The Journal of Open Source Software

Alpaga: A Python package for automated analysis of Second Harmonic Generation polarization experiments

Fabien Rondepierre [1*], **Maxime Fery**[1], **Oriane Bonhomme** [1], **and Guillaume Le Breton** [1*¶]

1 Institut Lumière Matière, UMR5306, Villeurbanne, France ¶ Corresponding author * These authors contributed equally.

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

Alpaga (AnaLyse en PolArisation de la Génération de second hArmonique) is a Python package designed for the automated analysis of Second Harmonic Generation (SHG) experimental acquisitions. The software provides a comprehensive workflow for processing spectroscopic measurements from Surface Second Harmonic Generation (SSHG) (Shen, 1989; Tran et al., 2017) and Second Harmonic Scattering (SHS) experiments, which are crucial techniques in surface science and nonlinear optics research.

The package implements a robust automated procedure that extracts Gaussian peak intensities from spectral measurements through three main steps: automatic file detection and organization, spectral cleaning and averaging with removing non-physical artifacts, and Gaussian fitting for intensity extraction. This automated approach significantly reduces the time and potential human error associated with manual data processing while providing consistent and reproducible analysis results.

Statement of need

SHG experiments involve long acquisition times, which generate large volumes of spectroscopic data often affected by significant noise, requiring careful processing to extract meaningful physical parameters. Researchers typically face several challenges when analyzing SHG data: (1) handling numerous acquisition files with varying experimental parameters, (2) removing non-physical artifacts, (3) averaging multiple acquisitions to improve signal-to-noise ratios, and (4) consistently fitting Gaussian profiles to extract peak intensities. These tasks are often performed manually or with custom scripts, leading to inconsistencies between research groups and potential analysis errors.

Alpaga addresses these challenges by providing a standardized, automated workflow specifically designed for SHG polarization analysis. The Figure 1 presents this robust data processing: going from experimental raw results to a research-oriented analysis, thanks to noise cleaning and peak intensity extraction. The software is particularly valuable for research groups working with SSHG and SHS experiments, where systematic analysis of polarization-dependent measurements is essential for understanding surface properties and molecular orientation. Python enables Alpaga to provide a user-friendly interface while leveraging efficient numerical libraries for computationally intensive operations.

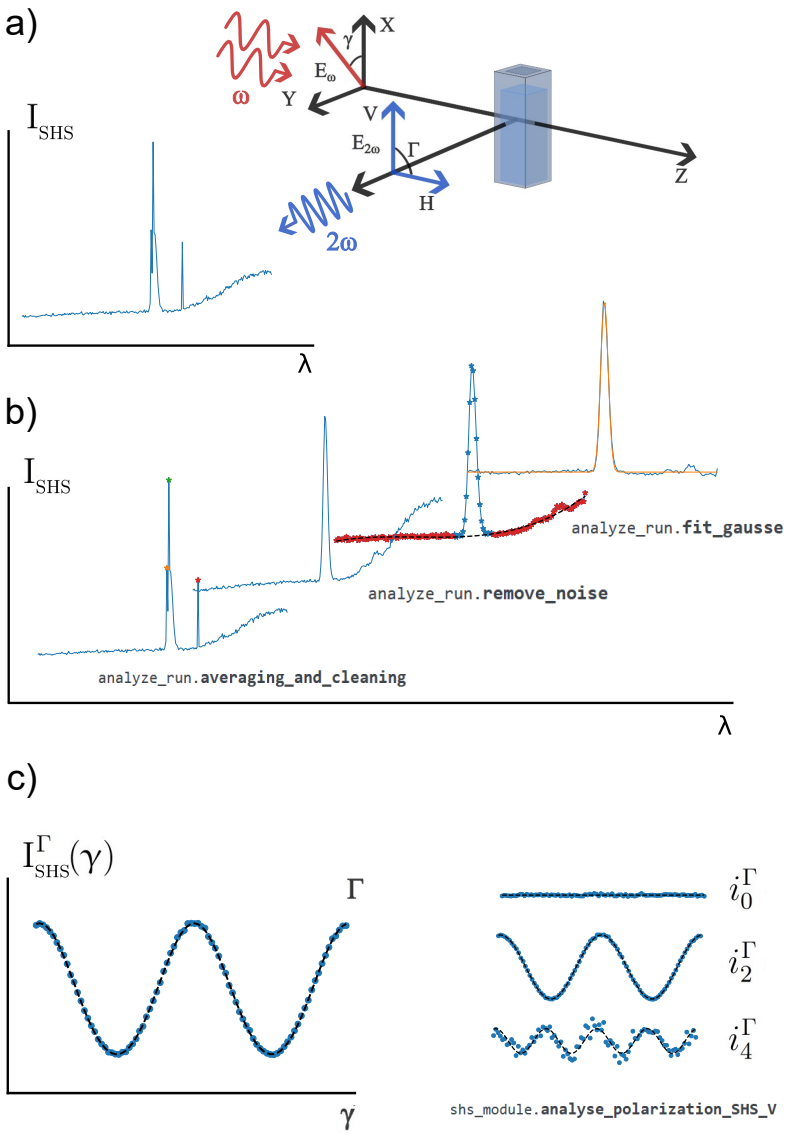

Figure 1: `Alpaga` goal: robustly extract from raw spectra meaningful information from **a)** Second Harmonic Generation experimental setups (here Second Harmonic Scattering). **b)** Sequence of second-harmonic signal processing steps. First, suppression of spurious peaks and averaging of multiple acquisitions performed with the `averaging_and_cleaning` function. Then, noise estimation and correction across distinct processing regions using the `remove_noise` function. Eventually, extraction of the second-harmonic component followed by Gaussian fitting with the `fit_gauss` function. **c)** Fourier analysis of polarization-resolved SHS measurements using the `analysis_polarisation_SHS_V` function.

`Alpaga` was designed to be used by both experienced researchers in nonlinear optics and students learning SHG analysis techniques. The automated nature, supported by comprehensive documentation (wiki) and tutorials, makes the workflow accessible to newcomers while providing the reliability and consistency required for research applications.

Key Features and Implementation

Alpaga is built on established Python scientific libraries, including NumPy, SciPy, and Matplotlib, providing reliable numerical operations and visualization capabilities. The software architecture follows a modular design with several key components:

Automated File Management: The software automatically identifies and organizes spectroscopic data files based on experimental parameters, streamlining the analysis workflow for large datasets with consistent naming conventions.

Spectral Cleaning and Averaging: Detection and removal of electronic noise and other non-physical spikes from spectra. The cleaning procedure includes configurable parameters for spike detection sensitivity and handles the averaging of multiple acquisitions with identical experimental parameters to improve signal-to-noise ratios.

Gaussian Fitting with Multiple Options: The package offers various fitting approaches for extracting peak intensities from Gaussian profiles. Users can select from different fitting algorithms and configure parameters such as baseline handling (fluorescence background) to optimize results based on their specific experimental conditions.

Domain-Specific Analysis Tools: Dedicated modules for SSHG and SHS analysis provide specialized functionality for extracting physical parameters relevant to surface science applications, including polarization-dependent analysis and orientation parameter extraction.

Related Work

The automated workflow of Alpaga sets it apart from the manual analysis methods often used in the field, offering greater consistency and efficiency for SHG research groups. In contrast to general-purpose spectroscopic packages such as rampy (Losq, 2025), Alpaga addresses the specific needs of the SHG community by providing dedicated tools for polarization-dependent measurements. It embeds domain knowledge of SHG experiments, including typical artifact patterns, polarization conventions, and the mathematical frameworks required for SSHG and SHS analysis. In addition, Alpaga features a file management system designed for handling multiple spectra per acquisition, a capability particularly relevant for polarization-based studies.

Usage and Impact

The Alpaga project started in 2022 at the Institut Lumière Matière laboratory (ILM), France, to merge all the different analysis tools established in our experimental group. The four authors of this publication participated in the code development and usage. To date, about a dozen scientists have used Alpaga to treat experimental data. The software has already enabled more efficient and consistent data analysis workflows for research groups working with SSHG and SHS experiments (at ILM or other labs, thanks to the automated file management procedure), and contributes to improved reproducibility in SHG research. This procedure has been used in multiple scientific communications and publications (Fery, 2025; Le Breton, 2022; Le Breton et al., 2023, 2024; Rondepierre, 2025; Rondepierre, Brevet, et al., 2025; Rondepierre, Salmon, et al., 2025), while not always being directly mentioned.

The package includes comprehensive documentation with detailed examples and parameter explanations, making it accessible to both experienced researchers and newcomers to SHG analysis. Installation is straightforward through standard Python package management tools, and the software is distributed under the LGPL v2.1 license to ensure broad accessibility.

The Journal of Open Source Software

Acknowledgements

We are grateful to Estelle Salmon, Emmanuel Benichou, and Pierre-François Brevet for their valuable help, insightful guidance, and for generously sharing their previous implementation, which has significantly shaped `Alpaga` current workflow. We also warmly acknowledge the former members of the ONLI group, Aurélie Bruyère, Lucile Sanchez, and Antonin Pardon ADD MORE?, whose pioneering contributions laid the foundation for the present development of the experimental workflow. This work was conducted at the Institut Lumière Matière (ILM).

References

Fery, M. (2025). *Sonder la structure de liquides par des méthodes optiques non linéaires: Des électrolytes concentrés en volume aux liquides confinés* [PhD thesis]. Université Claude Bernard-Lyon I.

Le Breton, G. (2022). *Second harmonic generation of non-resonant liquid: From the molecular response to the experimental measurement* [PhD thesis]. Université Claude Bernard-Lyon I.

Le Breton, G., Bonhomme, O., Benichou, E., & Loison, C. (2023). Liquid water: When hyperpolarizability fluctuations boost and reshape the second harmonic scattering intensities. *The Journal of Physical Chemistry Letters*, *14*(18), 4158–4163.

Le Breton, G., Loison, C., Vynck, K., Benichou, E., & Bonhomme, O. (2024). Microscopic view on the polarization-resolved s-SHG intensity of the vapor/liquid interface of pure water. *The Journal of Chemical Physics*, *161*(15).

Losq, C. L. (2025). *Rampy: Python software for spectral data processing (IR, raman, XAS…)*. https://github.com/charlesll/rampy.

Rondepierre, F. (2025). *Corrélations orientationnelles des liquides explorées par diffusion de second harmonique* [PhD thesis]. Université Claude Bernard Lyon I.

Rondepierre, F., Brevet, P.-F., & Duboisset, J. (2025). Asymptotic underscreening in concentrated electrolytes measured by optical second harmonic scattering of water. *The Journal of Physical Chemistry Letters*, *16*(11), 2690–2694.

Rondepierre, F., Salmon, E., Jonin, C., Duboisset, J., & Brevet, P.-F. (2025). Polarization resolved second harmonic scattering of neat water in the right angle and forward scattering geometries. *The Journal of Chemical Physics*, *162*(3), 034201.

Shen, Y. (1989). Surface properties probed by second-harmonic and sum-frequency generation. *Nature*, *337*(6207), 519–525.

Tran, R. J., Sly, K. L., & Conboy, J. C. (2017). Applications of surface second harmonic generation in biological sensing. *Annual Review of Analytical Chemistry*, *10*(1), 387–414.