

Observing the Sun from start to finish: The Tunable magnetograph's design, calibration, data reduction and scientific exploitation.

PhD dissertation by

Pablo Santamarina Guerrero

Instituto de Astrofísica de Andalucía (IAA-CSIC)

Programa de Doctorado en Física y Matemáticas (FisyMat)
Universidad de Granada

A thesis submitted in fulfillment
of the requirements of the degree of
Doctor of Philosophy

October 22, 2024

PhD thesis supervised by

Dr. David Orozco Suárez

Dr. Julián Blanco Rodríguez



INSTITUTO DE
ASTROFÍSICA DE
ANDALUCÍA



EXCELENCIA
SEVERO
OCHOA



UNIVERSIDAD
DE GRANADA

CONTENTS

1	Introduction	3
1.1	Background	3
1.2	Motivation of our work	4
1.3	Introduction	6
1.4	A brief introduction to spectropolarimeters.	7
1.4.1	Imaging and optical quality.	7
1.4.2	Spectroscopy	13
1.4.3	Polarimetry	13
1.4.4	What do spectrpolarimeters tell us about the Sun?	16
2	Sunrise III and TuMag: Design and calibration.	19
2.1	Sunrise III	20
2.1.1	Observatory's design	21
2.1.2	Science with Sunrise.	22
2.2	The Tunable magnetograph: TuMag	24
2.2.1	TuMag's design and light path.	24
2.2.2	Instrument performance and verification.	26
2.2.2.1	Imaging performance.	27
2.2.2.2	Spectral performance.	29
2.2.2.3	Polarimetric performance.	33
3	TuMag's pipeline and data.	37
3.1	TuMag's observing modes	38
3.1.1	Calibration modes	40
3.1.1.1	Flat-field observations	40
3.1.1.2	Dark-current observations	41
3.1.1.3	Linear polarizer and micropolarizers observations.	41
3.1.1.4	Pinhole Observations.	42
3.1.1.5	Straylight target.	42
3.1.1.6	Prefilter scans.	42
3.1.1.7	Phase diversity.	43
3.2	Timelines	43
3.3	Pipeline	47

3.3.1	Standard data reduction process.	47
3.3.2	Extra calibration blocks.	50
3.3.3	Image reconstruction.	50
3.3.3.1	Linear polarizer calibration.	50
3.3.3.2	Micropolarizers calibration.	50
3.3.3.3	Prefilter scans.	50
4	Challenges in data reduction: <i>Etalon cavity map</i>	51
4.1	One device, two configurations.	52
4.1.1	Collimated configuration	53
4.1.2	Telecentric configuration	55
4.1.2.1	Telecentric imperfect configuration	56
4.2	Sunspot observation simulation	57
4.2.1	Simulated data.	58
4.2.2	Observations simulation.	59
4.2.3	Flat-field correction.	61
4.2.4	Velocity maps	63
4.2.5	Magnetic field maps.	66
4.3	Fitting algorithm	67
4.3.1	Initial approximations and observations simulation.	68
4.3.2	Fitting algorithm	70
4.3.3	Test scenarios and results	74
4.3.3.1	Impact of the noise level	74
4.3.3.2	Impact of the object approximation	76
4.3.3.3	The crossover case	78
4.3.3.4	Real data	80
4.3.4	Algorithm applied at the Sunspot	83
4.4	Conclusions.	83
5	Scientific exploitation.	87
5.1	Persistent Homology in Solar Magnetograms.	87
5.1.1	Persistent Homology	87
5.1.2	Persistent Images	90
5.1.3	Data	92
5.1.3.1	Quiet Sun observations	92
5.1.3.2	Active regions observations	92
5.1.4	Analysis and results	93
5.1.4.1	Quiet Sun results	95
5.1.4.2	Active region results	98
5.1.4.3	'Interacting' Diagram	99
5.1.5	Conclusions	103

rals for page numbers

CHAPTER 1

INTRODUCTION

1.1 Background

In June 2009, the first Sunrise observatory (Barthol et al., 2011) was launched from Kiruna, Sweden, aboard a stratospheric balloon. Equipped with a 1-m aperture telescope, a multi-wavelength UV filter imager, and IMaX, a Fabry-Pérot-based magnetograph, Sunrise was the most complex payload carried by a solar stratospheric balloon to date. Aimed at studying the magnetic fields of the Sun and the dynamics of solar plasma convective flows, the mission was an outstanding success. It resulted in the publication of over a hundred peer-reviewed scientific articles in numerous high-impact journals, including *Astronomy and Astrophysics* (A&A), *The Astrophysical Journal* (APJ), and *Solar Physics*, among others.

Following the success of its inaugural flight, Sunrise embarked on a second journey (Solanki et al., 2017) on June 13, 2013. The primary objective of this subsequent flight was to investigate the active regions of the Sun, as it remained completely *quiet* throughout the entirety of the initial flight. Despite minimal alterations to the instrumentation aboard the observatory, the variance in solar activity during this second flight yielded fresh perspectives and valuable data, ultimately securing the mission success, despite encountering some technical challenges.

Given the success of the first two flights, a third iteration of the Sunrise mission was planned, featuring an updated design. For this third edition, the telescope was equipped with three post-focal instruments: SUSI, a UV spectrograph; SCIP, an infrared spectrograph; and TuMag, the evolution of the IMaX magnetograph. Sunrise III was initially scheduled to fly during the summer of 2020 but was postponed to 2022.

The third launch of Sunrise plays a crucial role in this dissertation. This thesis, initiated in 2020, was centered on the development of the data reduction pipeline for the TuMag instrument, which was entirely developed by the Spanish solar physics consortium. According to the original plan, the first half of the thesis was dedicated to the calibration of the instrument and the preparation of the data pipeline. This way, once the mission was launched, the second half of the thesis could focus on the correction and scientific analysis of the data produced during this third flight. However, this plan (and thus the scope of the thesis) encountered a setback on July 10, 2022, when the third flight of the Sunrise observatory had to be aborted just a few hours after the launch due to a mechanical failure during

the ascent phase.

The observatory was recovered days later after a brief stay in the Scandinavian Alps. Both the telescope and the instruments were found to be in good condition, allowing for the recovery of the observatory and providing hope for a second attempt. However, the process of retrieving the instruments, disassembling, calibrating, and verifying their condition before relaunching the mission is lengthy, and it was not until this year, 2024, that a second attempt became feasible.

In the absence of data produced by Sunrise to process, analyze, and exploit, the scientific work conducted within the framework of this thesis has been compelled to slightly shift its focus. Over these years, we have focused on delving deeper into image correction techniques for data obtained from Fabry-Pérot interferometers, such as TuMag and IMaX. As well as conducting several studies using data products from other instruments, such as the Polarimetric and Helioseismic Imager aboard Solar Orbiter (SO/PHI) and HMI.

It wasn't until the 10th of July of 2024 that Sunrise III got its second chance to fly, and this time, the opportunity was not wasted. After a very successful flight that lasted 6 days, the observatory landed in the northern region of Canada on the 16th of July. Figure 1.1 shows the trajectory our favourite solar observatory followed over these days. The recovery process started immediately after landing, and we were able to lay hands on the data for the first time on September 2024.

In the following chapters, we will present the work undertaken during the calibration and commissioning of TuMag, conducted in 2021, 2022, and 2024. Additionally, the research carried out between the first and second flights of Sunrise III, which has resulted in the publication of two articles as the main author — one published in APJ and the other in A&A — will also be detailed in this manuscript, as well as other studies that have not yet been published in any scientific journal.

1.2 Motivation of our work

In experimental sciences, there is a very strong relation between technological and scientific advances due to the simple fact that we cannot draw conclusions from what we cannot see. We believe it is important for experimental scientists, and more specifically, for observational astronomers, to know the limitations and capabilities and understand the functioning of the instruments we use.

This philosophy is one of the pillars of this thesis, which covers topics ranging from the design and calibration of scientific instruments to the exploitation of the data they produce. With this thesis, we aim to provide a broad, yet detailed, view of the various stages of a scientific mission, from its conception and objectives through its design and calibration, data reduction and preparation for scientific exploitation, and finally, the studies and conclusions derived from it.

In particular, we will detail this process within the framework of solar physics through the development of TuMag, the magnetograph aboard Sunrise III. We will present the scientific objectives of the mission and attempt to link the design concepts with the scientific

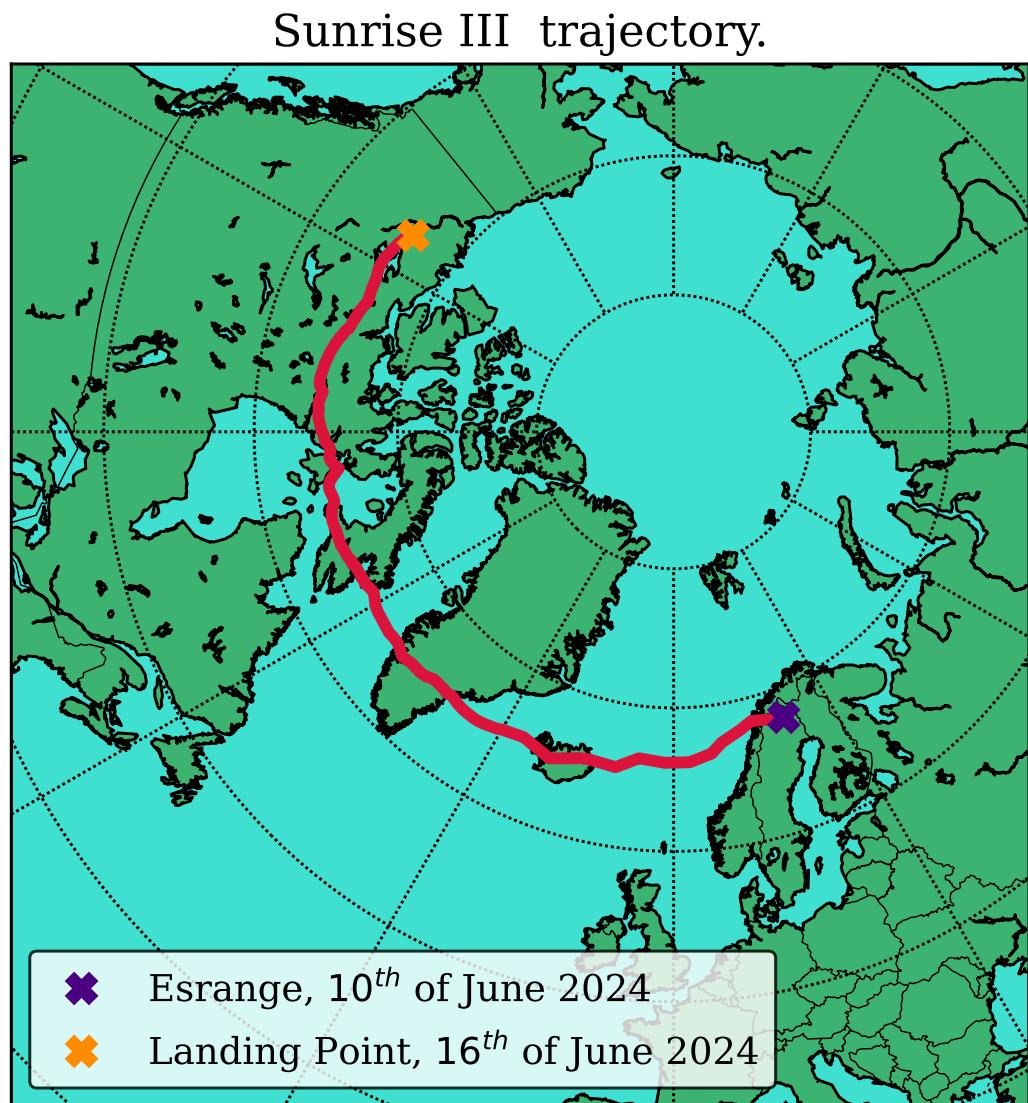


Figure 1.1 Sunrise III trajectory. Aqui poner tambien un foto de sunrise con el globo.

questions we aim to answer. We will address the challenges encountered in data correction due to the technical or instrumental limitations, a subject of ongoing debate within the community and of current relevance. And finally, we also aim to offer a brief dip into the scientific exploitation that can be carried out with the final data product.

With this thesis we aim to clarify the following points:

- ✿ Scientific objectives of TuMag.
- ✿ Instrumental ways of achieving the scientific purposes
- ✿ Open problems for data reduction. Flat-fields, etalon effects in the data.
- ✿ Offering an example of data exploitation with a study case. Persistent Homology.

1.3 Introduction

Intro general de la tesis.

Astronomy is one of the broadest fields of knowledge. It studies everything from the smallest astronomical objects, such as the small asteroids that inhabit our solar system, to the global structure and evolution of the universe, including the study of planetary systems, stars, black holes and the galaxies in which they are found. However, despite the diversity of disciplines—ranging from stellar astronomy, radio astronomy, and cosmology, to extragalactic astronomy, astrobiology, and solar physics—they all share a common tool for studying the cosmos: light. Since the very beginning of astronomy, the astronomer's work has been to learn how to modify and measure the properties of the photons that reach us in order to infer the characteristics of the observed object. Although recent advancements have provided astronomers with new lenses to *see* the cosmos, like gravitational waves (**REFERENCIA**) or neutrinos (**REFERENCIA**), among others, light remains as our main resource. Our understanding of the cosmos has always gone hand-in-hand with our ability to design and develop new (or more efficient) and clever ways to dissect the light, spanning from the first solar clocks, passing through Newton's first telescope to the modern-day spaceborne telescopes like the Hubble, James Webb or Solar Orbiter.

Solar physics is no different from other astronomical disciplines in this regard. Our main tool to *see* the Sun is through light. Contrary to what one may think, solar physicists are as photon-starved as any other astronomer. Even though our star is closer and (apparently) brighter than any other astronomical object, our requirements regarding resolution and sensitivity are so high that we are as dependent on extremely optimized instrumentation as any other discipline. Thus, the development of instrumentation employing state-of-the-art technology and techniques plays an important role in modern solar physics.

1.4 A brief introduction to spectropolarimeters.

As suggested by the name, spectropolarimeters are instruments designed to measure both the spectral and polarimetric properties of light. In other words, they assess the polarization state of light as a function of wavelength. Their use is widely extended in astrophysics, owing to the substantial amount of information that can be inferred about the light source from these properties.

There are two distinct types of spectropolarimeters, distinguished by their approach to spectroscopy: slit-based spectrographs and narrow-band tunable filtergraphs. The latter preserve spatial resolution by capturing two-dimensional images of the solar scene at the expense of sacrificing spectral resolution. Conversely, slit-based spectrographs provide excellent spectral resolution but have a limited spatial resolution.

Regardless of the method employed for spectroscopy, spectropolarimeters must be able to measure the polarization state of light, in other words, they must determine the Stokes parameters of the incident light. These four parameters, typically expressed as a pseudo-vector, $[I, Q, U, V]$, were introduced by Stokes (Stokes, 1851) as a formalism to fully describe the polarization state of light. The first parameter, I , denotes the total intensity, while Q and U provide information on the intensity of linearly polarized light at 0° and 90° , respectively. Lastly, V represents the intensity of circularly polarized light.

Outstanding polarimetric sensitivity and spectral resolution are rendered ineffective if the optical capabilities of the instrument are not up to par. The design of these instruments must achieve diffraction-limited imaging, coupled with a signal-to-noise ratio that ensures a polarimetric sensitivity for stokes Q , U and V of approximately 1000. Additionally, it must provide the best spatial resolution that the telescope allows, all while maintaining high spectral resolution and minimizing observation time. Consequently, instrument design requires a careful balance among these three properties: spectral, optical, and polarimetric capabilities.

In the following sections, we will examine each of these aspects in greater detail, with a particular emphasis on filtergraphs. Tunable filtergraphs play a significant role in this thesis, as the primary blocks of scientific work presented herein have been conducted for TuMag, a tunable magnetograph, as well as etalon-based instruments in general. Consequently, our description will be tailored to these types of instruments. It is important to note that much of the information provided will be generic and applicable also to slit-based spectrometry; however, certain specific behaviors will be unique to filtergraphs.

1.4.1 Imaging and optical quality.

Filtergraphs are, first and foremost, imagers. The high-resolution imaging that filtergraph instruments are capable of is one of the pivotal reasons for their extended use. The ability to capture a two-dimensional scene of the solar surface makes them ideal for studying solar plasma structures, which require resolutions close to 100 km on the solar photosphere. These instruments must be able to ensure an image quality and resolving power enough to measure these structures. For this reason, we will begin our description of the filtergraphs

with a brief explanation of image formation and image quality assessment.

Let us assume that the extended source we are observing has an intensity distribution in the image plane given by $O(\xi_0, \eta_0)$. Then, if we assume a linear optical system and incoherent illumination, the intensity distribution measured at a point ξ, η of the image plane is given by :

$$I_j(\xi, \eta; \lambda_s) = \iint O(\xi_0, \eta_0) S(\xi_0, \eta_0; \xi, \eta; \lambda_s) d\xi_0 d\eta_0, \quad (1.1)$$

where $S(\xi_0, \eta_0; \xi, \eta; \lambda_s)$ represents the imaging response of the instrument, also referred to as the Point Spread Function (PSF). The PSF describes the normalized intensity distribution in the image plane when observing a point source, which, due to diffraction and inherent imperfections in any imaging system, cannot be imaged as an ideal point.

The PSF is crucial in the assessment of image quality and resolving power of an instrument since it defines how fine detail will be imaged into the detector. One particularly relevant metric for image quality assessment that can be derived from the PSF is the optical transfer function (OTF), which is the Fourier transform of the PSF (Vargas Dominguez, 2009a).

$$OTF(\nu) = \hat{S}(\xi_0, \eta_0; \xi, \eta; \lambda_s), \quad (1.2)$$

where the operator $\hat{\cdot}$ is the Fourier transform, and ν represents the frequency vector in the Fourier domain.

The OTF describes how different spatial frequencies are transferred from the object to the image, thus characterizing the system's ability to resolve fine details. However, since imaging systems measure intensities, we are primarily concerned with how the intensity pattern of an object is transferred to the image. A key metric for quantifying this transfer is modulation, or contrast, which is defined as the ratio between the peaks and valleys of intensity at a given spatial frequency:

$$M_\nu = \frac{I_{max}^\nu - I_{min}^\nu}{I_{max}^\nu + I_{min}^\nu}. \quad (1.3)$$

The function that encodes the dependency of the modulation with spatial frequencies is called the modulation transfer function (MTF), and is strictly related to the OTF as the ratio of the modulation of the object MTF_{obj} and that of the image MTF_{im} can be computed from the magnitude of the OTF (Gaskill, 1978):

$$MTF = \frac{MTF_{im}(\nu)}{MTF_{obj}(\nu)} = |OTF(\nu)|. \quad (1.4)$$

From this definition, it is evident that a perfect optical system would have an $MTF = 1$ at all spatial frequencies, meaning that all details are perfectly transferred from the object to the image. However, real optical systems exhibit a decrease in MTF as spatial frequency increases. In practice, the resolution of an optical system can, and is often defined as the spatial frequency at which both the MTF and, consequently, the OTF reach zero (Tyson,

2000). This threshold frequency marks the limit beyond which the system can no longer resolve finer details.

Another key concept for assessing the imaging performance is the phase error or wavefront. The wavefront of an optical system is defined as the deviation in phase at any point within the image from that of an ideal spherical wavefront (Snyder, 1975). Such deviations arise from various optical imperfections within the imaging system, and their impact on image quality depends on the specific nature of the aberration. For instance, imperfections in mirror shape or lens configuration can result in spherical aberrations, leading to a broadening of the PSF and a subsequent reduction in resolution. Other common aberrations include astigmatism, where the focal point varies along different axes, producing distorted images, and comatic aberrations (coma), which can occur due to misalignment of optical elements and manifest as tail-like distortions in the images of point sources.

It is common to see requirements or assessment of the optical quality of optical instruments in terms of the root mean square (rms) of the variance of the wavefront, $\Delta\phi(\xi, \eta)$, usually referred to as the wavefront error (rms WFE) or simply WFE:

$$WFE = \sqrt{\frac{1}{A} \int_A (\Delta\phi(\xi, \eta))^2 d\xi d\eta}, \quad (1.5)$$

where A is the area of the aperture.

This value, essentially the standard deviation of the wavefront across the aperture, is closely tied to beam propagation quality. In fact, it can be demonstrated that the wavefront variance can be derived from the Strehl ratio, or viceversa. The Strehl ratio is defined as the ratio of the peak intensity of a point source in an aberrated system to that of an ideal system operating at the diffraction limit. It is one of the most widely used metrics for assessing the optical quality of a system, ranging from 1, for a perfect, unaberrated system, to 0. For small aberrations, the Strehl ratio (SR) and WFE are related by the following expression (Snyder, 1975):

$$SR \simeq \exp \left[- \left(\frac{2\pi WFE}{\lambda} \right)^2 \right]. \quad (1.6)$$

Although the Strehl ratio and rms WFE provide a concise measure of the optical quality of a system, the WFE contains additional information regarding imaging performance. Rather than relying solely on a single averaged value (such as the standard deviation), the wavefront can be represented as a two-dimensional map projected onto a plane normal to the light path, typically the image plane. To carry out such a representation analytically, it is essential to select an appropriate mathematical framework. Given the widespread use of circular apertures in telescopes, mirrors, lenses, and other optical components, it is advantageous to approach the problem using polar coordinates, ρ and θ , and in particular, to employ an orthonormal basis for the interpretability of the results. Among the multiple (infinite) sets of polynomials that fulfill these requirements, the Zernike polynomials (Zernike, 1934) offer some distinct advantages. The Zernike polynomials are a sequence of polynomials that compose an orthonormal basis over a unit circle. Given an arbitrary wavefront,

$(W(\rho, \theta))$, the expansion in terms of the Zernike polynomials can be expressed as:

$$W(\rho, \theta) = \sum_{n,m} C_n^m Z_n^m(\rho, \theta), \quad (1.7)$$

where Z_n^m are the Zernike polynomials, C_n^m are the amplitudes of the coefficients in the expansion and n and m are the radial order and angular frequency, respectively. The Zernike polynomials can be obtained from:

$$\left. \begin{array}{l} Z_n^m(\rho, \theta) = R_n^m(\rho) \cos(m\theta), \text{ for } m \geq 0, \\ Z_n^{-m}(\rho, \theta) = R_n^m(\rho) \sin(m\theta), \text{ for } m < 0, \end{array} \right\} \quad (1.8)$$

where $R_n^m(\rho)$ are the radial functions given by:

$$R_n^m(\rho) = \sum_{l=0}^{(n-m)/2} \frac{(-1)^l (n-l)!}{l! [\frac{1}{2}(n+m)-l]! [\frac{1}{2}(n-m)l-]!} \rho^{n-2l}. \quad (1.9)$$

This representation of the wavefront is particularly valuable because each mode, defined by a specific pair of n and m values, corresponds to a distinct aberration in the wavefront, with the associated coefficient representing the contribution to the rms WFE for that specific aberration. Furthermore, the orthogonality of the Zernike basis ensures that adding additional terms to the expansion does not influence the values of previously calculated coefficients. In other words, the Zernike polynomial expansion enables the wavefront to be expressed as the sum of individual aberrations, providing a clear decomposition of the wavefront errors.

Figure 1.2 presents an example of a simulated wavefront, including a two-dimensional cross-section and the individual Zernike components of the simulation. The simulation incorporates only the first ten Zernike polynomials, corresponding to polynomials with $n \leq 3$, which account for aberrations such as defocus, astigmatism, coma, and trefoil, among others. For a comprehensive overview of the Zernike expansion in wavefront characterization, we direct the reader to Lakshminarayanan & Fleck (2011).

Although properties such as the PSF and WFE provide us with a lot of information to evaluate the instrument's performance, they are not known a priori. Instruments need to be equipped with means to measure these quantities in order to assess the performance and correct for any defects. The strategy to follow depends on the specific characteristics of the instrument. For instance, ground-based instruments can be equipped with adaptive optics systems, where the wavefront is directly measured and corrected in real time (REFERENCES). However, this approach is not feasible for all instruments due to the technical demands required to implement such corrections. In these cases, alternative techniques, such as the phase diversity method, are necessary.

The phase diversity algorithm (Childlaw et al., 1979) is a method to infer the aberrations present in an optical system by obtaining, at least, two simultaneous, or quasi-simultaneous, images of the same object where an additional and known aberration is introduced to one of the images.

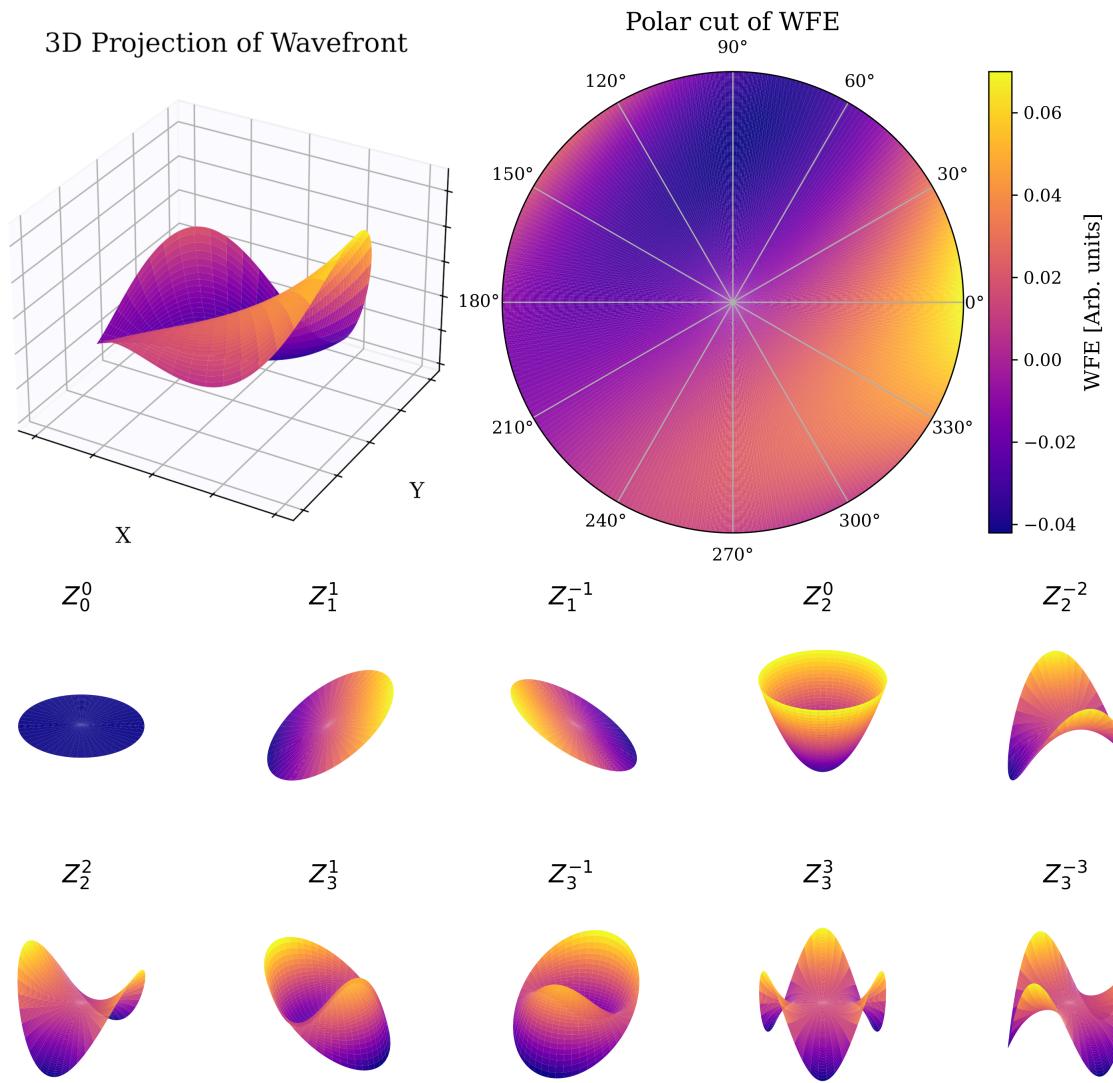


Figure 1.2 Simulation of a wavefront employing all Zernikes with $n \leq 3$. The top left panel shows the 3-dimensional representation of the wavefront and the top right panel shows a cut in a plane normal to the direction of travel. The bottom two rows show the shape of the individual Zernike polynomials included in the simulation.

The algorithm works by minimizing a cost function that depends on the OTF of the system which can be parametrized by the Zernike expansion (Paxman et al., 1992):

$$\mathcal{L}(C) = \sum_k \frac{|\hat{I}_1(\nu)\hat{S}_2(\nu, C) - \hat{I}_2(u)\hat{S}_1(\nu, C)|^2}{|\hat{S}_1(\nu, C)|^2 + |\hat{S}_2(\nu, C)|^2}, \quad (1.10)$$

where k represents the pairs of aberrated (subindex 1) and unaberrated (subindex 2) images, I stands for the intensity distributions, and S for the system's OTF expressed in terms of a Zernike expansion with coefficients C.

By finding the coefficients of the zernike expansion that minimize \mathcal{L} , we are able to characterize the wavefront and identify the aberrations present in the optical system. Thus, determining the OTF, and consequently the PSF.

Our interest in determining the wavefront and OTF is not only to evaluate the instrument's performance but also to enable image restoration by mitigating the aberrations introduced during the imaging process. The procedure for removing the effects of the aberrations consists on removing the influence of the PSF on the final intensity distribution. In other words, the goal is to *deconvolve* the PSF from the image.

Coming back to equation (1.1), we can simplify the integrals to a convolution operator assuming an spatial invariance of the PSF. In that case, the observed intensity can be expressed by:

$$I(\xi, \eta) = O(\xi, \eta) * S(\xi, \eta) + N(\xi, \eta) \quad (1.11)$$

where we added a term accounting for the noise present in real meaurmeents $N(\xi, \eta)$.

The treatment of the problem is easier in the Fourier domain, where the convolution operator becomes a product of the fourier transforms of the corresponding functions. Therefore, in the Fourier domain, eq. 1.11 becomes:

$$\hat{I}(\nu) = \hat{O}(\nu)\hat{S}(\nu) + \hat{N}(\nu), \quad (1.12)$$

Since neither $\hat{O}(\nu)$ nor $\hat{N}(\nu)$ are known, it is not possible to analytically derive the object, even if the system's response, is known. Therefore, statistical approaches must be employed to deconvolve PSF.

One such approach is the Wiener-Helstrom filter (Helstrom, 1967), which proposes that the estimated object, \tilde{O} , can be computed as:

$$\tilde{O}(\nu) = \frac{\hat{S}^*(\nu)\hat{I}(\nu)}{|\hat{S}(\nu)|^2 + P_N(\nu)/P_O(\nu)}, \quad (1.13)$$

where the term $P_N(\nu)/P_O(\nu)$ represents the ratio between the power spectral densities of the noise and the object. Although this factor is unknown, it can be estimated based on the expected S/N in the data.

1.4.2 Spectroscopy

Among the tunable filtergraphs, Fabry-Pérot Interferometers (FPIs), also known as etalons (used interchangeably), represent one of the most prevalent forms of narrow-band tunable spectrographs. Composed by a resonant optical cavity formed by two distinct optical media, these devices allow only the passage of light with wavelengths corresponding to constructive interference within the cavity.

The transmission profile of an etalon, being produced by an interference phenomenon, is characterized by a series of narrow and periodic transmission peaks. The wavelengths at which these resonance peaks are located, their width, and their separation are determined solely by the physical properties of the etalon. In fact, it is not difficult to demonstrate (Bailén et al., 2019) that a resonant cavity produces a periodic transmission profile, with maxima occurring at a wavelength λ such that:

$$\lambda = \frac{2nd \cos \theta}{m}, \quad (1.14)$$

where n is the refractive index of the medium inside the cavity, d is the distance between the mirrors, θ is the angle of incidence of the incoming light ray and m is the interferential order ($m \in \mathbb{Z}$).

With Eq. (1.14) in mind, it is clear that an etalon allows for tuning the wavelengths of the transmission peaks by either changing the distance between the mirrors or by altering the refractive index. Although changing the angle of incidence also results in a wavelength shift, it introduces other issues, such as ghost images or profile broadening in telecentric configurations, among other effects. Consequently, the angle is not used for wavelength tuning.

To isolate a single wavelength, or a narrow band surrounding it, it is necessary to select one transmission peak, often referred to as the main order. This is typically achieved by using a pre-filter with a small bandwidth that only allows light with wavelengths near the desired measurement region to pass through. This prefilter prevents light from other interference orders than the main one, known as secondary orders, to reach the instrument.

Figure 1.3 shows a simulation of the spectral behavior of this optical setup. The order-sorting pre-filter is shown with a shaded purple area and the unaltered transmission profile of the etalon is shown in dashed lines for different values of the refractive index. In solid lines, the resulting transmission profile is shown, that is, the transmission allowed through both the pre-filter and etalon at the same time.

1.4.3 Polarimetry

A polarimeter must determine the full Stokes components of the incoming light; however, these properties cannot be directly measured, as only the intensity of light is observed, not its intrinsic characteristics. Thus, polarimeters derive the Stokes parameters, rather than measure them. In order to do so, a series of multiple, simultaneous or quasi-simultaneous observations are taken, in which the polarization state of the incoming light is systemat-

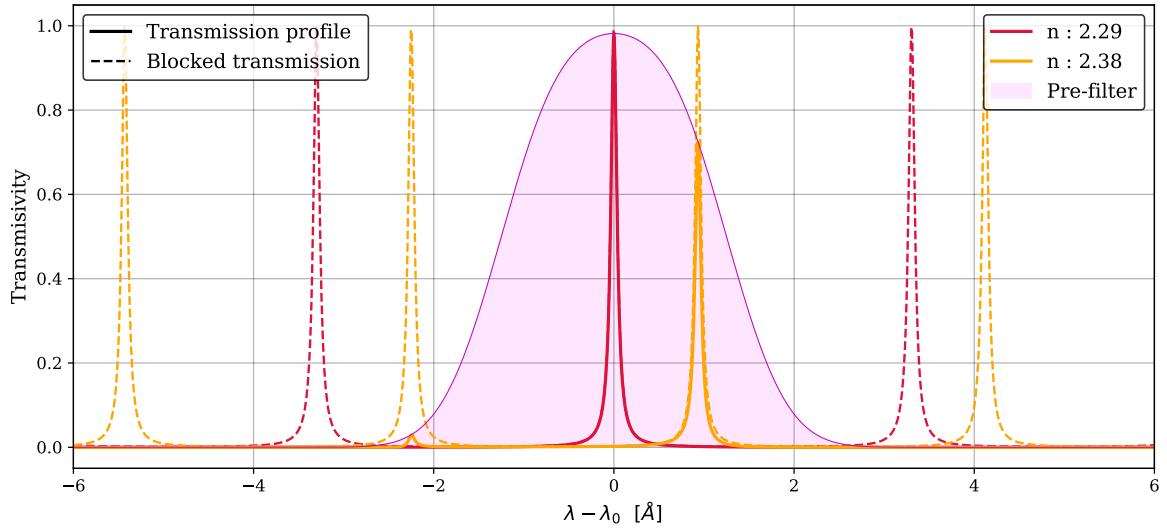


Figure 1.3 Transmission profiles of the same etalon with varying refractive indices (n). The dashed lines represent the original transmission profile, while the solid lines indicate the portion of the transmission profile that passes through the order-sorting pre-filter (shaded purple area).

ically altered. These different measurements, commonly referred to as modulations, are generated by inducing a known alteration in the polarization. The Stokes parameters are then reconstructed by combining the information from all measurements through a process known as demodulation.

In order to understand how polarimeters derive the stokes components we need to briefly model how the different modulations are generated. Mathematically, the effect on polarization of a linear and finite system can be treated as a combination of linear transformations on the Stokes vector and, therefore, can be represented by a matrix in \mathbb{R}^4 , known as the *Mueller Matrix*. Let \mathbf{M} be the matrix that describes these transformations, then the polarization state that reaches the detector follows:

$$\mathbf{I}_{out} = \mathbf{M}\mathbf{I}_{in}, \quad (1.15)$$

where \mathbf{I}_{in} and \mathbf{I}_{out} are the Stokes vectors of the light that reaches the instrument, and the detector, respectively. However, since we only measure intensities, the actual quantity measured by our CCD is:

$$I_{obs} = m_{00}I_{in} + m_{01}Q_{in} + m_{02}U_{in} + m_{03}V_{in} , \quad (1.16)$$

where m_{0i} is the i-th element of the first row of the Mueller Matrix. This means that the intensity we measure is a linear combination of the different polarization states of the incoming light. To determine the values of the individual parameters I_{in} , Q_{in} , U_{in} , and V_{in} , further independent measurements are necessary, which can be achieved by modifying the Mueller matrix. In particular, it is easy to see that four independent measurements are re-

quired in order to construct a system of equations that allows us to determine the full Stokes vector. This process is known as modulation, and the four independent measurements are the different modulations.

If we denote each of the modulations by I_j with $j \in \{1, 2, 3, 4\}$, we can construct the following system of equations:

$$\begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{pmatrix} = \underbrace{\begin{pmatrix} m_{01}^1 & m_{02}^1 & m_{03}^1 & m_{04}^1 \\ m_{01}^2 & m_{02}^2 & m_{03}^2 & m_{04}^2 \\ m_{01}^3 & m_{02}^3 & m_{03}^3 & m_{04}^3 \\ m_{01}^4 & m_{02}^4 & m_{03}^4 & m_{04}^4 \end{pmatrix}}_{\mathbf{O}} \begin{pmatrix} I_{in} \\ U_{in} \\ Q_{in} \\ V_{in} \end{pmatrix} \quad (1.17)$$

where the superindex in m_{oi}^j denotes the values of the Mueller Matrix for each modulation. Through straightforward algebra, it is easy to see that the stokes vector of the incoming light can be determined by $\mathbf{I}_{in} = \mathbf{D}\mathbf{I}_{obs}$, where \mathbf{D} is the demodulation matrix, the inverse of the modulation matrix, \mathbf{O} , and \mathbf{I}_{obs} is the vector containing the 4 measured modulations.

Carefully determining \mathbf{O} , and consequently \mathbf{D} , during the instrument calibration process is crucial, as the accuracy of the determination of the Stokes components depends entirely upon it. It can be proven (del Toro Iniesta & Collados, 2000) that the optimum modulation scheme—the values of \mathbf{D} that enable the Stokes vector to be computed with minimal uncertainty—satisfies the conditions:

$$\varepsilon_1 \leq 1, \text{ and } \sum_{i=2}^4 \varepsilon_i^2 \leq 1, \quad (1.18)$$

where the polarimetric efficiencies for each stokes parameter ($i = 1, 2, 3, 4$), ε_i , are defined as:

$$\varepsilon_i = \left(N_p \sum_{j=1}^{N_p} D_{i,j}^2 \right)^{-1/2}, \quad (1.19)$$

where N_p is the number of independend modulations.

When designing the modulation scheme for a given instrument, it is essential to satisfy the efficiency conditions given in Equation (1.18) to ensure optimal polarimetric accuracy for all Stokes components. Furthermore, for equal sensitivities in the measurements of Stokes parameters Q, U, and V, the corresponding efficiencies should all be equal, with a value of $1/\sqrt{3}$. This is a very important result beacause polarimetric efficiencies are directly related to the smallest measurable polarimetric signals, the polarimetric sensitivity—essentially the inverse of the signal-to-noise ratio (SNR). This relation can be expressed as (del Toro Iniesta & Collados, 2000):

$$\left(\frac{S}{N} \right)_i = \frac{\varepsilon_i}{\varepsilon_1} \left(\frac{S}{N} \right)_1, \quad i = 2, 3, 4. \quad (1.20)$$

From equations (1.20) and (1.18) it is clear that the sensitivities for computing Stokes Q, U, and V will always be lower than that of Stokes I, as their corresponding efficiencies are

smaller. To achieve an SNR of 10^3 in Stokes measurements, which is the sensitivity required to detect weak polarization signals, an SNR of at least $(S/N)_0 \gtrsim 1700$ is necessary in the measurement of Stokes I for a quasi-optimal modulation scheme.

Spectropolarimeters ultimately combine measurements in polarization, spectral, and spatial (image) domains. Consequently, the final observed intensity depends on all three properties simultaneously. By integrating the spectral behavior of the etalon and pre-filter with the polarimetric measurements, and taking into account the spatial dependence of these measurements, we can revisit equation (1.1) and rewrite it for FPI-based spectropolarimeters. In that case, the observed intensity for a modulation j at any point of the focal plane η, ξ when the etalon is tuned at a wavelength λ_s is determined by:

$$I_j(\xi, \eta; \lambda_s) = g(\xi, \eta) \int_0^\infty T(\lambda) \iint O_j(\xi_0, \eta_0; \lambda) S(\xi_0, \eta_0; \xi, \eta; \lambda - \lambda_s) d\xi_0 d\eta_0 d\lambda, \quad (1.21)$$

where $T(\lambda)$ accounts for the presence of the order-sorting pre-filter, $S(\xi_0, \eta_0; \xi, \eta; \lambda - \lambda_s)$ accounts for the imaging response of the instrument when the etalon is tuned at the wavelength λ_s , $g(\xi, \eta)$ represents a spatial gain factor that accounts for any wavelength independent pixel-to-pixel intensity fluctuations occurring in the focal plane, and $O_j(\xi_0, \eta_0; \lambda)$ is the intensity distribution of the incoming light for a modulation j and is given by:

$$O_j(\xi_0, \eta_0; \lambda) = m_{00}^j I_{in}(\xi_0, \eta_0; \lambda) + m_{01}^j Q_{in}(\xi_0, \eta_0; \lambda) + m_{02}^j U_{in}(\xi_0, \eta_0; \lambda) + m_{03}^j V_{in}(\xi_0, \eta_0; \lambda) \quad (1.22)$$

1.4.4 What do spectrpolarimeters tell us about the Sun?

Spectropolarimeters are often referred to as magnetographs (e.g., TuMag), suggesting they measure magnetic fields directly. However, this is not entirely accurate. In astrophysics, the physical properties of the light source are inferred by correlating them with the observed properties of the light, rather than measuring them directly. By evaluating the polarization of sunlight at different wavelengths, spectropolarimeters enable us to infer the magnetic field and estimate plasma velocities on the solar surface.

The simplest calculation we can carry out that provides us with physical quantities of the Sun is that of the line-of-sight (LOS) velocities. Given the spectral shift of a specific absorption or emission spectral line, $\Delta\lambda$, with respect to its rest position, λ_0 , the LOS velocities can be computed with the Doppler formula:

$$v_{\text{LOS}} = \frac{\Delta\lambda}{\lambda_0} c, \quad (1.23)$$

where c stands for the speed of light in vacuum.

The polarization properties of light come into play when determining the magnetic fields. Due to Zeeman and Hanle effects, the polarity and spetcroscopy of spectral lines can be altered when formed in the presence of magnetic fields. Due to the Zeeman effect, the spectral lines widen or split into different polarized components when a strong magnetic

field is present (del Toro Iniesta, 2003), such as in the surroundings of sunspots and active regions. In the other hand, the Hanle effect is sensitive to weaker fields, and can be used to study, for example, the magnetic structure of solar prominences or turbulent fields in the solar photosphere (Bianda et al., 1998) where the fields are not strong enough to leave an imprint through the Zeeman effect.

One simple strategy to employ polarization and spectral data to derive the magnetic fields is through the center-of-gravity method. According to Uitenbroek (2003), the LOS strength of the magnetic field can be obtained through:

$$B_{\text{LOS}} = \frac{\lambda_+ - \lambda_-}{2} \frac{4\pi mc}{eg_L \lambda^2} , \quad (1.24)$$

where m and e are the electron mass and charge respectively, g_L stands for the Landé factor and λ_+ and λ_- are the centroids of the right and left circularly polarized line components, respectively, and are computed by:

$$\lambda_{\pm} = \frac{\int \lambda [I_{\text{cont}} - (I \pm V)] d\lambda}{\int [I_{\text{cont}} - (I \pm V)] d\lambda} , \quad (1.25)$$

where the subindex "cont" stands for the wavelength at the continuum.

The vector magnetic field (*i.e.*, strength, azimuth and inclination), and not only the LOS strength can also be derived. However, the derivation of these quantities has to be achieved through inversions of the radiative transfer equation (RTE). The applicability of the different methods to carry out this inversion is an extensive topic as there are some assumptions that can be applied in some cases but not in others, such as the weak-field or Milne-Eddington approximations, among others. For an extended discussion of this topic, we refer the interested reader to del Toro Iniesta & Ruiz Cobo (2016).

CHAPTER 2

SUNRISE III AND TUMAG: DESIGN AND CALIBRATION.

Observing the Sun with the highest possible quality and resolution is crucial for deepening our understanding of the physical processes governing its behavior. This necessity drives the continuous development of new state-of-the-art observatories and advanced instrumentation. Each new instrument or telescope builds upon the technical achievements of its predecessors, integrating past knowledge while introducing innovations that push the boundaries of solar observation.

An example of such advancements is the third edition of the Sunrise observatory, which marks the culmination of a collaborative effort among several international institutions. Spearheaded by the Max Planck Institute for Solar System Research (MPS) in Göttingen, Germany, the international consortium is also composed by the Spanish Solar Space Consortium (S³PC), the National Astronomical Observatory of Japan (NAOJ), the Kiepenheuer Institute for Solar Physics (KIS) in Freiburg, Germany, and the Johns Hopkins University Applied Physics Laboratory (APL) in the United States.

Following an initial unsuccessful flight in 2022, which was aborted six hours after launch, Sunrise III was granted a second opportunity in the summer of 2024. On July 10th, 2024, at 04:22:40 UTC, the observatory was successfully launched by the Columbia Scientific Balloon facility (CSBF-NASA) from Esrange, a scientific facility operated by the Swedish Space Corporation in Kiruna, Sweden. After reaching a stable altitude of 37.5 km, the commissioning phase began, marking the official start of the observation campaign. Observations commenced shortly thereafter and continued until the campaign concluded on July 16th at 18:20:54 UTC, when the flight was terminated.

Among the payload instruments aboard Sunrise III, TuMag holds particular significance for this thesis. TuMag, an imaging magnetograph developed by the Spanish Solar Space Consortium (S³PC) under the leadership of the Instituto de Astrofísica de Andalucía (IAA-CSIC) in Granada. The S³PC also includes the Instituto Nacional de Técnica Aeroespacial (INTA), the Instituto de la Riva (IDR-UPM) at the Universidad Politécnica de Madrid, the Universitat de València (UV), and the Instituto de Astrofísica de Canarias (IAC). TuMag is central to this thesis, as the core of the work focuses on its calibration, operations, and data reduction processes.

In this chapter, we present an overview of the Sunrise III mission, with a particular focus on the Tunable Magnetograph (TuMag). We will first outline the scientific motivations behind its development and the design choices. This will be followed by a detailed discussion of the technical specifications of both the mission and TuMag.

2.1 Sunrise III

Equipped with a telescope with a 1m aperture, two slit-based spectropolarimeters and an imaging magnetograph, the Sunrise III observatory is the most complex solar telescope to ever leave the ground. The coordination of three different scientific instruments allows Sunrise to simultaneously perform narrow-band polarimetric imaging in the visible while carrying out spectropolarimetry in the near-UV and near-IR, from the advantageous point of observation of ~ 36 km of altitude without the annoying interference of the atmosphere's turbulence.

The three instruments aboard Sunrise III have been carefully designed to complement each other and address the scientific purposes of the mission. The Tunable Magnetograph (TuMag, REF), developed by the S³PC, carries out high-spatial-resolution imaging spectropolarimetry in the visible range of light. Able to tune to three different spectral lines, namely the highly Zeeman-sensitive iron lines at 525.02 and 525.06 nm, and the Mg I b₂ line, TuMag can probe the photosphere and low chromosphere quasi-simultaneously.

The absence of atmosphere allows the Sunrise Spectropolarimeter and Imager (SUSI, Ref), developed by MPS and NAOJ, to observe in the near-UV, performing imaging and spectropolarimetry in the range of 309-417 nm. The high polarimetric sensitivity and large number of spectral lines present in this range, many of which are sensible to the Hanle effect, allows SUSI to sample many heights in the solar atmosphere at the same time while measuring the weak magnetic fields.

In the same way that the atmosphere complicates the observations of the UV range from the ground, observations of many lines in the near-IR are also unfeasible with ground-based instruments due to the telluric contamination of the atmosphere. The Sunrise Chromospheric Infrared spectro-Polarimeter (SCIP, REF), co-developed by NAOJ and IAA-CSIC, takes advantage of the absence of atmosphere and observes two of the Ca II triplets lines. Spectropolarimetry measurements of these lines provides information of the 3-D structure of the chromosphere and its magnetic fields, derived thanks to the high zeeman sensitivity of the selected lines. Furthermore, the large number of available photons at these wavelengths ensures a high S/N and polarimetric sensitivities.

The ability to probe simultaneously the near-IR, the visible and the near-UV, performing high-resolution polarimetric imaging and spectroscopy makes Sunrise III an unique observatory, capable of studying the connection and interaction of the small-scale phenomena occurring at different layers of the solar's atmosphere with unprecedented detail and completeness.

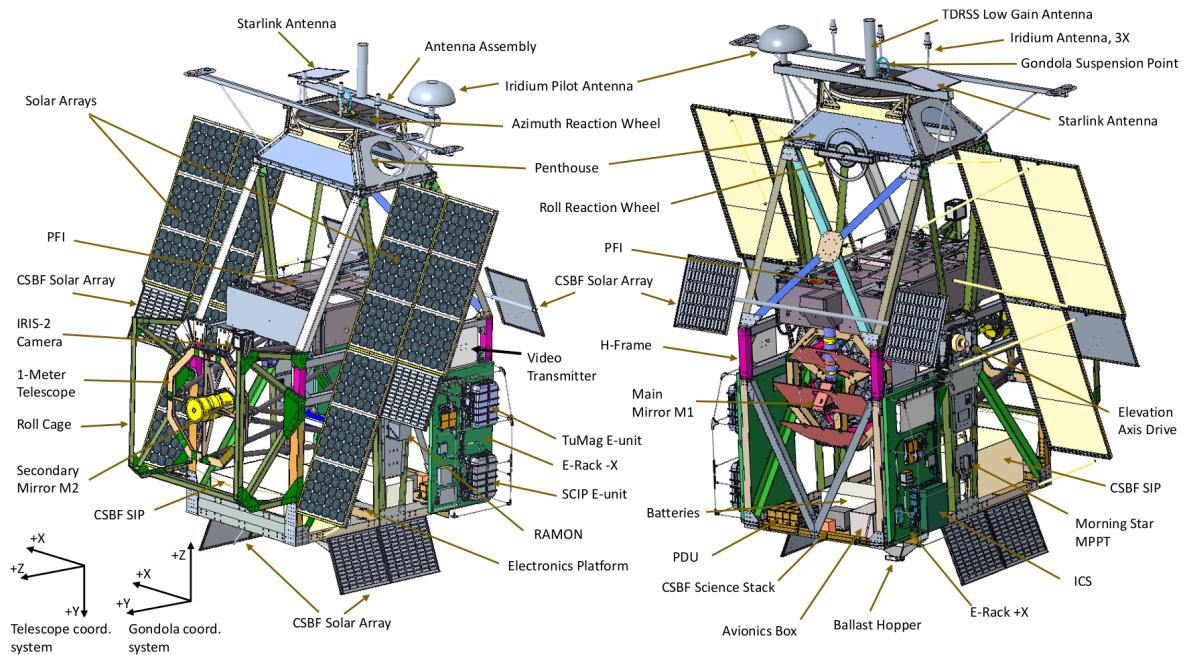


Figure 2.1 Drawing design of the Sunrise III observatory. Image reproduced from XXXXX with permission. **No he pedido permiso para esta aún, que es del paper de Sunrise III que mandó Andi, pero creo que quedaría bien algo enseñando el observatorio. Al no estar publicado el paper de Sunrise no se muy bien cual coger.**

2.1.1 Observatory's design

While the scientific instruments are central to the research performed by Sunrise, several additional subsystems play a crucial role, each contributing to the overall success of the mission.

The structural framework housing all components of the observatory, as well as the interface connecting the observatory to the flight apparatus, is provided by the gondola. This gondola, engineered by APL, is not only tasked with safeguarding the instruments and telescope during ascent and landing, but also with ensuring the stability of the pointing system. Given that the observatory is suspended from a balloon, it is subject to wind-induced motion and pendulum-like oscillations, which threaten the stability required for prolonged observations. The gondola's pointing control system (PCS) must actively counteract these disturbances in real time by making precise adjustments to the telescope's elevation and azimuth. In conjunction with the Correlating Wavefront Sensor (CWS), developed by KIS, which is responsible for image stabilization and autofocus, the PCS was required to achieve a pointing accuracy better than 0.005" root mean square (rms) over extended periods to facilitate long-duration observations.

The Sunrise III telescope is a Gregory-type reflector with a 1-meter aperture, featuring a 234 mm central obscuration and an effective focal length of 24.2 meters. This configuration provides a field of view (FoV) of 3.4', corresponding to approximately 150 Mm on the solar

surface. The telescope directs light to the post-focus instrumentation platform (PFI), located above the telescope. The PFI houses the three scientific instruments and the CWS, and is responsible for distributing light among these four instruments. This distribution, performed by the Image Stabilization and Light Distribution Unit (ISLiD), must efficiently separate the different wavelengths in a photon-efficient manner to provide the highest number of available photons to each instrument.

While all the subsystems discussed thus far directly influence the optical performance, it is equally important to recognize the crucial role played by other subsystems, such as the electronics and software control. In particular, the Instrument Control System (ICS) is responsible for the management of the observatory, gathering housekeeping and issuing commands to the electronic units of each instrument. As will be elaborated in the following chapter, Sunrise III observations were designed to operate in a semi-autonomous manner through the use of pre-programmed timelines. This approach requires that all electronic systems function in synchrony, with minimal human intervention.

2.1.2 Science with Sunrise.

The absence of Earth's atmosphere opens the window for IR and UV observations and offers a level of image stability that cannot be achieved in ground-based observatories due to atmospheric seeing. However, these advantages are also present in spaceborne missions, such as the Spectro-Polarimeter and Narrowband Filter Imager aboard the Solar Optical Telescope of the Hinode mission (Kosugi et al. 2007, Tsuneta et al. 2008), or the Polarimetric and Helioseismic Imager aboard Solar Orbiter (SO-PHI) (Solanki et al. 2020, Müller et al. 2020), among many others. Nonetheless, spaceborne missions have strong restrictions regarding payload, mass and data rate.

The absence of these restrictions in balloon-borne observatories allows for a more complex and versatile instrumentation than the one present in space missions. The combination of these two factors, namely the absence of atmosphere and the complex and advanced instrumentation they can carry, places observatories such as Sunrise in an unique position, and provides them with unique perspectives on solar phenomena.

Many aspects of the physical processes driving our Sun remain unresolved. The mechanisms underlying various solar phenomena are still the subject of debate, ranging from the origin and removal of magnetic flux in the solar photosphere to the processes responsible for heating the chromosphere and corona, as well as the small-scale dynamics of solar plasma. The three instruments aboard Sunrise work in consonance to provide novel insights into these phenomena, and aim at helping the community solve some of the open questions of solar physics.

The magnetic field, present across multiple scales and heights, is the principal driver of solar activity. Understanding the magnetic field is essential for comprehending the processes that govern solar phenomena, energy distribution, and plasma dynamics. Numerous works direct their efforts to the study of the structures and evolution of magnetic fields. For instance, several works study the emergence of magnetic flux in the photosphere, such as, Danilovic et al. (2010) and Guglielmino et al. (2012), where they utilized Sunrise I IMaX

data to examine small-scale flux emergence events occurring within solar granules. Likewise, the processes responsible for magnetic flux removal are not fully known. Several studies, such as Zwaan (1987) and Guglielmino et al. (2012), have proposed mechanisms for flux removal in the photosphere; however, no model is favoured over the other by current observations.

A thorough 3-dimensional analysis of the magnetic fields is essential to study these events. The combination of spectropolarimeters and vector magnetographs aboard Sunrise, which are capable of measuring strong magnetic fields through the Zeeman effect, and detecting weaker, more turbulent (Bellot Rubio & Orozco Suárez, 2019) magnetic fields using the Hanle effect - particularly sensible in the UV - can provide a new and complete perspective of these events.

One of the scientific targets of Sunrise is the study of the upper atmosphere, whose dynamics and heating mechanisms are not yet completely understood. In fact, the transfer of energy from the lower layers to the chromosphere and corona is one of the open problems in stellar astrophysics. Several studies propose mechanisms in which the magnetic field plays a central role in this energy transfer. Some works suggest upward currents generated by the slow motion of plasma in the photosphere as a driving mechanism (Parker (1983), Pontin & Hornig (2020)), while others highlight heating processes induced by jets (Shibata et al., 2007) or magnetic vortex phenomena, such as twisted magnetic fields known as solar tornadoes (Wedemeyer-Böhm et al., 2012).

Although some observational signatures of these processes have been detected, the detailed characterization of these events requires higher spatial and temporal resolutions than those currently available. The high-cadence UV observations, where several spectral lines sensitive to the weaker magnetic fields of the chromosphere are present, combined with magnetic field maps of the photosphere and lower chromosphere provided by TuMag, and complementary observations of the Ca II infrared lines, provide Sunrise with the necessary tools to investigate these phenomena with unprecedented detail.

Sunrise also aims to provide novel insights into small-scale plasma dynamics. The Sun is highly dynamic, with structures evolving on timescales of minutes. Several studies propose that the magnetism in the quiet Sun is driven by the turbulent small-scale dynamics of the plasma (Petrovay & Szakaly (1993), Hotta et al. (2015), Rempel et al. (2023), among others). However, investigating these processes requires high spatial and temporal resolutions, which are often unattainable in ground-based observations. Similarly, other approaches to plasma dynamics, such as helioseismology (Gizon et al., 2010), also demand such high-resolution data. To address these challenges, Sunrise III conducted extended, highly stable due to the absence of atmospheric seeing, and uninterrupted observations lasting up to 6 hours, with the highest temporal resolution permitted by the (S/N) requirements.

In addition to these objectives, Sunrise will explore new and exciting areas, including the measurement of the polarized solar spectrum in the UV. This spectral region remains largely unexplored due to the technical challenges associated with its observation. Atmospheric absorption makes it impossible to observe this band from ground-based observatories, and it has yet to be measured by any space mission. SUSI is the first UV spectropolarimeter to acquire high-resolution data in this wavelength range.

Requirements	Value
Field of view	63'' x 63''
RMS wavefront error	$W \sim \lambda/14$
Spatial sampling	3 × 3 pixels
Plate scale	0.0378'' / pixel
Polarimetric efficiencies	$\epsilon_{1,2,3} \lesssim \frac{1}{\sqrt{3}}$
SNR ratio	$(\frac{S}{N})_0 \gtrapprox 1700$
Spectral resolution	< 9 pm
Spectral lines	Fe I 5250.2 Å, Fe I 5250.6 Å and Mg I b_2 5172.7 Å.
Time for a two-line observation	< 90 s

Table 2.1 Tumag scientific requirements.

2.2 The Tunable magnetograph: TuMag

TuMag is an FPI-based tunable imaging spectropolarimeter, capable of measuring the full Stokes vector across various spectral lines. This tunability allows TuMag to probe the magnetic field in both the photosphere and lower chromosphere, with high resolving power, thanks to its near-diffraction-limited imaging capabilities. The design of TuMag is inherited from IMaX, the imaging spectropolarimeter that flew aboard previous Sunrise missions. However, TuMag incorporates several advancements over its predecessor, including the addition of filter wheels for tunability between spectral modes and calibration systems, along with newly designed cameras and modulation packages.

In this section, we present an overview of the design of TuMag and its performance. It is important to note that this discussion will primarily focus on the instrument's optical design and performance—specifically its polarimetric, spectral, and imaging properties. The thermal design and control software, while crucial to the instrument's functionality, will not be covered here to avoid excessive length. Nonetheless, it is essential to acknowledge that these aspects are integral to TuMag's overall performance and represent a substantial portion of the effort contributing to its success.

2.2.1 TuMag's design and light path.

As a polarimeter, TuMag must be able to measure the full Stokes vector of the incoming light. To achieve this, it must generate four distinct modulation states and measure them in an almost simultaneously manner. As a spectrometer, TuMag must possess the capability to select specific wavelengths along multiple spectral lines. This selection process involves, first filtering the light with a pre-filter, which selects a "broad" spectral range, followed by an etalon that further narrows the bandpass within the selected range. Throughout this procedure, stringent requirements regarding polarimetric sensitivity (efficiency), spectral resolution, and imaging quality must be maintained. A summary of these requirements is presented in Table 2.1.

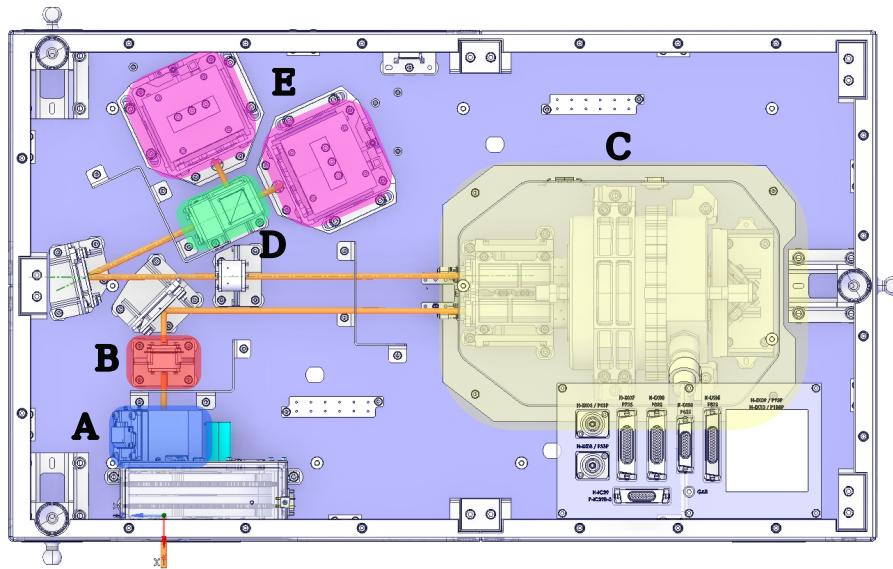


Figure 2.2 Schematic representation of the Tumag instrument. Some relevant optical devices in the light path (yellow line) are highlighted with a colored box and labeled with letters from A to E: A) Filter wheel, B) PMP, C) Etalon oven, D) beam-splitter and E) cameras. Image taken from TUMAG PAPER REF, reproduced with permission.

Light is delivered to TuMag by the ISLiD system and subsequently re-imaged onto two cameras where the images are recorded. Before reaching the cameras, the light passes through all the different subsystems of the optical unit. The first components encountered by the light are a blocking prefilter and the filter wheels (box A in Fig. 2.2). The blocking prefilter, with a wide bandpass centered at 520 nm, is employed to eliminate unwanted spectral ranges. The filter wheels are comprised by a double-disk system (Sánchez et al., 2022) that houses the prefilters for selecting specific spectral lines and a series of calibration modules. Specifically, three prefilters are mounted on the second disk of the filter wheel, corresponding to the spectral lines Fe I 5250.2 Å, Fe I 5250.6 Å, and Mg I b_2 5172.7 Å. Additionally, the filter wheels include a PD plate, which is used to introduce a known defocus into the final image to facilitate image reconstruction techniques, along with a linear polarizer, a plate of micropolarizers, and a pinhole set, employed during calibration observations.

After passing through the filter wheels, the light is directed into the Polarization Modulation Package (PMP), a subsystem derived from the SO/PHI instrument (Álvarez-Herrero et al. 2018a, Solanki et al. 2020), highlighted with the red box in Fig.2.2. The PMP's primary function is to modulate the light to produce the different polarization states required to deduce the Stokes components. This is achieved using two liquid crystal variable retarders (LCVRs), which are oriented with their fast axes at 45° relative to each other. These LCVRs induce a retardance on the transmitted light that varies with the voltage applied across the crystals. The system can operate in two distinct modulation schemes: a vector modulation scheme, which generates four independent linear combinations of equally-weighted Stokes components across consecutive observations, allowing for the retrieval of the full Stokes

vector after demodulation; and a longitudinal modulation scheme, which generates only two modulations, providing information solely on the intensity and circular polarization.

Following modulation, the light is directed into a LiNbO₃ Fabry-Pérot etalon, highlighted in yellow in Fig.2.2 (box C). Likewise IMaX, the etalon operates in a collimated setup and with a double pass configuration (Álvarez-Herrero et al., 2006). In this configuration, after the light passes through the etalon once, it is redirected by a pair of mirrors to pass through the etalon a second time. This double-pass configuration significantly enhances spectral resolution by narrowing the transmission profile. The LiNbO₃ etalon tunes the resonance wavelength by varying the refractive index of the cavity through the application of high voltages (ranging from -4000 V to 4000 V). Compared to air-gapped etalons, these kind of etalons offer the advantage of having no moving parts, which is particularly beneficial for spaceborne or balloon-borne instruments. However, this advantage comes with the need for precautions to prevent discharges caused by air ionization.

The final optical element the light encounters before reaching the cameras is a polarizing beam splitter (green box C in Fig.2.2). At this stage, the light beam is divided into two orthogonal, linearly polarized components, each directed towards a different camera. This dual-beam configuration (Lites, 1987) is designed to minimize spurious signals induced by jitter of the gondola (see del Toro Iniesta (2003) for an extended discussion), as it effectively cancels fluctuations from Stokes I to the other Stokes parameters that may arise due to image motion or solar evolution (*i.e.* cross-talk).

Light then reaches the cameras, shown with pink boxes (boxes E) in the scheme, where images from both are recorded and stored. After mission recovery, the data is processed on-ground to combine images from the different cameras, modulation states, and spectral lines, ultimately deriving the scientific products. This processing and reduction of the data is accomplished using software specifically developed for TuMag, which will be extensively discussed in Chapter 3.

2.2.2 Instrument performance and verification.

All subsystems within the TuMag light path function collaboratively to deliver high-resolution spectroscopic data of the solar spectrum. To ensure data quality, TuMag underwent multiple verification and calibration processes, during which its spectral, polarimetric, and imaging properties were meticulously tested. These procedures, commonly referred to as end-to-end (E2E) calibration tests (see Álvarez Herrero et al. (2022) for a detailed description of the tests), were conducted at various stages of the mission. Specifically, they were performed during the assembly, integration, and verification (AIV) activities with the stand-alone instrument at INTA facilities in Madrid, Spain; during the AIV phase of the PFI platform at MPS facilities in Göttingen, Germany; and during the TuMag AIV phase in the Sunrise III mission at ESRANGE facilities in Kiruna, Sweden. These tests were designed not only to validate the instrument's capabilities but also to measure critical parameters such as the tuning constant of the etalon, modulation matrices, and best-focus position—each of which is vital for the optimal operation of TuMag and the subsequent data processing. We will now delve into the details of the imaging, spectral and polarimetric properties of

the instrument as well as the verification processes and results, as the two are intimately related.

2.2.2.1 Imaging performance.

TuMag captures photons using two custom-made cameras (Orozco Suárez et al., 2023) equipped with GPIXEL back-illuminated GSENSE400BSI detectors, each featuring a $2k \times 2k$ pixel array, and specifically designed to meet TuMag's scientific requirements. These cameras provide a broad FoV of $63'' \times 63''$, sufficient to encompass an entire medium-sized active region, with a plate scale of $0.0378''/\text{pixel}$.

In order to fulfill the requirement of the wavefront error of $W \sim \lambda/14$, the instrument must have means to correct for the additional aberrations introduced by the telescope, the image stabilization and light distribution (ISLiD) system and uncorrected jittering. For this purpose, TuMag is equipped with a PD plate in the filter wheel that allows for the assessment of the PSF during the observations to apply image restoration techniques during the data processing.

The imaging E2E tests involved projecting several targets at the F4 focus, including a USAF test target, star targets, and a grid, observed both with and without the PD plate. These targets were utilized to evaluate the MTF and to assess the resolving power of TuMag. The PD measurements enabled verification of the wavefront error (WFE) derived from the MTF and an evaluation of the image quality following image restoration.

The USAF target^{*} consists on a series of horizontal and vertical line pairs (lp) arranged in sets of three with varying resolutions. Identifying the highest resolution group observable with TuMag allows for a fast diagnostic of the instrument resolution and performance. In fig. 2.3, measurements of group 4 and 5 (and higher) of the USAF target are shown for both cameras and the three pre-filters. The second set of group 5 (highlighted in a white box), which corresponds to 35.9 lp/mm in the target and 24.3 lp/mm in the image, is of special interest since its close to the Airy disk radius (26.4 lp/mm) and therefore close to TuMag's resolution limit.

The results show a better optical performance for the 517 nm pre-filter than the other two pre-filters. The USAF 5.2 set is clearly resolved fo this pre-filter in both cameras showing almost no difference between vertical and horizontal resolutions. However, results for the 525 nm prefilters exhibit a worsening of the resolution, with the same set being hardly resolved in the horizontal direction in both prefilters.

However, a more precise evaluation of the optical performance can be achieved from the MTFs. Figure 2.4 shows the MTFs computed with a slit target (see Huang et al. (2013) for a description of the MTF computation method) during the E2E tests performed in December 2021 at INTA facilities. These results agree with the diagnostic carried with the USAF tests: the 517 nm pre-filter shows a good performance in both directions, with values above the expected behaviour. Meanwhile, 525 pre-filters exhibit a large difference between different directions with an important drop in vertical resolution in both cases. This observed astigmatism is attributed to the etalon and physical deformations of the pre-filters caused by

^{*}The 1951 USAF target from Thorlabs Inc, model: R1DS1N.

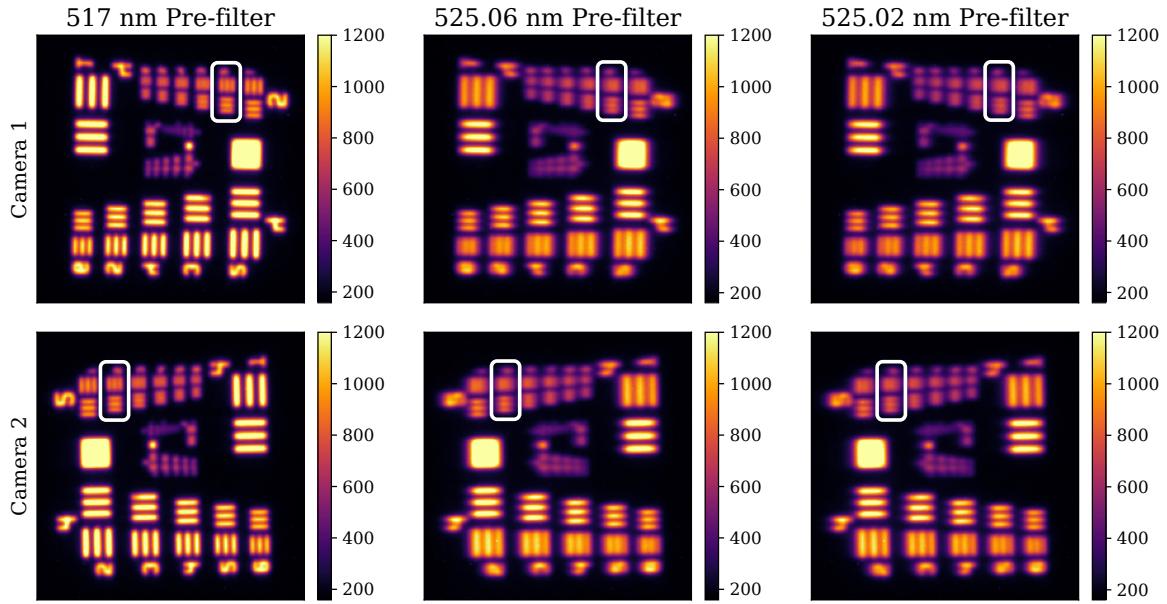


Figure 2.3 USAF target measurements for both cameras and the three pre-filters performed during E2E tests at INTA facilities on December 2021. The white boxes highlight the second element of the test group 5 (35.9 lp/mm). The scale of the images is set in digital counts.

the mechanical method used to secure and tilt them. This effect is particularly noticeable in the iron pre-filters due to the higher angles of incidence required for their tuning.

The comparison of the obtained MTF and the diffraction-limited one allows for an estimation of the Strehl ratio, and consequently the wavefront error (see section 1.4.1).

Table 2.2 shows the results for the Strehl ratios and WFE derived from this computation. All values, except for the horizontal resolution in camera 1 of the 517 nm prefilter are lower than the $\lambda/14$ set as a requirement. However, images can always be restored if $WFE \gtrsim \lambda/5$ (Vargas Dominguez, 2009b) if the PSF is known, thus the need for the inclusion of PD capabilities in the instrument. Furthermore, PD techniques not only allow us to enhance the optical performance of the instrument but also evaluate the optical performance during the calibrations in order to verify the results obtained through the computation of the MTF.

Figure 2.5 shows the measurements and results of the PD analysis for the 517 nm pre-filter and the camera 1. The measurements were carried out during the final E2E tests performed at Kiruna on April 2024 using the random dot target (left and central columns of the figure). The measurements consist on 5 sets of focused-defocused pairs of images. The PD algorithm is run over a zoomed-in region of 600 pixels in sub-patches of 128x128 pixels. The mean Zernike coefficients are shown in the top right panel, where the error has been computed as the standard deviation between different sub-patches. A 2D representation of the rms WFE is also shown in the bottom right panel.

The PD analysis indicates a small amplitude for most aberrations, with coefficients beyond Z15 approaching zero, except for the spherical aberration (Z_{11}, Z_4^0) which is the dominant contribution to the rms wfe. However, the results exhibit significant dispersion, as

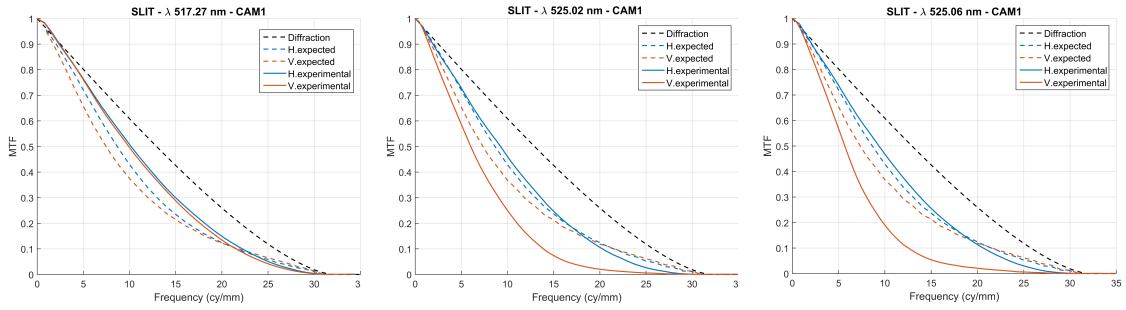


Figure 2.4 MTFs derived for camera 1 (very similar results for camera 2) for the three pre-filters from measurements of the stand-alone AIV phase performed at INTA in December 2021.

Pre-filter and camera	Strehl ratio	Strehl ratio	WFE	WFE
	Vertical	Horizontal	Vertical	Horizontal
517 nm - Cam 1	0.782	0.826	$\lambda/12.7$	$\lambda/14.5$
517 nm - Cam 2	0.761	0.806	$\lambda/12.1$	$\lambda/13.5$
525.02 nm - Cam 1	0.436	0.725	$\lambda/6.9$	$\lambda/11.1$
525.02 nm - Cam 2	0.405	0.726	$\lambda/6.6$	$\lambda/11.1$
525.06 nm - Cam 1	0.451	0.764	$\lambda/7$	$\lambda/12.1$
525.06 nm - Cam 2	0.444	0.736	$\lambda/7$	$\lambda/11.3$

Table 2.2 Optical performance evaluated from the MTFs obtained with the slit target at December 2021 E2E tests.

reflected by error bars that reach values up to 0.025λ for the first coefficients. Both the defocus and astigmatism are pretty low (Zernike indexes 4, 5 and 6, Z_2^0 , Z_2^{-2} and Z_2^2 , respectively), agreeing with the results obtained from the MTF analysis which showed a good resolution in both vertical and horizontal directions. The overall rms WFE obtained from this analysis is $\lambda/11.4$. It is important to note that the PD analysis shown here was carried out at the final stages of the calibration campaign, with TuMag mounted on the PFI and the light being fed to the instrument through the telescope and ISLiD system, whereas the MTF determination was conducted in the stand-alone AIV phase, without the aberrations introduced by these systems. Nevertheless, both analyses agree on a WFE better than $\lambda/10$, indicating very high optical quality, despite the fact that the FPI of TuMag operates in a collimated configuration, which is known to degrade optical performance (Scharmer, 2006).

2.2.2.2 Spectral performance.

TuMag filters wavelengths through a sequential process, beginning with a broad blocking pre-filter that eliminates unwanted regions of the solar spectrum, and followed by a second narrow-band pre-filter that is tuned to the three selected spectral lines. Finally, the LiNbO₃ Fabry-Pérot etalon is charged of selecting a very narrow band around specific

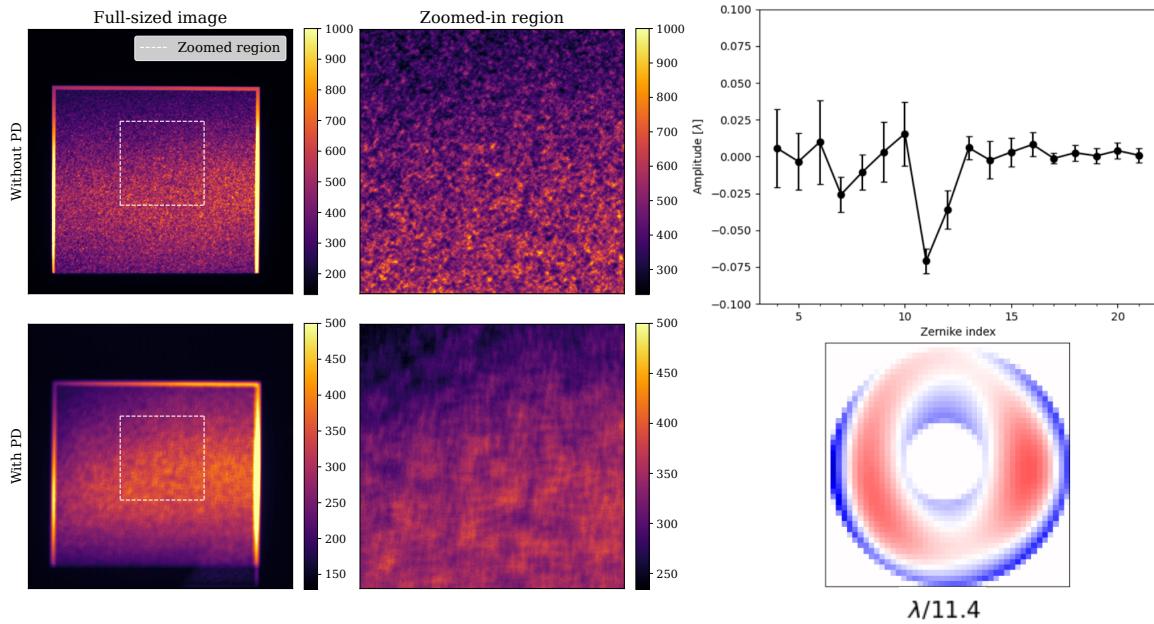


Figure 2.5 Random dot target measurements of the 517 nm pre-filter with the camera 1 and without the PD plate (left and central columns) taken during the Sunrise III AIV phase in Kiruna on April 2024. The right column shows the Zernike coefficients obtained from the PD analysis in the top panel and the 2D representation of the rms WFE. The PD analysis has been carried out by F. J. Bailén, reproduced with permission.

wavelengths along the spectral lines. The narrow-band pre-filter and the etalon are critical to TuMag's spectroscopic performance and require careful evaluation during calibration.

The three TuMag pre-filters were custom-manufactured by MaterionTM and have a full width at half maximum (FWHM) close to 1 nm. They are centered near the rest wavelength of the three spectral lines at normal incidence, with a peak transmission exceeding 80% in all cases. Each pre-filter was tuned by adjusting the incidence angle to align the peak transmission wavelength with the spectral line core. This process was performed using a coelostat at the INTA facilities, where the rest positions in volts of the spectral lines were determined. The Fe I 5250.2 Å line was found at 2129 V, the Fe I 5250.6 Å line at -2507 V, and the Mg I b_2 5172.7 Å line at -2245 V. While this tuning was successful, particularly for the iron lines, the spectral position of the pre-filters was found to be highly sensitive to illumination conditions. This sensitivity was evident from the shifts observed in the pre-filter measurements during the various stages of the assembly process. As illustrated in the left column of Fig. 2.6, the variation in the spectral position of the pre-filters is not sufficient to cause the spectral line to be blocked by the pre-filter, but it may result in the spectral line falling on the wing of the pre-filter during observations.

TuMag's etalon (see Table 2.3) operates in a collimated setup with a transmission profile with a FWHM of 0.87 pm (in the double-passs configuration), thus achieving a spectral resolution that exceeds the required 9 pm. Observations of an iodine cell illuminated with

Property	Value
Reflectivity	0.892
Thickness	281 μm
FWHM (double-pass)	0.8
Tuning Constant	3300 V/ \AA

Table 2.3 Tumag Fabry-Pérot specifications.

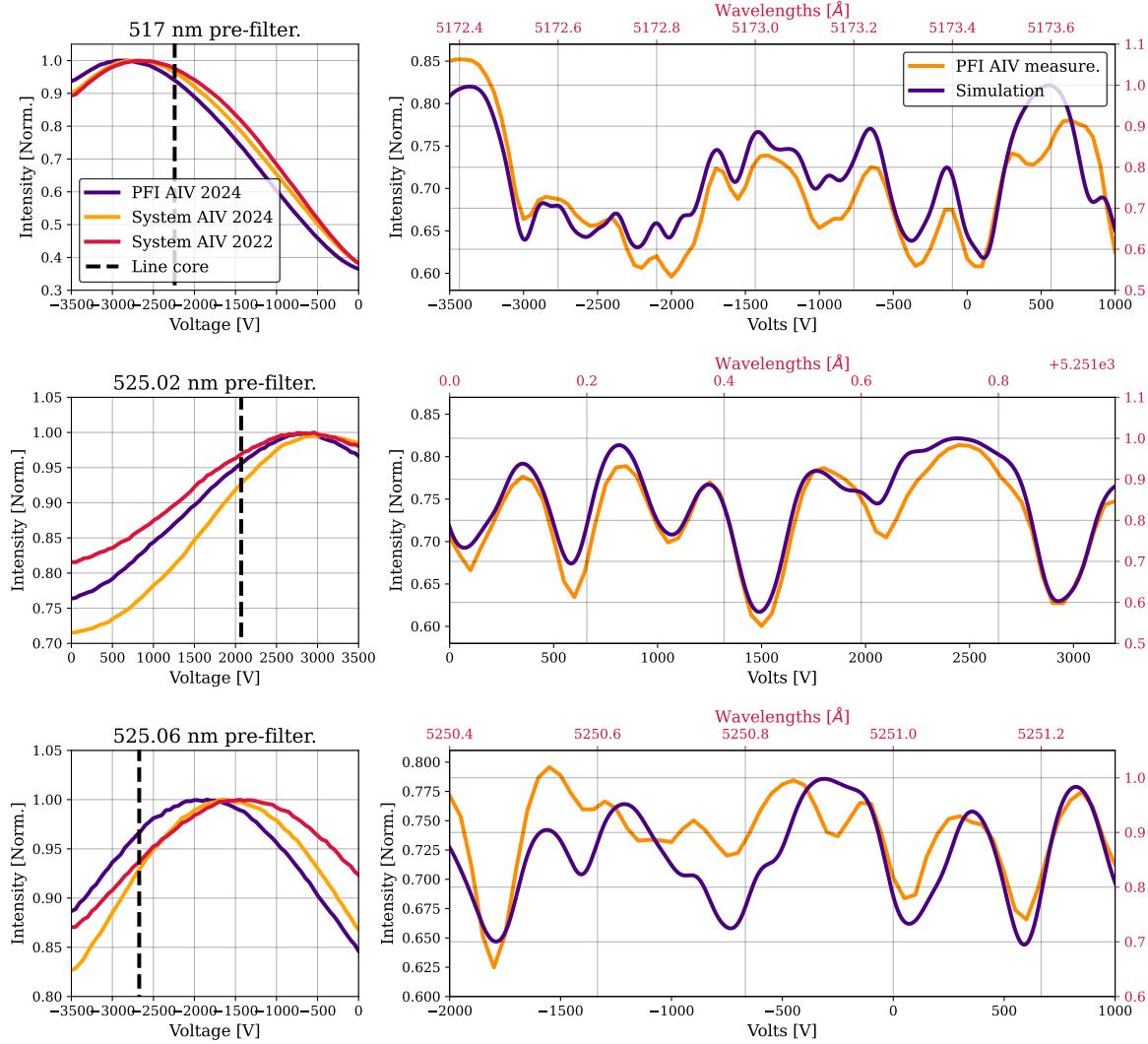


Figure 2.6 TuMag spectroscopic calibration results. Each row shows results for the 517 nm, 525.02 nm and 525.06 nm pre-filters, from top to row. The left column shows measurements of the pre-filters carried out with a flat LED on different stages of the AIV phases. The right column shows the fit of the I_2 cell observation with a simulation employing an etalon with a reflectivity of 0.892 (FWHM ~ 0.87). Note that the absolute value of the wavelengths of the simulation (red axis) might be shifted with respect to real values due to unknown conditions of the reference.

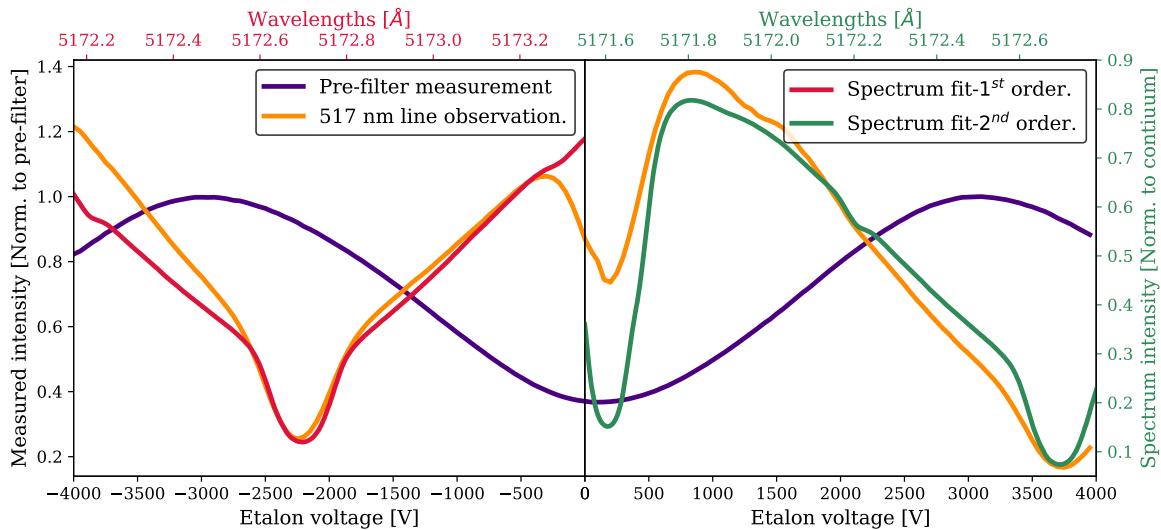


Figure 2.7 Results of the spectroscopic calibration during the end-to-end calibrations of the AIV phase of 2021. The dark blue curve represents the measurement of the 517 nm pre-filter, alongside an observation of the magnesium line using the coelostat at INTA facilities, shown in orange. Two different fits of the solar spectrum are overplotted on the figure. The red line represents a fit to the primary etalon order (negative voltages), while the green line corresponds to a fit to the second etalon order (positive voltages).

a diode were conducted to verify the transmission profile's shape and accurately assess the tuning constant. The right column of Fig. 2.6 presents, in orange, the iodine cell measurements obtained during the assembly, integration, and verification (AIV) phase of TuMag's integration into the Post Focal Instruments (PFI) platform, which took place at the Max Planck Institute for Solar System Research (MPS) in Göttingen, Germany, in November 2023. Additionally, the dark blue line in the figure represents a simulation of the iodine spectrum observations. This simulation was generated using an analytical model of the transmission profile of collimated etalons (see section 4.1.1 for a detailed overview of the model). The results confirm that the spectral resolution achieved in the iodine cell observations is consistent with the estimated 0.87 pm resolution. Furthermore, these observations enabled the calculation of the etalon's tuning constant by identifying the corresponding line cores between the simulation and observation and applying a least squares fitting to establish the relationship, which was measured in 3300 V/Å.

An observation of the solar spectrum with the 517 nm pre-filter, conducted at INTA facilities in December 2021 during the end-to-end calibration tests, is presented in Fig. 2.7, along with the corresponding pre-filter measurement. The magnesium line core is detected at approximately -2200 V using the primary order of the etalon and reappears around 3750 V with a secondary order. A fitting of the solar spectrum[†] is also shown for both orders. These results reveal significant contamination from the secondary order near the pre-filter's min-

[†]Reference

Spectral lines	Modulation	Vectorial				Longitudinal	
		I1	I2	I3	I4	I1	I2
525 & 517 nm	LCVR1 retardance	225°	225	315°	315°	180°	180°
525 & 517 nm	LCVR2 retardance	234.74°	125.26°	54.74°	305.26°	90°	270°
525 nm	LCVR1 voltage	2.291	2.533	1.992	1.947	2.761	2.761
	LCVR2 voltage	2.375	3.360	6.433	2.016	4.723	2.186
517 nm	LCVR1 voltage	2.343	2.580	2.031	1.972	2.797	2.797
	LCVR2 voltage	2.371	3.416	6.548	2.051	4.77	2.206

Table 2.4 Tumag LCVR retardances and corresponding voltages for both modulation schemes and the three pre-filters. Note that a single value is provided for both iron pre-filters.

imum transmittance. At around 0 volts, the observed spectrum (orange line) is a composite of contributions from both the primary (red line) and secondary (green line) orders. This contamination is particularly relevant for data processing, as continuum measurements of the magnesium line are typically conducted at -80 V. The broader profile of the magnesium line necessitates continuum measurements farther from the line core, making it more susceptible to this contamination. In contrast, the narrower iron lines do not require such extensive offsets for continuum measurements and are thus less affected.

2.2.2.3 Polarimetric performance.

TuMag modulates the incoming light through a PMP composed of two anti-parallel LCVRs. These devices can modify the phase retardance induced to the light that goes through them by changing the alignment of their molecules when subject to a voltage potential. Their advantages for airborne instruments lie in their lightweight and compact design, the low voltage required for operation ([0 – 10]V), and their efficiency in producing either four linearly independent modulation states for full-Stokes polarimetry or only two states for measuring the longitudinal component of the magnetic field through Stokes V. This versatility is a specific advantage of LCVRs, not found in quarter-waveplate-based PMPs (Pillet et al., 2004).

TuMag's polarimetric measurement approach is divided into the two modulation schemes already mentioned: vectorial and longitudinal. In the vectorial scheme, four linearly independent modulation states are generated in rapid succession by the PMP, enabling the calculation of the full Stokes vector. Conversely, the longitudinal approach generates only two modulation states, providing information on just two components. This modulation is designed to compute Stokes V by determining the quantities $I \pm V$.

Both modulation schemes are required to operate under an optimal modulation scheme. Such a scheme is defined by a modulation matrix with the following polarimetric efficiencies: $\epsilon_{opt} \geq [1, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$. The selected modulation scheme was based on the retardances outlined in Table 2.4. A thorough calibration of the liquid crystal variable retarders (LCVRs) was conducted to accurately determine the voltages necessary to produce the specified re-

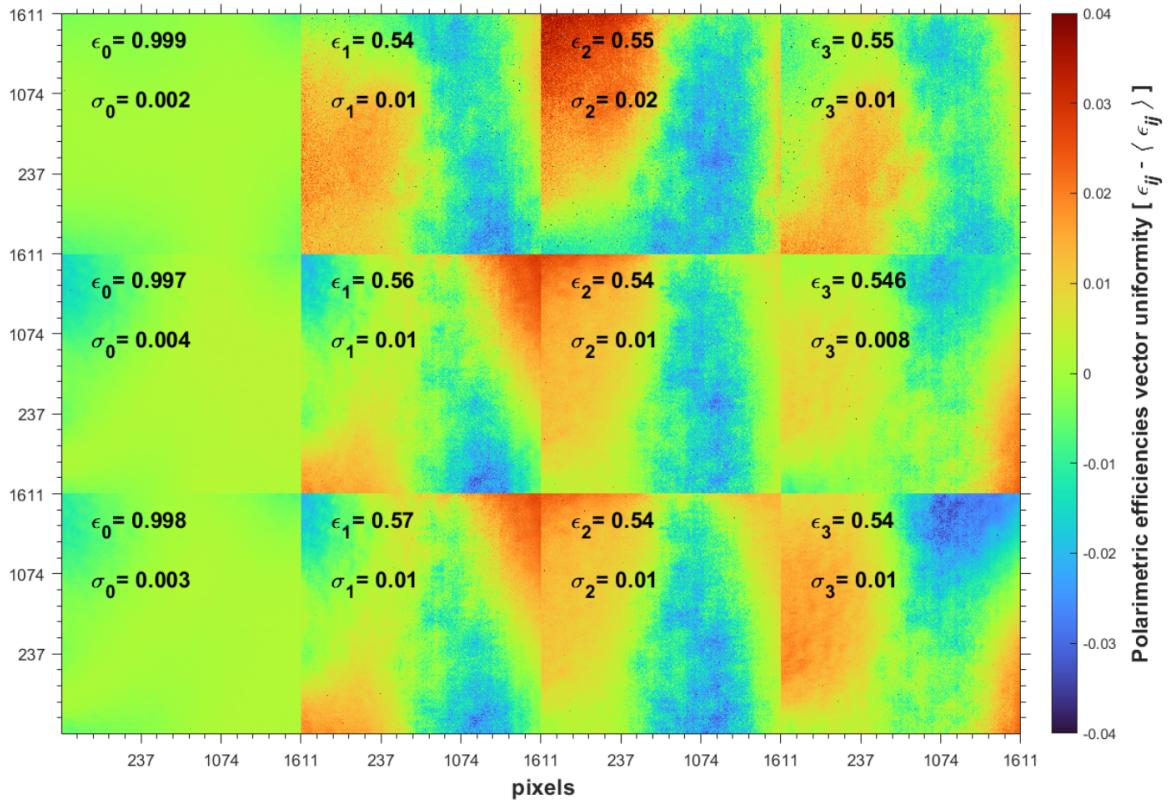


Figure 2.8 Polarimetric efficiencies for camera 1 and the three pre-filters (from top to bottom, the different rows show the results for 517 nm, 525.02 nm, and 525.06 nm). The different columns correspond to the efficiencies of the different Stokes components. The colormap measures the differences in efficiencies along the FoV. Results obtained during the E2E tests performed at INTA in December 2021, during the stand-alone AIV phase.

tardances (Álvarez-Herrero et al., 2018b).

Considerations on the (S/N) are critical for ensuring the required polarimetric sensitivity. Achieving an S/N of 10^3 in the Stokes measurements imposes a requirement of $S/N \approx 1200$ for each modulation measurement per camera. This calculation assumes near-optimal polarimetric performance, and takes into account the dual-beam polarimetry technique, which increases the S/N by a factor of $\sqrt{2}$ when combining data from the two cameras. A single shot of the cameras is insufficient to reach these S/N values, as the sensors do not have enough capacity in their electron wells. To address this, multiple exposures are captured and subsequently summed during each observation. This *accumulation* strategy, extensively tested and employed in various polarimeters (e.g., Elmore et al. 1992, Martínez Pillet et al. 1999, Lites et al. 2001), has proven compatible with image reconstruction techniques (González & Kneer, 2008; Van Noort & van der Voort, 2008). It allows for adjusting S/N levels depending on the scientific objectives of the observation, balancing between velocity and polarimetric sensitivity.

However, in order to fulfill the polarimetric sensitivity requirements, the modulation

matrix of the instrument must be carefully addressed during the polarimetric calibrations. Any deviation in the computation of the modulation matrix, will introduce spurious signals in the polarization measurements, known as cross-talk. The polarimetric calibration involves a series of measurements using a light beam with a known polarization state, generated by a rotating linear polarizer and a rotating quarter-waveplate. By varying the positions of these two devices, 40 different input polarization states were produced and measured with three pre-filters. These measurements allowed for the precise determination of the modulation matrix by solving the system of equations (1.15), where the only unknown is the modulation matrix \mathbf{M} , as both the measured modulation and the Stokes components of the incoming light are known. The modulation matrices for both cameras (indicated through the subindex) and all pre-filters that were determined through this process during the polarimetric E2E tests conducted at the Sunrise III AIV phase in Kiruna, Sweden, in 2022, are:

$$\begin{aligned} M_0^{517} &= \begin{bmatrix} 0.951 & -0.612 & 0.474 & 0.459 \\ 0.955 & -0.331 & -0.758 & -0.382 \\ 1.058 & 0.456 & 0.562 & -0.712 \\ 1.036 & 0.747 & -0.260 & 0.600 \end{bmatrix} & M_1^{517} &= \begin{bmatrix} 1.054 & 0.763 & -0.394 & -0.524 \\ 1.036 & 0.497 & 0.793 & 0.306 \\ 0.953 & -0.282 & -0.475 & 0.683 \\ 0.958 & -0.585 & 0.320 & -0.613 \end{bmatrix} \\ M_0^{525.02} &= \begin{bmatrix} 0.954 & -0.694 & 0.406 & 0.414 \\ 0.969 & -0.390 & -0.803 & -0.368 \\ 1.042 & 0.418 & 0.495 & -0.705 \\ 1.035 & 0.710 & -0.266 & 0.612 \end{bmatrix} & M_1^{525.02} &= \begin{bmatrix} 1.059 & 0.771 & -0.449 & -0.433 \\ 1.024 & 0.449 & 0.723 & 0.335 \\ 0.965 & -0.344 & -0.543 & 0.650 \\ 0.953 & -0.606 & 0.191 & -0.641 \end{bmatrix} \\ M_0^{525.06} &= \begin{bmatrix} 0.951 & -0.687 & 0.403 & 0.424 \\ 0.962 & -0.373 & -0.800 & -0.339 \\ 1.048 & 0.415 & 0.500 & -0.728 \\ 1.038 & 0.736 & -0.236 & 0.601 \end{bmatrix} & M_1^{525.06} &= \begin{bmatrix} 1.060 & 0.777 & -0.403 & -0.463 \\ 1.032 & 0.471 & 0.754 & 0.290 \\ 0.960 & -0.306 & -0.497 & 0.681 \\ 0.948 & -0.620 & 0.205 & -0.619 \end{bmatrix} \end{aligned}$$

The results of the polarimetric calibration performed during the end-to-end (E2E) tests at INTA in December 2021 are presented in Fig. 2.8. The figure shows the results for camera one; however, camera two demonstrated nearly identical efficiencies. The polarimetric efficiencies across the entire field of view (FoV) exceed the required thresholds $\varepsilon_{req} \geq [0.95, 0.45, 0.45, 0.45]$, and approach the optimal values. Furthermore, efficiency variations along the FoV are generally low, with standard deviations lower than 0.01. This homogeneity of the polarimetric performance makes the data reduction easier as no special treatment is required for specific regions.

CHAPTER 3

TUMAG'S PIPELINE AND DATA.

The 2024 observational campaign of the third edition of the Sunrise observatory was an outstanding success. In contrast to previous flights, where technical challenges severely limited the number of useful observations, all subsystems performed exceptionally well during this third flight, allowing for nearly continuous instrument operation over more than six days. From TuMag's perspective, the campaign yielded approximately 10 terabytes of data, consisting of over 40 scientific observation blocks and 250 calibration observations.

The substantial volume of data recorded by the three instruments, of which TuMag captured the least (in digital space) due to the instrument's nature, required that it be physically recovered on-site, as it could not be broadcasted from the observatory to the operations center. Recovery activities began immediately after landing and lasted until early August, during which all surviving components, along with the data vaults, were transported to Yellowknife, Canada, the nearest city to the landing site. The data vaults arrived at MPS in early August, where a backup was created before the data associated with each instrument was sent to the respective IP institution. TuMag's data arrived at IAA in late August, marking the official start of the reduction process.

The reduction process began by labeling all images and identifying the more than 600 000 images captured by TuMag. Once the observations were correctly identified, the reduction process commenced and, at the time of writing, remains ongoing. Due to the relevance of the pipeline development and results for this thesis, this chapter will provide an overview of TuMag's data and the state of its pipeline, although the results remain preliminary.

The discussion will begin by introducing TuMag's various observing modes, both scientific and calibration, followed by a brief review of the observation campaign, outlining the different observation programs and their scientific objectives. The chapter will conclude with an examination of the data reduction process, detailing the pipeline and presenting some initial results. It is important to note that, due to the late arrival of the data, this thesis had to be written in parallel with the reduction process. Therefore, the results presented here are preliminary, and the final product may differ as additional reduction steps are incorporated.

Observing mode	Spectral lines	N_λ	N_P	N_a	N_c	$t_{eff}(s)$	(S/N)
0s	Mg I b_2 5172.7 Å	12	1	2	1	6.3	500
0p	Mg I b_2 5172.7 Å	12	4	16	1	37.62	1000
1	Mg I b_2 5172.7 Å	10	4	16	1	31.81	1000
2	Fe I 5250.2 Å, Fe I 5250.6 Å	8	4	16	1	23.4	1000
3	Fe I 5250.2 Å, Fe I 5250.6 Å	5	2	20	1	10.04	1000
4	Mg I b_2 5172.7 Å	3	4	10	10	54.01	2500
5	Fe I 5250.2 Å, Fe I 5250.6 Å	3	4	10	10	53.60	2500

Table 3.1 Scientific observing modes. From left to right, the columns are: observing mode identifier, measured spectral lines, number of wavelengths, of modulations, of accumulations, of cycles, the total time and the polarimetric SNR.

3.1 TuMag's observing modes

With the purpose of simplifying the operation activities, TuMag operates through a series of so-called observing modes. The observing modes are a list of pre-configured settings tailored for various observations, including both calibration and scientific purposes. Each mode is designed to fulfill the specific objectives of the corresponding observation and enables nearly automatic operation of the instrument during flight.

A summary of the properties for each observing mode is provided in Table 3.1. There are four distinct modes designed to observe the magnesium line. Mode 0s performs a fast, extended scan of the spectral line using 12 wavelength samples: [-40, -30, -20, -10, 0, 10, 20, 30, 40, 50, 60, 65]*, with one modulation and two accumulations to maximize scanning speed. Mode 0p is similar to mode 0s but employs a full-vector modulation scheme, requiring 16 accumulations to ensure the required SNR. Mode 1 provides a shortened scan of the magnesium line, with measurements taken at [-30, -20, -10, -5, 0, 5, 10, 20, 30, 65], also utilizing a vectorial modulation scheme. Finally, mode 4 is a "deep" magnetic mode, featuring a highly reduced scan with only three samples at [-10, 0, 10], but with increased accumulations and cycles to enhance polarimetric sensitivity.

Three observing modes are configured for the iron lines. Mode 2 employs a vectorial modulation scheme applicable to both iron lines, with sampling at [-12, -8, -4, 0, 4, 8, 12, 22] pm. Mode 3 uses a longitudinal modulation scheme, measuring only Stokes I and V, with samples taken at [-8, -4, 4, 8, 22] pm. Lastly, mode 5 closely resembles mode 4, but is configured for the iron lines, with sampling at [-8, 0, 8] pm. The only difference between these two modes is the sampling scheme.

Although most of the parameters are set up by the observing mode and cannot be changed, there are some configurable parameters that allow to slightly modify the observing modes to fit the specific goal of a particular. These parameters are the following:

Solo estos? o había más?

- ✿ λ_{rep} : A parameter that allows to repeat all the observations carried out at every

*Sampling positions are given relative to the line core.

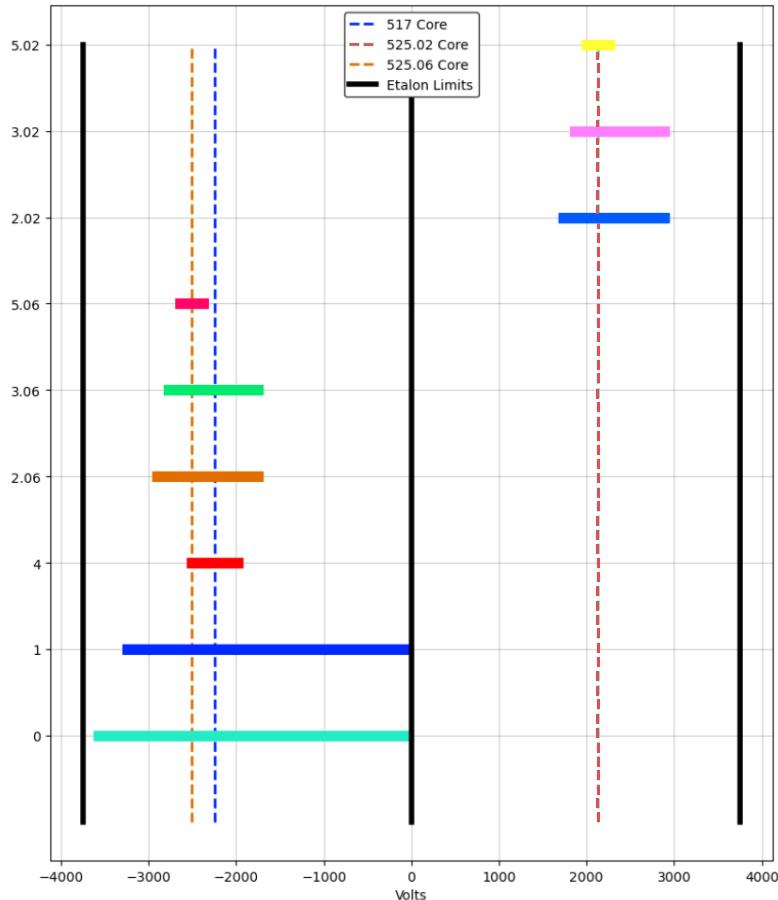


Figure 3.1 Schematic representation of the voltage range covered by all observing modes. The dashed lines indicate the position of the line core as measured during the E2E tests performed at INTA in December 2021. The black lines represent the voltage limits that cannot be crossed in an observing mode.

spectral position before changing wavelength. This parameter is employed for flat-field observations (see the following section). By default is set to 1.

- Etalon offset : A parameter that allows for the introduction of a global shift to the spectral sampling by offsetting the absolute voltages (and thus, wavelengths) of the scan. This parameter was used to center the spectral line in shorter observing modes affected by solar rotation or other effects that might shift the spectral position. The default value is set to 0 V.
- N_a : Even though the number of accumulations is fixed in nominal observing modes, this parameter was set as configurable in order to allow modifications for faster observations when needed. The default value depends on the observing mode.

Figure 3.1 presents a schematic representation of the voltage ranges for the observing modes when converting spectral sampling to volts. The black lines indicate the voltage boundaries that cannot be surpassed during an observation due to technical constraints. These limits are set at ± 3750 V as the maximum and minimum values, with an additional limitation at 0 V, since a polarity change poses technical challenges that could not be addressed within an observation mode. These restrictions are significant in two cases: firstly,

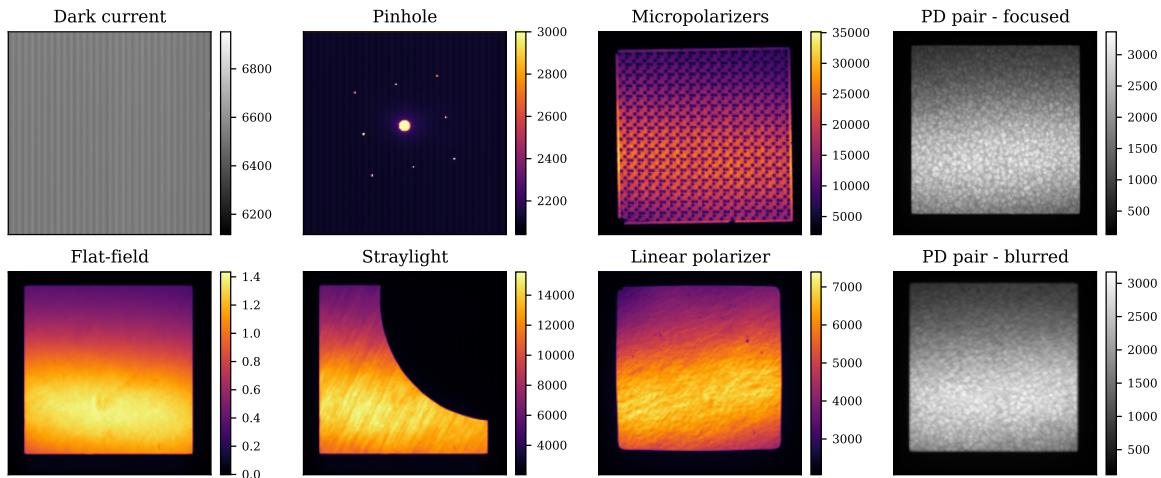


Figure 3.2 Examples of calibration observations. All images, with the exception of the flat-field, are presented in their raw format, without any manipulation or correction applied. The flat-field observation depicted corresponds to the first modulation of the continuum measurement obtained during a flat-field observation corresponding to observing mode 1. All data are belong to camera 1, and the colorbar is calibrated in digital counts save for the flat-field which is normalized to its mean value.

for Magnesium observation modes, specifically modes 1 and 0, where the continuum measurement is positioned as far from the core as possible, at -80 V, due to the 0 V crossing limitation. Secondly, these constraints are relevant when applying an etalon offset to shift the spectral positions of a particular observing mode, as the offset cannot cause the final positions to exceed these boundaries.

3.1.1 Calibration modes

An additional type of observing modes are also designed aimed at carrying out calibration observations. These calibration observing modes are more flexible than scientific ones, and allow for the configuration of several parameters to match the observations to the aim of the scientific observation.

3.1.1.1 Flat-field observations

One of the essential calibration procedures in any telescope-based astronomical observation is the acquisition of flat-field images. These observations are designed to measure intensity variations across the FoV, which arise from several factors, including intensity gradients induced by the etalon, dust particles, or pixel efficiency variations, among other sources. The aim is to capture a region with no discernible structure, ideally producing a uniformly flat intensity distribution. However, achieving such flat-field observations is not always straightforward, particularly for certain instruments. While ground-based tele-

scopes can utilize twilight periods to observe areas of the sky devoid of stars, space-borne or balloon-borne solar telescopes, such as Sunrise III, are unable to the same and must look for alternative methods.

In Sunrise III, flat-field images are generated by deliberately blurring the solar image through rapid movements of the mirror. This process effectively removes any solar structure from the FoV when averaging out multiple blurred observations, resulting in a flat-field image devoid of solar features.

In the case of TuMag, flat-field observations are performed using a modified version of the nominal observing mode, where the λ_{rep} is set to 4. Additionally, multiple consecutive instances (N_{reps}) of these observations are executed, typically 5 or 7. During data processing, a single flat-field is generated for each wavelength position and modulation state by averaging all corresponding observations.

Figure 3.2 shows an example of a flat-field observation, for one camera, modulation and wavelength (bottom left panel). The image shows a clear deviation from flatness in the measurement, primarily due to the etalon intensity gradient, which accounts for the change in intensity between the brighter bottom half and the darker top half, and some minor inhomogeneities over the FoV.

3.1.1.2 Dark-current observations

A second critical calibration procedure for any observation involving electronic cameras is the measurement of dark current. In the absence of incident photons, electrons within the camera's wells can still be randomly excited. This spontaneous excitation can be incorrectly interpreted as photon-induced counts when analyzing the data. Dark current observations are designed to characterize these random electronic excitations, which are primarily influenced by the camera's physical conditions, particularly temperature, so that they can be accurately subtracted from the final images.

For TuMag, dark current calibration involved capturing a series of 50 images with $N_a = 50$ with no light entering the instrument. As with flat-field observations, a single dark current frame for each camera is generated by averaging all individual observations. In the top left panel of fig. 3.2 a dark current shot is depicted, characterized by the vertical strips pattern.

3.1.1.3 Linear polarizer and micropolarizers observations.

TuMag's filter wheels are equipped with two targets designed to assess the instrument's polarimetric performance: a linear polarizer and a set of micropolarizers. Both targets are situated in the first filter wheel and are used in conjunction with the three distinct prefilters located in the second filter wheel. The linear polarizer serves to evaluate the polarimetric calibration, particularly by quantifying the level of cross-talk, as no circular polarization should be detected when using this target. The micropolarizers provide a more complete assessment, as they consist of multiple linear polarizers oriented at different angles.

Observations with this targets are carried with the three prefilters, at a single wavelength, located in the continuum of each line. For each measurement, a vectorial modula-

tion scheme is employed that allows for the derivation of the four stokes parameters. In the third column of figure 3.2 observations of both targets are shown.

3.1.1.4 Pinhole Observations.

Another calibration target included in the filter wheels is the pinhole target. This target blocks most of the light reaching the instrument, except for a few small holes arranged in a square-like pattern across the FoV, as shown in the top panel of the second column of figure 3.2. A larger hole is located at the center of the FoV, surrounded by eight smaller holes that trace a square with the central hole at its midpoint. These observations serve various purposes, including image alignment, detecting the presence of ghost images, or identifying etalon reflections, among other uses.

Pinhole observations are conducted similarly to those with polarizers, that is, in combination with the three prefilters at a single wavelength (the continuum of each line), but without applying any modulation.

3.1.1.5 Straylight target.

Not all the light that reaches the detector is necessarily the intended signal for a given observation. Some unwanted light, primarily originating from internal reflections along the optical path, may also reach the instrument. This unwanted contribution, known as straylight, contaminates the measurements by reducing contrast, lowering the S/N, and generally degrading the spectral, optical, and polarimetric performance of the instrument.

To address this contamination, TuMag performed a series of observations using a target that blocks part of the FoV (see the bottom panel of the second column of figure 3.2). By analyzing the dark region in these observations, it becomes possible to measure and model the straylight reaching the instrument, allowing for its subsequent removal from the data.

3.1.1.6 Prefilter scans.

TuMag observations are very sensible to spectral shifts either from the pre-filters or from the observed spectral line position. The shift of the pre-filters can happen due to changes in the physical conditions of the filter wheels such as changes in temperatures which spectrally shift the behavior of the pre-filters. The position of the pre-filter greatly affect the measurements as it reduces the intensity of the measurements that are obtained in the wings of the pre-filter. Due to solar rotation, or changes in the conditions of the etalon, although these are less likely, the spectral position at which the spectral lines are recorded can change. This effect is specially important in observing modes that require great spectral accuracy, such as the deep modes, where only three spectral positions close to the line core are employed.

In order to verify the spectral behaviour of the prefilter, as well as the position of the spectral line, a series of observations were carried out, usually before and after the scientific observations, where a spectral scan with a rich spectral smapling was taken for all the pre-filters employed in the observation. These scans, measure the voltage range of the specific line with a sampling of 100V and without modulating.

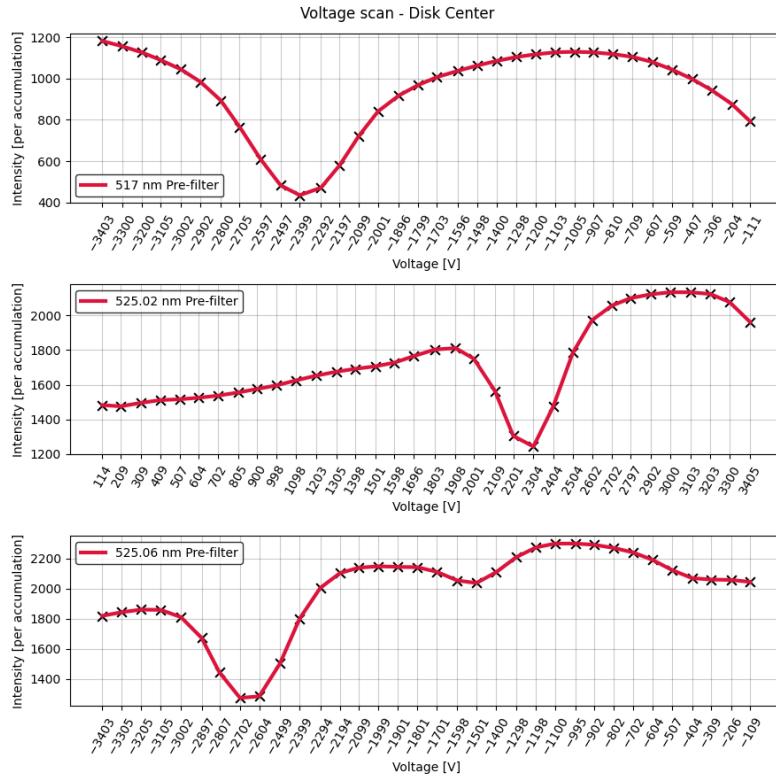


Figure 3.3 Schematic representation of the voltage range covered by all observing modes. The dashed lines indicate the position of the line core as measured during the E2E tests performed at INTA in December 2021. The black lines represent the voltage limits that cannot be crossed in an observing mode.

3.1.1.7 Phase diversity.

Lastly, TuMag is equipped with the capability to perform phase diversity for image reconstruction. As discussed in previous chapters, applying image reconstruction techniques is essential to meet the optical quality requirements. To this end, TuMag includes a PD plate in the first filter wheel that introduces a known defocus in the images. Capturing images with and without this plate enables the computation of the instrument's PSF, which can then be deconvolved from the data.

PD measurements require quasi-simultaneous pairs of aberrated and unaberrated images. Therefore, TuMag's PD observations consist of a series of 32 or 40 rapid, non-accumulated shots with the PD plate, followed by a corresponding series without the PD plate. The feasibility of this sequential scheme for phase diversity techniques has been confirmed in Bailén et al. (2022). A pair of focused-defocused images of quiet-sun observations is shown in the last column of figure 3.2.

3.2 Timelines

The operations of Sunrise III were designed to be nearly autonomous to ensure synchronization between the scientific instruments, the telescope, and the CWS. Given the limited time available for the observation campaign, this autonomy also helps to speed up operations, thus enabling more observation programs to be accommodated within the mission's

duration.

The Sun is a highly dynamic system, exhibiting a wide range of behaviors and phenomena, from large-scale structures such as active regions, sunspots, and flares, to smaller, quiet Sun structures where interactions at small scales drive the evolution of magnetic flux. This diversity, observable in various spectral lines and across different regions of the solar disk, demands multiple observations with distinct characteristics.

Prior to the first flight of Sunrise III in 2022, a series of timelines were developed to program both calibration and scientific observation blocks. These timelines were carefully designed by the Sunrise Science Team, taking into account the 70 observing proposals submitted for Sunrise, in order to prioritize observations that met the requirements of the majority of these proposals.

Observing proposals that could be fulfilled by targeting the same solar feature, while considering its disk position, were grouped into a single timeline. Each timeline included not only the necessary scientific observation blocks but also the required calibration observations to ensure data accuracy. Thus, timelines consist of a sequence of scientific and calibration observation blocks. The observing blocks within a timeline could vary in content depending on the scientific objectives and the status of the other instruments involved.

For simplicity, operations related to SCIP and SUSI will be excluded from the discussion, save for a few important remarks. In the case of TuMag, each observing block was composed either of a combination of two observing modes executed consecutively, or a single observing mode repeated throughout the block.

The timelines of the Sunrise III observation campaign can be grouped in the following blocks:

- ✿ Quiet Sun observations at disk center (QSDC), as the name implies, focus in regions near the solar disk center that are free from significant solar activity. These timelines typically involve long series of observations aimed at studying the small-scale structure and magnetic flux evolution in the quiet Sun.

There are four distinct timelines in this category: three standard timelines, which employ the nominal observing modes, and a different timeline, the (QSDC_HC). This timeline employs high-cadence variations of the standard observing modes specifically designed to enhance the temporal resolution between images, which is crucial for helioseismology techniques.

- ✿ Sunspot observations (SP) are specifically designed to study sunspots. There are four different timelines for this purpose. Some of these are short programs used to track the same sunspot over multiple days, with the goal of studying the evolution and decay of the sunspot. Others are more extensive programs aimed at examining, in greater detail, the magnetic activity of sunspots and their penumbral structures.
- ✿ Polar observations (PL) target the region close to the limb in both poles of the Sun. These areas are of special interest due to their distinct magnetic behavior compared to the disk center. Additionally, these regions provide the opportunity to measure faint signals outside the main solar disk, such as spicules in the lower chromosphere,

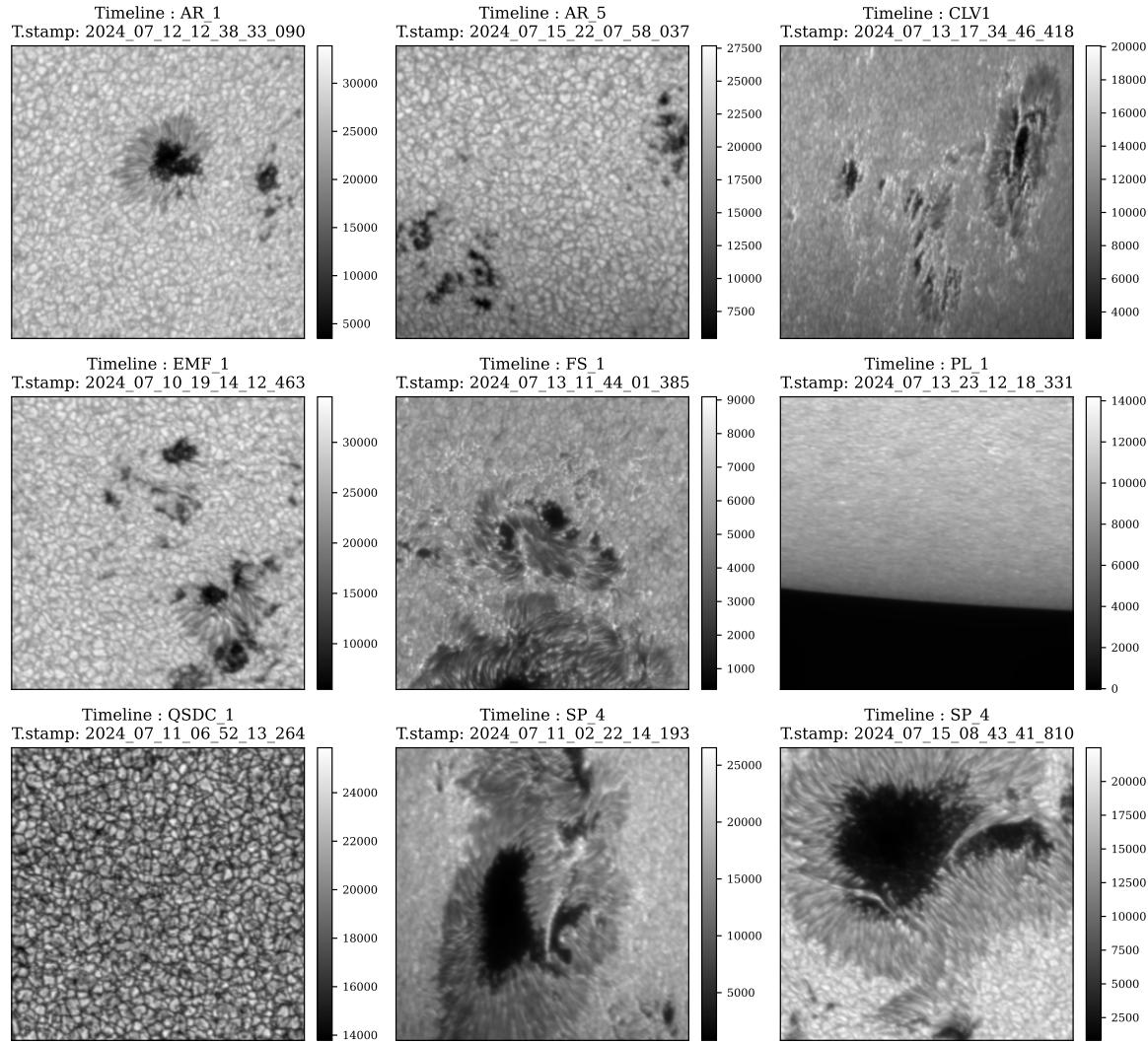


Figure 3.4 Continuum images of the first modulation of the first observation mode from different timelines. The image shown has been flat-fielded and dark-corrected. The timestamp provided corresponds to the first image of the observation mode. The colorbar is given in counts.

observed outside the continuum disk of iron. Two different instances of these timelines are conceived in the observation campaign, mostly differentiated in TuMag by the selected spectral lines.

- East and West limb (EW) observations are designed to target the equatorial regions of the solar limb. In addition to exhibiting magnetic structures distinctly different from those observed at the poles, the reason for having a separate timeline from the PL timelines lies in the orientation and technical constraints related to SCIP and SUSI's slits. The relative positioning of the regions and the inclination of the telescope introduce unique challenges. In these EW observations, the spectrometer slits are aligned parallel to the limb, contrasting with the PL timelines, where the slit is positioned perpendicular to the limb.
- Active regions (AR) observations are designed to study areas exhibiting solar activity, excluding those specifically focused on in the sunspot programs. These observations typically consist of two-hour series, employing the standard combination of the iron 525.02 nm and magnesium lines, using modes 1 and 2.02, which represent the most common observation block for TuMag. Although five different AR timelines were planned for the Sunrise campaign, only three were executed.
- Emergence flux (EMF) programs are specifically designed to study active regions that exhibit a large flux emergence. For TuMag, the observation blocks are shared with those of the AR programs, namely, the combination of mode 1 and 2.02 for series of around 2 hours.
- Full spectral scan (FS) observations are primarily designed for SUSI and SCIP, where their complete set of spectral bands is utilized. These scans are intended to be carried out in both quiet Sun and active regions. For TuMag, FS observations consist of long series focusing on the iron spectral line in quiet Sun regions, while in active regions, they include a combination of iron and magnesium observations.
- The flares programs (FL) were designed for target opportunities of a flaring region. These programs were intended to be activated only when an active region showed signs of flaring. For TuMag, the observations during these programs consist of the standard combination of iron 525.02 nm and magnesium spectral lines.
- Center-to-limb variation (CLV) observations were intended to target regions of the solar disk characterized by μ values that had not been previously observed. The parameter μ , defined as the cosine of the angle between the surface normal and the observer's line of sight, serves as a useful indicator of a region's proximity to the disk center. Specifically, μ ranges from 1 at the disk center to 0 at the limb (Thompson, 2006). Conducting CLV observations at previously unmeasured μ values enables us to capture data from different regions across the disk, facilitating studies of how observational features vary with their position on the disk.

During the Sunrise III observation campaign, 38 timelines were run, including calibration timelines in addition to the scientific programs presented here. Some examples of the different targets employed during the campaign are shown in fig. 3.4. A detailed record of TuMag's observations can be found online both in the pipeline's repository and in TuMag's official data website[†].

3.3 Pipeline

Before data can be employed for scientific purposes, it has to undergo a process where all the instrumental and spurious effects are removed and the necessary computations for the scientific aim are carried out. This process, usually referred to as data reduction, has to be specifically designed for each instrument, as the particular properties of the instrument come into play. Being a spectropolarimeter, TuMag's data reduction pipeline must, in addition to removing the instrumental artifacts, compute the Stokes components of the incoming light.

In this section we introduce the software that has been developed for TuMag's data processing. Due to the proximity of the data's arrival to the end of this thesis, its important to note that the data reduction is still in development, with some calibration steps still undeveloped. In the following, we present the current status of the pipeline, along with a few examples of the results. The pipeline is publicly available in a GitHub repository[‡].

3.3.1 Standard data reduction process.

The specific steps that have to be applied to a particular observation depends on the observation mode, and scientific aim of the observation, as different observations may require additional steps prior to the scientific exploitation. However, any observing mode that employs either two of the modulation schemes, share a series of common steps that must be followed. Taking into account that save for the high cadence timelines, and the observing mode 0s, all observations follow this scheme, this is the process that the majority of TuMag data has to undergo.

These steps include the basic corrections that have to be applied to all images, namely flat-fielding and dark current corrections. But additionally, the final steps must combine the data from both cameras, and compute the Stokes components. This process is illustrated in the block diagram in fig. 3.5 and consists of the following steps:

1. Dark current processing.
2. Flat-fielding processing.
3. Camera's alignment computation.

[†]https://www.uv.es/jublanro/tumag_data_test.html

[‡]https://github.com/PabloSGN/TuMags_Reduction_Pipeline

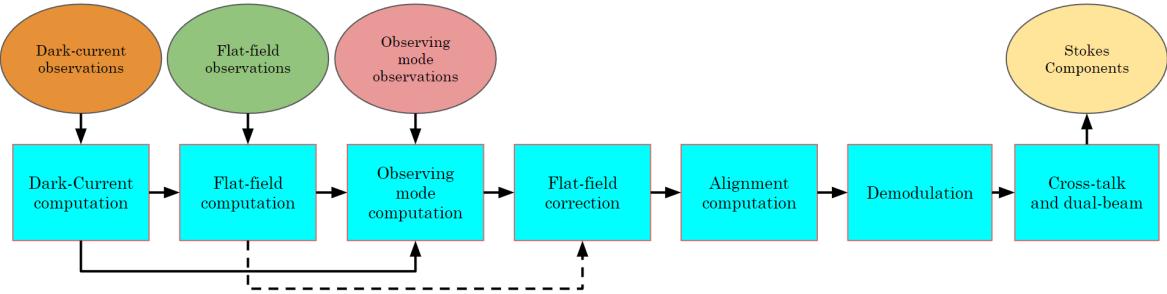


Figure 3.5 Block diagram of the standard reduction process: Blue boxes represent the individual steps that make up the reduction, while ellipses indicate the different sets of observations and the final product (yellow ellipse).

4. Processing of observing mode images (dark-current corrected).
5. Apply flat-field correction.
6. Demodulation.
7. Cross-talk correction and cameras combination.

The data reduction process begins with the dark-current processing, which involves averaging all individual dark frames within a specific set to generate a single dark-current frame per camera. This dark current is then subtracted from all images used in the reduction, including flat-fields, pinholes, and scientific observations, after rescaling the dark-current to the appropriate number of accumulations.

The second step is the flat-field computation. These observations, as previously mentioned, are a modified version of the nominal mode with an increased λ_{rep} and are repeated N_{rep} times. The processing involves averaging all images taken at a specific wavelength for the same modulation. Thus, $\lambda_{\text{rep}} \times N_{\text{rep}}$ images are averaged to produce a single flat-field frame. To maintain spectral line information, flat-fields are normalized to their mean value, as flat-fields taken at the line core have lower intensity than those in the continuum. The goal is to correct intensity variations within a single frame without altering relative intensities across different spectral points.

Having computed the dark-current and the flat-fields, the scientific observations can be corrected by subtracting the dark-current and dividing the resulting image by the flat-field of the corresponding wavelength and modulation.

After flat-fielding and dark current correction, the different modulations and the two cameras are combined to infer the stokes components for every measured wavelength. However, in order to combine the data from different cameras and modulations, the images have to be aligned beforehand. The alignment is carried out with the already corrected images and consists on a sequential process where the four modulations of the camera one are first aligned, and then the images of the second camera are aligned with respect to the image of the camera one and the corresponding modulation. The alignment is performed

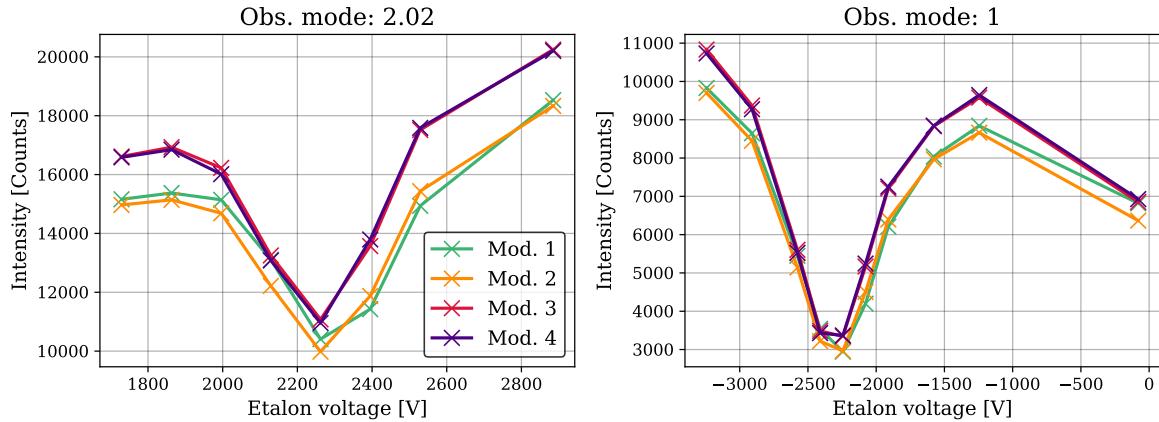


Figure 3.6 Block diagram of the standard reduction process: Blue boxes represent the individual steps that make up the reduction, while ellipses indicate the different sets of observations and the final product (yellow ellipse).

at subpixel accuracy employing the methods described in ?, where the alignment is computed employing a two-dimensional cross-correlation in the Fourier domain.

After alignment, the Stokes components are derived through a process known as demodulation. This process is carried independently at different wavelengths and consists on a matrix multiplication employing the demodulation matrix and the four modulations in the vectorial scheme:

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \underbrace{\begin{pmatrix} d_{00} & d_{01} & d_{02} & d_{03} \\ d_{10} & d_{11} & d_{12} & d_{13} \\ d_{20} & d_{21} & d_{22} & d_{23} \\ d_{30} & d_{31} & d_{32} & d_{33} \end{pmatrix}}_{\mathbf{D}} \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{pmatrix}, \quad (3.1)$$

where the matrix \mathbf{D} is the inverse of the modulation matrix of the corresponding camera and pre-filter that were computed during the calibrations (see sec. 2.2.2.3). And for the longitudinal scheme:

$$\begin{pmatrix} I + V \\ I - V \end{pmatrix} = \begin{pmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}. \quad (3.2)$$

From the matrix multiplication the Stokes components for each camera are derived. However, the demodulation is never perfect, due to small deviations of the demodulation matrix from the real one, or instrumental artifacts that have been uncorrected by the flat-fielding. These defects result in a contaminated Stokes components, where information from one component appears in other components, typically from Stokes I into Q , U , and/or V . Nonetheless, these contamination, known as cross-talk, can be extracted from the data.

The cross-talk correction is a manual process as different data sets require different corrections. For instance, some data sets may not show contamination from I to U while others do. Moreover, the correction might require to be applied using the information from

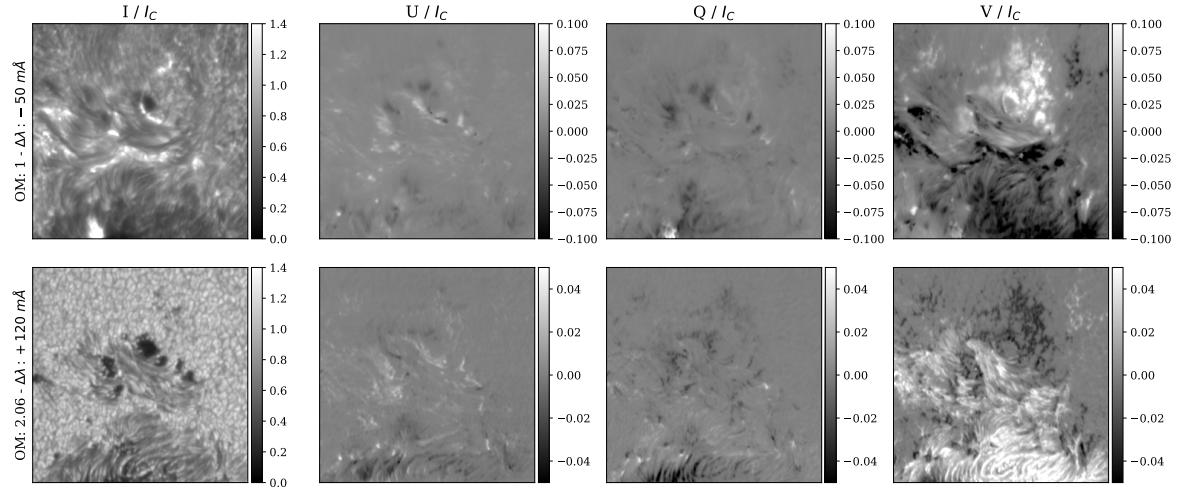


Figure 3.7 Hola hola karakola

the whole FoV or only of a small region. Thus, this is a correction that has to be carefully applied and it's hard to standardize to the whole set of observations. Nonetheless, the concept of the correction is the same.

The correction of the cross-talk from I to any other component starts by measuring the relation between the two components by fitting through a least-squares method, a polynomial of first order in order to compute the tendency, if there is one. Once this relation has been established, and the strength of the cross-talk (*i.e.* the value of the slope) measured, the correction is applied by simply, removing the tendency of the data:

$$S_{corr} = S_{orig} - (I_{orig} * a + b), \quad (3.3)$$

where the relation between the stokes component S_{orig} and stokes I I_{orig} has been fitted to the line: $S_{orig} = I_{orig} * a + b$

3.3.2 Extra calibration blocks.

3.3.3 Image reconstruction.

3.3.3.1 Linear polarizer calibration.

3.3.3.2 Micropolarizers calibration.

3.3.3.3 Prefilter scans.

CHAPTER 4

CHALLENGES IN DATA REDUCTION

ETALON CAVITY MAP

Fabry-Pérot interferometers (FPIs) are widely employed in the field of solar physics. Their spectroscopic and tunability properties make them especially suitable for selecting a narrow spectral band of incoming light. They also offer a two-dimensional view of the solar scene, hence allowing for the implementation of powerful and widespread image post-processing reconstruction techniques, such as phase diversity (?) and multi-object multi-frame blind deconvolution (MOMFBD; ?), which are difficult to implement in slit-based spectrographs (?). Many state-of-the-art instruments use FPIs as narrowband tunable filters. Among others, these instruments include the spaceborne Polarimetric and Helioseismic Imager (Solanki et al., 2020) aboard the Solar Orbiter mission (Müller et al., 2020) (SO/PHI); the Imaging Magnetograph Experiment (IMaX) instrument (?), which flew on the first two flights of the balloon-born SUNRISE observatory (Barthol et al. 2011, Solanki et al. 2017); and the Tunable Magnetograph (TuMag) instrument for its third edition. These instruments are based on solid LiNbO₃ etalons. Regarding ground-based instruments, some examples include the Crisp Imaging Spectro-Polarimeter (CRISP) at the Swedish 1-m Solar telescope (?) at the Observatorio del Roque de los Muchachos in La Palma, Canary Islands; the GREGOR Fabry-Perot Interferometer (GFPI; ?, ?) at the Observatorio del Teide in Tenerife, Canary Islands; the Visible Tunable Filter (VTF; ?) developed for the *Daniel K. Inouye* Solar Telescope (DKIST; ?) of the Haleakalā Observatory in Hawaii; and the future Tunable Imaging Spectropolarimeter (TIS) of the European Solar Telescope (?), all of which are based on air-gapped etalons.

Thus far, we have focused on TuMag, its properties, and the correction of its data due to its central role in this thesis. However, there are other tunable spectropolarimeters, each with its own specific characteristics, advantages, and challenges. In this chapter, we shift our focus away from TuMag and turn our attention to another class of magnetographs, specifically those with their etalon in a telecentric configuration.

Currently, there is an ongoing debate within the scientific community regarding the optimal optical configuration for the etalon in spectropolarimeters, whether it should be

collimated or telecentric. This debate has garnered significant interest, as evidenced by the numerous studies addressing the properties of each configuration. Notable contributions include the four works by Bailén et al. (Bailén et al. 2019, ?, ?, ?), an extensive analysis on topics such as the impact of defects, applications for instrumentation, and the analytical formulation of the etalon's transmission profiles. Other examples include the studies by (?) and (Scharmer, 2006), which analyze the optical performance of FPIs, as well as the works of ? and ?, which discuss the properties of the etalons in the Telecentric Etalon Solar Spectropolarimeter (TESOS) and the Interferometric BIdimensional Spectrometer (IBIS), respectively. Additionally, the review of ? provides a broad and comprehensive overview of this topic. The reality is that it is challenging to provide a definitive answer to this question, as each configuration has its own advantages and disadvantages.

The primary argument against collimated mounts is the amplification of wavefront errors resulting from the large footprint of the light rays on the etalon plates, which includes both large and small scale defects (?). This amplification deteriorates the optical quality of these setups and has promoted the use of telecentric setups, where the footprints of the light rays are much smaller and the optical performance is less affected by large scale defects.

Despite its potential benefits, the telecentric configuration is no exempt of challenges. In this configuration each light ray traverses the etalon through different regions, thus being affected differently at each point from the various small scale defects present in the etalon. These defects, compose the so-called, cavity-map and its correction can be one of the main challenges of employing this setups.

This chapter aims to provide a comprehensive examination of telecentric configurations for etalons within the context of solar spectropolarimeters. We will begin by exploring the analytical formulation of the transmission profiles and spatial PSF of FPI's in both collimated and telecentric configurations. Following this, we will delve into the effects of the cavity error in solar observations, through a simulation of an observation of a Sunspot. Then, we will present a strategy to derive the cavity map from flat field observations in an attempt to improve the data correction. This strategy is presented and tested in a forthcoming academic paper under the title of "Correcting Fabry-Perot etalon effects in solar observations" *, and the contents of the corresponding section will be extracted from there.

4.1 One device, two configurations.

We established in section ?? that the observed intensity distribution at the coordinates ξ , η of the focal plane of any etalon-based instrument tuned to a wavelength λ_s obeys the following expression (Bailén et al., 2019):

$$I(\xi, \eta; \lambda_s) = g(\xi, \eta) \int_0^{\infty} T(\lambda) \iint O(\xi_0, \eta_0; \lambda) S(\xi_0, \eta_0; \xi, \eta; \lambda - \lambda_s) d\xi_0 d\eta_0 d\lambda, \quad (4.1)$$

*DOI : <https://doi.org/10.1051/0004-6361/202449825>

where $T(\lambda)$ accounts for the presence of an order-sorting pre-filter, $O(\xi_0, \eta_0; \lambda)$ represents the brightness distribution of the observed object at the point (ξ_0, η_0) , $S(\xi_0, \eta_0; \xi, \eta; \lambda - \lambda_s)$ accounts for the imaging response of the instrument when tuned at the wavelength λ_s , and $g(\xi, \eta)$ represents a spatial gain factor that accounts for wavelength independent pixel-to-pixel intensity fluctuations occurring in the focal plane due to differences in the detectors' sensitivity.

The imaging response of the instrument coincides with the PSF of the instrument when the optical response is invariant against translations. In such a case, we can substitute the last two integrals by the convolution operator, but this is not strictly true in etalon-based instruments, where the response varies pixel to pixel either because of etalon irregularities or because of variations in the illumination across its clear aperture.

Deriving S typically requires determining the electric field of the polychromatic wave in the image plane. According to Bailén et al. (2019), this field can be calculated by summing all the electric fields ($E^{(t)}$) across the pupil, such that the electric field at any point (ξ, η) of the image plane follows the expression:

$$E_{im}^{(t)}(\xi, \eta) = \frac{1}{\pi R_{pup}^2} \int \int_{pupil} E^{(t)}(x, y) e^{-ik(\xi x/f + \eta y/f)} dx dy, \quad (4.2)$$

where (x, y) are the coordinates in the pupil, f stands for the focal length, and R_{pup} is the radius of the pupil. It can also be shown that the vector electric fields transmitted by the etalon can be expressed as:

$$E^{(t)}(\xi, \eta) = \frac{\sqrt{\tau}}{1-R} \frac{e^{i\delta/2} - Re^{-i\delta/2}}{1+F \sin^2(\delta/2)} E^{(i)}, \quad (4.3)$$

where τ is the transmission factor for normal incidence, R stands for the reflectivity of the etalon mirrors, F is a factor defined as $F \equiv 4R(1-R)^{-2}$, $E^{(i)}$ are the incident electric fields, and δ stands for the phase differences between the transmitted and incident rays.

The transmission profile of the etalon (S) is defined as the average ratio between the transmitted and incident intensities, calculated from the complex conjugate of the corresponding electric fields. The shape of δ varies depending on the illumination setup of the etalon, whether collimated or telecentric. Consequently, each configuration has unique solutions and characteristics and is differently affected by inhomogeneities (defects) in the etalon's properties.

4.1.1 Collimated configuration

Collimated mounts are characterized by having the etalon located at the pupil plane and therefore receive a collimated beam from each point of the observed object. As illustrated in the schematic on the left side of fig. 4.1, light from any point on the object will fall on the same area of the etalon. Consequently, any local defects on the etalon crystals or on the plates' parallelism is averaged all over the clear aperture, thus making the optical quality

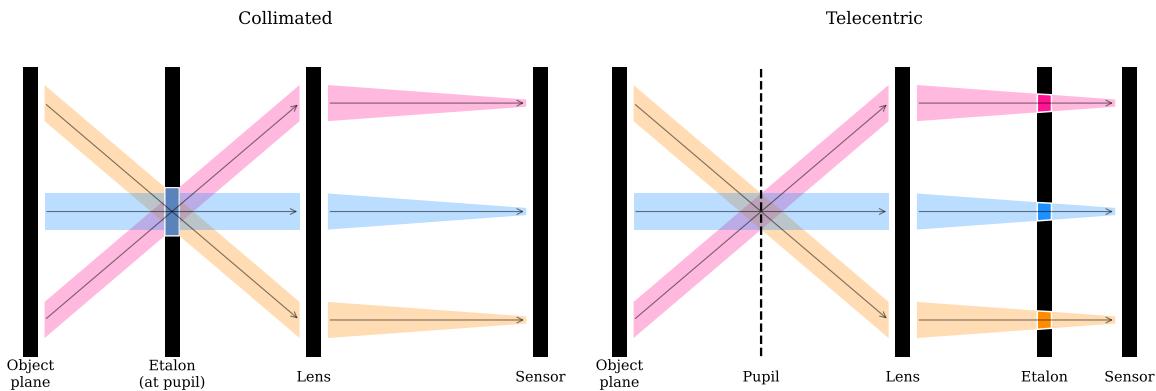


Figure 4.1 Schematic representation of the two optical setups of an FPI, collimated (left) and telecentric (right). The different colors represent distinct light rays originating from various points on the object plane. The white boxes in the etalon highlight the sections that are traversed by the light rays.

constant along the FoV. However, the angle of incidence of the light beam varies along the FoV, thus shifting the spectral transmission profile.

Analytical solutions of equation (4.2) are impossible to obtain if δ has a dependence on the pupil coordinates, as is the case of collimated etalons with a presence of local defects. In that case, the transmission profile of the etalon has to be evaluated numerically.

However, in order to study the spectral behavior of the transmission profile, we can, as a first-order approximation, disregard the spatial PSF and focus solely on the spectral transmission profile (ψ). Under this assumption, the phase difference between the incident and transmitted rays of an ideal collimated etalon can be expressed as:

$$\delta(\xi, \eta, \lambda) = \frac{4\pi}{\lambda} nd \cos(\theta(\xi, \eta)), \quad (4.4)$$

where n stands for the refractive index of the etalon cavity, d is the distance between the mirrors and θ is the angle of incidence. In such a case, it can be shown that the spectral transmission profile follows the expression:

$$\psi(\xi, \eta, \lambda) = \frac{\tau}{1 + F \sin^2(\delta(\xi, \eta, \lambda)/2)}. \quad (4.5)$$

The spectral behaviour of the transmission profile, such as the spectral position of the resonance peaks and the distance between them (the free spectral range), is encoded in the parameter δ , which is a function of the refractive index of the etalon cavity, the distance between mirrors, and the angle of the incident beam. The reflectivity R of the mirrors determines the width of the resonance peaks through the parameter F , $F \equiv 4R(1 - R)^{-2}$.

Local defects in the collimated configuration are averaged out, which means that d and n respectively represent the mean values of the thickness and refractive index across the clear aperture of the FPI, and thus remain constant for every pixel. Yet, they produce a broadening of the transmission profile and worsen the optical quality of the instrument.

However, the spatial dependence of ψ naturally arises from θ , which varies from pixel to pixel.

Assessing the spatial PSF of the FPI is more challenging, as it can only be determined analytically for monochromatic light and in the absence of defects. We will not delve into the equations for this specific scenario as it lies beyond the scope of this thesis. However, interested readers are referred to the work of Bailén et al. (2019), where this topic is extensively discussed.

4.1.2 Telecentric configuration

In the telecentric configuration, the etalon is placed very close to an intermediate focal plane, while the pupil is focused at infinity. As shown in the sketch on the right side of fig. 4.1, in this setup, the etalon is illuminated by cones of rays that are parallel to each other, thus passing through different sections of the interferometer. Local inhomogeneities (defects or cavities) on the etalon produce differences in the transmission profile across the FoV, which are directly mapped into the image plane. This means that the optical response and the transmission profile shift locally on the image sensor.

In telecentric configurations, δ always depends on the coordinates of the pupil, even in the absence of defects, since each point in the etalon sees a cone of rays coming from different parts of the pupil. Thus, as stated before, solutions to equation (4.2) are to be found numerically. Nonetheless, if we neglect the spatial PSF once again, we can derive an analytical expression for an ideal telecentric etalon, where all light cones impact perpendicularly.

After some messy algebra and clever variable changes, one can recast eq. (4.2) in terms of the radial coordinates of the pupil and analytically solve the equations. The resulting spectral transmission profile of an ideal telecentric etalon is given by (?):

$$\Psi(\xi, \eta, \lambda) = \Re [E(a(\xi, \eta, \lambda), b(\xi, \eta, \lambda))]^2 + \Im [E(a(\xi, \eta, \lambda), b(\xi, \eta, \lambda))]^2, \quad (4.6)$$

with $E(a(\xi, \eta, \lambda), b(\xi, \eta))$ being:

$$\begin{aligned} E(a(\xi, \eta, \lambda), b(\xi, \eta)) = & 2\sqrt{\tau} \left\{ \frac{1}{\alpha_1} [\arctan(\gamma_1) - \arctan(\gamma_2)] + \right. \\ & \left. i \frac{1+R}{1-R} \frac{1}{\alpha_2} \left[\ln \left(\frac{(1+\gamma_3)^2 + \gamma_4^2}{(1-\gamma_3)^2 + \gamma_4^2} \right) - \ln \left(\frac{(1+\gamma_3)^2 + \gamma_5^2}{(1-\gamma_3)^2 + \gamma_5^2} \right) \right] \right\}, \quad (4.7) \end{aligned}$$

where, the auxiliary functions are defined as:

$$\begin{aligned}
 a(\xi, \eta, \lambda) &\equiv \frac{2\pi}{\lambda} n(\xi, \eta) d(\xi, \eta) , \\
 b(\xi, \eta) &\equiv \frac{1}{8n(\xi, \eta)^2 (f\#)^2} , \\
 \alpha_1 &\equiv 2ab\sqrt{F} , \\
 \alpha_2 &\equiv 2\alpha_1\sqrt{F+1} , \\
 \gamma_1 &\equiv \sqrt{F} \sin a , \\
 \gamma_2 &\equiv \sqrt{F} \sin(a[1-b]) , \\
 \gamma_3 &\equiv \sqrt{\frac{F}{F+1}} , \\
 \gamma_4 &\equiv \frac{\tan(a/2[1-b])}{\sqrt{F+1}} , \\
 \gamma_5 &\equiv \frac{\tan(a/2)}{\sqrt{F+1}} .
 \end{aligned} \tag{4.8}$$

The parameter a has the same role as δ for the collimated case. However, the dependence on the image plane coordinates in this case is caused by potential variations in n and/or d , as each light beam traverses different sections of the etalon. These variations constitute the "cavity map" of a telecentric etalon and must always be taken into account during data reduction processes.

The parameter b accounts for the contribution of the focal ratio, $f\#$, and has an impact on the spectral resolution and the apodization of the pupil as seen from the etalon (?). Thus, the resolution is now affected by both F and $f\#$, through the parameters a and b .

4.1.2.1 Telecentric imperfect configuration

The equations shown in Sect. 4.1.2 are valid whenever the incident cone of rays is perpendicular to the etalon mirrors. We refer to this situation hereinafter as "perfect telecentrism". However, real instruments are likely to present deviations from such an ideal case. These deviations can be caused by an intentional tilt of the etalon to suppress ghost images on the detector (Scharmer, 2006), by an accidental tilted angle of incidence caused by deviations from the ideal paraxial propagation of rays within the instrument, or simply because of misalignment of the optical components. In the three cases, the incident cone of rays is no longer perpendicular to the etalon, and hence, we consider these scenarios to have imperfections in the telecentrism degree. One important consequence of the loss of telecentrism is an asymmetrization of the transmission profile that must be accounted for when modeling the instrument response.

The transmission profile in this case is influenced by the angle of incidence of the chief ray at each point of the clear aperture of the etalon, in addition to the parameters mentioned in the previous sections. Unfortunately, the equations for the transmission profile in these

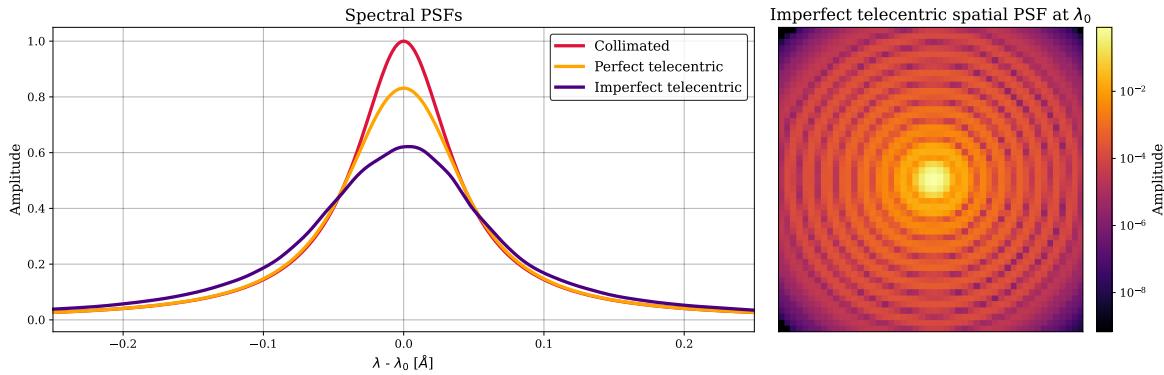


Figure 4.2 *Left*: central peak of the etalon’s spectral transmission profile for the three different configurations. *Right*: Spatial PSF of the imperfect telecentric etalon at λ_0 . The parameters of the etalon are $R = 0.92$, $n = 2.29$, $d = 251 \mu\text{m}$, $f\# = 56$, $\theta = 0^\circ$ (collimated and perfect telecentric), and $\Theta = 0.3^\circ$ (imperfect telecentric).

configurations are much more complicated than in the ideal case, and can no longer be analytically solved. We must revisit equation (4.2) and solve the integrals numerically.

Figure 4.2 shows on the left the transmission profile corresponding to the three different scenarios: collimated illumination of the etalon, perfect telecentrism, and imperfect telecentrism. The etalon parameters have been selected to coincide with those of SO/PHI’s etalon. In both the collimated and perfect telecentric configurations, a normal incidence ($\theta = 0$) scenario is shown, whereas in the imperfect telecentric case, we assumed an angle of incidence of the chief ray, Θ , of 0.3° . The parameter a has been adjusted slightly in order to tune the transmission profile at λ_0 .

Regarding the properties of each profile, note that the telecentric configurations achieve lower peak transmissions than the collimated case. In addition, the telecentric profiles are wider due to the different incidence angles across the illuminating cone of rays. Such a broadening increases with decreasing f-ratios. Lastly, non-normal incidence of the chief ray in the telecentric configuration further widens and shifts ($\sim 4 \text{ m\AA}$ for $\Theta = 0.3^\circ$) the profile, making it asymmetrical.

4.2 Sunspot observation simulation

The analytical formulations equip us with the tools necessary to investigate the impact of the cavity map on observational data and the consequences of incorrect or incomplete corrections. To this end, we will simulate a series of observations of a sunspot using a telecentric FPI and attempt to correct the observations utilizing generated flat-fields. Our aim is to quantify the spurious effects of the cavity map on scientific products, specifically the line-of-sight (LOS) velocities and magnetic field strength.

We have chosen to employ a sunspot for this study since they are one of our main *laboratories* for the study of the magnetic field and dynamics of the solar photosphere. Sunspots

are dark regions of the solar surface where strong magnetic fields gather. The innermost part of the spot, the umbra, with a brightness around the 20% of its surroundings, hosts very strong magnetic fields oriented almost perpendicular to the solar's surface. Surrounding the umbra in fully developed sunspots is the penumbra, a brighter region characterized by a radial filamentary structure with a nearly horizontal outflow of the plasma known as the Evershed flow (?).

The physical mechanisms driving the phenomena in these regions are not yet fully understood and have been a topic of discussion for many years, and still are. Processes such as the Evershed flow lack a universally accepted theoretical model, and numerous studies attempt to provide a theoretical basis for the observations, often presenting conflicting interpretations. Some examples of these works are (?), or (?), where they interpret the evershed flow as a hot gas confined to magnetic flux tubes that rise due to convection processes. In contrast, alternative models such as the one proposed by ? and ?, and later updated in ?, suggest that the observed filaments are field-free gaps where standard convection occurs. According to these models, the convective cells are elongated as the upward flow is redirected by the inclined magnetic field characteristic of the penumbra, resulting in the outward flow.

Accurate measurements of velocities and magnetic fields are of paramount importance for deciphering these processes, as the weak and small-scale manifestations of these effects may be crucial for their understanding. Examples for this are the works of ? or ?, where they found and study lateral downflows at the edge of penumbral filaments, supporting the model of penumbral convection from filament dynamics. However, these downflows are very challenging to detect, and require very high spatial resolution and accurate velocity measurements. For this reason, we will focus on the study of the spurious effects of the cavity error in the measurements of velocities and magnetic fields, with a special emphasis on penumbral flows.

The target of our observations will be a Magnetohydrodynamic (MHD) simulation of a sunspot in the photosphere. We will observe it through the eyes of a telescope of **XX meters of aperture** and an FPI similar to the one flying aboard the SO as part of the PHI instrument. This section begins with an overview of the simulated data (Sect.4.2.1), followed by a detailed explanation of the observation simulation (Sect.4.2.2). Subsequently, we present and discuss the different flat-fielding strategies in section 4.2.3. The section concludes with the discussion of the measurements of the LOS velocities and magnetic field strength in sections 4.2.4 and 4.2.5, respectively.

4.2.1 Simulated data.

Description of the simulation.

Structure of the simulation.

In order to obtain the magnetic field, we first need to simulate a modulation-demodulation scheme as the ones employed in spectropolarimeters like TuMag or SO/PHI. For this purpose, four different modulations have been generated with etc.

4.2.2 Observations simulation.

All the instruments built around the use of an etalon as a wavelength filtering element operate in a very similar way. They scan a spectral line by tuning the etalon (by changing the distance between mirrors and/or modifying the refractive index) to a desired number of wavelengths along the spectral line. At each spectral position, the solar scene is recorded. The measured intensity is approximately given by Eq. (4.1), with the etalon's transmission profile centered at the desired wavelength.

We have carried out two sets of simulations, one with a perfect telecentric configuration, where the incidence is normal, and one with an angle of incidence of the chief ray is 0.5° . For each of these sets, a simulated observation of the sunspot has been carried out employing 45 wavelengths for the scan.

We also need to simulate flat-field observations in order to correct the data. These observations have been computed with the same configuration than the one employed for the sunspot observation. For the flat-field observations we have generated an *ideal* flat-field where an average profile computed over a region of quiet-sun was replicated at every pixel of the generated image. One of the primary objectives of these simulations is to assess the effects of cavity errors in telecentric setups. Therefore, we conducted these simulations under two scenarios: one incorporating a cavity map and another assuming a defect-free etalon, which serves as our reference.

Although, in practice, it is not often possible to fully characterize the pre-filter, we assumed it has a rectangular shape centered at the wavelength of the observed spectral line (λ_0) and a width of $2\Delta\lambda$ such that only one order of the etalon passes through. With this consideration, equation (4.12) can be written as follows:

$$I(\xi, \eta; \lambda_s) = g(\xi, \eta) \int_{\lambda_0 - \Delta\lambda}^{\lambda_0 + \Delta\lambda} \iint O(\xi_0, \eta_0; \lambda) S(\xi_0, \eta_0; \xi, \eta; \lambda - \lambda_s) d\xi_0 d\eta_0 d\lambda. \quad (4.9)$$

For every tuned wavelength, the intensity in a given pixel is given by equation (4.9), where the observed object is the sunspot simulation or the generated *ideal* flat-field.

The imaging response of the FPI is computed by solving the integrals in equation (4.2). Given that we will account for the PSF in these simulations, numerical methods are necessary to evaluate the integrals. The computational burden is significant, as the integrals must be evaluated $N_{pixel}^2 N_\lambda N_{mods} N_{sets} N_{cavities} > 2 \times 10^8$ times[†] for both sunspot and flat field observations. This extensive computation results in excessively long simulation times. To address this issue, a neural network was developed to reduce the computational time from 12 s to 5 ms, more than a factor 2×10^3 faster. The neural network was trained with known solutions to the integrals for a fixed set of etalon parameters, ensuring that its outputs are accurate to within 0.01% for parameters within the training set.

In real-world scenarios, variations in pixel sensitivities (gain), in addition to the etalon imperfections (cavity map), are unavoidable. Including the effect of the gain into the simulation is straightforward as it only acts as a multiplicative factor that modifies the final

[†] $N_{pixels} = 561, N_\lambda = 45, N_{mods} = 2, N_{sets} = 2, N_{cavities} = 2$

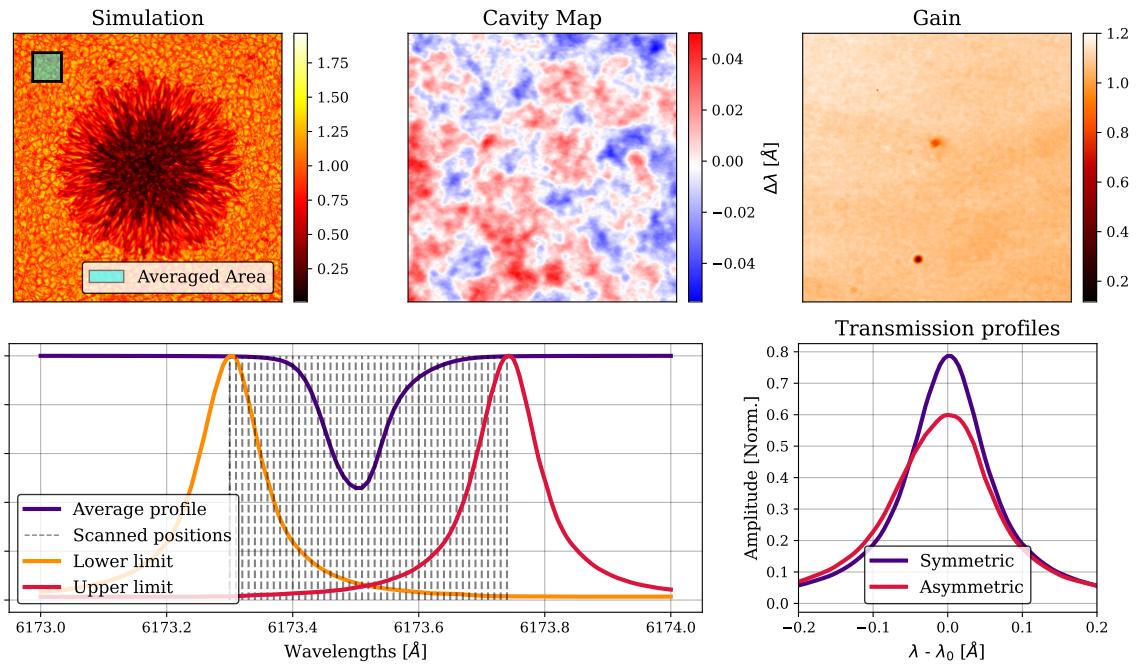


Figure 4.3 Inputs for the simulation of the sunspot observation. The top row shows, from left to right, the continuum of the MHD simulation, the cavity map expressed as the corresponding shift in \AA and the gain map. The bottom row shows, again from left to right, a representation of the quiet-sun average profile with all the scanned wavelengths, and the transmission profile of the two different FPI configurations, the symmetric (perfect telecentric) and asymmetric (imperfect telecentric). The parameters employed for the etalon are: $R = 0.892$, $n = 2.3268$, $d = 281 \mu\text{m}$, $f = 60$, $\Theta = 0^\circ$ (symmetric), $\Theta = 0.5^\circ$ (asymmetric).

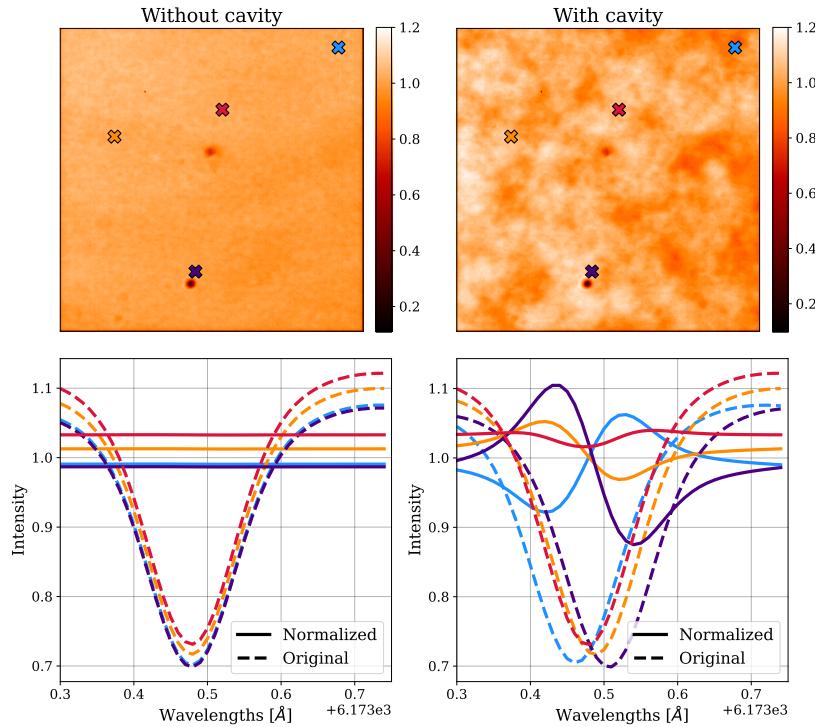


Figure 4.4 The top row shows the observed flat-fields at the core of the line for both the simulations without (left) and with (right) cavity. The bottom row shows the spectral profile of the pixels marked with an "X" of the corresponding color, for both scenarios, and before and after the normalization.

intensity. In contrast, the effects of the cavity map are more cumbersome. Pixel-to-pixel variations in etalon defects shift the transmission profile of the FPI and must be accounted for when computing the imaging response. Both the gain and cavity map introduced in the simulations have been selected to resemble the real-case scenario of the SO/PHI instrument. The gain map (Fig. 4.3 top right panel) utilized in the simulations is derived from a flat field of the SO/PHI-HRT instrument, and the cavity map employed (Fig. 4.3 top central panel) is sourced from the cavity map of PHI's etalon.

4.2.3 Flat-field correction.

In principle, the flat-field correction is straightforward: first, the measured flat fields are normalized to preserve the spectral line information; and second, the observations for all tuned wavelengths are divided by the corresponding normalized flat field. However, in telecentric setups, this process is more challenging because flat-field observations are influenced by cavity errors, which can introduce additional artifacts into the measurements rather than (or in addition to) correcting them.

Figure 4.4 shows the flat-fields observed at the core of the line for the simulations with and without cavity and the spectral profiles for a series of pixels. The effect of the cavity map in the flat-field observations is clearly displayed in both representations. In the flat-field itself, higher values of the cavity map appear brighter, as any displacement from the line core to either the blue or the red will result in an increase of the final intensity. In the spectral profiles, the pixels corresponding to higher values of the cavity error show a

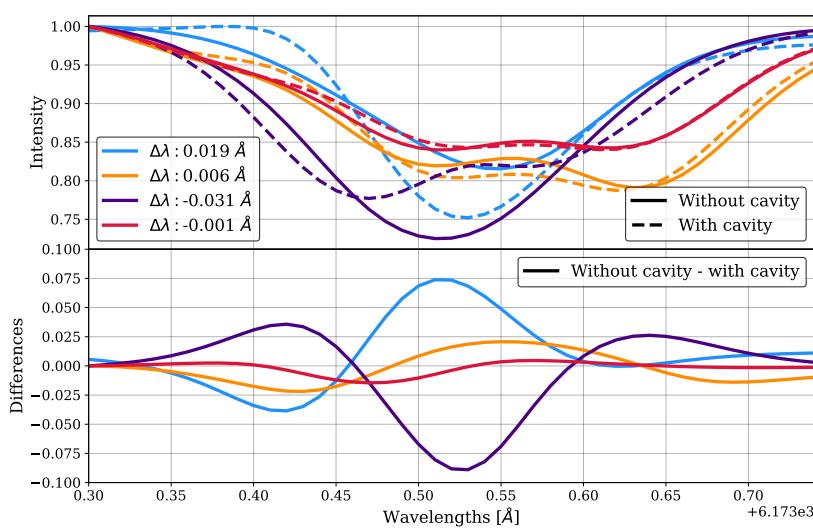


Figure 4.5 The top panel shows four spectral profiles after the flat-field correction for the two scenarios, with and without cavity, and for the symmetric FPI. The selected pixels are the same than in Fig. 4.4, and correspond to different values of the cavity map. The bottom panel shows the difference between the profiles with and without cavities.

noticeably shift with respect to the rest position of the line core.

From the normalized profiles for the scenario involving a cavity, it is evident that the flat-field profiles are not flat, as expected, given their role in correcting the observed profiles from the cavity-induced shift. However, the amplitude of the variations from flatness in these profiles is excessively large. This significant deviation presents a problem, as it will not only shift the other profiles but also alter their shape

This issue is illustrated in Fig. 4.5, which displays the profiles after flat-field correction for both simulations, with and without the cavity map. Ideally, if the flat-field correction effectively compensates for the FPI cavity errors, the profiles from both scenarios should be identical or at least very similar. However, upon comparison, significant differences emerge, particularly for pixels with high cavity map values (light blue and dark blue lines). These profiles are not only shifted but also exhibit alterations in their shape.

This challenge is why some data reduction pipelines adopt a different approach. Instead of using the flat-fields of the corresponding wavelengths, only the continuum flat-field is employed for the correction. When the continuum is measured sufficiently far from the spectral line, it remains unaffected by any spectral shift. By using this measurement, one can correct for spurious effects unrelated to spectral shifts, such as pixel sensitivity variations or dust grains. However, this method does not correct for the effects of the cavity.

If the cavity map is known, the velocity associated with the spectral shift caused by the cavity can be calculated. This "velocity-error" map can then be used to correct the Doppler velocity map derived from the observations that have not been corrected for the cavity map. Theoretically, this approach should address the cavity effects in velocity calculations. Nonetheless, potential impacts of the cavity map on magnetic field calculations remain uncorrected, as the relationship between wavelength shifts and magnetic fields is not as straightforward as in the case of velocity.

4.2.4 Velocity maps

We begin our analysis by studying the LOS velocities and the errors associated with their computation. These velocities are derived from the spectral shift of each profile relative to the rest position of the spectral line, using the Doppler formula (Eq. (1.23)). Given that this computation relies exclusively on the spectral shift, it is particularly susceptible to errors caused by cavity effects.

According to the center-of-gravity method (Uitenbroek, 2003), the central wavelength of a given spectral profile, λ_{COG} , can be computed from:

$$\lambda_{COG} = \frac{\int \lambda (I_{cont} - I) d\lambda}{\int (I_{cont} - I) d\lambda}. \quad (4.10)$$

Once λ_{COG} has been determined, the spectral shift of the corresponding profile will be given by $\Delta\lambda = \lambda_0 - \lambda_{COG}$.

We have computed the velocities employing the two flat-field correction strategies mentioned in the previous section: the "standard" approach, where each observation is corrected with the flat-field of the corresponding wavelength; and the "continuum" approach, where only the continuum flat-field is employed. Figure 4.6 shows the continuum intensity and the velocity maps obtained for both strategies in the top row.

At first glance, the velocities derived for each approach might look identical, but under closer inspection, some differences arise. In order to highlight these differences and to assess the accuracy of the obtained velocities we will compare them to the velocity obtained in the scenario where no cavity map was introduced. Figure 4.6 shows this comparison in the bottom row, along with the cavity map with the umbra and penumbra overplotted to be able to align the error with the cavity.

The error in the velocity computation shows that the effects of the cavity map remain uncorrected in the observations. Both correction approaches exhibit errors with structures corresponding to the cavity map. For the standard correction, this effect is pervasive throughout the field of view (FoV), whereas for the continuum approach, it is primarily concentrated in the umbra, although other regions also exhibit this behavior.

Moreover, the errors are significant, given that regions of the penumbra show errors of around 10%. Errors in the umbra are similarly significant in percentage terms, if not worse; however, studies of these regions are less frequent due to their low SNR resulting from their inherent darkness. Conversely, as discussed in the introduction, penumbral flows are of particular interest, prompting us to prioritize our discussion on the results in that area.

A more quantitative analysis is shown in Fig. 4.7, where the velocity error along a circular path is shown, for both approaches, in addition to the associated spectral shift of the cavity values at identical positions. This representation clearly demonstrates the correlation between the structural features of the cavity map and the velocity error. For both approaches, with a more pronounced effect observed in the standard approach, the absolute value of the error escalates with increasing spectral shifts of the cavity.

It is evident from the previous analyses that the flat-field correction approach significantly affects the results, as differences are observed between various correction methods.

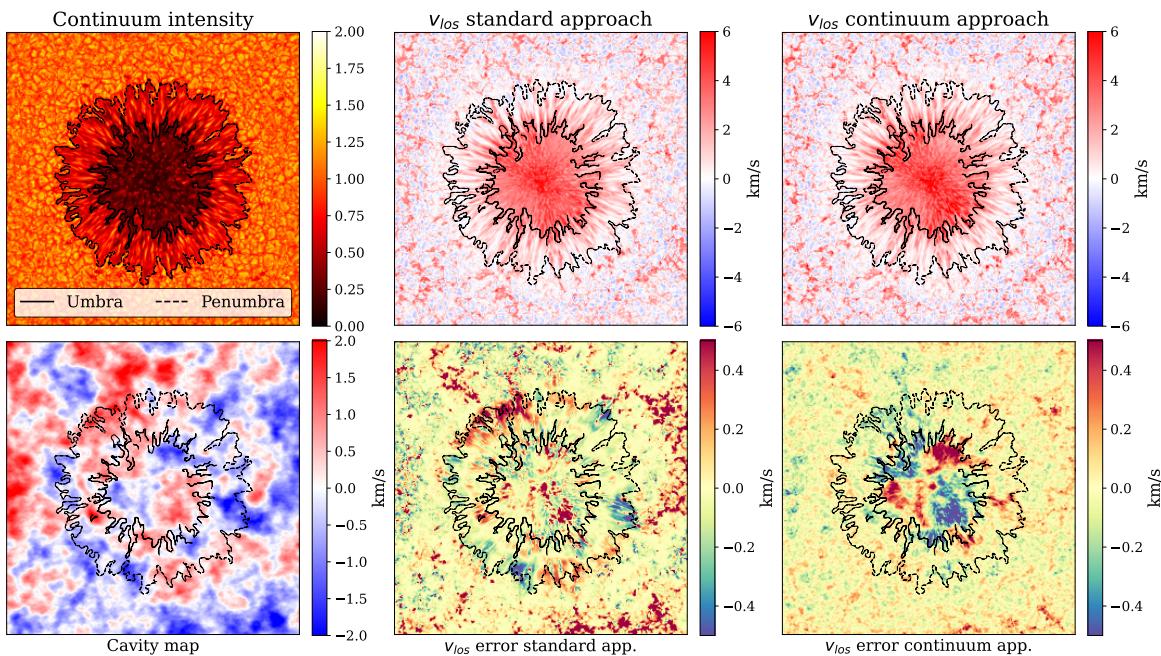


Figure 4.6 The top row shows, from left to right, the observed intensity in the continuum, the LOS velocity for the standard approach and the LOS velocity for the continuum approach, all three for the symmetric FPI and belonging to the simulations with cavity map. The bottom row shows, again from left to right, the velocity error associated with the cavity map, the error for the LOS velocity for the standard approach and the error for the LOS velocity of the continuum approach. The error is computed by subtracting the measurement to the reference. All representations show the boundaries of the umbra and penumbra for an easier identification of the regions in the velocity and cavity maps.

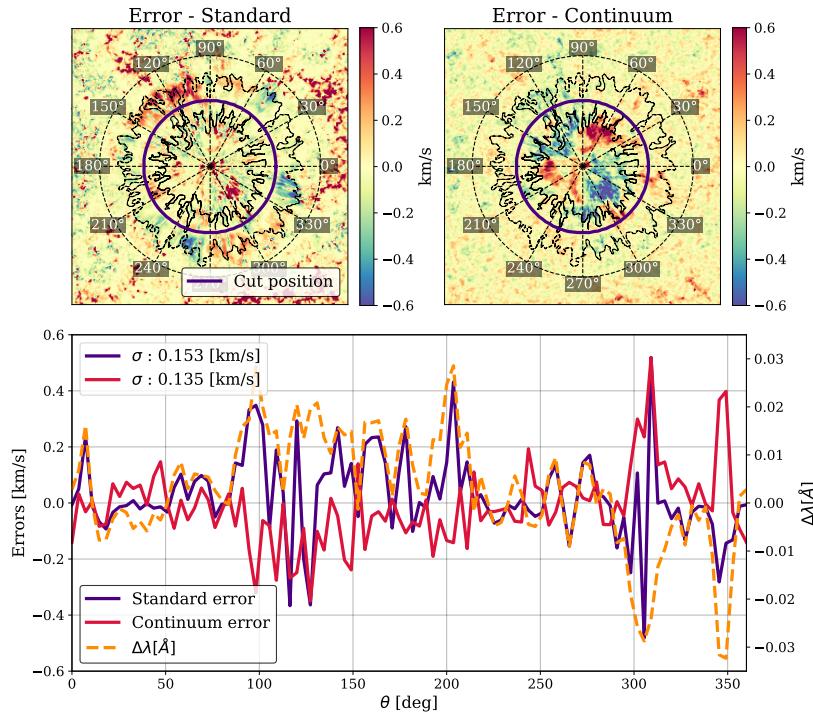


Figure 4.7 The top row shows the velocity error maps for both the standard and continuum approaches, along with the cut position and an overlay of the degrees for easier interpretation of the plotted profiles. The bottom panel shows the velocities along the cut for both approaches along with the value of the cavity map at the corresponding position.

However, the impact of the FPI's transmission profile remains to be addressed. It is common practice to neglect the asymmetric nature of real telecentric mounts in most data reduction pipelines, as the potential effects of profile asymmetries are typically assumed to be negligible.

The main effect of the assymetry is a systematic shift of the profile measurements towards the blue wing of the spectral profile, due to its higher transmissivity than its blue counterpart. This shift results in an average difference of around -0.2 km s^{-1} that can be seen over the whole FoV.

This systematic shift can be removed from the velocities computed with the assymmetric

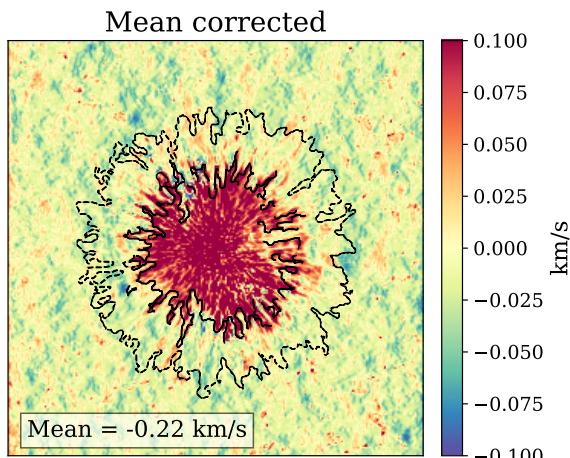


Figure 4.8 Difference in LOS velocities between the symmetric and assymmetric FPIs (symmetric - asymmetric). The mean value of the difference has been substracted for the whole FoV.

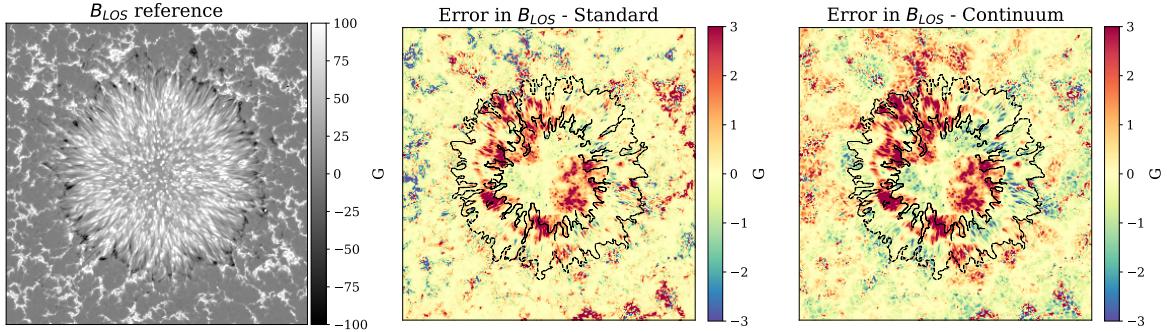


Figure 4.9 From left to right, B_{LOS} for the simulation with no cavity and a symmetric FPI, error in B_{LOS} computation for the standard approach with a symmetric FPI, and error in B_{LOS} computation for the continuum approach with a symmetric FPI.

FPI by subtracting the average difference between symmetric and asymmetric velocities. Figure 4.8 depicts the differences in the computed velocities for the symmetric and asymmetric FPIs, after applying this correction. Despite accounting for the systematic shift, discrepancies are still present, predominantly within the umbra, though they are observed across the entire FoV. The umbra is again more sensible than other regions, displaying the highest error in the whole map, mainly exhibiting a global shift showing smaller velocities for the asymmetric scenario. Outside the umbra, differences are smaller but still prevalent throughout the entire FoV. These differences can reach up to 100 m s^{-1} and typically exhibit a trend towards higher velocities in the case of the asymmetric scenario.

The fact that the umbra exhibits different behaviors compared to the rest of the FoV, along with the difference noted between the two types of FPIs, suggests that the amplitude and behavior of these effects are sensitive to the shape of the spectral profile and the FPI's transmission profile.

4.2.5 Magnetic field maps.

It is expected that cavity errors significantly impact velocity calculations, as these are derived solely from the analysis of spectral shifts. However, the effect of these errors on the calculation of the magnetic field is not as straightforward as its determination involves not only the spectroscopy, but also polarization.

To calculate B_{LOS} , we need to determine the circularly polarized component of the light. This requires demodulating the simulated observation using a process analogous to that described in the TuMag pipeline (ref ??). Once the stokes components have been determined, the LOS strength of the magnetic field is computed using the center-of-gravity method (eq. (1.24)).

Figure 4.9 shows the LOS strength of the magnetic field for the simulation with no cavity (the one employed as reference) and the errors in its calculation for both flat-fielding approaches. In the continuum approach, the structures observed in the error map closely

resemble those in the cavity map. This similarity is expected since no correction has been applied to account for the cavity error. While it is possible to compute the velocities associated with the spectral shift caused by the cavity map, this is not feasible for the magnetic field. Consequently, any effect of the cavity remains in the data when using the continuum approach.

As a result, errors in the magnetic field computation appear larger or more prevalent across the entire field of view (FoV) for the continuum approach. However, for both correction strategies, the errors remain relatively small, reaching values of approximately 3 G, compared to magnetic field signals of around ± 100 G. Although small, these residuals of the cavity map in the magnetic field maps are one of the main problems of employing the continuum approach in real data.

4.3 Fitting algorithm

Results from the previous section highlight the relevance of carefully addressing the data correction of FPI-based instruments. Although some pipelines address these deviations by correcting them only at first order through a regular flat-fielding procedure with no further computations, the residues left in the data can lead to errors as high as those shown in the previous section, depending on the FPI's properties. For this reason, different approaches have been taken to address these additional corrections. Typically, each instrument has a specific data reduction pipeline where the corrections are carried out taking into account the individual properties of the instruments, hence giving rise to different methods. An example of a more detailed calibration pipeline can be found in ?. By using a simplified analytical expression for the etalon transmission profile, Schnerr et al. were able to extract from the flat fields the contributions caused by the FPI cavity errors and variations in reflectivity for the CRISP instrument, which are then taken into account in the data calibration.

For the first order, this is a good approximation, but since asymmetries in the transmission profiles naturally arise in telecentric instruments (Bailén et al., 2019), deviations originating from them cannot be fully corrected. Thus, knowledge of the exact shape of the transmission profile is necessary to fully account for these deviations. ? already allowed asymmetries to be dealt with in the pipeline of the CRISP instrument, but the asymmetries originated from wavelength shifts induced by two distinct etalons, or from an angular dependence along the FoV introduced to approximate the behavior in telecentric setups.

We aim to further extend and generalize the strategy employed in ? and ? by including the exact shape of the transmission profile for the telecentric configuration in the analysis. By doing so, we can fit and take into account the presence of asymmetries in the profile that arise in these setups due to an asymmetric pupil apodization. We have developed a method for deriving the etalon properties from the data in such a way that prior knowledge of the distribution and magnitude of the defects is not needed. We do not make any distinction between defects associated with the mirror's flatness or separation, refraction index, or cavities. This way, we ensure the applicability of the algorithm to all types of FPIs (air and solid). Our technique also differentiates the effects associated with the etalon from other

corrections not related to it, such as pixel-to-pixel variations of the gain across the detector. By comparing the theoretical prediction, obtained through the analytical expressions of the transmission profile, with the measured data, we can disentangle the etalon properties from the flat-field observations.

In this section, we introduce the algorithm and evaluate its performance. We begin by outlining the approximations made and detailing the algorithm in Sections 4.3.1 and 4.3.2, respectively. The method was tested on a series of simulated observations with controlled configurations to evaluate the impact of different observational properties on the accuracy of the results. Finally, the algorithm was applied to a dataset from the HRT telescope of the SO/PHI instrument to assess its performance in real scenarios. The results for both the simulated scenarios and real data are discussed in Section 4.3.3.

4.3.1 Initial approximations and observations simulation.

As initial approximations, we will adopt the same assumption regarding the prefilter as the one employed for the sunspot simulation; specifically, that it has a rectangular shape with a width such that only one order of the etalon is let through. Additionally, we will disregard the spatial PSF of the FPI and assume that the spatial dependence can be represented by a Dirac delta function to simplify the equations. Therefore, if we assume that the image response of the FPI follows the expression:

$$S(\xi_0, \eta_0; \xi, \eta; \lambda - \lambda_s) = \delta(\xi_0 - \xi, \eta_0 - \eta) \Psi(\xi, \eta; \lambda - \lambda_s), \quad (4.11)$$

equation (4.9) simplifies into:

$$I(\xi, \eta; \lambda_s) = g(\xi, \eta) \int_{\lambda_0 - \Delta\lambda}^{\lambda_0 + \Delta\lambda} O(\xi, \eta; \lambda) \Psi(\xi, \eta; \lambda - \lambda_s) d\lambda. \quad (4.12)$$

The explicit shape of Ψ varies depending on the optical configuration of the instrument, whether collimated or telecentric. In particular, we will employ three types of transmission profiles: one for the collimated configuration, one for a perfect telecentric setup (normal incidence), and one for an imperfect telecentric setup. The first two have analytical expressions for the transmission profile in the absence of the spatial PSF, as detailed in Sections 4.1.1 and 4.1.2, respectively. For the imperfect telecentric configuration, the transmission profile must be determined numerically (see Section 4.1.2.1).

We have tested the performance of the algorithm on a series of simulations of a spectral line observation in different conditions. We used the Kitt Peak FTS-Spectral-Atlas as the reference (?) and, specifically, the Fe I spectral line at 6173.3 Å. Each observation was composed of N_λ wavelengths, where the measured intensity was recorded. At every wavelength λ_s , the corresponding transmission profile of the etalon Ψ^{λ_s} was computed, and the "observed" intensity $I_{\text{obs}, i}^{\lambda_s}$ corresponding to a specific spatial location (ξ, η) , represented hereinafter by the pixel i , was calculated using Eq. (4.12). Additionally, we took into account the presence of additive Gaussian noise. This noise does not necessarily respond

to any parameter fluctuation within our analytical expressions or photon noise but comes from any unexpected variations that may not have been modeled in the theoretical scheme.

Additionally, we included the presence of defects arising from irregularities or inhomogeneities on either the cavity thickness d , the refractive index n , or from deviations of the angle of incidence θ . In order to simulate this, we introduced a relative perturbation Δa into the etalon equation that accounts for any local deviation of the value of a with respect to its nominal value. This parameter changes from pixel to pixel differently for the collimated and telecentric configurations. In the former, the profile shifts across the FoV only because of the different incidence angles of the light beam on the etalon. In the latter, local variations of n and/or d are mapped directly onto the detector. We also note that variations in the incidence angle must be considered as well when the degree of telecentrism varies along the detector. Analytically, the parameter a at each i -th pixel is given by $a'_i = a\Delta a_i$, where $a = (2\pi/\lambda)nd \cos \theta$ is constant along the FoV. Note that rewriting the equations for the transmission profiles (sections 4.1.1 and 4.1.2) using this definition of a is straightforward. In collimated configurations, the parameter δ (eq. (4.4)) is simply $2a$. In perfect telecentric configurations, θ is always 0, so the given transmission profile is already expressed in terms of a .

We let $n_i^{\lambda_s}$ be the noise contribution at the i -th pixel and wavelength λ_s . Thus, the observed intensity at that pixel when the etalon is tuned at λ_s , $I_{\text{obs},i}^{\lambda_s}$ is given by

$$I_{\text{obs},i}^{\lambda_s} = g_i \frac{\int_{\lambda_0-\Delta\lambda}^{\lambda_0+\Delta\lambda} O(\lambda)\Psi^{\lambda_s}(\lambda, \Delta a_i)d\lambda}{\int_{\lambda_0-\Delta\lambda}^{\lambda_0+\Delta\lambda} O(\lambda)\Psi^{\lambda_c}(\lambda, \Delta a_i)d\lambda} + n_i^{\lambda_s}, \quad (4.13)$$

with λ_c being the continuum wavelength. From a practical point of view, the integration limits are set in such a way that only a single resonance (or order) of the etalon is included with the limits, thus, acting akin to the sorting pre-filter commented on previously. We note that the denominator strictly corresponds to the intensity at the continuum of the line in the absence of the transmission profile or if the continuum wavelength is far enough from the spectral line. In any other case, the transmission should be taken into account as well to normalize the observations to the local continuum, which is necessary since we work with relative measurements. An example of a spectral line measurement is displayed in Fig. 4.10.

For both the collimated and telecentric configurations, we modeled etalon and gain imperfections over a $100 \times 100 \text{ px}^2$ image. Pixel-to-pixel variations in the sensor efficiency were modeled following a random spatial distribution, as shown in Fig. 4.11 (top panel). Additionally, we included a set of pixels with very low gain values, which represent a group of dead pixels or dust grains.

We modeled the etalon defects as changes in Δa in such a way that the maximum displacement reaches 3 pm. The spatial distribution of the values of Δa follows an increasing radial distribution, as shown in Fig. 4.11 (bottom panel). Such a spatial distribution coincides with the expected one in collimated etalons due to the change in the incidence angle across the FoV. Telecentric mounts do not exhibit a spatial distribution of their defects such as this, but using the same spatial distribution in the two cases allowed us to compare the

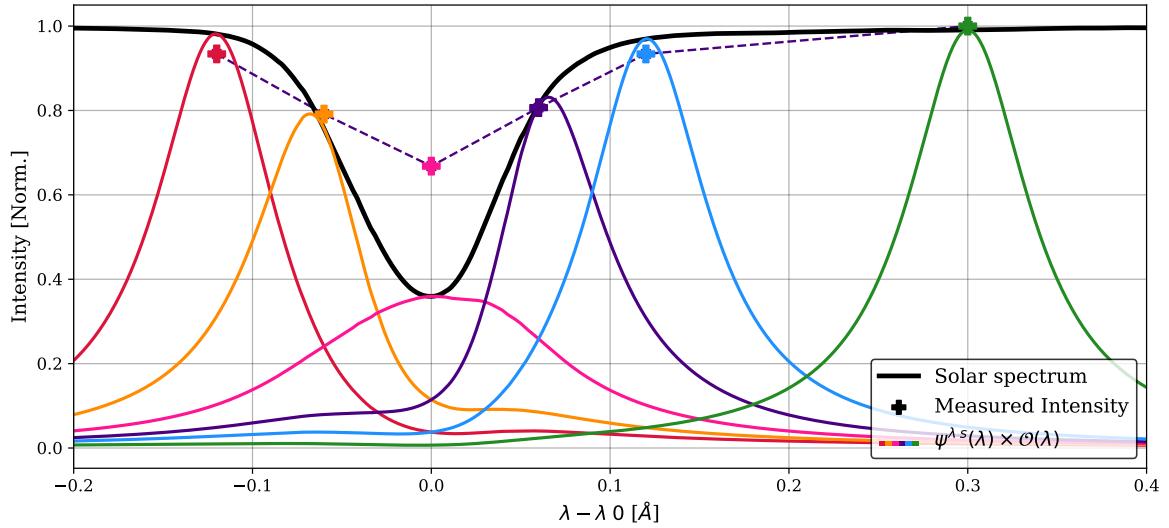


Figure 4.10 Simulated observation of the Fe I spectral line ($\lambda_0 = 6173.3\text{\AA}$) using a collimated mount and a total of $N_\lambda = 6$ wavelengths that have been equally distributed along the spectral line, with the exception of the continuum measurement (light blue), which is selected at 300 m\AA{} from the blue of the line core. The measured intensity is the result of computing the value given by Eq. 4.13 at each wavelength and with $g = 1$.

performance of the method for both setups in a systematic way. Since Δa accounts for relative perturbations, it is by definition an adimensional parameter. However, to grant it a physical meaning, we express the values of Δa in \AA{}, representing the associated shift of the transmission profile with respect to the original position determined by a .

4.3.2 Fitting algorithm

We have developed an algorithm able to extract the distribution of the etalon defects and the gain map from data taken by etalon-based instruments, which enables the correction of the two contributions separately. The algorithm works by minimizing a given merit function that depends on the gain and the etalon defects.

In particular, we have defined an error metric, ε^λ , at each tuned wavelength, computed by comparing the measured intensity with the theoretical prediction. If we let $I_{i,\text{obs}}^{\lambda_s}$ be the measured intensity at a given i pixel for an etalon tuned to the wavelength λ_s , the error metric at each wavelength is given by

$$\varepsilon^{\lambda_s}(\Delta a_i, g_i) = I_{i,\text{obs}}^{\lambda_s} - g_i \frac{\int_{\lambda_p}^{\lambda_q} O(\lambda) \Psi^{\lambda_s}(\lambda, \Delta a_i) d\lambda}{\int_{\lambda_p}^{\lambda_q} O(\lambda) \Psi^{\lambda_c}(\lambda, \Delta a_i) d\lambda}. \quad (4.14)$$

The merit function we employed is then the quadratic summation of the error metric over all tuned wavelengths:

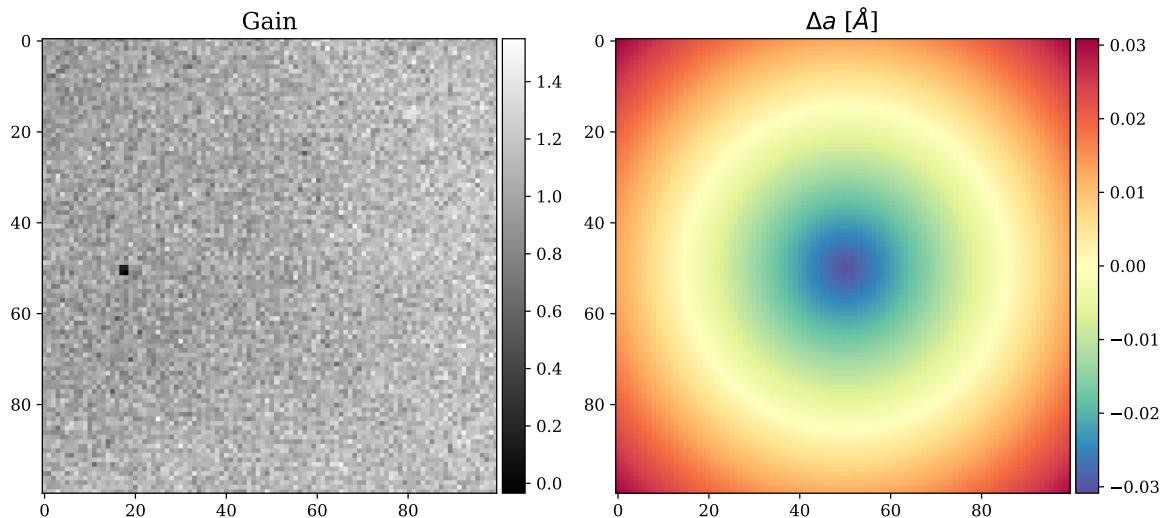


Figure 4.11 Input maps introduced when simulating the observations. The left panel represents the gain generated as white Gaussian noise, with values ranging from 0.8 to 1.2. A dust speck was introduced by creating a group of four pixels with low values of $g = 0.2$ for the gain. The right panel shows the spatial distribution of the defects in the etalon. The distribution follows a radial pattern starting from the center of the FoV. The defects vary from 0 % deviation to up to 5×10^{-4} %, which corresponds to a shift of 3 pm. Both possible directions for the deviations have been considered. The sign of the deviation is negative at the very center, which introduces a redshift, while it is positive at the corners, causing a shift of the profile into the blue.

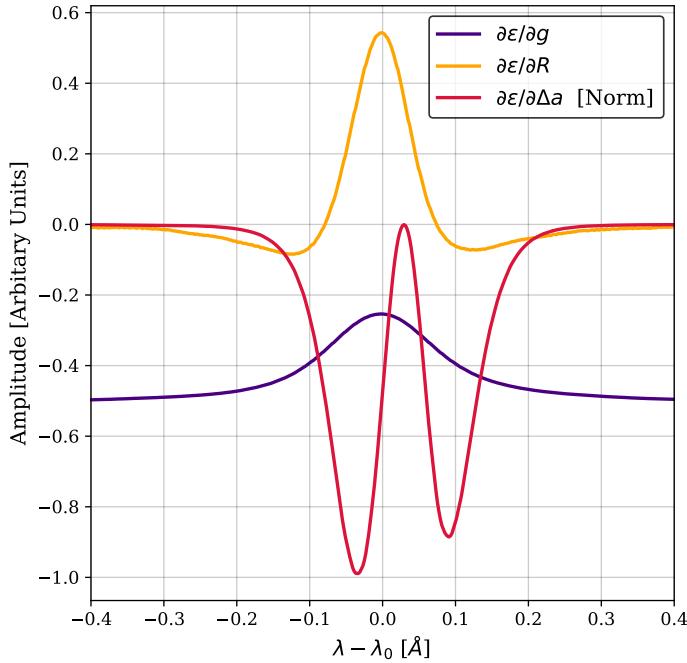


Figure 4.12 Derivatives of the error metric as a function of wavelength. The derivative with respect to Δa has been normalized in order to fit the three curves in the same plot.

$$f(\Delta a_i, g_i) = \sum_{s=0}^{N_\lambda} \left(\varepsilon^{\lambda_s}(\Delta a_i, g_i) \right)^2. \quad (4.15)$$

Both the camera gain and the defects of the etalon change from one pixel to another, which is why we address each pixel independently, but they remain constant at every wavelength. Hence, the transmission profile of the etalon varies at different points of the FoV; but at a given pixel, it is constant at all tuned wavelengths. Therefore, the algorithm is able to better obtain the etalon properties as we increase the number of wavelengths.

Figure 4.12 shows the derivatives of the error metric, Eq. (4.14), as a function of wavelength, that is, before computing the summation over s of the merit function, with respect to the gain, the reflectivity, and Δa . The curve corresponding to the Δa derivative is different from the others, whereas the derivatives of both the gain and the reflectivity exhibit similar shapes. Hence, variations in either the reflectivity or the gain introduce similar changes in the merit function, which can produce a trade-off between these two parameters, especially when the spectral line is sampled in only a few points. Given that discrepancies arising from errors in reflectivity are assimilated by gain maps, we did not take into account reflectivity errors when computing our simulations, as they have no impact on cavity map calculations.

A few key aspects arise when analyzing the merit function and its applicability on real data. The first point to bear in mind is that the shape of the object, $O(\lambda)$, is not known a priori. Therefore, we needed to provide a guess for it. The method works by assuming that differences between the prediction and the observation are caused exclusively by the etalon defects or the gain. If the object used during the fitting process differs considerably from the real one, the prediction and observation will have differences that will erroneously

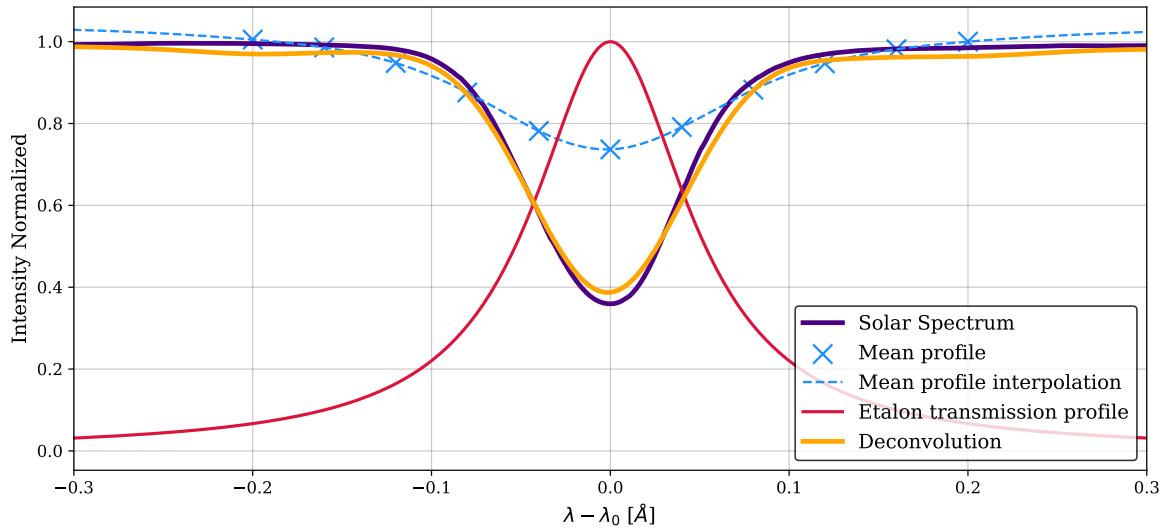


Figure 4.13 Deconvolution of the object with a measurement of the Fe I spectral line using $N_\lambda = 9$. All points of the FoV have been used to compute the average profiles (blue crosses). The deconvolution (orange) is the result of deconvolving the mean profile interpolation (dashed line) with the displayed etalon transmission profile (red).

be identified as etalon defects or gain variations. This is the main source of errors for the method when applied to real data. Two approaches can be followed in order to address this issue. The first one consists of assuming the solar atlas profile as the object. This is a good approximation, provided the data to which the algorithm is applied to lack information about solar structure, either because they are observations of long integration times of the quiet sun or produced by averaging several quiet sun observations (flat fields). If this condition cannot be met, this approach is not valid. The second approach involves deriving an approximated object from the data themselves by deconvolving the mean profile of the observation with the etalon's transmission profile. This approach can account for any difference the real object may have with the solar atlas and thus has a greater resemblance to the real object. Nevertheless, the process of deconvolving is prone to errors when the sampling is insufficient and can introduce additional noise into the problem. We have tested both approaches to compare their performances on different scenarios in order to assess when to use one or the other.

We employed Newton's method to minimize the merit function, Eq. (4.15), as it has been proven to quickly converge (in five iterations or fewer, usually). The method begins by assuming an initial guess for the gain g_j and Δa_j parameters. Then, provided the initial guess is sufficiently close to the solution and that the merit function is continuous and differentiable, the gain g_j and Δa_j encoded in the vector, \mathbf{x}_j , can be updated iteratively at each iteration, j , as

$$\mathbf{x}_{j+1} = \mathbf{x}_j - \mathcal{H}^{-1} \mathcal{J}^T f(\mathbf{x}_j) , \quad (4.16)$$

where \mathcal{H} and \mathcal{J} are the Hessian and Jacobian matrices of the merit function f , respec-

tively, calculated for $\mathbf{x}_j = [g_j, \Delta g_j]^T$, and T stands for the transpose. Hence, the transmission profile of the etalon and its derivatives have to be computed for every wavelength and every pixel at each iteration. This can be computationally costly, especially when using imperfect telecentric configurations, where numerical integrals are involved. All derivatives needed for the algorithm are calculated analytically, except when simulating imperfect telecentrism. A detailed formulation of these derivatives is provided in the appendix.

Regarding the object $O(\lambda)$, if we assume it is given by the solar atlas, no additional computations are needed. However, when using the deconvolution approach, the object has to be calculated in each iteration. In this case, the algorithm works as follows: First, we compute the average profile across the whole FoV, and we force the continuum intensity to be the same on both sides of the spectral range to reduce the boundary effects of the deconvolution. This step is only necessary in case the spectral line is sampled in only a few positions, as is the case of the SO/PHI, IMaX, or TuMag instruments, where only a continuum point, either at the red or the blue side of the spectrum, is recorded. Both the object and transmission profile require a good spectral sampling to accurately compute the integrals of Eq. (4.14). Second, a cubic spline interpolation is applied to the generated average profile to artificially improve the spectral sampling, if necessary. Finally, the interpolated profile is then deconvolved by means of a Wiener filter with the etalon's transmission profile. The result of this deconvolution is the object, $O(\lambda)$, used in the minimization algorithm. The deconvolution of the object is done every time the etalon defects are updated in order to improve the resemblance of the deconvolved object to the real one. Figure 4.13 shows an example of this process in a simulated observation using nine scanned points and a collimated configuration. The deconvolution manages to reproduce the original signal, with only some minor differences in the line core and the beginning of the wings.

4.3.3 Test scenarios and results

The aim of the simulations carried out in this section was to characterize the role of the noise $\delta_i^{\lambda_s}$, the spectral sampling, the selection of the object $O(\lambda)$, and the accuracy of the method for both the collimated and telecentric configurations. All simulations were run for different choices of the number of scanned wavelengths, ranging from $N_\lambda = 5$ to $N_\lambda = 21$.

4.3.3.1 Impact of the noise level

We first assumed that the spectrum of the observed object is given by the solar atlas. This way, all errors in the derivation of the gain and etalon defects only come from the noise introduced into the measurement. We refer to this as the "ideal case." Since we were combining different measurements taken at different wavelengths, we considered a worst-case scenario and simulated three different signal-to-noise ratios: 100, 150, and 200.

We restricted imperfections in the telecentrism to arise only for one scenario, $S/N = 200$, since simulating imperfections requires a high computational effort due to the lack of a theoretical expression for both the transmission profile and its derivatives. We also assumed that the degree of telecentrism (0.3°) is known in this case.

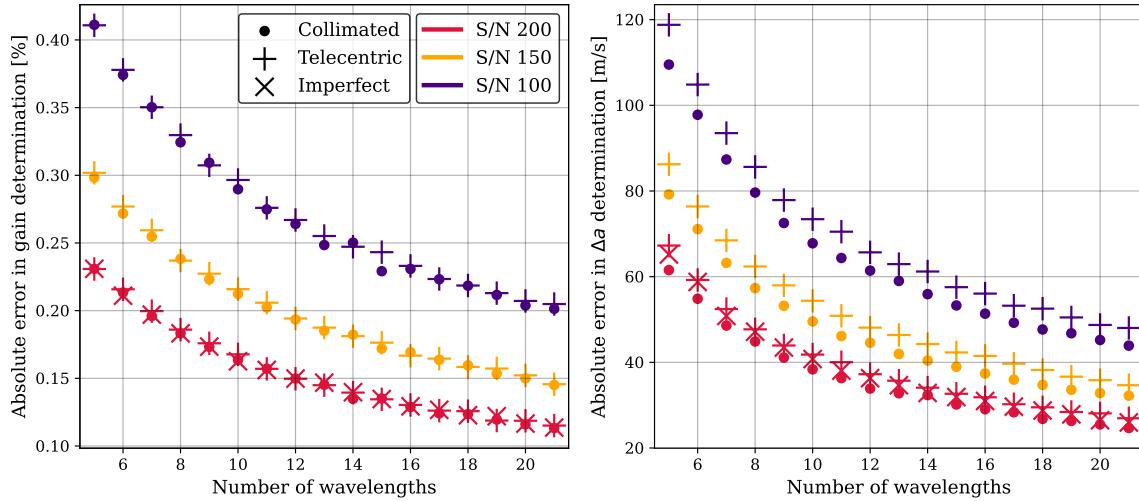


Figure 4.14 Absolute errors of the gain (left) and etalon defect (right) derivations averaged over all the FoV. The number of wavelengths corresponds to the parameter N_λ of wavelengths used to scan the profile.

Figure 4.14 shows the average absolute error in g (left panel) and in Δa (right panel) over the whole FoV as a function of the wavelength sampling, N_λ . The error in g is expressed as a percentage of its real value. Errors in Δa are given in meters per second since they are mostly responsible for shifting the profile. Errors in Δa can be translated into velocity errors by computing the doppler velocity (Eq. (1.23)) associated to the spectral shift of the transmission ppeak produce by the error in Δa . Figure 4.14 shows the average absolute error in g (left panel) and in Δa (right panel) over the whole FoV as a function of the wavelength sampling, N_λ . The error in g is expressed as a percentage of its real value. Errors in Δa are given in meters per second since they are mostly responsible for shifting the profile. Errors in Δa can be translated into velocity errors by computing the doppler velocity (Eq. (1.23)) associated to the spectral shift of the transmission ppeak produce by the error in Δa .

All the scenarios exhibit a similar behavior as far as their dependence on the spectral sampling is concerned, namely, the absolute errors decrease monotonically when the wavelength sampling increases. The reason for this is simply that a larger number of wavelength samples increases the amount of available information that the algorithm can use, thus making the fitting for g and Δa more precise. These results highlight the importance of properly sampling the targeted spectral line. A modest sampling of only $N_\lambda = 5$ can produce errors as large as 120 ms^{-1} in the worst-case scenario ($S/N = 100$).

The noise level also plays an important role in the accuracy of the results. Scenarios with a lower S/N always have larger errors, for a given N_λ , in both the gain and Δa computations. The difference in the performance of the algorithm due to the noise also changes with the spectral sampling; scenarios with a poor spectral sampling suffer from larger differences in the accuracy between the different S/N (50 ms^{-1} for $N_\lambda = 5$ between S/N = 200 and S/N =

100) than those with higher samplings (35 ms^{-1} for $N_\lambda = 21$).

The optical configuration of the etalon has a very small impact on the accuracy of the algorithm. Results for the three setups are very similar, particularly in the gain calculation, for which the results are almost identical for all configurations. Retrieval of Δa is slightly better for the collimated mount, though.

Figure 4.15 shows the spatial distribution of the errors in the retrieval of Δa for different choices of N_λ . There are no signs of a radial distribution in the maps shown in the figure, contrary to the actual distribution of the Δa parameter, as shown in Fig. 4.11, bottom panel. This means that the precision of the method is similar no matter the amplitude of the defects, that is, we achieve the same accuracy in the retrieval of defects associated with shifts of 3 pm ($\sim 1450 \text{ ms}^{-1}$, near the corners of our FoV), which correspond to cavity errors of around 1.5 nm or incidence angles of approximately 0.4 degrees, and in the retrieval of regions where no defect is present (radius of 20 pixels from the center of the FoV approximately). Instead of a radial distribution, the errors follow a Gaussian-like distribution (shown at the bottom panels in 4.15) similar to the one followed by the noise contribution.

The standard deviation of the errors for both the gain and Δa computations are also reduced with an increase in spectral sampling. The last row of Fig. 4.15 displays the error distributions in the calculation of Δa for the three optical configurations and different spectral samplings. These results illustrate how the three configurations yield practically identical results and how the distribution narrows as N_λ increases, thereby improving the results. Specifically, the standard deviation decreases from 50 ms^{-1} for $N_\lambda = 5$ to 20 ms^{-1} for $N_\lambda = 21$. In the case of the gain determination, the standard deviation ranges between 0.2 % and 0.1% for the scenarios with the poorest and highest spectral sampling, respectively.

4.3.3.2 Impact of the object approximation

To infer the error of the algorithm when the object is unknown, we compared the performance of the ideal case, that is, when the object used to generate the observations is known, with the one achieved when deconvolving the object from the data. Only the collimated setup was simulated in order to focus exclusively on the errors introduced by the deconvolution. The data has been degraded by Gaussian noise with an S/N = 200 in both scenarios.

Figure 4.16 shows the results for the two approaches. Interestingly, the error in the gain for the deconvolution approach does not decrease with a larger number of wavelengths, unlike the ideal case. Nevertheless, the average error of the calculation is below 0.4 %, with a dispersion (1σ) of $\pm 0.3\%$. The deconvolution approach is prone to higher errors when deriving the gain due to the normalization of the profiles. The reason for this is two-fold. first, if the continuum is far enough from the spectral line, the normalization is strictly the integral over the transmission profile because the object is flat along the integration interval. However, this is not strictly true since the wings of the transmission profile can reach the spectral line (see Fig. 4.10), hence modifying the normalization of the profile when the object changes at each iteration. Second, should the continuum intensity of the derived object vary with respect to its real value due to the deconvolution process (e.g., due

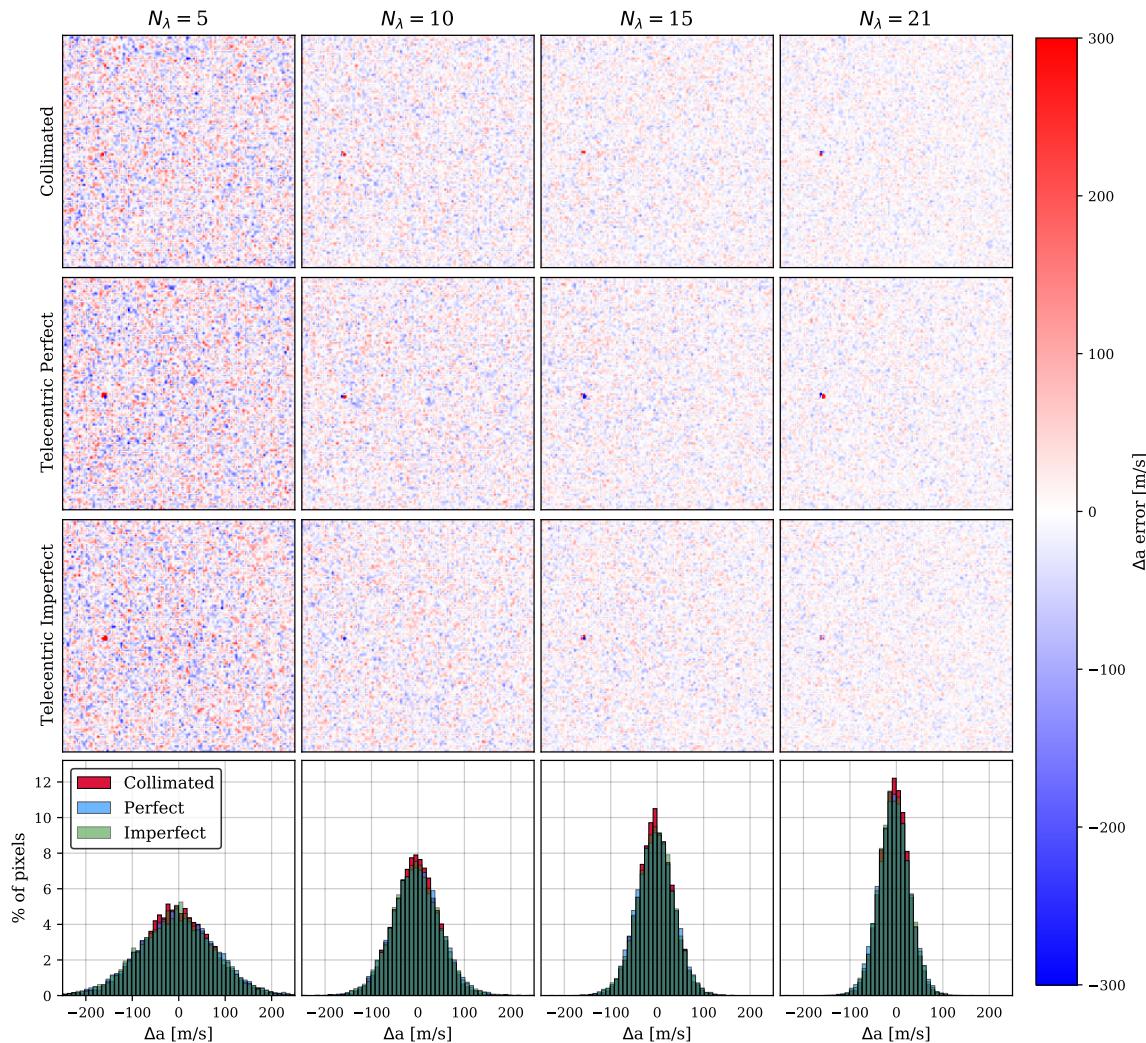


Figure 4.15 Distribution of the errors in the Δa computation for the three configurations (first three rows) and different spectral samplings (columns). In the bottom panels of each column, the error distribution for the corresponding spectral sampling is shown for the three configurations.

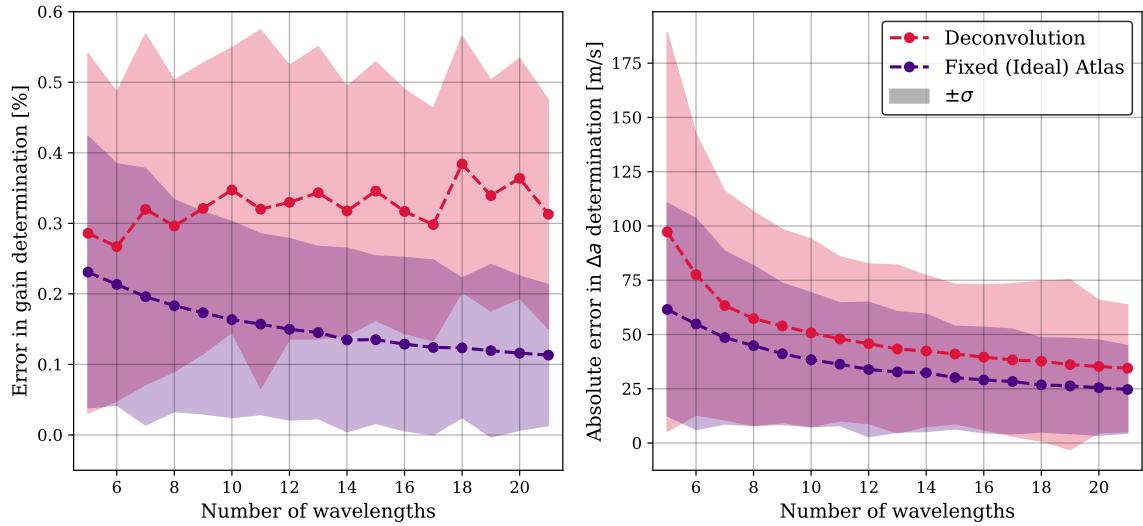


Figure 4.16 Errors in gain determination and etalon properties averaged over all the FoV with a signal-to-noise ratio of 200 and a collimated configuration.

to boundary effects), there will be a shift in the intensity of the whole profile induced by the normalization process. These two effects seem to dominate the accuracy on the gain determination, regardless of the chosen sampling.

For Δa , the performance of the method is slightly worse than for the ideal case when using the deconvolution approach. Unlike the gain determination, errors in the Δa derivation show a strong dependence on the spectral sampling. Differences between both approaches range from 10 ms^{-1} to 40 ms^{-1} and increase with decreasing N_λ . The sensitivity with N_λ is especially high up to $N_\lambda = 8$. A modest increase of N_λ from five to six improves the determination of $\Delta a \sim 20 \text{ ms}^{-1}$, whereas at better spectral samplings, the difference between each simulation decreases more slowly, without any relevant improvement as the sampling increases. In any case, differences are all well within $\pm 1\sigma$.

4.3.3.3 The crossover case

The fact that the sensitivity of the model to the gain and to the Δa parameters are different guarantees (to some extent) that the parameters can be separated from each other. The treatment of the problem is very different between etalon configurations, and therefore full knowledge of the setup is critical. However, this is not always feasible due to the unavoidable presence of errors, misalignment, and imperfections on the instrument. Approximations to describe the optical setup are also common in the pipeline of an FPI instrument because they reduce computational efforts. For instance, telecentric mounts are usually simplified as collimated setups, as the f-numbers employed in solar instruments are usually very large. Imperfections of telecentrism are commonly neglected, too. In this section, we analyze the impact of assuming an incorrect etalon mounting. To do so, we repeated the previous exercise, starting from a perfect and imperfect telecentric configuration but

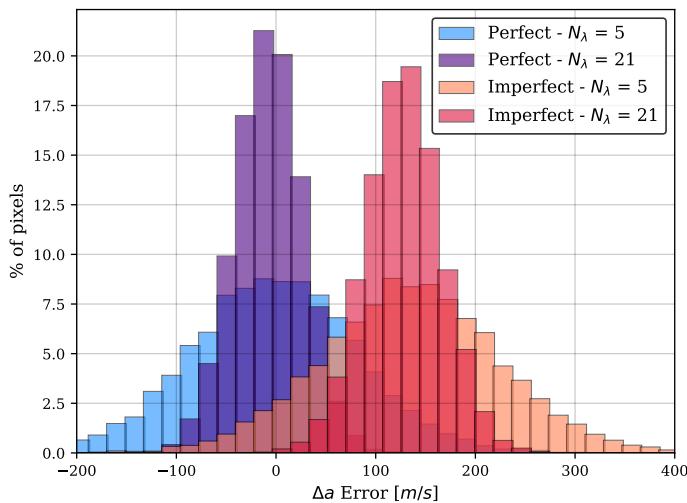


Figure 4.17 Distribution of the errors in the determination of Δa for the crossover scenarios (different configuration in the observation generation and minimization algorithm) for both perfect and imperfect configurations. Only results for the two extreme spectral samplings ($N_\lambda = 5$ and $N_\lambda = 21$) are shown.

assuming that the transmission profile shape corresponds to a collimated one.

In this exercise, we assumed that we have an instrument with an FPI in a telecentric mount, as in the previous sections, in both perfect and imperfect configurations and an S/N = 200. We also considered that the object is given by the spectral solar atlas. The shift of the perfect telecentric transmission profile with respect to the collimated one was corrected using Eq.52 from Bailén et al. (2019) to avoid the emergence of spurious velocity signals. Imperfections in the telecentrism shift the profile more. This additional displacement was left uncorrected intentionally so we could study its effects.

Figure 4.17 shows the error distributions for Δa when the model assumes a collimated configuration for $N_\lambda = 5$ and $N_\lambda = 21$ and for both perfect and imperfect configurations. The amplitude and dispersion of the error distributions are very similar for the two mounts and are comparable to the results obtained in the ideal case (Fig. 4.15, bottom panel). The main difference between the perfect and imperfect scenarios is a shift of 130 ms^{-1} for the reason mentioned above. We note that this shift can easily be accounted for since it is a known and measurable effect.

The similarity in the error distributions for the calculation of Δa in the two scenarios demonstrates that the error incurred when assuming a collimated etalon does not significantly impact the determination of the cavity maps of the etalon. This is because changes in Δa mostly induce a shift of the transmission peak by an equal amount in both cases.

We note, however, that the amplitude, width, and shape of the transmission profile differ significantly between the telecentric and collimated configurations, leading to an expected higher error in gain calculation. Figure 4.18 shows the absolute errors in gain and Δa calculations for both crossover scenarios and the ideal case after correcting the wavelength shift between the different mounts. For Δa , the performance of the method is very similar in the three setups, as also observed earlier. This behavior is nevertheless anticipated since the properties selected for simulating the imperfect etalon were chosen to mirror those of the SO-PHI etalon, which were adjusted to closely resemble the behavior of a collimated etalon to the greatest extent possible.

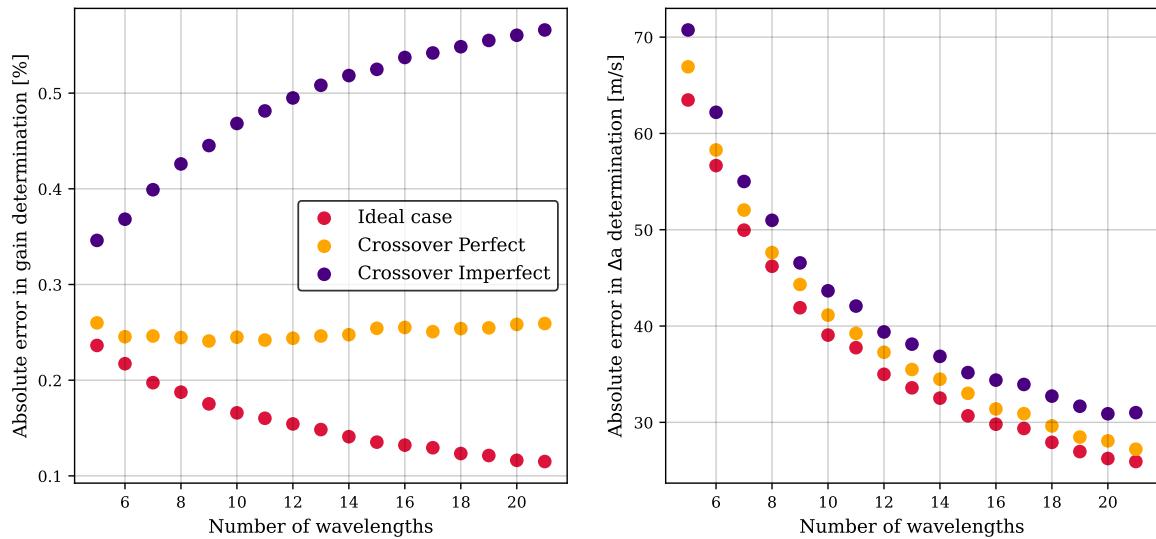


Figure 4.18 Average errors of the gain (left) and etalon defect (right) calculations over all the FoV for the two crossover cases and the standard case (also shown in Fig. 4.14 as the collimated case with S/N = 200) for reference. All Δa errors have been computed by correcting differential offsets of the transmission profile between the different mounts.

The differences are larger for the gain determination. Not only are the errors higher in the crossover cases, but the trend is entirely different. Instead of decreasing when increasing the number of wavelengths, gain errors remain the same for the perfect case and increase with the number of samples for the imperfect case. Similar to the deconvolution case (Fig. 4.16), the difference between transmission profiles introduces an error in the normalization process that systematically affects the rest of the measurements. This effect becomes more pronounced as the number of wavelengths increases, given that this error is introduced more frequently, and it is even more prominent in the imperfect case, as not only are the profiles different in this scenario, but they are also asymmetric. This asymmetry results in an imbalance in the measurement of the profile, as one wing of the spectral line has a higher transmissivity and is observed with greater intensity than the other.

Our results suggest that assuming an etalon in a collimated configuration for instruments with telecentric mounts can be a good first-order approximation for cavity map calculations, provided that the level of asymmetry of the transmission profile is known. However, achieving an accurate knowledge of the degree of telecentrism is often challenging in real instruments, as it usually varies across the FoV. Meanwhile, the results highlight that this approximation leads to a considerable increase in the error in the gain determination, which increases when increasing the spectral sampling. This contrasts with the standard philosophy of solar instrumentation, which requires a high number of points to better scan the spectral line.

4.3.3.4 Real data

While simulated tests are crucial for understanding the algorithm's behavior as a function of the different parameters of the problem, tests with real data are required to validate the effectiveness of the method. In this section, we evaluate the results obtained by the algorithm when implemented on observations acquired using the High-Resolution Telescope (HRT) of the SO/PHI instrument.

The observation we used corresponds to a flat-field observation employing six points to scan the 6173 Å line (five along the spectral line and an additional continuum measurement) conducted on March 9, 2022. The FPI aboard SO/PHI is illuminated in a telecentric mount, with an expected degree of telecentrism of about 0.3°. The tests conducted with this dataset have the same procedure as the ones employed for simulated data, that is, we only fit the gain and $\Delta\alpha$ parameters, and we employed the deconvolution approach (Sect. 4.3.3.2).

Figure 4.19 shows the results obtained when applying the algorithm to HRT-SO/PHI data. The top row shows, from left to right, the observed flat at the continuum, the cavity map measured under laboratory conditions (on ground), and a selected region of the cavity map. In the bottom row, the same cases are shown but for the results inferred by the algorithm. That is, the flat-field (i.e., the gain) and the cavity map (i.e., $\Delta\alpha$).

The results obtained for the cavity map closely resemble the laboratory measurements. The overall structure is preserved, maintaining the gradient from the upper left to the lower right. Similar structures are also discernible, such as the vertical line in the lower-right corner and the ring-like patterns evident in the lower section of the image. However, closer scrutiny of the observed structures revealed some differences. This deviation between results and measurements is expected, considering that the experiment conducted here represents a first-order approximation to assess the algorithm's capability to generate coherent results. Several factors contribute to these differences. Firstly, the spatial resolution of the cavity measured on ground is lower. Secondly, the degree of telecentrism has a predicted variation of 0.23° across the FoV. Thirdly, we expected an error for the six-point scan, employing the deconvolution strategy, as large as a hundred meters per second.

Concerning the results for the gain determination, the comparison depicted in Fig. 4.19 suggests that pixel-to-pixel gain variations can also be determined accurately with our method, as both small-scale and large-scale variations are reproduced with detail.

A more quantitative analysis of the results is depicted in Fig. 4.20. The upper panel shows a comparison between the fitted and observed profiles for three distinct cases with varying degrees of precision, while the bottom panel shows the distribution of χ^2 for all pixels. In the worst-case scenario, small differences between the real profile and the fit can be seen in both the line wing and the line core. In any other case, the fittings are rather satisfactory.

4.3.4 Algorithm applied at the Sunspot

No estoy seguro de añadir esto. En principio sale algo mejor pero los resultados no son muy distintos y tampoco quería alargar demasiado el capítulo.

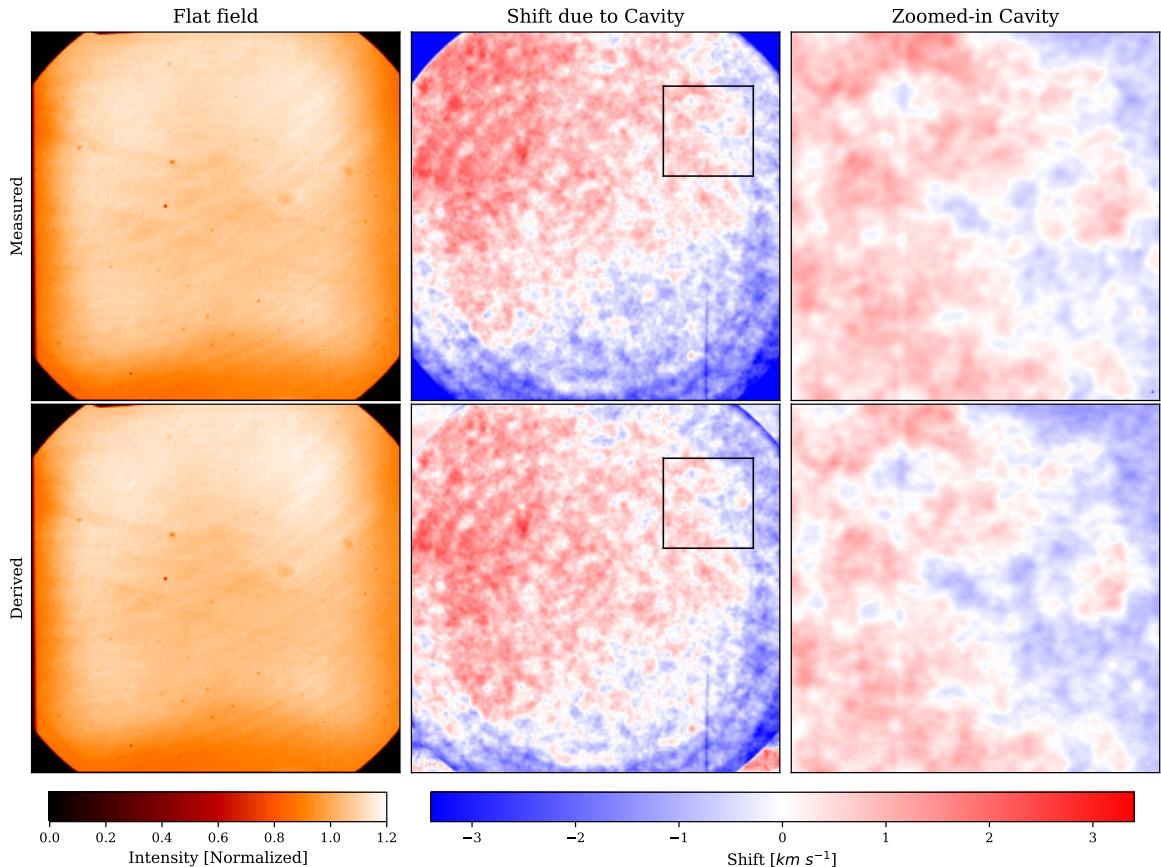


Figure 4.19 Comparison between the observed flat-field and cavity map (top row, left and center columns, respectively) and the derived flat-field (gain) and cavity map (Δa ; bottom row, left and center columns, respectively). The right column in both rows showcases a zoomed-in region of each cavity map. The area corresponding to the zoomed-in region is indicated with a black square in the full cavity map to its left. The properties used to simulate SO/PHI's etalon are $R = 0.925$, $n = 2.29$, $d = 251.63 \mu m$, $f\# = 60$, $\Theta = 0.23^\circ$.

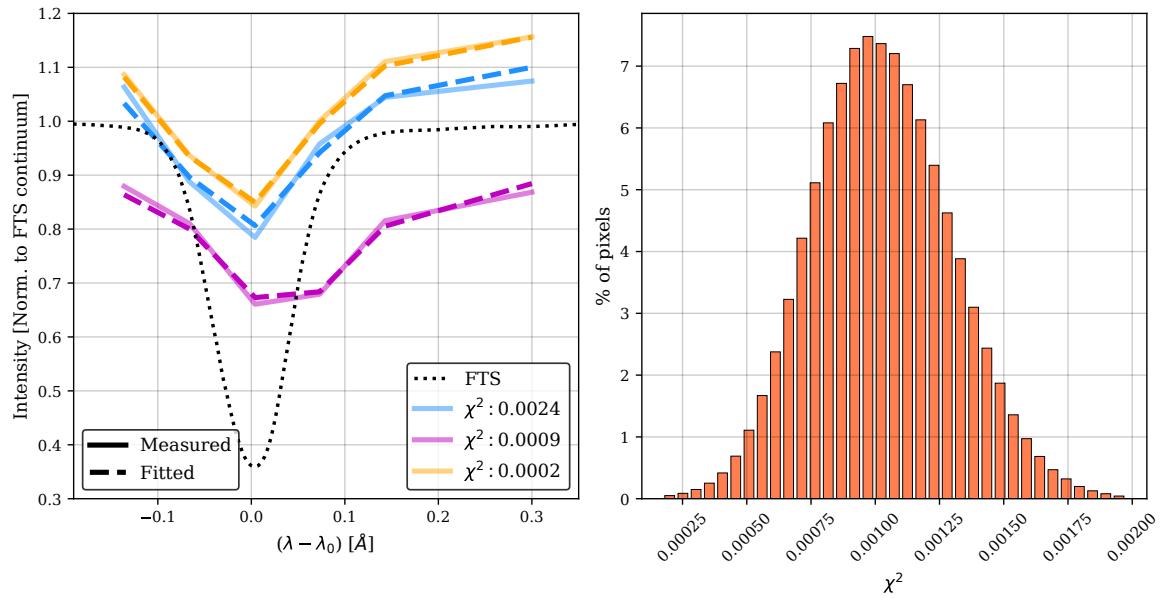


Figure 4.20 Comparison between the measured (solid lines) and fitted profiles (dashed lines) for three pixels, each representing varying degrees of accuracy (top panel). The FTS is shown as a reference. The bottom panel displays the distribution of χ^2 values for all pixels. Among the three selected cases, one demonstrates an average fit (depicted in pink, with χ^2 close to the mean value), while the other two correspond with extreme cases—one with a notably good fit and the other with a poor fit (depicted in yellow and blue, respectively). The value for χ^2 has been computed employing equation (4.15)

4.4 Conclusions.

In this chapter we have highlighted the relevance of carefully addressing the data reduction of etalon-based instruments, specifically those with the FPI in telecentric setups. Equipped with the analytical modelling of the FPIs, we have been able to address the impact of the etalon defects on the observations. We have quantified the error committed in the computations of LOS velocities and magnetic fields when the cavity errors are (only) partially corrected through the simulation of an observation of a Sunspot. Additionally, we have provided a new strategy for the data correction of etalon-based spectrometers that separates the effects of the FPIs on the flat-fields from other effects in an attempt to improve the flat-field correction. Lastly, we have tested the performance of this new method on a series of simulated observations and real data to verify its applicability.

The simulation of the sunspot observation in section 4.2 demonstrates the significant impact of improperly corrected cavity errors on the physical measurements. The discussion is centered around the flat-fielding procedure, which presents one of the main challenges in these setups, as cavity errors also affect flat-field observations. We show that these flat-fields alter and distort the observed spectral profiles when applied. The correlation between the structures found in the velocity calculation errors after flat-field correction and those of the cavity map suggests that the flat-fielding has not (fully) corrected these errors. In the penumbra, errors can reach up to 400 ms^{-1} for both flat-fielding approaches, which would invalidate any measurements, as penumbral flows are of similar magnitudes.

Furthermore, we found that the shape of the transmission profile affects the measurements, with velocities derived using an asymmetrical transmission profile differing from those obtained with a symmetrical one. This, combined with the distinct behavior of the umbra compared to the rest of the field of view (FoV), implies that these measurements are sensitive to both the shape of the transmission profile and the spectral profile of the object.

Similar effects are observed in the magnetic field calculations, where the error structures correspond to those of the cavity map. However, in the case of magnetic fields, this effect is less significant, as the error values are small relative to the observed values (3 G of error in a 100 G field).

These results motivate the pursuit of methods to account for and correct these effects. One such method is the algorithm we developed to disentangle the FPI-based spurious effects on the flat-fields from those of other origins. The method, presented in Section 4.3, is capable of extracting the cavity map from the observations by employing the analytical modeling of the FPIs in different configurations.

We evaluated the performance of the algorithm through a series of simulated tests designed to assess various key aspects of its applicability to real data. In the first place, the performance of the algorithm was tested as a function of the S/N of the observations, the spectral sampling, and the configuration of the etalon, assuming the spectrum of the object is known. We also studied the possibility of deriving the observed object from the data through a deconvolution process. We tested the performance of this approach as a function of the spectral sampling. Finally, we simulated a scenario in which the etalon configuration was different from the real one. The aim of these tests was to assess the impact of neglecting

some properties of the etalon, such as the asymmetries in telecentric imperfect scenarios.

The algorithm has shown a high sensitivity to the noise level of the data. Nonetheless, the worst-case scenario considered in this work corresponds to much worse performances than current instruments can achieve in terms of S/N, and the errors are still within the necessary limits. In addition, the performance of the algorithm is the same for regions of the FoV where the etalon properties are far from their nominal values and regions where there is no deviation of the etalon properties.

Results for the three different optical configurations of the etalon are very similar. The telecentric configuration showed a worse performance than the collimated one in terms of cavity map retrieval. The loss of precision is nonetheless very small, below $\sim 5 \text{ ms}^{-1}$ in most cases.

The deconvolution of the object during the calculations was shown to yield a performance close to the ideal case in terms of cavity map determination. The performance when computing the gain showed a different behavior, although the average error never exceeds 0.4%. Concerning the dependence on the spectral sampling, the results show that with a scanning of at least six points, the additional error of deconvolving the object will not be larger than 25 ms^{-1} . Overall, we estimated a total error smaller than 100 ms^{-1} for the worst-case scenario, where only five wavelengths are used to scan the line using the deconvolution approach for an S/N of 200.

In real instruments, light travels through different optical elements before reaching the etalon. Along this light path, the observed object might be modified in such a way that it no longer resembles a fixed reference. It is in fact very difficult to assess the error associated with this assumption (i.e., assuming that the object reaching the etalon is that of the FTS atlas profile) since these deviations cannot be measured. By deriving the object from the data, we expect additional errors due to the deconvolution process of around 10 ms^{-1} , which we expect to be smaller than the ones produced by selecting a fixed object that differs considerably from the real one. Additionally, the deconvolution of the object allowed us to apply the algorithm to data where a solar structure is partly present.

Knowledge of the exact shape of the transmission profile has proven to be relevant to ensuring the accuracy of the algorithm. The results obtained for the crossover scenario have demonstrated that approximating the FPI's transmission profile by another can serve as a first-order approximation. However, the determination of the gain has proven to be much more sensitive to changes in the transmission profile. The errors observed in the crossover scenario are higher overall than those of the ideal case. In addition, the unaccounted asymmetries of the imperfect configuration paradoxically increase the error when improving the spectral sampling.

The results obtained when applying the algorithm to real observations taken with SO/PHI reinforce the validity of the algorithm. Indeed, we have been able to extract the contribution of the cavity map from the flat-field observations. Comparison of the derived cavity map with lab measurements suggests that the algorithm can successfully extract the cavity map from the flat-field observations. In future works, we aim to allow the algorithm to modify the angle of incidence across the FoV and validate the results by implementing them in the SO/PHI pipeline.

CHAPTER 5

SCIENTIFIC EXPLOITATION.

5.1 Persistent Homology in Solar Magnetograms.

5.1.1 Persistent Homology

Persistent homology stands out as a prominent technique within the topological data analysis toolkit, primarily for its capacity to capture the shape and distribution information of a dataset. The algorithm is rooted in the mathematical framework of homology groups. In topology, these groups measure the number of n -dimensional holes in a data set, or in other words, the number of connected components for a 0^{th} dimensional analysis, holes or rings for a 1^{st} dimensional analysis, spherical voids for the 2^{nd} dimensional analysis, and so on.

The primary objective of persistent homology is not only to compute the homology groups of a given dataset but also to study how they vary at different scales. To achieve this, the input data undergoes a process of division into a series of sequential subspaces, with each subspace encompassing the previous one. This sequential process, known as filtration, begins with a starting subspace comprising a single point from the original dataset. Subsequent subspaces are then constructed by incrementally adding points to the previous subspace until the final subspace includes all points of the original dataset.

After the filtration process is performed, persistent homology algorithms shift their focus to analyzing the evolution of topological features across the different subspaces. Specifically, they record the filtration value at which a new feature appears, meaning that it is absent in the previous subspace, and when it disappears, meaning that it is no longer present in the following subspaces. These two events are known as the birth and death of a topological feature, respectively.

In a nutshell, the n -dimensional persistent homology of a dataset with a given filtration can be described as the aggregation of all n -dimensional features (homology groups) that were created (birth) and subsequently eliminated (death) during the filtration process (?).

When applying persistent homology on a greyscale image, our focus lies in filtering the data according to the pixel values. Multiple filtering approaches exist, with the most extended ones being sublevel and superlevel filtrations, both based on the concept of thresholding. In these filtrations, the image is cropped to a specific value, forming a subspace that includes all pixels with values higher than this value in a superlevel filtration, or lower in

a sublevel filtration. This cropping value (i.e. the filtration value) is systematically varied from the lowest to the highest values of the image, or vice versa, thus generating a different subspace for each value. As a result, the persistence homology analysis captures and examines the evolution of topological features across different thresholds, enabling insights into the image's structural properties at various scales (?).

A more formal way of defining these filtrations can be done by considering an image as a discrete representation of a function f , defined over a two-dimensional space \mathbb{X} , such that:

$$f : \mathbb{X} \longrightarrow \mathbb{R} . \quad (5.1)$$

Let \mathbb{S}_ϕ be the subspace of \mathbb{X} for a filtration value of ϕ . In such a case, a filtration can be expressed as:

$$\mathbb{X} : \mathbb{S}_{\phi_0} \subset \mathbb{S}_{\phi_1} \subset \mathbb{S}_{\phi_2} \subset \dots \subset \mathbb{X} . \quad (5.2)$$

With this formulation, a topological feature with birth-death coordinates:

$$(B, D) = (\phi_I, \phi_{II}) , \quad (5.3)$$

corresponds to a feature that appears for the first time during the filtration process at the subspace \mathbb{S}_{ϕ_I} , and *persists* until the subspace $\mathbb{S}_{\phi_{II}}$, where it ceases to exist.

In a sublevel filtration, each subspace can be expressed as:

$$\mathbb{S}_\phi = f^{-1} ((-\infty, \phi]) , \quad (5.4)$$

where ϕ_0 is selected as the lowest value for any given pixel and its value is increased until the subspace includes all pixels. On the contrary, in a superlevel filtration, the subspaces can be expressed as:

$$\mathbb{S}_\phi = f^{-1} ([\phi, \infty)) , \quad (5.5)$$

where ϕ_0 is selected as the highest value for any given pixel and its value is decreased along the filtration.

Various methods exist for representing the information derived from a persistent homology analysis, including Betti numbers, persistence bars, and persistent diagrams (PDs) (?), among many others. For this study, we will utilize the PDs as our chosen approach due to their straightforward interpretation and extended use. A n -dimensional PD is a multiset of Birth-Death pairs, (B_i, D_j) , with multiplicity k , where each pair measures the number (k) of n -dimensional components that have been born at the filtration subspace \mathbb{X}_i and died in \mathbb{X}_j , that is usually represented in a 2D scatter plot.

The process of generating a PD of a greyscale image is as follows. We start by selecting the filtration direction (sublevel or superlevel) and the dimension of the analysis (either 0 or 1). We initialize a threshold as the highest or lowest value from the image, depending on the choice of filtration. We then perform the filtration by systematically adjusting the threshold and creating a binary image for each threshold. This process divides the image into two sets: pixels with values above the threshold and pixels with values below it. The choice of filtering determines which of the two sets makes up the subspace. We then look for the existing topological features within each of these subdivisions. The specific process

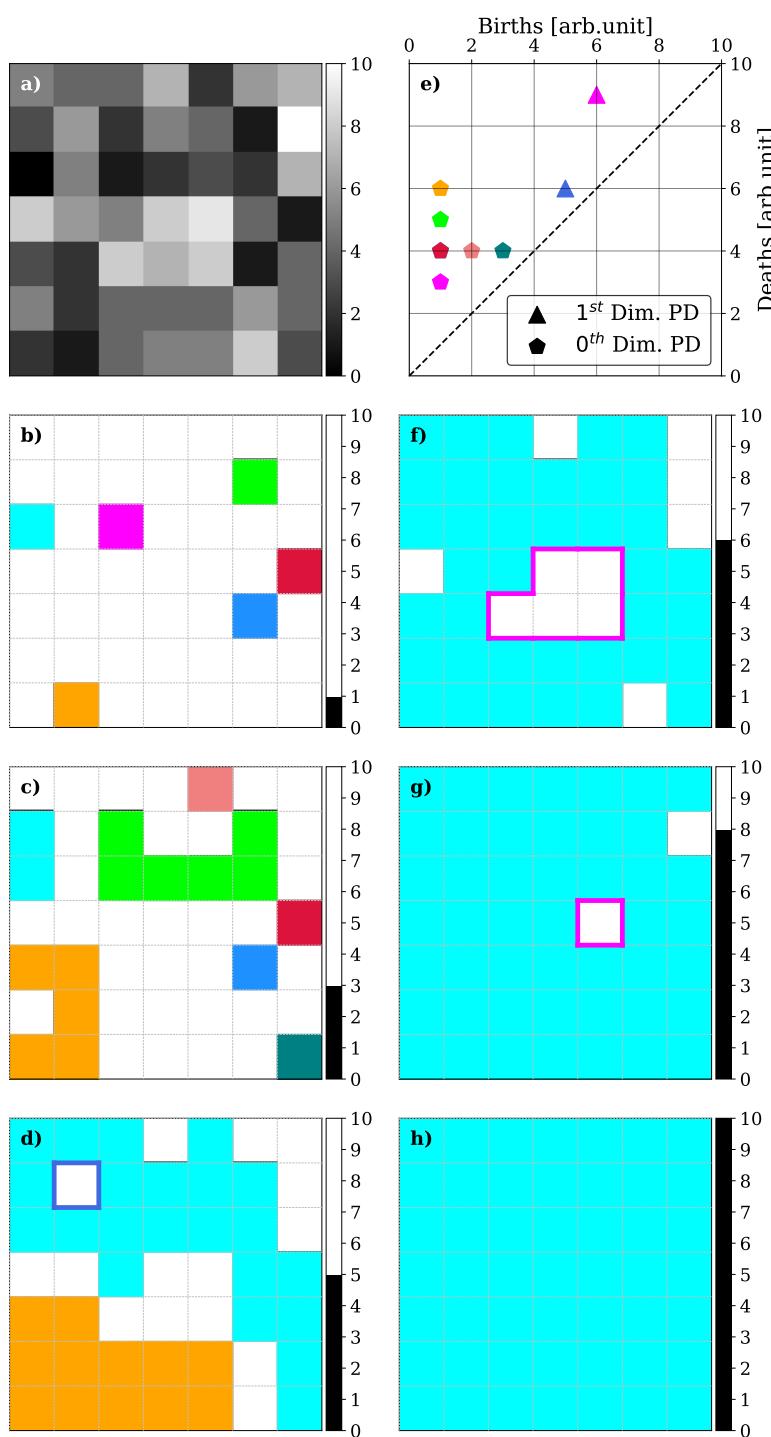


Figure 5.1 Sublevel filtration of a greyscale image and PDs of the 0^{th} and 1^{st} dimensions. Panel a) shows the input data. In panels b), c), d), f), g), and h) different snapshots of the filtration process are shown. The value of the filtration parameter, ϕ is shown in the color bar at the right of each image. Only pixels with a value lower than the filtration value (colored pixels) belong to the subspace shown in each snapshot. Different homology groups are represented with different colors at each snapshot. For connected components (0^{th} dimensional homology groups) the whole pixel is shown with the corresponding color. For rings, (1^{st} dimensional homology groups), only the border of each hole is colored. In panel e) the PDs of both dimensions are shown. The color of each point in the diagram is the same as the one used to plot the corresponding topological feature in other panels. An animation depicting the whole filtration process is also provided as part of the article.

by which these features are identified is detailed in the next paragraph, where the structures corresponding to both dimensions are illustrated using the example shown in Fig. 5.1. We repeat this process until the threshold reaches the opposite limit to that from which it started. Along this process, we follow the appearance, merging, and disappearance of connected components. When two components merge, the longer-lived one (*i.e.* the first to appear along the filtration process) absorbs the younger one, thus resulting in the death of the second (?). We determine the birth and death for each component based on the thresholds at which these events occur. Finally, we construct a scatter plot where the horizontal axis represents the birth values and the vertical axis represents the death values. Each point on this plot corresponds to a persistence point, whose coordinates reveal the scales at which the corresponding topological feature is present.

An example of this process with a sublevel filtration is shown in Fig. 5.1. Panel a) displays the input data, panels b), c), d), f), g), and h) show some of the key steps of the filtration process, and, lastly, panel e) displays the PD for a 0th and 1st dimensional analyses. These plots illustrate how connected components and rings are born and then die as we increase the filtration level. As these components (shown in different colors) increase in size and come into contact with other components, one absorbs the other, thus resulting in the death of the second. This phenomenon is shown in panels b) to d), where we observe the progression of the components until only the blue and orange connected components remain. Additionally, the diagrams also reveal the appearance of two rings in the data (panels f) and g)). These rings are found when pixels that do not belong to the subspace are surrounded by a connected component, and die when those pixels are included in the component as the threshold increases (blue ring in panel f)). Finally, the PD (panel e)) displays the birth and death values (*i.e.* the filtration value) of all the features, of dimensions 0 and 1, that have been identified (birth) and subsequently eliminated (death) throughout the filtration process.

5.1.2 Persistent Images

The PD displayed in Fig. 5.1 contains only a limited number of points due to the simplicity of the input image. However, when analyzing real data, these diagrams can consist of hundreds or even thousands of birth-death pairs with high multiplicities, simply due to the size of the images. Additionally, features not only representing the genuine behavior of the data but also reflecting the distribution of noise appear on the diagrams. To address this complexity, several strategies have been developed to simplify the information from PDs, such as persistence curves (?), persistence landscapes (?), or persistence images (PI) (?). In this study, we will focus on the latter, due to its noise filtering capabilities and because the representation of the results remains in a Birth-Death diagram, allowing for easy interpretation of the results, similar to a persistence diagram.

PIs are a condensed form of a persistence diagram, offering a concise and easy-to-understand representation of its topological features. They capture the spatial distribution and persistence information of these features, allowing for the enhancement of the most relevant ones and filtering of the others. A PI is constructed using the concept of persis-

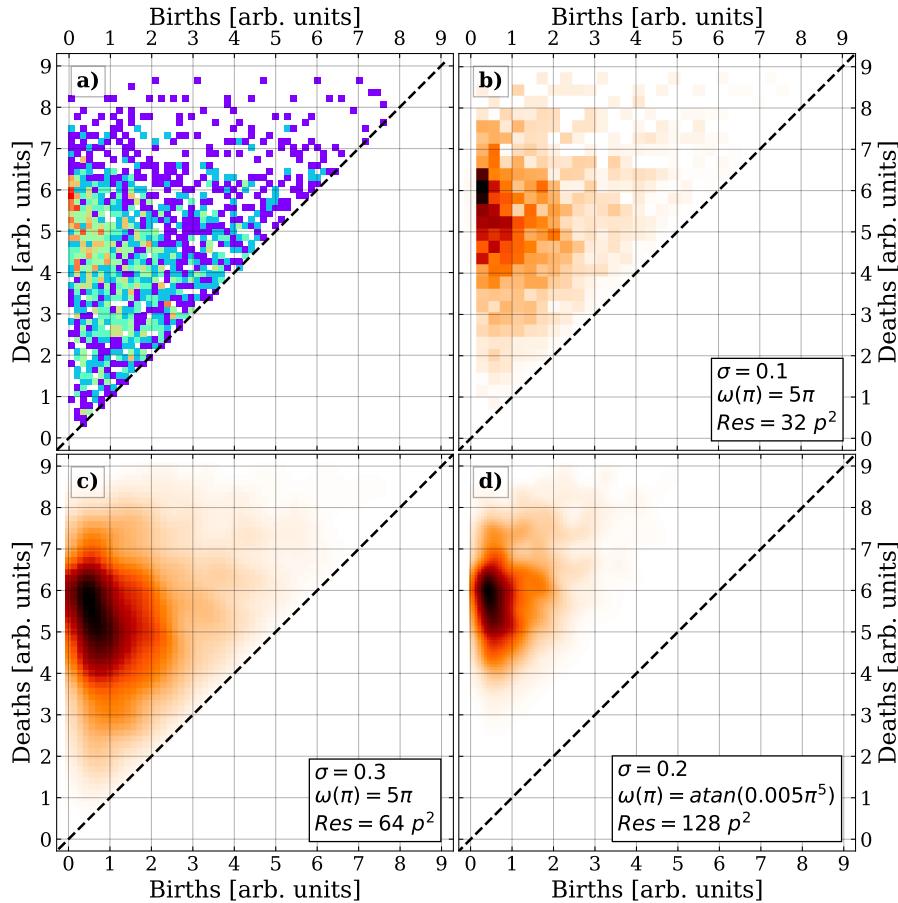


Figure 5.2 Panel a) shows an example of a persistence diagram as a 2D histogram. The color of each bin represents its multiplicity, with the red spots corresponding to higher values. Panels b), c), and d) show three different examples of PIs. The three parameters given in the legends of the PIs are: the standard deviation (σ) of the Gaussian kernel ($K(z)$), the weighting function, and the resolution of the image PI in pixels (p).

tence. Each topological feature, represented by a point in a PD, has a persistence, π , of:

$$\pi = D - B \quad (5.6)$$

where (B, D) are the corresponding birth-death coordinates in the diagram. A feature with a large persistence is present at different scales in the data and therefore is more likely to represent the real behavior of the data. On the contrary, short-lived features are typically associated with the noise distribution and usually do not provide much information about the data.

When constructing a PI, a weighting function, $\omega(\pi)$, is employed to assign weights to each point in the diagram, ensuring that longer-lived features have greater weights than shorter-lived ones. There are multiple choices for the shape of the weighting function, which are entirely dependent on the aims of the study and data type. The simplest example

is often a linear or power-law relation ($\omega(\pi) = a\pi^b$), where a and b can be tuned to assign progressively higher weights to higher persistencies, thus focusing the study on the longer-lived components. On the other hand, if the objective is to filter out noise while assigning similar weights to all non-noise points so that all points have a similar relevance in the analysis, the chosen function is usually an arc-tangent.

The PI is then generated by dividing the persistence diagram plane into a grid with a desired resolution. Within each grid region (or pixel), the weighted features of the diagram within the region are added up using a kernel density estimation. The kernel function, $K(z)$, can be tuned to suit the nature and objectives of the analysis, with Gaussian functions being the most common approach.

The resultant PI is a 2D matrix, wherein each pixel corresponds to a specific area in the persistence diagram, and its value represents the cumulative weight of the topological features found within that area. In Figure 5.2, three examples of PI (panels b), c), and d)) are presented for the same persistence diagram (panel a)), where distinct choices of resolution, kernel function, and weighting function have been applied to each image.

All the PDs, PIs, and the rest of the analysis tools presented in this work, have been computed using the Homcloud python package (?).

5.1.3 Data

In this work, we study the results of applying persistent homology to different regimes of solar activity by applying the analysis to both quiet Sun and active region magnetograms.

5.1.3.1 Quiet Sun observations

The study of quiet Sun regions requires high magnetic spatial and temporal resolutions and sensitivities to be able to capture the small-scale evolution of the magnetic structures due to their weak signals and short time scales (Bellot Rubio & Orozco Suárez, 2019). For this reason, we employ observations taken by the Solar Optical Telescope (SOT; Tsuneta et al. 2008) aboard the *Hinode* satellite (Kosugi et al., 2007), a space-borne solar observatory. In particular, we employ observations from Hinode's Operation Plan (HOP) 151. These observations consist of long (≥ 20 h) and mostly uninterrupted sequences of measurements of the Narrowband Filter Imager of the Na I D1 line at 5896 Å taken with a cadence of 50–70 s. The data correction of the selected observation sets has been carried out in (?).

5.1.3.2 Active regions observations

We employ observations of active regions (ARs) taken by the Helioseismic and Magnetic Imager (HMI; ?, ?) on board the Solar Dynamics Observatory (?). HMI provides a continuous observation of the Sun where a full-disk magnetogram, as well as Dopplergrams, are provided at all times. The full-disk, uninterrupted observations of HMI make it a very suitable instrument to study the evolution of active regions as the formation and development of active regions can be fully captured.

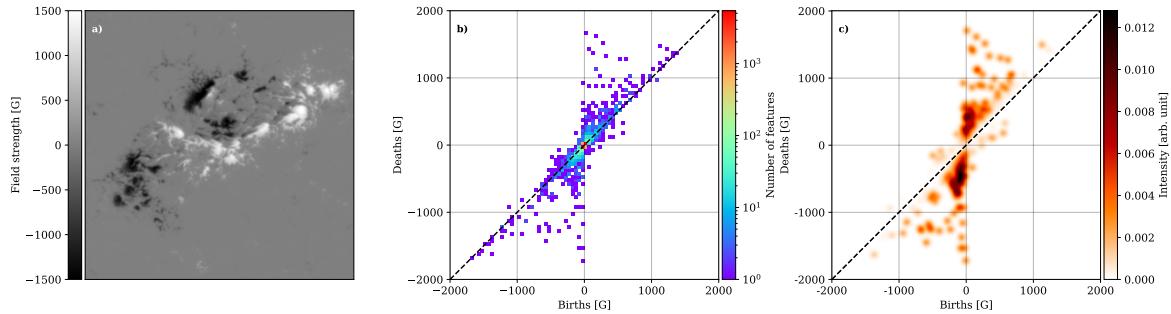


Figure 5.3 (a) SDO/HMI magnetogram taken on 2011-02-13 depicting an active region (NOAA AR 11158). (b) The corresponding PD combining superlevel and sublevel filtrations. (c) PI generated from the PD in panel b) with the following configuration: Resolution = 1000 pixels² (4 G per pixel), weighting function: $\omega(\pi) = \arctan(5 \times 10^{-8} \pi^3)$ and a gaussian kernel with $\sigma = 40$ G.

We focus the analysis on a series of newly-emerging ARs identified in (?). In particular, we employed the 12-minute cadence observations taken during the period from May 2010 to June 2011, which corresponded to a period of low solar activity.

5.1.4 Analysis and results

The application of persistent homology to a specific dataset can vary depending on the aims of the study. Different dimensions of the analysis and various types of filtrations focus on distinct features within the data. It is crucial to have prior knowledge of the expected structures and relevant features to be captured in the analysis in order to determine the appropriate approach. In this section, we aim to outline the most appropriate approach for studying the particular case of solar magnetograms.

The solar magnetograms employed here represent the longitudinal component of the magnetic field on the photosphere and are typically presented as greyscale images, as shown in Figure 5.3, panel a). The polarity of the line-of-sight magnetic field is indicated by the sign of each pixel, where positive and negative signals correspond to field lines pointing towards and away from the observer. Applying a single filtration to a greyscale image only displays features corresponding to one polarity (positive or negative) in a PD. However, to conduct a comprehensive study of the magnetic field, both polarities are essential, thus necessitating the use of two separate filtrations with different filtration directions.

We determined that a combination of superlevel and sublevel filtrations with a 1st-dimensional persistent homology analysis was the most suitable approach for studying solar magnetograms. This choice is based on two main reasons. Firstly, the 1st dimensional analysis allows us to identify the most prominent features in a PD, which is not the case in the 0th dimensional analysis where the strongest feature does not appear in the diagram as it never dies (see Fig. 5.1). Secondly, by combining superlevel and sublevel filtrations, we can display the results of both filtrations in a single diagram. Features corresponding to dif-

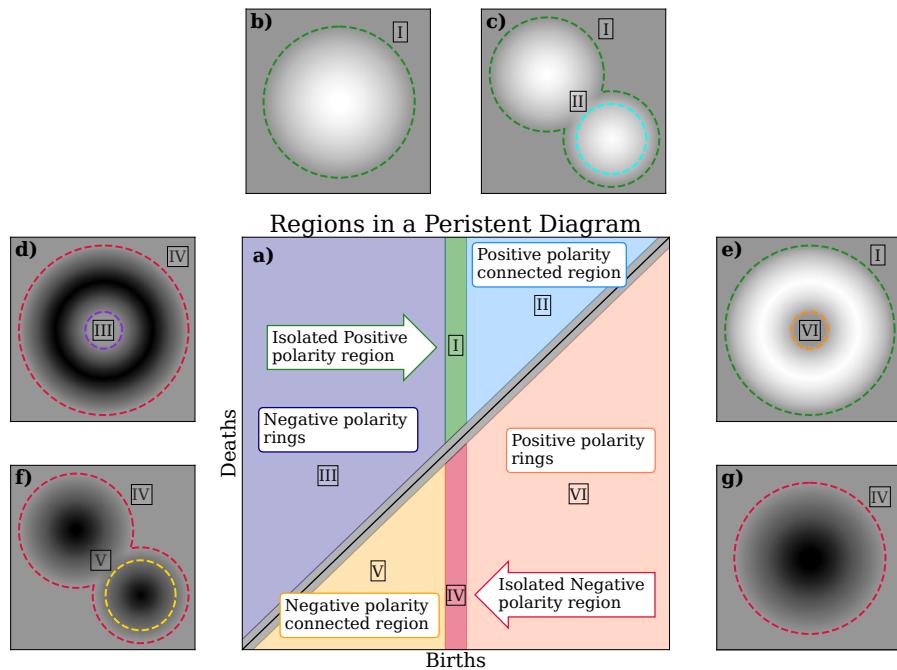


Figure 5.4 Schematic representation of the PD and the different regions (panel a)). The magnetic structures corresponding to the topological features found in the different regions are shown in panels b) to g), with each feature identified by a ring with the corresponding color.

ferent filtrations will have persistencies with opposite signs (features found in a superlevel filtration will be born at higher filtration values than their death, resulting in a negative lifespan). This enables us to construct a PD in which all features displayed above the identity line (with positive persistencies) correspond to the sublevel filtration, and those below the line correspond to the superlevel filtration (see panel b) in Fig. 5.3).

The PDs, and consequently the PIs, offer valuable insights into the magnetic structures present in the magnetograms. The location of a topological feature on the diagram is completely determined by the properties of the corresponding magnetic structure. Specifically, this position is influenced by factors such as the maximum intensity of the magnetic field, its proximity to other magnetic structures, and its geometric shape, including the presence of holes or pores within the structure. These characteristics allow us to partition the diagram into distinct regions, where topological features within each region correspond to different types of magnetic structures.

We distinguished between six distinct regions in the diagram. Figure 5.4 illustrates these regions in panel a) and provides schematic representations of the corresponding magnetic structures in panels b) to g). The regions are as follows: first, topological features located above the identity line (positive persistencies) with birth values close to 0 (region I in the diagram). Features within this region represent isolated magnetic structures of positive polarity, that is, patches of positive magnetic flux fully enclosed by an absence of any magnetic field. The threshold defining what is considered “close to 0” is determined by the data’s properties. To identify isolated structures, we set the threshold as a function of the statistical properties of the background signal (i.e. areas of the magnetogram with little magnetic flux). Specifically, the limits for this region are set as $(-5\sigma_{bg}, 5\sigma_{bg})$, where σ_{bg} denotes the standard deviation of the background signal found in a 15×15 pixels box devoid of strong magnetic structures.

The second region (II in the diagram) comprises features above the identity line with positive birth values, representing connected structures with positive polarities, that is, positive magnetic field structures in contact with another positive structure but not fully merged. The third region (region III) contains topological features above the identity line with negative birth values, which corresponds to magnetic structures of negative polarity exhibiting a ring-like attribute, namely, structures with pores or holes. These three regions of the diagram have counterparts with negative persistencies. Thus, features associated with isolated structures of negative polarities are found in the region with a birth value close to 0 but below the identity line (region IV), features for connected negative structures are also located below the line but with negative birth values (region V). Lastly, features arising from positive magnetic structures with ring-like attributes are found below the identity line but with positive birth values (region VI).

5.1.4.1 Quiet Sun results

Quiet Sun regions are characterized by the presence of weak, small-scale magnetic field signals that exhibit rapid evolution. This rapid evolution leads to a multitude of signals in a single snapshot that evolve quickly from one frame to another. The small scale and rapid changes of the processes of the quiet Sun make data analysis techniques desirable for their

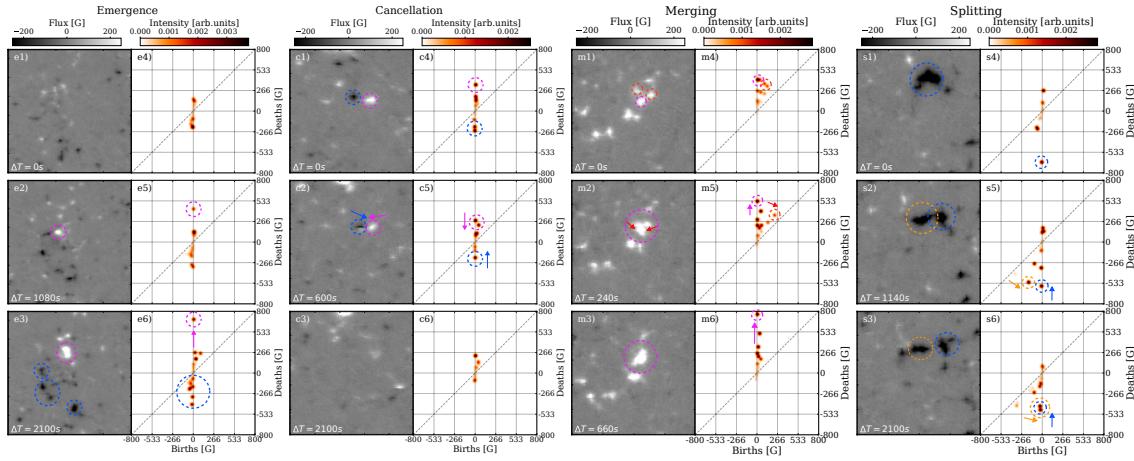


Figure 5.5 Examples of flux emergence (left column), flux cancellation (central left column), merging (central right column), and splitting (right column) events in the quiet sun. The time interval shown in the magnetograms is always with respect to the first frame (panel a) for emergence and g) for cancellation). The PIs correspond to the magnetogram to their left. The parameters for their computation are: resolution = 1000 pixels² (1.6 G per pixel), weighting function: $\omega(\pi) = \arctan(5 \times 10^{-7} \pi^3)$ and a Gaussian kernel with $\sigma = 16$ G.

study due to the complexity of such endeavors.

When examining the evolution of signals across the entire field of view, we observe a dynamic process characterized by numerous regions interacting destructively while new signals emerge throughout the whole region. Despite these continuous changes, the overall structure of the magnetogram appears stable, with consecutive snapshots exhibiting strikingly similar properties. This apparent equilibrium state is also evident when studying the PIs, as consecutive frames show minimal differences in their representations.

Due to the apparent equilibrium state of the overall structure of the magnetograms, it is necessary to narrow down the field to which we apply the analysis. We found that when the number of signals in the studied region is lower, we can observe flux cancellation and emergence events, as well as merging and splitting events, through the changes induced in the persistence diagram. In cancellation events, two regions of magnetic flux with opposite polarities interact destructively, nullifying each other. On the contrary, in flux emergence events, we observe signals of both polarities suddenly emerge from a region with little magnetic flux. In splitting and merging events, only one polarity is involved. Two or more distinct structures of equal polarity merge together in merging events, and a single structure is divided into two or more for splitting events.

In Figure 5.5, an example of an emergence event is shown in three snapshots through the magnetograms (panels e1 to e3) and their corresponding PIs (panels e4 to e6). When we focus on the second snapshot (panel e2 and e5), we see that a new positive and isolated feature (birth ~ 0) that was not present in the previous snapshot has appeared in the PI and stands out from the rest of the signals (highlighted with a pink circle in both magnetogram and PI), while simultaneously, the density of negative polarity features also begins

to increase. In the last frame, we see that in the case of positive polarity, the majority of the signal has concentrated in a single structure, as shown by the increase of the death value of the corresponding feature in the PI. A few connected features are also seen, but these are less significant. Meanwhile, the negative polarity signal has been distributed into multiple structures instead of concentrating in a single one, as evidenced by the absence of a prominent feature in the PI and the increased density of connected features (highlighted with blue circles both in the PI and the magnetogram).

A very similar analysis can be carried out to analyze cancellation events. The evolution of the persistence image is very similar to that of the emergence events but in the opposite direction. Figure 5.5 also shows an example of a flux cancellation event through magnetograms (panels c1 to c3) and PIs (panels c4 to c6). In the beginning, the PI shows the presence of features of opposite polarities. When the corresponding structures approach each other and begin to interact, the magnetic signal starts to decrease, which is observed in the PI as a simultaneous movement of the features towards the center of the diagram. This reduction in signal continues until both features reach the center of the diagram, which corresponds to the moment when they will have completely canceled each other out (panels c3 and c6).

For the events that only involve one polarity, namely merging and splitting events, the same behavior is seen in the PI for positive and negative features, but on their respective sides of the PI.

An example of a merging event of positive polarity structures is shown in Fig. 5.5, in panels m1) to m6). These events start with multiple isolated, or interacting structures (as shown in panels m1 and m4), that are moving towards each other. As the structures cluster, two movements are seen in the PI: firstly, the features corresponding to the structures with the weakest field (red features in panels m2 and m5) move towards the identity line; and secondly, the feature corresponding to the main structure (pink feature in panels m2 and m5) experiences an increase in its absolute death value due to the increase in magnetic flux coming from the rest of the structures. When the structures are fully merged, only a single feature appears in both the magnetogram and PI (pink feature in panels m3 and m6), in the isolated region (region I or IV, depending on the polarity).

This process is reversed for splitting events, as shown in panels s1) to s6) of Fig. 5.5. We see how an initially isolated feature in the PI (blue feature in panels s1 and s4) evolves into two (or multiple) features. When the process has started, but the two parts have not yet completely separated, a second feature appears in the diagram in the region corresponding to the connected structures (regions III or V in the diagram), as shown in panels s2) and s5). As the two structures continue to separate, this second feature gradually approaches the region for isolated structures (regions I and IV) as the magnetic field surrounding it in the magnetogram diminishes (panel s5). Eventually, when both structures are completely separated (i.e. with no signal around them), they will both appear on the diagram as two isolated features with a lower death value compared to the initial one (panels s3 and s5). This reduction occurs because the magnetic flux is distributed between the two structures.

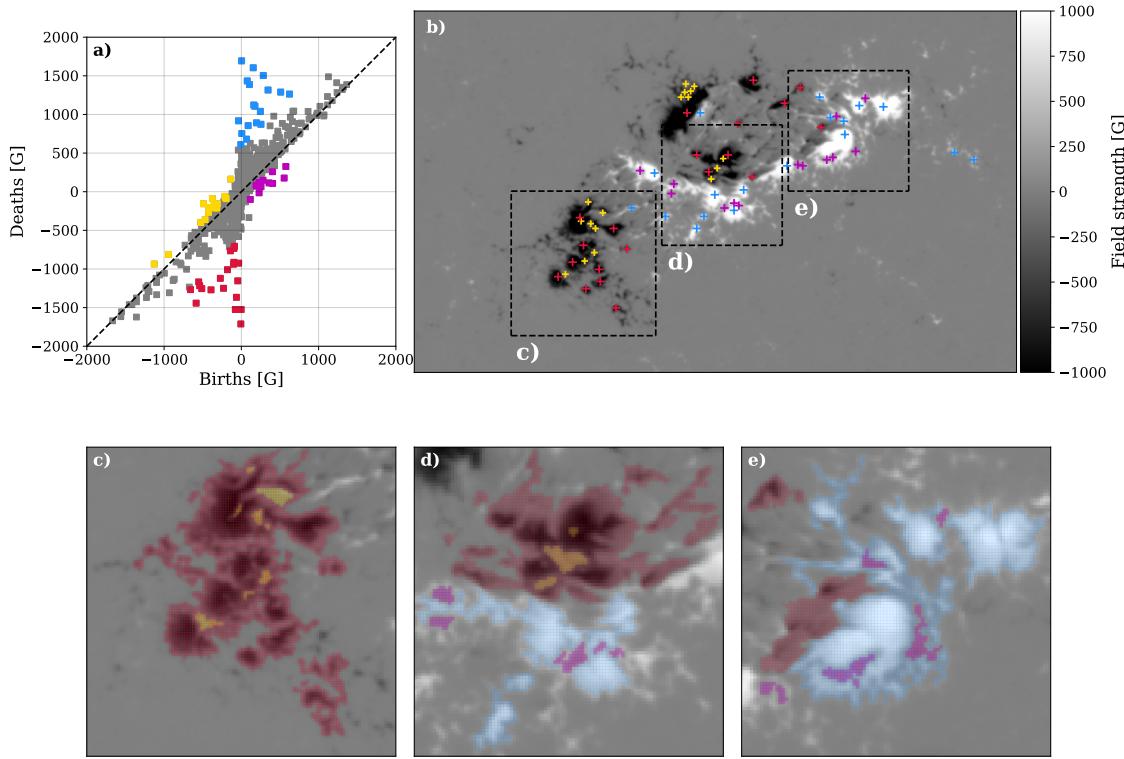


Figure 5.6 SDO/HMI magnetogram taken on 2011-02-13 at 06:34 UT depicting an active region (NOAA AR 11158) and the corresponding persistent diagram (panel a)). Panels c) to e) display a zoomed region of the active region, corresponding to the labeled region with the same letter in panel b). Features of different types are shown in different colors in the diagram. The corresponding structures are shown in the same colors in the magnetograms; as colored crosses in panel b) and as a colored transparent overlay displayed over the whole pixel in panels c) to e) to show the extent of the structure.

5.1.4.2 Active region results

It is important to understand which types of magnetic structures can be identified through persistent homology and establish the correspondence between these structures and the position of the corresponding topological feature in a persistence diagram. To achieve this understanding, Fig. 5.6 displays both a magnetogram with complex morphology (panel b) and its corresponding persistence diagram (panel a), along with three zoomed-in regions of the magnetogram (panels c to e). In all panels, some topological features or their corresponding magnetic structures have been color-coded based on their types, or equivalently, based on their positions in the persistence diagram. In the complete magnetogram (panel b), structures have been marked with a cross, indicating the pixel where the structure died during the filtration process. Meanwhile, in panels c to e, all pixels composing each structure have been colored. It is noteworthy that nearly all pixels appear colored because we have selected the most significant structures—those with longer lifetimes (Eqn. (5.6)). Consequently, these structures encompass all less significant structures that are absorbed and

incorporated into the former during the filtration process.

The analysis of the persistence diagram allows us to deduce several properties of the magnetogram. Firstly, the persistence diagram provides a rapid assessment of the intensity of the magnetic flux since the death value of the topological feature coincides with the maximum flux (in absolute value) within the corresponding structure. In the case of Figure 6, we observe that several structures exhibit maximum (absolute) values surpassing 1500 G, with multiple structures falling within the range of 1000 G to 1500 G.

Secondly, we can infer how the magnetic signals are distributed by examining the number of isolated and connected structures in the diagram (structures highlighted in blue and red depending on their polarity in Fig. 5.6). The complex morphology of the structure displayed in the magnetogram is evident in the high number of connected structures (regions II and V in the diagram) and the absence of prominent isolated structures (regions I and IV).

Lastly, the presence or absence of ring-like structures provides insights into how the magnetic structures are connected. These features can only be found in regions where connected structures create highly complex morphologies with gaps between them, as illustrated in the magnified regions of the magnetogram in Figure 6 (panels c to e).

An example of how the three features allow us to classify ARs depending on their morphologies is shown in Figure 5.7, where three different ARs and their corresponding PIs are displayed. Although at first sight, the PIs appear to be very similar, especially the ones shown in panels d) and f), upon closer inspection, it is possible to find the differences when focusing on the three features mentioned previously. The first AR (panel a)), also shown in Fig. 5.6, shows a very complex morphology, where the magnetic field of both positive and negative polarities is distributed in multiple connected structures. This behavior is displayed in the PI through the high density in the isolated and connected features in equal proportions (i.e. with no prominent features) and with the presence of ring-like features in both polarities. In contrast, the second AR (panel b) shows a simpler magnetic structure with weaker signals. The PI for this case shows an absence of ring-like features in both polarities and a very low density in the regions for connected and isolated features. Lastly, panel c) shows an AR where the positive magnetic field is concentrated in one main area whereas the negative polarity magnetic field shows a fragmented structure. Comparing the corresponding PI with that of the initial case (panels f and d, respectively), a similarity is evident in the region corresponding to negative polarities, observed in both the density of connected components and the presence of ring-type structures. Nevertheless, when examining the positive polarity, it becomes evident that, unlike its negative counterpart, there are only a few prominent isolated structures and a notably low density of connected structures. Furthermore, this asymmetry between the positive and negative distributions is underscored by the absence of ring-like structures in the former.

5.1.4.3 ‘Interacting’ Diagram

So far, we have shown how persistent homology is capable of identifying the various morphologies of active regions and the types of structures that can be identified through persistent images or diagrams. Nevertheless, these structures are either isolated or regions of

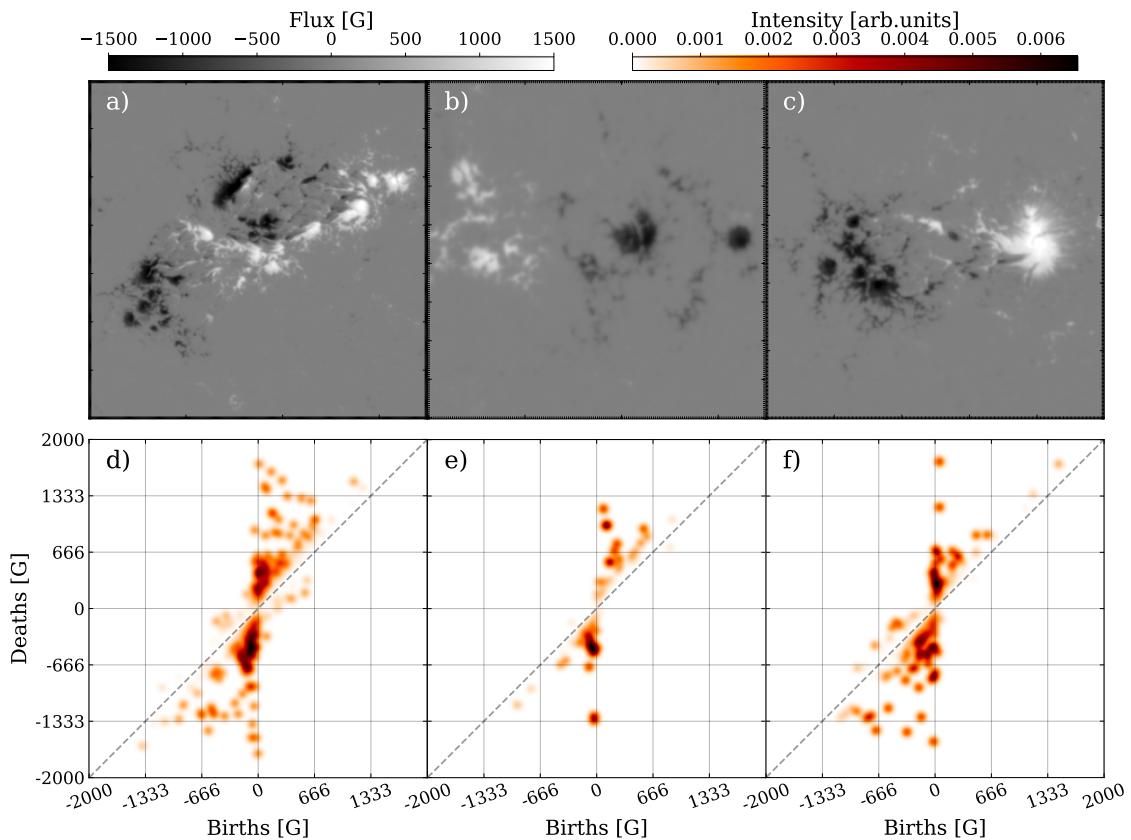


Figure 5.7 SDO/HMI magnetograms of three different active regions and their corresponding persistence images. a) NOAA AR 11158, date: 2011-02-13 at 06:34. b) NOAA AR 11098, date: 2010-08-12 at 20:58. c) NOAA AR 11072, date: 2010-05-22 at 20:58. All persistence images have been generated with the following parameters: Resolution = 1000 pixels² (4 G), weighting function: $\omega(\pi) = \arctan(5 \times 10^{-8} \pi^3)$ and a gaussian kernel with $\sigma = 40$ G.

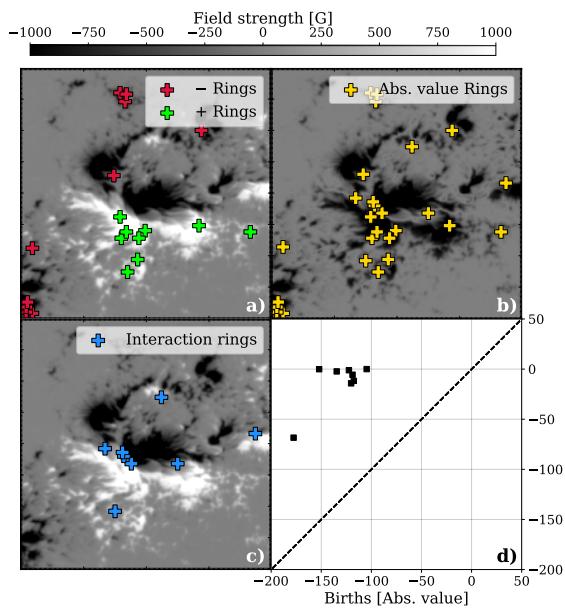


Figure 5.8 Depiction of the different steps carried out for the generation of the interaction diagram (panel d)) for a zoomed-in region of an SDO/HMI magnetogram of NOAA AR 11158 observed on the 2011-02-13 at 18:34. The crosses point to the position at which each of the ring-like structures die, which is always located inside the perimeter of the hole. Panel a) shows the original magnetogram and the position where rings of positive and negative polarities are found. Panel b) shows the magnitude of the field strength in negative values along with the rings found in this representation. Panel c) shows the original magnetogram again, but now featuring the interaction rings, identified when comparing the position of the rings found in the previous steps. Lastly, panel d) shows the ‘interaction diagram’ generated using only the interacting ring structures.

the same polarity that interact with each other. Due to the nature of the filtration process, persistent homology is unable to detect structures where magnetic fields of opposite polarities form a joint structure. However, ARs where there is a significant interaction between magnetic fields of opposite polarities are of greater interest due to their association with flare production.

This issue is illustrated in the analysis carried out in the previous section of NOAA Active Region 11158 (see Fig. 5.6). Although we can capture the complexity of both positive and negative magnetic structures through the persistent diagram, the δ -spots present in the central region remain unnoticed. However, the large number of flare events associated with this region including an X2.2-class event are thought to be related to the abundance of these structures (e.g. ?, ?). In this section, we describe a way to efficiently detect and quantify these structures using persistence homology, with only a few additional steps in the analysis.

We start by tracking the position in the magnetogram where rings are detected. Following this, we modify the magnetogram by inverting the sign of one polarity, ensuring that all pixels are either negative or positive. This way, we construct a second ‘magnetogram’ in which we only have information about the intensity, in absolute value, of the magnetic field. This allows us to identify features formed by any combination of signals, whether they are of the same polarity or opposite. Using this new magnetogram, we repeat the analysis and record the positions in the magnetograms at which we find ring-like features. It is worth noting that in this analysis, all the rings identified in the initial step, which were formed

by structures of equal polarities, are still detected, although their position in the persistent diagram may have changed due to the change in polarity, hence the relevance of tracking the pixel in the magnetograms. However, only in this second analysis can we identify rings formed between opposite polarities. By selectively considering the ring-like features exclusively identified in this second analysis, we can effectively identify and characterize the structures that are formed by the interaction between different polarities.

Figure 5.8 depicts an example of the interaction analysis for NOAA AR 11158, which exhibits a strong interaction between opposite-polarity fields. Panel a) displays the magnetogram, indicating also the positions where ring-like features have been identified for both positive and negative polarities. Only features with absolute persistencies greater than $5\sigma_{bg}$ and whose birth occurs in the range: $(5\sigma_{bg}, \infty)$, for positive polarities, and: $(-\infty, -5\sigma_{bg})$ for negative polarities, have been recorded. We now repeat the same analysis but using only the magnitude of the signal. To do this, we invert the positive polarity and, once again, register the positions of the ring-like features in this new magnetogram, as shown in panel b). As observed in panel c), the majority of the interaction rings (*i.e.* those exclusively identified in the second analysis) are located in areas characterized by strong interaction, where the δ -spots were found. In these areas, both polarities interact, resulting in the formation of ring-like structures comprising positive and negative magnetic fields due to their close proximity. These areas, such as δ -spots, are of particular interest, as they typically harbor magnetic cancellation, reconnection, and flux emergence. However, some points appear to be situated in uni-polar fields. These points, despite what may appear at first sight, are found by this analysis due to a structure that requires the other polarity to close completely and thus form a ring. It is noteworthy that the occurrences of such cases are quite limited when compared to the rings observed in highly interacting zones. While their presence does not necessarily indicate intense interaction, it does imply a certain level of interaction between the two polarities. These features can be represented in a persistence diagram in an analogous way to the standard results of a persistent homology analysis. This is what we have referred to as ‘interacting diagram’ and it is worth noting that only the magnitude of the birth and death coordinates are relevant parameters since the sign will be the one matching the polarity selected at the second step. We have chosen to invert the positive polarity so that these features have positive persistencies, as it is more common in persistent homology studies. However, it is important to emphasize that this decision is completely arbitrary and has no impact on the results of the analysis.

It is useful to determine the information conveyed by the interaction diagram regarding the structures themselves. By taking into account the position and quantity of the interaction rings, along with the temporal evolution of the diagram, we can discern the moment and location where these highly interacting structures develop. Therefore, interaction diagrams could be a new tool to identify, through their topological properties, the strong-gradient polarity inversion lines that characterize δ -spots. To achieve this, it is necessary to incorporate into the analysis the temporal evolution of these structures and study the properties that can be extracted from the δ -spots through these diagrams, which goes beyond the scope of this work but represents the next (necessary) step to assess the predicting capabilities of persistent homology in the field of solar physics.

5.1.5 Conclusions

In this study, we investigate the most adequate approach for the application of persistent homology algorithms to the analysis of solar magnetograms. By combining different filtrations in a single one-dimensional persistent homology analysis, we can effectively capture structures corresponding to both polarities of the magnetic field. We have applied this analysis to observations of the quiet Sun and active regions, taken with both Hinode/SOT and SDO/HMI, respectively. Lastly, we have analyzed the results and identified the features of the data that can be found through persistent diagrams and images, and also show some examples of applications of the algorithms.

Our proposed approach to persistent homology algorithms involves the integration of sublevel and superlevel filtrations within a single analysis, enabling the creation of a comprehensive persistence diagram that encompasses features from both positive and negative magnetic structures. Through the examination of the positions of these identified features within the resulting persistence diagram, we can discern the diverse magnetic features present in the magnetograms. This approach has demonstrated its efficacy in capturing the intricate complexity of magnetic structures, with a particular emphasis on active regions. Through this method, we have achieved successful differentiation between the various morphologies present in active regions by analyzing the presence or absence of specific features in the corresponding persistence images.

On the other hand, the persistent images obtained from quiet Sun observations exhibit significant similarity to each other. This indicates a lack of overall evolution in the magnetic structures within these regions. In quiet Sun areas, small regions of magnetic flux interact with each other in small-scale events, while the overall structure remains relatively static. These small-scale events become more apparent in persistent images when the field of view is reduced. These small-scale events, such as flux emergence or cancellation, can be observed through persistent images as a joint movement of negative and positive features. In cancellation events, the features move toward the center of the image, while in emergence events, they move away from the center.

Additionally, we have successfully identified interactions between opposite-polarity magnetic fields by detecting ring-like features formed by these two polarities. To achieve this, we introduced a method for calculating an ‘interaction diagram’ that selectively displays features resulting from the interaction between polarities. This interaction diagram is generated by comparing the ring-like features identified in an analysis using only the absolute value of the signal with those found in the standard analysis. This approach enables us to detect the presence of δ -spots and quantify the level of interaction between polarities, which is one of the critical factors for the understanding and prediction of flare eruptions.

In conclusion, our application of persistent homology to solar magnetograms has provided a comprehensive and insightful framework for studying magnetic structures on the solar surface. The topological features derived from magnetograms serve as a foundation for classifying active regions based on their morphology and level of interaction, as certain topological features may have inherent connections to solar atmospheric phenomena. For instance, the presence of interaction rings in active regions might be correlated with flare production, while the interaction of signals from opposite polarities observed in a per-

sistent diagram in the quiet Sun could be linked to small-scale reconnection events or the separation of signals associated with flux emergence. The exploration of these relationships and the assessment of the presented tools in achieving precise active region classification and their potential as predictive tools are topics of our upcoming research. Moreover, we have introduced new tools, such as the interaction diagram, which facilitates the detection and quantification of structures interacting with opposite polarities, like δ -spots, addressing a crucial aspect of flare prediction. The findings presented in this article lay a solid foundation for future studies, emphasizing the potential of persistence images as valuable inputs for machine learning algorithms and contributing to advancements in space weather forecasting.

Lastly, it is important to emphasize that in this study we have focused primarily on static images in order to provide a solid basis for future investigations. The next logical step in this study is to complete the analysis of active regions, which includes examining their temporal evolution. This approach allows for the simultaneous consideration of two key factors in understanding flare eruption processes: morphological complexity, whose analysis is intrinsic to persistent homology, and the study of their temporal evolution through the analysis of the evolution of persistence and interaction diagrams.

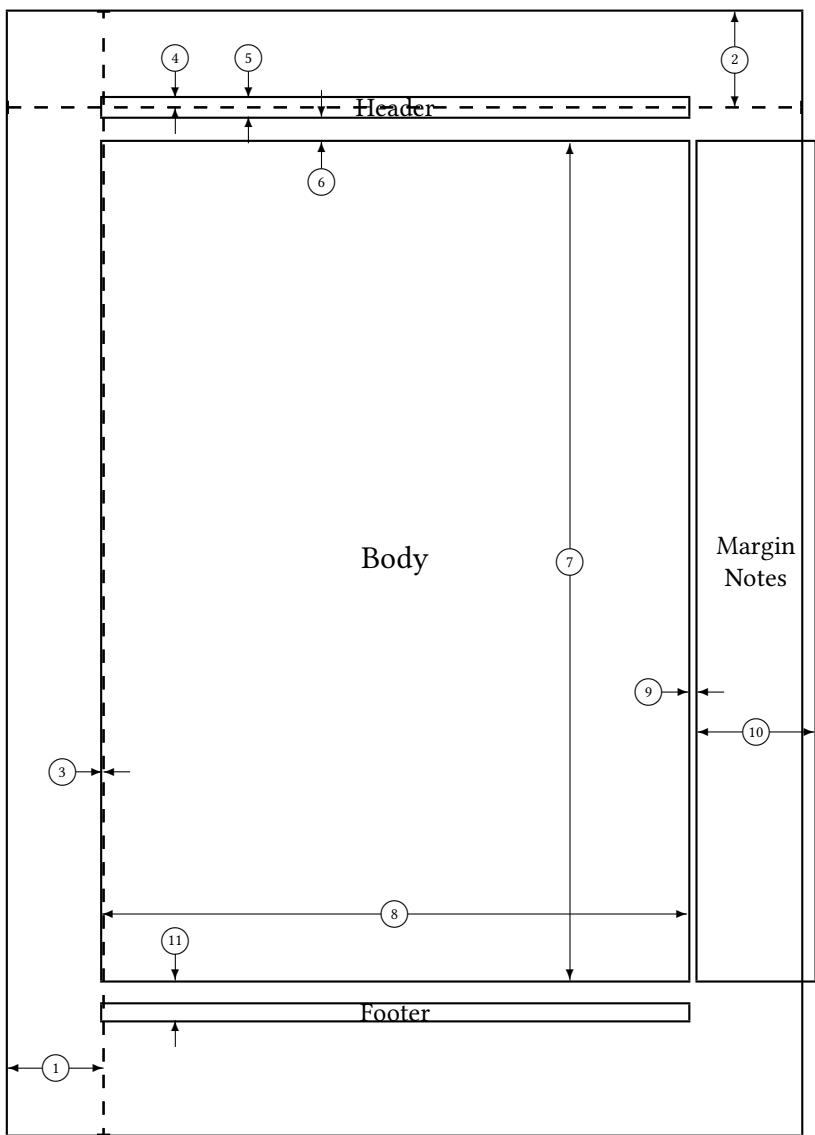
BIBLIOGRAPHY

- Álvarez-Herrero, A., Belenguer, T., Pastor, C., et al. Lithium niobate fabry-perot etalons in double-pass configuration for spectral filtering in the visible imager magnetograph imax for the sunrise mission. *Space Telescopes and Instrumentation I: Optical, Infrared, and Millimeter*, SPIE, 797–812 (2006).
- Álvarez-Herrero, A., Parejo, P. G., & Silva-López, M. Fine tuning method for optimization of liquid crystal based polarimeters. *Optics Express* **26**, 12038 (2018).
- Álvarez-Herrero, A., Parejo, P. G., & Silva-López, M. Fine tuning method for optimization of liquid crystal based polarimeters. *Optics Express* **26**, 12038 (2018).
- Álvarez Herrero, A., Garranzo-García, D., Núñez, A., et al. End-to-end tests of the tumag instrument for the sunrise iii mission. *Ground-based and Airborne Instrumentation for Astronomy IX*, SPIE, 791–802 (2022).
- Bailén, F. J., Orozco Suárez, D., Blanco Rodríguez, J., & del Toro Iniesta, J. C. Performance of Sequential Phase Diversity with Dynamical Solar Scenes. *Astrophys. J. Suppl.* **263**, 7 (2022).
- Bailén, F. J., Suárez, D. O., & del Toro Iniesta, J. On fabry–pérot etalon-based instruments. i. the isotropic case. *The Astrophysical Journal Supplement Series* **241**, 9 (2019).
- Barthol, P., Gandorfer, A., Solanki, S. K., et al. The sunrise mission. *Solar Physics* **268**, 1 (2011).
- Bellot Rubio, L., & Orozco Suárez, D. Quiet sun magnetic fields: an observational view. *Living Reviews in Solar Physics* **16**, 1 (2019).
- Bianda, M., Solanki, S., & Stenflo, J. Hanle depolarisation in the solar chromosphere. *Astronomy and Astrophysics*, v. 331, p. 760-770 (1998) **331**, 760 (1998).
- Childlaw, R., Devaney, A., & Gonsalves, R. Analytical studies of phase estimation techniques, volume 1. *Final Technical Report* (1979).
- Danilovic, S., Beeck, B., Pietarila, A., et al. Transverse component of the magnetic field in the solar photosphere observed by sunrise. *The Astrophysical Journal Letters* **723**, L149 (2010).

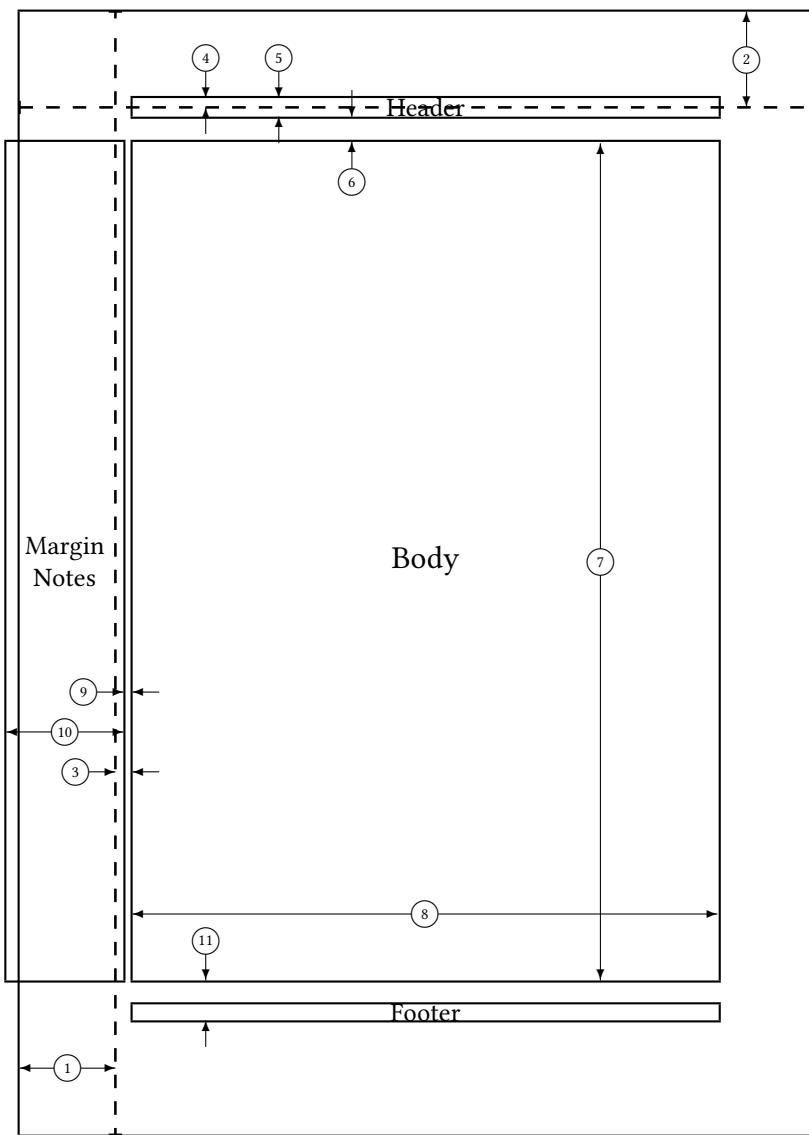
- del Toro Iniesta, J. C. (2003), Introduction to spectropolarimetry (Cambridge university press).
- del Toro Iniesta, J. C., & Collados, M. Optimum modulation and demodulation matrices for solar polarimetry. *Applied Optics* **39**, 1637 (2000).
- del Toro Iniesta, J. C., & Ruiz Cobo, B. Inversion of the radiative transfer equation for polarized light. *Living Reviews in Solar Physics* **13**, 4 (2016).
- Elmore, D. F., Lites, B. W., Tomczyk, S., et al. Advanced stokes polarimeter: a new instrument for solar magnetic field research. *Polarization analysis and measurement*, SPIE, 22–33 (1992).
- Gaskill, J. D. (1978), Linear systems, Fourier transforms, and optics, Vol. 56 (John Wiley & Sons).
- Gizon, L., Birch, A. C., & Spruit, H. C. Local helioseismology: three-dimensional imaging of the solar interior. *Annual Review of Astronomy and Astrophysics* **48**, 289 (2010).
- González, N. B., & Kneer, F. Narrow-band full stokes polarimetry of small structures on the sun with speckle methods. *Astronomy & Astrophysics* **480**, 265 (2008).
- Guglielmino, S., Pillet, V. M., Bonet, J., et al. The frontier between small-scale dipoles and ephemeral regions in the solar photosphere: emergence and decay of an intermediate-scale dipole observed with sunrise/imax. *The Astrophysical Journal* **745**, 160 (2012).
- Helstrom, C. W. Image restoration by the method of least squares. *Josa* **57**, 297 (1967).
- Hotta, H., Rempel, M., & Yokoyama, T. Efficient small-scale dynamo in the solar convection zone. *The Astrophysical Journal* **803**, 42 (2015).
- Huang, K.-Y., Chia, C.-M., & Chang, E. Optical resolution measurement system for small lens by using slanted-slit method. *Optical Measurement Systems for Industrial Inspection VIII*, SPIE, 705–712 (2013).
- Kosugi, T., Matsuzaki, K., Sakao, T., et al. The hinode (solar-b) mission: An overview. *Solar Physics* **1**, 3 (2007).
- Lakshminarayanan, V., & Fleck, A. Zernike polynomials: a guide. *Journal of Modern Optics* **58**, 545 (2011).
- Lites, B., Elmore, D., & Streander, K. The solar-b spectro-polarimeter. *Advanced solar polarimetry—theory, observation, and instrumentation*, 33 (2001).
- Lites, B. W. Rotating waveplates as polarization modulators for stokes polarimetry of the sun: evaluation of seeing-induced crosstalk errors. *Applied Optics* **26**, 3838 (1987).

- Martínez Pillet, V., Collados, M., Sánchez Almeida, J., et al. Lpsp & tip: full stokes polarimeters for the canary islands observatories. *High Resolution Solar Physics: Theory, Observations, and Techniques*, 264 (1999).
- Müller, D., Cyr, O. S., Zouganelis, I., et al. The solar orbiter mission-science overview. *Astronomy & Astrophysics* **642**, A1 (2020).
- Orozco Suárez, D., Álvarez García, D., López Jiménez, A. C., et al. Spgcam: A specifically tailored camera for solar observations. *Frontiers in Astronomy and Space Sciences* **10**, 1167540 (2023).
- Parker, E. Magnetic neutral sheets in evolving fields-part two-formation of the solar corona. *Astrophysical Journal, Vol. 264, P. 642, 1983* **264**, 642 (1983).
- Paxman, R. G., Schulz, T. J., & Fienup, J. R. Joint estimation of object and aberrations by using phase diversity. *JOSA A* **9**, 1072 (1992).
- Petrovay, K., & Szakaly, G. The origin of intranetwork fields: a small-scale solar dynamo. *Astronomy and Astrophysics, Vol. 274, p. 543 (1993)* **274**, 543 (1993).
- Pillet, V. M., Bonet, J. A., Collados, M. V., et al. The imaging magnetograph experiment for the sunrise balloon antarctica project. *Optical, Infrared, and Millimeter Space Telescopes*, SPIE, 1152–1164 (2004).
- Pontin, D. I., & Hornig, G. The parker problem: existence of smooth force-free fields and coronal heating. *Living Reviews in Solar Physics* **17**, 5 (2020).
- Rempel, M., Bhatia, T., Bellot Rubio, L., & Korpi-Lagg, M. J. Small-scale dynamos: from idealized models to solar and stellar applications. *Space Science Reviews* **219**, 36 (2023).
- Sánchez, A., Gonzalo, A., Garranzo, D., et al. High precision and thermally controlled filter wheel. *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation V*, SPIE, 1120–1135 (2022).
- Scharmer, G. Comments on the optimization of high resolution fabry-pérot filtergraphs. *Astronomy & Astrophysics* **447**, 1111 (2006).
- Shibata, K., Nakamura, T., Matsumoto, T., et al. Chromospheric anemone jets as evidence of ubiquitous reconnection. *Science* **318**, 1591 (2007).
- Snyder, A. W. (1975), in Photoreceptor optics (Springer), 38–55.
- Solanki, S., Riethmüller, T., Barthol, P., et al. The second flight of the sunrise balloon-borne solar observatory: overview of instrument updates, the flight, the data, and first results. *The Astrophysical Journal Supplement Series* **229**, 2 (2017).
- Solanki, S. K., del Toro Iniesta, J., Woch, J., et al. The polarimetric and helioseismic imager on solar orbiter. *Astronomy & Astrophysics* **642**, A11 (2020).

- Stokes, G. G. On the composition and resolution of streams of polarized light from different sources. *Transactions of the Cambridge Philosophical Society* **9**, 399 (1851).
- Thompson, W. Coordinate systems for solar image data. *Astronomy & Astrophysics* **449**, 791 (2006).
- Tsuneta, S., Ichimoto, K., Katsukawa, Y., et al. The solar optical telescope for the hinode mission: an overview. *Solar Physics* **249**, 167 (2008).
- Tyson, R. K. (2000), Introduction to adaptive optics, Vol. 41 (SPIE press).
- Uitenbroek, H. The accuracy of the center-of-gravity method for measuring velocity and magnetic field strength in the solar photosphere. *The Astrophysical Journal* **592**, 1225 (2003).
- Van Noort, M., & van der Voort, L. R. Stokes imaging polarimetry using image restoration at the swedish 1-m solar telescope. *Astronomy & Astrophysics* **489**, 429 (2008).
- Vargas Dominguez, S. Study of horizontal flows in solar active regions based on high-resolution image reconstruction techniques. *Ph. D. Thesis* (2009).
- Vargas Dominguez, S. Study of horizontal flows in solar active regions based on high-resolution image reconstruction techniques. *Ph. D. Thesis* (2009).
- Wedemeyer-Böhm, S., Scullion, E., Steiner, O., et al. Magnetic tornadoes as energy channels into the solar corona. *Nature* **486**, 505 (2012).
- Zernike, F. Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. *Monthly Notices of the Royal Astronomical Society*, Vol. 94, p. 377-384 **94**, 377 (1934).
- Zwaan, C. Elements and patterns in the solar magnetic field. IN: *Annual review of astronomy and astrophysics. Volume 25 (A88-13240 03-90)*. Palo Alto, CA, Annual Reviews, Inc., 1987, p. 83-111. **25**, 83 (1987).



1 one inch + \hoffset	2 one inch + \voffset
3 \oddsidemargin = -1pt	4 \topmargin = -7pt
5 \headheight = 14pt	6 \headsep = 19pt
7 \textheight = 631pt	8 \textwidth = 441pt
9 \marginparsep = 7pt	10 \marginparwidth = 88pt
11 \footskip = 30pt	\marginparpush = 7pt (not shown)
\hoffset = 0pt	\voffset = 0pt
\paperwidth = 597pt	\paperheight = 845pt



1 one inch + \hoffset	2 one inch + \voffset
3 \evensidemargin = 13pt	4 \topmargin = -7pt
5 \headheight = 14pt	6 \headsep = 19pt
7 \textheight = 631pt	8 \textwidth = 441pt
9 \marginparsep = 7pt	10 \marginparwidth = 88pt
11 \footskip = 30pt	\marginparpush = 7pt (not shown)
\hoffset = 0pt	\voffset = 0pt
\paperwidth = 597pt	\paperheight = 845pt