Università degli Studi di Torino
Scuola di Dottorato

Dottorato in Fisica

A Machine Learning approach for astrophysical X-ray polarimetry with Gas Pixel Detectors

Candidate: Nicoló Cibrario

Supervisor: Prof. Raffaella Bonino

Contents

Introduction

In astrophysics four different techniques are used to study the information carried by electromagnetic radiation: photometry, imaging, spectroscopy, and polarimetry. Photometry refers to the measurement of the flux or intensity of light emitted by celestial objects, whereas imaging focuses on analyzing spatially resolved representations of these objects. Spectroscopy aims at studying the radiation properties as a function of energy, while polarimetry examines the geometrical orientation of electromagnetic field oscillations. X-ray astrophysics has studied and demonstrated the dynamic nature and complex structures of celestial sources through the combined use of the first three techniques, whereas polarimetry has played a marginal role until recent years. Still in the early years of the 21st centuries, the challenging and time-consuming nature of X-ray astrophysical polarization measurements lagged behind the advancements in spectroscopy and imaging.

The development of the Gas Pixel Detector (GPD) [1], and with it the opportunity to exploit the photoelectric effect for polarization measurements, changed this narrative, achieving unprecedented sensitivity in the 1-10 keV energy range. The interest of the astrophysical community for the X-ray polarimetry reached a climax with the launch of the Imaging X-ray Polarimetry Explorer (IXPE) on December 9th, 2021, a NASA mission designed to perform astrophysical X-ray polarimetry in the 2-8 keV energy range [2]. Since the beginning of data acquisition, IXPE demonstrated the ability of X-ray polarimetry to access new and groundbreaking information about the properties of a wide variety of celestial objects.

IXPE employs three GPDs to determine the polarization properties of incident photons, by analyzing the tracks of photo-electrons (PEs) generated through the interaction of the photons with the gas contained in the detector. The state-of-the-art IXPE data analysis relies on the analytic reconstruction of these tracks [3]. While reading the early chapters of this thesis, it will become evident how this kind of reconstruction well aligns with the use of Machine Learning (ML) algorithms. The application of these techniques to IXPE data reconstruction has already been explored by other few groups, leading to interesting and promising results [4; 5; 6]. However, these works demonstrated that there is substantial room for improvement as well, for both the analysis of IXPE data and for future X-ray polarimetry missions.

This thesis addresses the scenario just described, presenting a new algorithm that leverages the promising performance of Convolutional Neural Networks (CNNs) while maintaining the core structure of the standard analytic algorithm. The goal is to create a stable and high-performing model, that could be applicable to current IXPE data, and also benefit future X-ray polarimetry missions and similar applications.

Chapter 1 introduces the physics of polarimetry, alongside with the main experimental techniques involved in astrophysical measurements. Chapter 2 starts with an overview of astrophysical X-ray polarimetry and then focuses on IXPE mission and instrument. The Gas Pixel Detector, its components and functioning, as well as the reconstruction

of the events, are discussed. Chapter 3 is dedicated to IXPE science, and particular attention is directed towards sources which are functional to test the application of the algorithm developed in this work. Chapter 4 gives a brief introduction to Machine Learning and Convolutional Neural Networks, describing in greater details the models and methods involved in the project of this thesis. Chapter 5 depicts the developing phase of the algorithm and reports the performance achieved with simulated data. Chapter 6 validates the results obtained with simulations by applying the algorithm to laboratory data. Finally, Chapter 7 reports examples of its application with real IXPE data.

Chapter 1

X-Ray Polarimetry

X-ray polarimetry is a relatively new field of study in astrophysics. Recent advancements in technology and innovative ideas have allowed for significant growth in this area over the past decade. The goal of X-ray polarimetry is to understand the physical processes that underlie the production of polarized photons by astrophysical sources and to expand our knowledge in this discipline.

The fundamentals of X-ray polarimetry will be discussed in the first section of this chapter, followed by an introduction of the techniques and the instrumentation which can be used to perform polarimetry measurements.

1.1 Polarimetry Fundamentals

Generally, polarization is a property of the light referring to the geometrical orientation of the oscillations of the electromagnetic fields. In classical physics, the light is an electromagnetic wave consisting of a coupled oscillating electric and magnetic field perpendicular to each other. Conventionally, the direction of the polarization corresponds to the direction of the electric field.

Completely polarized radiation

In this section the focus is on *monochromatic plane waves*, i.e. waves consisting of a single frequency or wavelength, with ideally flat and planar fronts. The polarization vector, or the electric field vector, can be decomposed along the two orthogonal x and y axes. Writing as $\hat{x}$ and $\hat{y}$ the corresponding unit vectors, the value of the polarization vector over time can be written as [7]:

$$\overrightarrow{E}(t) = (\hat{x}E_1 + \hat{y}E_2)e^{i\omega t}. \tag{1.1}$$

A representation of the $\overrightarrow{E}$ vector at a fixed time t, alongside with the parameters used to describe the quantities linked to the polarization vector in this section, is represented in Fig. 1.1.

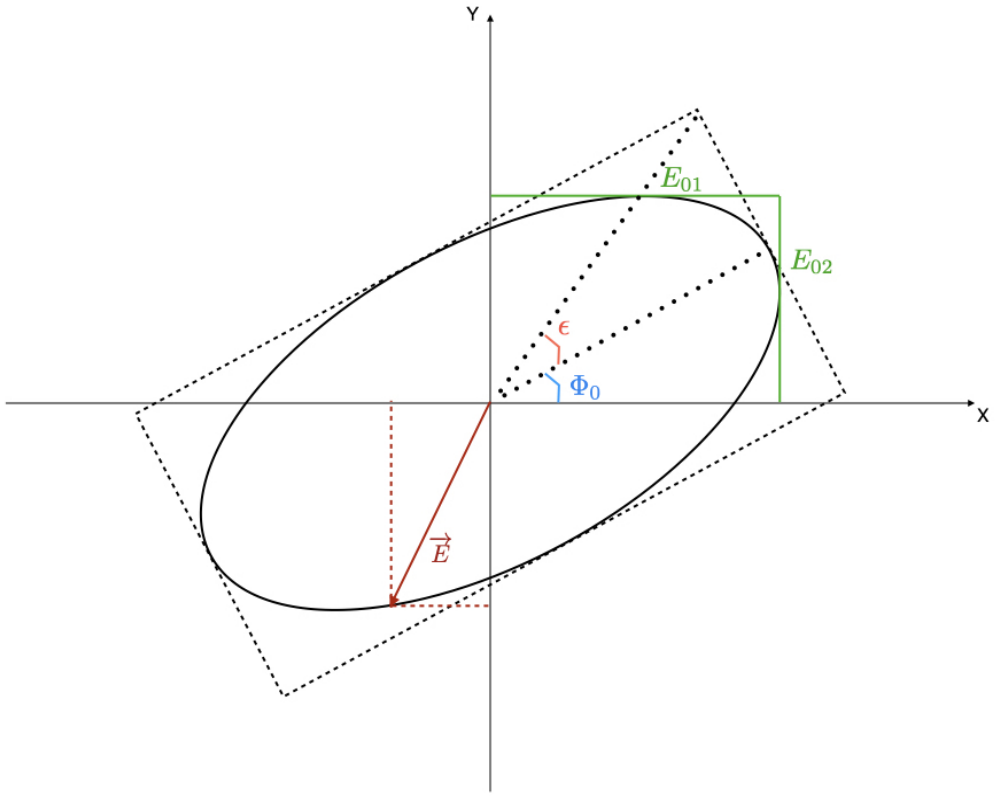

Figure 1.1: **Polarization ellipse.**

By introducing arbitrary phase angles ϕ_1 and ϕ_2, the E_1 and E_2 components can be expressed as:

$$E_1 = E_{01}e^{i\phi_1} \tag{1.2}$$

$$E_2 = E_{02}e^{i\phi_2}. \tag{1.3}$$

The physical components of the polarization vector $\vec{E}$ along $\hat{x}$ and $\hat{y}$ are defined by its real part:

$$E_x(t) = E_{01}\cos(\omega t + \phi_1) \tag{1.4}$$

$$E_y(t) = E_{02}\cos(\omega t + \phi_2). \tag{1.5}$$

These equations allow to describe the variation over time of the x and y components of the polarization vector. The oscillations can be visualized as an ellipse in the plane perpendicular to the wave's propagation direction, as shown in Figure 1.1. The rotation angle Φ_0 and the angle ϵ, which is linked to the ellipticity of the polarization, can be written as a function of the ϕ_1 and ϕ_2 angles as:

$$\tan 2\Phi_0 = \frac{2E_{01}E_{02}\cos\alpha}{E_{01}^2 - E_{02}^2} \tag{1.6}$$

$$\sin 2\epsilon = \frac{2E_{01}E_{02}\sin\alpha}{E_{01}^2 - E_{02}^2} \tag{1.7}$$

where $\alpha = \phi_2 - \phi_1$. The polarization is called *right-handed* if $\overrightarrow{E}$ traces the ellipse clockwise, while it is called *left-handed* if counter-clockwise.

Equations 1.4 and 1.5 can be easily elaborated to obtain [7]:

$$\frac{E_x^2}{E_{01}^2} - 2\frac{E_xE_y}{E_{01}E_{02}}\cos\alpha + \frac{E_y^2}{E_{02}^2} = \sin^2\alpha. \tag{1.8}$$

Since the period of the wave oscillation is significantly shorter than the time of observation, Eq. 1.8 is still valid considering the time average over the wave period of the E_x and E_y values:

$$\frac{<E_x^2>}{E_{01}^2} - 2\frac{<E_xE_y>}{E_{01}E_{02}}\cos\alpha + \frac{<E_y^2>}{E_{02}^2} = \sin^2\alpha. \tag{1.9}$$

By multiplying Eq. 1.9 for $4E_{01}^2E_{02}^2$ and, being E_x and E_y sinusoidal waves, calculating the time average $<E_{x(y)}^2> = \frac{1}{2}E_{01(2)}^2$ we find:

$$2E_{02}^2E_{01}^2 + 2E_{01}^2E_{02}^2 - (2E_{01}E_{02}\cos\alpha)^2 = (2E_{01}E_{02}\sin\alpha)^2 \tag{1.10}$$

that can be rearranged to obtain:

$$I^2 = U^2 + Q^2 + V^2 \tag{1.11}$$

where I,U,Q and V are called *Stokes parameters* and are defined as:

$$I = E_{01}^2 + E_{02}^2 = E_{00}^2 \tag{1.12}$$

$$Q = E_{01}^2 - E_{02}^2 = E_{00}^2\cos(2\epsilon)\cos(2\Phi_0) \tag{1.13}$$

$$U = 2E_{01}E_{02}\cos\alpha = E_{00}^2\cos(2\epsilon)\sin(2\Phi_0) \tag{1.14}$$

$$V = 2E_{01}E_{02}\sin\alpha = E_{00}^2\sin(2\epsilon) \tag{1.15}$$

where $E_{00}^2 = E_{01}^2 + E_{02}^2$ is the total intensity of the radiation. The Stokes parameters have important physical interpretations: the I parameter represents the total intensity of the radiation, Q and U characterize the intensity of linearly polarized radiation along different axes, and V describes the intensity of circularly polarized radiation. A representation of their interpretation is reported in Fig. 1.2.

Two key equations which link the Stokes parameters to the polarization properties of the radiation are:

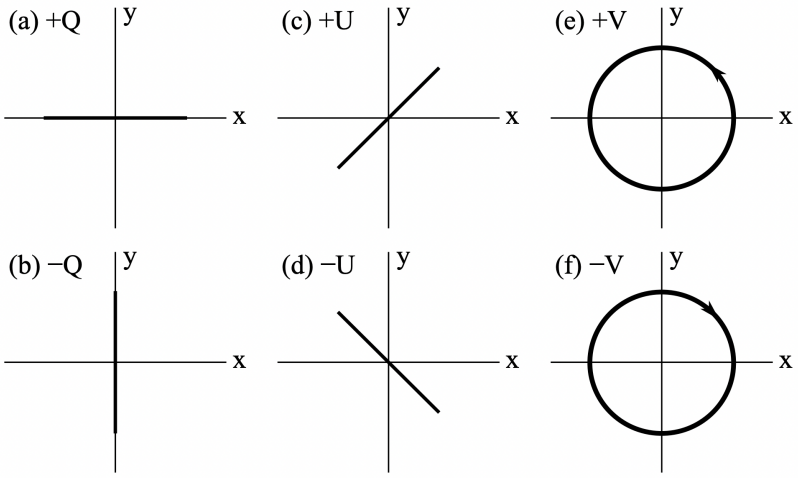

Figure 1.2: **Examples of polarization vectors for Stokes parameters nominal values.** Image credits: [8]

$$\tan(2\Phi_0) = \frac{2E_{01}E_{02}\cos\alpha}{E_{01}^2 - E_{02}^2} = \frac{U}{Q} \tag{1.16}$$

$$\sin(2\epsilon) = \frac{2E_{01}E_{02}\sin\alpha}{E_{01}^2 - E_{02}^2} = \frac{V}{I}. \tag{1.17}$$

Equations 1.16-1.17 indicate that the Stokes parameters can fully describe the polarization ellipse properties, with Q and U defining the rotation angle Φ_0, while V and I being related to the ellipticity ϵ.

Partially polarized radiation

As mentioned before, the monochromatic waves that have been discussed are referred to as *completely* or *fully* polarized, as the polarization vector demonstrates a consistent and non-random directional pattern over time. Nonetheless, in astrophysics as in other real-world scenarios, we never encounter a single monochromatic component, but radiation which can be decomposed as a superposition of many wave components, each one with its own polarization.

By introducing the concept of *quasi-monochromatic* waves, i.e. waves for which amplitudes and phases vary *slowly* with time, it is possible to derive the general relation which links the Stokes parameters [7]:

$$I^2 \geq U^2 + Q^2 + V^2. \tag{1.18}$$

We know that the sign equal is true for completely polarized radiation, while for completely unpolarized radiation there is not a preferred orientation for the polarization vector on the (x,y) plane, and the Eq. 1.18 becomes:

$$Q = U = V = 0. \tag{1.19}$$

This short digression helps introducing the concept of *partially polarized radiation*. We said that radiation can be considered as a superposition of single independent wave components, for which Stokes parameters are additive. Therefore, for k wave components the following equation holds:

$$X = \sum_k X^{(k)} \tag{1.20}$$

where X represents each Stokes parameter. The set of the Stokes parameters for a general electromagnetic wave can be written as [7]:

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} \sqrt{Q^2+U^2+V^2} \\ Q \\ U \\ V \end{bmatrix} + \begin{bmatrix} I - \sqrt{Q^2+U^2+V^2} \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{1.21}$$

which permits to define partially polarized radiation as a superposition of a completely polarized component with a completely unpolarized one. In Eq. 1.21 the first term of the sum represents the completely polarized component of intensity $\sqrt{Q^2+U^2+V^2}$, while the second term represents the completely unpolarized component of intensity $I - \sqrt{Q^2+U^2+V^2}$. Eq. 1.21 allows one also to define the degree of polarization p as the ratio of the intensity of the polarized component to the total intensity I:

$$p = \frac{I_{pol}}{I} = \frac{\sqrt{Q^2+U^2+V^2}}{I}. \tag{1.22}$$

Of particular interest is the radiation for which $V = 0$, which is referred to as linear polarization. From equations 1.12-1.15, it is possible to infer that the case $V = 0$ corresponds to an ellipticity $\epsilon = 0$, which results in a polarization vector that lies in a single direction, from which the name *linear polarization*. This particular type of polarization is indeed measurable, thanks to the instruments and techniques that will be described in the following section.

Summarizing, in this first section we have covered essential concepts related to the basics of polarimetry. The possibility of describing radiation as a superposition of fully polarized and unpolarized waves was presented, and Stokes parameters were introduced, which serve as a powerful tool for quantifying and managing radiation properties, especially those tied to polarization. These basic concepts will be widely used and partially expanded in the following sections, as they are key to describe the instrumentation and data analysis in the field of astrophysical X-ray polarimetry.

1.2 Techniques and instrumentation

This section focuses on the techniques and instrumentation that could be used to perform polarization measurements in the X-ray band. X-ray spectrum could be further smoothly subdivided according to the polarization detection technique which can be exploited. In Fig. 1.3 a schematic representation of the energy intervals with the respective detection techniques is reported [9]. It is important to highlight that the energy ranges are not rigidly defined, rather, they serve as a practical point of reference.

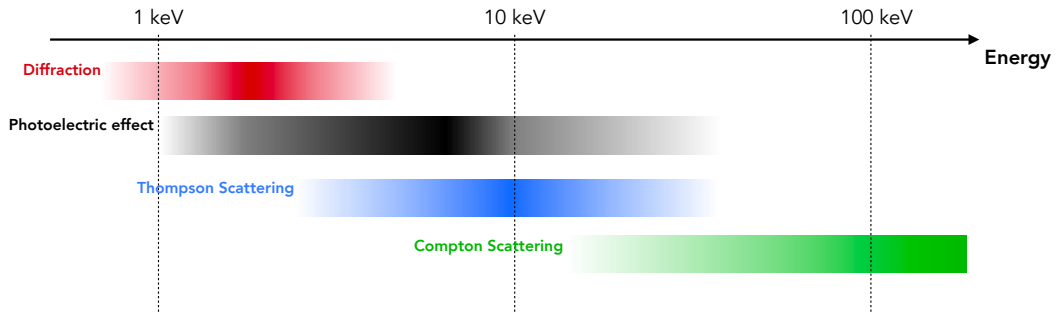

Figure 1.3: **Schematic representation of the available X-ray polarimetry techniques as a function of energy.**

In the following subsections the different polarimetry techniques will be described, and particular attention will be directed towards the photoelectric effect detection, as it is the working principle of the Gas Pixel Detector.

1.2.1 Diffraction

Bragg diffraction polarimeters are based on the use and properties of crystals, and they are mainly employed for photons with energies below a few keV. They were employed for the first detection of X-ray polarization of an extra-terrestrial X-ray source, the Crab Nebula [10].

In Fig. 1.4 the outline of the working principle of the Bragg diffraction is reported. A crystal with lattice spacing d is represented, alongside with an X-ray beam with energy E and incident angle θ. The atomic structure of the crystal causes constructive interference of the emitted light only if the energy of the incident radiation follows the so called Bragg's Law [11]:

$$E = \frac{nhc}{2d\sin(\theta)} \tag{1.23}$$

where n is the diffraction order, h is the Planck constant and c is the speed of light. If the incoming radiation is unpolarized, its polarization can be considered as comprising two components: the π component, which lies parallel to the incidence plane, and the σ component, which is perpendicular to this plane.

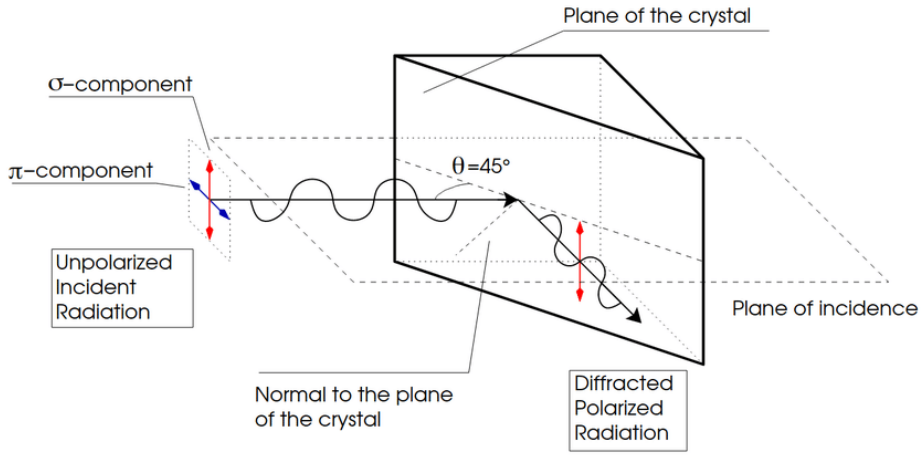

Figure 1.4: **Image describing Bragg diffraction.**

This crystal geometry preferentially reflects the X-rays whose electric vectors are perpendicular to the plane defined by the incident and reflected photons. If the diffraction angle is $\theta = 45^o$ only the σ component survives and the out-coming beam is 100% polarized, orthogonally with respect to the incidence plane.

If the crystal and the detector combination is rotated around the line of sight to a linearly polarized beam, the detected signal is sinusoidally modulated at twice the crystal rotation frequency. The degree of polarization is directly proportional to the amplitude of modulation, and the position angle of the electric vector is determined from the phase of the modulated signal. The downside of this polarization detection technique is that a specific crystal efficiently diffracts only a very narrow energy interval of the incident radiation: if this radiation is characterized by a continuous energy spectrum, as the typical astrophysical sources, most of the photons would not be diffracted at all.

Nonetheless, this technique is relevant in this thesis work, as well as in the recent development of X-ray polarimetry experiments, particularly in relation to the calibration of instrumentation. Bragg diffraction permits to generate polarized monochromatic beams out of unpolarized radiation, through the physical process just described. As a consequence, crystals find extensive utilization for beam polarization in laboratory settings. Their application involves the testing and calibration of detectors designed for X-ray polarimetry tasks, such as the Gas Pixel Detector (e.g. [12], [13]). This aspect will be discussed more in detail in Chapter 6.

1.2.2 Scattering

In scattering processes between radiation and free charged particles, there is a relation between the polarization of the incident photons and the differential cross section of the scattering process. Depending on the energy of the incident radiation, it is possible to

define two different types of interaction: *Thompson scattering* occurs when low-energy photons (compared to the rest mass of the electron) interact with charged particles, while *Compton scattering* involves higher-energy photons and a relativistic collision with a charged particle.

Thompson scattering

The Thompson scattering process consists in the interaction between an electromagnetic wave[1] and a free charged particle. After the collision, photons are scattered with a higher probability in the direction perpendicular to the polarization direction of incident photons:

$$\frac{d\sigma}{d\Omega} = r_e^2(1 - \sin^2\theta\cos^2\phi) \tag{1.24}$$

where θ is the angle between the direction of the incident and scattered photon and ϕ is the angle between the scattering plane defined by the directions and the polarization vector of the incident photon, as reported in Fig. 1.5.

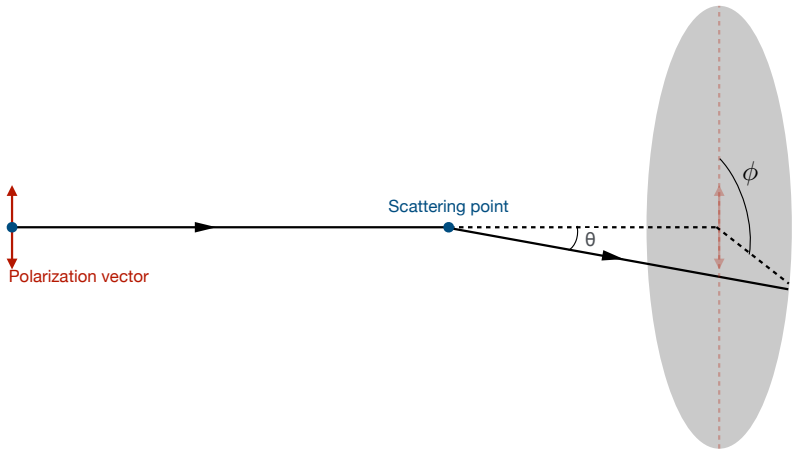

Figure 1.5: **Representation of Thompson scattering.**

Being a coherent process, i.e. the energy of the scattered radiation is the same as the incident one, there is no deposit of energy in the scatterer element, and it is difficult to separate the contribution of the signal from the background. Moreover, this process competes with the photoelectric absorption, therefore light scattering materials should be preferred for detectors: higher is the atomic number, higher is the probability for the photon to undergoes photoelectric absorption.

[1] Thompson scattering can be accurately described using classical electromagnetic theory without considering the quantum nature of particles, as $E << m_e c^2$.

POLIX is an instrument onboard of the *XPoSat* (X-ray Polarimeter Satellite), a space telescope launched on January 1st 2024, which measures the polarization of astrophysical sources in the energy band of 8-30 keV, exploiting Thompson scattering [14]. The instrument is made of a collimator, a scatterer and four X-ray proportional counter detectors. The scatterer is made of low atomic mass material which causes anisotropic Thomson scattering of incoming polarized X-rays. It is surrounded by xenon filled X-ray proportional counters as X-ray detectors, which collect the scattered X-ray photons. This mission serves as a valuable complement to IXPE, offering a broader energy range despite a reduction in sensitivity.

Compton scattering

Measurements of polarimetry through Compton scattering exploit the dependence on the polarization of the incident photon of the Klein-Nishina cross section [15]. This cross section refers to the scattering between a photon and a free electron:

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2}{2}\left(\frac{E'}{E}\right)^2\left(\frac{E'}{E} + \frac{E}{E'} - 2\sin^2\theta\cos^2\phi\right) \tag{1.25}$$

where r_0 is the classical electron radius, E is the incident photon energy, E' is the scattered photon energy, θ is the scattering angle, ϕ is the azimuthal angle, which value is zero for the direction parallel to the polarization vector of the incoming photon, and:

$$\frac{E}{E'} = \frac{1}{1 + \frac{E}{m_e c^2}(1 - \cos\theta)} \tag{1.26}$$

We notice how for the non-relativistic limit ($E << m_e c^2$), the quantity E/E' would be approximately 1 (coherent scattering), and the Klein-Nishina cross section would result in the Thompson cross section. As for the Thompson scattering, from Eq. 1.25 we observe that the photon is scattered with higher probability perpendicularly to the direction of the polarization of the incident photon. In Compton polarimeters, scatterers are typically made of light elements that are capable of scintillation, such as plastics or crystals. Also in the case of Compton scattering, as reported in Fig. 1.3, the process must compete with photo-absorption, so it is necessary to use materials that can effectively scatter light without absorbing it. An example of a Compton scattering polarimeter is the PENGUIN-M instrument, on board the CORONAS-PHOTON mission [16]. Its main purpose is to measure linear polarization of the radiation emitted by solar flares, in the energy band of 20-150 keV.

In conclusion, both Compton and Thomson scattering polarimeters play important roles in the field of polarimetry, and the choice of the technique depends on the specific requirements of the measurement being made and the properties of the scatterer being used.

1.2.3 Photoelectric effect

The photoelectric effect consists in the interaction between a photon and an atomic electron, resulting in the absorption of the photon and in the subsequent emission of the electron, which is usually called photo-electron (PE). The cross section of the photoelectric effect is calculated by summing the contributions of each electron shell of the atom. However, the contribution of the inner shells is dominant compared to the one of the outer shells when the energy is sufficient to extract electrons from them. For X-rays it is correct to assume that only the inner shells are involved.

X-ray polarimetry measurements can be significantly enhanced by taking advantage of the strong correlation between the photoelectric effect and the polarization of the incoming radiation. The correlation can be highlighted in the dependencies of the differential cross section of the photoelectric effect for a K-shell electron [15]:

$$\frac{d\sigma}{d\Omega} = Z^5 E^{-7/2} \quad \frac{\sin^2\theta \cos^2\phi}{(1+\beta\cos\theta)^4} \tag{1.27}$$

where θ is the photo-electron polar angle of emission, while ϕ is the azimuthal angle of emission with respect to the polarization direction of the incident photon. A representation of the interaction is reported in Fig. 1.6.

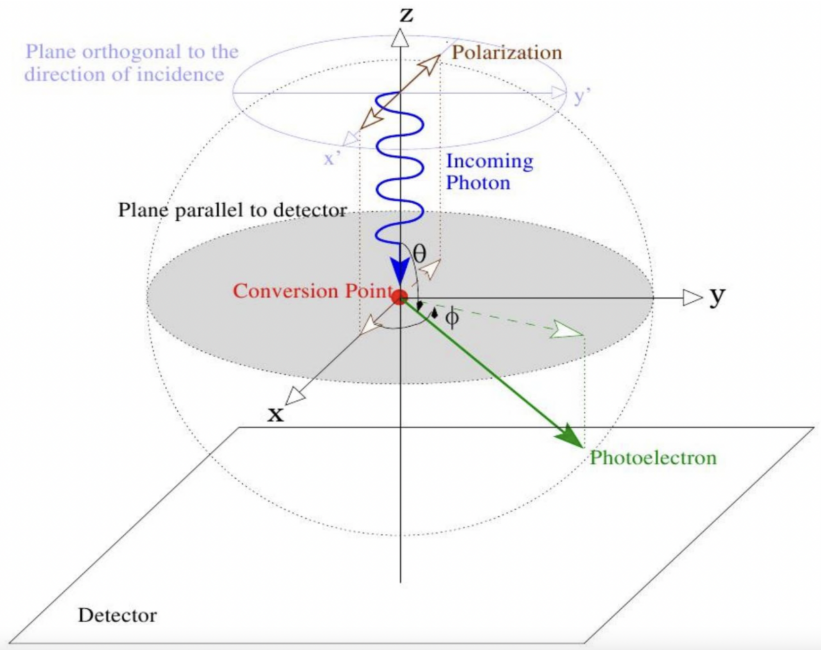

Figure 1.6: **Schematic representation of the photoelectric effect**. The quantities describing the relation between the incident photon polarization and the emitted photo-electron are highlighted. Image credits: [9]

It's important to note that Eq. 1.27 is a simplified form of the differential cross-section

and more sophisticated models may include additional terms to account for the effects of electron-electron interactions and other factors. The exact form of the differential cross-section for a K-shell PE would depend on the specifics of the atomic system being considered. However, the main aspect is that the photo-electron emission probability is modulated by a $\cos^2(\phi)$ term. This means that the PE is more likely to be emitted in a direction parallel to the electric vector of the incoming photon. Specifically, the number of emitted photo-electrons is modulated as a $\cos^2(\phi)$ function, and with an amplitude being proportional to the polarization degree of the incident radiation; while the phase angle of the distribution is related to the direction of the polarization. In Fig. 1.7 an example of the azimuthal distribution of the photo-electron emission directions for a 100% linearly polarized radiation (with the polarization direction forming a 70° angle with the reference axis) is reported.

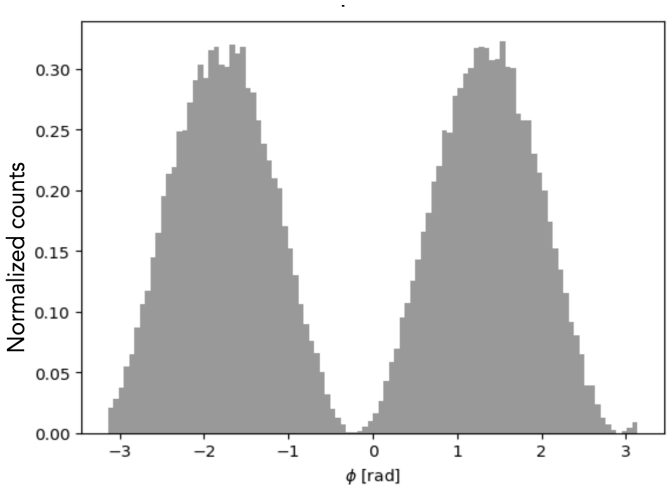

Figure 1.7: **Normalized azimuthal distribution of the emitted photo-electrons directions for a 100% linearly polarized incident radiation.** The polarization direction forms a 70° angle with the reference axis.

Fig. 1.7 also highlights how the distribution is completely modulated for 100% polarized photons, with no photo-electrons emitted in the direction perpendicular to the polarization one.

It should be mentioned that the inner shell vacancy caused by the PE emission is filled by an electron from the outer shell. The energy characterizing the transition is negligible compared to the binding energy of the inner shell. However, it results in an emission of fluorescence light, or otherwise in the emission of a secondary electron, called *Auger* electron.

As the aim of a photoelectric effect-based polarimeter is to reconstruct the emission direction of the photo-electrons, their path must be spatially resolved in the detector. At the energies of a few keV, the typical length of the PE path is about 1 μm in solids and almost 3 orders of magnitude higher in gas. The development of solid state devices

for this kind of measurement is a very active field of research, but the technology is not fully developed yet at the time of writing this thesis. For these reasons gas detectors are the state-of-the-art detectors to perform polarimetry measurements of astrophysical sources in the 1-10 keV energy band. Specifically, the main technology available to perform polarimetry by exploiting the photoelectric effect is the Gas Pixel Detector [17]. The GPD is the detector onboard of the IXPE telescope, and will be described in the following chapter.

Chapter 2

IXPE observatory

The history of the observational X-ray polarimetry in astrophysics began with two missions: Ariel-5 [18] and OSO-8 [19]. Launched respectively in 1974 and 1975, they were the first two satellites with an X-ray polarimeter onboard, based on the mechanism of Bragg diffraction. They retained their status as the sole X-ray polarimeter-equipped satellites for the following 45 years.

The OSO-8 satellite conducted an extensive observation of the Crab Nebula, providing a significant measurement of polarized radiation at energies of 2.6 keV and 5.2 keV [20]. Notably, the measured polarization exceeded what is known as the Chandrasekhar limit, offering compelling evidence for the presence of a non-thermal component in the X-ray emission, and suggesting the involvement a probable synchrotron mechanism at play [21]. Unfortunately, most of the results from the two experiments had a lower impact on the scientific community: while the Crab data generated significant interest, observations of other sources revealed that X-ray polarimetry was still challenging and intricate, despite the consistency of the data with existing theories.

In the following years, the sensitivity gap between imaging and polarimetry quickly grew, and incorporating a polarimeter on an imaging/spectral instrument would have introduced a substantial increase in complexity and demanded a significant portion of the observation time. Moreover, it would have been necessary to conduct polarimetry only on a limited number of brighter sources, which were no longer at the forefront of X-ray astronomy. All these considerations resulted in the absence of X-ray polarimetry missions in the following decades.

Since the early days of X-ray astronomy, some laboratories made attempts to develop a polarimetry technique based on the photoelectric effect [22; 23]. These works focused on detecting modulations in the non-uniform angular distribution of emitted photo-electrons. Results were not conclusive, and the idea that Bragg diffraction was the most effective technique still persisted for a long time. At the end of the 20th century, a partnership between INAF and INFN achieved the first significant detection of an effect viable for polarimetry in the range 2-8 keV [24]. This result triggered a collective effort to combine the whole ingredients needed to detect the photo-electron emission

direction within the detector volume. This effort ultimately matured into the Gas Pixel Detector.

The first astrophysical measurements with GPD were performed on the Cubesat mission *PolarLight* [25], launched in 2018. This not only established the success of the employed technology, but also allowed to achieve some relevant scientific results: Crab observations confirmed the findings of the OSO-8 mission 40 years after the first detection [26]. However, the first NASA mission dedicated to the study of the polarization of X-rays from many different astrophysical sources was the *Imaging X-ray Polarimeter Explorer* (IXPE). IXPE enabled the detection of the polarization properties of several sources, thereby substantially increasing our knowledge of the structure and emission processes of these celestial objects. In this chapter, an overview of the IXPE mission and its instrument will be provided.

2.1 IXPE mission and technical overview

The IXPE mission was announced on January 3rd, 2017 and launched on a Falcon-9 rocket from Kennedy Space Center on December 9th, 2021. It was placed into an equatorial orbit at a nominal altitude of 600 km and at a nominal inclination of 0°.

Figure 2.1: **Picture of IXPE launch from Cape Canaveral on December 9th, 2021, winner of the Astronomy Picture of the Day (APOD).** Credits: Jordan Sirokie.

IXPE is a *Small Explorer* (SMEX) mission, i.e. a mission which carries out the most highly focused space science investigations at the lowest cost. The maximum

cost for such a mission is set to be US $120 million. It was initiated by NASA, in collaboration with the Italian Space Agency (ASI). The project management, system engineering, and safety and mission assurance oversight were assigned to the Marshall Space Flight Center (MSFC). This center was also responsible for the Mirror Module Assembly (MMA) fabrication and for the Science Operations Center (SOC), which performs science operations and data processing [2].

The italian team, involving INAF, INFN and the industrial partner OHB-I, coordinated by ASI, contributed with the instruments at the heart of the mission: the polarization sensitive detection systems. The team was responsible for the design, assembly, and delivery of the Detector Units (DUs), which include the Gas Pixel Detectors, the associated readout electronics and their mechanical housing with thermal control and electrical harnessing to the satellite. It performed the characterization and calibration of the DUs, and also provided the full detector simulation and track-reconstruction framework. Since IXPE launch, italian institutions, among which the University of Torino, have been very active in the data analysis and monitoring of the telescope performance.

Communication with IXPE is established through a primary ground station located in Malindi, Kenya, supplemented by a backup station in Singapore. The frequency of data downloads depends on the specific observing program, but it consists on average of approximately 6-7 contacts per day. Data are stored at NASA's High-Energy Astrophysics Science Archive Center (HEASARC) at Goddard Space Flight Center (GSFC) and then released within 1 week to the scientific community on the online archive [27] (the embargo time was extended to 1 month only during the first 3 months of the mission). The first scientific target of IXPE was Cas-A, which data were released on January 22nd, 2022 [28].

During the scheduled first two years of observations, the collaboration assumed the responsibility of target selection. The list of selected source was publicly available. Additionally, the scientific community had the opportunity to propose *Target of Opportunities* (ToOs), which are observations requested in response to events that may occur at an unknown time. ToOs are generally used to follow-up transient phenomena, such as unpredictable astronomical events or unexpected variabilities of specific sources. Following the completion of IXPE's prime mission in January 2024, NASA initiated a General Observer (GO) program. Individuals need to submit proposals for observations with IXPE, with the possibility to obtain an exclusive-use period of up to 6 months of the data. The current observational plan covers up to the end of June 2024 and is available online [29].

In Figure 2.1, a schematic view of the IXPE observatory along with its crucial payload elements is represented. The payload consists in three identical X-ray telescopes, each equipped with a Mirror Module Assembly (MMA), featuring a polarization-sensitive Detector Unit (DU) positioned at its focus, and fixed X-ray shields and collimators. The Detector Service Unit (DSU), located beneath the upper deck of the spacecraft, manages data from the detectors. When stored before launch, the MMAs were not

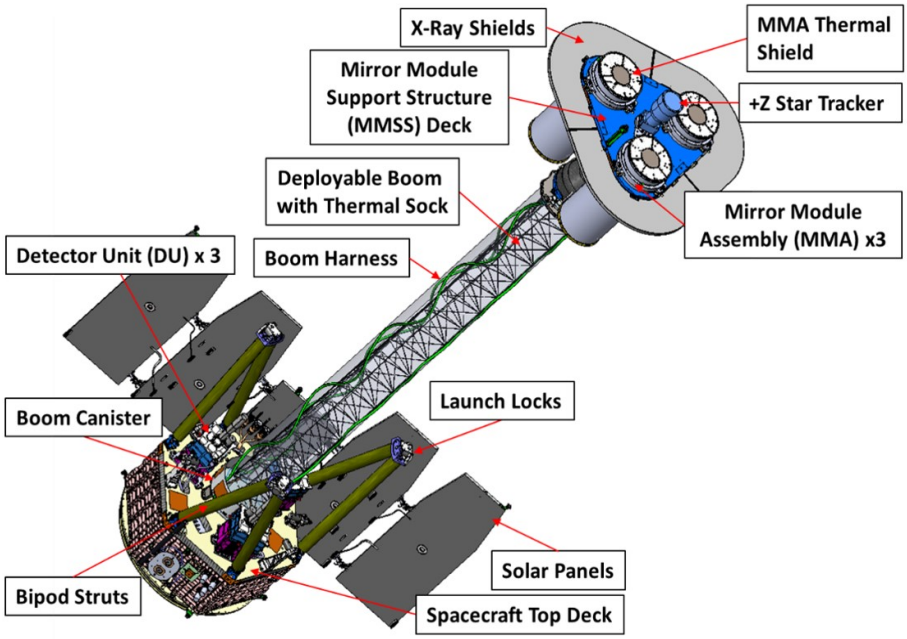

Figure 2.2: **Schematic view of the IXPE telescope, alongside with its payload elements.**

aligned with the DSU: a lightweight coilable boom was deployed in flight to establish the necessary focal length and to position each MMA above its respective detector. The following sections discuss the science payloads of the IXPE observatory, the Mirror Module Assembly and the Detector Unit. Particular attention is given to the core component of the IXPE Detector Unit, the Gas Pixel Detector, and to its events reconstruction.

2.2 Mirror Module Assembly (MMA)

The Mirror Module Assembly consists of co-aligned grazing incidence mirrors. The primary objective of the MMA is to collect X-ray photons emitted by celestial sources, and then redirect them towards X-ray detectors. This assembly consists of multiple mirror shells arranged concentrically, which are precision-crafted to reflect and focus X-rays onto the DU, enhancing the detection sensitivity. A schematic view of the working principle is reported in Fig. 2.3. IXPE telescope is equipped with three MMAs, each comprised of 24 concentrically nested mirror shells. Each mirror shell has a *Wolter 1* geometry with both the parabolic (P) and hyperbolic (H) segments [30]. The IXPE shells are uniform in length at 600 mm, with thickness ranging from 0.179 mm to 0.254 mm and diameters spanning from 165 mm to 277 mm, and have a focal length of 4 m. While usually built using pure nickel, the shells are fabricated by electroforming using a nickel and cobalt alloy, in order to achieve higher strength. A picture of the

mirrors of a single MMA is reported in Fig. 2.4. The mirrors are incorporated in the ancillary hardware to mount and protect them, and dampen their motion during launch vibration. Thermal insulation is facilitated by shields covering the entrance and exit apertures of the MMA. Shields are composed of 1.4 μm thick polyimide films coated with 50 nm of aluminum, and effectively reflect optical and thermal radiation while maintaining high transmissivity within the IXPE 2-8 keV bandpass [31].

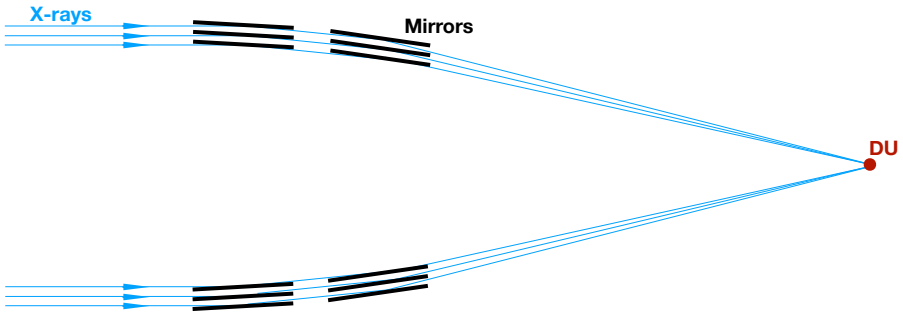

Figure 2.3: **Schematic working principle of the MMA to focus the X-rays onto the detector unit.**

Figure 2.4: **IXPE Mirror Module Assembly rear view showing the 24 nichel/cobalt nested mirrors.**

2.2.1 MMA Calibration

A precise calibration of the modules was crucial before launch, as being IXPE an *explorer* mission, there is no standard measurements to compare to. The two key parameters which characterized the MMA calibration phase, as majorly involved in the IXPE scientific results, were the **effective area** and the **Half Power Diameter**

(HPD). The calibration phase of the modules was carried out at the MSFC Stray Light Test Facility (SLTF). The facility includes an extensive beam-line of approximately 100 meters length, and a primary instrument chamber measuring 14 m in length and 3 m in diameter, where the tested MMA was located.

The **effective area** was measured using a finite-source-distance in a dedicated facility. A correction was applied to the finite-source-distance area to give the on-orbit infinite-source-distance effective area. Two cross-calibrated Silicon Drift Detectors (SDDs) were located in the instrument chamber: one was positioned to capture the flux emerging from the optics and the other one was placed before the mirror module. This configuration granted to determine the effective area of the MMAs as a function of the energy, as reported in Fig. 2.5.

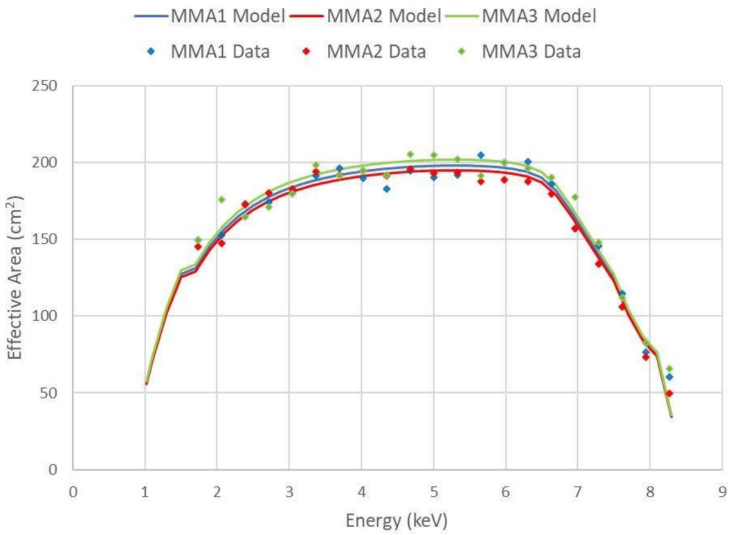

Figure 2.5: **Effective area (measured and best-fit model) for a point source at infinity as a function of energy for the three modules.** Image credits: [2]

The **Half Power Diameter** refers to the diameter which encloses half of the focused X-rays. A Charge-Coupled Device camera (CCD) was placed at the focus of the optics, in order to evaluate the HPD of the mirror modules, which was then converted to an angle. It is important to emphasize that the MMA Half Power Diameter is the dominant contribution to the total HPD of the instrument: nominal values at 2.3 keV for each module are reported in Tab. 2.1, alongside with the mean HPD of the three DUs. The Detector Units HPD is approximately three times smaller than the MMA ones.

2.3 Detector Unit (DU)

The core of the IXPE payload is the Detector Unit, where the GPD is located. Fig. 2.6 provides a schematic representation and a picture of a single DU. In addition to

	MMA 1	MMA 2	MMA 3
HPD (arcsec)	19	25	27.6
	DUs (mean value)		
HPD (arcsec)		$\lesssim 6.5$	
	Total		
HPD (arcsec)		22.2	

Table 2.1: **Half Power Diamater values for each mirror module at 2.3 keV**. The mean HPD of the three DUs and the total measured HPD of the telescope (Total) at the same energy are reported too.

accommodating the GPD, this unit also contains all the back-end electronics required for event acquisition, signal control, processing and transmission, alongside with high-voltage power supplies to operate the detectors. Each DU houses a filter and calibration wheel for supporting onboard calibrations and a passive collimator for reducing stray X-rays background.

The basic structure of a GPD consists of a gas-filled chamber, a Gas Electron Multiplier (GEM), an integrated ASIC of pixelated anodes for charge collection and signal processing, and a readout electronics system for signal digitization, control and dispatch. A schematic representation of the GPD is shown in Fig. 2.7.

Here, a brief description of the working principle is depicted, but further details of the GPD components will be provided later in this section. X-ray photons pass through a beryllium (Be) window entering the gas-filled chamber. Here, they interact with the gas through photoelectric effect, and the emitted PE ionizes the gas atoms and creates a cloud of electrons and positive ions. The primary ionization electrons are drifted toward the GEM, which amplifies the signal, that is then collected in the readout Application Specific Integrated Circuit (ASIC), made of a pixelated anode, which enables to visualize the track of the photo-electron. Two examples of PE tracks are reported in Fig. 2.8.

The detection of polarization of incident radiation is possible on a statistical basis, by reconstructing the emission directions of the photo-electron tracks measured by the detector. Eq. 1.27 highlighted that, if the incident radiation is linearly polarized, the azimuthal direction of the emitted photo-electron is not random, but preferably lies on the oscillation plane of the X-ray electric field, i.e. the polarization direction ($\frac{d\sigma}{d\omega} \propto \cos^2\phi$).

Moreover, the GPD is also a good imaging device, as from the track images it is possible to reconstruct the absorption point of the incident photon. It provides information about spectroscopy and timing as well. In general, the GPD is capable of capturing and analyzing all the properties of incident photons: direction, arrival time, energy and polarization.

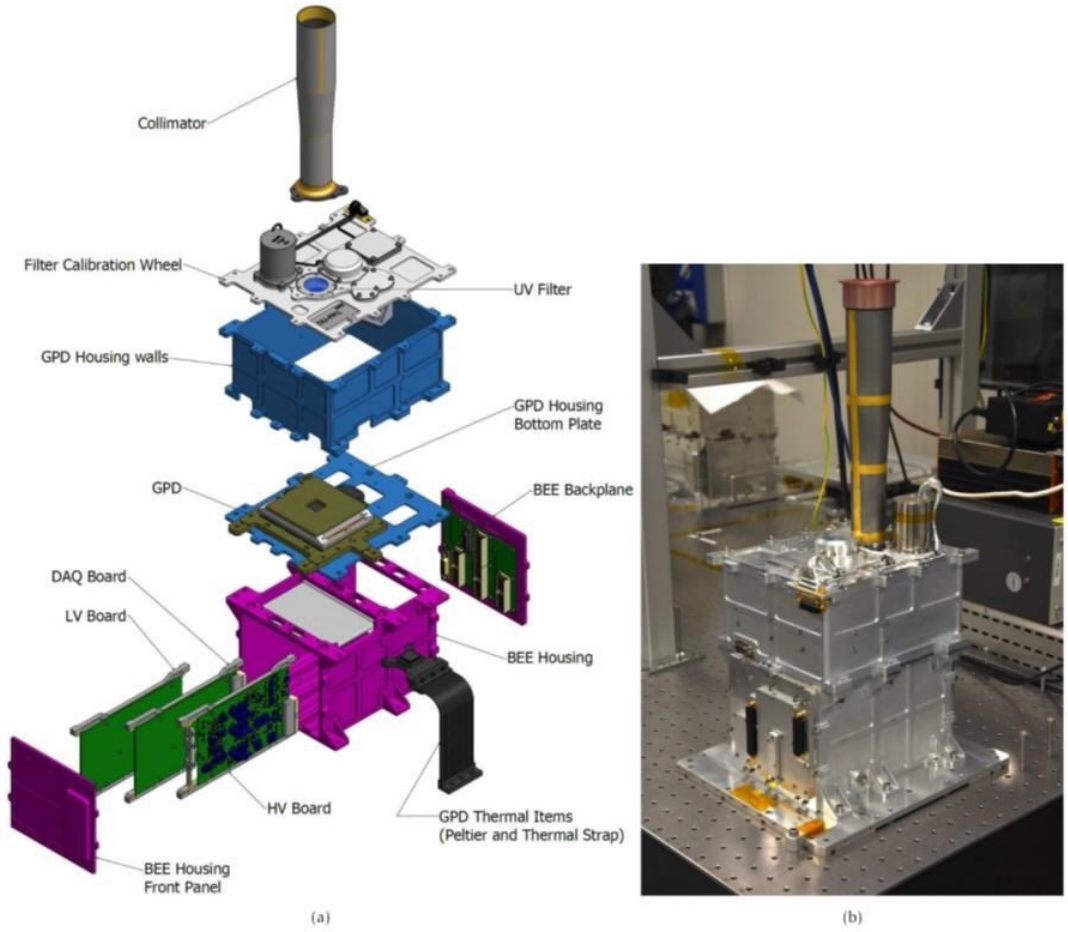

Figure 2.6: **Detector Unit onboard IXPE telescope.** *Left panel*: exploded view of a single Detector Unit. *Right panel*: picture of one DU before IXPE launch.

2.3.1 GPD design

Gas cell

The choice of the gas that serves as an absorber was a crucial phase of the GPD design, because it involved multiple factors. Quantum efficiency and stopping power were the two main quantities which were considered when selecting the gas mixture. Heavy elements are preferred for their quantum efficiency, whereas light ones offer a more favorable stopping power, resulting in longer tracks. Longer tracks can improve the accuracy of the emission direction reconstruction, which is key for polarization measurement (see sections 1.2.3, 2.3.2). Nonetheless, low-Z atoms facilitate the emission of the Auger electron, which makes more challenging reconstructing the correct emission angle. To balance these aspects, pure dimethyl-ether (DME) was chosen as an optimal compromise due to its quenching properties, low risk of accidental discharges,

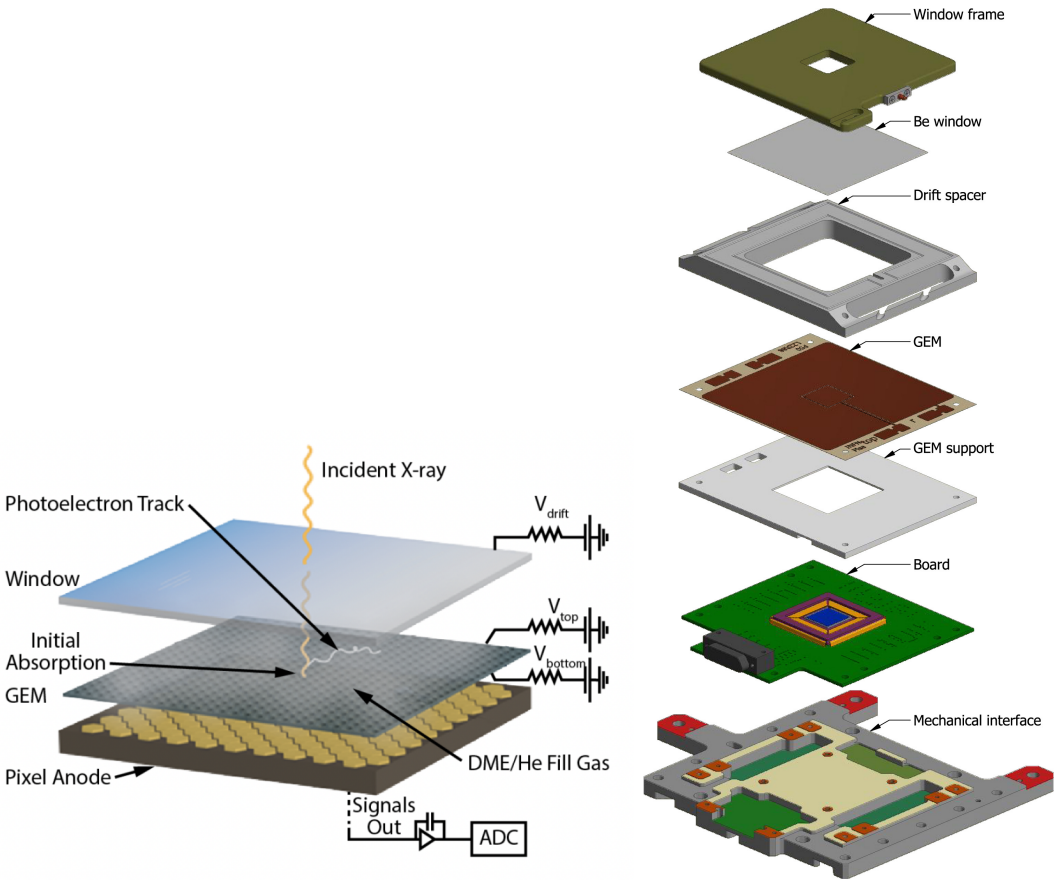

Figure 2.7: **Gas Pixel Detector.** *Left panel*: schematic and compact view of the Gas Pixel Detector and its working principle. The typical values for the voltages are: V_{drift}=-2800V, V_{top}=-870V, V_{top}=-400V, V_{ASIC}=0V (image credits: [1]). *Right panel*: exploded view of the GPD components and its mechanical interface. The colors of the mechanical interface identifies planes at the same height (image credits: [17]).

and low transverse diffusion coefficient. Moreover its use helps minimizing the blurring of tracks due to diffusion, which is a significant limit in the reconstruction of the emission direction of photo-electrons.

The performance of the GPD is largely determined by the thickness of the absorption gap. While a thicker gap may result in a higher quantum efficiency, it also leads to a longer drift length for primary ionization, causing a decrease in the reconstruction performance due to transverse diffusion of the track. The gas cell's geometry optimization is also dependent on the gas pressure. In Fig. 2.9 the polarization reconstruction performance as a function of these two physical quantities is reported, at a benchmark energy of 3 keV.

Fig. 2.9 shows that there is not a unique best choice for the two quantities, but the working point was set at an absorption thickness of 10 mm and a gas pressure of 800

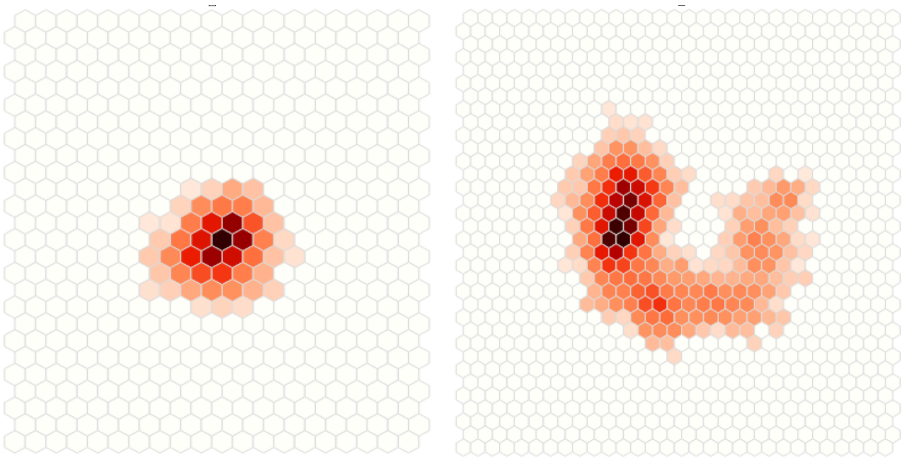

Figure 2.8: **Two examples of simulated photo-electron tracks**. The color intensity is proportional to the charge deposited by the photo-electron inside the gas cell and collected by the hexagonal anodes. On the *left panel* a track generated by a low energy photon (2.5 keV) is displayed, while on the *right panel* the track is generated by a high energy photon (7.5 keV).

mbar [17]. The GPD gas cell was filled and sealed by Oxford Instruments Technologies Oy in Espoo, Finland.

Gas Electron Multiplier (GEM)

The GEM is a specialized device used in particle and radiation detectors to amplify the signals generated by charged particles when crossing slabs of gas, and is used for the same purpose also in the GPD [32]. It consists of a thin polymer foil with a high density of small holes. As the primary electrons generated by the collisions of the PE with the gas move towards the anode, they pass through an electric field created by the voltage difference $V_{top} - V_{bottom}$ inside the holes of the GEM (see Fig. 2.7). This field causes the electrons to gain energy and causes further ionizations which lead to avalanche multiplication. Unlike typical GEM devices used in high-energy physics [33], those developed for the IXPE mission have a very fine pitch: this is necessary to preserve the photoelectron track morphology and match the sampling capabilities of the readout plane. The holes pattern of the GEM follows a hexagonal grid that corresponds to the pitch of the ASIC: the horizontal pitch is 43.3 μm, the vertical pitch is 50.0 μm, while the holes diameter is 30.0 μm. The active area of the GEM is slightly larger than the readout ASIC, by approximately 0.5 mm on all four sides, to account for potential misalignment during assembly.

During its manufacturing process, conducted by SciEnergy in Japan in collaboration with RIKEN, a laser is used to drill the holes into the dielectric substrate. This step is achieved by using a 1.8 mm-wide laser beam sweeping the GEM active surface

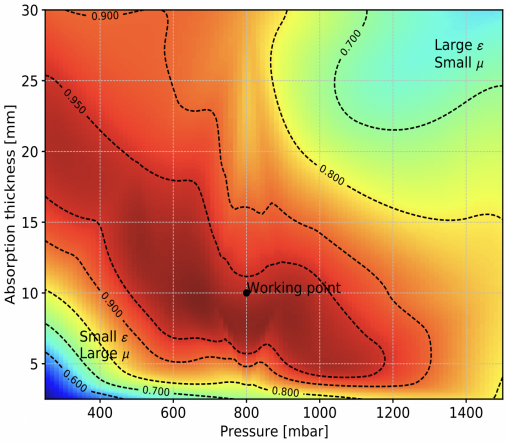

Figure 2.9: **Quality factor F_Q as a function of absorption thickness and gas pressure.** F_Q quantifies the polarization reconstruction performance, and will be discussed and better defined in Sec. 2.3.2. Image credits: [17]

multiple times to cover the entire area, with a small overlap of about 100 μm between successive passes both in the vertical and horizontal direction. This process results in the formation of 8 thin horizontal stripes and 8 thin vertical stripes (spaced by 1.8 mm) at the overlap positions of adjacent laser sweeps, which are detectable by naked eye, as shown by 2.10. This is particularly relevant as it affects the polarization measurements of the detector by causing a systematic effect known as *spurious modulation*, which holds a significant impact on the GPD data analysis and will be widely discussed in the following chapters.

Application Specific Integrated Circuit (ASIC)

The ASIC acts as a readout anode for the GPD, and was entirely developed by the INFN in Pisa [34]. It is composed of a matrix of 105 600 pixels (300 columns at 50.00 μm pitch and 352 rows at 43.30 μm pitch, matching the GEM holes) organized in a hexagonal pattern, which results in a 15 $\times$ 15 mm^2 active area. Each pixel consists in a hexagonal metal electrode, which is linked to a charge-sensitive amplifier, realized directly in the silicon substrate at the foundry, followed by the signal shaping circuit and a multiplexer for sending the analog signal to the outside ADC. The system has the capability of self-triggering, and it is equipped with a signal processing for automatic localization of the event [17]. The choice of the hexagonal shape instead of the squared one for the pixels was implemented to avoid a possible asymmetry and biases at 90°, and it gives a more uniform response. However, residual effects could still be present, with a 60°periodicity.

The ASIC is connected to a Back-End Electronics (BEE) system, which permits both the controlling of the readout chip and the generation and commanding of the voltages

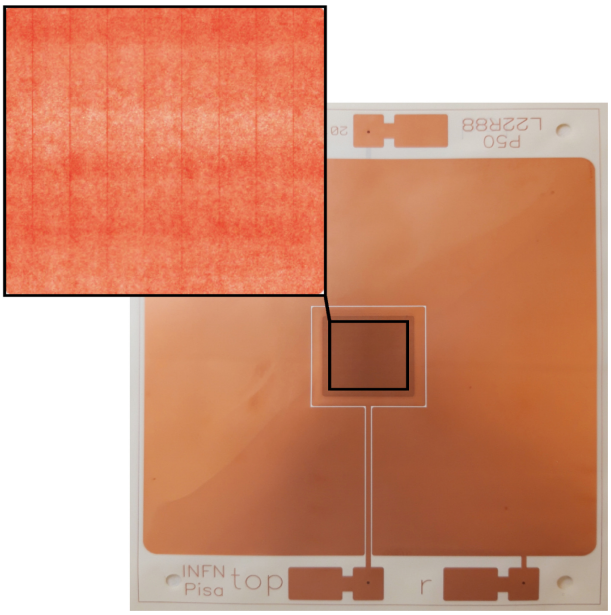

Figure 2.10: **Image detailing the fine structure of the GEM inside the detector.** The thin stripes due to the manufacturing process are visible in the zoomed panel. Image credits: [17]

needed for the detector functioning. Moreover, the BEE is responsible for the handling and acquisition of the science data.

2.3.2 Event Reconstruction

Once the GPD provides the images of the photo-electron track, the main core of the data analysis consists in the reconstruction of the physical quantities needed to infer the information about the polarization of the incident radiation. Specifically, as widely mentioned in this thesis, the emission direction reconstruction contains these information, and determines the performance of the algorithm used for the analysis.

Emission direction reconstruction

After its production inside the GPD gas cell, the photo-electron undergoes scattering interactions with the gas atoms, which are responsible for a progressive randomization of the track direction. Distinguishing between the initial and final parts of the track is crucial for extracting the features about the incident photon polarization. This can be achieved by analyzing the density distribution of the collected charge.

The energy loss of photo-electrons due to ionization in the gas cell is inversely proportional to their energy [35]:

$$-\frac{dE}{dx} \propto \frac{1}{E} \ . \tag{2.1}$$

As a consequence, the photoelectron experiences progressively larger energy losses as it travels through the gas. Eventually, this leads to the generation of the Bragg peak upon reabsorption. Consequently, an asymmetry in the charge distribution within the track arises, enabling the identification of the initial segment of the track. The randomization of the PE path together with the potential production of the Auger electron make the identification of the initial part of the track a non-trivial task.

Currently, the state-of-the-art analysis for the reconstruction of the emission direction was developed by the INFN team, and it is called moment analysis [3]. This approach is indeed based on the study of the momenta of the charge distribution, in order to reconstruct the track properties, and in particular the PE emission direction. The analysis can be divided in four main steps, represented schematically in Fig. 2.11:

1. Firstly, the barycenter of the charge distribution is calculated, as:

$$x_{\mathrm{b}} = \frac{\sum_{\mathrm{i}} q_{\mathrm{i}} x_{\mathrm{i}}}{\sum_{\mathrm{i}} q_{\mathrm{i}}} \quad y_{\mathrm{b}} = \frac{\sum_{\mathrm{i}} q_{\mathrm{i}} y_{\mathrm{i}}}{\sum_{\mathrm{i}} q_{\mathrm{i}}} \tag{2.2}$$

where q_i is the charge collected in each i_{th} pixel, and (x_i, y_i) are respectively the horizontal and vertical coordinates of each pixel center on the readout plane. The barycenter is used to evaluate the second moment of charge distribution $M_2(\phi)$, where ϕ is the angle with respect to a reference axis, which is conventionally the x axis:

$$M_2(\phi) = \frac{\sum_{\mathrm{i}} q_{\mathrm{i}} [(x_{\mathrm{i}} - x_{\mathrm{b}}) \cos(\phi) + (y_{\mathrm{i}} - y_{\mathrm{b}}) \sin(\phi)]^2}{\sum_{\mathrm{i}} q_{\mathrm{i}}} \ . \tag{2.3}$$

The direction of maximum and minimum elongation of the track correspond to the maximum and minimum values of $\mathrm{M}_2(\phi)$. These can be obtained by setting $\frac{\mathrm{dM}_2}{\mathrm{d}\phi} = 0$. (In Fig. 2.11, the blue dashed line represents the maximum elongation direction of the track.)

2. Secondly, the third moment of the charge distribution is evaluated, as it is an indicator of the skewness of the track, and it is crucial for identifying the initial part of the track.

$$M_3(\phi) = \frac{\sum_{\mathrm{i}} q_{\mathrm{i}} [(x_{\mathrm{i}} - x_{\mathrm{b}}) \cos(\phi) + (y_{\mathrm{i}} - y_{\mathrm{b}}) \sin(\phi)]^3}{\sum_{\mathrm{i}} q_{\mathrm{i}}} \ . \tag{2.4}$$

$\mathrm{M}_3(\phi)$ is used to determine an horseshoe region which encompasses the initial part of the track (see Fig. 2.11, second panel).

3. Indicating the pixels in the horseshoe region with the j subscript, the barycenter of the initial part of the track is calculated as:

$$x_{\mathrm{b,init}} = \frac{\sum_{\mathrm{j}} q_{\mathrm{j}} x_{\mathrm{j}}}{\sum_{\mathrm{j}} q_{\mathrm{j}}} \quad y_{\mathrm{b,init}} = \frac{\sum_{\mathrm{j}} q_{\mathrm{j}} y_{\mathrm{j}}}{\sum_{\mathrm{j}} q_{\mathrm{j}}} \ . \tag{2.5}$$

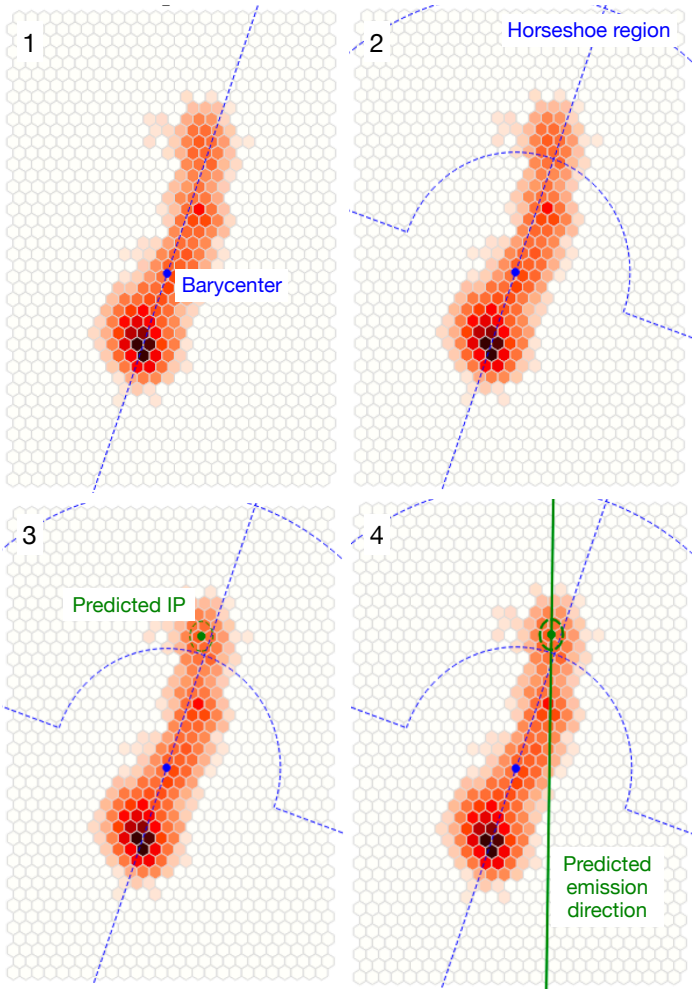

Figure 2.11: **Reconstruction of a photo-electron track with moment analysis.** The four panels schematically represent the steps required by the analytic algorithm to reconstruct the photo-electron emission direction (see text for description).

This new point is used to evaluate weights for each pixel in the whole track, as:

$$w_{\mathrm{i}} = \mathrm{e}^{-\frac{\mathrm{d}_{\mathrm{b,i}}}{\mathrm{d}_{\mathrm{s}}}} \tag{2.6}$$

where $d_{\mathrm{b,i}}$ is the distance between each track pixel and the position of the barycenter of the initial part of the track, and d_{s} is a scale parameter. The impact point is then defined as (green dot in Fig. 2.11):

$$x_{\mathrm{IP}} = \frac{\sum_{\mathrm{i}} w_{\mathrm{i}} x_{\mathrm{i}}}{\sum_{\mathrm{i}} w_{\mathrm{i}}} \quad y_{\mathrm{IP}} = \frac{\sum_{\mathrm{i}} w_{\mathrm{i}} y_{\mathrm{i}}}{\sum_{\mathrm{i}} w_{\mathrm{i}}} \ . \tag{2.7}$$

4. Finally, the second moment of the charge distribution is again calculated, this time with respect to the location of the impact point $(x_{\mathrm{IP}}, y_{\mathrm{IP}})$ and with the weighted pixels:

$$M_2'(\phi) = \frac{\sum_{\mathrm{i}} w_{\mathrm{i}}[(x_{\mathrm{i}} - x_{\mathrm{IP}})\cos(\phi) + (y_{\mathrm{i}} - y_{\mathrm{IP}})\sin(\phi)]^2}{\sum_{\mathrm{i}} w_{\mathrm{i}}} \tag{2.8}$$

and the emission direction is obtained by evaluating the angle ϕ which maximizes $M_2'(\phi)$ (green solid line in Fig. 2.11).

From emission angles to polarization properties

In Sec. 1.1 the Stokes parameters were introduced as a powerful mathematical instrument to fully describe and characterize the polarization state of the radiation. For monochromatic completely polarized radiation holds Eq. 1.11, where the values of the Stokes parameters are reported in Eqs. 1.12-1.15. It was also specified that for a partially polarized wave, thanks to the additive property of the Stokes parameter, it is possible to define the polarization degree as:

$$p = \frac{\sqrt{Q^2 + U^2 + V^2}}{I}. \tag{2.9}$$

We also discussed that the state-of-the-art polarimetry techniques measure and quantify linear polarization only. From now on, the polarization of the radiation will be assumed to be linear ($V = 0$). Eq. 1.11 simplifies to:

$$I^2 = U^2 + Q^2 \tag{2.10}$$

where Q and U are defined as:

$$Q = I\cos(2\Phi) \tag{2.11}$$

$$U = I\sin(2\Phi) \tag{2.12}$$

where Φ is the angle between the polarization vector and the x-axis. The polarization degree can be consequently defined as:

$$p_l = \frac{\sqrt{Q^2 + U^2}}{I} \tag{2.13}$$

and the polarization angle as:

$$\tan(2\Phi) = \frac{U}{Q}. \tag{2.14}$$

Using this information about the Stokes parameters, it is possible to derive the polarization properties of the incident radiation starting from the PE reconstructed emission angles. Assuming that we collected the radiation from a polarized source with a GPD, and that we reconstructed the tracks emission directions with the standard moment analysis, we would dispose of a data set made up of N angles $\{\phi_n\}$ with $n = 1, ..., N$, which represent the emission directions of the PE tracks with respect to the x-axis.

Starting from the cross section for the photoelectric effect (Eq. 1.27), and by accounting for the polarization degree of the incident radiation and the uncertainties in the reconstruction of the emission direction, the equation that describes the distribution of the $\{\phi_n\}$ angles can be written as [36]:

$$f(\Phi) = \frac{1}{2\pi}\{1 + p_0 \mu \cos[2(\Phi - \Phi_0)]\} \tag{2.15}$$

where Φ_0 is the phase of the distribution, p_0 is the true polarization degree of the incident radiation and μ is called **modulation factor**. The modulation factor is defined as the response to a 100% polarized radiation, i.e. the reconstructed amplitude of the azimuthal modulation measured for a 100% polarized beam ($0 \leq \mu \leq 1$). It can be used to estimate the sensitivity of the instrument to detect the polarization properties of the radiation. Its value depends both on the intrinsic properties of the detector and on the quality of the reconstruction algorithm. An ideal polarimeter followed by a perfect reconstruction algorithm could recover the whole polarization fraction of the incident beam ($\mu = 1$).

For each reconstructed emission angle ϕ_n, we define the Stokes parameter as [8]:

$$i_n = 1 \tag{2.16}$$

$$q_n = \cos 2\phi_n \tag{2.17}$$

$$u_n = \sin 2\phi_n. \tag{2.18}$$

For the additive property of the Stokes parameter, the same parameters for the whole data set can be defined as:

$$I = \sum_n i_n = N \tag{2.19}$$

$$Q = \sum_n q_n \tag{2.20}$$

$$U = \sum_n u_n \tag{2.21}$$

and their normalized values as:

$$\widetilde{Q} = \frac{Q}{I} \tag{2.22}$$

$$\widetilde{U} = \frac{U}{I} \tag{2.23}$$

For a large number of events $\widetilde{Q}$ and $\widetilde{U}$ are independent, Gaussian distributed variables with equal standard deviations. From Eq. 2.15 we evaluate the expected values for $\widetilde{Q}$ and $\widetilde{U}$ [8]:

$$<\widetilde{Q}> = \frac{1}{2} p_0 \mu \cos 2\Phi_0 \tag{2.24}$$

$$<\widetilde{U}> = \frac{1}{2} p_0 \mu \sin 2\Phi_0 \tag{2.25}$$

from which we find:

$$<\widetilde{Q}>^2 + <\widetilde{U}>^2 = \frac{p_0^2 \mu^2}{4} \tag{2.26}$$

$$\frac{<\widetilde{Q}>}{<\widetilde{U}>} = \tan 2\Phi_0. \tag{2.27}$$

From these expectation values we can reconstruct the polarization degree of the incident beam and its polarization vector direction as:

$$p_r = \frac{2}{\mu} \sqrt{\widetilde{Q}^2 + \widetilde{U}^2} \tag{2.28}$$

$$\Phi_r = \frac{1}{2} \arctan \frac{\widetilde{U}}{\widetilde{Q}} \tag{2.29}$$

Once the modulation factor of the combination of the detector and the reconstruction algorithm is known, these equations can be used to find the polarization properties of the incident radiation starting from the measurements of the PE emission angles. The reconstructed polarization fraction p_r (Eq. 2.28) exhibits an additional $2/\mu$ factor compared to the actual polarization fraction p_l (Eq. 2.13). The observed Stokes parameters are obtained from the distribution of emission angles of photo-electrons, exhibiting a sinusoidal distribution with a 180° period, where minima and maxima align to the polarization angle of the incident photons. Consequently, the sinusoidal behaviour leads to a reduction of the Stokes parameters derived from them by a factor of 2 compared to the true Stokes parameters of the incident photons. Additionally, the observed polarization fraction is reduced by a factor μ due to the modulation factor of the system [detector + reconstruction algorithm]. An estimate of the uncertainties for the reconstructed polarization fraction and angle can be obtained by propagating errors [8]:

$$\sigma(p_r) = \sqrt{\frac{2 - p_r^2 \mu^2}{(N-1)\mu^2}} \tag{2.30}$$

$$\sigma(\Phi_r) = \frac{1}{p_r \mu \sqrt{2(N-1)}} \tag{2.31}$$

Modulation Factor recovery

Eqs. 2.28-2.29 can be used to evaluate the polarization properties of the incident radiation, once the modulation factor of the combination of the instrument and algorithm is known. Firstly, the performance of the emission direction reconstruction algorithm is evaluated using Monte Carlo simulations of the detector response to 100% polarized radiation. Starting from Eq. 2.28 the modulation factor can be evaluated as:

$$\mu = 2\sqrt{\widetilde{Q}^2 + \widetilde{U}^2} \tag{2.32}$$

Once the modulation factor is evaluated on the whole energy range using Monte Carlo simulations, a validation phase of the algorithm is needed. The same process is repeated in laboratory using 100% polarized beams, in order to establish that the response of the algorithm to MC simulations is the same as to real data. These two steps will be described in details in Chapter 5 and 6.

The modulation factor is a useful figure of merit to evaluate, as well as to compare, the performance of reconstruction algorithms. However, if we want to take into account the properties of the observed sources too, we need to introduce a more complete figure of merit. Starting from the modulation factor value, the sensitivity of a polarimeter can be quantified defining the **Minimum Detectable Polarization (MDP)**. The MDP is the minimum degree of polarization of the incident radiation which can be detected by the polarimeter at a certain confidence level. Conventionally the 99% confidence level is used, and the MDP is given by [37]:

$$\mathrm{MDP}_{99} = \frac{4.29}{\mu \mathrm{R_S}} \sqrt{\frac{\mathrm{R_S} + \mathrm{R_B}}{\mathrm{T}}} \tag{2.33}$$

where R_S is the source rate, R_B is the background rate and T is the observation time. This figure of merit is particularly useful when planning observations, as it can be used to evaluate the observational time required to detect the expected polarization degree. If the background rate is negligible compared to the source, the MDP_{99} can be evaluated as:

$$\mathrm{MDP}_{99} \simeq \frac{4.29}{\mu\sqrt{\mathrm{N}}} \tag{2.34}$$

As an example, if an observation is expected to reach a minimum detectable polarization of 1% (we should expect a higher signal from the source), and for our reconstruction setup μ=0.5, we would need a number of counts from the source equal to:

$$N \simeq \left(\frac{4.29}{0.5 \cdot 0.01}\right)^2 = 720\,000 \tag{2.35}$$

The same quantity can also be expressed in terms of the detector quantum efficiency ϵ, which takes into account both the Be window transparency and the DME absorption efficiency, and the source flux F as [11]:

$$\mathrm{MDP}_{99} \propto \frac{1}{\mu\sqrt{\epsilon}\sqrt{\mathrm{F}}} \tag{2.36}$$

where the quantity $\mu\sqrt{\epsilon} = Q$ is defined as *quality factor*, and was previously mentioned in describing Fig. 2.9.

2.3.3 DU Calibration

As for the Mirror Module Assembly, we report here a brief description of the calibration phase of the GPD. Specifically, its energy resolution and modulation factor are of particular interest. Importantly, the spatial resolution of the GPD has already been documented in Table 2.1.

Ground calibration was performed at the laboratory of INAF-IAPS in Rome [13]. This calibration procedure consisted in exposing each Detector Unit to both polarized and unpolarized monochromatic beams. With this approach it was possible to evaluate the response of the detector to source beams with known spectral and polarization properties. Moreover, radioactive sources are onboard IXPE simulating the same conditions of astrophysical observations to investigate the detector response over time. The specific sources and beam configurations employed in the ground calibration will be detailed in Chapter 6.

As widely discussed in Sec. 2.3.2, the GPD is mainly a polarization detector, but the PE track preserves the information about its energy too. The charge deposited in the gas inside the GPD cell is indeed proportional to the energy of the incident photon.

In Fig. 2.12, the distribution of total ADC counts (referred to as Pulse Height Amplitude, PHA) for tracks generated by a simulated 3 keV energy beam is displayed. The majority of events are included in a Gaussian distribution centered around the PHA value corresponding to the energy of 3 keV. By evaluating the DUs response to different energy beams, it was possible to determine a correlation between the energy of incident photons and the total ADC counts of PE tracks. The energy resolution is estimated by fitting with a Gaussian the PHA peak and dividing the measured Full Width Half Maximum (FWHM) by its mean. Results for the three DUs are reported in the left panel of Fig. 2.13.

However, as clearly shown in Fig. 2.12, there is a tail of events with lower PHA values, indicating occurrences where photon conversions happened outside the gas cell or in close proximity to its edges. For these events, not all of the photoelectron energy was contained within the gas cell, resulting in ADC counts that do not correspond to the entire energy of the incoming photon. These events do not carry the correct information about the incident photon polarization properties, but they are not easily distinguishable in IXPE data: there are several works and ongoing studies aimed at weighting or excluding these events from the analysis [38; 39].

The reconstruction of tracks produced by lower energy photons is generally more challenging. This observation becomes even clearer in the right panel of Fig. 2.13, where

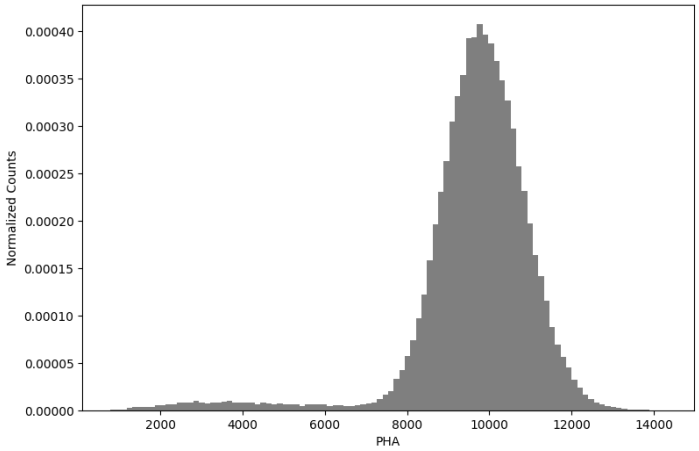

Figure 2.12: **Pulse Height Amplitude (PHA) distribution for a simulated 3 keV energy beam.**

the modulation factor is presented as a function of energy. At 8 keV, approximately 50% of the total modulation is reconstructed, whereas at 2 keV, only ∼15%. This phenomenon is attributed to the characteristics of lower-energy tracks, which are generally smaller and less elongated than higher-energy tracks. Two examples of two simulated low-energy tracks (2.5 keV) are reported in Fig. 2.14. This presents a challenge in reconstructing the emission direction of photo-electrons, as the absence of pronounced track skewness complicates the identification of the initial part of the track and of the photon impact point. Consequently, this difficulty leads to a decrease of the modulation factor values at lower energies.

2.4 IXPE systematics

In this section two IXPE systematics will be discussed, as majorly involved in the context of this work. The former (A) has been detected and characterized during the ground calibration phase, whereas the latter (B) after the telescope launch.

A) Spurious modulation

During the ground calibration of the GPD, an unexpected spurious instrumental signal, referred to as *spurious modulation*, was detected. When exposed to unpolarized radiation, an ideal polarimeter would register a minimal modulation amplitude, due to the Poisson distribution of photo-electrons emission directions, that is compatible with a null amplitude and that decreases with increasing counts. However, this expectation is not observed in the case of the IXPE GPD [40]. Especially at lower energies, in fact, a non-zero modulation of the emission angles distribution is detected for unpolarized monochromatic beams.

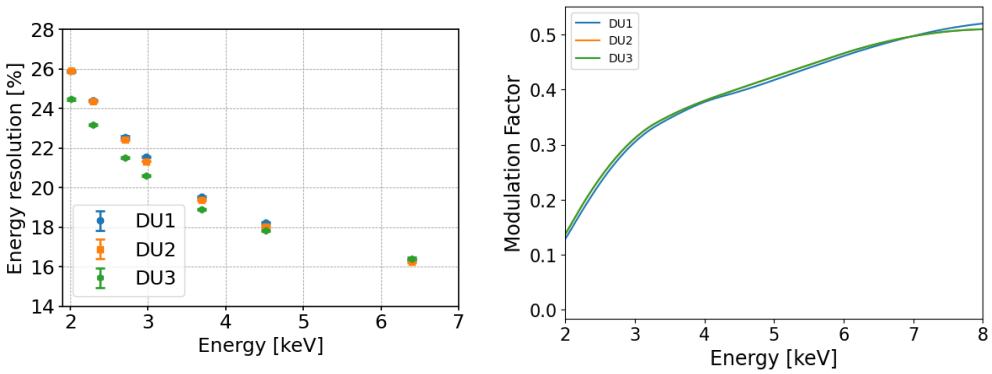

Figure 2.13: **Detector Unit calibration results.** *Left panel*: energy resolution of the three DUs for monochromatic beams in the 2-8 keV energy range. *Right panel*: modulation factor as a function of energy for the three DUs.

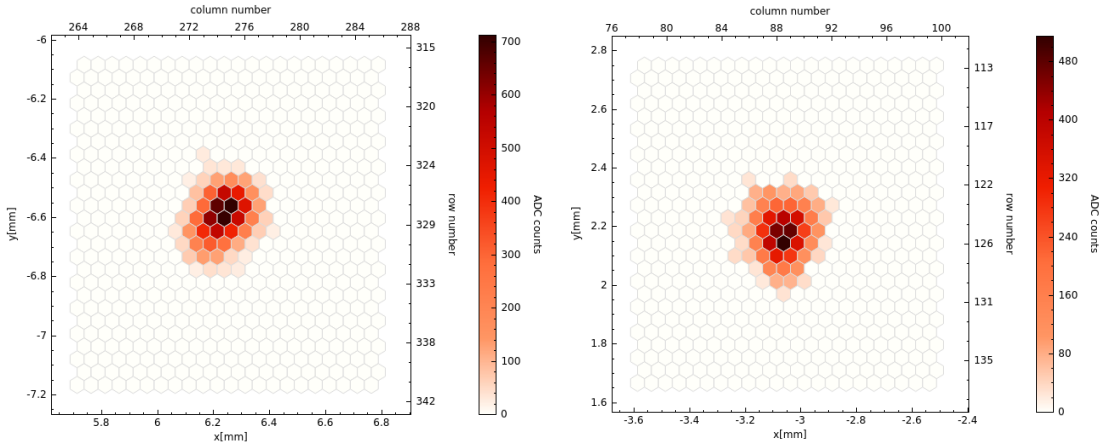

Figure 2.14: **Examples of PE tracks generated by low-energy photons**. These tracks are generated by simulating the GPD response to a 2.5 keV energy beam.

The source of this systematic effect has not been conclusively identified yet, but a reliable explanation has been suggested by examining the spatial dependency of the residual polarization on the GPD map. In Fig. 2.15 the residual normalized Q and U values for an unpolarized 2.7 keV energy beam as a function of the impact point position on the GPD plane are reported in the *upper panes*. The spatial variation observed in the residual modulation suggests a correlation with the presence of the Gas Electron Multiplier (GEM) in the detector. The features in the Q/I map[1] reveal a distinct pattern that mimics the physical characteristics of the GEM, which are the consequence of its production process, highlighted in the *lower panel* of Fig. 2.15 and described

[1]Q/I values highlight the polarization patterns along the horizontal and vertical directions, while U/I values along the 45° directions with respect to the reference axes.

in Sec. 2.3.1. Nevertheless, a comprehensive understanding of the relation between these physical features and the spurious modulation is still pending clarification. It is important to highlight that the spurious modulation essentially mimics an extra $\cos^2 \phi$ component, sharing the same frequency as the modulation induced by an effective source polarization. Due to this characteristic, the presence of spurious modulation may appear hidden within the polarized modulation curve, yet it is indeed present, and could amplify or reduce the observed modulation.

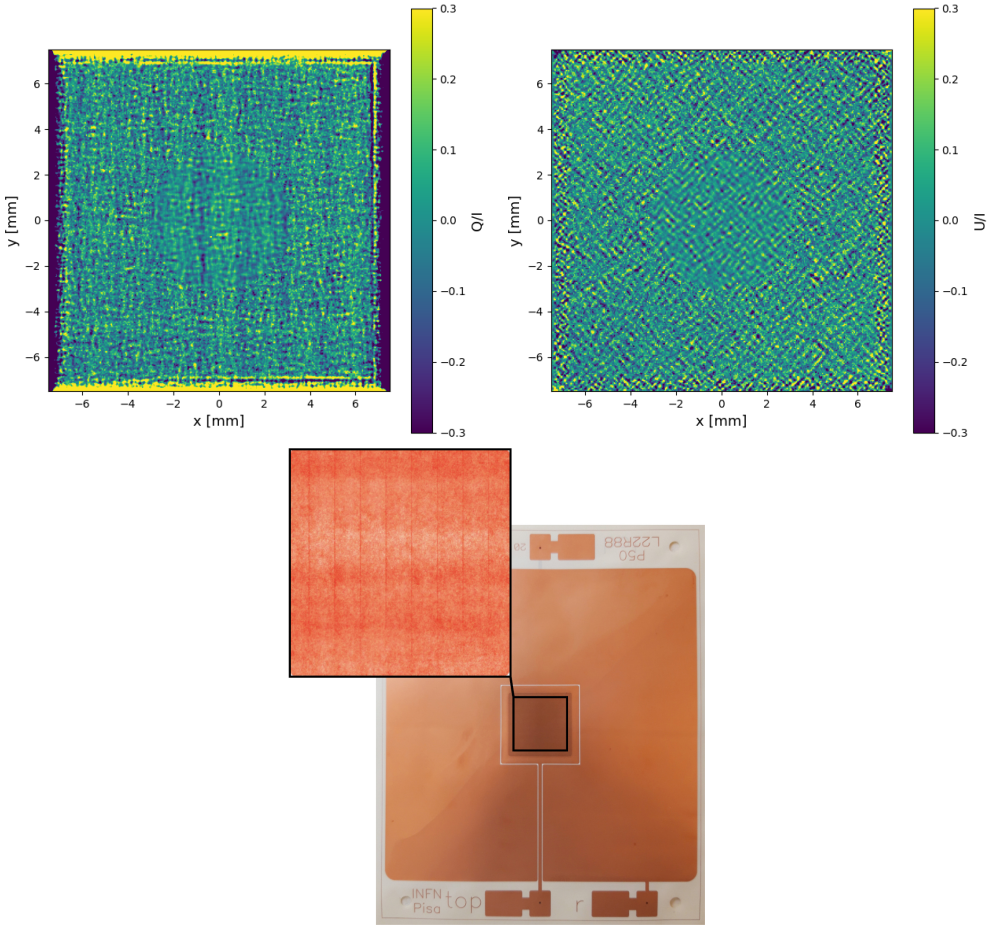

Figure 2.15: **Binned residual normalized Stokes parameters**. Q/I (*left panel*) and U/I (*right panel*) are achieved by detecting photons of a 2.7 keV unpolarized energy beam, which are binned according to the impact point position on the GPD plane. The *lower panel* reports again the picture of the GEM highlighting the structures caused by its manufacturing process.

This systematic was characterized and quantified during the IXPE ground calibration campaign, and all the analyses of the astrophysical sources results that will be discussed in the following section are corrected in order to account for this effect. This particular calibration phase will be described with more details in Chapter 6.

B) Polarization leakage

An additional unexpected systematic was detected by the collaboration soon after IXPE launch: a radial polarization pattern was observed for unpolarized calibration sources located onboard. Subsequently, a similar pattern was identified in unpolarized celestial point sources as well. This effect is caused by a failure in reconstructing the correct photon impact point, and was denominated *polarization leakage* [41].

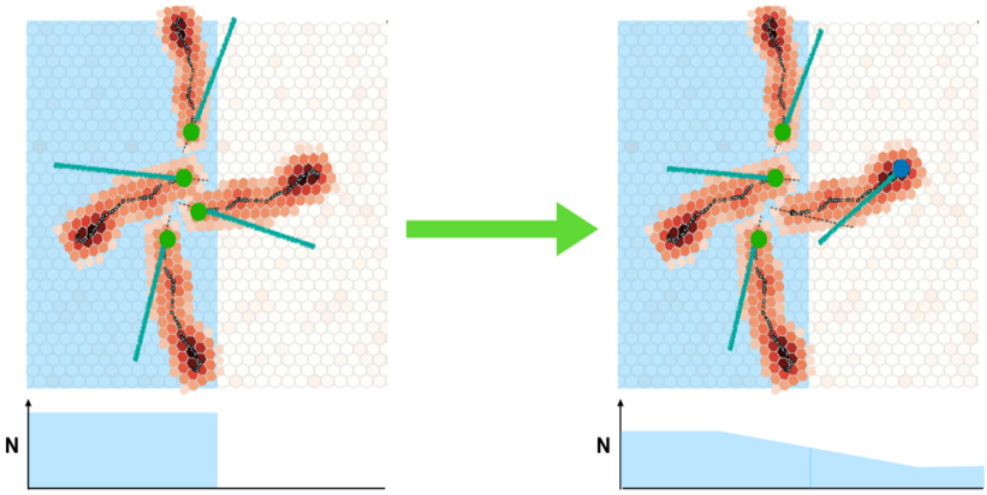

Figure 2.16: **Illustration of the polarization leakage effect for an extended unpolarized source**. The *light blue region* represents a source, alongside with four events detected on the edge of the same source. On the *left panel*, the true IPs and emission directions of the PEs are reported, each event is placed within the boundary, with none extending beyond it, as shown in the *lower panel* that displays event counts. On the *right panel*, the absorption point of the track that extends beyond the source's edge has been wrongly reconstructed at the track's end (*blue dot*). Consequently, there are fewer events reconstructed within the edge, and more events placed outside of it, as indicated in the event count represented in the *lower panel*. Image credits: [41]

Fig. 2.16 provides a simplified representation of the underlying reason behind this systematic effect. The light blue region represents the GPD portion involved by a celestial source. Four events, located near the edge of the source image, are depicted with their respective tracks. The true absorption points and emission directions are denoted by green dots and segments in the left panel, whereas the reconstructed are reported in the right panel. For the photons coming from the edge of the source there is the chance of placing the reconstructed impact point of the track outside the actual source boundary, alongside the axis of the track elongation, as represented for a single event in the right panel of Fig. 2.16. This will result in both a blurring of the source's

edges and a polarization pattern perpendicular to its boundary. Even when a source lacks a pronounced sharp edge, but instead exhibits a gradual change in flux, a gradient of polarization could emerge. This leads to the formation of a polarized halo encircling any source, even in cases where intrinsic polarization is null.

For IXPE observations, leakage becomes critical in the evaluation of the polarization properties of extended sources, where characterizing and removing the effects due to polarization leakage is extremely challenging. Its contribution mostly depends on the geometry of the source and its intrinsic polarization, and currently there is no available comprehensively effective method to correct it. Additionally quantitative information about the polarization leakage and its repercussion for point and extended sources observations will be provided in Chapter 5.

Chapter 3

IXPE science

IXPE mission aims to expand our understanding of high-energy astrophysical processes and sources, aligning with NASA's primary astrophysics objective: "Discover how the universe works". Through its unique capability of performing polarimetric measurement in the X-ray band the mission adresses key science objectives: determining the radiative processes and detailed properties of specific cosmic X-ray sources or categories and exploring general relativistic and quantum effects in extreme environments. Polarization indeed provides a distinctive method for investigating physical quantities as ordered magnetic fields, aspheric matter distributions, or general relativistic coupling to black-hole spin, that are otherwise challenging to measure. Therefore, IXPE complements existing investigations in high-energy astrophysics by introducing a crucial and relatively unexplored dimension to the parameter space.

The detailed science objectives for IXPE encompass a broad range of X-ray sources. In the context of **radio-loud AGN**, IXPE goal is focused on the study of relativistic jets, with special attention to the blazar subclass, where these jets either contribute significantly or dominate emission across all frequencies [42]. Highly polarized synchrotron emission measured in radio was expected to possibly extend into the X-ray band, originating from high-energy electrons near the jet base or downstream of the shock. IXPE aimed to distinguish between these two models by detecting the magnetic field alignments with respect to the orientation of the jet axis [43; 44]. This study extends to the phenomenon of polarization angle rotations, observed in certain AGNs in the optical band, and possibly linked to helical magnetic fields or moving emission regions [45]. Collaborative efforts involve optical monitoring of blazars observed by IXPE, by exploring the correlation between optical and X-ray emissions. In other scenarios, where Compton scattering dominates the X-ray band, IXPE aimed at differentiating between synchrotron self-Compton (SSC) and external Compton (EC) mechanisms [46].

Pulsar Wind Nebulae (PWN) emit synchrotron X-rays as a result of the interaction between an ultrarelativistic pulsar wind and the surrounding medium, and are a relevant target of IXPE observations. X-ray polarimetric imaging allows for the ex-

ploration of the magnetic-field configuration in the proximity of the termination shock and other structures within the PWNe. Moreover, X-ray polarization mappings facilitate a distinctive exploration of the acceleration and cooling mechanisms at play. A comparative analysis with optical data aims to discern whether the particles emitting optical and X-ray radiation constitute distinct populations [47].

Supernova remnants (SNRs) represent special targets for IXPE to study the magnetic field properties within the regions of synchrotron emission. X-ray spectroscopy can locate these non-thermal regions, but only X-ray polarimetric imaging can detect the magnetic field direction and uniformity. One of the primary goals is to investigate whether the X-ray polarization matches the radio observations, in which the polarization direction changes according to the SNR age [48].

X-ray binaries, compact objects (black holes or neutron stars) accreting material from a companion star, exhibit robust X-ray outbursts and relativistic jets. IXPE aims at unraveling the specific emission sites, be it the accretion disk, corona, or jet. In systems where the presence of a stellar mass black hole is likely, cyclic transitions between two primary X-ray spectral states is observed [49]: the luminous soft state characterized by a thermal accretion disk and a steep power-law tail, and the hard state dominated by a predominantly flat power-law spectrum. Comprehending the power-law emission involves considerations such as Comptonization of thermal disk emission in a hot corona [50], synchrotron or synchrotron self-Compton emission from the relativistic jet [51], or thermal disk radiation up-scattered in the jet [52]. Each of these scenarios was expected to produce polarized radiation, with distinct polarization signatures [53]. IXPE, through its observations across different spectral states, is able to delineate the specific origins and processes governing X-ray production in these systems.

Magnetars are young and highly magnetized neutron stars. Their magnetic field can extend up to $\sim 10^{15}$ G. The *magnetar model* [54] predicted a persistent thermal emission in IXPE band produced by low-level seismic heating of the interior and up-scattered in the external magnetosphere, according to the Resonant Compton Scattering (RCS) scenario [55]. The polarization is expected to be phase and energy dependant [56]. Surface emission was predicted to exhibit linear polarization along two distinct modes, ordinary (O) and extraordinary (X), with the polarization vector aligned either parallel or perpendicular to the plane defined by the photon trajectory and the local magnetic field. The polarization degree heavily relies on the physical conditions of the outermost layers of the star. IXPE aims to measure the non-linear QED prediction of *vacuum polarization in a strong magnetic field* through phase-resolved observations [57].

3.1 IXPE results

In this section some of the scientific results obtained by IXPE in the years 2022 and 2023 are briefly discussed. It must be stated that this is not a complete list of all the

results accomplished by the collaboration, but it provides an overview to present the impact of the mission on the astrophysics community. Specifically, greater attention will be dedicated to the groundbreaking results, to targets that can be exploited to point out the main features of the algorithm presented in this work, or, where explicitly indicated, to observations and analysis to which I personally contributed.

Polarimetry requires long observational times, that, depending on the target source, can range from hours to many days. The mission observing plan consists in pointing known X-ray sources over multiple IXPE orbits, until the observation is complete. As an example, in Fig. 3.1 the targets of the first year of IXPE observations are reported.

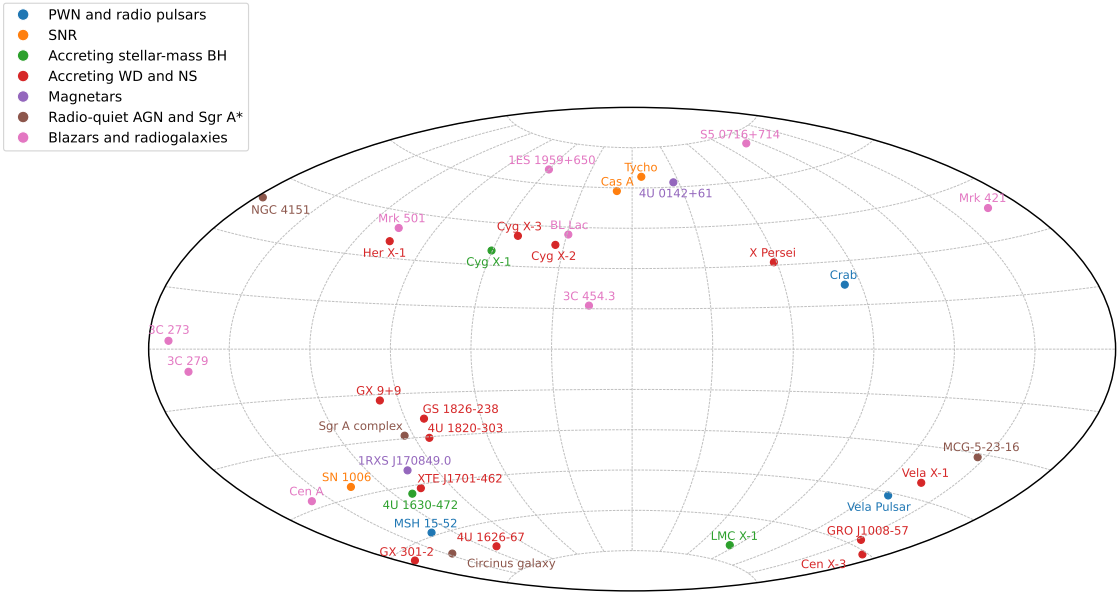

Figure 3.1: **List of sources observed during the first year of IXPE mission.** Different colors corresponds to different celestial objects.

Pulsar Wind Nebulae

IXPE results regarding Pulsar Wind Nebulae are presented here as they are functional to discuss the *polarization leakage* effect in the following chapters (e.g., see Sec. 5.4.2). As mentioned earlier in this thesis, the **Crab Nebula** is the only astrophysical source for which a previous significant measurement of polarization in the soft X-ray band was detected [20; 10; 26]. Hence, the observation plan for IXPE promptly included the Crab Pulsar and Nebula for two observations, in February and March 2022.

IXPE allowed the first detection of the phase- and space-resolved polarization of the source in the soft X-ray band [58]. Both showed some discrepancies with the existing models.

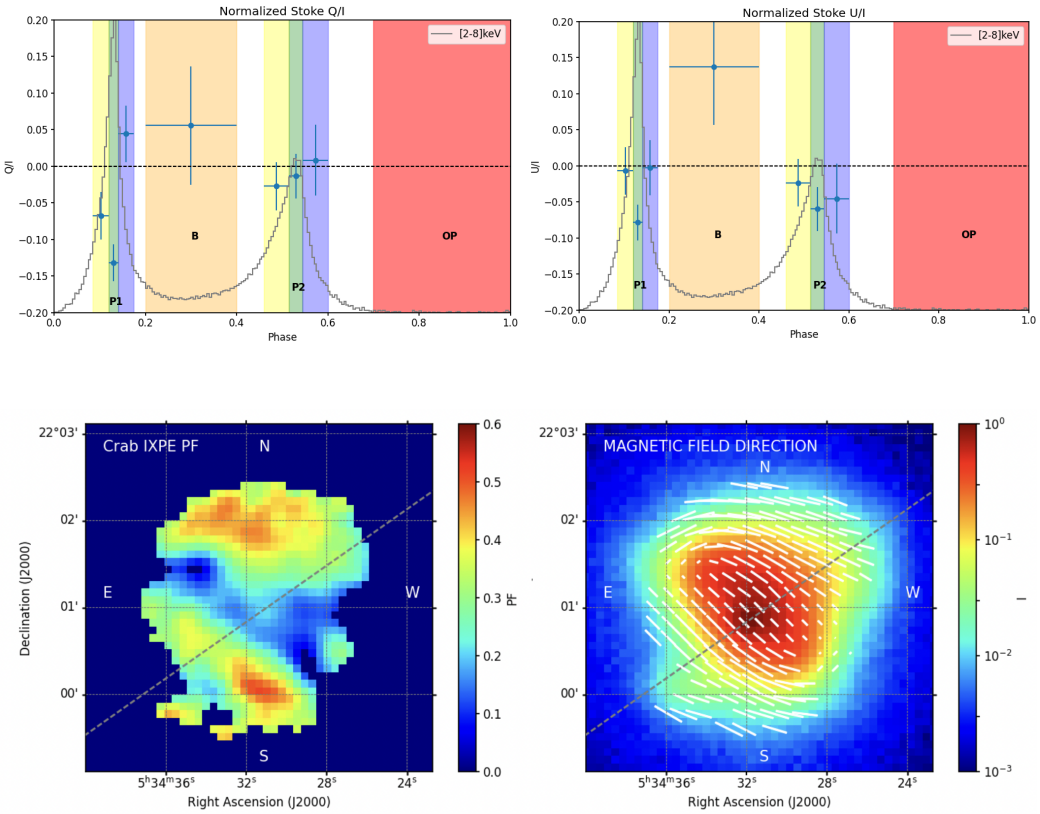

Figure 3.2: **Phase- and space-resolved IXPE observation of Crab nebula.** *Upper panels*: normalized Stokes parameters for the phase-resolved observation. Phase bins are represented by colors: Off-pulse (OP, *red*), Peak 1 and 2 (P1 and P2, *green*), left and right wings (*yellow and blue*), Bridge region (B, *orange*). Error bars correspond to 1σ errors. *Lower-left panel*: map of the polarization degree (PF) of the Crab Nebula. *Lower-right panel*: colors of the image represent the intensity map (I) in the 2-8 keV energy band, while the *white lines* show the binned polarization direction. Image credits: [58]

Regarding the phase-resolved detection, significant polarization was observed exclusively in the core of the main peak (PD $= 15.4 \pm 2.5\%$), while it was below the MDP threshold for the second peak and the bridge region. Upper panels of fig. 3.2 report the detected Q/I and U/I for the phase-resolved observation. The relatively low average polarization stands in contrast to the prevailing characteristics of the majority of established pulsar (PSR) models (e.g. [59]). These models predict the polarization fraction in pulsed emission to fall within the range of 40-80%, with a particularly high

polarization degree in the bridge region. Peaks, on the other hand, tend to experience depolarization. Analytical models of striped-wind emission propose the possibility of lower polarization in bridge, while also predicting complete unpolarization in the first peak [60]. Moreover, the measured polarized fraction for the integrated pulsed emission in hard X-rays exceeded 20%-30%, and a decrease of the polarization degree in the soft X-ray band is not predicted by any model. IXPE results suggested that current pulsar models overlook crucial physical factors, such as micro-turbulence, which could result in significant depolarization [61], as well as short time-scale variability [62].

Regarding the nebula emission, the integrated polarization degree was found to be in agreement with the previous observations, while the polarization angle differs significantly from other estimates. This reflects the spatial variation of the polarization degree, firstly detected by IXPE, or possible temporal variability. In the lower section of Fig. 3.2 the space-resolved polarized structure of the nebula is shown in the *left panel*, as well as the count map in the *right panel*. The spatially resolved PD reaches a maximum of $\sim 50\%$, which is significantly higher than what expected from predictions based on synchrotron turbulent modelling of the torus and the inner ring luminosity profile, which were calibrated on the OSO-8 results [63]. This indicates the possibility that the turbulence level and its evolution within the nebula are not as intense as initially predicted. The lower polarization degree in the inner region of the nebula, and the higher polarization degree observed in the outer regions, are in contrast with the optical estimates, where the higher fraction is observed in the inner region [64]. This result confirms previously proposed models suggesting the possibility that optical and X-ray emitting particles undergo acceleration in different locations, sampling distinct regions of the nebula [47], and suggests a highly ordered magnetic field at the edge of the torus of the nebula.

The Crab was not the only PWN which IXPE observed during the first year. The **Vela nebula** was also an important target of the mission, as previous radio observations showed a very high polarization degree ($\sim 60\%$) [65]. Also in this case IXPE could perform a space-resolved polarization analysis, as reported in Fig. 3.3 [66]. The pulsar contribution to the flux in the 2-8 keV is very low (less than 10% of the counts), so the polarization of the pulsed source was not measured by IXPE, and its contribution to the nebula emission is negligible. Fig. 3.3 shows the very high polarization degree detected from the nebula emission, with some bins exceeding 60%. This value is close to the highest polarization degree achievable for synchrotron emission, suggesting a strong uniformity of the magnetic field across the emission area. Due to the rapid cooling of electrons that emit synchrotron X-rays, the emitted photons originate in close proximity to the acceleration zone. Consequently, this observation challenges the plausibility of turbulence-driven diffusive shock acceleration and instead suggests alternative processes being responsible for accelerating the particles within the pulsar wind nebula (PWN) at the termination shock.

For both Crab and Vela nebulae the polarization leakage effect was considered when discussing the results. This effect could contribute approximately up to 10% and 5% to

the polarization fraction detected in some of the regions of the respective observations. Since both sources exhibit a high degree of polarization, the impact on the results was mild, and the main physical considerations could be considered largely unaffected. However, it is crucial to acknowledge that polarization leakage can notably influence the analysis of extended sources, potentially imposing limitations on the achievement of significant results.

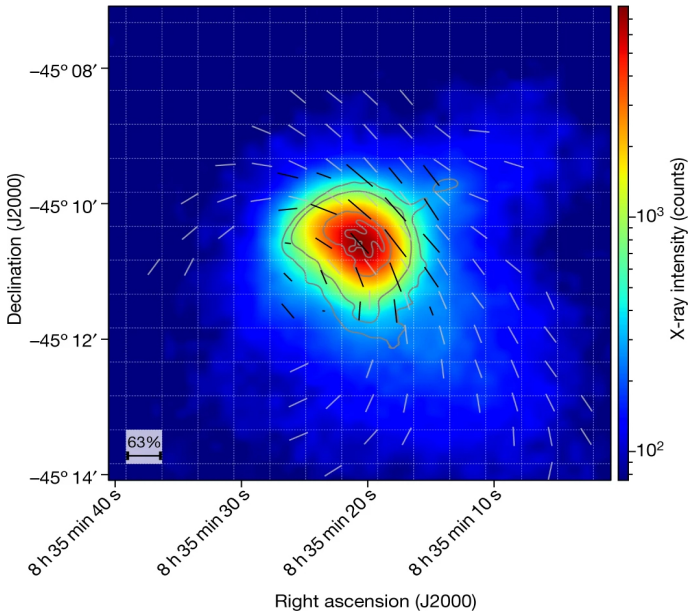

Figure 3.3: **Intensity map of the Vela PWN, alongside with the binned X-ray polarization direction in the central part of the nebula (*black lines*)**. The length of the lines is proportional to the polarization degree of the same bin. The *thinner gray lines* represent the polarization direction in the radio band. The *grey contours* are obtained from Chandra observations in the same 2-8 keV range. Image credits: [66]

Active Galactic Nuclei (AGNs)

I am part of the AGN Topical Working Group (TWG) of IXPE, and I collaborated to the analysis of some of the IXPE observations. Active Galactic Nuclei (AGN) are galaxies characterized by very bright cores and, in some cases, relativistic jets, fueled by material accretion onto a central supermassive black hole. These sources have been extensively studied in the past, revealing diverse observational behaviors across the electromagnetic spectrum [67; 68]. Various studies attempted to categorize them based on a limited set of physical parameters. The most widely accepted and straightforward classification considers two key physical parameters of AGN: the inclination of the torus

to the line of sight and the source luminosity across the spectrum [69; 70; 71]. Fig. 3.4 offers a schematic representation of this classification.

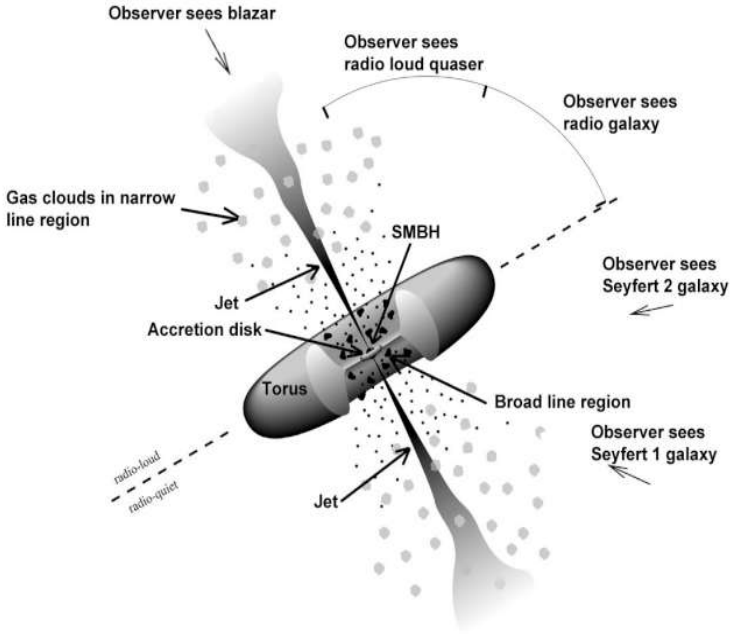

Figure 3.4: **Illustration of the Unified AGN Model**. The image shows the diverse classes distinguished by the observer orientation with respect to the accretion disk, the dusty torus, and the jet. Image credits: https://fermi.gsfc.nasa.gov/science/eteu/agn/

Blazars are AGN whose jets, which emit across a spectrum ranging from radio to extremely high-energy gamma rays, are directed very nearly towards our line of sight. The study of the magnetic field structure and emission mechanisms within these jets relies heavily on the polarization measurements across multiple wavelengths. However, before IXPE launch, the field of sensitive polarization observations remained confined to the radio, infrared, and optical domains, resulting in a lack of information concerning the environmental dynamics encountered by the most energetic particles.

The first detection of linear polarization in blazars was found by IXPE observation of **Mrk-501**, a blazar characterized by extremely variable emission in the gamma-ray band [72]. A polarization degree PD = $(10 \pm 2)\%$ was measured in the 2-8 keV band, with a polarization direction aligned with the jet [73]. The polarization degree in this band is more than a factor 2 higher than the optical one, and more than a factor 6 than the radio one, indicating an acceleration scenario where the higher energy particles emit from more magnetically ordered regions, and points to a shock front as the source of particle acceleration. This scheme is well reproduced by an energy-stratified jet model [74], where the X-rays are emitted closer to the site of acceleration of the particles, while particles emitting from radio to optical wavelengths cover a more extensive downstream region due to their extended radiative cooling times.

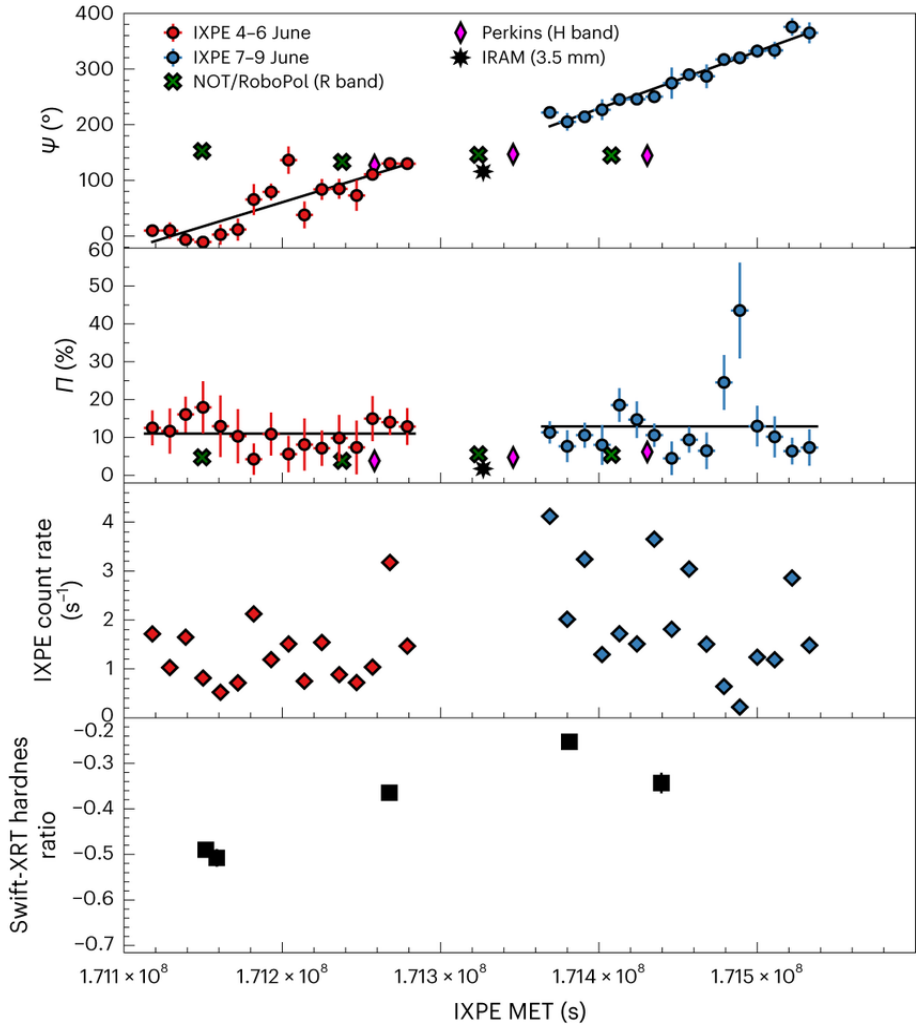

Figure 3.5: **Mrk-421 multi-wavelength results during IXPE observation.** *From top to bottom*: polarization angle, polarization degree, IXPE photon count rate and Swift-XRT hardness ratio as a function of IXPE Mission Elapsed Time (MET). IXPE data are reported as *red and blue dots and diamonds*. Swift-XRT hardness ratios are shown as *black squares*. Simultaneous radio (IRAM, *black stars*), infrared (Perkins, *magenta diamonds*) and optical (NOT/RoboPol, *green crosses*) observations are reported too. Image credits: [75]

A second interesting source observed multiple times by IXPE was the blazar **Mrk-421**. During its first observation in May 2022, a PD = $(15 \pm 2)\%$ was detected, with a constant polarization angle $\Phi = (35 \pm 4)^\circ$ [76]. As for Mrk-501, the polarization degree measured in the X-ray band was notably higher than those observed in the optical, radio, and infrared bands. The proposed scenario, therefore, followed a similar path to the one described for Mrk-501, suggesting a shock as the mechanism at play for accelerating the particles emitting in the X-ray band. However, during the second observation of Mrk-421, conducted in June 2022, we discovered a polarization angle rotation in the data [75]. Across the 5 days of observations, the polarization angle rotated by more than 360°, by keeping a constant PD = $(10 \pm 1)\%$. IXPE results are reported in Fig. 3.5, alongside with simultaneous multi-instruments measurements. The analysis confirmed the energy-stratified jet scenario, as the polarization degree was constantly higher than in other energy bands, and no significant polarization angle rotation was detected at longer wavelengths. Our analysis suggested that the rotation of the polarization vector was unlikely caused by random walks of the polarization angle. We also excluded the possibility that such rotation is a consequence of a magnetic reconnection process, as they would also cause γ-ray flares [77], which were not observed from this source during IXPE observation, and the polarization degree in the optical band should be comparable with the X-ray one [78]. The rotation of the polarization angle involves more probably a specific model where an off-axis emission feature progresses towards the observer along a helical magnetic field [79; 80]. The observed rotation rate is established by the time the feature, such as a magnetosonic shock, takes to complete an orbit around the jet axis. The process is schematically illustrated in Fig. 3.6, from both the perspective of the host galaxy reference frame (*left panel*) and of an observer whose line of sight aligns with the jet axis (*right panel*). This model is compatible with the absence of rotation observed in the radio, infrared and optical observations. A third observation of Mrk-421, conducted in December 2022, contributed to the interpretation of the emission processes of the source. Firstly, it further confirmed the energy-stratified jet model scenario by achieving a polarization degree of $(14 \pm 1)\%$ [81]. In contrast to the second observation, a clear rotation of the polarization vector was not observed. However, its deviation from a constant value necessitated a multi-component model fit. This fit encompassed both a constant and rotating polarization component [Pacciani et al., in prep]. This observation provided additional insights into the internal geometry of the jet, emphasizing the complex interactions between coexisting stable and rotating magnetic field structures.

In all the Mrk-421 IXPE observations I contributed to the time-resolved analysis, by evaluating the statistical significance of the time variability of the observations, and investigating possible energy-resolved variations. Moreover, I evaluated the impact of the background subtraction to the results. Specifically, the analyses were conducted according to two different approaches: a χ^2 test-based analysis, and an unbinned event-based maximum likelihood method [82]. I focused on the first approach, by combining both a flux-dependent and independent analysis, and sharing the outcomes and their

possible interpretation to the principal investigators and the blazars TWG members.

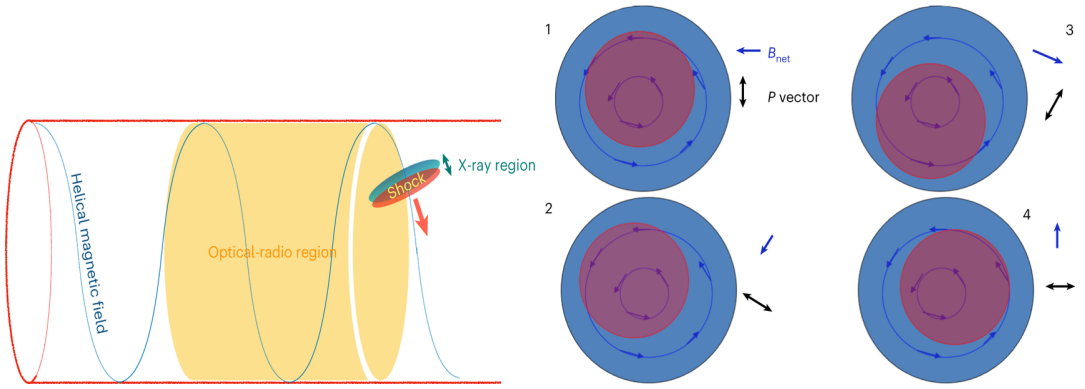

Figure 3.6: **Representation of an off-axis emission feature, like a magnetosonic shock, moving along the helical magnetic field lines down the jet of a blazar and its emission characteristics**. In the *left panel* the representation of the scenario in the blazar reference frame is illustrated. In the *right panel*, the perspective of a distant observer aligned with the jet shows the appearance of the emission feature (*red circles*), magnetic field, and polarization vector at four azimuthal positions along its spiral path. Image credits: [75]

Supernovae remnants

As for Pulsar Wind Nebulae, also Supernova Remnants (SNR) results are presented as they will be functional for the polarization leakage characterization within the context of this work (e.g., see Sec. 7.3). Supernovae represent the explosive endpoint in the life cycle of massive stars. Following a supernova event, an expanding remnant shell is left behind, therefore named a supernova remnant. SNRs play a crucial role in the enrichment of the interstellar medium. The material ejected during a supernova contains heavy elements synthesized in the stellar core, and particles can be accelerated to energies of hundreds of TeV [83].

Various observational methods, including X-ray and radio astronomy, have enabled the detection and study of SNRs. While they are considered to be the primary generators of galactic cosmic rays through diffusive shock acceleration, numerous uncertainties persist regarding the specific conditions existing at shock fronts, specifically in the context of the magnetic fields structure in proximity to the particle acceleration areas. Studies of the polarization of supernova remnants within the radio frequency range determined the synchrotron origin of the observed radiation, giving significant insights into the magnetic-field configurations and their association with particle acceleration. The direction of the radio polarization generally indicates a large-scale magnetic-field structure which shows a pattern depending on the SNR age [84]. Younger supernovae

showed a radial magnetic field with respect to the shock front, while for older SNRs the radio polarization map showed a tangential direction.

The high energy electrons responsible for X-ray emission, due to their rapid energy loss, exhibit a significantly shorter lifetime compared to radio-emitting electrons. Consequently, they are confined to smaller regions within sub-parsec scales from the acceleration sites. On spatial scales accessible in X-rays, there are two competing ideas regarding the magnetic-field topology in supernova remnants. One perspective suggests that shock compression enhances the magnetic-field component parallel to the shock front, resulting in a predominantly tangential magnetic field [85]. On the other hand, various processes proposed to explain the radial magnetic field in the radio band may already be in action close to the shock [86], where X-ray photons are produced. Therefore, through X-ray polarimetry it is possible to access information about the orientation (through the polarization angle) and turbulence (through polarization fraction) of the magnetic field close to the particle acceleration sites.

IXPE detected similar space-resolved polarization signals from three SNRs: **Cas-A** [28], **Tycho** [87] and the North-East region of **SN1006** [88]. By rotating each photon predicted polarization direction the analyses demonstrated that, as in radio, the polarization angle is perpendicular to the radius of the SNR, indicating a predominantly radial magnetic field. The models considering the processes responsible for the radial magnetic field observed in the radio band already at work on the sub-parsec scales at which the X-rays are emitted are reinforced by these observation. The different polarization fractions calculated from the three SNRs ($\sim$2 % for Cas-A, $\sim$12 % for Tycho and $\sim$22 % for SN1006) indicate different levels of turbulence in the acceleration regions.

Accreting compact objects

X-ray binaries results are discussed here as some observations will be exploited to compare the results of the standard analysis with the algorithm introduced in this work (e.g., see Sec. 7.3). An X-ray binary system forms when gas is torn from a star and accretes onto a compact object, as a black hole or a neutron star, which heats the gas sufficiently to emit X-rays. An artist representation of this kind of system is shown in Fig. 3.7. IXPE observed several systems, here some benchmark examples are reported.

Observations of X-ray accreting neutron star **Hercules X-1** detected a degree of linear polarization of $\sim$ 10% [89], which is significantly lower than the theoretical expectations for this source. Caiazzo & Heil [90] predicted indeed a polarization degree in the 1-10 keV energy band of $\sim$60-80 %, by modeling the pulse shape of the main peak, as well as the modulation of the cyclotron line with phase. This surprisingly low polarization observed by IXPE cannot be explained by depolarization through the way from the source to the observer. In general, it can not be interpreted in framework of existing models. However, some plausible scenarios were presented in Doroshenko

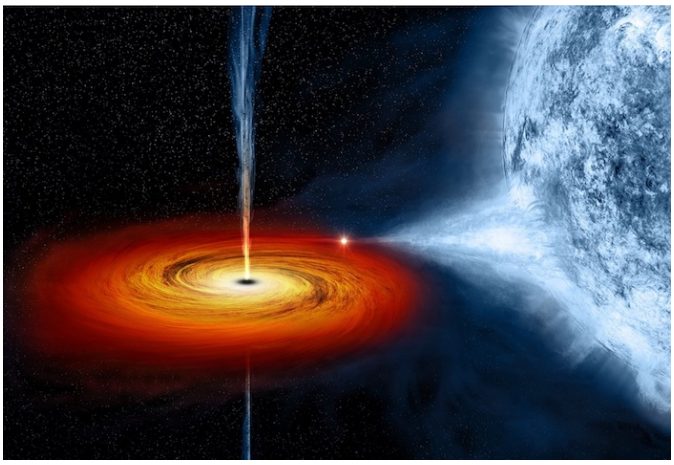

Figure 3.7: **Artist's impression of an X-ray binary system**. This illustration depicts the torn of the outer layers of the companion star, and the accretion disk which is formed around the compact object. A relativistic jet is also represented. Image credits: NASA/CXC/M.Weiss

et al. [89]. Firstly, the average polarization fraction is computed by considering instances when the south pole and the north pole of the pulsar are oriented toward the observer [91]. Since these poles exhibit distinct polarization properties, blending the two could lead to depolarization in the time-averaged observation. This phenomenon may partially account for the low polarization degree detected by IXPE. Nevertheless, the polarization fraction remains consistently low even during time intervals characterized by the emission from a single pole. The low polarization degree likely stems from a complex interplay of various mechanisms. A comprehensive understanding of the polarization properties of Her X-1, as well as other X-ray pulsars, needs further studies of accretion physics and emission mechanisms in these celestial objects. The low polarization degree observed by IXPE serves as a valuable input for theoretical models needed to clarify the emission from magnetized neutron stars undergoing accretion.

An energy-dependant analysis was performed for the accreting neutron star **GX 5-1**, observed by IXPE twice in March and April 2023 [92]. For these observations, as depicted in Fig. 3.8, a $\sim 20°$ shift of the polarization angle was observed between low energy and high energy photons. Its emission could be modelled with a contribution from the disk and a harder boundary layer (BL) or spreading layer (SL) emission, which could also be reflected by the same disk [93]. A simulation of the polarization signal from such a model, and the spectro-polarimetric analysis of the source, revealed that the disk and the BL-SL emissions could have different and non-orthogonal polarization angles, and thus could explain the energy-dependent PA shift. The same twofold nature of GX 5-1 emission could also account for the slight increase of the polarization fraction with energy, although the measurement is not significant enough. This particular detection will be discussed in more details in Chapter 7.

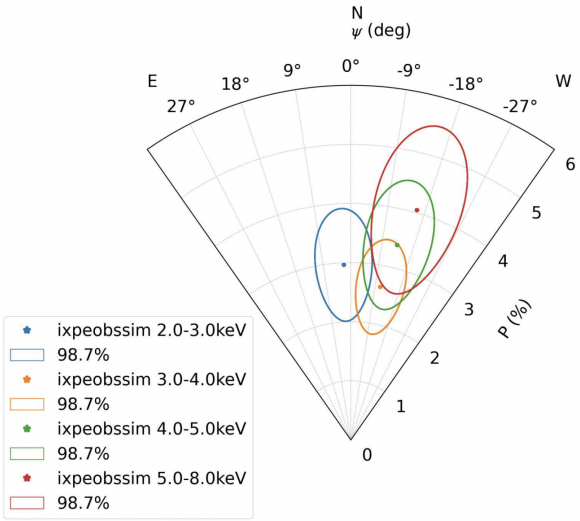

Figure 3.8: **IXPE results for the accreting neutron star GX 5-1**. The polar plot reports the energy-resolved polarization properties (degree and angle) of GX 5-1 IXPE observation. *Ellipses* represent the 3σ confidence interval. Image credits: [92]

Concerning accreting black holes systems, IXPE provided numerous interesting and unexpected results with its observations. In particular, IXPE detected a 4% polarization degree in the Black Hole X-ray Binary **Cyg X-1** [94]. This polarization, which aligns with the orientation of the radio jet, imposes constraints on the spatial arrangement of the hot corona. The observed high degree of polarization suggests that the X-ray bright region is viewed at a greater inclination than the inclination of the binary orbit. This is not common in binary systems, as stellar-mass black holes originate from supernovae events, and the inner accretion disk inclination usually matches the orbital inclination [95]. The supernova in Cyg X-1 may have resulted in a black hole with a spin that is not aligned. Gravitational effects could have potentially aligned the angular momentum vector of the inner accretion flow with the spin vector of the black hole [96].

While the observations of Cyg X-1 was conducted during hard state, the system **4U 1630-47** was observed in the high-soft state. The analysis revealed an unexpectedly high polarisation degree which changes with energy: a 6% polarization fraction was measured at 2 keV, rising to 10% at 8 keV [97]. In Fig. 3.9 the X-ray polarization measurement of 4U 1630-47 as a function of the energy is reported.

This measurement is not compatible with standard models of thin accretion discs, as they predict a lower polarization and no variability with energy [21]. A comprehensive explanation for the observed polarization degree and position angle involves a combination of a low black hole spin and a highly-ionized atmosphere characterized by a high optical depth. Fluctuations in the optical depth or variations in the outflow velocity of such an atmospheric configuration could account for the observed variability with

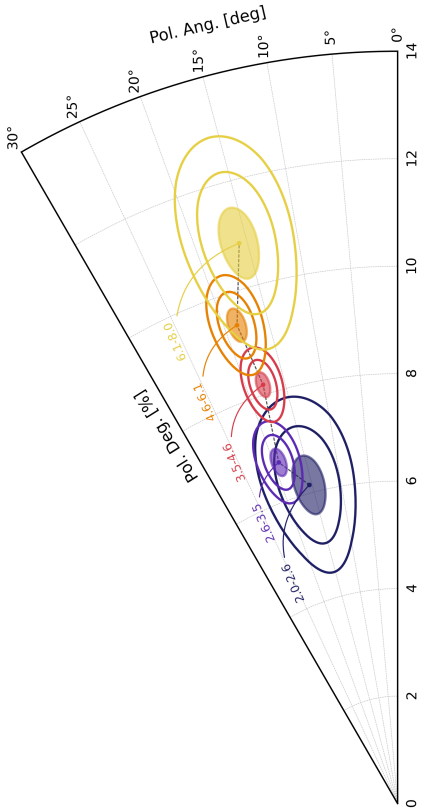

Figure 3.9: **Polarization degree and angle of 4U 1630-47 for each energy interval**. The *ellipses* represent the 1-2-3σ confidence levels. Image credits: [97]

energy.

Finally, **Cyg X-3** is an X-ray binary that IXPE observed in two different states: a hard X-ray radio quiescent state in October and November 2022, and during its transition to the soft state in December 2022. A high degree of polarization was detected during the hard state observation (PD $\sim$ 25%, constant with energy), and its modelling suggested that only reflected and scattered light is observed, and the main X-ray source is obscured by a thin funnel-like structure [98]. The direct observation of the X-ray source would have resulted in a variability of the polarization degree with energy, while the geometry and inclination of the system [99; 100], combined with the observed polarization properties, constrained the funnel-like shape of the envelope.

Magnetars

As for X-ray binaries, magnetars results are reported in this section as their observations are exploited to present some of the results of this thesis (e.g., see Sec. 7.2). IXPE observed four magnetars so far: **4U 0142+61** [101], **1RXS J170849.0-400910**

[102], **SGR 1806-20** [103] and **1E 2259+586** [104]. The first polarized X-ray signal from a magnetar was detected with the observation of 4U 0142+61, which is among the brightest persistent magnetars [105]. Both a phase- and energy-dependant analysis was conducted, and results are reported in Fig. 3.10.

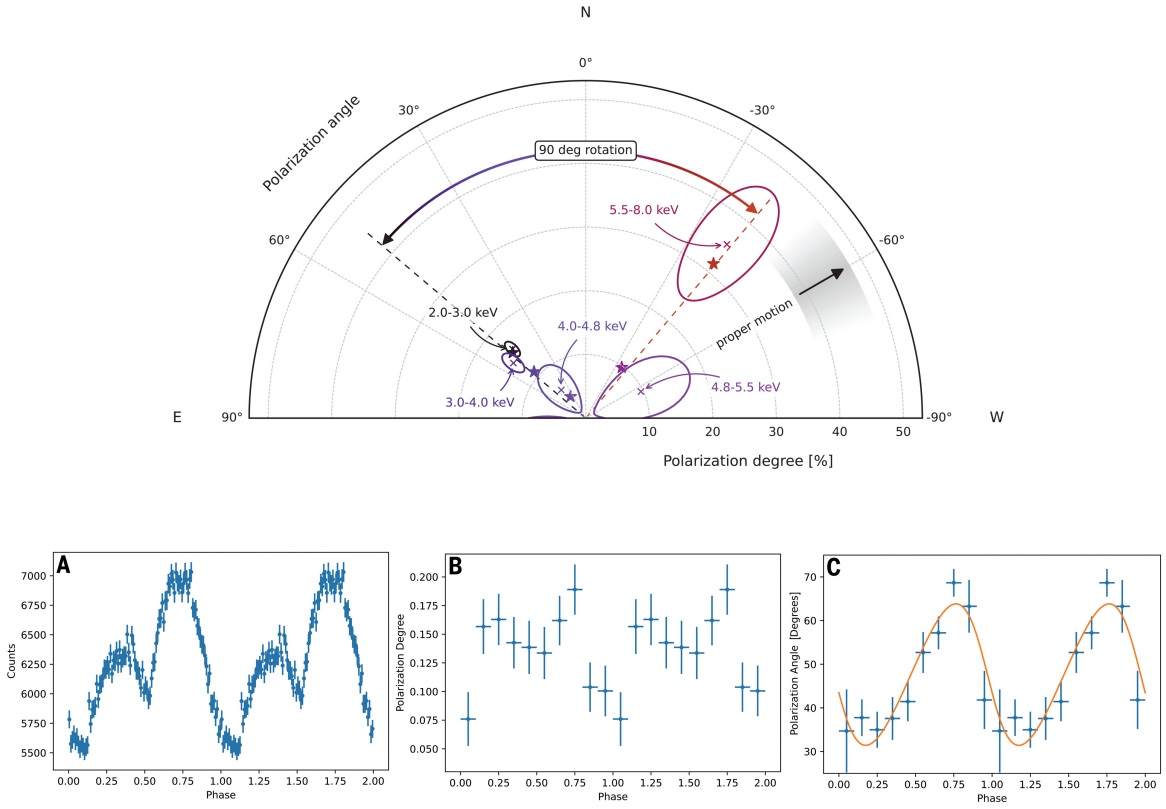

Figure 3.10: **Energy- and phase-resolved analysis of 4U 0142+61**. *Upper panel*: polar plot reporting the polarization degree and angle for five energy bins. The contours indicate the 1σ confidence level region. Stars indicate the corresponding PD and PA for the RCS scenario. The two *dashed lines* show the change in polarization angle from 2-3 keV bin (*black*) to the 5.5-8 keV bin. The black arrow and gray area represent the proper motion direction of the source and its uncertainty. *Lower panels*: (A) IXPE counts as a function of the magnetar phase. Error bars represent the 1σ confidence level. (B) Polarization degree as a function of the phase. Error bars indicate $\Delta\log L = 1$, where L is the unbinned likelihood. (C) Polarization angle as a function of the phase. Error bars are calculated accordingly to the polarization degree ones. The orange curve shows the best-fitting rotating vector model. Image credits: [101]

The energy-resolved analysis revealed distinct polarization behaviors in the low (2-4 keV) and in the high (5.5-8 keV) energy ranges, whereas at ∼4-5 keV the polarization degree drops below the MDP. Measurements of polarization angles strongly indicated a 90° shift between the low and the high energy bins. Results are in agreement with a

Resonant Compton Scattering scenario, by considering a hotter belt on the star surface close to the magnetic equator. 3D simulations assuming a magnetic field strength of $\sim 10^{14}$ G predict a predominance of O-mode photons at low energies (2-4 keV) with PD $\sim 10-15\%$ and an excess of X-mode photons at high energies with PD $\sim 35-40\%$ [101]. Results for such an emission model are reported as stars in the *upper panel* of Fig. 3.10.

The phase-resolved analysis aimed at observing the vacuum birefringence predicted by QED. According to QED effects, photon polarization vectors align with the star magnetic field up to a certain distance, known as the *polarization-limiting radius*, typically around 100 times the stellar radius for a magnetar [106]. Simulations of this scenario reproduced the observed polarization degree and angle in the phase-resolved analysis. Therefore, although the total polarization degree measured in 4U 0142+61 is not high enough to definitively confirm the role of QED vacuum birefringence, its influence on the radiation polarization was considered key for the results interpretation.

Gamma Ray Bursts

I had the opportunity to participate to the only IXPE analysis so far of a Gamma Ray Burst (GRB). GRBs are extremely energetic events which have been observed in outer galaxies. These phenomena exhibit an initial *prompt* release of gamma-rays, which represent the burst most luminous stage, and subsequently a fading *afterglow*, which can last for several days or even spans years, emitting radiation across the entire electromagnetic spectrum.

IXPE didn't plan to observe any GRB, since its reaction time is relatively slow (2-3 days). However, on October 9th, 2022 an exceptionally bright GRB occurred, and it was recorded as the brightest of all time [107]. Thanks to a ToO, IXPE was pointed towards **GRB-221009A** on October 11, to observe for the first time the 2-8 keV X-ray polarization of a GRB afterglow [108]. We detected no significant polarization, but an upper limit of 13.8% on the polarization degree of the afterglow emission was established. This outcome places restrictions on factors such as the jet opening angle and the viewing angle of the GRB. Interestingly, IXPE detected also halo-rings composed of dust-scattered photons, which represent echoes of the GRB prompt phase. This singular IXPE pointing serves as both the primary evaluation of X-ray polarization in a GRB afterglow and the first step towards a more comprehensive understanding of these phenomena.

Chapter 4

Machine Learning and Convolutional Neural Networks

Machine learning is a subset of artificial intelligence that focuses on enabling computers to learn from data, and make predictions or decisions without explicit programming. It draws inspiration from the human brain ability to recognize patterns and react to new information. Machine learning is a field which includes several techniques: among them, Neural Networks (NNs) are one of the most versatile and powerful tools, and today are employed in many contexts both in the industry and in the academy. Neural Networks include again a large variety of algorithms, developed and optimized to perform different types of task. Between them, Convolutional Neural Networks (CNNs) have emerged as a groundbreaking innovation that has revolutionized the field of computer vision. CNNs are a specialized class of Neural Networks designed to excel in problems as image recognition, object detection, and image generation. With their ability to automatically learn and extract meaningful features from raw data, CNNs have found applications in a wide range of domains, including the field of research in physics.

In this chapter, an overview of the main concepts and functioning of the Neural Networks are depicted, specifically with a focus on the CNNs. A greater attention is given to a particular type of CNN, the DenseNet, as it is the fundamental building block of the algorithm developed in this work.

4.1 Neural Networks overview

Hystorical overview

A first prototype of the concept of Neural Network was introduced by the neurophysiologist W. McCulloch and the mathematician W. Pitts back in 1943. Their work presented a simplified computational model inspired by how biological neurons cooperate in animal brains to perform complex computations using propositional logic [109].

Their model of the biological neuron, which later became known as an *artificial neuron*, was made up of one or more binary (on/off) inputs and one binary output. Its working principle was straightforward: the artificial neuron output would activate if a predefined threshold number of its inputs were active.

A slightly more elaborated version of the artificial neuron, the *Linear Threshold Unit (LTU)*, was defined in 1958 by F. Rosenblatt and became the building block of the *Perceptron*, which could be considered the first example of an artificial Neural Network [110]. A schematic representation of the LTU and the perceptron is reported in Fig. 4.1.

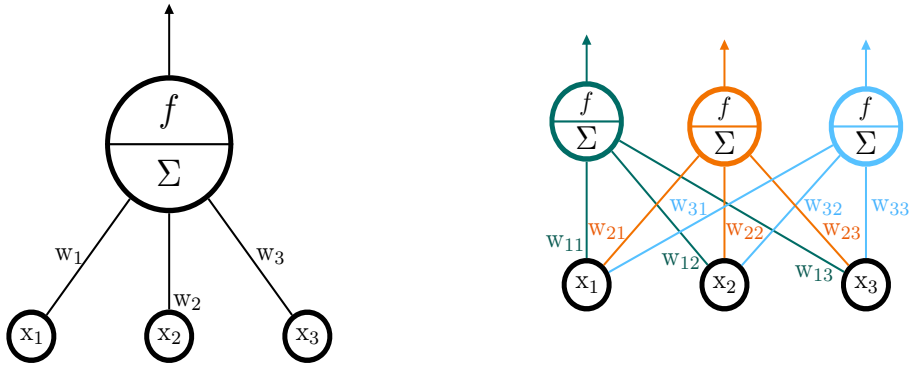

Figure 4.1: **Linear Threshold Unit and Perceptron.** *Left panel*: schematic representation of the Linear Threshold Unit. *Right panel*: example of a Perceptron made up of three LTUs.

The on/off values of the inputs and outputs were replaced by numerical values x_i and each input connection was associated with a weight w_i. The LTU computed a weighted sum of its inputs ($\Sigma = \mathbf{w}^T \cdot \mathbf{x}$), and then applied an **activation function** that outputs the result (denoted f in Fig. 4.1). The activation function acts upon the weighted sum of inputs to the neuron, and determines the neuron output. Activation functions are necessary to enable complex mappings between inputs and outputs. Common activation functions include the Rectified Linear Unit (ReLU) and the Sigmoid [111]. The former outputs the input for positive values and zero otherwise. ReLU is a key factor for mitigating the *vanishing gradient* problem, which will be discussed in the following sections. Sigmoid function maps input values to a bounded range, from 0 to 1, and is often employed in binary classification tasks.

A **Perceptron** is composed of a single layer of LTUs, each one connected to all the inputs (see Fig. 4.1, *right panel*). It had the ability to learn and make binary classifications based on a linear decision boundary, however, there were several trivial tasks that the Perceptron couldn't solve effectively [112]. One notable limitation of the Perceptron was its inability to handle problems that were not linearly separable. For example, the Perceptron would struggle with a simple XOR (exclusive OR) problem,

where the data points of two classes cannot be separated by a single line [113]. To address this limitation and enable Neural Networks to handle more complex tasks, the **Multi-Layer Perceptron (MLP)** was designed [114]. The MLP consists of multiple layers of neurons (typically an input layer, one or more hidden layers, and an output layer) interconnected by weighted connections. By introducing these hidden layers and the associated activation functions (such as the sigmoid or ReLU functions), MLPs became capable of learning and approximating non-linear functions. This made them suitable for solving a wide range of problems, including those the Perceptron could not handle. In Fig. 4.2, a schematic view of the MLP is represented.

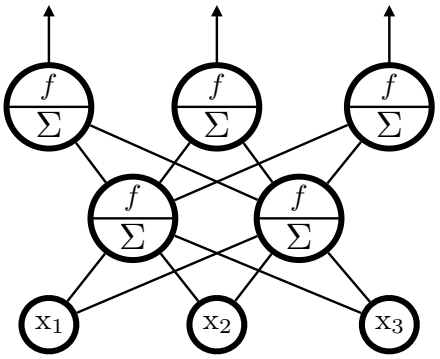

Figure 4.2: **Multi-Layer Perceptron**. Example of an MLP made up of an input layer, a single hidden layer, and the ouput layer.

The transition from Multi-Layer Perceptrons to Deep Neural Networks marked a significant step in the advancement of artificial NNs. **Deep Neural Networks (DNNs)** are MLPs with more complex architectures, and could have additional layers and specialized structures [115; 116]. They are designed to handle hierarchical representations of data, and learn spatial and temporal dependencies between them. Deeper networks can often achieve better performance with fewer parameters compared to simple Multi-Layer Perceptrons, and they outperformed them when using large data sets. However, it's essential to note that the choice between an MLP and a more complex DNN depends on the specific problem and the available data. For relatively simple tasks with limited data, an MLP might be sufficient and more computationally efficient. The choice of the right architecture involves factors like data size, complexity, computational resources, and the desired level of performance.

Building the algorithm

Neural Networks are a component of the machine learning category known as *supervised learning*. This is a machine learning paradigm where an algorithm learns to map

input data to corresponding output labels or target values. The primary goal of supervised learning is to learn from the labelled data a mapping function that can make accurate predictions for new, unseen data.

In order to achieve this goal all types of Neural Network, from MLPs to more complex architectures, need a **training phase**. The technique used to train NNs is called *Backpropagation*, and it was introduced in 1986 [117]. Backpropagation involves two key steps:

- **Forward step** The process begins with the input data, which is typically a single sample or a batch of samples from the training data set. Each sample consists of features or attributes. The input data is passed through the layers of the Neural Network from the input layer to the output layer. At each layer and for each neuron, the weighted sum and consequently the activation function are computed, and the process is repeated until the input has traversed all layers and reached the output layer. The final layer produces the network prediction or output. This prediction represents the model estimate for the given input.

  Once the network prediction is calculated, the next step is to evaluate the loss or error associated with this prediction. This is done by comparing the predicted output to the actual target values for the given input. The choice of loss function (e.g., *mean squared error* for regression, *cross-entropy* for classification) determines how this comparison is performed.

- **Backward step** At this point the algorithm calculates the contribution to the loss of each neuron in the last hidden layer, then quantifies the contribution of each neuron of the previous layer and so on, till it reaches the input layer. Thanks to the choice of an **Optimizer**, a tool which changes the weights, it is possible to decrease the loss and improve the predictions.

  Different types of Optimizers were developed through the years, but most of them are base on the technique of **gradient descent**. Defining θ as the vector of the N network parameters, i.e. the weights, and F as the chosen loss function, the gradient of the loss function with respect to each parameter is computed:

$$\nabla_\theta F(\theta) = \begin{bmatrix} \frac{d}{d\theta_0} F(\theta) \\ \frac{d}{d\theta_1} F(\theta) \\ \vdots \\ \frac{d}{d\theta_N} F(\theta) \end{bmatrix} \tag{4.1}$$

  where $\frac{d}{d\theta_i} F(\theta)$ are the partial derivatives of the loss function. This gradient represents the direction and magnitude of the steepest ascent of the loss function. The gradient is calculated using the chain rule of calculus, which breaks down the contribution of each parameter to the overall loss. Once the gradient is evaluated, the parameters are updated as:

$$\theta_{\mathbf{new}} = \theta - \eta \nabla_\theta F(\theta) \tag{4.2}$$

where η is called **Learning Rate**, and it is a numerical value chosen according to the type of network. As shown in equation 4.2, the parameters of the model are changed proportionally to the inverse of the gradient of the loss function: in this way it is possible to decrease the loss. The goal is to reach the minimum of the loss in order to get the best prediction.

The algorithm performs these two steps for every instance of the training set (or every batch of instances) making the loss smaller at every loop. If the loss is still large once the instances have all been passed through the network, it is possible to restart and feed it again with all the training set. Every time the network is trained with all the set, an **epoch** is completed. Concretely, the vector θ can be initially filled with random values, and it is then improved gradually, one step at the time, proportionally to the value of the learning rate.

A crucial step of the training phase involves the **validation** process. Rather than calculating the loss directly on the training set, an independent set known as the validation set is utilized for evaluation. This evaluation occurs after each epoch or batch of epochs. The primary purpose of this process is to mitigate the risk of **overfitting**. Overfitting occurs when a model learns properties of the training data that do not well generalize to new unseen data. This phenomenon remains undetectable if the evaluation of the loss is conducted on the training set. By calculating the loss on a distinct validation set after each epoch, it is possible to identify signs of overfitting, manifested as an increase in the loss values. It serves as an early warning system, allowing to adjust the model complexity or apply regularization techniques to enhance generalization performance. Importantly, during the validation phase, the network parameters remain unchanged, preserving the integrity of the evaluation process.

The validation process is also key to decide when to stop training a network. A standard method is not available and different techniques can be used according to the model, the data set and the computational resources at disposal. However, **early stopping** can be considered a baseline method to determine when to end the training phase [118]. It consists in monitoring a validation metric (e.g., the loss) and interrupt training when it stops improving. A second possible approach consists in setting a predetermined number of epochs for training. This approach is suitable when there is a reasonable expectation on the number of epochs required for the specific task, and often involves a large amount of tests and optimization processes.

4.2 Convolutional Neural Networks

In the previous section, we mentioned the distinctions between Deep Neural Networks and Multi-Layer Perceptrons, emphasizing that DNNs exhibit more complex architectural designs. The Convolutional Neural Networks are a particular class of DNNs.

The key difference with respect to the standard Neural Network architectures was the introduction of two specific layers: the convolutional and the pooling ones.

A **convolutional layer** is a fundamental building block of CNNs, designed specifically for processing grid-like data, such as images. Its main purpose is to detect patterns like edges, textures, or more complex structures within the input data, while preserving some degree of shift, scale and distortion invariance. The core operation of a convolutional layer is the convolution itself. It involves sliding a *kernel* across an image to output a final product for each position of the kernel. Kernels are small, learnable matrices that determine which features the convolutional layer should focus on. Each element within the matrix is multiplied for the corresponding pixel value, and the resulting products are summed to yield a singular output. Subsequently, the kernel shifts to a new position in the image, iterating the process. A representation of the convolution operation is reported in Fig. 4.3.

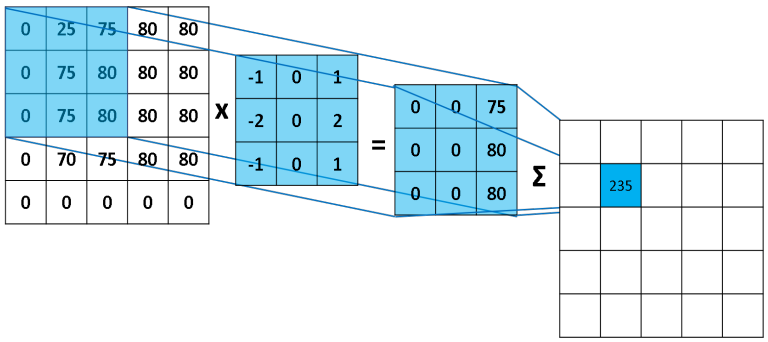

Figure 4.3: **Representation of the working principle of the convolution operation.** Credits: [119].

While the values characterizing the kernel are learnt by the network during the training process, some specifications of the convolution operation, as the kernel dimensions, must be tuned manually: the *padding* is a layer of pixels of some unchanging values wrapped around the original image so that the kernel can align conveniently to the image itself, while the *stride* represents the number of pixels the kernel should slide across the image (both horizontally and vertically) each iteration. Following a convolution operation, a series of feature maps is obtained. Each of these maps corresponds to the activation pattern produced by a distinct set of kernels applied to the input data. These feature maps effectively represent hierarchical features. The early layers of the network usually focus on identifying straightforward and basic patterns, while the deeper layers excel at elaborating more intricate and complex representations. This hierarchical organization enables the CNN to systematically capture the varying levels of complexity inherent in the input information.

Pooling layers primary goal is to reduce the spatial dimensions of the input volume, effectively downsampling the feature maps generated by the convolutional layers. Common pooling methods include *Max Pooling* and *Average Pooling*. Similar to convolution, a pooling layer involves a sliding window moving across the input feature maps. At each position, the pooling operation (max or average) is applied within the window, producing a single output for that region. A representation of the pooling operation is reported in Fig. 4.4.

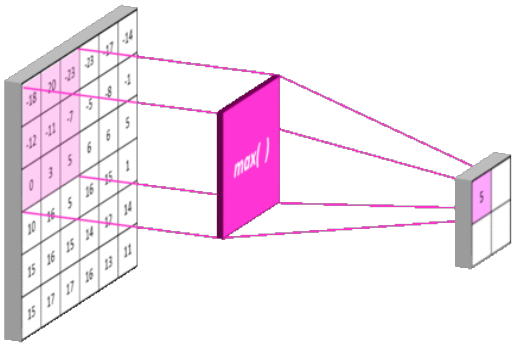

Figure 4.4: **Representation of the working principle of the Max Pooling operation.** Credits: [119].

The same parameters as the convolution operation (size, stride and padding) need to be tuned manually for each pooling layer.

It is common to add a pooling layer after each convolutional step. This combination serves two main purposes: firstly, convolutional layers pinpoint and highlight important features, and secondly, pooling layers reduce the size of the data, keeping the crucial information and making calculations simpler. This coordinated ensemble of convolutions and pooling layers generates a lower dimension representation of the image, which is finally flattened and given as an input to a fully connected Neural Network, which provides the final output of the CNN.

The first CNN structure was developed by the computer science researcher Yann LeCun and published in 1998 and called *LeNet-5* [120]. The schematic representation of its architecture is reported in Fig. 4.5. However, the first successful application of a Convolutional Neural Network occurred only 14 year later, when the *AlexNet* [121] architecture achieved an unprecedented classification efficiency of the ImageNet dataset, a popular image classification competition [122].

The delayed adoption of CNNs was in part attributed to computational constraints. The quite complex architectures involved in training Deep Neural Networks, especially the convolutional operations, demanded substantial computational power. During the initial years, conventional Central Processing Units (CPUs) struggled to handle the computational intensity of CNNs. The availability and widespread use of Graphics Processing Units (GPUs) finally allowed to exploit their full potential. GPUs are de-

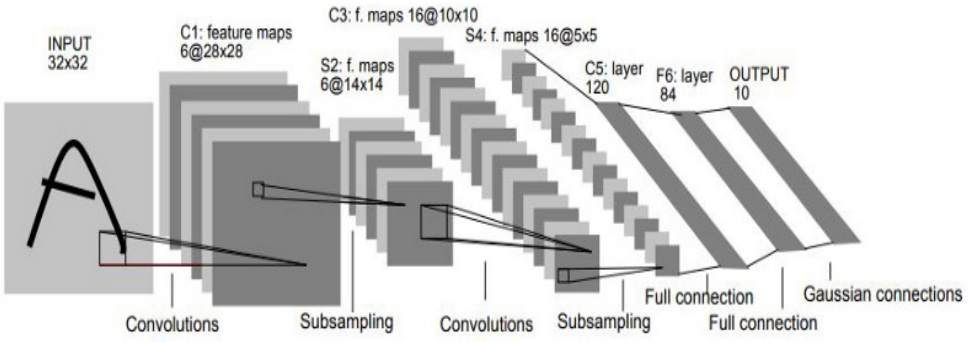

Figure 4.5: **LeNet-5 architecture.** Credits: [120]

signed for parallel processing tasks, and proved to be a game-changer for NNs computations. Their architecture, well suited for handling the parallel nature of convolutional operations, significantly accelerated the training process [123]. This newfound computational efficiency, coupled with advancements in model architectures like AlexNet, is the key for the breakthrough that revolutionized image classification.

Following AlexNet success, there was a growing interest in CNNs for different applications in computer vision. Researchers and practitioners started to develop and experiment with novel CNN architectures, refining designs for specific tasks. Google's *GoogLeNet*, also known as *Inception* [124], introduced the concept of inception modules, facilitating the training of deeper networks. The Visual Geometry Group's *VGGNet* proposed simpler, yet deeper architectures, demonstrating that depth was a crucial factor for improving performance [125]. Finally, Residual Networks, or *ResNets*, introduced skip connections that alleviated the vanishing gradient problem, which will be discussed in the following section, enabling the training of extremely deep networks [126]. An alternative structure which was designed to address the vanishing gradient problem was the *DenseNet* architecture [127], which at the time of writing this thesis represents the state-of-the-art CNN for some of the image classification and recognition tasks.

Dense Network

Due to the growth in depth and complexity of the CNNs, a new training issue called *vanishing gradient* was found. During the CNN training, the gradients that are back-propagated through the layers diminish exponentially, while the depth of the network increases. This diminishing gradient challenges the ability of the network to effectively update the weights of early layers, impeding their learning, *vanishing* the information contained in the original image and leading to suboptimal training results. As already mentioned, ResNets architecture, as other similar models [128; 129], addressed the problem by establishing shorter pathways connecting the initial layers with the later

ones.

The key innovation in DenseNets is the dense connectivity pattern, where each layer receives direct inputs from all the preceding layers and passes its own feature maps to all the subsequent layers. A representation of the architecture is reported in Fig. 4.6.

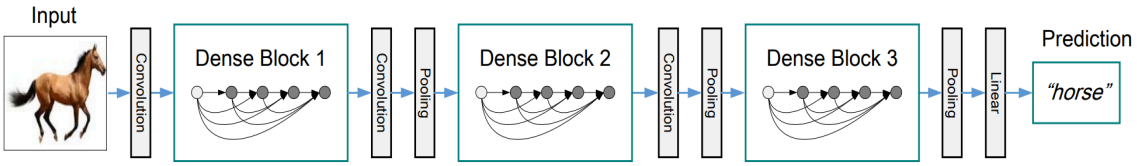

Figure 4.6: **DenseNet architecture with three dense blocks.** The layers between two adjacent blocks are the representation of the transition blocks. Credits: [127]

Essentially, the input to the n_{th} layer is the concatenation of the feature maps $\mathbf{x}_i$ of all the previous layers:

$$\mathbf{x}_n = F_n([\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{n-1}]) \tag{4.3}$$

where F_n represents a composite function, usually including an activation function and a convolutional operation. This connectivity pattern has several advantages. Firstly, it encourages feature reuse and enhances gradient flow through the network, mitigating the vanishing gradient problem. Secondly, it reduces the number of parameters as each layer receives the feature maps of all preceding layers, promoting parameter sharing.

The concatenation of the feature maps is possible only when they have the same dimensions. However, as discussed in Sec. 4.2, the downsampling of the image while passing through the network is a key feature of the CNNs to reduce the number of parameters. DenseNets solution is to divide the network into different densely connected *dense blocks*, and incorporating *transition blocks* to control the growth of the number of feature maps and manage computational complexity. These transition layers typically include a 1×1 convolutional layer followed by a downsampling operation, such as average pooling. Transition blocks ensure that the network remains computationally efficient while still capturing complex features.

The dense connections in the network should in principle increase the number of parameters of the network. If each function F_n generates k features map, the n_{th} layer will have as input $k_0 + k \cdot (n-1)$ maps, where k_0 is the number of channels of the input images. The key advantage of the DenseNets in terms of number of parameters, is that a relatively low value of k, compared to other CNN architectures, allows one to achieve the same performances in different tasks [127].

The low parameter count in DenseNets can be attributed to their adoption of an implicit form of *deep supervision*. This concept was introduced in Deeply Supervised Nets [130], where classifiers were attached to each hidden layer of the network, forcing them to predict significative features. While DenseNets do not follow the exact

approach of attaching classifiers to each layer, a single classifier at the network's top directly supervises all layers through at most two or three transition layers, simulating the process of deep supervision.

One of the simplest DenseNet architecture is called **DenseNet-121**, and it consists of 4 dense blocks, with a total of 120 convolutions and 4 average pooling operations. In the traditional architecture, before entering the initial dense block, the input images undergo a convolution operation with 16 output channels. This operation holds significant importance in the context of this work, and receives increased attention in the following chapter. The final number of parameters in the network depends on the choice of the *growth rate* variable k.

Summarizing, DenseNets are parameter-efficient networks due to their dense connectivity pattern, which enables feature reuse and sharing across layers. Moreover, the dense connectivity facilitates smoother gradient flow during training, addressing the vanishing gradient problem.

Chapter 5

A hybrid CNN-analytic approach for X-ray Polarimetry

Several groups have already shown the potential of Machine Learning to reconstruct the photo-electron track and determine the incident photons properties. For example, Kitaguchi et al. [4] trained a CNN to classify the PE tracks into polarization angle bins, and the same Neural Network also predicts the impact point position on the track. Moriakov et al. [5] tested the use of CNNs through regression instead of classification, again to infer the impact point and the PE initial emission direction. Peirson et al. [6] and Peirson & Romani [39] also used CNNs to predict energy, impact point, and polarization direction, evaluating prediction uncertainties for each event thorough a deep ensemble technique [131]. Despite the progresses made by these works, all of them have encountered some form of systematic error, which could stem from either unusual photo-electrons emission angles distributions for unpolarized sources, to discrepancies between simulated and laboratory data outcomes. In addition, all these works used standard cartesian convolutional blocks to process images with hexagonal pixels.

The main goal of my project is to develop an algorithm that leverages the strengths of Convolutional Neural Networks in enhancing performance in photo-electron tracks reconstruction, and the stability and reliability of the state-of-the-art analytical algorithm both with simulations and experimental data. The results presented in this section were published in April 2023 [132].

In the first section of the chapter, the dataset used for the developing of the algorithm is described. In Sec. 5.2 the relevance of the impact point parameter is highlighted in the context of IXPE. In Sec. 5.3 we describe the structure of the adopted CNN and illustrate its training and optimization processes. Finally, Sec. 5.4 discusses the results.

5.1 Dataset

In order to train and evaluate the performance of the algorithm, full-detector Monte Carlo simulations were used. These simulations were generated using a software called

ixpesim, which is built on top of GEANT4 and incorporates a customized version of the *Livermore Polarized* physics list [133]. We created various categories of simulations, but all of them shared the parameters listed in Table 5.1.

Gas chamber size	15(x) x 15(y) x 10(z) [mm]
Gas mixture	Dimethil Ether (DME)
Gas pressure	720 mbar
Gas ionization energy	24.5679 eV
Filling temperature	293.15 K

Table 5.1: **Summary of the simulated physical properties of the detector.**

Different categories of simulations were used for different tasks:

- **Training and validation data set:** 2 million events were generated from an unpolarized beam, with a flat energy spectrum in the range 1-9 keV. As already mentioned in the previous sections, this range includes the energies of the highest sensitivity (mostly due to the higher effective area) of current and future experiments adopting this X-ray polarization technique. This data set is used to train and validate our CNN.

- **Test data set:** Test samples were obtained simulating 100 000 events for each set at fixed energies: two sets (100% polarized and unpolarized) for 13 different mono-chromatic beams (between 2 keV and 8 keV with 0.5 keV step). These events are used to evaluate and compare the performance of the algorithms for different energies of the incident X-rays.

- **Point sources:** Three sets of 500 000 events were simulated to reproduce three different unpolarized point sources, with typical spectral shapes of astrophysical sources. Specifically, we simulated two power law spectra *PL1* and *PL2* (with a -0.7 and -1.7 spectral indices respectively) and one black body spectrum *BB* (kT = 3 keV). This data set is used to evaluate and study the systematic effect denominated *polarization leakage*.

- **Crab Nebula:** A simulation of the Crab Nebula as observed by a single Detector Unit of IXPE (DU1) was generated. The radiation is simulated to be unpolarized, and the observational time is based on the observation of the Crab Nebula performed by IXPE between February and March 2022 [58]. This set of data is used to evaluate the polarization leakage and its effect on the polarization measurements of extended sources.

For each simulated event the software provides the true energy, the true impact point position and the true emission angle. This information is used to label the events, making the supervised learning process of the CNN possible.

5.2 A key parameter: the impact point

Before delving into the description of the algorithm, I will go through a brief digression about the importance of the reconstruction of the incident photon impact point projected on the pixelated anode. It may not be directly related to the photon polarization properties, but it does have a significant role in the reconstruction of the emission angle with the moment analysis, as already mentioned in Sec. 2.3.2. Its evaluation is indeed the third phase of the standard analysis, and it is used in equation 2.8, which is also reported here for convenience:

$$M_2'(\phi) = \frac{\sum_i w_i[(x_i - x_{IP})\cos(\phi) + (y_i - y_{IP})\sin(\phi)]^2}{\sum_i w_i} \tag{5.1}$$

As a reminder, (x_{IP}, y_{IP}) are the predicted impact point coordinates, (x_i, y_i) the track pixel positions and w_i are the weights of each pixel (assigned according to the skewness of the track). By maximizing $M_2'(\phi)$ the predicted emission direction is achieved (see Sec. 2.3.2). In this section we illustrate two issues which highlight the importance of a correct estimation of the impact point.

5.2.1 Modulation factor recovery

To emphasize the importance of the impact point, we can refer to the modulation factor, i.e. the fraction of reconstructed polarization for a 100% polarized source (see Sec. 2.3.2). Specifically we wanted to investigate whether an improvement in the precision of the impact point determination can lead to a better reconstruction of the emission angle, and ultimately, to more accurate polarization measurements of incident photons.

To test this hypothesis, the standard reconstruction is applied but replacing the predicted IP with the Monte Carlo (true) one in Eq. 5.1. The modulation factor is then evaluated for the fixed energy polarized beams using this substitution, and it is compared with the outcome of the standard approach. The result of this comparison is shown in Fig. 5.1, where the modulation factor is plotted as a function of energy, by employing the true (*red*) or the predicted (*black*) impact point. Clearly, this substitution is infeasible with experimental data, due to the lack of knowledge about the true impact point. However, by comparing the two curves, it becomes evident how refining the accuracy of the IP estimation could have a direct impact on the performance of polarization measurements. Additionally, the red curve could also be interpreted as the upper limit of enhancement that could be achievable through improved accuracy in impact point estimation.

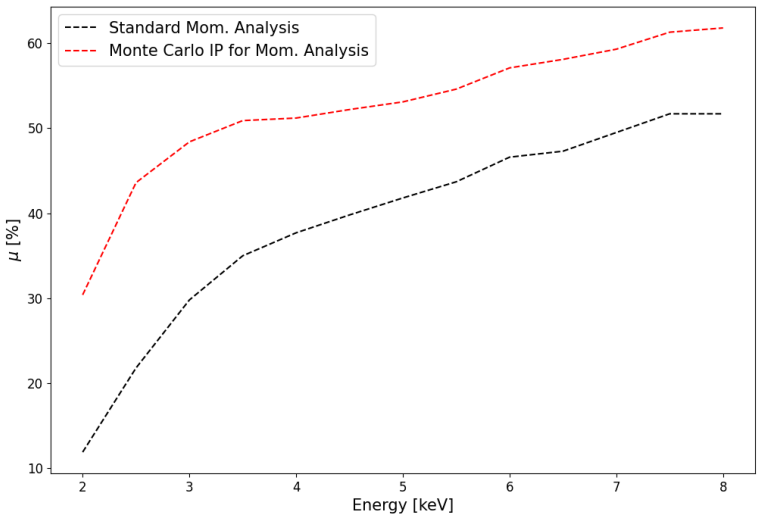

Figure 5.1: **Modulation factor as a function of energy.** Results for the standard moment analysis are represented by the *black dashed line*. Results achieved by the same analysis but employing the true impact point instead of the predicted one are shown with the *red dashed line*.

5.2.2 Polarization leakage

In Sec. 2.4 the polarization leakage effect was introduced. Some astrophysical sources exhibit an induced polarization pattern associated to intensity edges and gradients, due to the incorrect estimation of the impact point position from photo-electron tracks. This effect can be easily quantified when analyzing unpolarized point sources, because any detection of non-zero radial polarization would be due to the systematic effect of the polarization leakage. Nonetheless, due to the radially symmetric pattern that characterizes a point source image, the polarization derived from the whole observation is expected to be null, since any residual radial polarization would be cancelled by events separated by 180 degrees.

To detect the potential induced radial polarization in the simulated unpolarized point sources, each predicted emission angle is projected onto the radial direction. The radial-alignment process consists in aligning the reference axis used to evaluate the emission angle to the radial direction. A schematic view of the process is reported in Fig. 5.2.

The *residual modulation* $\mathcal{M}$ is then evaluated as:

$$\mathcal{M} = 2\sqrt{\widetilde{Q}^2 + \widetilde{U}^2} \tag{5.2}$$

where $\widetilde{Q}$ and $\widetilde{U}$ are the normalized radial-aligned Stokes parameters. $\mathcal{M}$ definition is similar to the modulation factor μ one, but it has a distinct physical interpretation.

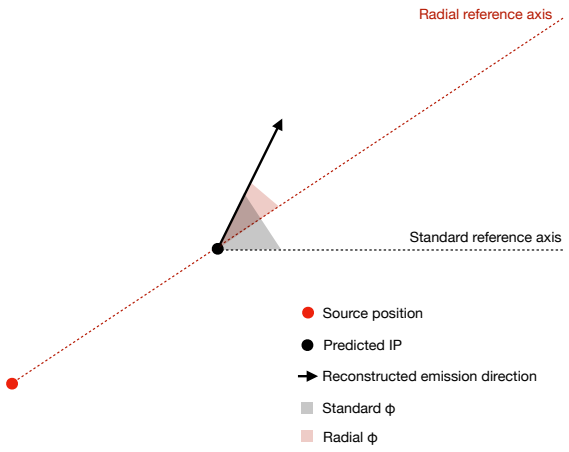

Figure 5.2: **Schematic view of the radial alignment process**. This is performed to evaluate the residual radial modulation of point sources, which is used to quantify the polarization leakage effect.

The modulation factor is specifically defined for a 100% polarized source, whereas the residual modulation is independent of the intrinsic polarization of the source. Moreover, $\mathcal{M}$ does not directly correlate with the polarization fraction, and it indeed does not take into account the modulation factor.

In Tab. 5.2 the residual radial modulation ($\mathcal{M}_R$) obtained for the simulated unpolarized point sources described in Sec. 5.1 with the standard moment analysis is reported.

Spectral Model	$\mathcal{M}_R$ [%]
Power Law, index -1.7 (PL1)	6.27 ± 0.31
Power Law, index -0.7 (PL2)	6.48 ± 0.26
Black Body, kT 3keV (BB)	6.58 ± 0.23

Table 5.2: **Summary of the residual radial modulation ($\mathcal{M}_R$) for the unpolarized point sources.**

Results in Tab. 5.2 show that, if we align the predicted emission angle to the radial direction, the moment analysis finds residual modulation for unpolarized sources up to $\sim 7\%$. To effectively illustrate this effect, the source images are spatially binned and normalized Stokes parameters Q/I and U/I are evaluated for each bin. This allows us to plot an interpolated binned graph, as shown in Figure 5.3, which reveals a distinct radial modulation pattern in the three unpolarized sources. By improving the quality of the impact point estimation this effect could be reduced, and improved accuracy of the measures of the polarization properties of the sources could be achieved.

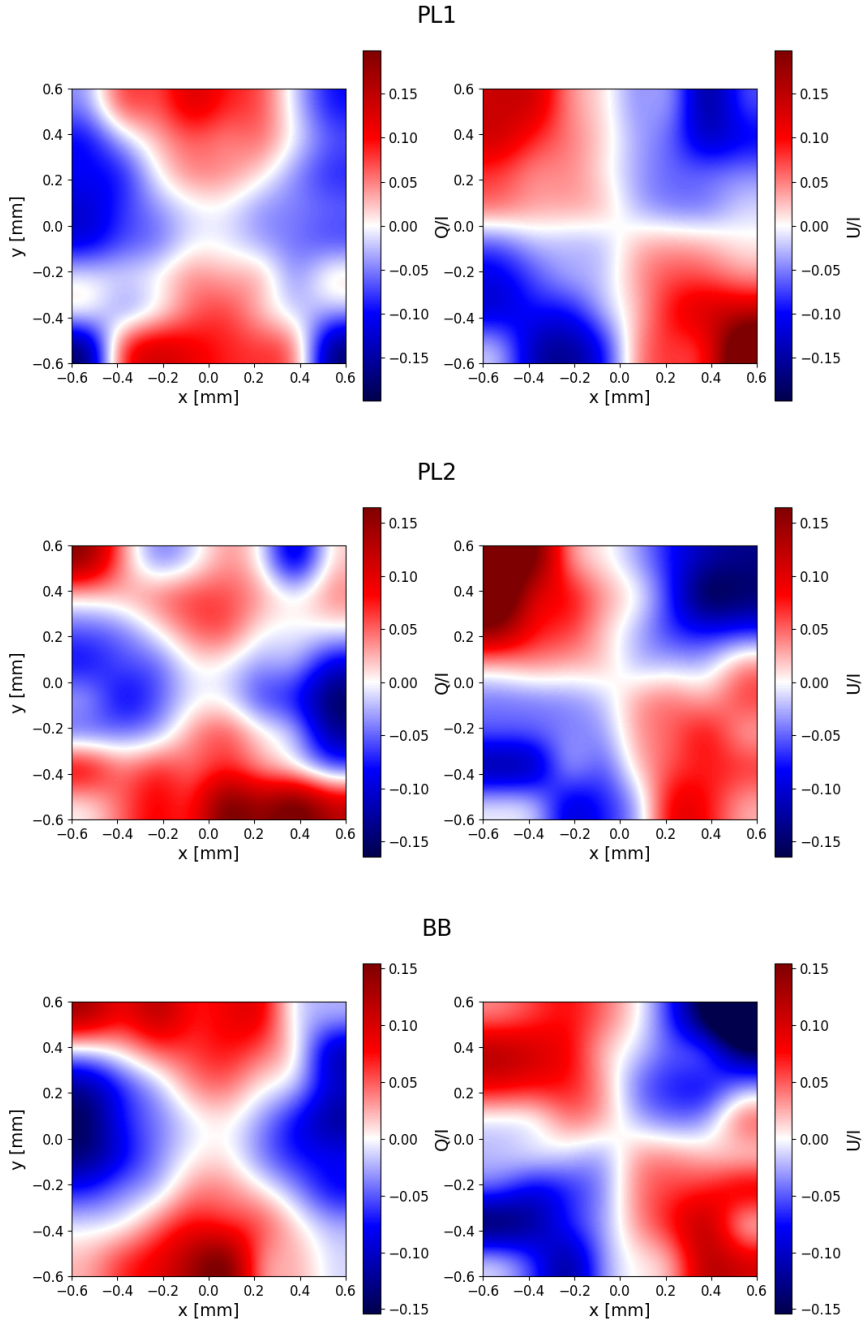

Figure 5.3: **Binned and interpolated calculation of Q/I and U/I.** Results are reported for the Power Law spectrum with index -1.7 source (PL1, *upper panel*), for the Power Law spectrum with index -0.7 source (PL2, *central panel*) and for the Black Body spectrum source (BB, *lower panel*). In all three cases, the source is located at the center of the GPD, $(x_S, y_S) = (0, 0)$. A radial polarization pattern results by these plots.

5.3 CNN structure

The previous section discussed about the importance of accurately reconstructing the impact point to generally improve the polarization measurements with GPDs. Additionally, as mentioned at the beginning of the chapter, CNNs have already shown promising results in reconstructing the PE tracks parameters.

Therefore, the primary objective of the approach that I have adopted and is discussed in the current section is to design a CNN that is specifically optimized for predicting the impact point location. The CNN-predicted impact point will be then introduced in the 4th step of the standard moment analysis in Eq. 2.8, with the purpose of both improving the general performance of the algorithm, i.e. enhancing the modulation factor, and of mitigating the polarization leakage effect.

As the proposed algorithm involves both a CNN-based and an analytic analysis, I will refer to it for the rest of this work as the *Hybrid* algorithm or method. In Fig. 5.4, the main steps of the hybrid method and its main difference compared to the standard moment analysis are represented.

5.3.1 Hexagonal convolution

We built a Convolutional Neural Network based on the DenseNet-121 [127] architecture, introduced in Chapter 4, which we modified to incorporate hexagonal convolutions at minimum performance cost. The hexagonal structure of the GPD anode pixels poses a significant challenge for ML-based analyses in this field. Convolution and pooling operations, which are commonly used with squared pixel images, are not directly applicable to hexagonal pixels, mainly because the CNN convolutional kernels are not spatially equivariant in hexagonal space. As a result, it becomes challenging to apply standard CNN structure to GPD tracks.

Steppa & Holch [134] developed a Python library extending the PyTorch deep learning framework with convolution and pooling operations on hexagonal grids. Unfortunately, employing this feature in our CNN would result in a significant slowdown of the algorithm. Peirson et al. [6] came up with a different approach: each track is pixel-shifted both by odd and even rows, generating two squared images for each track. To obtain a single training sample, the square conversion pair is stacked as done for color channels in an RGB image CNN problem, but with only two channels involved. This process used only cartesian convolutional layers, while avoiding polarization biases due to the hexagonal-to-squared conversion. However, equivariance in hexagonal space is still not fully achieved with this method.

Our solution aims to apply convolutional kernels which maintain the equivariance in hexagonal space, as for Steppa & Holch software, but which do not significantly compromise the computational time performance of the algorithm compared to the standard cartesian convolutional layers. Both these conditions are achieved by implementing hexagonal convolution and pooling layers as a C++ extension for PyTorch.

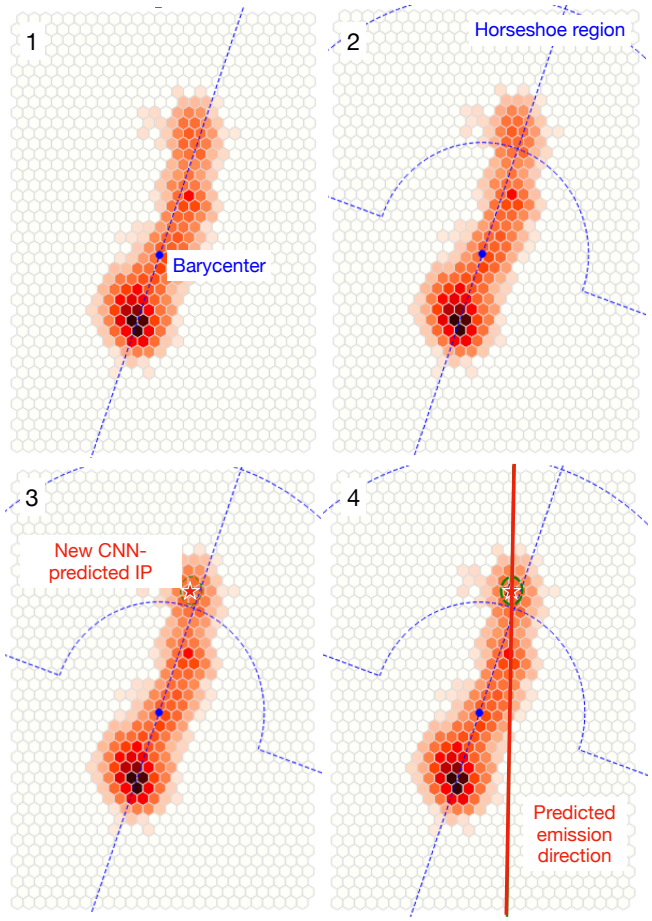

Figure 5.4: **Reconstruction of a photo-electron track with the hybrid method.** The four panels schematically represent the steps required by the hybrid algorithm to reconstruct the photo-electron emission direction. The main and only difference compared to the standard moment analysis is the employment of the CNN-predicted impact point, as highlighted in the third panel.

This particular section of the code was developed by the data scientist N. Moriakov, which collaborated to the developing of the algorithm during its initial stages. Specifically, the standard cartesian kernel is replaced by a hexagonal kernel. It consists of rectangular sub-kernels, each corresponding to a set of equal length columns of the hexagonal kernel, as represented by the white kernels in the left panel of Fig. 5.5. The single hexagonal convolution operation therefore consists in several cartesian convolutions with rectangular sub-kernels. In order to account for the shift that the original hexagonal image undergoes to match the Pytorch tensor structure, a complex scheme of padding and slicing of the input tensor is applied. The results of the sub-convolution operations are then merged and added, composing the final convolved hexagonal data in tensor format. A schematic representation of the hexagonal convolution operation

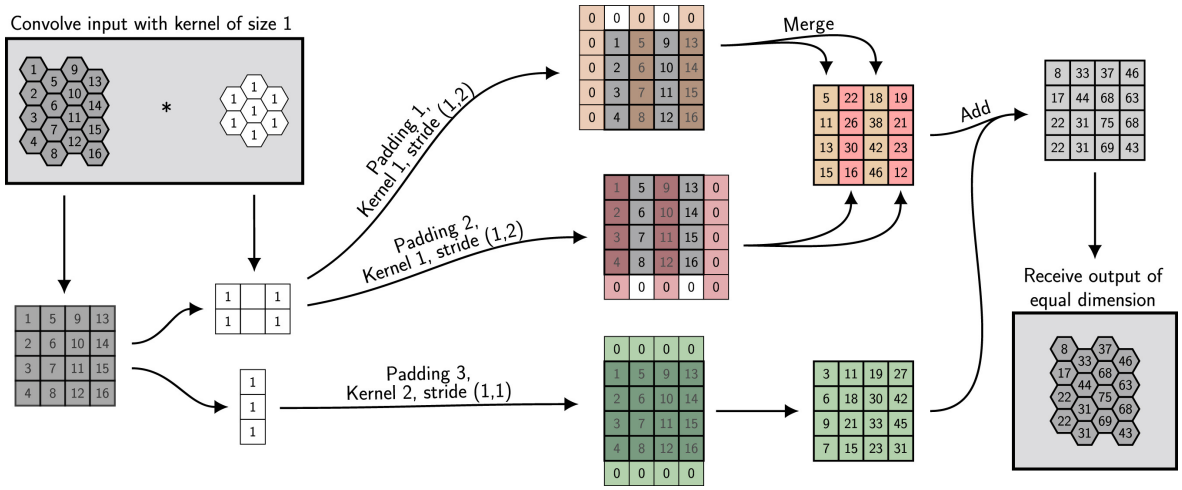

Figure 5.5: **Schematic view of the hexagonal convolution algorithm.** Image credits: [134].

is reported in Fig. 5.5, with the kernel dimension equal to 1. For hexagonal average pooling, a subgrid of the original grid with a stride of 2 is taken, and the average of the 7 values around each element of the hexagonal subgrid is computed. By relying on C++ PyTorch API to call the cartesian convolution subroutines, thus reducing the Python overhead, the hexagonal convolution operation results ∼4 times faster than the original python-based implementation, by using a *GPU NVIDIA Tesla V100* [136].

5.3.2 Architecture and Hyperparameters

General structure

Our CNN model adopts the DenseNet architecture with four dense blocks in its simplest configuration, denoted as DenseNet-121. The same DenseNet architecture was successfully applied to ImageNet images sized at 224x224 [122]. ImageNet dataset was also analyzed increasing the number of parameters achieving slightly improved performance. A smaller version of DenseNet architecture, with only 3 dense blocks, was employed for benchmark datasets consisting of smaller images, as CIFAR and SVHN (32x32 pixels) [127]. Given that GPD track images are commonly larger than the images of CIFAR and SVHN datasets, but smaller than ImageNet ones, DenseNet-121, the architecture consisting of four dense blocks and the lowest number of parameters, is anticipated to be well-suited for their analysis. A simple scheme of the network we built is reported in Fig. 5.6.

Hexagonal convolutional layers were used in the first convolutional block of the CNN architecture. The block consists of three stacked hexagonal convolutional layers, each with 64 filters, a stride of 1, and a kernel radius of 1. After each hexagonal convolutional layer, a batch normalization layer and a ReLU activation function are applied. For

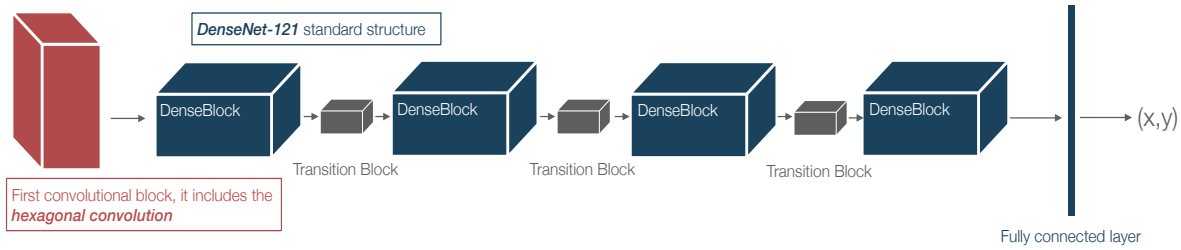

Figure 5.6: **Schematic view of the CNN of the hybrid algorithm**. The *red block* represents the hexagonal convolution layers, while the *blue and grays blocks* follow the standard Densenet-121 architecture [127].

reducing the computational time, we perform a transition from the hexagonal grid to the cartesian grid after the initial hexagonal convolutional block. This is done by applying a hexagonal convolutional layer on a cartesian subgrid of the hexagonal grid with a stride of 2. The transition convolution has a kernel radius of 1 and 64 filters, allowing it to gather image features from the entire hexagonal grid. A batch normalization layer and a ReLU activation function are applied after the transition convolution too. Subsequently, we switch to the standard DenseBlocks of the DenseNet-121, which consist of only cartesian 2D convolutions, as described in Sec. 4.2. A fully connected layer for impact point regression is applied to the final CNN feature map, which has dimensions of 6×6. Dropout layers with a dropout probability of 10% are used throughout the network inside the DenseBlocks.

A major and time-consuming aspect to take into account while building the model is the choice of the *hyperparameters* of the network. We use Adam optimizer as it is commonly used in similar regression problems, and we define the following loss function:

$$L(x_{true}, y_{true}|x_{pred}, y_{pred}) = |(x_{true}, y_{true}) - (x_{pred}, y_{pred})| \tag{5.3}$$

As for the choice of the learning rate and the number of epochs, we perform a fine-tuning by traning multiple times the CNN with sub-samples of the traning set, and testing its learning curve by looking at the loss function and the performance on the validation set. The final network is trained for 60 epochs, with a decaying learning rate starting from 10^{-4}. Moreover, in order to improve the general performance of the network, a form of Online Hard Example Mining (OHEM) [135] is employed from the 30th epoch of the training phase, i.e. in each batch only 50% of samples with worst performance is taken into account during gradient computation. The training phase of the network was run on a single *GPU NVIDIA Tesla V100* [136], and required ∼15 days to complete the process.

Output nodes

Initially, the output nodes of the network were configured to represent the impact point position on the (x, y) plane, approximated to the nearest integer pixel value. Accordingly, we observed that as the network trained over several epochs, it began to consistently predict the IP position as an integer or half-integer value. However, this represents a limitation in the CNN performance and could introduce potential biases in the results. As an example, in the left panel of Fig. 5.7 the distance between the true and the predicted impact point after the training of the network is reported, in comparison with the distance resulting from the standard moment analysis.

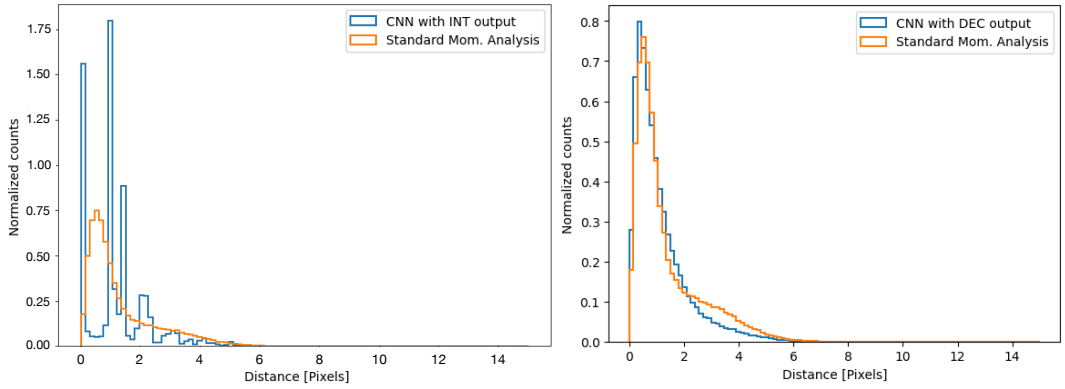

Figure 5.7: **Distance between the true impact point and the reconstructed one.** *Left panel*: The *blue histogram* refers to a CNN trained to reconstruct the IP to the nearest integer value. The *orange histogram* refers to the standard moment analysis. *Right panel*: The *blue histogram* refers to a CNN trained with decimal pixels as output. The *orange histogram* refers to the standard moment analysis. The test sample is an unpolarized 3.5 keV energy beam.

To avoid this potential bias, the output of the network is set as an impact point position in decimal pixels. Reconstructed-true IP distance for the decimal pixels-CNN and standard moment analysis are shown in the right panel of Fig. 5.7.

5.3.3 Input images

As stated in the previous sections, the network is trained with images of tracks generated through the *ixpesim* simulation framework. Each image is associated to a matrix that represents the pixel-wise counts, reflecting the energy loss of the photo-electron during its path within the GPD gas. An example of a raw image and the corresponding pixel values are reported in Fig. 5.8. Several operations were performed on the input images before using them to train the network.

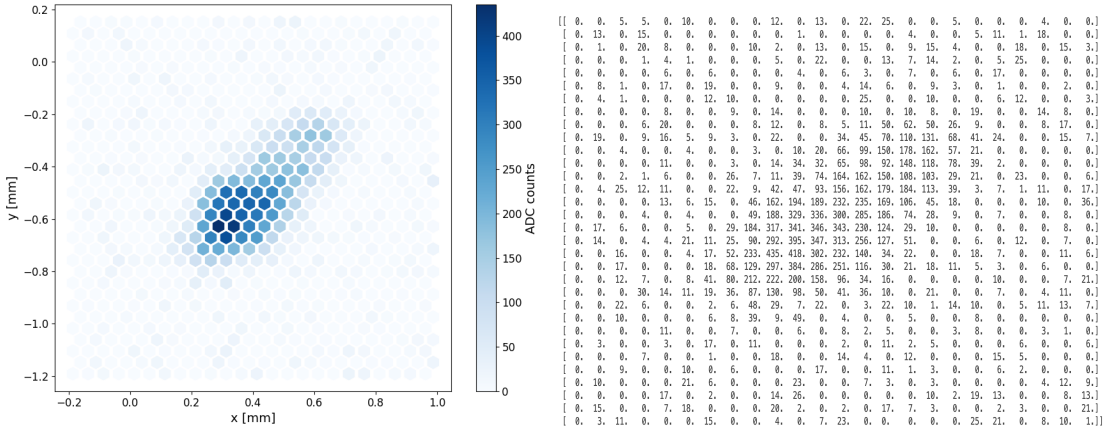

Figure 5.8: **Example of a simulated photo-electron track during its pre-processing phase (1).** *Left panel*: Track image before the pre-processing phase. *Right panel*: matrix of the pixel values of the image.

A) Pixel clipping

The initial step consists in removing from the training set any tracks that do not correspond to a valid physical event. For each image, all the pixels with a number of counts lower than 20 are clipped to 0, as such a few counts are most likely generated by electric noise (20 ADC counts $\sim$ 45 electrons) [17]. Subsequently, only tracks with more than 3 pixels with ADC counts higher than 0 are then retained within the training set, as for lower numbers of pixels it would be impossible to obtain meaningful results, besides being events most likely generated by instrumental background.

As noticeable in the right panel of Fig. 5.8, even after the 20 ADC counts clipping, some pixels disconnected from the photo-electron track may still exhibit non-zero (i.e. $>$ 20) ADC count values. These counts are presumably generated by electric noise as well and could potentially impact the performance of the CNN. A clustering algorithm based on the DBSCAN model [137] is thus applied, being it a very popular model designed to recognize clusters of arbitrary shape.[1] In Fig. 5.9 the result of this second preprocessing phase is reported for the same track of Fig. 5.8.

Concluding, the 20 ADC counts threshold and the clustering algorithm clip all the pixels not belonging to the main photo-electron track to zero.

B) Artificial sharpening

A key feature of the pre-processing phase of the tracks is the artificial sharpening of the images. As the aim of the CNN is to precisely locate the photon IP on the detector plane, an algorithm for increasing the resolution of the images, while keeping the hexagonal symmetry of the pixel matrix, is introduced.

[1] The same procedure is also applied in the standard moment analysis

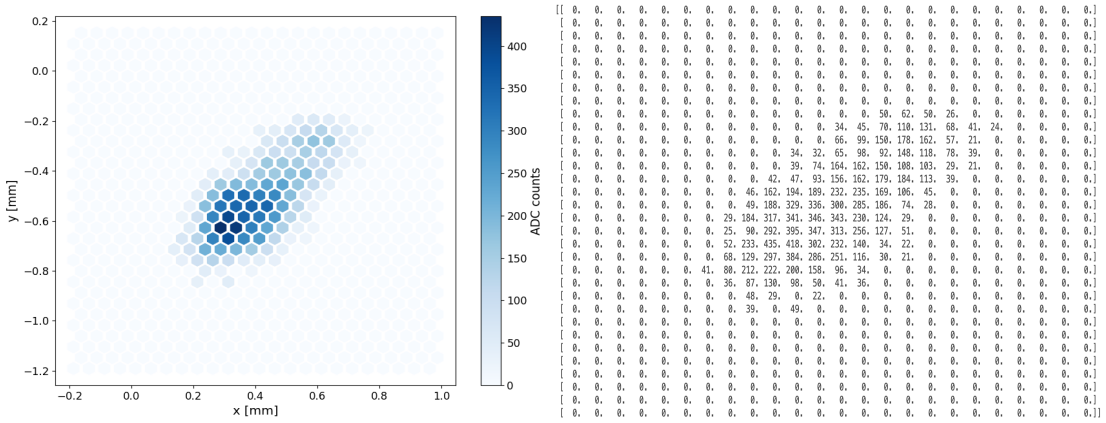

Figure 5.9: **Example of a simulated photo-electron track during its pre-processing phase (2).** *Left panel*: track image after the pixel-clipping and clustering process. *Right panel*: matrix of the pixel values of the image.

For each couple of adjacent pixels of the original image, a new pixel is added halfway between them, and its ADC counts value is calculated as the linear interpolation of the two neighbour pixels. This procedure preserves some important physical parameters of the original image. Firstly, as previously mentioned, the resulting image remains a hexagonal matrix of pixels, but with a finer pixel pitch. Specifically, the horizontal distance between pixels is reduced from a pitch of 50 μm to 25 μm, while the vertical pitch decreases from 43.3 μm to 21.65 μm.

As the information about the IP location is in millimeters in the simulations, and in number of pixel for the CNN output, it is crucial to establish a mapping between these two coordinate systems and to adjust it accordingly when increasing the images resolution. The two coordinate systems on a schematic view of the readout plane are represented in Fig. 5.10.

By utilizing information such as the number of rows (N_{row}) and columns (N_{col}), as well as the horizontal (p_{row}) and vertical (p_{col}) distances between the centers of the hexagonal pixels, it is possible to establish a mapping that converts the pixel coordinates (i, j) into the corresponding position in millimeters (x, y):

$$x = [i - \frac{1}{2}(N_{\text{col}} - \frac{3}{2} + j \text{mod} 2)] \cdot \text{p}_{\text{col}} \tag{5.4}$$

$$y = [\frac{1}{2}(N_{\text{row}} - 1) - j] \cdot p_{\text{row}}. \tag{5.5}$$

In order to account for the sharpening process of the images, it is possible to change the variables describing the number of rows and columns, and their respective pitches:

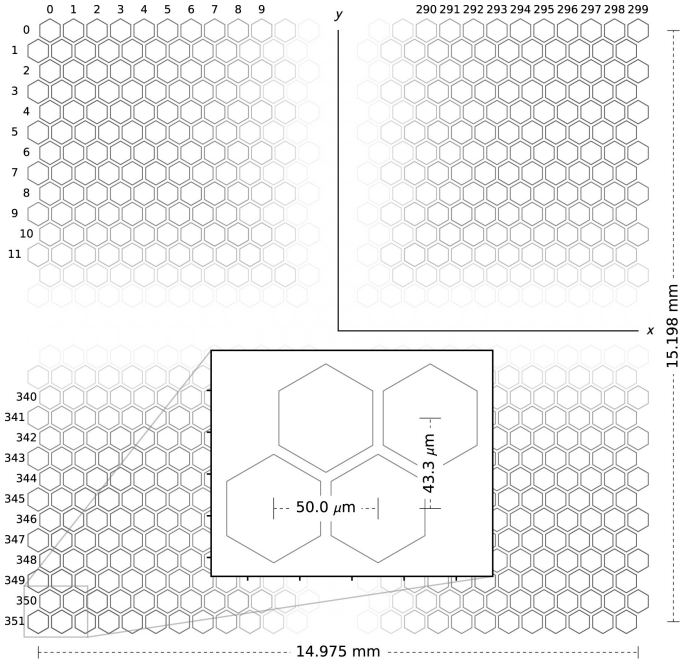

Figure 5.10: **Scheme of the readout plane of the GPD.** The (x, y) coordinates system has its origin in the center of the GPD, while the numbers on the borders of the pixel matrix refer to the (i, j) coordinates system. The horizontal and vertical distance between the center of the pixels is reported too. Image credits: [17]

$$\begin{aligned} N_{\text{col}} \longrightarrow N'_{\text{col}} = 2N_{\text{col}} \quad & N_{\text{row}} \longrightarrow N'_{\text{row}} = 2N_{\text{row}} \\ p_{\text{col}} \longrightarrow p'_{\text{col}} = \frac{1}{2}p_{\text{col}} \quad & p_{\text{row}} \longrightarrow p'_{\text{row}} = \frac{1}{2}p_{\text{row}} \end{aligned} \tag{5.6}$$

$$i \longrightarrow i' = 2i - 1 \quad j \longrightarrow j' = 2j. \tag{5.7}$$

The (x, y) position on the readout plane with respect to the new quantities is:

$$x = [i' - \frac{1}{2}(N'_{\text{col}} - 1 + j' \text{mod} 2)] \cdot \text{p}'_{\text{col}} \tag{5.8}$$

$$y = [\frac{1}{2}(N'_{\text{row}} - 2) - j'] \cdot p'_{\text{row}}. \tag{5.9}$$

The second important physical parameter is the information on the total charge deposited in the track, that is preserved, as the total counts in the original images are related to the new ones as:

$$\text{PHA}_{\text{old}} = \frac{1}{4}\text{PHA}_{\text{new}}. \tag{5.10}$$

The CNN in this work would not be affected by the non-conservation of the charge. However, this characteristic may hold significance if the sharpening process is employed in different contexts. In Fig. 5.11 the same track of Fig. 5.8 and 5.9 is reported after the sharpening process.

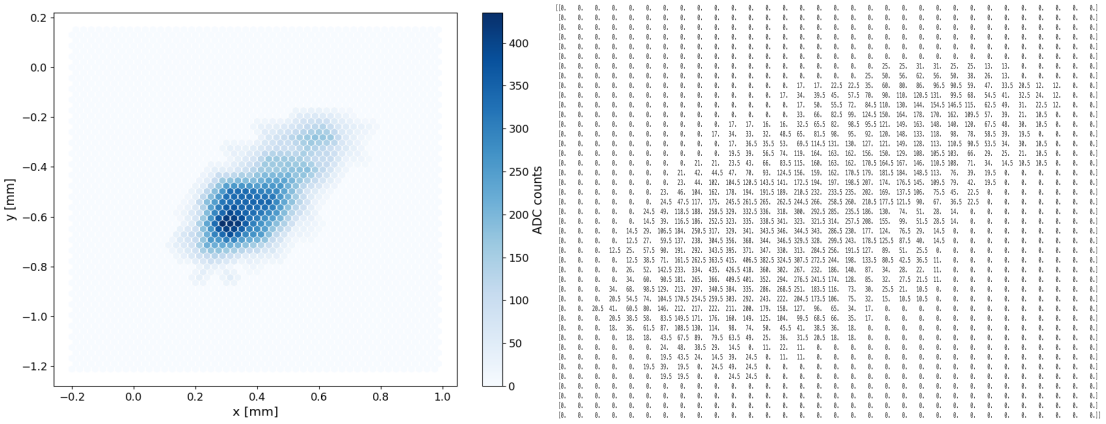

Figure 5.11: **Example of a simulated photo-electron track during its pre-processing phase (3).** *Left panel*: Track image after the artificial sharpening process. *Right panel*: matrix of the pixel values of the image, zoomed around the track.

C) Image cropping

Before feeding images into the network, an essential factor to consider is their dimensions. In the simulation files, as also in IXPE data, each image size is correlated with the track size. However, the CNN needs uniform image dimensions, preferably square-shaped (N×N). Choosing the appropriate image dimension N for input to the CNN involves finding a balance between capturing the necessary information in the images and managing computational resources. Firstly, it is necessary to determine the maximum track length which is needed to be captured in the images, which depends on the energy of the incident photons. A longer track, generated by a higher energy photon, will require a larger image dimension to ensure it is fully contained. At the same time it is crucial to evaluate the available computational resources too, such as memory, processing power, and time constraints. Larger image dimensions require more memory for storage and processing, and they can increase computational time.

After fine-tuning the parameter N, the image dimensions is set to 72×72 pixels. However, at very high energies, not all tracks are fully contained within the image. Figure 5.12 shows the MC impact points for 100 000 tracks generated by 7.5 keV photons. The red square represents the 72×72 pixel frame, and approximately 10% of the impact points are not contained in the image. In such cases, we still consider the image for the network training phase, providing the information about the impact point to the network with negative pixel values or values higher than 72.

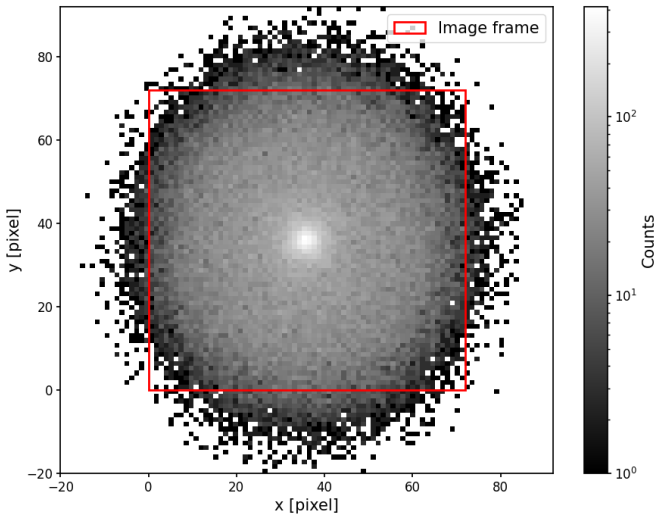

Figure 5.12: **True impact points distribution for a 7.5 keV energy beam.** The IP positions are reported as (x_{pix}, y_{pix}) on a 72×72 pixels image (*red square*).

After training the network, we carry out a test to evaluate its performance with tracks not fully contained in the image. If an image does not contain the true impact point within the 72×72 frame, the network is still able to reconstruct the impact point for most images outside the frame, although with lower accuracy. For impact points inside the frame, the mean distance between the true and reconstructed impact points is approximately 1.6 pixels, whereas for impact points not contained in the image, it is approximately 4.5 pixels.

To mitigate the issue of impact points not being contained within the frame, we introduce a new feature during the testing phase of the network: each test image is rotated by 60° and 120°[2], for each image the impact point is predicted, and finally the prediction is rotated back. The final reconstructed IP position is the mean between the three predictions. This process serves two purposes. Firstly, three different predictions for the same image could enhance the precision of impact point reconstruction and mitigate potential biases. Secondly, although the original image may not fully contain the track and the photon impact point, after at least one of the two rotations the track might be fully contained within the image. In this case, out of the three images generated for each event, only the ones with the impact point inside the frame are considered for calculating the reconstructed impact point. By applying this procedure, the percentage of tracks with the impact point included in at least one image frame for the 7.5 keV beam increases from approximately 90% to 96%.

[2]For these angles the hexagonal symmetry of the images is preserved.

Finally, the pixel values are normalized between 0 and 1. Although this normalization process results in the loss of information regarding the total deposited charge, it is not relevant for reconstructing the impact point location.

5.4 Results

Once the network is trained, its performance are tested with the test samples described in Sec. 5.1. The CNN takes 1 second of GPU time to process approximately 200 tracks (3 images per track, 1 for each rotation angle). Compared to the standard reconstruction, it would mean that for each track an additional processing time of $\sim 5 \cdot 10^{-3}$ s must be considered when using the CNN-predicted impact point.

Firstly, we evaluate the quality of the reconstruction of the impact point location, comparing the results to the standard moment analysis. Secondly, we evaluate the performance of the hybrid algorithm on the polarization measurements, by introducing the CNN-predicted impact points in Eq. 2.8.

5.4.1 Impact point reconstruction results

We test the impact point reconstruction performance with the 13 mono-chromatic unpolarized samples, in order to evaluate the performance as a function of energy in the 2-8 keV range. The polarization of these beams is not relevant for this study, as the figures of merit refer to the impact point reconstruction only. In order to evaluate the repercussions of the artificial sharpening process we also train a CNN with standard images, i.e. with not-resolution-augmented images (48×48 pixels frame), and we compare the results for the two CNNs and for the standard moment analysis.

Three different figures of merit are introduced to estimate the impact point reconstruction performance. Firstly, the mean distance between the predicted and true impact point is evaluated. Moreover, the percentage of events for which the distance between the true IP position and the predicted one is smaller than one pixel and two pixels are calculated too. The unit *pixel* we report in these figures of merit indicates the standard pixel dimension, not the sharpened one, for both the moment analysis and the two CNNs. The results for the three figures of merit are presented in Fig. 5.13: the *black line* corresponds to the standard moment analysis, the *orange line* corresponds to the CNN trained with not-sharpened images (Standard CNN), the *red line* illustrates the results for the CNN trained with sharpened images (Sharpened CNN), and the *black dashed line* indicates the mean distance between the true IP and the barycenter of the track.

The comparison between the Standard and Sharpened CNN demands some considerations. Firstly, the *upper panel* in Fig. 5.13 shows that the mean distance between the reconstructed and true IP is lower for the Sharpened CNN, for energies up to 7 keV. The *central and lower panel* curves support this result: as an example, for the

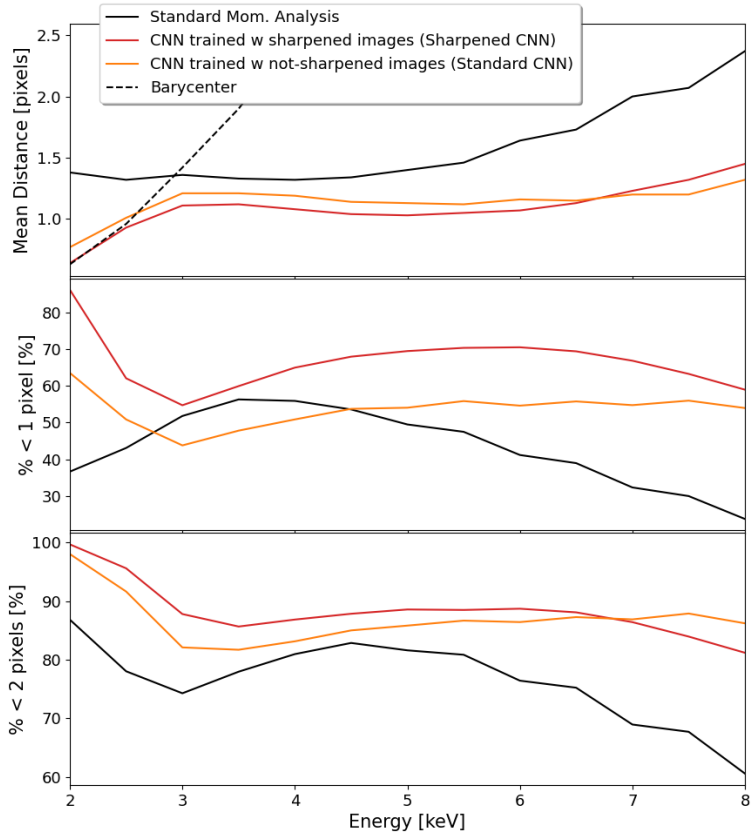

Figure 5.13: **Impact point reconstruction results.** Comparison of the IP position reconstruction among the sharpened-images CNN (*red line*), a not-sharpened-images CNN (*orange line*) and the moment analysis (*black line*). In the *top panel*, the mean distance between the true IP position and the predicted one is reported. The *black dashed line* represents the distance between the true IP and the barycenter of the track. In the *middle and bottom panel*, the percentage of events for which the distance between the true and the predicted IP position is lower than respectively 1 and 2 pixels is reported.

Sharpened CNN at 3 keV(6 keV) the $\% < 1$ *pixel* is up by $\sim 11\%(16\%)$ and the $\% < 2$ *pixel* is up by $\sim 6\%(2\%)$. However, it is important to note that the Standard CNN overcomes the performance of the Sharpened CNN for energies higher than approximately 7 keV in two out of three figures of merit. This reversal in trend is attributed to the sharpening process that causes an increased number of very long tracks to be cropped. As an example, as it was also reported in the previous section, at 7.5 keV approximately 96% of the true impact points are contained within the 72×72 frame for sharpened images, while this percentage exceeds 99% for the 48×48 frame used with not-sharpened images. Although this issue could be addressed by extending the frame dimension to accommodate longer tracks, it would inevitably slow down the algorithm. Given that IXPE's effective area is considerably low at very high energies, prioritizing algorithm speed is deemed more crucial. Nonetheless, if this approach needs to be applied in similar applications, careful attention should be given to this aspect.

Regarding the comparison between the CNN and the moment analysis, all the figures of merit show a significant improvement in the reconstruction of the IP location performed by the CNN. It should be noticed how between 3 and 4 keV training the CNN with not-sharpened images would not allow to improve the performance in the $\% < 1$ *pixel* compared to the standard moment analysis. However, this is achieved with the sharpened images. For example, the Sharpened CNN compared to the standard moment analysis at 3 keV(6 keV) has the $\% < 1$ *pixel* up by $\sim 3\%(29\%)$ and the $\% < 2$ *pixel* up by $\sim 13\%(12\%)$. It is interesting to notice how the $\% < 2$ *pixel* for the CNN is consistently higher than 80%, showing how, differently from the standard moment analysis, the network is also accurate in the identification of the IP position when the tracks become longer. This feature is key in order to reduce the systematic induced by the polarization leakage effect.

The *top panel* of Fig. 5.13 also provides information about the average distance between the track barycenter and the true IP position. For tracks with very low energy, the true IP position is usually very close to the barycenter, while the moment analysis predictions tend to place the IP towards the outer edges of the track. Consequently, in the standard moment analysis, substituting the predicted IP with the barycenter could enhance the precision at low energies, as shown in the top panel of Fig. 5.13. However, when working with data that lack MC information, as laboratory or IXPE data, determining the energy threshold from which the barycenter should be used as the predicted IP is not straightforward, and thus, it is not employed in the standard analysis. On the other hand, the CNN automatically tends to place the predicted IP close to the barycenter at low energies.

Fig. 5.14 reports an example of the reconstructed IPs 2D distributions around the true ones for the 5.5 keV unpolarized energy beam, for both the standard moment analysis on the *left panel* and the Sharpened CNN on the *right panel*. This illustration underlines the absence of biases in determining the IP position along either axis for the hybrid method. The same outcome is achieved with all the energy beams. Moreover, once more, the enhancement in the IP reconstruction precision achieved with the CNN

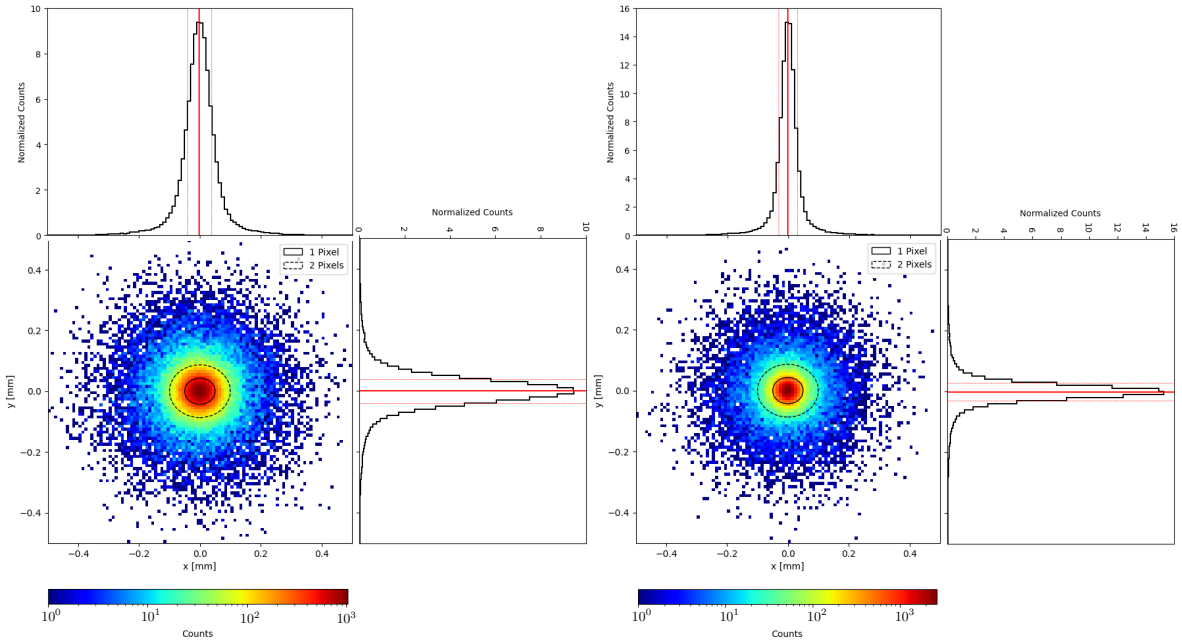

Figure 5.14: **2D distributions and relative x and y projections of the reconstructed impact point around the true one for the 5.5 keV energy beam.** *Left panel* reports the results for the standard moment analysis, whereas *right panel* the hybrid method ones. The *solid (dashed)* black lines in the 2D histograms enclose the events included within 1(2) pixel(s). The *red solid lines* in the 1D histograms represent the mean of the distributions, whereas the *red dashed lines* the Full Width Half Maximum (FWHM).

is highlighted.

5.4.2 Polarization estimation results

In this section, the results regarding the polarization estimation performance are reported. Specifically, the data set made up of 13 mono-chromatic polarized samples is used to evaluate and compare the modulation factor as a function of energy, while the three unpolarized simulated point sources are used to evaluate the performance of the algorithms in case of unpolarized radiation and to detect potential systematic uncertainties.

All these events are passed through the CNN, which assigns an impact point to each of them. Subsequently, this CNN-predicted impact point replaces the one predicted by the standard moment analysis in Eq. 2.8 to evaluate the emission angle for the hybrid method. All the results are compared to the ones obtained with the standard moment analysis.

Polarized radiation

The figure of merit taken into account to evaluate the capability of an algorithm to correctly reconstruct the polarization degree of the incident radiation is the modulation factor μ. For each mono-chromatic polarized beam, μ is evaluated and the results as a function of energy are reported in Fig. 5.15.

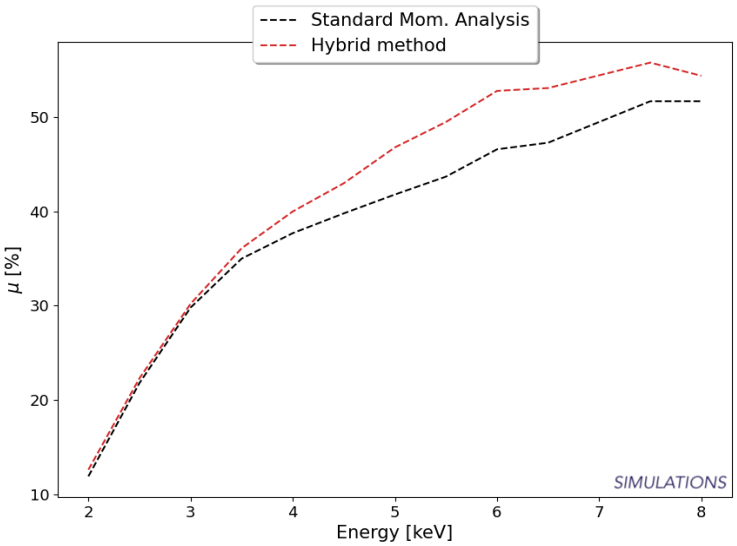

Figure 5.15: **Modulation factor (μ) as a function of energy for the simulated polarized beams**. Results for the standard moment analysis (*black line*) and for the hybrid method (*red line*) are reported.

Despite the substantial improvement in the reconstruction of the IP position, the improvement on the modulation factor amounts to only 1% up to 3.5 keV and increases up to a maximum of $\sim 6\%$ at 6 keV. This means that at low energies the new CNN-predicted impact point is not close enough to the true one to allow a significant improvement in the reconstruction of the emission angle. An even higher precision would be needed to reach the upper μ values (see Fig. 5.1): given the nearly circular symmetry of the tracks, a sub-pixel shift in the predicted impact point leads to a significantly different emission angle reconstruction. At higher energies, even if the algorithm is overall less precise compared to the lower energies, the improvement in the IP reconstruction permits to obtain significant improvement in the reconstruction of the emission direction, which determines higher modulation factors compared to the standard moment analysis. Being the tracks more elongated at higher energies, a $\sim$1 pixels shift of the predicted IP with respect to its true position does not largely affect the emission angle reconstruction.

Unpolarized radiation

The three simulated unpolarized point sources are used to evaluate the algorithm performance in case of unpolarized radiation. Unlike the data set of unpolarized monochromatic beams, these simulated sources offer a more realistic evaluation of the algorithms response to polarization leakage, as they mimic typical spectra observed in astrophysical sources.

Firstly, we study the resulting emission angle distribution as a function of the azimuthal angle ϕ. This analysis is performed to check for potential biases in the reconstruction algorithm, which would result in a non-flat angle distribution and/or in a non-zero residual modulation. In Fig. 5.16 the distribution of the emission angles is reported for the three different sources, for both the hybrid algorithm in red and for the standard moment analysis in black, alongside with the table of the residual modulation values (Tab. 5.3).

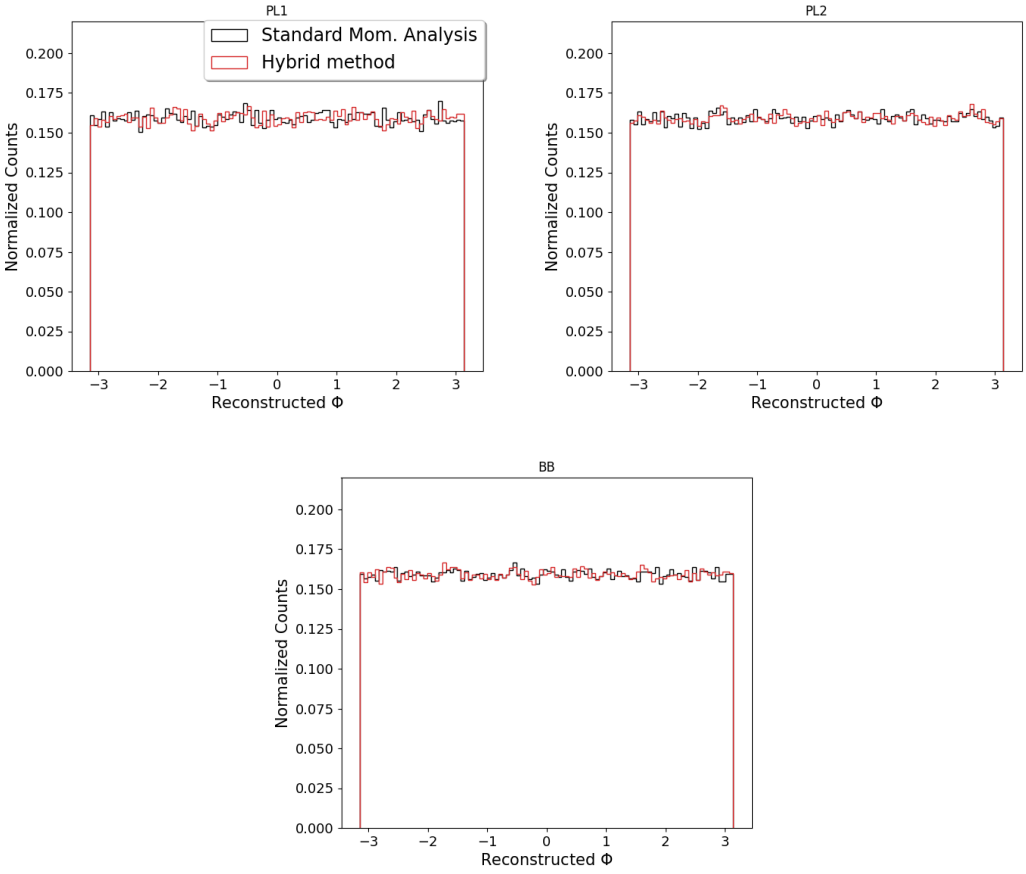

Figure 5.16: **Emission angles distributions for the three simulated unpolarized point sources.** Results are reported for the hybrid method (*red histogram*) and for the standard moment analysis (*black histogram*).

Spectral Model	$\mathcal{M}_L$ [%]	
	Mom. Analysis	Hybrid
Power Law, index -1.7 (PL1)	0.19 ± 0.31	0.15 ± 0.31
Power Law, index -0.7 (PL2)	0.22 ± 0.26	0.17 ± 0.26
Black Body, kT 3keV (BB)	0.13 ± 0.23	0.18 ± 0.23

Table 5.3: **Summary of the linear residual modulation ($\mathcal{M}_L$) for the unpolarized point sources**. The results show no significant residual polarization, as all the modulation values are compatible with zero within 1σ.

Fig. 5.16 and Tab. 5.3 demonstrate that both the algorithms do not detect residual linear modulation $\mathcal{M}_L$. All the modulation values are compatible with zero within 1σ, and ϕ distributions of all the simulated sources do not show any unexpected features.

However, measuring the linear modulation only is not enough to completely evaluate the possible systematics linked to the response of the algorithm to unpolarized photons, because, as depicted in Sec. 5.2.2, in order to evaluate the polarization leakage effect with point sources, an alignment of the predicted emission angles with respect to the radial direction is required. For each source we perform the alignment and calculate the radial residual modulation $\mathcal{M}_R$. In Fig. 5.17 and in Tab. 5.4 the results with their uncertainties for the standard moment analysis, for the same analysis but employing the barycenter as the predicted IP under 3 keV (see Sec. 5.4.1), and for the hybrid method are reported.

Spectral Model	$\mathcal{M}_R$ [%]		
	Mom. Analysis	Mom. Analysis + bary	Hybrid
Power Law, index -1.7 (PL1)	6.27 ± 0.31	5.22 ± 0.31	3.53 ± 0.31
Power Law, index -0.7 (PL2)	6.48 ± 0.26	5.62 ± 0.26	3.55 ± 0.26
Black Body, kT 3keV (BB)	6.58 ± 0.23	6.16 ± 0.23	4.14 ± 0.23

Table 5.4: **Summary of the radial residual modulation ($\mathcal{M}_R$) of the three unpolarized point sources.**

As already discussed in Sec. 5.2.2, the standard moment analysis finds residual modulations that deviate significantly from zero, even when the simulated point sources are unpolarized. By employing the barycenter as the IP for low energies it is possible to slightly mitigate this effect, but as already mentioned in Sec. 5.4.1, it is not a trivial substitution to perform when handling data without the Monte Carlo information about the photon energy, and in any case it is not resolutive. By employing the hybrid method, we can achieve residual modulation values which are significantly mitigated, by a factor of approximately ~ 2, compared to the standard moment analysis.

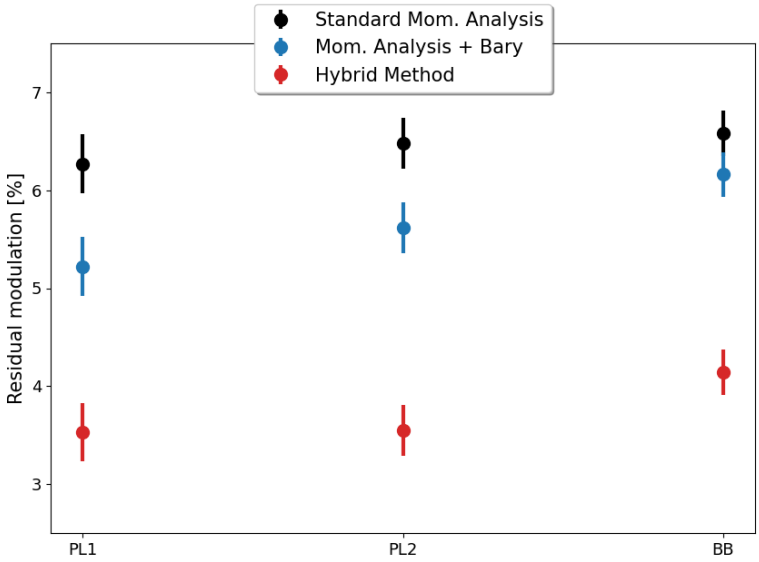

Figure 5.17: **Radial residual modulation for the simulated unpolarized point sources.** Results are reported for the standard moment analysis in *black*, the moment analysis that employs the barycenter as IP prediction in *blue*, and our hybrid method in *red*.

Results vary slightly depending on the source spectra. Firstly, the performance disparity between the standard moment analysis and the hybrid method changes with energy, and the accuracy of the IP reconstruction is the foundation of the polarization leakage effect. Additionally, polarization leakage is induced by photons across the whole energy range. At low energies the barycenter and the CNN-predicted IP closely align (refer to Fig. 5.13), but in any case employing the CNN-predicted IP results in a more pronounced reduction of the polarization leakage. Above 3 keV the magnitude of this systematic effect increases and widens the difference between the residual modulations measured with the two methods.

An additional analysis can be performed to prove that the residual modulation is linked to the polarization leakage effect. For this discussion, only the data of the Black Body unpolarized source are considered, but similar results are obtained with the other two point source simulations.

We select only the events for which the predicted IP is contained in the portion of GPD represented in the left panel of Fig. 5.18. For these events, the distribution of the emission angles (not radial-aligned) is reported in the right panel of Fig. 5.18. The resulting distribution is the signature of the polarization leakage effect: for a *real* polarized signal, the probability of a photo-electron to be emitted in the azimuthal direction ϕ is equivalent to the probability of being emitted in the $\phi \pm \pi$ direction

($d\sigma/d\Omega \propto \cos^2\phi$, see Sec. 1.2.3). Consequently, the distribution function of the azimuthal angles of the tracks would be characterized by a sinusoidal shape (m=2 modes). On the other hand, the tracks which cause the polarization leakage signal tend to align towards the center of the source. This results in the distribution function of the azimuthal angles of the tracks being characterized by an m=1 mode. Considering the GPD portion in Fig. 5.18, the polarization leakage effect would produce an excess of predicted directions which form an angle of $\sim -112°$ respect to the reference axis, as depicted by the red arrow reported in the left panel. Right panel shows indeed a peak around $\sim -112°$, which is more marked for the standard moment analysis compared to the hybrid method.

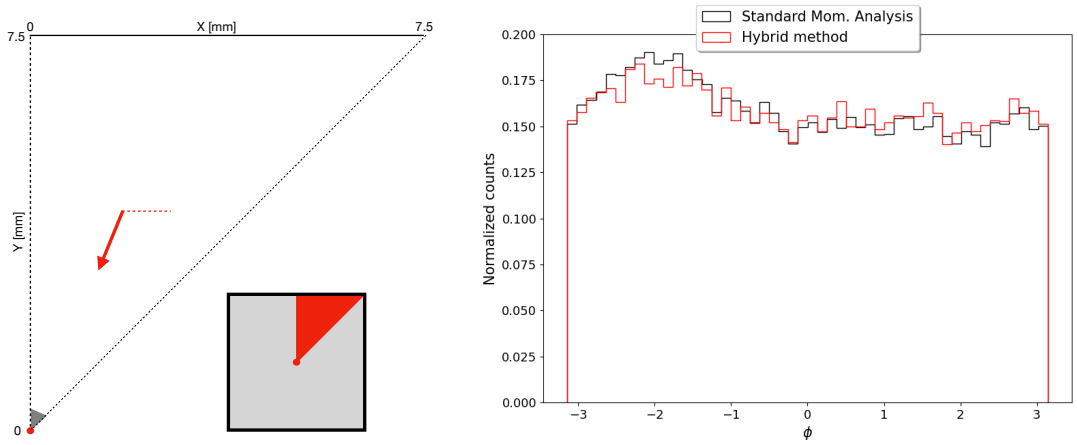

Figure 5.18: **Emission angle distribution for a specific region of the GPD**. *Left panel*: Portion of the GPD considered for selecting the events which generate the distribution in the right panel. The *red dot* is the source position $(x_S, y_S) = (0, 0)$. *Right panel*: Distribution of the emission angle for the events for which the impact point is reconstructed in the GPD portion represented in the left panel. Reconstructed ϕs for standard moment analysis are reported in *black*, while the ones for the hybrid method are reported in *red*.

Extended source

As discussed earlier in this chapter, measuring polarization leakage with point sources is an useful tool for quantifying the effect. However, in actual observations of point sources, polarization leakage is often not a concern. This is because we rarely focus on the radial polarization of such sources, and even if we did, the effect would be easily correctable. Dealing with extended sources, on the other hand, presents a greater challenge. The non-trivial geometry of emission areas, coupled with irregular intensity edges and gradients, complicates the handling and prediction of polarization leakage signal.

We conducted a simulation of a 92 ks observation of an unpolarized source with intensity and properties similar to the Crab Nebula with a single IXPE detector. The simulation is based on the *Chandra X-ray Observatory* observation of the nebula [138]. This means that, as the source is unpolarized, every significant sign of residual polarization is due to the polarization leakage effect. This procedure is quite standard when analyzing extended sources, because it permits to quantify the influence of the leakage on the observed polarization. In Fig. 5.19 the Q/I and U/I binned interpolated values (*upper and middle panel*), and the binned polarization fraction (*lower panel*) for the standard moment analysis in the *left panels* and for the hybrid method in the *right panels* are reported. In the lower panels, only the bins with a significant detection ($>3\sigma$) are represented.

Examining the outcomes obtained with the standard moment analysis, the consequences of the polarization leakage effect are clearly detectable. The Q/I and U/I plots exhibit polarization patterns similar to those observed in the case of point sources (refer to Fig. 5.3), and the lower panel reports bins displaying a significant polarization degree up to approximately 25% in the external region of the nebula. On the other hand, the mitigation of the polarization leakage achieved by the hybrid method, and discussed for the point sources in the previous section, has a substantial impact on the observation of the extended sources too. The Q/I and U/I plots reported in the right panel of the same figure show in fact a reduced manifestation of polarization leakage, and the bottom panel features only four bins (instead of 67) where a significant measure of polarization is detected, with reduced intensities (the highest polarization fraction achieved with the hybrid method is 14%).

These results highlight the potential of this new algorithm on the observation of some extended sources: the hybrid method could in fact enhance the significance of the signal, leading to a more robust detection.

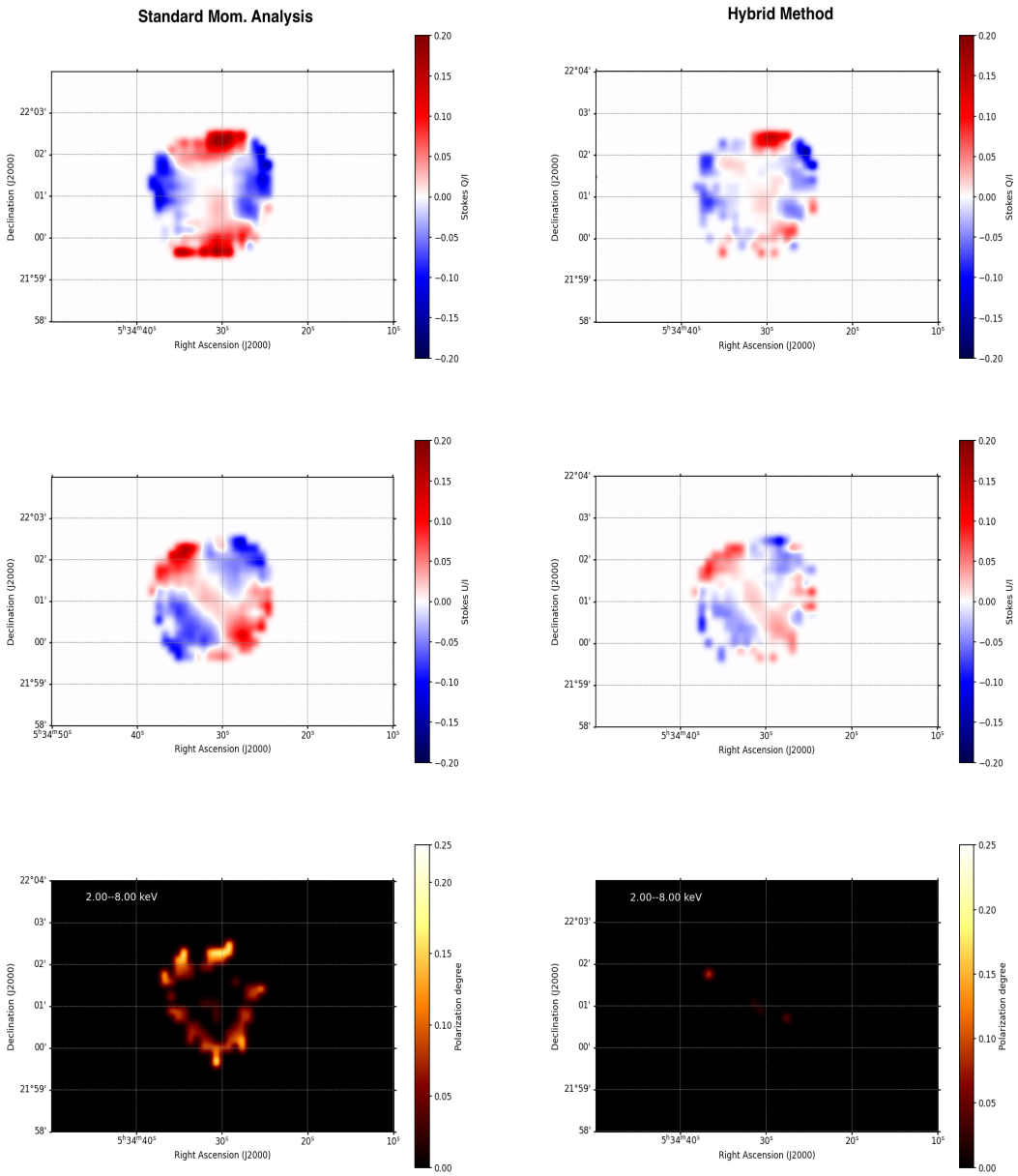

Figure 5.19: **Polarization leakage effect on a simulated extended source.** *Upper panels*: Q/I binned stokes parameter for the unpolarized simulation of the Crab nebula with a single IXPE detector, for the standard moment analysis on the *left* and for the hybrid method on the *right*. *Middle panels*: same as upper panels, but for the U/I stokes parameter. *Lower panels*: Polarization degree significantly ($>3\sigma$) detected by the standard moment analysis on the *left* and by the hybrid method on the *right*. In the *top and middle panels*, only the bins with a number of counts which is $> 96\%$ of the highest value are shown.

Chapter 6

Validation of the algorithm with experimental data

The previous chapter discussed the structure and performance of a new CNN-based algorithm for tracks reconstruction of events acquired with Gas Pixel Detectors. As reported in Sec. 5.1, the data set we used to develop the algorithm and to present the results consisted of Monte Carlo simulations. A rigorous and comprehensive validation of new models conceived for the analysis of scientific data, such as IXPE data, is of utmost importance. This validation process consists in testing the algorithm with laboratory data, which provide the unique opportunity to control and arrange the properties of data, i.e. the actual energy and polarization properties of the measured radiation. Moreover, for missions like IXPE, which are called discovery missions, standard astrophysical sources are indeed not available for calibration measurements, hence laboratory data are the only viable option.

For Machine Learning based algorithms this process is particularly important, as they could be very sensitive to the differences between simulations and real data. As a matter of fact, previous CNN-based algorithms, that have been conceived for IXPE data reconstruction, showed the biggest limitations in this context. For instance, although the Peirson et al. network [6] exhibited very promising results with simulations, when tested by the IXPE calibration team using laboratory data, it revealed two distinct systematic issues [139]. Firstly, the energy reconstruction performed by the neural network introduced a systematic bias, resulting in an overpopulation of events at higher energies compared to the standard moment analysis reconstruction, which was deemed more robust. Secondly, the modulation curve of unpolarized beams exhibited dominant contributions at higher frequencies than expected from spurious modulation and the intrinsic polarization of the beam, as depicted in Fig. 6.1. While the first systematic bias is not relevant to this new algorithm, given that the energy reconstruction mirrors the standard moment analysis, meticulous attention must be given to the modulation curves achieved with the hybrid method.

The validation process I have conducted for the hybrid algorithm involves measure-

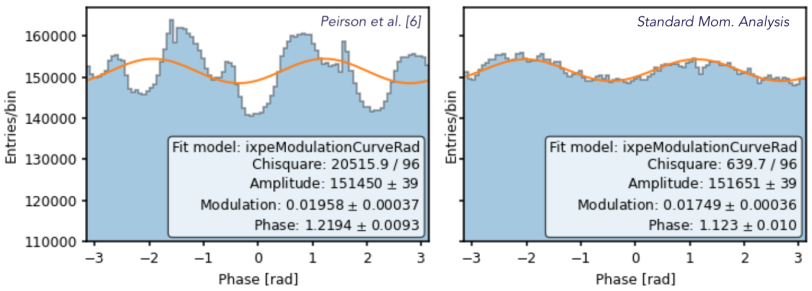

Figure 6.1: **Modulation curves for the calibration phase of previous algorithms**. The *left panel* reports the distribution of the emission angles achieved by the Peirson et al. [6] algorithm for an unpolarized 2.7 keV energy beam as a *blue histogram*. The *orange line* represents the modulation curve best-fit model. Analogously, the *right panel* shows the results for the standard moment analysis. Image credits: [139].

ments of both unpolarized and polarized radiation. Unpolarized beams are used to evaluate the ability of the algorithm to not introduce biases in the measurement, similarly to what it was discussed and verified with simulations in the previous chapter. Additionally, the unpolarized calibration data could be employed to evaluate the systematic effect of the spurious modulation, which was shortly described in Sec. 2.4. On the other hand, polarized radiation is used to validate the modulation factor curve we obtained with simulations: we expect to achieve values of modulation factors similar to those obtained with simulated data.

To validate the performance of the hybrid algorithm, I employ the data acquired during the IXPE ground calibration campaign performed at INAF-IAPS in Rome [13], which were used to calibrate the GPD performance and validate the standard moment analysis before IXPE launch. Throughout the calibration process, the IXPE detector units were exposed to both unpolarized and polarized beams, spanning various energies in the 2-8 keV range. The resulting FITS files contain the information regarding photo-electron tracks that is essential for subsequent data analysis. I process these data both with the hybrid method and with the standard moment analysis to achieve comprehensive results for comparison purposes and for validating the new algorithm.

Sec. 6.1 describes the employed data set. Results concerning unpolarized radiation, alongside with the spurious modulation maps characterization, are reported in Sec. 6.2. Sec. 6.3 depicts the results regarding polarized radiation. Finally, Sec. 6.4 is a brief appendix describing the X-ray Calibration Facility (XCF), a X-ray detectors calibration setup we are developing at the University of Torino.

6.1 Calibration data set

The laboratory setup employed to collect the data set we use to validate the performances of the algorithm is shown in Fig. 6.2. In this work only a summary of the main

characteristics of the setup, which are key for the validation process we performe, is depicted, but a detailed description is available in Muleri et al. [12].

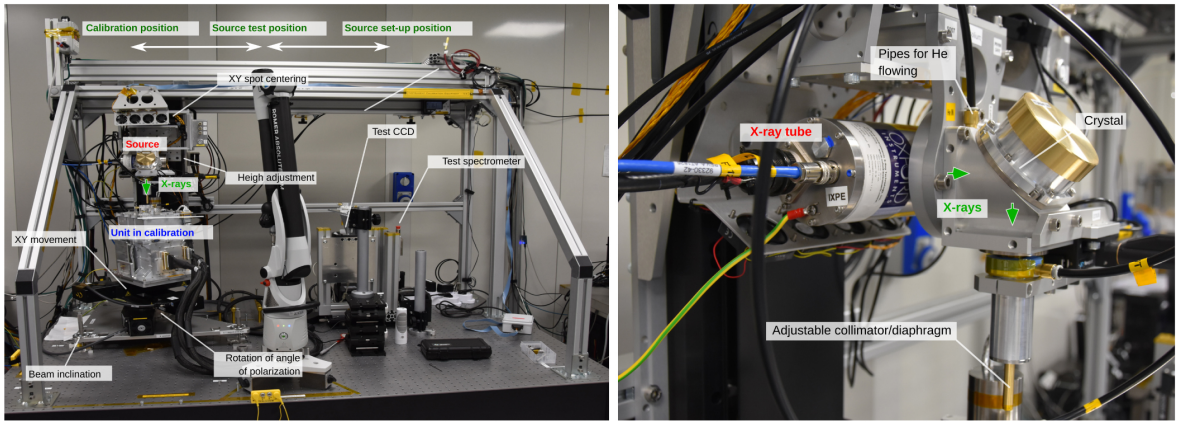

Figure 6.2: **Experimental setup at INAF-IAPS in Rome used to produce unpolarized and polarized beams devoted to IXPE ground calibration**. Image credits: [12].

The setup shown in Fig. 6.2 is called Instrument Calibration Equipment (ICE), and was originally used to calibrate the Detector Units now onboard of IXPE. The setup was specifically built for this purpose: an internally developed system offered significant flexibility, being consistently accessible for the calibration phase and enabling specialized measurements for IXPE DUs. It comprises calibration sources, which can generate both ∼100% polarized or unpolarized beams in the 2-8 keV range, alongside with the mechanical equipment needed to efficiently handle the rotation and alignment processes of the GPD. A set of manual and motorized stages were indeed designed for shifting the DU orthogonally to the incident beam, to move the same beam across the sensitive area of the detector. Importantly, the DU could also be rotated on the azimuthal plane to test the response at different polarization angles, and the inclination of the detector plane with respect to the beam could be adjusted as well.

Before calibrating the DUs, the test sources were characterized with a Charge-Coupled Device (CCD) imager, to image the beam spot, and with a Silicon Drift Detector (SDD) spectrometer and photometer, to measure the source flux and spectrum. A collimator and a diaphragm were used for polarized sources to achieve well defined beam spots.

Unpolarized sources

Different types of unpolarized sources were employed in the ICE facility to span the energy range of IXPE. Both a radioactive source and X-ray tubes are used to generate the beams. This tube consists of a cathode, which is heated to generate electrons through a process known as thermionic emission. These electrons are subsequently

accelerated and directed toward the anode, a metallic disk designed to initiate the production of fluorescence lines and a continuum Bremsstrahlung emission.

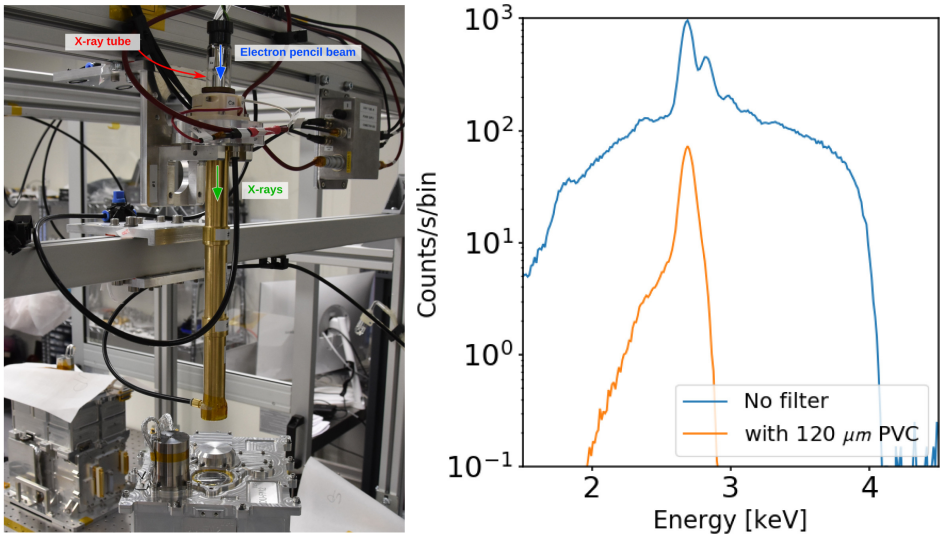

Figure 6.3: **ICE setup features.** *Left panel*: schematic representation of X-ray production on the head-on X-ray tube of ICE. *Right panel*: Example of Rh anode beam spectrum with (*orange*) and without (*blue*) the employment of a PVC filter. Image credits: [12].

A schematic representation of the process is reported in the *left panel* of Fig. 6.3. The emitted photons have two distinct pathways: they can either be collected directly by the detector after a potential suppression of the continuum emission with X-ray filters, or they may collide a target that, in response, emits fluorescence light, which is then measured by the GPD. An example of the Bremsstrahlung suppression performed by the filters is reported in the *right panel* of Fig. 6.3. Tab. 6.1 reports the sources and the methodology used to produce X-ray unpolarized beams at different energies in the IXPE range.

The calibration with unpolarized beams was conducted employing two data collection strategies. Firstly, the emitted radiation covered the whole sensitive area of the GPD with the so-called *Flat Field (FF)* measurements. Additionally, the central region (within 3.3 mm radius) was irradiated with higher sensitivity with the *Deep Flat Field (DFF)* measurements. The latter were aimed at better characterizing the central area of the detector, which is mostly involved in the astrophysical measurements with IXPE.

A crucial property of the unpolarized data set we employ for the validation of the hybrid algorithm must be underlined. We had access to the calibration data regarding only the DFF measurements of IXPE Detector Unit n° 1 (DU1), and not DU2 and DU3. This allowed us to characterize the spurious modulation maps of the hybrid method only for DU1 and only for its central region. These data are adequate to test and validate the functioning of the algorithm, but, if the hybrid method is to be inserted

Configuration	Energy [keV]
Fluorescence of Zr target illuminated by Rh anode	2.04
Fluorescence of Mo target illuminated by Ag anode	2.29
Direct X-rays with Rh anode✣	2.70
Direct X-rays with Ag anode✣	2.98
Direct X-rays with Ca anode	3.69
^{55}Fe nuclide	5.89

Table 6.1: **Summary of configurations and targets used to produce X-ray unpolarized beams at the reported energies.** The configurations marked by ✣ presented a polarized Bremsstrahlung continuum emission which was partially suppressed by the employment of X-ray filters (see right panel of Fig. 6.3).

into the official pipeline, it will be necessary to complete this study by considering the entire area of DU1 and the other 2. Additional considerations about this limitation will be discussed in the following sections.

Polarized sources

In order to produce polarized beams, the same X-ray tube described in the previous section and represented in Fig. 6.2 is employed. However, in this case, the X-rays produced by the anode are polarized by Bragg diffraction at nearly 45 degrees using a specific crystal. A brief introduction to this mechanism was depicted in Sec. 1.2.1. When hitting the crystal, only the radiation component which polarization is orthogonal to the plane of diffraction is efficiently diffracted. The resulting beam is 100% polarized if the crystal is at 45° with respect to the incident beam direction. This process permits to achieve fully polarized radiation from an unpolarized beam. Bragg's law is reported here again for convenience:

$$E = \frac{nhc}{2d\sin(\theta)} \tag{6.1}$$

E is the photon energy, d is the crystal lattice pitch, n is the diffraction order and θ is the incidence angle with respect to the crystal surface. Eq. 6.1 shows that for a fixed incident angle, and according to the selected crystal (d), only photons with a specific energy are diffracted (at the first order n=1). As a consequence, for each energy, a distinct combination of X-ray tube and crystal is used. The choice of diffraction angle is carefully adjusted to match the crystal lattice step to the energy of the selected fluorescence line generated by the X-ray tube. In this configuration, the incoming radiation can be effectively considered monochromatic at the energy of the fluorescence line, precisely defining the required diffraction angle. In cases where fluorescence lines are not available at the desired energy, the X-ray tube continuum radiation is employed

for diffraction: only the radiation components with the energy meeting the Bragg's law requirements are effectively diffracted. For this specific case, also the higher diffraction orders (n>1) are diffracted.

Moreover, for the monochromatic beams, X-rays which satisfy Bragg condition produce an arc on the detector. A schematic representation of this process is reported in Fig. 6.4. As already mentioned, a collimator first, and then a diaphragm placed at the center of the Bragg arc, were used to select a narrow beam spot with a diameter of approximately 1 mm.

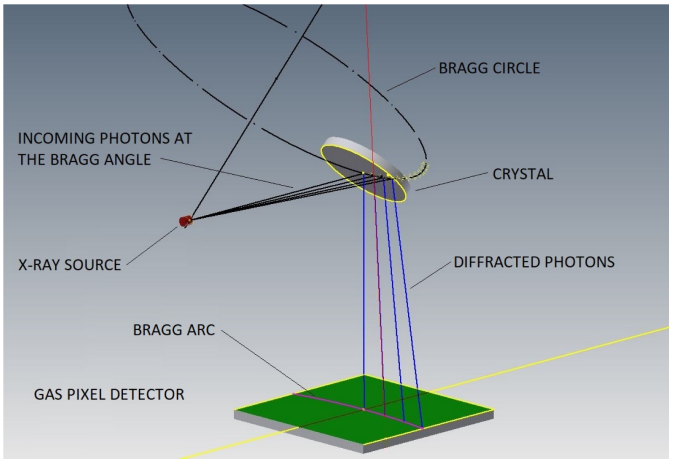

Figure 6.4: **Schematic view of the Bragg diffraction for monochromatic radiation.** Image credits: [12].

The resulting beams are not precisely 100% polarized, as the diffraction angle is not always precisely 45°: when evaluating the response of the GPD to the polarized beams, corrections factors need to be applied to account for this effect. Tab. 6.2 reports the choice of incident radiation and crystal used to produce polarized beams at different energies, alongside with the effective polarization degree of the final beam.

6.2 Unpolarized radiation results

In Sec. 5.4.2 we showed that both the standard moment analysis and the new hybrid algorithm found no residual linear polarization for unpolarized simulated point sources. When analyzing unpolarized calibration data, we would expect the same outcome to verify that the reconstruction algorithm does not introduce any non-physical biases. The data set described in Sec. 6.1 was used to validate the response of the hybrid algorithm to unpolarized radiation. For each energy beam, ∼20M events are analyzed employing the two algorithms, evaluating the resulting track emission angles ϕ. Starting from these angles, the residual modulation can be evaluated through Stokes parameters, as previously described for simulations.

Incident radiation + crystal	Energy [keV]	Polarization Degree [%]
Continuum + PET(002)	2.01	100.0
Mo (lα) + InSb (111)	2.29	99.3
Rh (lα) + Ge (111)	2.70	99.5
Ag (lα) + Si (111)	2.98	95.1
Ca (kα) + Al (111)	3.69	99.4
Ti (kα) + Si (220)	4.51	99.5
Fe (kα) + Si (400)	6.40	100

Table 6.2: **Summary of configurations of incident radiation and crystal used to produce X-ray polarized beams at the reported energies.**

In Fig. 6.5 the distributions of the predicted emission angles for three different energy beams are reported. Each plot is produced considering a subsample of ∼700 000 events. Results show that a residual linear modulation has been detected with both the standard moment analysis and the hybrid algorithm. The angle distribution is indeed characterized by a sinusoidal shape that is easily recognizable, while an ideal polarimeter should measure only a very small fluctuation of modulation due to the Poisson distribution of photo-electrons, as obtained with simulated data (see Fig. 5.16). Moreover, very high-frequency small peaks ($\nu = 6, 12$) appear in the hybrid method distribution. These will be discussed later in this section. In Tab. 6.3, the measured residual modulations as a function of energy are reported. Compatible values are achieved with the standard moment analysis and the hybrid method.

Energy [keV]	Standard Mom. Analysis $\mathcal{M}$ [%]	Hybrid Method $\mathcal{M}$ [%]
2.04	1.5 ± 0.2	1.6 ± 0.2
2.29	1.2 ± 0.2	1.3 ± 0.2
2.70	1.6 ± 0.2	1.6 ± 0.2
2.98	4.0 ± 0.2	3.9 ± 0.2
3.69	0.8 ± 0.2	0.7 ± 0.2

Table 6.3: **Residual modulation ($\mathcal{M}$) values obtained by analyzing unpolarized beams with the standard moment analysis and with the hybrid method.**

This residual modulation was already measured and studied with the standard moment analysis before IXPE launch, and it is determined by two factors. Firstly, the continuum Bremsstrahlung emission, which characterizes some of the configurations described in Tab. 6.1, is partially polarized, and a small residual modulation could persist despite the filters employed to suppress it (from approximately 0 to 4%, see Tab. 6.4). Secondly, most of the systematic signal is due to the *spurious modulation*

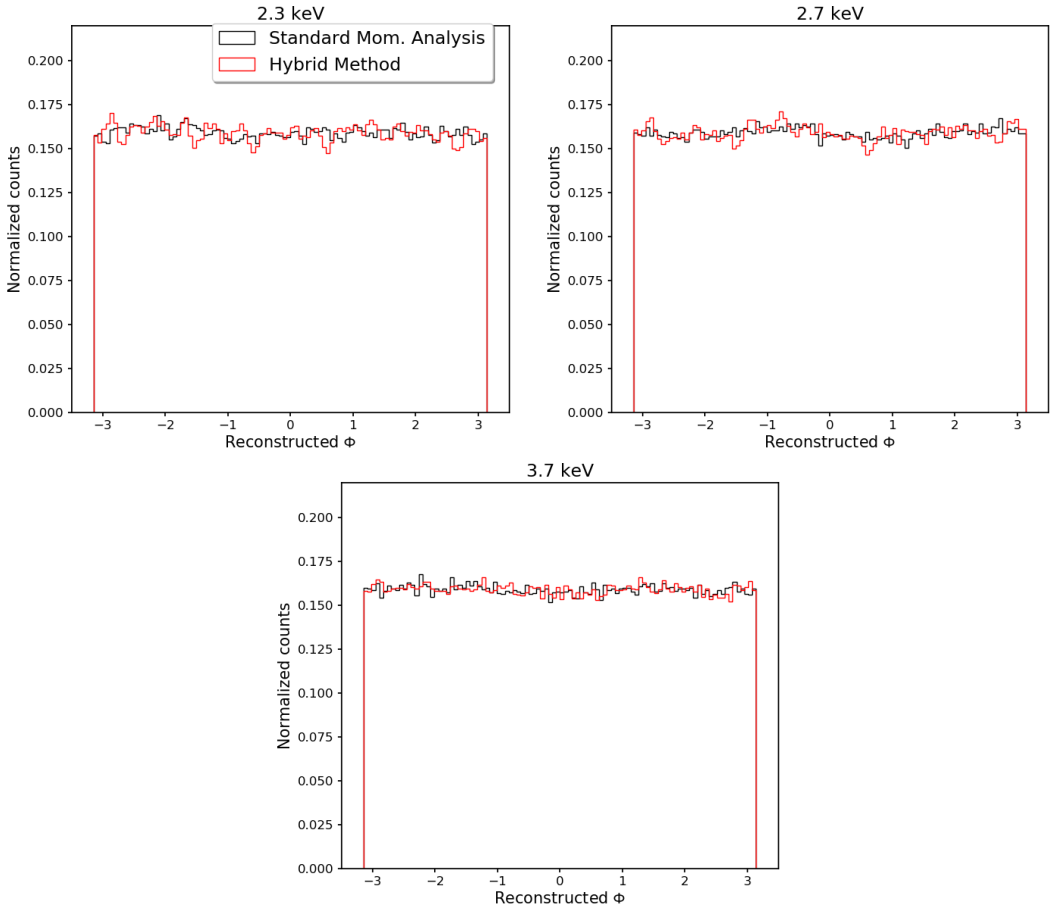

Figure 6.5: Emission angles ϕ distributions for three unpolarized beams (Upper left: 2.29 keV; Upper right: 2.70 keV; Lower center: 3.69 keV). Results are reported for the hybrid method (red histogram) and for the standard moment analysis (black histogram).

effect. As already mentioned in Chapter 2, spurious modulation is essentially thought to be caused by the GEM and operates as an extra cosine-squared component, i.e. it shares the same frequency as the modulation caused by authentic source polarization. Consequently, the combined effect of genuine and spurious modulation will continue to exhibit a $\cos^2$ pattern, and could amplify or reduce an authentic modulation. It is thus crucial to properly determine and characterize the spurious modulation.

Anomaly in the hybrid ϕ distribution

Before delving into the characterization of spurious modulation, it is important to address the high-frequency small peaks observed in the hybrid method distribution of emission angles at low energies. In Fig. 6.6, the distribution of the hybrid method

reconstructed emission angles for the 2.7 keV beam, with the relative power spectrum, is reported. As already mentioned, excluding the ν=2 peak due to the joint contribution of residual polarization and spurious modulation, higher frequency peaks (ν=6,12) are detectable. These peaks, absent in the standard moment analysis distribution, might initially suggest a systematic bias in the CNN impact point prediction, but this is not the case. In order to understand their origin, two previously mentioned features of the hybrid method need to be revisited.

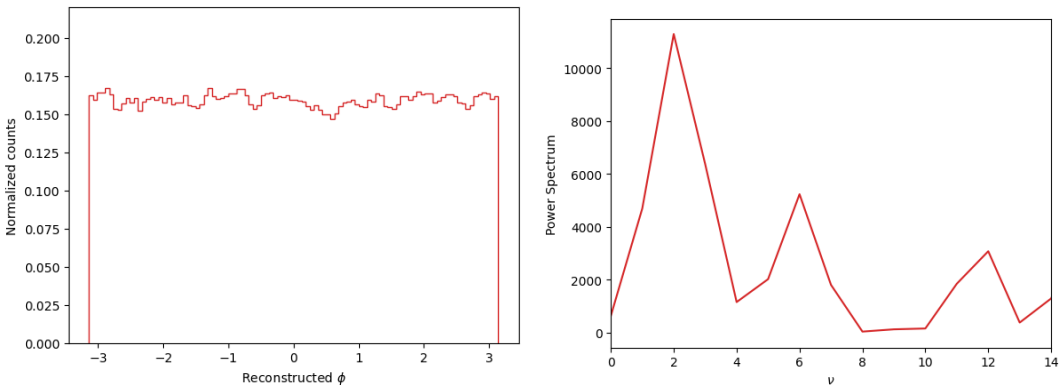

Figure 6.6: **Normalized counts and power spectrum of the hybrid method ϕ distribution.** *Left panel*: normalized distribution of the hybrid method predicted emission angles for the 2.7 keV beam subsample. *Right panel*: power spectrum of the emission angles distribution.

Firstly, results regarding the IP reconstruction (see Fig. 5.13) emphasized the CNN inclination to predict the impact point close to the barycenter of the track for energies below 3 keV. For low-energy tracks, in fact, the barycenter closely aligns with the true photon impact point.

Secondly, the final emission direction is assigned weighting the pixels charge according to their distance from the barycenter of the horseshoe region, both in the standard moment analysis and in the hybrid method (see Eqs. 2.5-2.8). The upper panel of Fig. 6.7 reports a track example with its barycenter in yellow, whereas the blue dot is the barycenter of the horseshoe region (marked by a dashed blue line). The lower panels show the distribution of the angles defined by the vector connecting the horseshoe barycenter (HS) and the track barycenter (BAR) for the 2.7 keV energy beam (dashed black line in the upper panel), with its relative power spectrum. This distribution is entirely independent of CNN results and exhibits the same frequencies $\nu = 6, 12$ as the hybrid method emission angles distribution, and could explain the observed high-frequencies peaks in Fig. 6.6. As the CNN-reconstructed IP starts deviating from the barycenter of the track from energies greater than ∼3 keV, the absence of these high frequency peaks for the 3.69 keV beam is also explained. Such peaks are absent in the standard moment analysis distribution as well, as the reconstructed IP is never

close to the barycenter of the track. While the origin of this biased alignment between the horseshoe barycenter and the track barycenter demands further exploration, it ultimately has no consequence on the determination of the Stokes parameters and, subsequently, on the polarization properties of the incident radiation.

Spurious modulation characterization

Spurious modulation characterization is a key process for the calibration phase of IXPE GPD. Its contribution must be subtracted to the measured polarization in order to achieve unbiased results. This section focuses on how to characterize and then subtract the spurious modulation.

It was previously mentioned that the residual modulation detected by analyzing the unpolarized beams is due to two instrumental systematics. Firstly, sources described in Tab. 6.1 should be completely unpolarized, but they actually could present an intrinsic residual polarization. Additionally, spurious modulation is also a contribution to the systematic, caused by the detector structure. Rankin et al. [40] established a method to disentangle the true polarization signal of the source from the spurious modulation. This method consists in conducting measurements by rotating the GPD around its center by 90°, i.e. by orienting the GPD in 2 different directions perpendicular to each other. In these two measurements, when viewed from the GPD reference frame, the phase of spurious modulation remains constant, while the phase of the intrinsic modulation produced by the source shifts by 90°.

A crucial aspect of the spurious modulation is that this effect is position and energy dependent [17]. For the standard moment analysis the calibration of the spurious modulation was performed for all the three IXPE DUs. In this work we report the results obtained for both the standard moment analysis and the hybrid method for a single DU, and only for the central part of the GPD. As previously mentioned, the available data set was indeed obtained by illuminating a 3.3 mm-radius circled area around the center of the detector.

For each energy bin, the Stokes parameters of each event are calculated, as reported in Eqs. 2.17-2.18, for both the reconstruction algorithms. The same procedure is performed for two different GPD orientations, rotated by 90°. The GPD is divided in 300×300 spatial macro bins, and each event is assigned to the appropriate bin on the basis of its reconstructed impact point. For each bin and for each configuration, the total normalized Q/I and U/I values are calculated as in Eqs. 2.22-2.23. At this point, for each energy bin, eight GPD maps are available: Q/I and U/I maps at 0° and 90° for both the standard moment analysis and the hybrid method. In Fig. 6.8, the Q/I and U/I maps at the 0° configurations are reported for both the reconstruction algorithms at the benchmark energy bin of 2.29 keV.

In Chapter 1, the additive property of the Stokes parameters was discussed, i.e. the Stokes parameters of waves superposition is the sum of the individual waves Stokes parameters. From this property, the Q/I and U/I values of these maps are determined

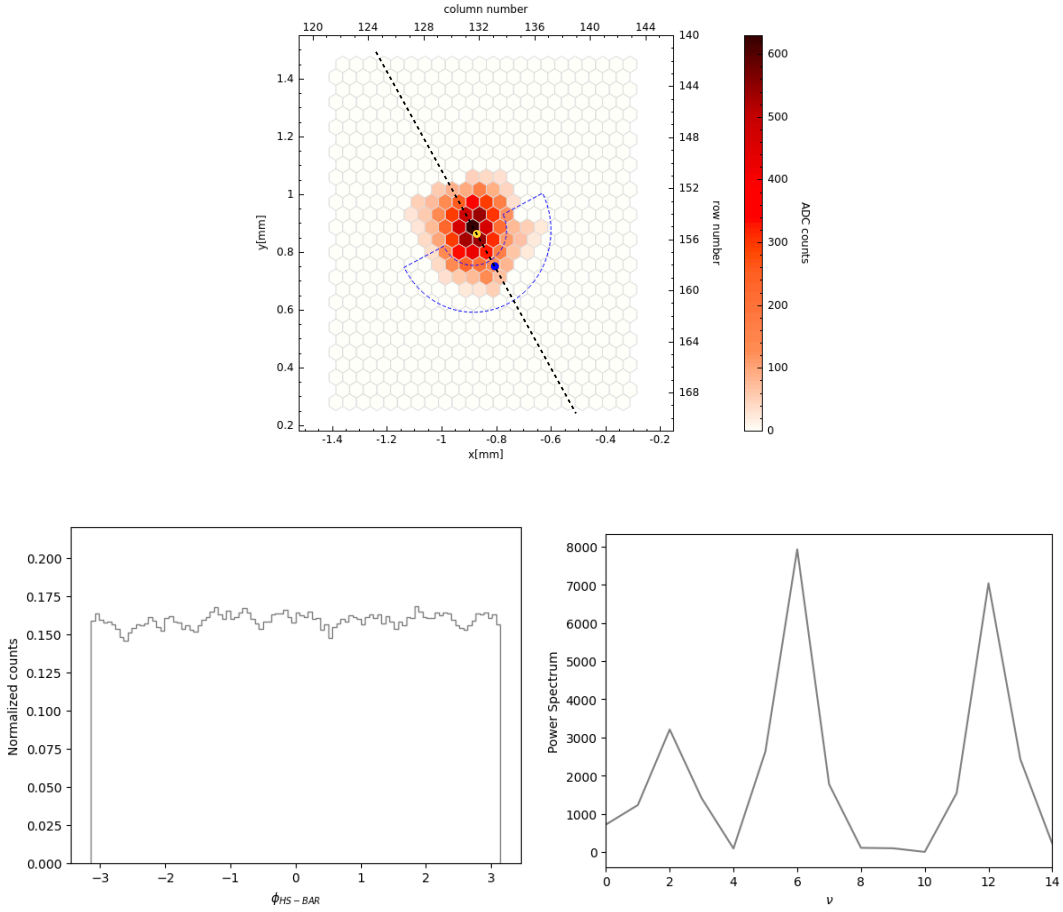

Figure 6.7: **Example of PE track, normalized counts and power spectrum of the ϕ_{HS-BAR} distribution.** *Upper panel*: Example of a photo-electron track. The *yellow dot* is the barycenter of the track, whereas the *blue dot* is the barycenter of the horseshoe region (marked by the *dashed blue line*). The *dashed black line* is the direction of the vector connecting the two points. *Left panel*: normalized distribution of the angles defined by the horseshoe barycenter (HS) and the track barycenter (BAR), denoted ϕ_{HS-BAR}, for the 2.7 keV beam subsample. *Right panel*: power spectrum of the ϕ_{HS-BAR} distribution.

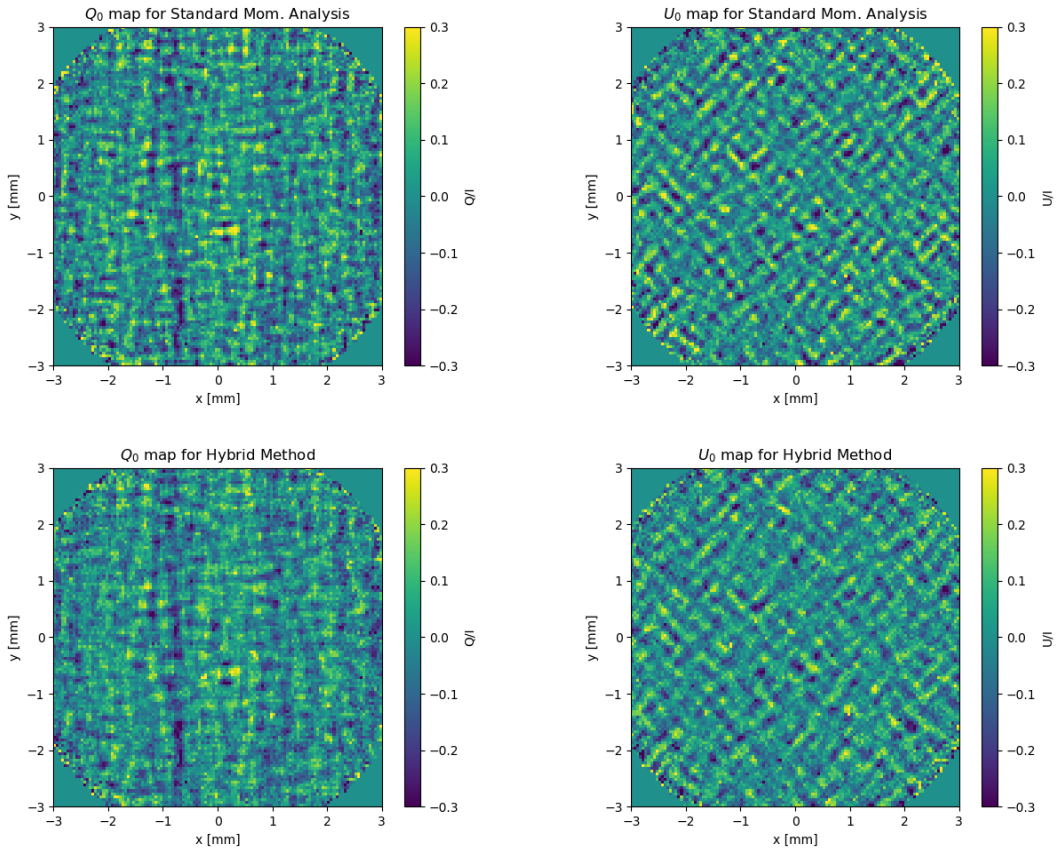

Figure 6.8: **Q/I and U/I binned maps for the central part of the GPD (DU1 on IXPE), at the 0° configuration, for the 2.29 keV energy beam.** *Upper panels*: results for the standard moment analysis. *Lower panels*: results for the hybrid method.

by (all these Stokes parameters are intended as normalized, i.e. divided by I):

$$\begin{cases} Q_0 = Q_{SM} + Q_{Source}(\text{GPD at } 0°) \\ Q_{90} = Q_{SM} + Q_{Source}(\text{GPD at } 90°) = \text{Q}_{\text{SM}} - \text{Q}_{\text{Source}}(\text{GPD at } 0°) \\ U_0 = U_{SM} + U_{Source}(\text{GPD at } 0°) \\ U_{90} = U_{SM} + U_{Source}(\text{GPD at } 90°) = \text{U}_{\text{SM}} - \text{U}_{\text{Source}}(\text{GPD at } 0°) \end{cases} \tag{6.2}$$

where *SM* indicates the spurious modulation contribution, while *Source* marks the intrinsic source polarization one, which depends on the angle configuration. By solving the equations for the *SM* contribution we find:

$$\begin{cases} Q_{SM} = \frac{Q_0 + Q_{90}}{2} \\ U_{SM} = \frac{U_0 + U_{90}}{2} \end{cases} . \tag{6.3}$$

Similarly, the Stokes parameters relative to the intrinsic residual polarization of the source can be obtained as:

$$\begin{cases} Q_{Source} = \frac{Q_0 - Q_{90}}{2} \\ U_{Source} = \frac{U_0 - U_{90}}{2} \end{cases} . \tag{6.4}$$

From Eq. 6.3 two spurious modulation maps (Q_{SM} and U_{SM}) for each energy bin can be evaluated, for both the standard moment analysis and the hybrid method. Results for the benchmark energy bin at 2.29 keV are reported in Fig. 6.9.

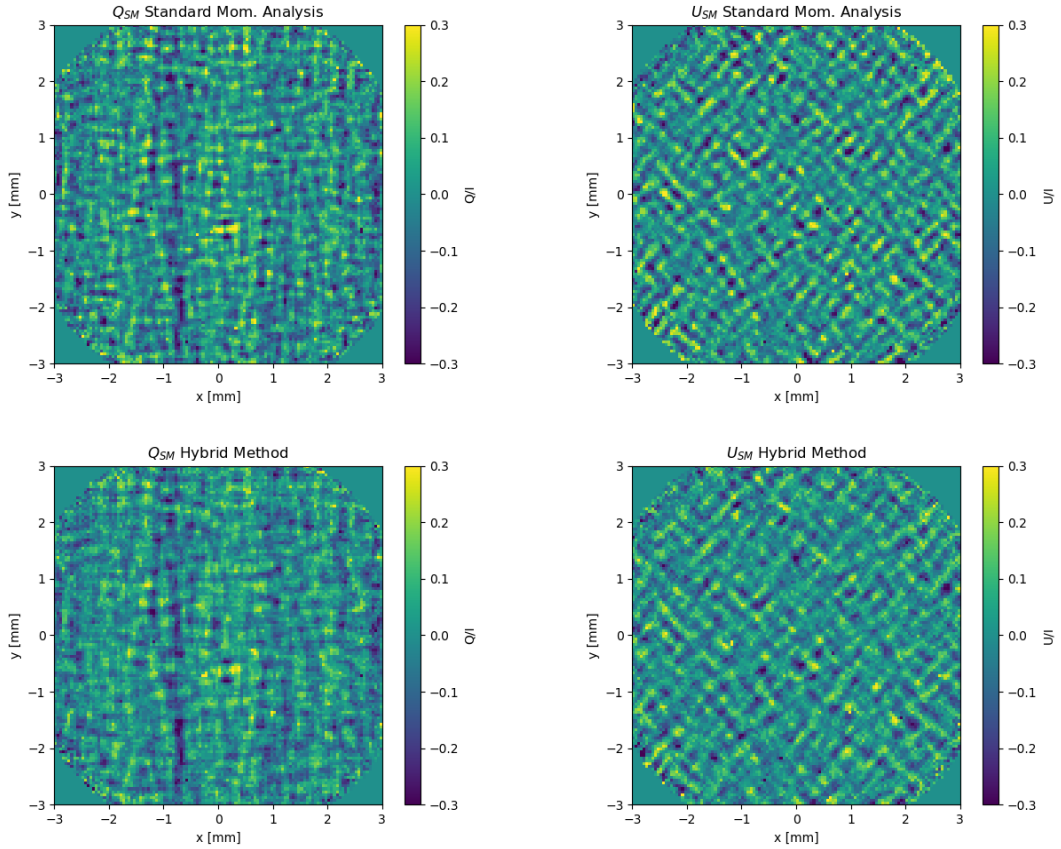

Figure 6.9: **Q/I and U/I binned spurious modulation maps for the central part of the GPD (DU1 on IXPE), for the 2.29 keV energy beam**. *Upper panels*: results for the standard moment analysis. *Lower panels*: results for the hybrid method.

It should be noticed the similarity between the results obtained through the standard moment analysis and the hybrid method, highlighting the presence of shared general structures. We expected to observe this resemblance, as the algorithms used for calculating the emission angle are very similar, with the primary distinction being

the position of the impact point. Our findings confirm that spurious modulation is a systematic caused by the detector, as it is observable with different reconstruction algorithms.

Once these maps are produced, they are used to correct the Q and U values obtained from the measured events, accounting for the influence of spurious modulation on polarization measurements. The procedure involves correcting each photon individually by subtracting the spurious modulation map value relative to the spatial bin where the photon is absorbed and to its measured energy. Since the facility setup and the outlined process yield distinct maps for discrete energy values, linear interpolation is employed bin by bin across different energies of the Stokes maps. This ensures an accurate adjustment of spurious modulation values even for events with reconstructed energies that do not precisely align with the values of the unpolarized beams listed in Table 6.1. Specifically, for each event, the correction is applied as:

$$\begin{cases} q_{\text{corrected}} = q_{\text{uncorrected}} - Q_{\text{SM}}[x][y][E] \\ u_{\text{corrected}} = u_{\text{uncorrected}} - U_{\text{SM}}[x][y][E] \end{cases} \tag{6.5}$$

where (x,y) is the reconstructed IP position and E is the reconstructed photon energy.

Finally, once the spurious modulation correction is applied photon by photon, it is important to validate the procedure by confirming that the remaining residual modulation is comparable to the intrinsic modulation of the unpolarized sources, that has not been taken into account yet. In Tab. 6.4 the comparison between the modulation of the sources ($\mathcal{M}_{Source}$) obtained as in Eq. 6.4 (without spacially binning the events) and the one obtained after correcting the 0° data as in Eq. 6.5 are reported for the hybrid method. Compatible values are achieved correcting the 90° data. For the standard moment analysis, this has already been verified during IXPE calibration campaign [40].

Energy [keV]	$\mathcal{M}_{Source}$ **from Eq. 6.4 [%]**	$\mathcal{M}$ **from Eq. 6.5 [%]**
2.04	0.16 ± 0.04	0.16 ± 0.04
2.29	0.16 ± 0.04	0.14 ± 0.04
2.70	1.62 ± 0.04	1.64 ± 0.04
2.98	4.02 ± 0.04	4.03 ± 0.04
3.69	0.05 ± 0.04	0.05 ± 0.04

Table 6.4: **Residual modulation values obtained by analyzing unpolarized beams**. The values on the first column are achieved by the Stokes parameters evaluated as in Eq. 6.4 without spacially binning the events. Results on the second column are achieved by subtracting the spurious modulation event by event to the 0° data, according to Eq. 6.5.

The compatibility of $\mathcal{M}_{Source}$ with the residual $\mathcal{M}$ achieved by correcting for the

spurious modulation validates the correction procedure. For each new measurement, following spurious modulation subtraction, we can confidently detect the intrinsic polarization of the incident radiation.

All the results presented in this section confirm the proper functioning of the hybrid method when used for analyzing unpolarized radiation in laboratory: the spurious map closely resembles the one obtained with the standard moment analysis, the residual intrinsic modulation of the sources obtained with two different methods is comparable, and the findings show no significant bias or unexpected behaviour which was undetected with simulations.

6.3 Polarized radiation results

Once the spurious modulation maps are generated, and the response of the GPD to unpolarized beams presents no sign of residual modulation (with the exception of the intrinsic modulation of the source), the polarized data set analysis can be addressed. Firstly, few considerations about the energy spectra of the radiation beams and the spatial distribution of their reconstructed impact points are necessary.

In Fig. 6.10, the spectra of the polarized radiation beams, as detected by the SDD spectrometer prior to the DU calibration, are shown. Almost all the configurations reported in Tab. 6.2 are characterized by the X-ray tube fluorescence emission coupled with the crystal. The 2.01 keV energy beam, however, is generated by the diffraction of continuum Bremsstrahlung radiation, and higher diffraction orders display non-negligible intensities compared to the first order, as shown in the upper plot of Fig. 6.10. Nonetheless, the GPD energy resolution is adequate to differentiate the lines, and only the events corresponding to the 2.01 keV line can be selected.

Fig. 6.10 also shows that for the other energy beams there is no need for such a selection, as the only lines present in the energy distributions are those that correspond to the desired energy.

Secondly, as discussed in the previous section, the data at my disposal allowed me to characterize the response of the GPD to the unpolarized radiation only for the central part of the detector ($R < 3.3$ mm, where R is the distance from the center of the GPD). Consequently, only the events of the polarized beams that impact in the characterized area of the GPD are considered in the analysis, as the correction for the spurious modulation is available only for that region. While this restriction to the central detector region marginally impacts the statistical aspects, the robustness of my validation procedure is not affected: there is no compelling reason to expect divergent results when considering the entire GPD surface.

The methodology applied to analyze polarized beams is equivalent to the one used with unpolarized beams: each event is corrected by the spurious modulation effect through Eqs. 6.5, and the modulation factor is evaluated the for the available energy beams, using both the standard moment analysis and the hybrid method. A correction

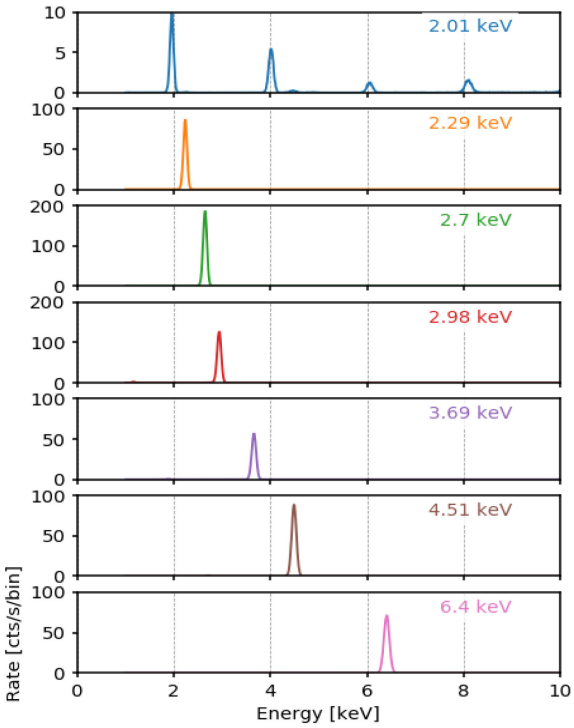

Figure 6.10: **Energy distribution of the polarized beams used for the validation phase of the algorithm**. The sources configurations are reported in Tab. 6.2. Image credits: [12]

factor needs to be finally applied to the achieved modulation factor values in order to account for the fact that some of the beams are not 100% polarized (see Tab. 6.2). Results are reported in the *upper panel* of Fig. 6.11.

The modulation factor comparison between the standard moment analysis and the hybrid method reported in the *upper panel* of Fig. 6.11 closely resembles the one obtained with simulations in the *lower panel*. The improvement is indeed marginal up to 3 keV, it is ∼1.5% at 3.7 keV, and peaks at ∼5% at 6.4 keV. A direct comparison between the results obtained with simulations and lab data, e.g. the μ values of the hybrid method with simulations (*red line* in the *lower panel*) and the ones with lab data (*red line* in *upper panel*), is not possible, because the gas pressure of the detector suffers a variation with time and has a direct impact on the detector performance. The pressure value adopted for the simulation and reported in Tab 5.1 could not correspond to the one at the time of the calibration campaign. The impact of the gas pressure variation on the polarimetric sensitivity is mild [2], but still could affect the values of modulation factors and prevents the comparison between effective data and simulations.

To complete the validation procedure and ensure the absence of biases, we evaluate the distribution of the reconstructed angles ϕ after the correction for spurious modulation. This information is reported in Fig. 6.12 for three distinct energy beams, and requires

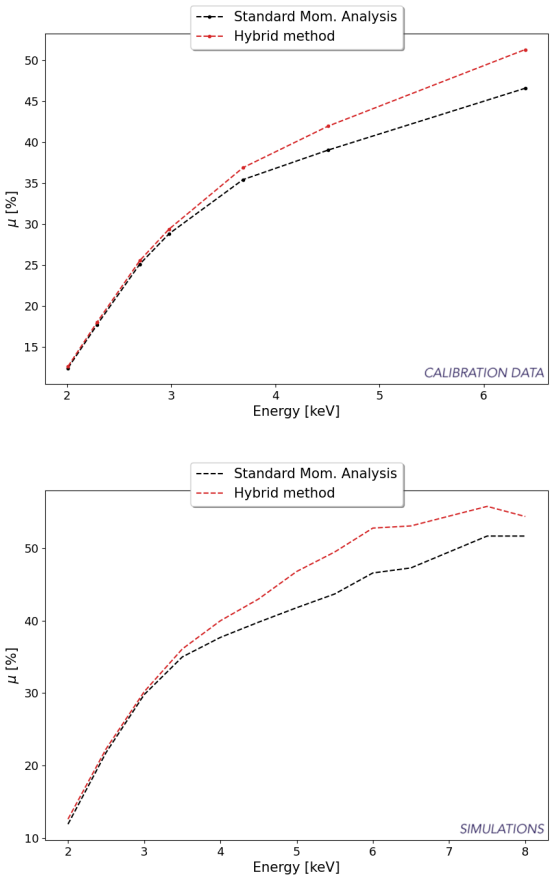

Figure 6.11: **Modulation factors for laboratory and simulated 100% polarized radiation beams.** The modulation factor values achieved with the polarized calibration sources are reported in the *upper panel*, for both the standard moment analysis (*black*) and for the hybrid method (*red*). As a comparison, the modulation factors for the simulated data set are reported again in the *lower panel*.

some clarifications. It should be noticed that the predicted ϕ values from both methods span the range of 0° to 360°. However, during the evaluation of Stokes parameters for spurious modulation correction, the distinction between events separated by 180° is lost. Consequently, recalculating the angle ϕ from Q and U restricts the ϕ distribution to the range 0°-180°. Moreover, after correction for the spurious modulation, q and u values for each event are affected, and they should be considered *pseudo*-Stokes parameters, as for example the relation $(q^2 + u^2 = 1)$ is not respected any more. Similarly, also the distributions reported in Fig. 6.12 concern *pseudo*-emission angles.

By interpolating and extrapolating the modulation factor values acquired for the 100% polarized beams, we can derive μ as a function of energy from 2 to 8 keV. Now, all the necessary physical quantities to conduct an astrophysical source data analysis

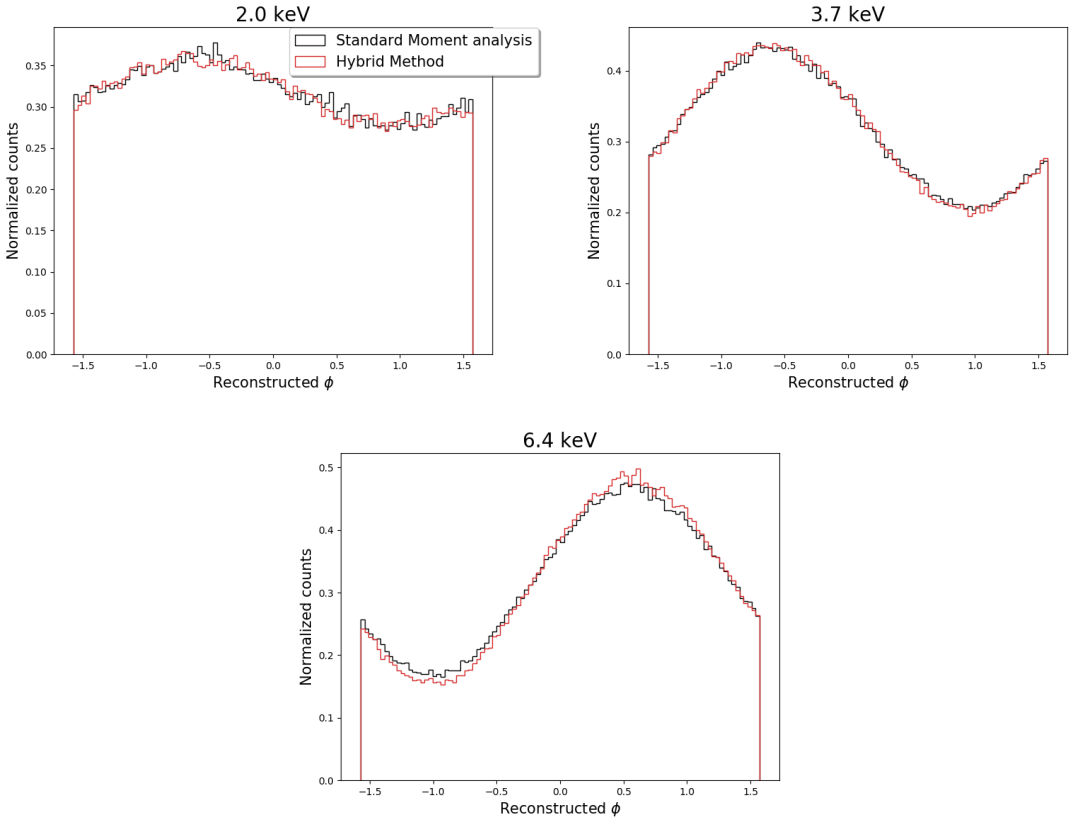

Figure 6.12: **Examples of predicted emission angles distributions for three benchmark energies** (*Upper left panel*: 2.7 keV | *Upper right panel*: 3.7 keV | *Lower panel*: 6.4 keV). Results are reported for both the standard moment analysis (*black*) and for the hybrid method (*red*).

using IXPE observations are available.

6.4 X-ray Calibration Facility (XCF)

As discussed throughout the chapter, the facility at INAF-IAPS in Rome was designated as the pre-launch calibration setup for IXPE. Meanwhile, at the University of Torino, we set-up a laboratory called X-ray Calibration Facility (XCF). XCF is currently employed to monitor the secular variations of IXPE GPD response over time and to study other new effects arisen during in-flight calibrations. Initially conceived as a calibration source to qualify GPD developed for astrophysical missions like IXPE, the XCF can easily satisfy evolving requirements to support R&D programs of innovative position- energy- and polarization-sensitive X-ray detectors. XCF is a compact, adaptable irradiation platform which can produce photon beams of different energy,

size, and polarization characteristics, ideal for detector calibrations and monitoring. A schematic view of the XCF design is presented in Figure 6.13.

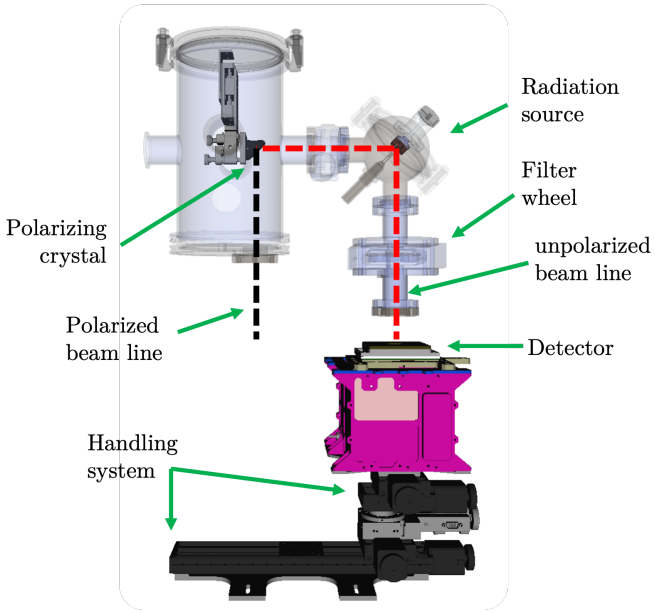

Figure 6.13: **Schematic view of the structure and functioning of the X-ray Calibration Facility**.

The radiation source is a McPherson Mod. 642 X-ray tube with a twofold output. The primary feature of this device is its multi-anode carousel, which can accommodate up to six anodes along with their holders. The multi-anode can generate fluorescence lines in the 2-10 keV range, in addition to the Bremsstrahlung component. A handle enables rotation of the carousel, facilitating the positioning of the desired anode in front of the filament. As mentioned before, radiation is emitted in two directions: the vertical beam is used for unpolarized measurements, whereas the horizontal beam is polarized through Bragg diffraction (see Sections 1.2.1 and 6.1).

This facility provides a unique opportunity to tailor the testing of our hybrid algorithm to specific needs. Presently, part of the data analysis of XCF relies on the hybrid algorithm. Additionally, considerable efforts for enhancing the Gas Pixel Detector performance (either through advancements in ASIC technology or the exploration of alternatives to the GEM) are currently underway. XCF will enable a comprehensive testing and characterization of the hybrid algorithm response and its adaptability to these future detectors.

Chapter 7

IXPE data analysis

In this chapter, the results of IXPE data analysis with the hybrid method are discussed. Firstly, Sec. 7.1 depicts an overview of IXPE data and in particular of the processing pipeline. In Sec. 7.2 the limits due to the characterization of only the central part of DU1 are discussed, and a solution for employing the whole active surface of all the DUs is presented. Finally, the comparison with the standard moment analysis is made explicit through two benchmark examples in Sec. 7.3.

7.1 IXPE data overview

IXPE data are publicly available on HEASARC, the NASA primary archive for missions studying photons of the entire energy spectrum [27]. On the archive two main types of IXPE data can be downloaded:

- **Level-1 files** Level-1 data (LV1) are FITS files containing raw information on the instrument and the spacecraft, and including the photo-electron track, reconstructed as two-dimensional image data array. They also contain information about the Pulse Height Amplitude (PHA) of the event, as well as the reconstructed impact point and emission angle in the detector reference frame. Events in LV1 fles are sorted in chronological order. They are employed for science analysis only in very few cases, as most of the useful information is also stored in the lighter Level-2 files.

- **Level-2 files** Level-2 data (LV2) are FITS files containing the filtered and calibrated event data produced by the instrument pipeline [140]. Specifically, they hold Pulse Invariant (PI) energy corrected for some instrumental systematics: a charging effect, due to the charge of the avalanche in the GEM which could affect the electric field in the holes and thus modify the gain, and factors accounting for the detector temperature and large-scale spatial variations. Additionally, they contain the emission angle of the photo-electrons resolved into Stokes parameters,

corrected for spurious modulation, and the IP position onto the detector plane projected to the J2000 tangent plane centered on the observed object. LV2 files only store events in Good Time Intervals (GTI) when data from the target source are expected to be dominant. Thus, the LV2 GTI includes intervals in which the detectors and the DSUs were properly configured and the source was tracked in the instrument field of view, whereas intervals of calibration, occultation, or slew are not included.

The IXPE archive provides also access to the files and information necessary for transitioning from LV1 to LV2 files. These resources are available both as downloadable housekeeping files and through the IXPE calibration database (CALDB) [141]. The former include primarily information related to the instrument, which is observation-dependent or changes over short timescales. The CALDB houses files that are either common to all observations or subject to changes over longer timescales. Examples include the spurious modulation maps and the modulation factors. This database undergoes periodic updates based on the results of in-flight calibration results.

As widely discussed throughout the text, the information we are interested in extracting from the IXPE analysis of an astrophysical source are the polarization degree and phase of the incident radiation. In Chapter 1, these quantities were derived and expressed as function of the modulation factor of the instrument and the Stokes parameters of the detected photons, as reported in equations 2.28 and 2.29, and repeated here again for clarity:

$$p_r = \frac{2}{\mu}\sqrt{\widetilde{Q}^2 + \widetilde{U}^2} \tag{7.1}$$

$$\Phi_r = \frac{1}{2}\arctan\frac{\widetilde{U}}{\widetilde{Q}} \tag{7.2}$$

where $\widetilde{Q}$ and $\widetilde{U}$ are the normalized measured Stokes parameters (i.e. Q/I and U/I). For standard moment analysis, the Stokes parameters are available for each event in the LV2 file, while the value of the modulation factor μ is available on IXPE CALDB as a function of energy.

Concerning the hybrid method, the Stokes parameters for each event must be evaluated from scratch starting from LV1 files: the photo-electron tracks are pre-processed as described in Sec. 5.3.3, and then are passed through the trained Neural Network, in order to achieve a new reconstructed impact point. The CNN-predicted IP replaces the standard moment analysis one in the fourth phase of the event reconstruction (see Sec. 2.3.2), achieving a new estimation of the emission direction. The same instrument pipeline adopted for the standard reconstruction algorithm is also used to process the hybrid method data to obtain the final corrected Stokes parameters. The modulation factor μ is interpolated and extrapolated from results in Fig. 6.11.

Fig. 7.1 schematically reproduces the entire procedure to process the original LV1 IXPE files with the standard moment analysis (*black path*) and the hybrid method

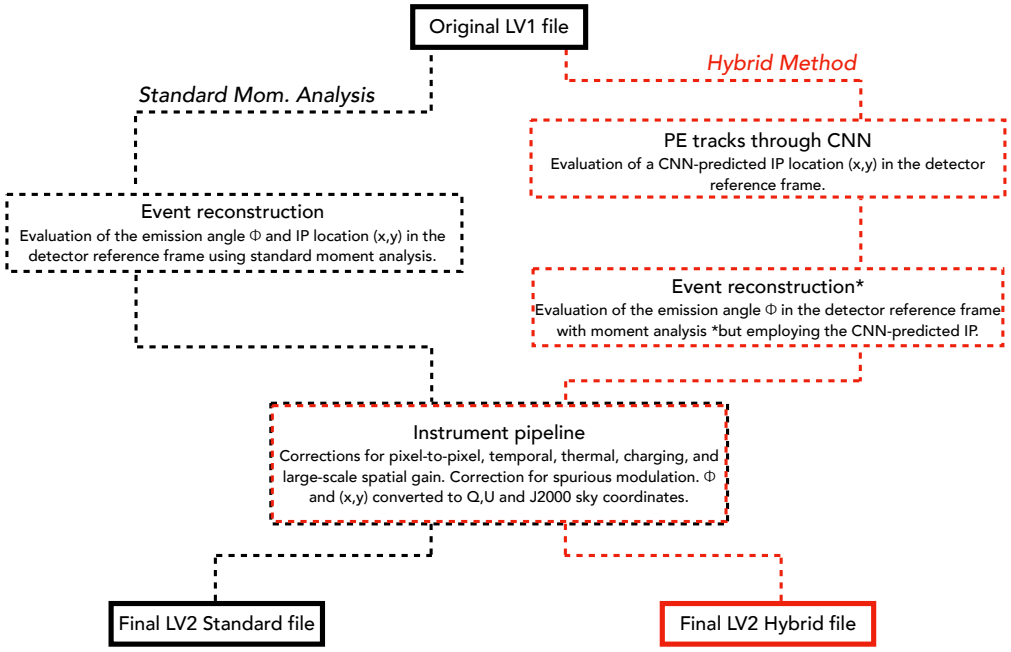

Figure 7.1: **IXPE pipeline schematic view**. Representation of the steps required from the standard moment analysis (*black line*) and the hybrid method (*red line*) to get to final LV2 data.

(*red path*). For IXPE data analysis with the standard reconstruction algorithm there is no need to go through this procedure, as LV2 files are already publicly available. However, we had to autonomously run the pipeline starting from LV1 file to obtain the the hybrid method LV2 files.

Level-2 files can be processed with an analysis framework specifically developed for IXPE, that is called *ixpeobssim*. The software is publicly available as well, and it enables spatial, spectral and temporal X-ray sources analysis, as well as complex polarimetry measurements and visualization tools [142].

7.2 Employing all three detectors for science data analysis

Section 6.2 highlighted that the available calibration data enabled us to evaluate the spurious modulation map with the hybrid method only in the central part of the GPD and for a single DU onboard of IXPE. Basically, results achievable with the hybrid

method would pertain only to data acquired under these conditions and could not encompass the whole observations performed by IXPE. This limitation would entirely prevent the possibility of analyzing extended sources, and would heavily constrain the observations of point sources as well.

However, we conducted a simple analysis of several point sources by using only the central region of DU1 data, and demonstrated that correcting the Stokes parameter obtained with the hybrid method for the spurious modulation maps computed with the standard moment analysis (H_{SM}), or with the hybrid method (H_{HM}), leads to compatible results. Quantitatively, for all the sources and for all the tested energy binning configurations, I verified the compatibility between $\widetilde{Q}_{H_{SM}}$ and $\widetilde{Q}_{H_{HM}}$, and independently between $\widetilde{U}_{H_{SM}}$ and $\widetilde{U}_{H_{HM}}$, with a Z-test: the measurements are always compatible within 1σ. The following figure of merit, which accounts for both the Stokes parameters, is employed to facilitate the representation of such compatibility:

$$\frac{\sqrt{(\widetilde{Q}_{H_{SM}} - \widetilde{Q}_{H_{HM}})^2 + (\widetilde{U}_{H_{SM}} - \widetilde{U}_{H_{HM}})^2}}{\sqrt{\sigma^2_{H_{SM}} + \sigma^2_{H_{HM}}}} \tag{7.3}$$

The obtained values are reported for four different point sources and for two energy binning configurations each in Fig. 7.2.

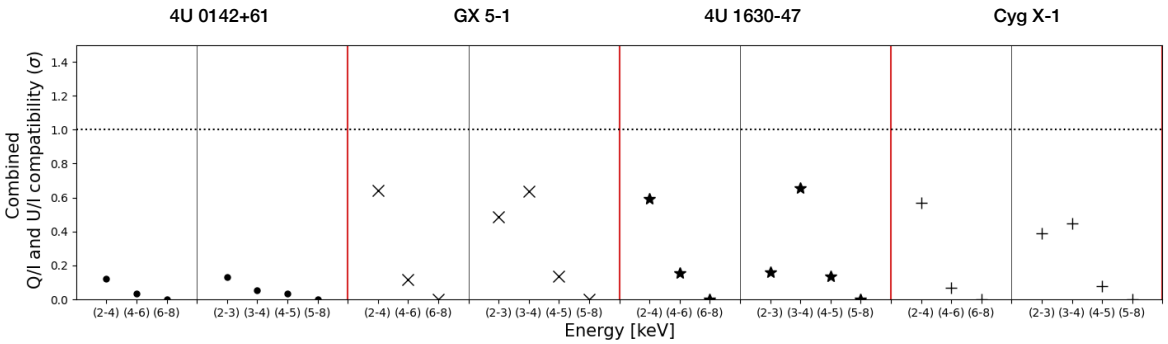

Figure 7.2: **Combined Q/I and U/I compatibility for two energy binning configurations for four different sources**. The results refer to the compatibility between the Stokes parameters obtained with the hybrid method and corrected for the spurious modulation maps computed with the standard moment analysis, or with the hybrid method (see Eq. 7.3). Two energy binning configurations are considered for each source (4U 0142+61: *dots*, GX 5-1: ×, 4U 1630-47: *stars*, Cyg X-1: +)

.

Two examples of such comparison as polar plots are reported for the magnetar 4U-0142+61 and the accreting neutron star GX 5-1 in Fig. 7.3 and in Tab. 7.1.

This compatibility allows us to correct the hybrid method Stokes parameters with the standard moment analysis spurious maps, thus accessing the possibility to perform

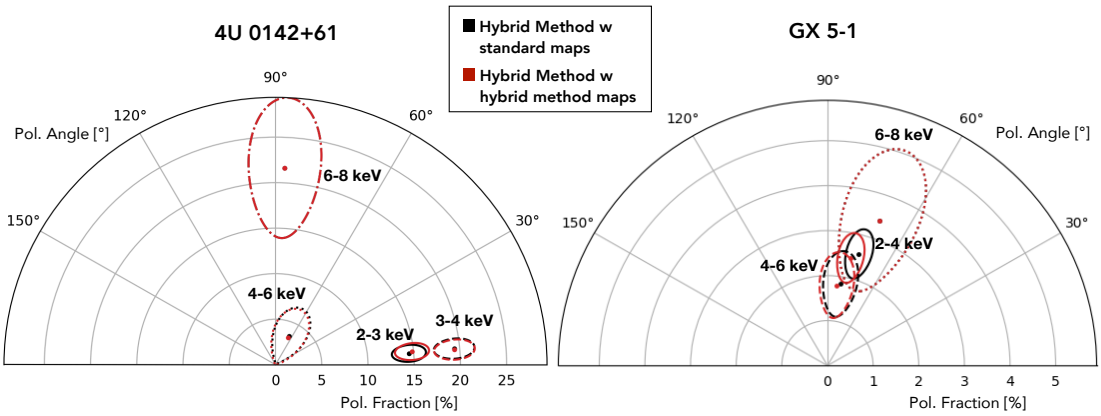

Figure 7.3: **Energy binned polarization polar plots**. Results are obtained with the hybrid method and corrected for the hybrid method spurious maps (*red*), and for the standard moment analysis spurious maps (*black*). The data refer to the magnetar 4U-0142+61 (*left panel*) and the accreting neutron star GX 5-1 (*right panel*), and consider only events impacting on the central area of DU1. The 6-8 keV bin results perfectly overlap, as the spurious modulation for these energies is negligible.

the analysis of the astrophysical sources by employing all the Detector Units, rather than being restricted to the central part of DU1.

7.3 Benchmark examples

In this section we assume that the compatibility just discussed is true for all observations and across all Detector Units. While minor differences may have a marginal impact on the results, we are confident that this assumption does not provide significantly misleading outcomes. However, it is essential to note that the results presented in this section are useful to give an indication about the hybrid method potential. For comprehensive and completely reliable outcomes, the spurious maps for all the three detectors and for their entire sensitive area must be computed. For this reason, the results and considerations reported in the next pages attain only to the comparison between the standard moment analysis and the hybrid method, and the potential impact of the latter in IXPE data analysis. The science-related results and implications are beyond the purpose of this section, and the reader can refer to the cited literature.

Point source: GX 5-1

The hybrid method, validated through both simulations and calibration data, demonstrated an enhancement in performance with respect to the IXPE standard reconstruction algorithm, particularly at medium and high energies. However, due to the power-

4U 0142+61		
Energy [keV]	H_{HM} **PD [%]**	H_{SM} **PD [%]**
2-3	14.8 ± 1.9	14.5 ± 1.9
3-4	19.4 ± 2.2	19.4 ± 2.2
4-6	3.3 ± 3.3	3.4 ± 3.3
6-8	21.6 ± 7.7	21.6 ± 7.7
	H_{HM} **PA [°]**	H_{SM} **PA [°]**
2-3	5.6 ± 3.6	4.9 ± 3.7
3-4	4.9 ± 3.3	5.2 ± 3.3
4-6	65.7 ± 29.4	64.8 ± 27.8
6-8	87.4 ± 10.2	87.4 ± 10.2
GX 5-1		
Energy [keV]	H_{HM} **PD [%]**	H_{SM} **PD [%]**
2-4	2.4 ± 0.5	2.6 ± 0.5
4-6	1.7 ± 0.7	1.7 ± 0.7
6-8	4.3 ± 1.5	4.3 ± 1.5
	H_{HM} **PA [°]**	H_{SM} **PA [°]**
2-4	81.0 ± 6.4	76.8 ± 5.9
4-6	81.4 ± 11.7	79.0 ± 11.4
6-8	65.9 ± 10.3	65.9 ± 10.3

Table 7.1: **Hybrid method polarization degree and angle comparison with different spurious modulation maps.** Hybrid method results for two IXPE observations (4U 0142+61 and GX 5-1) are reported by using only the central region of DU1 data. Stokes parameters are corrected for the spurious map computed with the hybrid method (H_{HM}), or with the standard moment analysis (H_{SM}).

law distribution characterizing the spectra of nearly all astrophysical X-ray sources, observations are often dominated by low-energy photons. For this reason I conducted the analysis of several point sources, focusing on the ones where energy-resolved polarization measurements have already been conducted by the IXPE collaboration. I thus divided data into energy bins and estimated the polarization of the selected sources by employing both the reconstruction algorithms, with the aim of highlighting the significance of the improvements introduced by the hybrid approach.

As a benchmark example, results from the analysis of the neutron star X-ray binary GX 5-1 are reported, but similar outcomes were achieved for other sources (specifically Cyg-X1 [94], 4U 0142+61 [101] and 4U 1630-47 [97]). GX 5-1 was observed twice by IXPE, from March 21st to 22nd and from April 13th to 15th, 2023, for a total observational time of approximately 100 ks. Results achieved with the standard algorithm

were shortly described in Sec. 3.1 and discussed in Fabiani et al. [92]. In Fig. 7.4 the IXPE lightcurve of GX 5-1 is reported. The two panels represent the two observational windows, while the colors refer to the single Detector Unit rates.

I performed the analysis for the two observations independently, and for the joint observation too, under the assumption that the geometry and the physical process producing polarization do not change.

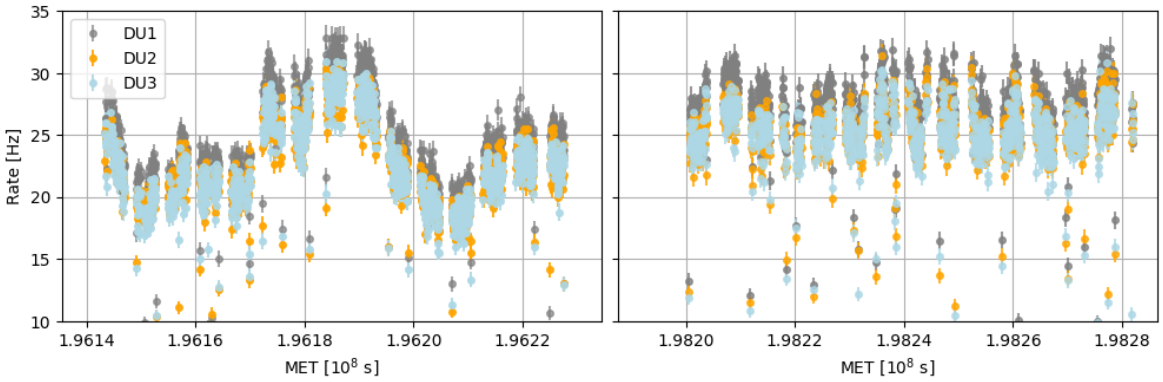

Figure 7.4: **Lightcurve of GX 5-1 IXPE observations**. *Left panel* refers to the first observation (March 21st to 22nd, 2023), while *right panel* refers to the second observation (April 13th to 15th, 2023). The three DUs rates are reported separately, and each DU is characterize by a different color (DU1: *gray* | DU2: *orange* | DU3: *light blue*). The Mission Elapsed Time (MET) is the amount of time recorded since IXPE launch.

I present here a model-independent analysis of the source polarization properties by employing 1 keV wide energy bins. The results are shown in Fig. 7.5. The *left panels* correspond to the first observation of GX 5-1, the *central panels* to the second observation, and the *right panels* show the results for the combined data. Results from the hybrid method are shown in *red*, while those from the standard moment analysis are shown in *black*.

		OBS 1	**OBS 2**	**OBS 1+2**
		m $[10^{-3} keV^{-1}]$	**m** $[10^{-3} keV^{-1}]$	**m** $[10^{-3} keV^{-1}]$
Q/I	Standard Mom. Analysis	1.9 ± 2.2 (p=0.86)	0.3 ± 1.1 (p=0.78)	1.2 ± 1.5 (p=0.42)
	Hybrid Method	3.5 ± 1.8 (p=0.05)	0.6 ± 2.3(p=0.79)	2.4 ± 1.6 (p=0.13)
U/I	Standard Mom. Analysis	-6.1 ± 3.2 (p=0.06)	-7.5 ± 2.1 (p<0.05)	-6.9 ± 1.4 (p<0.05)
	Hybrid Method	-5.7 ± 2.5 (p<0.05)	-7.5 ± 2.0 (p<0.05)	-6.7 ± 1.4 (p<0.05)

Table 7.2: **Angular coefficients of the best-fit linear model obtained for the Q/I and U/I plots**. Results are reported for the standard moment analysis and for the hybrid method (y = **m**x + q), with the relative *p-values*. Results flagged by are not compatible with m=0 at 1σ.

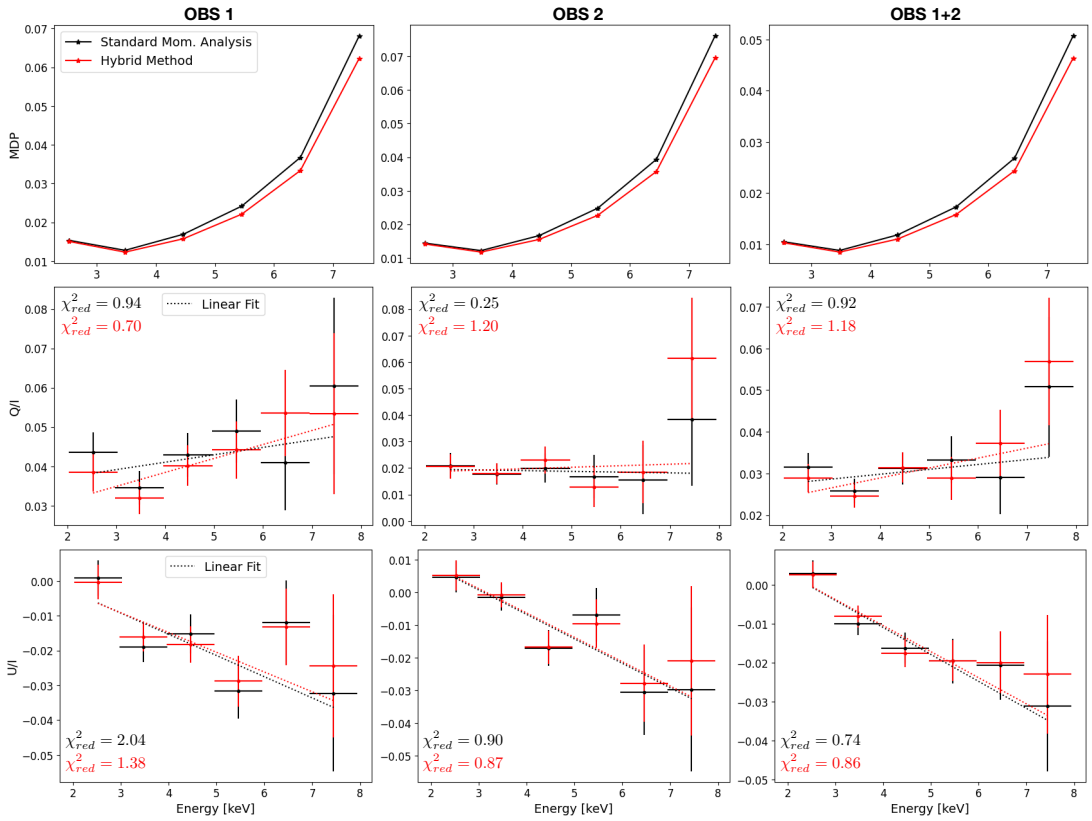

Figure 7.5: **Energy-binned results for the IXPE analysis of GX 5-1 employing the standard moment analysis (*black*) and the hybrid methiod (*red*).** Column-wise, the results are divided by observations (1st obs. | 2nd obs. | combined obs.). The *upper panels* report the MDP as a function of energy. The *middle and lower panels* display respectively the Q/I and U/I values as a function of energy. The *dashed lines* correspond to the best-fit linear model. Reduced χ^2 values are reported as well, with $n_{dof} = 4$. All χ^2 values fall within the 5-95% confidence interval.

Initially, the Minimum Detectable Polarization (MDP) at a 99% confidence level is evaluated for each energy bin. The obtained values for both the standard moment analysis and the hybrid method are compared in the *top panels* of Fig. 7.5. A slight improvement in MDP at medium and high energies is observed when employing the hybrid method. $\mathrm{MDP_{HYB}}/\mathrm{MDP_{STD}}$ ranges from a minimum of 0.98 in the 2-3 keV bin up to a maximum of 0.90 in the 6-7 keV bin. This indicates that the same MDP would require slightly shorter observations using the hybrid algorithm.

In the *middle and lower panels*, the Stokes parameters as a function of energy are presented, together with the best-fit linear model (*dashed lines*). I consider the compatibility of the angular coefficient m of the linear fit with zero: the non-compatibility would result in the Q/I and U/I parameters significantly deviating from being constant

with energy. The achieved coefficients are reported in Table 7.2, alongside the probabilities of m being compatible with zero (*p-value*). Both the standard moment analysis and the hybrid method agree on the findings for the U parameter, showing a negative angular coefficient significantly deviating from zero at 1σ. However, setting the confidence level at 95%, only the hybrid method presents significant results for all the three observations. Concerning the Q values fits, while the standard moment analysis finds angular coefficients compatible with zero at 1σ for both individual and combined observations, the hybrid method indicates positive values that are not compatible with zero at 1σ for the first and the combined observations. However, the dependence of the Q parameter on energy is mild, as the results are not significant at the confidence level of 95%. Similar outcomes are achieved using 5 instead of 6 bins and spacing them equally in logarithmic scale to increase statistics at higher energies. Results are presented here only for the 6-bin scenario as a greater number of bins is advantageous for a linear fit.

Both methods suggest a polarization direction changing with energy, but the hybrid method gives a slightly stronger indication of an increasing polarization degree with energy, albeit with low significance. The polarization degree and angle are illustrated in Fig. 7.6 as functions of energy, for the combined observations and for both methods.

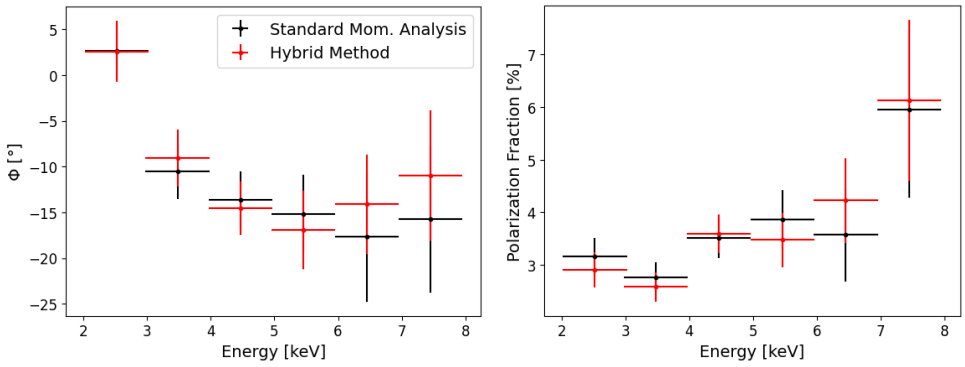

Figure 7.6: **Energy-binned polarization angle (*left panel*) and degree (*right panel*) for IXPE combined observation of GX 5-1**. Results for the standard moment analysis are reported in *black*, while results for the hybrid method are reported in *red*. Error bars represent the 1σ uncertainty.

Employing the hybrid method for the analysis of point sources where a dependency of the polarization properties on energy is detectable, could allow to shorten the observations, thanks to the achievement of lower MDP values. For the same observational time, hybrid method could improve the sensitivity of the polarization measurements at medium and high energies, potentially enhancing the significance of a result. For the specific case of GX 5-1, as already discussed in Sec. 3.1, in the paper published by the collaboration is mentioned the possibility of an increasing polarization fraction with energy, though the results are not significant enough [92]. A slightly more robust

evidence of this could be possible by employing the hybrid method.

Extended sources: G21.5

As discussed in the previous chapters, the polarization leakage effect mainly contributes to the systematics of the extended sources observations, especially when sharp intensity gradients and edges are involved. When sources are very bright and present a high polarization degree (as discussed for Crab and Vela nebulae), polarization leakage could be considered negligible, while for other sources it could represent a substantial limitation to the analysis. Chapter 5 demonstrated the potential impact of the hybrid method on the reduction of the polarization leakage in the analysis of extended sources with simulated data. Here an example of the potential improvement on experimental data is reported, specifically for the case of G21.5, a supernova remnant observed by IXPE during October 2023, for a total observational time of $\sim$ 350ks [Di Lalla et al., in prep]. An interpolated intensity map of the source as seen by IXPE is reported in Fig. 7.7

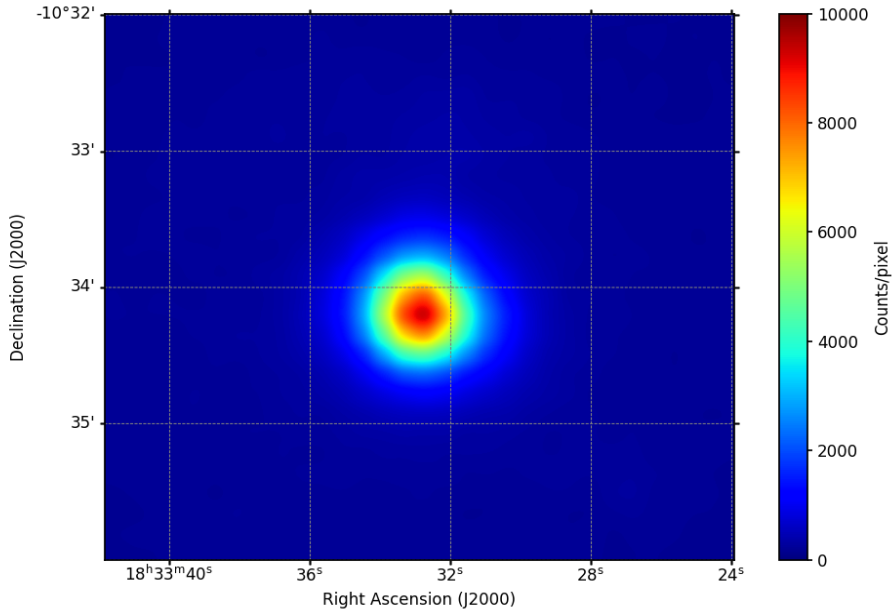

Figure 7.7: **Interpolated count map of G21.5 as detected by IXPE**. All three detector units are considered for computing this map.

Firstly, in order to quantify the contribution of the polarization leakage effect to the measurements, an observation of a G21.5-like unpolarized source is simulated through *ixpesim* software. This is a very common key process in IXPE extended sources analysis: the source is simulated as unpolarized, so that every hint of significant residual

polarization resulting from the analysis is attributed to the leakage. In Fig. 7.8 the binned residual polarization detected by the standard moment analysis (*left panel*), and by the hybrid method (*right panel*), are reported. Only the bins for which the measured polarization is significant at least at 3σ are shown. A strong evidence for polarization leakage effect emerges from the moment analysis results. A notable radial polarization is observed where the intensity gradient of the source is most prominent. As for the case of the unpolarized Crab nebula (see Fig. 5.19), a significant reduction of the polarization leakage effect is achieved with the employment of the hybrid method.

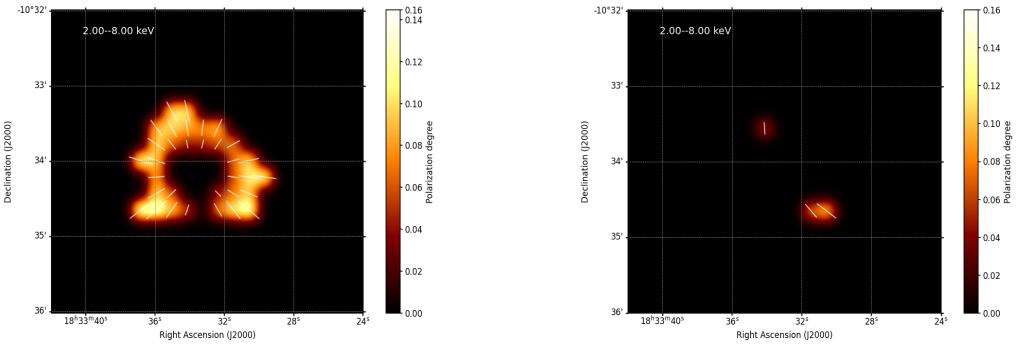

Figure 7.8: **Binned polarization map of the unpolarized simulation of G21.5** Polarization degree significantly ($\geq 3\sigma$) detected by the standard moment analysis (*left panel*) and by the hybrid method (*right panel*) for the unpolarized simulation of G21.5 in the 2-8 keV energy interval. The thin white lines indicate the measured polarization direction.

The radial polarization fraction of the simulated unpolarized source is then evaluated by integrating the signal from several circular annuli around the center of the source, and then reported as a function of the distance from the center in Fig. 7.9 as a *shaded band*. The color of the band refers to the polarization direction of the signal: a *redder* color stands for a polarization orientation parallel to the radial direction, while a *blue* color stands for a perpendicular direction. The band width represents the 1σ uncertainty. Comparing the band in the *left panel* (standard moment analysis) with the one in the *right panel* (hybrid method) it is possible to infer again a reduction of the polarization leakage. At distances of 0.6-0.8′ from the center, the standard moment analysis finds residual polarization degree up to ∼15%, while the hybrid method PD never exceeds 10%.

IXPE data from the observation of G21.5 are processed as usual with both the standard moment analysis and with the hybrid method. The radial polarization fraction is evaluated for two annuli, defined by $0.2' < \mathrm{R} < 0.6'$ and $0.4' < \mathrm{R} < 0.8'$. Results are reported as markers with error bars in Fig. 7.9, with the color again referring to the orientation of the polarization direction. According to the simulation outcomes a radial polarization would be expected in case of unpolarized source.

In the *left panel*, results for the standard moment analysis are presented. A radial polarization trend is indeed detected, although lower than the expected value due to polarization leakage only (*shaded band*). One plausible explanation could be the presence of a tangential polarization, which partially balances the contribution of the radial polarization caused by leakage. However an explicit detection of such tangential polarization is lacking, and any assumption regarding its presence remains speculative.

On the *right panel*, the results for the hybrid method are shown: a tangential polarization clearly emerges from the analysis. The measurement has a relatively low significance ($\sim 2\sigma$), but makes explicit a detection which could have only been suggested with the standard moment analysis. In the *lower panel* of Fig. 7.9, the polar plots for the radial polarization degree and angle for the radiation coming from the annulus defined by $0.2' < \mathrm{R} < 0.6'$ are reported as well, with the relative 1σ uncertainties. These lower panel plots quantify what expressed through colors in the upper panel plots: for the standard moment analysis the predicted polarization direction is almost parallel to the radial direction, while for the hybrid method it is almost perpendicular.

At the beginning of this section, it was emphasized that the reported results are plausible, but represent only an indication of the measurements that could be obtained using the hybrid algorithm, pending the utilization of accurate spurious modulation maps. Nevertheless, these findings underscore the algorithm potential in the context of extended sources. Improved reconstruction of the impact point position could be highly beneficial in determining the polarization characteristics of this source and others, where polarization leakage significantly influences observations.

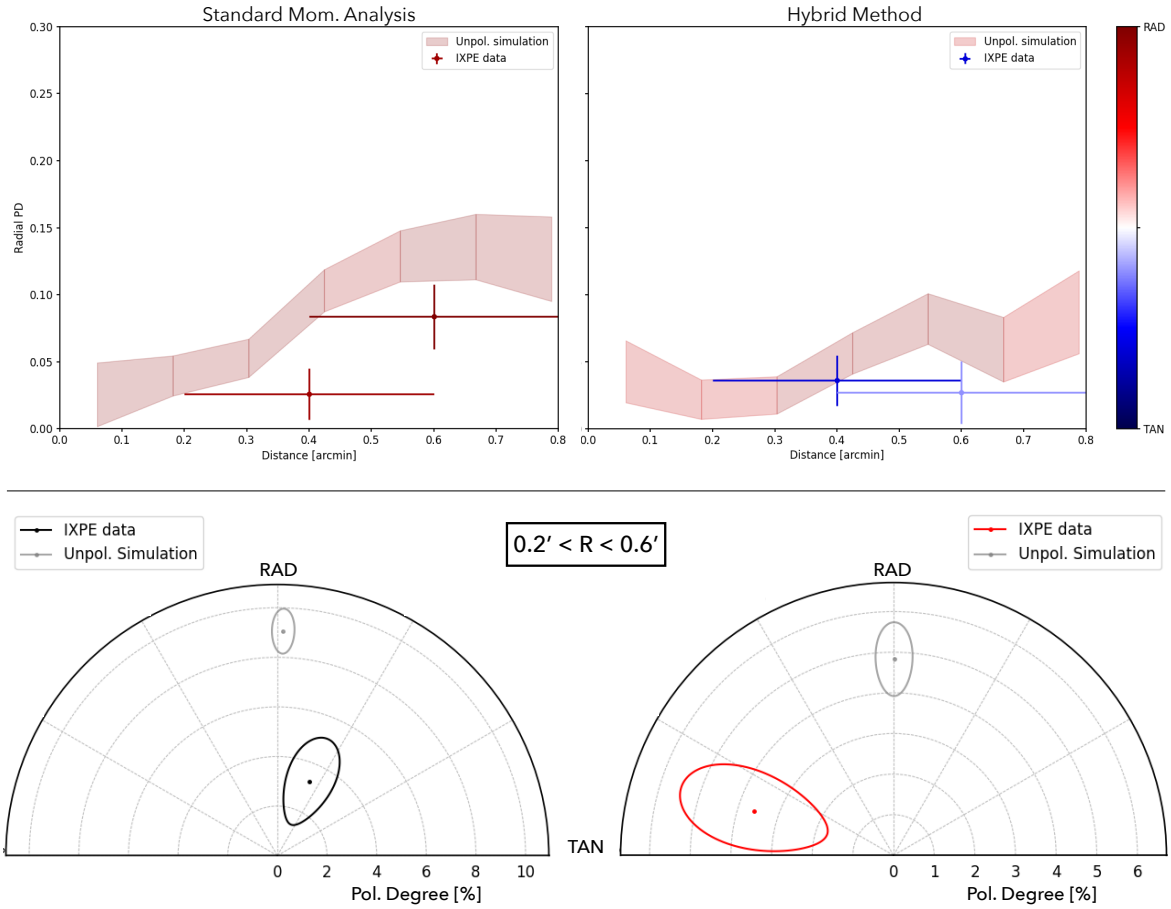

Figure 7.9: **G21.5 IXPE data analysis results.** *Upper panels*: radial polarization for the G21.5 unpolarized simulation (*shaded band*) and for the G21.5 IXPE data (*markers with error bars*) as a function of the distance from the center of the source. The vertical error bars represent the 1σ uncertainty. The color of both the band and the markers refer to the orientation of the predicted polarization respect to the radial direction. *Lower panels*: Polar plots of the radial polarization detected for the annulus defined by $0.2' < \mathrm{R} < 0.6'$. The *gray point* refers to the simulated data, while the *black* (standard moment analysis, *left panel*) and the *red* (hybrid method, *right panel*) points refer to IXPE data. *Ellipses* represent the 1σ uncertainties.

Conclusions

X-ray polarimetry has emerged as a pioneering field within recent astrophysics. With the advent of IXPE, it has unveiled new scientific findings, which both confirmed and reshaped our understanding of astrophysical phenomena. This field undergoes continuous evolution in scientific paradigms and data analysis methodologies, guiding the development of the future X-ray polarimetry missions.

This thesis presented a novel hybrid analytic-Machine Learning approach for the reconstruction of photo-electron tracks in Gas Pixel Detectors, the detectors onboard IXPE. We developed a Convolutional Neural Network for predicting the photon impact point position on the detector, starting from the analysis of the track image. Our architecture is based on DenseNet-121 with hexagonal convolutional layers implemented as a C++ extension for PyTorch.

We demonstrated significant improvement in the IP reconstruction compared to the standard state-of-the-art analytic algorithm, denoted as moment analysis. An artificial sharpening process applied to images further improved the performance. Compared to the standard moment analysis, the mean distance between the true and the predicted impact point was reduced by $\sim$20% at 3 keV and by $\sim$50% at 8 keV. By employing the CNN-predicted impact point in the moment analysis, we observed an enhancement in the modulation factor, from $\sim$1% at low energies to $\sim$6% at higher energies, demonstrating the importance of precise determination of the impact point in the reconstruction process. Our hybrid method also significantly mitigated the systematic effect denominated *polarization leakage*, both in point sources and extended sources observations. While there is not a standardized metric for quantifying this improvement, we can state that the residual radial modulation of unpolarized point sources was reduced by a factor of $\sim$2 compared to the standard moment analysis.

The validation process using ground laboratory calibration data further confirmed the algorithm performance. Measurements of unpolarized beams showed no unexpected residual polarization after the correction for the spurious modulation, indicating the absence of biases in unpolarized radiation detection. Moreover, the modulation factors for 100% polarized beams align with simulation trends, consolidating the algorithm reliability with experimental data.

IXPE data analysis showed the potential improvement achievable with the hybrid method. In the context of extended sources, it can enhance the significance of results where the standard moment analysis is limited by the polarization leakage effect. Concerning point sources, the improved performance in the modulation factor recovery at medium and high energies could shorten the observational times or reduce the uncertainties linked to the polarization detection in the 4-8 keV energy band.

In summary, our hybrid analytic-ML approach offers advancements in GPD-based X-ray polarimetry, enhancing performance, mitigating systematic effects, and demonstrating robustness with experimental data. While initially conceived for future X-ray polarimetry missions, IXPE data analysis results demonstrated the value of this new

approach also for the current measurements, and the collaboration is considering the possibility to integrate it in the official IXPE pipeline, which now employs the standard moment analysis only.

Acronyms

ADC Analog to Digital Converter

AGN Active Galactic Nucleus

ASI Agenzia Spaziale Italiana

ASIC Application-Specific Integrated Circuit

BEE Back-End Electronics

BB Black Body

CCD Charge-Coupled Device

CNN Convolutional Neural Network

CPU Central Processing Unit

DAQ Data AcQuisition

DNN Dense Neural Network

DME DiMethyl Ether

DSU Detector Service Unit

DU Detector Unit

FITS Flexible Image Transport System

GEM Gas Electron Multiplier

GPD Gas Pixel Detector

GPU Graphics Processing Unit

HEASARC High Energy Astrophysics Science Archive Research Center

HPD Half-Power Diameter

IAPS Istituto di Astrofisica e Planetologia Spaziali

ICE Instrument Calibration Equipment

INAF Istituto Nazionale di Astrofisica

INFN Istituto Nazionale di Fisica Nucleare

IP Impact Point

IXPE Imaging X-ray Polarimetry Explorer

LTU Linear Threshold Unit

MC Monte Carlo

MDP Minimum Detectable Polarization

MET Mission Elapsed Time

ML Machine Learning

MLP Multi-Layer Perceptron

MMA Mirror Module Assembly

MSFC Marshall Space Flight Center

NASA National Aeronautics and Space Administration

NN Neural Network

OSO-8 Orbiting Solar Observatory 8

PA Polarization Angle

PD Polarization Degree

PE Photo-Electron

PHA Pulse Height Amplitude

PL Power Law

PSF Point Spread Function

PWN Pulsar Wind Nebula

ReLU Rectified Linear Unit

SDD Silicon Drift Detector

SMBH Super Massive Black Hole

SNR SuperNova Remnant

ToO Target of Opportunity

XCF X-ray Calibration Facility

List of Figures

Bibliography

[1] C. Sgrò et al. The gas pixel detector on board the IXPE mission. In *UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XX*. SPIE, 2017.

[2] M. C. Weisskopf et al. The Imaging X-Ray Polarimetry Explorer (IXPE): Pre-Launch. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2021.

[3] R. Bellazzini et al. Novel gaseus X-ray polarimeter: data analysis and simulation. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2003.

[4] T. Kitaguchi et al. A convolutional neural network approach for reconstructing polarization information of photoelectric X-ray polarimeters. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2019.

[5] N. Moriakov et al. Inferring astrophysical X-ray polarization with deep learning. *arXiv e-prints*, 2020.

[6] A. L. Peirson et al. Deep ensemble analysis for Imaging X-ray Polarimetry. *Nuclear Instruments and Methods in Physics Research A*, 2021.

[7] G. B. Rybicki and A. P. Lightman. *Radiative Processes in Astrophysics*. Wiley, 1979.

[8] F. Kislat et al. Analyzing the data from X-ray polarimeters with Stokes parameters. *Astroparticle Physics*, 2015.

[9] S. Fabiani. Instrumentation and Future Missions in the Upcoming Era of X-ray Polarimetry. *Galaxies*, 2018.

[10] R. Novick et al. Detection of X-Ray Polarization of the Crab Nebula. *The Astrophysical Journal Letters*, 1972.

[11] S. Fabiani and F. Muleri. *Astronomical X-ray Polarimetry*. Aracne, 2014.

[12] F. Muleri et al. The IXPE instrument calibration equipment. *Astroparticle Physics*, 2022.

[13] A. Di Marco et al. Calibration of the IXPE Focal Plane X-Ray Polarimeters to Polarized Radiation. *The Astronomical Journal*, 2022.

[14] M. Chandreyee and P. Biswajit. Prospect of polarization measurements from black hole binaries in their thermal state with a scattering polarimeter. *Monthly Notices of the Royal Astronomical Society*, 2011.

[15] W. Heitler. *The Quantum Theory of Radiation*. 1954.

[16] V.A. Dergachev et al. Hard X-ray compton polarimetry with the PENGUIN-M instrument in the spaceborne experiment CORONAS-PHOTON. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Fizicheskaya*, 2009.

[17] Baldini L. et al. Design, construction, and test of the Gas Pixel Detectors for the IXPE mission. *Astroparticle Physics*, 2021.

[18] G. F. Carpenter et al. Ariel 5 Hard X-Ray Measurements of Galactic and Extra-galactic X-Ray Sources. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 1976.

[19] M. C. Weisskopf et al. The X-ray polarization experiment on the OSO-8. In *X-ray binaries*, 1976.

[20] M. C. Weisskopf et al. A precision measurement of the X-ray polarization of the Crab Nebula without pulsar contamination. *The Astrophysical Journal Letters*, 1978.

[21] S. Chandrasekhar. *Radiative Transfer*. 1960.

[22] G. R. Rieglerand others. A low-energy X-ray polarimeter. *Nuclear Instruments and Methods in Physics Research*, 1970.

[23] P. W. Sanford, A. M. Cruise, and J. L. Culhane. *Techniques for Improving the Sensitivity of Proportional Counters Used in X-ray Astronomy*. Springer Netherlands, 1970.

[24] P. Soffitta et al. Sensitivity to x-ray polarization of a microgap gas proportional counter. *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 1995.

[25] H. Feng and P. Kaaret. The HaloSat and PolarLight CubeSat Missions for X-Ray Astrophysics. In *Handbook of X-ray and Gamma-ray Astrophysics*. Springer Nature Singapore, 2022.

[26] H. Feng et al. Re-detection and a possible time variation of soft X-ray polarization from the Crab. *Nature Astronomy*, 2020.

[27] IXPE archive. URL: https://heasarc.gsfc.nasa.gov/docs/ixpe/archive/.

[28] J. Vink et al. X-Ray Polarization Detection of Cassiopeia A with IXPE. *The Astrophysical Journal*, 2022.

[29] IXPE Collaboration. For Scientists: The Long-Term Plan. URL: https://ixpe.msfc.nasa.gov/for_scientists/ltp.html.

[30] J. Harvey, P. Thompson, and A. Krywonos. Hyperboloid-hyperboloid grazing incidence x-ray telescope designs for wide-field imaging applications. In *X-Ray Optics, Instruments, and Missions III*, 2000.

[31] S. D. Bongiorno et al. Assembly of the IXPE mirror modules. In *Optics for EUV, X-Ray, and Gamma-Ray Astronomy X*, 2021.

[32] F. Sauli. Gem: A new concept for electron amplification in gas detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1997.

[33] C. Caputo. Physics Studies for the CMS muon system upgrade with triple-GEM detectors. *Journal of Instrumentation*, 2014.

[34] R. Bellazzini et al. Direct reading of charge multipliers with a self-triggering CMOS analog chip with 105k pixels at 50μm pitch. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2006.

[35] M. S. Longair. *High energy astrophysics*. Cambridge University Press, 2011.

[36] H. Krawczynski et al. Scientific prospects for hard X-ray polarimetry. *Astroparticle Physics*, 2011.

[37] T. E. Strohmayer and T. R. Kallman. On the Statistical Analysis of X-Ray Polarization Measurements. *The Astrophysical Journal*, 2013.

[38] A. Di Marco et al. A Weighted Analysis to Improve the X-Ray Polarization Sensitivity of the Imaging X-ray Polarimetry Explorer. *The Astronomical Journal*, 2022.

[39] A. L. Peirson and R. W. Romani. Toward Optimal Signal Extraction for Imaging X-Ray Polarimetry. *The Astrophysical Journal*, 2021.

[40] J. Rankin et al. An Algorithm to Calibrate and Correct the Response to Unpolarized Radiation of the X-Ray Polarimeter Onboard IXPE. *The Astronomical Journal*, 2022.

[41] Bucciantini, N. et al. Polarisation leakage due to errors in track reconstruction in gas pixel detectors. *Astronomy and Astrophysics*, 2023.

[42] R. D. Blandford and R. L. Znajek. Electromagnetic extraction of energy from Kerr black holes. *Monthly Notices of the Royal Astronomical Society*, 1977.

[43] A. P. Marscher. Relativistic jets and the continuum emission in QSOs. *The Astrophysical Journal*, 1980.

[44] A. P. Marscher and W. K. Gear. Models for high-frequency radio outbursts in extragalactic sources, with application to the early 1983 millimeter-to-infrared flare of 3C 273. *The Astrophysical Journal*, 1985.

[45] N. Fraija et al. Long-term Optical Polarization Variability and Multiwavelength Analysis of Blazar Mrk 421. *The Astrophysical Journal Supplement Series*, 2017.

[46] A. L. McNamara, Z. Kuncic, and K. Wu. X-ray polarization in relativistic jets. *Monthly Notices of the Royal Astronomical Society*, 2009.

[47] T. Schweizer et al. Characterization of the optical and X-ray properties of the north-western wisps in the Crab nebula. *Monthly Notices of the Royal Astronomical Society*, 2013.

[48] J. Vink. Supernova remnants: the X-ray perspective. *Astronomy and Astrophysics Reviews*, 2012.

[49] R. A. Remillard and J. E. McClintock. X-Ray Properties of Black-Hole Binaries. *Annual Review of Astronomy and Astrophysics*, 2006.

[50] X. Hua and L. Titarchuk. Comptonization Models and Spectroscopy of X-Ray and Gamma-Ray Sources: A Combined Study by Monte Carlo and Analytical Methods. *The Astrophysical Journal*, 1995.

[51] L. Maraschi, G. Ghisellini, and A. Celotti. A Jet Model for the Gamma-Ray–emitting Blazar 3C 279. *The Astrophysical Journal Letters*, 1992.

[52] C. D. Dermer, R. Schlickeiser, and A. Mastichiadis. High-energy gamma radiation from extragalactic radio sources. *Astronomy and Astrophysics*, 1992.

[53] J. Poutanen and O. Vilhu. Compton scattering of polarized light in two-phase accretion discs. *Astronomy and Astrophysics*, 1993.

[54] R. C. Duncan and C. Thompson. Formation of Very Strongly Magnetized Neutron Stars: Implications for Gamma-Ray Bursts. *The Astrophysical Journal Letters*, 1992.

[55] C. Thompson, M. Lyutikov, and S. R. Kulkarni. Electrodynamics of Magnetars: Implications for the Persistent X-Ray Emission and Spin-down of the Soft Gamma Repeaters and Anomalous X-Ray Pulsars. *The Astrophysical Journal*, 2002.

[56] R. Taverna et al. Probing magnetar magnetosphere through X-ray polarization measurements. *Monthly Notices of the Royal Astronomical Society*, 2013.

[57] M. Van Adelsberg and D. Lai. Atmosphere models of magnetized neutron stars: QED effects, radiation spectra and polarization signals. *Monthly Notices of the Royal Astronomical Society*, 2006.

[58] N. Bucciantini et al. Simultaneous space and phase resolved X-ray polarimetry of the Crab pulsar and nebula. *Nature Astronomy*, 2023.

[59] J. Dyks and A. K. Harding. Rotational Sweepback of Magnetic Field Lines in Geometric Models of Pulsar Radio Emission. *The Astrophysical Journal*, 2004.

[60] J. Pétri. Phase-resolved polarization properties of the pulsar striped wind synchrotron emission. *Monthly Notices of the Royal Astronomical Society*, 2013.

[61] B. Cerutti, A.A. Philippov, and G. Dubus. Dissipation of the striped pulsar wind and non-thermal particle acceleration: 3D PIC simulations. *Astronomy and Astrophysics*, 2020.

[62] D. M. Scott, M. H. Finger, and C. A. Wilson. Characterization of the timing noise of the Crab pulsar. *Monthly Notices of the Royal Astronomical Society*, 2003.

[63] N. Bucciantini et al. Modeling the effect of small-scale magnetic turbulence on the X-ray properties of Pulsar Wind Nebulae. *Monthly Notices of the Royal Astronomical Society*, 2017.

[64] P. Moran et al. Optical polarimetry of the inner Crab nebula and pulsar. *Monthly Notices of the Royal Astronomical Society*, 2013.

[65] R. Dodson et al. The radio nebula surrounding the Vela pulsar. *Monthly Notices of the Royal Astronomical Society*, 2003.

[66] F. Xie et al. Vela pulsar wind nebula X-rays are polarized to near the synchrotron limit. *Nature*, 2022.

[67] I. Robson. *Active galactic nuclei*. 1996.

[68] D. E. Osterbrock and G. J. Ferland. *Astrophysics of gaseous nebulae and active galactic nuclei*. 2006.

[69] R. Antonucci. Unified models for active galactic nuclei and quasars. *Annual Review of Astronomy and Astrophysics*, 1993.

[70] C. M. Urry and P. Padovani. Unified Schemes for Radio-Loud Active Galactic Nuclei. *Publications of the Astronomical Society of the Pacific*, 1995.

[71] C. M. Urry. The AGN Paradigm for Radio-Loud Objects. *Active Galactic Nuclei: from Central Engine to Host Galaxy*, 2003.

[72] V. Acciari et al. Spectral Energy Distribution of Markarian 501: Quiescent State vs. Extreme Outburst. *Astrophysical Journal*, 2010.

[73] I. Liodakis et al. Polarized blazar X-rays imply particle acceleration in shocks. *Nature*, 2022.

[74] F Tavecchio et al. Probing dissipation mechanisms in BL Lac jets through X-ray polarimetry. *Monthly Notices of the Royal Astronomical Society*, 2018.

[75] L. Di Gesu et al. Discovery of X-ray polarization angle rotation in the jet from blazar Mrk 421. *Nature Astronomy*, 2023.

[76] L. Di Gesu et al. The X-Ray Polarization View of Mrk 421 in an Average Flux State as Observed by the Imaging X-Ray Polarimetry Explorer. *The Astrophysical Journal Letters*, 2022.

[77] H. Zhang et al. Radiation and Polarization Signatures from Magnetic Reconnection in Relativistic Jets. II. Connection with γ-Rays. *The Astrophysical Journal*, 2022.

[78] L. Di Gesu et al. Testing particle acceleration models for BL Lac jets with the Imaging X-ray Polarimetry Explorer. *Astronomy and Astrophysics*, 2022.

[79] V. M. Larionov et al. The Outburst of the Blazar S5 0716+71 in 2011 October: Shock in a Helical Jet. *The Astrophysical Journal*, 2013.

[80] H. Zhang et al. Polarization signatures of relativistic magnetohydrodynamic shocks in the blazar emission region. I. force-free helical magnetic fields. *The Astrophysical Journal*, 2016.

[81] D. E. Kim et al. Magnetic field properties inside the jet of Mrk 421 - Multiwavelength polarimetry, including the Imaging X-ray Polarimetry Explorer. *Astronomy and Astrophysics*, 2024.

[82] H. L. Marshall. Analysis of polarimetry data with angular uncertainties. *The Astronomical Journal*, 2021.

[83] V.L. Ginzburg and S.I. Syrovatskii. Chapter IV - The origin of Cosmic Rays. In *The Origin of Cosmic Rays*. 1964.

[84] G. Dubner and E. Giacani. Radio emission from Supernova Remnants. *The Astronomy and Astrophysics Review*, 2015.

[85] B. Jun and M. L. Norman. On the Origin of Strong Magnetic Fields in Young Supernova Remnants. *The Astrophysical Journal*, 1996.

[86] V. N. Zirakashvili and V. S. Ptuskin. Diffusive Shock Acceleration with Magnetic Amplification by Nonresonant Streaming Instability in Supernova Remnants. *The Astrophysical Journal*, 2008.

[87] R. Ferrazzoli et al. X-Ray Polarimetry Reveals the Magnetic-field Topology on Sub-parsec Scales in Tycho's Supernova Remnant. *The Astrophysical Journal*, 2023.

[88] P. Zhou et al. Magnetic Structures and Turbulence in SN 1006 Revealed with Imaging X-Ray Polarimetry. *The Astrophysical Journal*, 2023.

[89] V. Doroshenko et al. Determination of X-ray pulsar geometry with IXPE polarimetry. *Nature Astronomy*, 2022.

[90] I. Caiazzo and J. Heyl. Polarization of accreting X-ray pulsars - II. Hercules X-1. *Monthly Notices of the Royal Astronomical Society*, 2021.

[91] S. Blum and U. Kraus. Analyzing X-Ray Pulsar Profiles: Geometry and Beam Pattern of Hercules X-1. *The Astrophysical Journal*, 2000.

[92] S. Fabiani et al. Discovery of a variable energy-dependent X-ray polarization in the accreting neutron star GX 5-1. *Submitted to Astronomy and Astrophysics*, 2023.

[93] R. Popham and R. Sunyaev. Accretion Disk Boundary Layers around Neutron Stars: X-Ray Production in Low-Mass X-Ray Binaries. *The Astrophysical Journal*, 2001.

[94] H. Krawczynski et al. Polarized x-rays constrain the disk-jet geometry in the black hole x-ray binary Cygnus X-1. *Science*, 2022.

[95] L. Li, R. Narayan, and J. E. McClintock. Inferring the inclination of a black hole accretion disk from observations of its polarized continuum radiation. *The Astrophysical Journal*, 2009.

[96] J. M. Bardeen and J. A. Petterson. The Lense-Thirring Effect and Accretion Disks around Kerr Black Holes. *The Astrophysical Journal Letters*, 1975.

[97] A. Ratheesh et al. The high polarisation of the X-rays from the Black Hole X-ray Binary 4U 1630-47 challenges standard thin accretion disc scenario. *Submitted to Nature Astronomy*, 2023.

[98] A. Veledina et al. Astronomical puzzle Cyg X-3 is a hidden Galactic ultraluminous X-ray source. *Submitted to Nature Astronomy*, 2023.

[99] Vilhu, O. et al. Orbital modulation of X-ray emission lines in Cygnus X-3. *Astronomy and Astrophysics*, 2009.

[100] A. J. Mioduszewski et al. A One-sided Highly Relativistic Jet from Cygnus X-3. *The Astrophysical Journal*, 2001.

[101] R. Taverna et al. Polarized x-rays from a magnetar. *Science*, 2022.

[102] S. Zane et al. A Strong X-Ray Polarization Signal from the Magnetar 1RXS J170849.0-400910. *The Astrophysical Journal Letters*, 2023.

[103] Roberto Turolla et al. IXPE and XMM-Newton Observations of the Soft Gamma Repeater SGR 1806-20. *The Astrophysical Journal*, 2023.

[104] J. Heyl et al. The detection of polarized X-ray emission from the magnetar 1E 2259+586. *Monthly Notices of the Royal Astronomical Society*, 2023.

[105] S. A. Olausen and V. M. Kaspi. The McGill Magnetar Catalog. *The Astrophysical Journal Supplement Series*, 2014.

[106] R. Taverna et al. Polarization of neutron star surface emission: a systematic analysis. *Monthly Notices of the Royal Astronomical Society*, 2015.

[107] S. Lesage et al. Fermi-GBM Discovery of GRB 221009A: An Extraordinarily Bright GRB from Onset to Afterglow. *The Astrophysical Journal Letters*, 2023.

[108] M. Negro et al. The IXPE View of GRB 221009A. *The Astrophysical Journal Letters*, 2023.

[109] W. Mcculloch and W. Pitts. A Logical Calculus of Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 1943.

[110] F. Rosenblatt. A probabilistic model for information storage and organization in the brain. *Psychological Review*, 1958.

[111] PyTorch Contributors. PyTorch Documentation. URL: https://pytorch.org/docs/stable/index.html.

[112] M. Minsky and S. Papert. *Perceptrons*. MIT Press, 1969.

[113] T. Nitta. Solving the XOR problem and the detection of symmetry using a single complex-valued neuron. *Neural Networks*, 2003.

[114] F. Murtagh. Multilayer perceptrons for classification and regression. *Neurocomputing*, 1991.

[115] R. M. Schmidt. Recurrent Neural Networks (RNNs): A gentle Introduction and Overview. *arXiv e-prints*, 2019.

[116] A. Vaswani et al. Attention is All you Need. In *Advances in Neural Information Processing Systems*, 2017.

[117] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 1986.

[118] N. Morgan and H. Bourlard. Generalization and Parameter Estimation in Feedforward Nets: Some Experiments. In *Advances in Neural Information Processing Systems*, 1989.

[119] R. Robinson. MLNotebook. URL: https://mlnotebook.github.io/post/CNN1/.

[120] Y. Lecun et al. Gradient-based learning applied to document recognition. In *Proceedings of the IEEE*, 1998.

[121] A. Krizhevsky et al. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*, 2012.

[122] J. Deng et al. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[123] W. J. Dally, S. W. Keckler, and D. B. Kirk. Evolution of the Graphics Processing Unit (GPU). *IEEE Micro*, 2021.

[124] C. Szegedy et al. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[125] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *International Conference on Learning Representations*, 2015.

[126] K. He et al. Deep Residual Learning for Image Recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[127] G. Huang et al. Densely Connected Convolutional Networks. *arXiv e-prints*, 2016.

[128] R. K Srivastava, K. Greff, and J. Schmidhuber. Highway Networks. *CoRR*, 2015.

[129] G. Larsson, M. Maire, and G. Shakhnarovich. FractalNet: Ultra-Deep Neural Networks without Residuals. *CoRR*, 2016.

[130] C. Lee et al. Deeply-Supervised Nets. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, 2015.

[131] B. Lakshminarayanan, A. Pritzel, and C. Blundell. Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles. *arXiv e-prints*, 2016.

[132] Cibrario, N. et al. Joint machine learning and analytic track reconstruction for X-ray polarimetry with gas pixel detectors. *Astronomy and Astrophysics*, 2023.

[133] N. Di Lalla. *Study and optimization of the IXPE polarization-sensitive detectors*. PhD thesis, University of Pisa, 2019.

[134] C. Steppa and T. L. Holch. HexagDLy-Processing hexagonally sampled data with CNNs in PyTorch. *SoftwareX*, 2019.

[135] A. Shrivastava, A. Gupta, and R. Girshick. Training Region-based Object Detectors with Online Hard Example Mining. *arXiv e-prints*, 2016.

[136] NVIDIA. NVIDIA Tesla V100 GPU architecture. URL: https://images.nvidia.com/content/volta-architecture/pdf/volta-architecture-whitepaper.pdf.

[137] M. Ester et al. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. AAAI Press, 1996.

[138] M. C. Weisskopf. The Chandra X-Ray Observatory (CXO): An Overview. *arXiv e-prints*, 1999.

[139] A. Di Marco et al. Validation of neural network software by using IXPE ground calibration data. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2022.

[140] K. Dietz, A. Tennant, and S. O'Dell. IXPE Science Operations Center User Guide to Software, 2022. URL: https://heasarc.gsfc.nasa.gov/docs/ixpe/analysis/IXPE-SOC-DOC-009-UserGuide-Software.pdf.

[141] IXPE Calibration Database (CALDB). URL: https://heasarc.gsfc.nasa.gov/docs/ixpe/caldb/.

[142] L. Baldini et al. ixpeobssim: A simulation and analysis framework for the imaging X-ray polarimetry explorer. *SoftwareX*, 2022.