Shear Mapping in Python (SMPy): Modular, Extensible, and Accessible Dark Matter Mapping

Georgios N. Vassilakis[1,2,3,4¶], **Jacqueline E. McCleary**[1], **Maya Amit**[1], **and Sayan Saha**[1]

1 Department of Physics, Northeastern University, Boston, MA, USA **2** Institute of Astronomy, University of Cambridge, Cambridge, UK **3** Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK **4** Kavli Institute for Cosmology, University of Cambridge, Cambridge, UK ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

Understanding the universe's large-scale distribution of dark matter is a central objective in the era of precision cosmology. A key technique for the study of dark matter is weak gravitational lensing: a phenomenon where light from distant galaxies is sheared as it passes through the gravitational field of a massive object, like a galaxy cluster. This shear, which manifests as a slight (weak) distortion of shapes over thousands of galaxies, allows astrophysicists to infer the distribution of total matter, including both luminous and dark matter.

Obtaining a mass distribution from a catalog of galaxy shears requires an intermediate step. A common tool for this step is the mapping of convergence (κ), which quantifies how much a gravitational lens converges the light from distant galaxies, resulting in a magnification of their shapes. This value is directly proportional to the projected mass density, enabling easy visualization of the overall mass distribution. For a comprehensive review of weak gravitational lensing, refer to (Umetsu, 2020).

The **Shear Mapping in Python (SMPy)** package provides a standardized, well-documented, and open-source solution for creating convergence maps from weak lensing galaxy shear measurements. `SMPy` was initially developed to support the Superpressure Balloon-borne Imaging Telescope (SuperBIT), a stratospheric, near-UV to near-IR observing platform which completed its 45-night observing run in spring 2023 with over 30 galaxy cluster observations (Gill et al., 2024; Sirks et al., 2023). `SMPy` has since evolved into a general-purpose tool suitable for analyzing the weak lensing data from cosmological surveys.

Statement of need

While mass maps are a key deliverable of many cosmological analyses (Jeffrey et al., 2021; Madhavacheril et al., 2024; Oguri et al., 2017), scientists are often left to make these maps from scratch. `SMPy` addresses an outstanding need for the lensing community: an accessible, well-documented, and extensible tool to construct publication-quality mass maps from galaxy shear data. Built on standard scientific Python packages, it provides an easy entry point for researchers new to mass mapping, while also being robust for more senior scientific use. There are currently three separate mapping methods implemented into `SMPy`: the classic Kaiser-Squires inversion (Kaiser & Squires, 1993), aperture mass mapping (Leonard et al., 2012; McCleary et al., 2020), and notably, to our knowledge, the first publicly available implementation of the `KS+` algorithm (Pires, 2020). `KS+` improves reconstruction quality by correcting for systematic effects including missing data, field borders, and reduced shear. `SMPy` also offers specialized and unique features valuable for mass mapping, such as flexible coordinate system support (both celestial and pixel space) and comprehensive signal-to-noise analysis with multiple noise randomization techniques. An example convergence map, created from simulated SuperBIT galaxy cluster observations (McCleary et al., 2023), is shown in Figure 1. SMPy is, to our knowledge, the first convergence mapping software to prioritize both accessibility and advanced features.

State of the field

The weak lensing community is served by publicly available mapping tools like lenspack and jax-lensing (Remy et al., 2022), each with their own strengths. jax-lensing excels at neural network-based approaches and deep learning methods, but applying it to survey-specific data products requires substantial additional development by the end user. lenspack provides well-documented, stand-alone mass-mapping functions, but does not provide an end-to-end convergence-mapping workflow (e.g., configuration management, coordinate handling, noise/SNR analysis, and plotting). In practice, many mass-mapping analyses still rely on bespoke codes built for a specific survey or science case. These gaps, combined with the need for a flexible, configuration-driven, and modular mapping framework, motivated the development of SMPy as an accessible and extensible alternative.

Software design

SMPy was built with the following design principles in mind:

1. **Accessibility:** SMPy is written entirely in Python and deliberately relies only on widely-used scientific Python packages (NumPy, SciPy, Pandas, Astropy, Matplotlib, and PyYAML). This choice of standard dependencies ensures that users can easily install the packages without complex dependency chains, and that the codebase is maintainable and familiar to the scientific Python community.
2. **Extensibility:** SMPy's modular architecture enables seamless addition of new mass mapping techniques, encouraging open-source contribution. The framework's standardized abstract base class architecture allows for the integration of different mapping methods with minimal effort.
3. **Usability:** Creating convergence maps with SMPy requires minimal input—users need to only provide a catalog of galaxies with their associated shear components and coordinates. This straightforward input requirement makes the tool accessible to researchers at all levels. A flexible configuration system is integrated via a single YAML file that defines file paths, convergence map algorithm settings, plotting parameters, and more. With this configuration file, the user can create convergence and SNR maps with one line, either via terminal or within code.
4. **Robustness:** SMPy is designed to be mathematically and algorithmically accurate, allowing the user to create convergence maps with any galaxy shear data. The coordinate system abstraction handles both celestial coordinates (with proper spherical geometry approximations) or pixel-based coordinates through a unified interface. To quantify the significance of the weak lensing detection, multiple noise realizations can be generated using either spatial shuffling (randomizing galaxy positions while preserving shear values) or orientation shuffling (randomizing shear orientations while preserving positions). These noise realizations are used to create a signal-to-noise map with the observed convergence.

These design decisions are reflected in SMPy's overall workflow: users provide a shear catalog and a single YAML configuration file, and SMPy produces convergence maps, signal-to-noise maps, and publication-quality plots with minimal boilerplate. The mapping framework is implemented through standardized abstract base classes, allowing different reconstruction methods to be interchanged while preserving a consistent interface and coordinate-system handling.

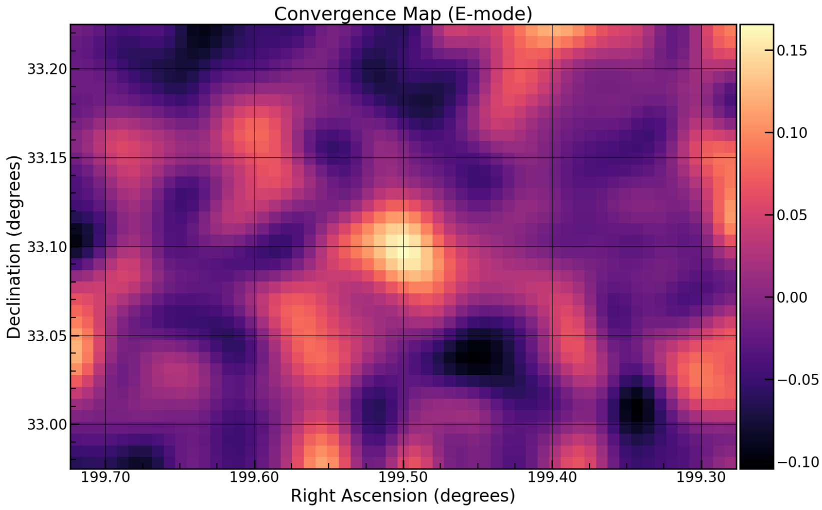

Figure 1: Example convergence map created with SMPy showing the mass distribution of a simulated galaxy cluster. The map was generated using the Kaiser-Squires inversion method on simulated weak lensing data from SuperBIT. The color scale represents the dimensionless surface mass density (convergence), with brighter regions indicating higher mass concentrations.

Research impact statement

`SMPy` was initially developed as part of the weak lensing analysis pipeline for SuperBIT, and it is currently being used to create the first convergence maps measured from the statosphere, derived from SuperBIT's 2023 science flight imaging. These convergence maps will be presented in a forthcoming SuperBIT weak-lensing analysis (Saha et al., in preparation, 2026; "Lensing in the Blue IV"). In addition, `SMPy` has been presented at the Jet Propulsion Laboratory, and we have received interest in exploring integration into Euclid weak-lensing analysis workflows.

AI usage disclosure

Generative AI was used in three ways in this work. First, Claude Sonnet 4 (with Extended Thinking enabled via the Claude.ai web interface) was used to assist in implementing the `KS+` algorithm in `SMPy`. The author reviewed all AI-assisted code line-by-line against the Algorithm 1 description in Appendix A of (Pires, 2020) and validated the implementation by injecting fiducial star masks into data and confirming that the reconstruction behaved as expected. Second, OpenAI's ChatGPT (5.0-Thinking and 5.2-Thinking) models have been used for generating unit tests for the codebase via the Codex CLI. Lastly, both OpenAI's ChatGPT models and Anthropic's Claude models have been used in the GitHub App harness to review pull requests and catch bugs. An example use case of this would include making a comment on an open pull request along the lines of '@codex/@claude review this PR'.

Software references

`SMPy` is written in Python 3.8+ and uses the following packages:

- `NumPy` (Harris et al., 2020)
- `SciPy` (Virtanen et al., 2020)
- `Pandas` (team, 2024)

- `Astropy` (Astropy Collaboration et al., 2013, 2018, 2022)
- `Matplotlib` (Hunter, 2007)
- `PyYAML` (Simonov, 2024)

Acknowledgements

This material is based upon work supported by a Northeastern University Undergraduate Research and Fellowships PEAK Summit Award.

References

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., Earl, N., Starkman, N., Bradley, L., Shupe, D. L., Patil, A. A., Corrales, L., Brasseur, C. E., Nöthe, M., Donath, A., Tollerud, E., Morris, B. M., Ginsburg, A., Vaher, E., Weaver, B. A., Tocknell, J., Jamieson, W., … Astropy Project Contributors. (2022). The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. *The Astrophysical Journal*, *935*(2), 167. https://doi.org/10.3847/1538-4357/ac7c74

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., Ginsburg, A., VanderPlas, J. T., Bradley, L. D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T. L., Cruz, K. L., Robitaille, T. P., Tollerud, E. J., … Astropy Contributors. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *The Astronomical Journal*, *156*(3), 123. https://doi.org/10.3847/1538-3881/aabc4f

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., … Streicher, O. (2013). Astropy: A community Python package for astronomy. *Astronomy & Astrophysics*, *558*.

Gill, A. S., Benton, S. J., Damaren, C. J., Everett, S. W., Fraisse, A. A., Hartley, J. W., Harvey, D., Holder, B., Huff, E. M., Jauzac, M., Jones, W. C., Lagattuta, D., Leung, J. S.-Y., Li, L., Luu, T. V. T., Massey, R., McCleary, J. E., Nagy, J. M., Netterfield, C. B., … Vitorelli, A. Z. (2024). SuperBIT superpressure flight instrument overview and performance: Near-diffraction-limited astronomical imaging from the stratosphere. *The Astronomical Journal*, *168*(2), 85. https://doi.org/10.3847/1538-3881/ad5840

Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jeffrey, N., Gatti, M., Chang, C., Whiteway, L., Demirbozan, U., Kovacs, A., Pollina, G., Bacon, D., Hamaus, N., Kacprzak, T., Lahav, O., Lanusse, F., Mawdsley, B., Nadathur, S., Starck, J. L., Vielzeuf, P., Zeurcher, D., Alarcon, A., Amon, A., … Collaboration, D. (2021). Dark energy survey year 3 results: Curved-sky weak lensing mass map reconstruction. *Monthly Notices of the Royal Astronomical Society*, *505*(3), 4626–4645. https://doi.org/10.1093/mnras/stab1495

Kaiser, N., & Squires, G. (1993). Mapping the dark matter with weak gravitational lensing. *The Astrophysical Journal*, *404*, 441–450. https://doi.org/10.1086/172297

Leonard, A., Pires, S., & Starck, J.-L. (2012). Fast calculation of the weak lensing aperture mass statistic. *Monthly Notices of the Royal Astronomical Society*, *423*(4), 3405–3412. https://doi.org/10.1111/j.1365-2966.2012.21133.x

Madhavacheril, M. S., Qu, F. J., Sherwin, B. D., MacCrann, N., Li, Y., Abril-Cabezas, I., Ade, P. A. R., Aiola, S., Alford, T., Amiri, M., Amodeo, S., An, R., Atkins, Z., Austermann, J. E., Battaglia, N., Battistelli, E. S., Beall, J. A., Bean, R., Beringue, B., … Zheng, K. (2024). The atacama cosmology telescope: DR6 gravitational lensing map and cosmological parameters. *The Astrophysical Journal*, *962*(2), 113. https://doi.org/10.3847/1538-4357/acff5f

McCleary, J. E., dell'Antonio, I., & Linden, A. von der. (2020). Dark matter distribution of four low-z clusters of galaxies. *The Astrophysical Journal*, *893*(1), 8. https://doi.org/10.3847/1538-4357/ab7c58

McCleary, J. E., Everett, S. W., Shaaban, M. M., Gill, A. S., Vassilakis, G. N., Huff, E. M., Massey, R. J., Benton, S. J., Brown, A. M., Clark, P., Holder, B., Fraisse, A. A., Jauzac, M., Jones, W. C., Lagattuta, D., Leung, J. S.-Y., Li, L., Luu, T. V. T., Nagy, J. M., … Tam, S. I. (2023). Lensing in the blue. II. Estimating the sensitivity of stratospheric balloons to weak gravitational lensing. *The Astronomical Journal*, *166*(3), 134. https://doi.org/10.3847/1538-3881/ace7ca

Oguri, M., Miyazaki, S., Hikage, C., Mandelbaum, R., Utsumi, Y., Miyatake, H., Takada, M., Armstrong, R., Bosch, J., Komiyama, Y., Leauthaud, A., More, S., Nishizawa, A. J., Okabe, N., & Tanaka, M. (2017). Two- and three-dimensional wide-field weak lensing mass maps from the hyper suprime-cam subaru strategic program S16A data. *Publications of the Astronomical Society of Japan*, *70*(SP1), S26. https://doi.org/10.1093/pasj/psx070

Pires, S. (2020). Euclid: Reconstruction of weak-lensing mass maps for non-gaussianity studies. *Astronomy & Astrophysics*, *638*, A141. https://doi.org/10.1051/0004-6361/201936865

Remy, B., Lanusse, F., Jeffrey, N., Liu, J., Starck, J.-L., Osato, K., & Schrabback, T. (2022). Probabilistic mass-mapping with neural score estimation. *Astronomy & Astrophysics*, *672*. https://doi.org/10.1051/0004-6361/202243054

Simonov, K. (2024). *PyYAML*. https://pyyaml.org/

Sirks, E. L., Massey, R., Gill, A. S., Anderson, J., Benton, S. J., Brown, A. M., Clark, P., English, J., Everett, S. W., Fraisse, A. A., Franco, H., Hartley, J. W., Harvey, D., Holder, B., Hunter, A., Huff, E. M., Hynous, A., Jauzac, M., Jones, W. C., … Vassilakis, G. N. (2023). Data downloaded via parachute from a NASA super-pressure balloon. *Aerospace*, *10*(11). https://doi.org/10.3390/aerospace10110960

team, T. pandas development. (2024). *Pandas-dev/pandas: pandas* (Version v2.2.2). Zenodo. https://doi.org/10.5281/zenodo.10957263

Umetsu, K. (2020). Cluster–galaxy weak lensing. *The Astronomy and Astrophysics Review*, *28*(1), 106. https://doi.org/10.1007/s00159-020-00129-w

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2