Boletim da Sociedade Astronômica Brasileira, **36**, no. 1, 65-69

SPARC4 Pipeline: automated data reduction at OPD for diverse scientific cases — from planets to galaxies

E. Martioli[1], C. V. Rodrigues[2], J. C. N. Campagnolo[3], F. J. Jablonski[2], A. C. Mattiuci[2], W. Schlindwein[2], F. Falkenberg[2], M. M. C. Mello[2], L. Fraga[1], L. de Almeida[1], I. J. Lima[1], & F. Monteiro[1]

[1] Laboratório Nacional de Astrofísica, Rua dos Estados Unidos, 154, Bairro das Nações, Itajubá, Minas Gerais, Brazil. e-mail: emartioli@lna.br
[2] Instituto Nacional de Pesquisas Espaciais, Avenida dos Astronautas, 1.758, Jardim Granja, São José dos Campos, São Paulo, Brazil
[3] Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rio de Janeiro, Rio de Janeiro, Brazil

Abstract. We introduce the SPARC4 Pipeline, an automated data reduction software package designed to process photometric and polarimetric data in four photometric bands (*g, r, i, z*), obtained with the new instrument SPARC4 installed on the 1.6 m telescope at the Pico dos Dias Observatory. The software is open source, written in Python, based on the ASTROPOP package and other widely used astronomical packages. We present the main strategies and techniques used for calibrating the images and obtaining aperture photometry and dual-beam polarimetry of point-like sources. Science-grade reduction products for observations obtained during commissioning and science verification runs are presented, covering a broad range of scientific applications. Specifically, we present the following results: astrometric and photometric calibration of a crowded field; high-precision, high-cadence photometric and polarimetric time series for a δ-Scuti variable star and a transit of the exoplanet WASP-78 b; polarimetric results from a standard star in the Flame Nebula; and a deep image of an extended object—the spiral galaxy pair NGC-4045.

Resumo. Apresentamos o SPARC4 Pipeline, um pacote de software automatizado para redução de dados, desenvolvido para processar dados fotométricos e polarimétricos em quatro bandas (*g, r, i, z*), obtidos com o novo instrumento SPARC4, instalado no telescópio de 1,6 m no Observatório do Pico dos Dias. O software é código aberto, escrito em Python, e baseado no pacote ASTROPOP e em outros pacotes astronômicos amplamente utilizados. Descrevemos as principais estratégias e técnicas empregadas para calibrar as imagens, realizar fotometria de abertura e realizar polarimetria de feixe duplo em fontes pontuais. São apresentados também os produtos de redução de nível científico para observações realizadas durante as missões de comissionamento e de verificação científica do instrumento, que abrangem uma ampla gama de aplicações científicas. Especificamente, apresentamos os seguintes resultados: calibração astrométrica e fotométrica de um campo rico; séries temporais fotométricas e polarimétricas de alta precisão e alta cadência para uma estrela variável δ-Scuti e um trânsito do exoplaneta WASP-78 b; resultados polarimétricos de uma estrela padrão na Flame Nebula; e uma imagem profunda de um objeto estendido — o par de galáxias espirais NGC-4045.

Keywords. Instrumentação: fotômetros, polarímetros, SPARC4 – Técnicas: processamento de imagens

1. Introduction

The Simultaneous Polarimeter and Rapid Camera in Four Bands (SPARC4, Rodrigues et al. 2012, 2024) is a new instrument installed on the Perkin-Elmer 1.6-meter telescope at the National Laboratory for Astrophysics (LNA) in the Pico dos Dias Observatory (OPD). It provides world-class rapid photometry and polarimetry in the four SDSS bands simultaneously, with a field of view of 5.7×5.7 arcmin^2. Developed by the National Institute for Space Research (INPE) in collaboration with LNA and other Brazilian universities and institutes, SPARC4 enhances observation efficiency and employs differential techniques in both photometry and polarimetry. These techniques are robust even under the observatory's variable weather conditions and are tailored to meet the primary scientific interests of its users.

Traditional data processing methods may struggle to handle the high data flow and multiple observation modes of SPARC4. A solution to improve both the observational and analytical efficiency is the implementation of automated data processing software, commonly referred to as a reduction pipeline. The advantages of using a pipeline are numerous, with key benefits including the uniformity of data processing, repeatability, ease of testing and evaluating the impact of different processing approaches, and the ability to establish an archive of processed data and a long-term calibration plan for the instrument. To this end, the SPARC4 instrument team has grouped a team of community experts to collaborate on developing the SPARC4 Pipeline.

2. The pipeline concept

The SPARC4 Pipeline is a suite of Python routines designed to process data from photometric and polarimetric observations taken with the SPARC4 instrument. It is based on the ASTROnomical Polarimetry and Photometry pipeline package (ASTROPOP, Campagnolo 2018, 2019)[1] and the `Astropy`[2] package (Astropy Collaboration et al. 2013), and it leverages widely used numerical and scientific libraries, such as NumPy[3] and SciPy[4], among others. The SPARC4 Pipeline codes are open-source and documented on GitHub[5].

The pipeline includes a main module, `sparc4_mini_pipeline.py`, located in the `scripts` directory. This module can be run from the command line and enables the automatic reduction of data from the four SPARC4 channels. Users can provide a configuration file using the

[1] https://github.com/juliotux/astropop/tree/main/astropop
[2] https://www.astropy.org/
[3] https://numpy.org/
[4] https://scipy.org/
[5] https://github.com/edermartioli/sparc4-pipeline

-params option, which overrides the default reduction parameters. This allows some parameters to be adjusted as needed, prioritizing either accuracy or performance. The pipeline also allows users to specify a target list file in the -target_list option. The purpose of the target list is to ensure that specified targets are included in the photometric and polarimetric catalogs generated by the pipeline, alongside other sources automatically detected. The pipeline automatically identifies the directory structure for the raw data obtained from the four channels of SPARC4 and initiates the reduction of each channel independently, though not simultaneously. The pipeline option -channels allows users to specify the sequence of channels to be reduced, with the default sequence set to "1,2,3,4". To parallelize the reduction, users can launch four separate terminal instances, each processing a different channel.

The pipeline organizes its five main libraries in the sparc4 directory, each with specific functionality, as summarized below:

- pipeline_lib.py: core library with the main pipeline execution routines and functions;
- db.py: library with routines for creating a simple database to manage input data;
- utils.py: utility library with routines for data reduction;
- products.py: library of input/output routines for SPARC4 reduction products;
- product_plots.py: library of routines for generating diagnostic plots of the reduction products.

Although the data processing is serialized in a main script, all of the pipeline routines can be executed independently in a Python session or Jupyter notebook by importing the appropriate libraries. This allows for customized reductions of SPARC4 data or data from any other instrument using similar techniques. The pipeline package also includes a notebook directory, which contains Jupyter notebooks with markdown examples.

3. The reduction methods

The pipeline reads the FITS images and creates a csv table as a night database per channel, which includes all image file names and their corresponding important keywords. This table allows an automatic identification of images by type (ZERO, FLAT, FOCUS, OBJECT), by science object name, by observing mode (PHOT or POLAR), by CCD mode, by waveplate type and position, etc. The pipeline then runs data reduction for each scientific object across all modes, producing science-grade products. These products are saved in FITS format, and visualization tools are provided to facilitate data access. In summary, the following reduction steps are performed for each channel and instrument mode for a given night of observations:

1. build a database;
2. combine zero images into a master zero;
3. combine flat-field images into a normalized master flat. A master flat is created for each instrument mode (PHOT or POLAR), for each retarder waveplate mode (L2 or L4), and for each position angle of the waveplate;
4. combine calibrated science images of the same object and same mode to create a stack image;
5. run a source detection algorithm to build a catalog of point-like sources in the stack image. Aperture photometry is performed on these sources, and an astrometric solution is obtained and saved in the WCS keywords of the stack image header. The photometry is calculated for a set of apertures and saved in different catalog extensions;
6. in POLAR mode, the two beams are matched in the stack frame and the photometric information for each beam are saved as separate catalog extensions;
7. each science image is corrected for CCD gain, subtracted from the master zero, and divided by the master flat, with all calibrations matching the instrument mode of the science image. The frame is then registered to the stack via cross-correlation. The stacked image catalog serves as a reference, with the frame offset applied before performing photometry on each source, ensuring consistency across all images. An astrometric solution is derived for each frame by applying an offset to the astrometric solution in the stack;
8. each science frame is saved as a processed image with pixel values expressed in units of electrons. Alongside these images, corresponding catalogs are generated, containing photometric measurements converted to instrumental magnitudes. The conversion uses the formula $-2.5 \log_{10}(f)$, where f represents the measured sum of electron counts within an aperture, divided by the exposure time. As a result, f is expressed in units of electrons per second;
9. in POLAR mode, the polarimetric sequences are identified, and dual-beam polarimetry is calculated either in linear polarization ($\lambda/2$) or linear and circular polarization ($\lambda/4$) modes using the methods developed by Magalhaes, Benedetti, & Roland (1984) and Rodrigues, Cieslinski, & Steiner (1998), and implemented by Campagnolo (2018) and Mattiuci et al. (2024). A polarimetry product encapsulates the measurements in a FITS file;
10. a photometric time series is produced, including photometry from all sources in the catalog. These data are also saved as a separate FITS product;
11. a polarimetric time series is also obtained in case of POLAR mode and the results are saved in a FITS product.

4. Scientific products

The SPARC4 pipeline reduces the raw data and saves the product of this reduction in FITS files. The description of these products is available in the SPARC4 Pipeline documentation. The notebook sp4_products.ipynb provides information and examples to familiarize the user with these products. The product names follow a convention adopted by the pipeline, as described in Table 1, and the headers of these images preserve the information from the raw image, in addition to adding information about the reduction procedures and results. The scientific products of the pipeline are summarized as follows:

- **Master Zero** (*_MasterZero.fits): combined image of zero exposures.
- **Master Flat** (*_MasterDomeFlat.fits): combined image of dome flat exposures.
- **Processed science image** (*_proc.fits) : calibrated science image (gain corrected, bias subtracted and flat corrected) with catalogs of detected sources and their aperture photometry for multiple apertures, and an updated astrometric solution saved in the WCS.
- **Science stack image** (*_stack.fits) : combined image of calibrated science exposures and photometric catalogs of sources.
- **Polarimetry** (**$\lambda/2$ or $\lambda/4$**) (*_polar.fits): polarimetry of all sources in the catalog for a polarimetric sequence.
- **Photometric time series** (*_lc.fits) : time series for photometric quantities from all sources in the catalog for all photometric apertures.

- **Polarimetric time series** (`*_ts.fits`) : time series for photometric and polarimetric quantities from all sources in the catalog.

5. Science verification

Whether conducting photometry or polarimetry through static imaging or time series, the scientific applications of the SPARC4 instrument are diverse. In this section, we illustrate the use of the pipeline with a few examples—though not an exhaustive list—of SPARC4's broad scientific applications, using data collected during the instrument's science verification and commissioning runs.

5.1. Astrometric and photometric calibration of SPARC4

Crowded fields have been observed during commissioning of SPARC4 to be used as photometric and astrometric standards (see Schlindwein et al. 2023). These fields are densely populated with stars spanning a range of magnitudes without significant overlap of sources. In particular, these fields include sources from calibrated surveys, such as SkyMapper[6] (Keller et al. 2007) and Gaia (Gaia Collaboration 2020), which can be used as references to calibrate our data. Here, we present results from observations of CR3, a field centered at coordinates 14:38:07.5 -42:35:12.6 (ICRS, J2000), observed on 2023-06-05 in PHOT mode and processed by the SPARC4 Pipeline. These data are available in the minidata package, which is linked on the pipeline's GitHub page.

Figure 1 shows the four stack image products of CR3 displayed with `SAOImageDS9`[7], with some sources retrieved from the Gaia DR3 catalog overlaid. The astrometric accuracy of these images in both directions is approximately 0.5 arcseconds, which is sufficient for matching objects with catalogs. Ongoing work aims to improve this accuracy.

Figure 2 shows the photometric calibration using detected sources in the CR3 field. The instrumental magnitudes are plotted against the SkyMapper magnitudes in the corresponding bands, and a linear fit is performed for each channel, resulting in zero-point magnitudes of -24.44 ± 0.09, -25.04 ± 0.03, -24.31 ± 0.04, and -23.66 ± 0.06 for the g, r, i, and z bands, respectively. The fit slopes deviate from the identity by 1% or less, indicating good agreement over the range of magnitudes (11 to 17 mag) covered by these data. A more detailed analysis of the SPARC4 photometric calibration, including higher-order terms, atmospheric extinction due to variable airmass, and color effects, will be presented in a forthcoming paper.

5.2. δ-Scuti variable star

On the night of 2024-06-18, while observing a potential exoplanet transit event under program OP2024A-004 (PI: E. Martioli), we observed the star ASASSN-V J140530.18+165743.8[8], which was found to be variable. Later, this object was identified as a V=13.8 pulsating star of δ-Scuti type, with a period of 1.44 hours. As illustrated in Figure 3, the differential photometry time series, or "light curves," obtained with the SPARC4 Pipeline achieve millimagnitude precision over durations of more than 4 hours, enabling us to capture the star's periodic variability in great detail, with amplitude variations on the order of 0.1 mag.

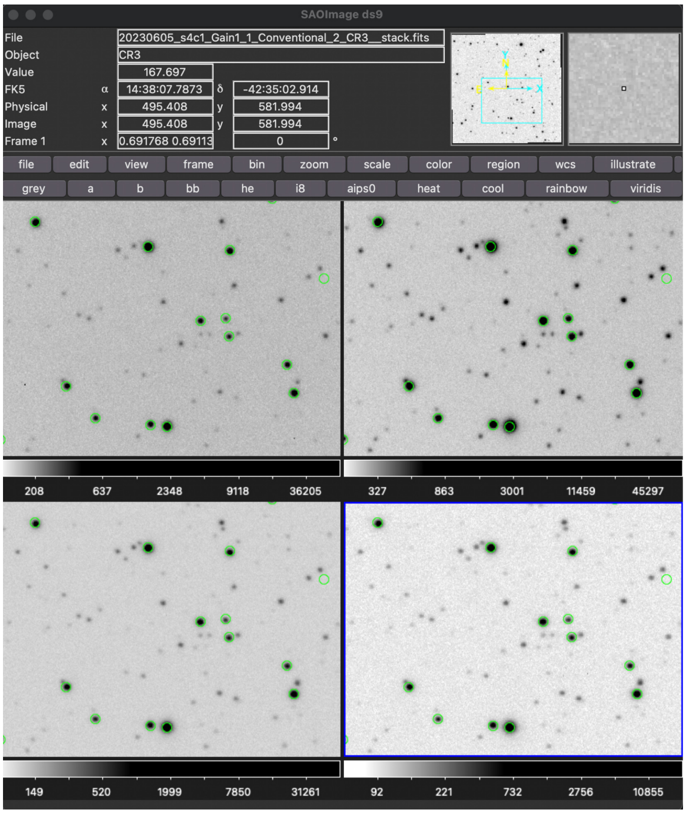

Figure 1. Stack image products of the field CR3, observed in the four channels of SPARC4, and displayed with `SAOImageDS9`. Green circles highlight the brightest sources retrieved from the Gaia DR3 catalog, showing a good match with the stars detected in the images.

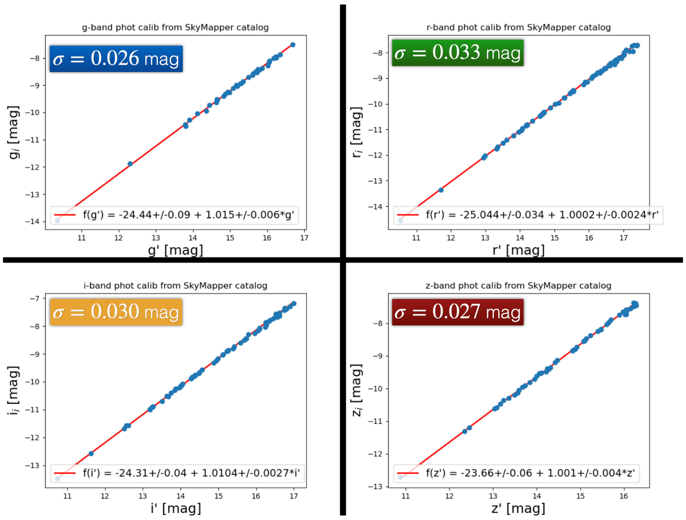

Figure 2. Photometric calibration of SPARC4 using magnitude data from the CR3 field, calculated by the SPARC4 Pipeline and compared against SkyMapper magnitudes in the same bands. The red lines show linear fits to the data, with the fit slopes also displayed. The standard deviations of the residuals are highlighted in each panel.

[6] https://skymapper.anu.edu.au/
[7] https://sites.google.com/cfa.harvard.edu/saoimageds9
[8] https://www.aavso.org/vsx/index.php?view=detail.top&oid=1509843

Table 1. File name convention for the SPARC4 Pipeline scientific products. The bracketed variables represent specific components, including `NIGHTDIR` (the night directory name), `CHANNEL` (the channel index, which can be 1, 2, 3, or 4), `DETECTORMODE` (keywords that identify the detector mode), `ORIGFILENAME` (the original file basename without the `.fits` extension), `OBJECT` (the value of the header keyword `OBJECT`, with spaces removed), and `POLARMODE` (the polarimetry mode, either L2 or L4).

Product	File name convention
Master Zero	{NIGHTDIR}_s4c{CHANNEL}_{DETECTORMODE}_MasterZero.fits
Master Flat	{NIGHTDIR}_s4c{CHANNEL}_{DETECTORMODE}_MasterDomeFlat.fits
Processed science image	{ORIGFILENAME}_proc.fits
Science stack image	{NIGHTDIR}_s4c{CHANNEL}_{DETECTORMODE}_{OBJECT}_stack.fits
Polarimetry ($\lambda/2$ or $\lambda/4$)	{ORIGFILENAME}_{POLARMODE}_polar.fits
Photometric time series	{NIGHTDIR}_s4c{CHANNEL}_{OBJECT}_lc.fits
Polarimetric time series	{NIGHTDIR}_s4c{CHANNEL}_{OBJECT}_polar_ts.fits

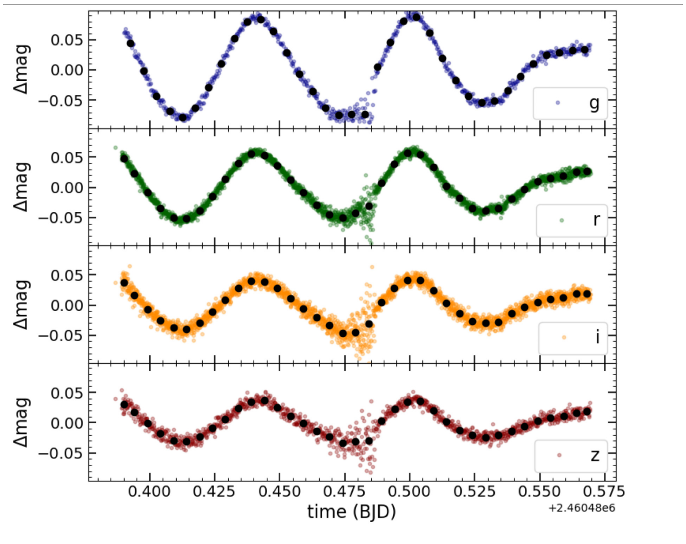

Figure 3. Differential photometry time series data of the δ-Scuti variable star ASASSN-V J140530.18+165743.8, observed in the four channels of SPARC4.

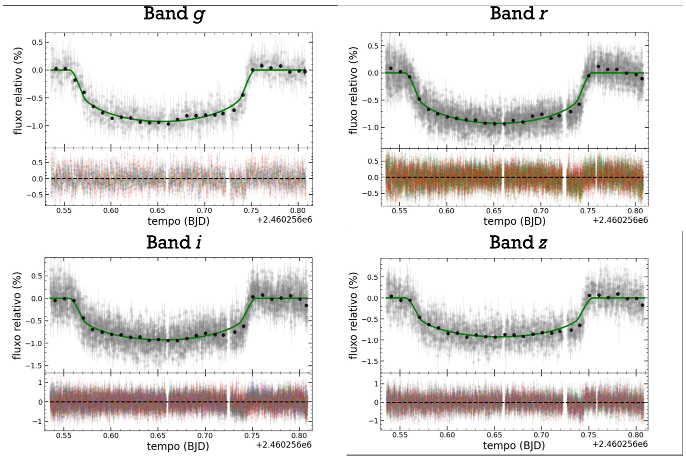

Figure 4. Transit of the exoplanet WASP-78 b observed with SPARC4 during the commission run on November 7, 2023. In the plot, grey points represent the differential photometry of WASP-78 with respect to other comparison stars in the field, while the green lines show the transit model. The colored points in the bottom part of each panel display the differential photometry for comparison stars.

5.3. Transit of the exoplanet WASP-78 b

A transit of the exoplanet WASP-78 b, which orbits a V=12.0 F8 star (Smalley et al. 2012), was observed with SPARC4 on 2023-11-07 in POLAR L2 mode, and the data were processed using the SPARC4 Pipeline. Figure 4 shows the light curve data along with a fitted transit model. The transit of WASP-78 b has a predicted depth of 0.63%. The final root mean square (RMS) of the photometric residuals is approximately 0.1%, demonstrating that the pipeline achieves a photometric precision in time-series data, even from polarimetric mode, that is sufficient for detecting and characterizing transiting exoplanets. A more detailed analysis of this transit along with several other transits of exoplanets will be presented in a dedicated paper.

5.4. Polarimetry of NGC 2024-1 in the Flame Nebula

Here, we present an example of linear polarization measurements of the point-like source NGC2024-1 in the Flame Nebula (NGC2024), performed using the pipeline. This field was observed by the SPARC4 team on 2022-11-16, during commissioning. The polarization of stars in this field are well known and stable, allowing us to calibrate the SPARC4 measurements against values reported in the literature. The scope of this paper is limited to present the pipeline results, while a separate paper in preparation will focus on the polarimetric characterization of SPARC4 (see Mattiuci et al. 2024). Figure 5 shows the four images of NGC 2024, with duplicated sources for dual beam differential polarimetry technique (Magalhaes, Benedetti, & Roland 1984). Figure 5 also shows the linear polarization measurements (L2 mode, Stokes Q and U) of the star NGC2024-1 at the center of the image. The polarimetry product of the pipeline stores measurements of the polarized flux fraction, $Z(\phi) = \frac{f_{\parallel}(\phi)-f_{\perp}(\phi)}{f_{\parallel}(\phi)+f_{\perp}(\phi)}$, as a function of the 16 position angles, ϕ, of the rotating $\lambda/2$-wave plate. The polarization quantities are then fitted to these data for all sources in the field, and the results are also stored in the polarimetry product. For the brightest source, NGC-2024-1, we obtained a total polarization of 9.06 ± 0.03%, 9.86 ± 0.02%, 9.53 ± 0.01%, 8.47 ± 0.01% for bands g, r, i, and z, respectively.

5.5. A snapshot of the pair of spiral galaxies NGC 4045

The SPARC4 instrument team has released a call for proposals for Science Verification (SV)[9] with the instrument during its commissioning phase in 2023. Here, we present a snapshot of observations of a pair of spiral galaxies, NGC 4045, obtained under program OPD2024A-017 (PI: Menéndez-Delmestre, K.) on the night of 2024-05-04 and reduced by the SPARC4 Pipeline. Figure 6 shows a composite false-color image of NGC 4045, created by combining the stacked images from the four channels

[9] https://www.gov.br/lna/pt-br/composicao-1/coast/obs/opd/informacoes/2023A-chamadaSPARC4

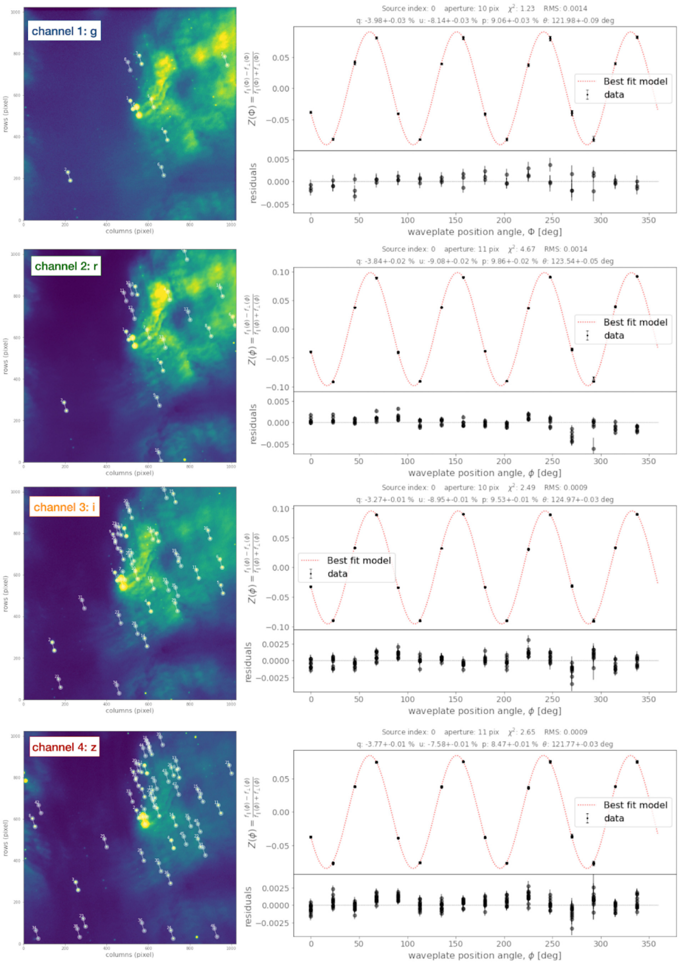

Figure 5. Polarimetry of NGC 2024-1 in the Flame Nebula. The left panels display SPARC4 polarimetric images in four bands, highlighting duplicated sources from dual-beam polarimetry. Photometric and polarimetric quantities are calculated and stored for each marked source. The right panels show $Z(\phi)$ as a function of the waveplate position angle ϕ, along with the best-fit polarization model for the brightest star, NGC 2024-1.

projected to the same reference frame, which are also displayed in this figure. Each stack is composed of 28 frames, each with an integration time of 150 s. The matching features for different colors in this image underscores the quality of reduction.

6. Conclusions

In this paper, we present the SPARC4 Pipeline, an open-source suite of Python routines designed to automatically reduce SPARC4 data and produce science-grade products. We provide a brief overview of the pipeline's concept and the methods used in the reduction process. A forthcoming paper (Martioli et al., in preparation) will offer a more detailed discussion of the pipeline's methods and scientific performance. To showcase the performance of both the SPARC4 instrument and the pipeline's ability to process its data, we present several results from different scientific cases. Starting with the crowded-field CR3, we demonstrate that the pipeline can achieve astrometric accuracy of 0.5 arcseconds and photometric accuracy at the percent level. We also present a time series of a δ-Scuti variable star with millimagnitude precision, demonstrating the pipeline's high-precision differential photometry capabilities at

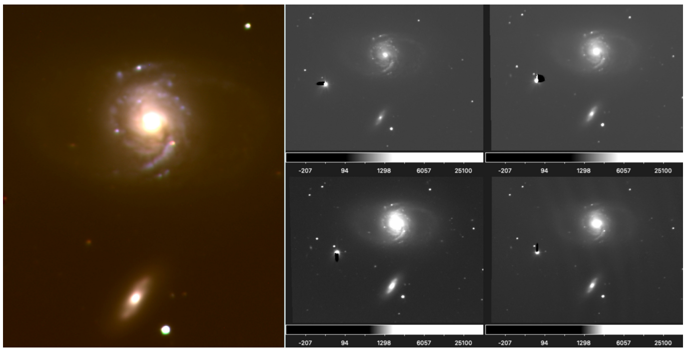

Figure 6. A pair of spiral galaxies, NGC 4045, observed in the four channels of SPARC4 (right panels), alongside a composite false-color image (left panel) created by combining the four images in different bands.

high cadence. Additionally, we show a transit of the exoplanet WASP-78 b, with a transit depth of approximately 0.7%, detected via differential photometry time series in polarimetric mode. Observations of the polarimetric standard star NGC2024-1 demonstrate the SPARC4 Pipeline's ability to obtain accurate polarization measurements, including from multiple sources within the same field. Finally, we present observations of the pair of spiral galaxies NGC 4045, showcasing the pipeline's ability to reduce deep imaging data obtained in photometric mode.

Acknowledgements. This paper uses data obtained with the Simultaneous Polarimeter and Rapid Camera in 4 bands (SPARC4), installed on the 1.6-m telescope at the Observatório do Pico dos Dias (OPD), managed by the Laboratório Nacional de Astrofísica (LNA) under the Ministério da Ciência, Tecnologia e Inovação (Brazil). SPARC4 was funded by Financiadora de Estudos e Projetos - Finep (Proc: 0/1/16/0076/00), Agência Espacial Brasileira - AEB (PO 20VB.0009), Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (Grant 2010/01584-8), Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG (APQ-00193-15 & APQ-02423-21), Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Grant #420812/2018-0) and INCT-Astrofísica. E.M. acknowledges funding from FAPEMIG under project number APQ-02493-22 and a research productivity grant number 309829/2022-4 awarded by the CNPq. C.V.R. thanks the support from AEB (PO 20VB.0009) and CNPq (Proc: 310930/2021-9).

References

Astropy Collaboration, Robitaille T. P., Tollerud E. J., Greenfield P., Droettboom M., Bray E., Aldcroft T., et al., 2013, A&A, 558, A33. doi:10.1051/0004-6361/201322068
Campagnolo J. C. N., 2018, ascl.soft. ascl:1805.024
Neves Campagnolo, J. C. 2019, PASP, 131, 024501. doi:10.1088/1538-3873/aaecc2
Gaia Collaboration, 2020, yCat, 1350
Keller S. C., Schmidt B. P., Bessell M. S., Conroy P. G., Francis P., Granlund A., Kowald E., et al., 2007, PASA, 24, 1. doi:10.1071/AS07001
Magalhaes A. M., Benedetti E., Roland E. H., 1984, PASP, 96, 383. doi:10.1086/131351
Mattiuci A. C., Martioli E., Vilega Rodrigues C., Campagnolo J. C. N., Schlindwein W., Falkenberg F., Bernardes D., 2024, BoSAB, 35, 199
Rodrigues C. V., Cieslinski D., Steiner J. E., 1998, A&A, 335, 979. doi:10.48550/arXiv.astro-ph/9805193
Rodrigues C. V., Taylor K., Jablonski F. J., Assafin M., Carciofi A., Cieslinski D., Costa J. E. R., et al., 2012, SPIE, 8446, 844626. doi:10.1117/12.924976
Rodrigues C. V., Gneiding C. D., Almeida L., Alves A. L. M., Andrade L. B. P., Bernardes D., Dos Santos J. B., et al., 2024, BASBr, 35, 44
Schlindwein W., Bernardes D. V., Martioli E., Rodrigues C. V., Campagnolo J. C. N., Figueiredo A. C. M., Fraga L., et al., 2023, BASBr, 34, 281
Smalley B., Anderson D. R., Collier-Cameron A., Doyle A. P., Fumel A., Gillon M., Hellier C., et al., 2012, A&A, 547, A61. doi:10.1051/0004-6361/201219731