gaia data release 1. pre-processing and source list creation. Astronomy & Astrophysics 595, 21 pp.

Farinella, P., et al., 1996. Asteroid families, old and young. Completing the Inventory of the Solar System, Astronomical Society of the Pacific Conference Proceedings 107, 45–55.

Fukugita, M., et al., 1996. The sloan digital sky survey photometric system. Astronomical Journal 111, 1748.

Gaia Collaboration, 2016. The gaia mission. Astronomy & Astrophysics 595, 36 pp.

Gaia Collaboration, 2022a. Gaia data release 3: Reflectance spectra of solar system small bodies. Astronomy & Astrophysics , 30 pp.

Gaia Collaboration, 2022b. Gaia data release 3: Summary of the content and survey properties. Astronomy & Astrophysics , 23 pp.

Gil-Hutton, R., et al., 2008. New cases of unusual polarimetric behavior in asteroids. Astronomy & Astrophysics 482, 309 – 314.

Gil-Hutton, R., et al., 2014. Polarimetric survey of main-belt asteroids, iv. new results from the first epoch of the casleo survey. Astronomy & Astrophysics 569, 6 pp.

Granvik, M., et al., 2017. Escape of asteroids from the main belt. Astronomy & Astrophysics 598, 13 pp.

Heck, P., et al., 2017. Rare meteorites common in the ordovician period. Nature Astronomy 1, 1–6.

Hirayama, K., 1918. Groups of asteroids probably of common origin. Astronomical Journal 31, 185–188.

Hsieh, H., et al., 2018. Asteroid family associations of active asteroids. The Astronomical Journal 155, 22 pp.

Ivezić, Z., et al., 2001. Solar system objects observed in the sloan digital sky survey commissioning data. The Astronomical Journal 122, 2749–2784.

Lazzarin, M., et al., 2005. Spectroscopic investigation of near-earth objects at telescopio nazionale galileo. Monthly Notices of the Royal Astronomical Society 359, 1575–1582.

Lazzaro, D., et al., 2004. S3os2: the visible spectroscopic survey of 820 asteroids. Icarus 172, 179–220.

van Leeuwen, F., 1997. The hipparcos mission. Space Science Reviews 81, 201–409.

Libourel, G., et al., 2017. Search for primitive matter in the solar system. Icarus 282, 375–379.

Lindegren, L., Bastian, U., 2010. Basic principles of scanning space astrometry. EAS Publications Series 45, 109–114.

Mahlke, M., et al., 2022. Asteroid taxonomy from cluster analysis of spectrometry and albedo. Astronomy & Astrophysics 665, 32 pp.

Masiero, J., Cellino, A., 2009. Polarization of asteroid (387) aquitania: The newest member of a class of large inversion angle asteroids. Icarus 199, 333–337.

Masiero, J., et al., 2011. Main belt asteroids with wise/neowise. i. preliminary albedos and diameters. The Astrophysical Journal 741, 20 pp.

Masiero, J., et al., 2013. Asteroid family identification using the hierarchical clustering method and wise/neowise physical properties. The Astrophysical Journal 770, 22 pp.

Masiero, J., et al., 2015. Asteroid family physical properties. Asteroids IV, University of Arizona Press, Tucson , 323–340.

McCord, T., et al., 1970. Asteroid vesta: Spectral reflectivity and compositional implications. Science 168, 1445–1447.

Michel, P., et al., 2015. Collisional formation and modeling of asteroid families. Asteroids IV, University of Arizona Press, Tucson , 341–354.

Migliorini, F., et al., 1995. Interlopers within asteroid families. Icarus 118, 271–291.

Milani, A., et al., 2014. Asteroid families classification: Exploiting very large datasets. Icarus 239, 46–73.

Milani, A., et al., 2016. Families classification including multiopposition asteroids. Asteroids: New Observations, New Models, Proceedings of the International Astronomical Union, IAU Symposium 318, 28–45.

Morbidelli, A., Henrard, J., 1991. Secular resonances in the asteroid belt - theoretical perturbation approach and the problem of their location. Celestial Mechanics and Dynamical Astronomy 51, 131–167.

Mothé-Diniz, T., Nesvorný, D., 2008. Tirela: an unusual asteroid family in the outer main belt. Astronomy & Astrophysics 492, 593–598.

Nesvorný, D., et al., 2002. Regular and chaotic dynamics in the mean-motion resonances: Implications for the structure and evolution of the asteroid belt. Asteroids III, W. F. Bottke Jr., A. Cellino, P. Paolicchi, and R. P. Binzel (eds), University of Arizona Press , 379–394.

Nesvorný, D., et al., 2003. Recent origin of the solar system dust bands. The Astrophysical Journal 591, 486–497.

Nesvorný, D., et al., 2005. Evidence for asteroid space weathering from the sloan digital sky survey. Icarus 173, 132–152.

Novaković, B., et al., 2022. Asteroid families: properties, recent advances, and future opportunities. Celestial Mechanics and Dynamical Astronomy 134.

Novaković, N., et al., 2011. Families among high-inclination asteroids. Icarus 216.

Parker, A., et al., 2008. The size distributions of asteroid families in the sdss moving object catalog 4. Icarus 198, 138–155.

Peale, S., 1976. Orbital resonance in the solar system. Annual review of astronomy and astrophysics 14, 215–246.

Perryman, M., 2012. The history of astrometry. The European Physical Journal 37, 745–792.

Perryman, M., et al., 2014. Astrometric exoplanet detection with gaia. The Astrophysical Journal 797.

Pfalzner, S., et al., 2015. The formation of the solar system. Physica Scripta 90.

Pravec, P., et al., 2005. Tumbling asteroids. Icarus 173, 108–131.

Radović, V., et al., 2017. An automatic approach to exclude interlopers from asteroid families. Monthly Notices of the Royal Astronomical Society 470, 576–591.

Reddy, V., et al., 2015. Mineralogy and surface composition of asteroids. Asteroids IV, Patrick Michel, Francesca E. DeMeo, and William F. Bottke (eds.), University of Arizona Press, , 43–63.

Riello, M., et al., 2018. Gaia data release 2. processing of the photometric data. Astronomy & Astrophysics 616, 19 pp.

Rivkin, S., et al., 2002. Hydrated minerals on asteroids: The astronomical record. Asteroids III, W. F. Bottke Jr., A. Cellino, P. Paolicchi, and R. P. Binzel (eds), University of Arizona Press , 235–253.

Rubincam, D., 1990. Drag on the lageos satellite. Journal of Geophysical Research: Solid Earth 95, 4881–4886.

Rubincam, D., 2000. Radiative spin-up and spin-down of small asteroids. Icarus 148, 2–11.

Sergeyev, A., Carry, B., 2021. A million asteroid observations in the sloan digital sky survey. Astronomy & Astrophysics 652, 12 pp.

Sergeyev, A., et al., 2022. Multifilter photometry of solar system objects from the skymapper southern survey. Astronomy & Astrophysics 658, 15 pp.

Spoto, F., et al., 2015. Asteroid family ages. Icarus 257, 275–289.

Sunshine, J., et al., 2008. Ancient asteroids enriched in refractory inclusions. Science 320, 514–517.

Tanga, P., et al., 2022. Gaia data release 3: The solar system survey. Astronomy & Astrophysics , 34 pp.

Tholen, D., 1984. Asteroid taxonomy from cluster analysis of photometry. PhD Thesis .

Tinaut-Ruano, F., et al., 2023. Asteroids' reflectance from gaia dr3: Artificial reddening at near-uv wavelengths. Astronomy & Astrophysics 669, 12 pp.

Tiscareno, M., Malhotra, R., 2008. Chaotic diffusion of resonant kuiper belt objects. The Astronomical Journal, 338.

Tsirvoulis, G., et al., 2014. Dynamical properties of the watsonia asteroid family. Asteroids, Comets, Meteors. Proceedings of the conference held 30 June-4 July, 2014 .

Vigna, A.D., et al., 2018. Detecting the yarkovsky effect among near-earth asteroids from astrometric data. Astronomy & Astrophysics 617, 16 pp.

Walsh, K., et al., 2013. Introducing the eulalia and new polana asteroid families: Re-assessing primitive asteroid families in the inner main belt. Icarus 225, 283–297.

Xu, S., et al., 1995. Small main-belt asteroid spectroscopic survey: Initial results. Icarus 115, 1–35.

York, G., et al., 2000. The sloan digital sky survey: Technical summary. The Astronomical Journal 120, 1579–1587.

Zappalà, V., et al., 1990. Asteroid families. i. identification by hierarchical clustering and reliability assessment. The Astronomical Journal 100, p. 2030.

Zellner, B., et al., 1985. The eight-color asteroid survey: Results for 589 minor planets. Icarus 61, 355–416.