



HAL
open science

ShapePipe: A modular weak-lensing processing and analysis pipeline

S. Farrens, A. Guinot, M. Kilbinger, T. Liaudat, L. Baumont, X. Jimenez, A. Peel, A. Pujol, M. Schmitz, J.-L. Starck, et al.

► **To cite this version:**

S. Farrens, A. Guinot, M. Kilbinger, T. Liaudat, L. Baumont, et al.. ShapePipe: A modular weak-lensing processing and analysis pipeline. *Astronomy and Astrophysics - A&A*, 2022, 664, pp.A141. 10.1051/0004-6361/202243970 . hal-03720964

HAL Id: hal-03720964

<https://hal.science/hal-03720964v1>

Submitted on 24 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SHAPEPIPE: A modular weak-lensing processing and analysis pipeline

S. Farrens¹, A. Guinot², M. Kilbinger¹, T. Liaudat¹, L. Baumont¹, X. Jimenez³, A. Peel⁴, A. Pujol¹,
M. Schmitz⁵, J.-L. Starck¹, and A. Z. Vitorelli¹

¹ AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, 91191 Gif-sur-Yvette, France
e-mail: samuel.farrens@cea.fr

² Université de Paris, CNRS, Astroparticule et Cosmologie, 75013 Paris, France

³ Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, 91190 Gif-sur-Yvette, France

⁴ Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland

⁵ Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France

Received 6 May 2022 / Accepted 20 June 2022

ABSTRACT

We present the first public release of SHAPEPIPE, an open-source and modular weak-lensing measurement, analysis, and validation pipeline written in Python. We describe the design of the software and justify the choices made. We provide a brief description of all the modules currently available and summarise how the pipeline has been applied to real Ultraviolet Near-Infrared Optical Northern Survey data. Finally, we mention plans for future applications and development. The code and accompanying documentation are publicly available on GitHub.

Key words. gravitational lensing: weak – methods: data analysis

1. Introduction

Weak gravitational lensing, the apparent distortion of the shapes of galaxies caused by the bending of light by mass along the line of sight, has been demonstrated to be a powerful probe of cosmology (Kilbinger 2015; Mandelbaum 2018). However, numerous steps are required in order to go from raw survey data to competitive constraints on cosmological parameters. Recent surveys, such as the Canada-France-Hawai'i Telescope Legacy Survey (Erben et al. 2013), the Hyper Suprime Cam (HSC) survey (Mandelbaum et al. 2018), the Kilo-Degree Survey (Kuijken et al. 2019), and the Dark Energy Survey (DES; Gatti et al. 2021), have carried out detailed weak-lensing analyses. Upcoming surveys, such as *Euclid* (Laureijs et al. 2011) and the *Vera C. Rubin* Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019), will also aim to tighten cosmological constraints with weak lensing. It is clear that weak lensing remains at the forefront of cosmological studies, and hence tools for weak-lensing analysis will remain in demand for years to come.

There are various pipelines available for simulating weak-lensing data (e.g. Kiessling et al. 2011; Harnois-Déraps & van Waerbeke 2015; Petri 2016). Various data processing pipelines also exist (e.g. Bosch et al. 2018), but these are generally very survey-specific and often private. We believe this leaves an ample gap in the market for an open-source end-to-end weak-lensing pipeline.

We present the first public release of SHAPEPIPE, an open-source and modular weak-lensing measurement, analysis, and validation pipeline written in Python. An earlier version of this pipeline was applied to real survey data in Guinot et al. (2022) with promising results.

The current version of SHAPEPIPE starts with reduced survey images and ends by providing shear measurements along with all of the information required for calibration. It includes a novel point spread function (PSF) modelling technique and various PSF validation tools. The code has been designed to facilitate the inclusion of new or improved processing steps to adapt to advances made in the coming years.

This paper is structured as follows. In the following section we describe the design of the software along with efforts we have made to ensure reproducibility and to encourage external contributions. In Sect. 3, we describe the various processing elements that comprise the current version of the pipeline. In Sect. 4, we summarise how the pipeline has been implemented thus far and what we have in mind with regards to future applications. Section 5 mentions our plans for future releases, and we end with some concluding remarks.

2. Software design

2.1. Modular architecture

A key feature dictating the design of SHAPEPIPE was flexibility. Weak gravitational lensing is a rapidly evolving field, with new analysis techniques and tools emerging regularly. A hard-coded weak-lensing pipeline would therefore have extremely limited applications and would become quickly outdated. It was therefore essential for us that a new pipeline be able to adapt in a way that was both maintainable and allowed the easy inclusion of new tools.

The SHAPEPIPE package comprises two principle sub-packages: `pipeline` and `modules`. The `pipeline` sub-package

manages job handling, logging, configuration, argument parsing, dependency handling, and the reading and writing of Flexible Image Transport System (FITS) files. This can effectively be seen as the core of SHAPEPIPE. The modules sub-package consists of a series of what we refer to as ‘pipeline modules’¹, which handle all of the data processing and analysis steps that make up a weak-lensing pipeline (see Sect. 3).

Pipeline modules are simply defined by writing a ‘module runner’, which is a standardised SHAPEPIPE script that defines what a given module should do. This could be as simple as calling an executable binary with some options or writing a full implementation in Python. In either case, once a module runner is in place, SHAPEPIPE can call this module in conjunction with any other. Inputs can be automatically passed or taken from modules, making it possible to chain a series of operations with relative ease.

This architecture will allow us to quickly add new modules to SHAPEPIPE as new tools come onto the market without disrupting the existing pipeline structure. It also means we can more easily rearrange or adapt the existing tools to new datasets.

SHAPEPIPE pipelines are launched via a command line execution script. Significant effort has been put into making this script user-friendly and flexible and providing useful and verbose logging information in order to track the success or failure of any given process.

2.2. Parallel computation

SHAPEPIPE was developed with the assumption that most weak-lensing survey data can be divided into tiles or patches, each of which can be processed independently. Therefore, SHAPEPIPE distributes jobs in an embarrassingly parallel way using JOBLIB (Joblib Development Team 2020) for multi-core machines, and MPI FOR PYTHON (Dalcin et al. 2005, 2008, 2011) on large computer clusters. This maximises the amount of data that can be processed simultaneously with the number of computer cores available to the user without having to parallelise each of the processing steps independently. Modules can also be configured to run in serial mode if the associated task requires access to all of the data, for example, for recombining tiles or patches.

2.3. Reproducibility

The ability to reproduce scientific results is a growing and continuing concern within academia. This is a particularly challenging problem in the context of a weak-lensing pipeline as there are numerous data treatment and processing steps, often involving various third-party tools that evolve over time. We have attempted to mitigate these problems as best as possible.

Firstly, we provide a dedicated SHAPEPIPE development and processing environment. This consists of a CONDA² environment where all Python packages, compilation tools (such as CMAKE), and other third-party software packages (see Appendix A) have fixed versions for a given SHAPEPIPE release. This also extends to all of the modules that make up the weak-lensing pipeline (see Appendix B). This not only ensures consistency across different platforms on which SHAPEPIPE is installed, but also isolates the tools used for processing from those already installed on a given platform (e.g. a different version of Python).

Secondly, we have built in an extensive logging system that tracks the versions of all the packages and modules used for

a given run of SHAPEPIPE, along with all the configuration options used. The logging system catches and reports all errors raised by each module run.

These measures should allow users to reprocess a given dataset in virtually identical conditions. However, we will continue to make reproducibility a priority for future releases of SHAPEPIPE.

2.4. Open-source development

SHAPEPIPE is a fully open-source project. All of the code is publicly available on GitHub³ along with extensive documentation⁴. We provide example configuration files demonstrating how to use the various tools in the context of our application use cases (see Sect. 4), and we additionally provide an extensive step-by-step tutorial for running SHAPEPIPE for processing Ultraviolet Near-Infrared Optical Northern Survey (UNIONS) data (see Sect. 4.1).

The GitHub repository is configured with template issues and pull requests, and includes detailed contribution guidelines. These features are intended to encourage and facilitate contributions from the wider community. Finally, we have also included a comprehensive code of conduct to ensure that contributors can interact in a safe and supportive environment.

3. Weak-lensing pipeline

In this section, we briefly describe the primary pipeline modules that are included in the first public release of SHAPEPIPE in the context of the weak-lensing processing steps they implement. A full list of the modules currently available is provided in Appendix B.

3.1. Masking

The current masking procedure in SHAPEPIPE takes inspiration from the THELI pipeline (Erben et al. 2005; Schirmer 2013) to mask out bright stars, diffraction spikes, Messier objects, and border regions. The mask module uses the CDSCLIENT software to download a catalogue of bright reference stars from version 2.2 of the Guide Star Catalog (GSC; Space Telescope Science Institute & Osservatorio Astronomico di Torino 2001). Alternatively, a star catalogue available on disk (with the same format as the GSC) can also be used, which is particularly relevant for cluster compute nodes that do not allow internet access.

The mask itself is generated using WEIGHTWATCHER (Marmo & Bertin 2008). An example of a mask produced using this module on a patch of UNIONS data is shown in Fig. 1.

3.2. Source detection

The identification of stars and galaxies in the survey images is performed using the community standard SExtractor (Bertin & Arnouts 1996). SHAPEPIPE allows modules such as SExtractor to be run multiple times with different configuration options, meaning that stars and galaxies can be targeted in different ways depending on the need. For example, one run can be made to target a large sample of galaxies for shape measurement, while another run can target a sample of stars for modelling the PSF.

¹ Not to be confused with Python modules.

² <https://docs.conda.io/>

³ <https://github.com/CosmoStat/shapepipe>

⁴ <https://CosmoStat.github.io/shapepipe/>

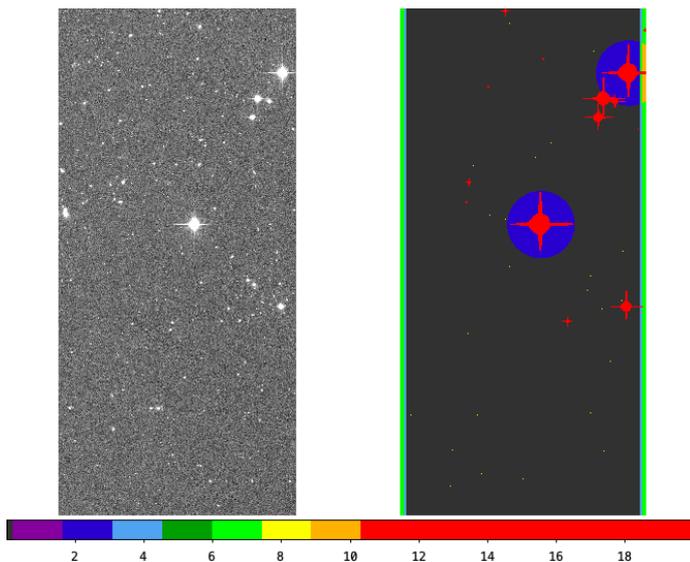


Fig. 1. Example of a mask produced by the ShapePipe mask module on a patch of UNIONS data. The colour bar shows the integer mask pixel values.

3.3. Source selection

The star and galaxy samples used for shape measurement can be refined using several SHAPEPIPE modules. One module implements the spread model method proposed by [Desai et al. \(2012\)](#) and [Mohr et al. \(2012\)](#), which tries to determine if a given object is more point-like or rather more extended in comparison with the PSF. Another module, called `setools`, can apply arbitrary cuts to the objects detected by `SEXTRACTOR`. Figure 2 shows some example plots of a refined star sample obtained after running the `setools` module in SHAPEPIPE.

3.4. PSF model

SHAPEPIPE includes two independent methods for modelling the PSF of the survey. The first method is the community standard PSFEX ([Bertin 2011](#)), which was developed to be used in conjunction with `SEXTRACTOR`. PSFEX produces independent models with polynomial variations of the PSF for each charge-coupled device (CCD) in the field and each exposure. SHAPEPIPE includes an additional module for interpolating the PSFEX PSF model from the positions of the star sample to the positions of the galaxy sample.

The second method is the multi-CCD (MCCD) approach of [Liaudat et al. \(2021\)](#). MCCD builds a model of the PSF for the entire focal plane, capturing both variations that span several CCDs and those that are localised on individual CCDs. As with PSFEX, there is a SHAPEPIPE module for interpolating MCCD PSF models to galaxy positions. SHAPEPIPE allows either or both methods to be run and subsequently used for the galaxy shape measurement.

3.5. PSF validation tools

SHAPEPIPE includes a variety of tools, such as plots of the PSF residuals and ρ -statistics ([Rowe 2010](#); [Jarvis et al. 2016](#)), for validating the quality of the PSF model. An example of how these PSF validation tools have been used for the analysis of UNIONS data can be found in [Guinot et al. \(2022\)](#).

3.6. Shape measurement

Galaxy shape measurement is currently performed using NGMIX ([Sheldon 2015](#)), which has been extensively applied and tested on DES data ([Jarvis et al. 2016](#); [Zuntz et al. 2018](#); [Gatti et al. 2021](#)). At present, NGMIX is implemented in SHAPEPIPE such that galaxy shapes are obtained from a fit of a Gaussian profile from Galsim ([Rowe et al. 2015](#)). The NGMIX module can be run with either of the PSF modelling methods described in Sect. 3.4. Additional NGMIX features and alternative shape measurement techniques are likely to be made available in future releases of SHAPEPIPE.

3.7. Calibration

The primary shear calibration method currently available in SHAPEPIPE is the so-called metacalibration framework ([Huff & Mandelbaum 2017](#); [Sheldon & Huff 2017](#)). Metacalibration is also implemented via the NGMIX module. This module does not directly provide calibrated shear measurements, but does provide all of the calibration images needed to calculate the shear response.

3.8. Miscellaneous

In addition to the essential and generic weak-lensing pipeline processing tools described in the preceding subsections, SHAPEPIPE includes various bookkeeping and file management tools. Some of these tools have been designed specifically for handling UNIONS data, but could be extremely useful for anyone wishing to reprocess these data. For example, SHAPEPIPE includes a module for matching objects detected by the pipeline with external catalogues and the possibility to generate random catalogues.

4. Applications

4.1. UNIONS

The primary focus and initial motivation for the development of SHAPEPIPE has been the analysis of UNIONS data. UNIONS is a collaboration between the Canada-France Imaging Survey (CFIS), Pan-STARRS, and the Wide Imaging with Subaru HSC of the Euclid Sky. One of the main objectives of UNIONS is to provide ground-based multi-band imaging data (u, g, r, i , and z bands collectively; see [Ibata et al. 2017a](#) for details on the u band) of the northern sky for *Euclid*. High quality ground-based photometry is fundamental for accurate photometric redshift estimation and is therefore critical to the success of the *Euclid* mission. In addition to this, there are various ongoing scientific studies with UNIONS data on stars and Milky Way physics ([Ibata et al. 2017b](#); [Fantin et al. 2019, 2021](#); [Thomas et al. 2018, 2019b](#); [Thanjavur et al. 2021](#)), globular clusters ([Thomas et al. 2020](#); [Jensen et al. 2021](#)), dwarf galaxies ([Thomas et al. 2019a](#)), galaxy evolution and mergers ([Bickley et al. 2021](#)), clusters of galaxies ([Roberts et al. 2022](#); [Spitzer 2021](#)), high redshift galaxies and active galactic nuclei ([Ellison et al. 2019](#)), PSF modelling ([Liaudat et al. 2021](#)), and gravitational lensing ([Guinot et al. 2022](#); [Chan et al. 2021](#); [Savary et al. 2021](#)).

[Guinot et al. \(2022\)](#) applied a preliminary version of SHAPEPIPE to 1 700 deg² of CFIS r -band data, obtaining shapes for around 40 million galaxies. The r band of CFIS is particularly well suited to weak-lensing studies given the average seeing of 0.65 and that it is complete down to magnitude 24.5.

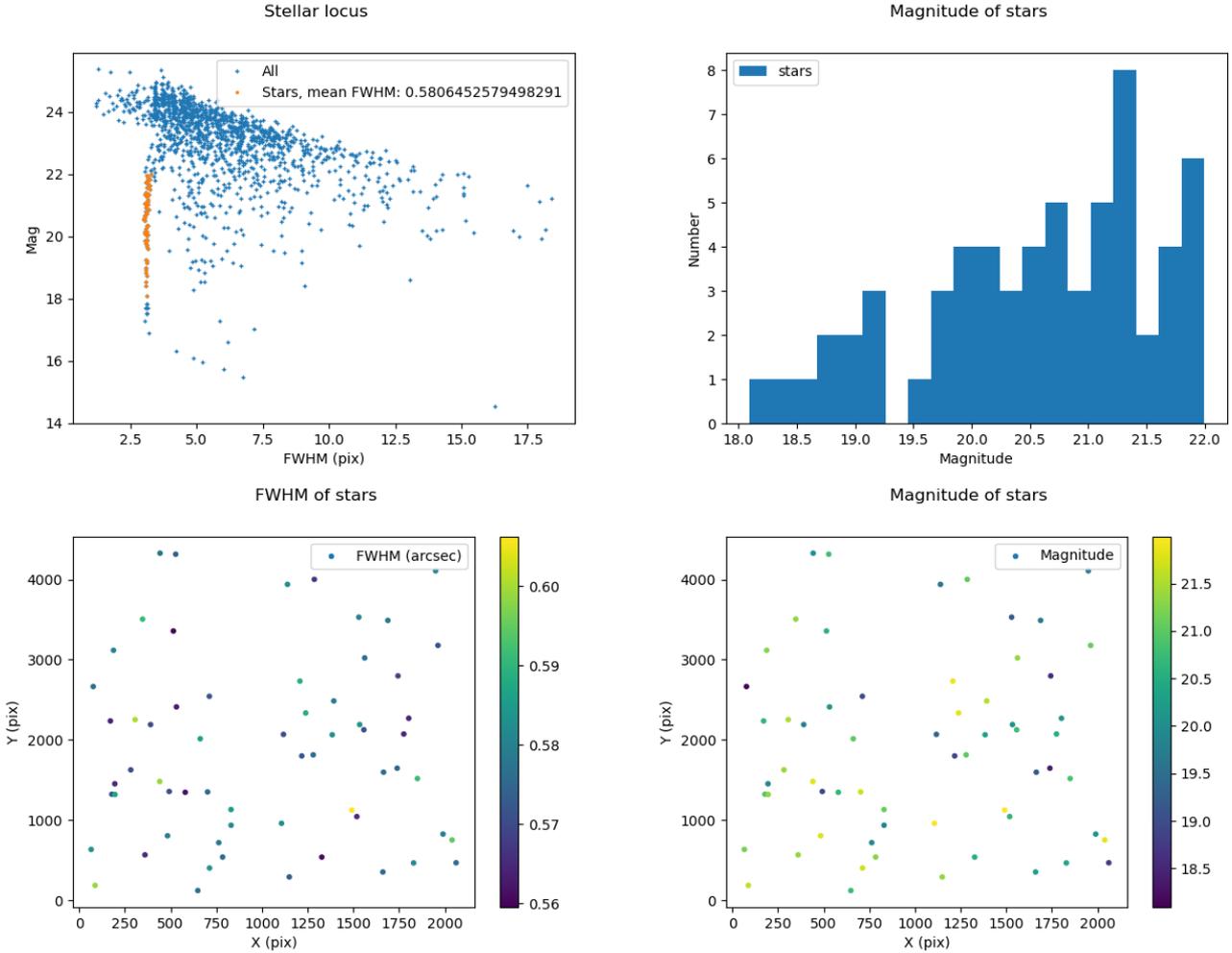


Fig. 2. Example of a refined star selection obtained with the `setools` module on a patch of UNIONS data. All of the plots shown are generated automatically by SHAPEPIPE. *Top left:* Stellar locus (orange points) in the size-magnitude plane. *Top right:* histogram of the magnitude of the objects. *Bottom left:* size of objects in the field. *Bottom right:* magnitude of objects in the field.

Guinot et al. (2022) demonstrate that SHAPEPIPE is capable of processing a large fraction of the sky (on multiple computing platforms) and achieving promising results. Each of the ~ 6800 images of size $0.5 \times 0.5 \text{ deg}^2$ was processed independently. The analysis of an image was performed on a virtual machine (VM) on the Canadian Advanced Network for Astronomical Research⁵. Where possible, individual processing steps were run in parallel on eight cores of the VM. The processing time of one image was around 4 hours. On average, 200 VMs were used in parallel, resulting in a total wall-clock run time of two weeks. The final shape catalogue was around 120 Gb in size. Despite the limitations of working with a single photometric band, Guinot et al. (2022) found: no B-mode signal, a good correlation between E-mode peaks and Planck clusters, and an overall low level of systematic bias.

Work is currently ongoing to process a larger CFIS area ($3\text{--}500 \text{ deg}^2$) with a more recent version of SHAPEPIPE (Kilbinger et al., in prep). This work takes advantage of some of the state-of-the-art features available in SHAPEPIPE, such as full focal plane modelling of the PSF using MCCD. The resulting catalogue will contain around 100 million galaxy shapes and will constitute one of the largest weak-lensing catalogues created to date. In addition to this, several follow-up scientific studies are planned (some are

already underway) using SHAPEPIPE-processed UNIONS data, such as 3×2 point analysis, peak statistics (Ayçoberry et al. 2022), and galaxy cluster lensing.

4.2. Other surveys

No concrete plans have been made at this stage for applying SHAPEPIPE to other datasets. Despite designing the architecture to be as flexible as possible, significant work would still need to be carried out to appropriately tune all of the processing steps to a different survey. Large cosmological surveys such as the DES, *Euclid*, and the Vera C. Rubin Observatory LSST already benefit from dedicated weak-lensing analysis pipelines highly tuned to the specific needs of each survey. However, independent reprocessing of some fraction of these datasets could be valuable for cross-validation and comparison. Additionally, some smaller surveys lack the workforce necessary to build a new weak-lensing pipeline, or indeed to adapt an existing pipeline to their data. Either or both of these use cases might be well suited for future SHAPEPIPE applications.

5. Future releases

We plan to provide regular patch and minor releases of SHAPEPIPE to fix bugs and make general improvements to

⁵ <https://www.canfar.net/>

ensure good maintenance of the code. We additionally plan one major release on a roughly yearly basis in which dependency versions will be updated and new or improved processing steps included. Specific improvements and additions that we currently have in mind include improved masking (new tools and use of *Gaia* data), metadetection, the handling of blended galaxies (see the preliminary work of Farrens et al. 2022), and the full handling of multi-band data.

6. Conclusions

In this paper, we have presented the first public release of the open-source and modular weak-lensing measurement, analysis, and validation pipeline SHAPEPIPE. We have described the objectives and design of the software. In particular, we have emphasised the flexibility of the code architecture, which was designed with the intention of facilitating the inclusion of new processing steps or the improvement of existing ones. This was done with the aim of keeping up with advances made in the field and the possibility of applying SHAPEPIPE to diverse datasets. We have also mentioned steps taken to ensure the reproducibility of results obtained with SHAPEPIPE.

We have provided a brief summary of the tools currently available for masking, source selection, PSF modelling, PSF validation, shape measurement, and calibration. We include links to more detailed documentation for each of these pipeline modules along with examples of how to use them in practice.

Finally, we have mentioned the applications of SHAPEPIPE to real UNIONS data that have been made so far along with plans for future works. We plan to submit follow-up papers in the coming years to inform the community of the advances that have been made in the development of SHAPEPIPE and its applications.

Acknowledgements. This work is based on data obtained as part of the Canada-France Imaging Survey, a CFHT large program of the National Research Council of Canada and the French Centre National de la Recherche Scientifique. Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA Saclay, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers (INSU) of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. This research is based in part on data collected at *Subaru* Telescope, which is operated by the National Astronomical Observatory of Japan. We are honored and grateful for the opportunity of observing the Universe from Maunakea, which has the cultural, historical and natural significance in Hawaii. Pan-STARRS is a project of the Institute for Astronomy of the University of Hawaii, and is supported by the NASA SSO Near Earth Observation Program under grants 80NSSC18K0971, NNX14AM74G, NNX12AR65G, NNX13AQ47G, NNX08AR22G, YORPD20_2-0014 and by the State of Hawaii. This work has made use of the CANDIDE Cluster at the Institut d'Astrophysique de Paris and made possible by grants from the PNCG, CNES, and the DIM-ACAV. This work was supported in part by the Canadian Advanced Network for Astronomical Research (CANFAR) and Compute Canada facilities. We gratefully acknowledge support from the CNRS/IN2P3 Computing Center (Lyon - France) for providing computing and data-processing resources. The authors wish to thank Jérôme Bobin, Jean-Charles Cuillandre, Sébastien Fabbro, Joel Gehin, Stephen Gwyn, François Lanusse, Alan McConnachie, Fred Maurice Ngolé Mboula, Erin Sheldon, Isaac Spitzer and Florent Sureau for various contributions made during the development of this package.

References

Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, **558**, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, **156**, 123

- Ayçoberry, E., Ajani, V., Guinot, A., et al. 2022, ArXiv e-prints, [arXiv:2204.06280]
Bertin, E. 2011, in *Astronomical Society of the Pacific Conference Series*, Astronomical Data Analysis Software and Systems XX, eds. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, 442, 435
Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
Bickley, R. W., Bottrell, C., Hani, M. H., et al. 2021, *MNRAS*, **504**, 372
Bosch, J., Armstrong, R., Bickerton, S., et al. 2018, *PASJ*, **70**, S5
Chan, J. E. A., Lemon, C., Courbin, F., et al. 2021, *A&A*, submitted
Dalcin, L., Paz, R., & Storti, M. 2005, *J. Parallel Distrib. Comput.*, **65**, 1108
Dalcin, L., Paz, R., Storti, M., & D'Elía, J. 2008, *J. Parallel Distrib. Comput.*, **68**, 655
Dalcin, L. D., Paz, R. R., Kler, P. A., & Cosimo, A. 2011, *Adv. Water Resources*, **34**, 1124
Desai, S., Armstrong, R., Mohr, J. J., et al. 2012, *ApJ*, **757**, 83
Ellison, S. L., Viswanathan, A., Patton, D. R., et al. 2019, *MNRAS*, **487**, 2491
Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, *Astron. Nachr.*, **326**, 432
Erben, T., Hildebrandt, H., Miller, L., et al. 2013, *MNRAS*, **433**, 2545
Fantin, N. J., Côté, P., McConnachie, A. W., et al. 2019, *ApJ*, **887**, 148
Fantin, N. J., Côté, P., McConnachie, A. W., et al. 2021, *ApJ*, **913**, 30
Farrens, S., Grigis, A., El Gueddari, L., et al. 2020, *Astron. Comput.*, **32**, 100402
Farrens, S., Lacan, A., Guinot, A., & Vitorelli, A. Z. 2022, *A&A*, **657**, A98
Gatti, M., Sheldon, E., Amon, A., et al. 2021, *MNRAS*, **504**, 4312
Guinot, A., Kilbinger, M., Farrens, S., et al. 2022, ArXiv e-prints, [arXiv:2204.04798]
Harnois-Déraps, J., & van Waerbeke, L. 2015, *MNRAS*, **450**, 2857
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, **585**, 357
Huff, E., & Mandelbaum, R. 2017, ArXiv e-prints, [arXiv:1702.02600]
Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017a, *ApJ*, **848**, 128
Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017b, *ApJ*, **848**, 129
Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
Jarvis, M., Bernstein, G., & Jain, B. 2004, *MNRAS*, **352**, 338
Jarvis, M., Sheldon, E., Zuntz, J., et al. 2016, *MNRAS*, **460**, 2245
Jensen, J., Thomas, G., McConnachie, A. W., et al. 2021, *MNRAS*, **507**, 1923
Joblib Development Team. 2020, Joblib: running Python functions as pipeline jobs <https://joblib.readthedocs.io>
Kiessling, A., Heavens, A. F., Taylor, A. N., & Joachimi, B. 2011, *MNRAS*, **414**, 2235
Kilbinger, M. 2015, *Rep. Progr. Phys.*, **78**, 086901
Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, *A&A*, **625**, A2
Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints, [arXiv:1110.3193]
Liaudat, T., Bonnin, J., Starck, J. L., et al. 2021, *A&A*, **646**, A27
Mandelbaum, R. 2018, *ARA&A*, **56**, 393
Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018, *PASJ*, **70**, S25
Marmo, C., & Bertin, E. 2008, in *Astronomical Society of the Pacific Conference Series*, Astronomical Data Analysis Software and Systems XVII, eds. R. W. Argyle, P. S. Bunclark, & J. R. Lewis, 394, 619
McKinney, W. 2010, in *Proceedings of the 9th Python in Science Conference*, eds. Stéfan van der Walt, & Jarrod Millman, 56
Mohr, J. J., Armstrong, R., Bertin, E., et al. 2012, *Software and Cyber Infrastructure for Astronomy II*
Petri, A. 2016, *Astron. Comput.*, **17**, 73
Roberts, I. D., Parker, L. C., Gwyn, S., et al. 2022, *MNRAS*, **509**, 1342
Rowe, B. 2010, *MNRAS*, **404**, 350
Rowe, B. T. P., Jarvis, M., Mandelbaum, R., et al. 2015, *Astron. Comput.*, **10**, 121
Savary, E., Rojas, K., Maus, M., et al. 2021, *A&A*, submitted
Schirmer, M. 2013, *ApJS*, **209**, 21
Sheldon, E. 2015, NGMIX: Gaussian mixture models for 2D images, *Astrophysics Source Code Library* [record ascl:1508.008]
Sheldon, E. S., & Huff, E. M. 2017, *ApJ*, **841**, 24
Shupe, D. L., Laher, R. R., Storrer-Lombardi, L., et al. 2012, in *SPIE Conf. Ser.*, **8451**, 84511M
Space Telescope Science Institute & Osservatorio Astronomico di Torino. 2001, *VizieR Online Data Catalog*: **I/271**
Spitzer, I., et al. 2021, *MNRAS*, submitted
Thanjavur, K., Ivezić, Ž., Allam, S. S., et al. 2021, *MNRAS*, **505**, 5941
The Pandas Development Team. 2020, <https://doi.org/10.5281/zenodo.6053272>
Thomas, G. F., McConnachie, A. W., Ibata, R. A., et al. 2018, *MNRAS*, **481**, 5223
Thomas, G. F., Annau, N., McConnachie, A., et al. 2019a, *ApJ*, **886**, 10
Thomas, G. F., Laporte, C. F. P., McConnachie, A. W., et al. 2019b, *MNRAS*, **483**, 3119
Thomas, G. F., Jensen, J., McConnachie, A., et al. 2020, *ApJ*, **902**, 89
Zuntz, J., Sheldon, E., Samuroff, S., et al. 2018, *MNRAS*, **481**, 1149

Appendix A: Third-party software

The principal third-party software packages used in SHAPEPIPE are shown in Table A.1. It should be noted that we are currently using a GitHub fork of NGMIX⁶ for technical reasons.

Where possible, we provide references to the papers in which the software is described in detail. More information concerning the CDSCLIENT package is available at the Strasbourg astronomical Data Center website⁷, and information on the SQLITEDICT package is available at the GitHub repository⁸.

Table A.1. List of third-party software packages used in ShapePipe.

Package name	Fixed version	References
ASTROPY	5.0	Astropy Collaboration et al. (2013, 2018)
CDSCLIENT	3.84	
GALSIM	2.2.5	Rowe et al. (2015)
JOBLIB	1.1.0	Joblib Development Team (2020)
MATPLOTLIB	3.5.1	Hunter (2007)
MCCD	1.2.1	Liaudat et al. (2021)
MODOPT	1.6.0	Farrens et al. (2020)
MPI4PY	3.1.3	Dalcin et al. (2005, 2008, 2011)
NGMIX	1.3.6	Sheldon (2015)
NUMPY	≥1.20	Harris et al. (2020)
PANDAS	1.4.1	McKinney (2010); The Pandas Development Team (2020)
PSFEX	3.21.1	Bertin (2011)
SEXTRACTOR	2.25.0	Bertin & Arnouts (1996)
SIP_TPV	1.1	Shupe et al. (2012)
SQLITEDICT	2.0.0	
TREECORR	4.2.6	Jarvis et al. (2004)
WEIGHTWATCHER	1.12	Marmo & Bertin (2008)

Appendix B: ShapePipe modules

All of the modules currently available in SHAPEPIPE that make up the weak-lensing pipeline described in Sect. 3 are shown in Table B.1. In-depth details on each module are available in the SHAPEPIPE documentation⁹.

Table B.1. List of ShapePipe modules.

Module name	Fixed Version	Brief description
find_exposures	1.1	Identify CFIS exposures used to build co-add tiles.
get_images	1.1	Download CFIS images from VOspace.
make_cat	1.1	Prepare the final shear catalogue.
mask	1.0	Mask out bright stars and other artifacts.
match_external	1.1	Match an external catalogue to objects identified with SEXTRACTOR.
mccd_fit	1.1	Run MCCD to fit a PSF model.
mccd_fit_val	1.1	Run and validate the MCCD PSF model.
mccd_interp	1.0	Interpolate the MCCD PSF model to galaxy positions.
mccd_plotsv	1.1	Produce plots for analysing the MCCD PSF model.
mccd_preprocessing	1.1	Pre-process inputs before running MCCD.
mccd_val	1.1	Validate the MCCD PSF model.
merge_headers	1.1	Merge the headers of split CCD files.
merge_sep_cats	1.1	Merge separate catalogues.
merge_starcat	1.1	Merge PSF model star catalogues.
ngmix	0.0.1	Run NGMIX to measure galaxy shapes and perform metacalibration.
pastecat	1.1	Combine multiple SEXTRACTOR output catalogues into a single file.
psfex_interp	1.1	Interpolate the PSFEX PSF model to galaxy positions.
psfex	1.0	Run PSFEX to build a PSF model.
random_cat	1.1	Create a random catalogue.
setools	1.1	Compute summary statistics, produce validation plots, and other useful tools.
sextractor	1.0.1	Run SEXTRACTOR to extract stars and galaxies.
split_exp	1.1	Split single-exposure images into separate files by CCD.
spread_model	1.1	Run the spread model method to refine the galaxy sample.
uncompress_fits	1.1	Uncompress FITS image files and save them as a single-HDU FITS file.
vignet_maker	1.1	Create postage stamps around galaxy positions.

⁶ <https://github.com/aguinot/ngmix>

⁷ <http://cdsarc.u-strasbg.fr/doc/cdsclient.html>

⁸ <https://github.com/RaRe-Technologies/sqlitedict>

⁹ <https://CosmoStat.github.io/shapepipe/>