The Journal of Open Source Software

merrypopins: A Python package for nanoindentation data science

Cahit Acar[1], **Anna Mercelissen**[1], **Hugo W. van Schrojenstein**[1], and **John M. Aiken**[1,2,3]¶

1 Utrecht University, The Netherlands **2** Expert Analytics, Norway **3** University of Oslo, Njord Centre, Norway ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Fei Tao

Reviewers:

- @MoonshinetheP
- @KennethKwabenaKenney

Submitted: 22 July 2025
Published: unpublished

Summary

`merrypopins` is a Python library to streamline the workflow of nanoindentation experiment data processing, automated pop-in detection, and statistical analysis of collections of pop-in events. Nanoindentation is a technique for experimental deformation of materials with the aim of characterizing material behaviour and quantifying mechanical properties from load-displacement data Oliver & Pharr (2004). Experiments performed with a spherical tip can also be used to construct stress-strain curves and by extension the determination of the yield point defining the transition from elastic to plastic deformation Pathak & Kalidindi (2015). Understanding the start of plasticity in materials at the microscale is crucial for various applications, including engineered materials and earthquake mechanics. A common feature during the loading part of nanoindentation experiments is the sudden increase of indentation depth at constant force, called "pop-in" events. Manually recognizing these characteristics is labor-intensive and subjective, emphasizing the importance of automated, reproducible detection approaches.

Statement of need

Detecting pop-ins is difficult because they appear in subtle, intermittent, and different ways within indentation curves. Historically, professional analysts have recognized pop-in occurrences manually. The researcher simply looks at either depth vs. time or stress-strain curves looking for sharp, localized changes. This approach suffers from subjectivity, labor intensity, and potential inconsistencies among multiple observers and big datasets. Modern nano-indentation machines can perform up to 12 indentations per second (Bruker, n.d.) and thus new tools to automate pop-in detection are necessary to be built in order to keep up with the increase in data. `merrypopins` marks the first attempt to automate pop-in detection.

Pop-ins are linked to dislocation in crystalline materials and are considered small-scale analogues of earthquakes Sato et al. (2020). Like real earthquakes, they follow statistical patterns, such as power-law distributions in size and time between events. Generally, the first pop-in during an indentation experiment coincides with the start of plasticity. The size of the indenter tip and the degree of pre-existing plastic deformation have a significant impact on the stress at which the first pop-in occurs Morris et al. (2011) and thus the yield hardness of the material. This effect is the result of a delayed plastic yielding when the volume stressed by the indenter tip does not contain any pre-existing dislocations for the initiation of plasticity. A smaller volume is more likely to be free of dislocations, especially when the material has a lower dislocation density, so the material will behave elastically up to higher load and stress. In contrast, larger tips sample a bigger volume, increasing the chance of hitting existing dislocations and causing the first pop-in at lower stresses. This size effect must be overcome to obtain yield hardness values applicable across scales or to other systems.

Primary users of `merrypopins` are students, researchers, and academics in the fields of material science, geology, nano-mechanics, and earthquake science. High-resolution indentation experiments are increasingly used to investigate plastic and fracture processes at the microscale. Despite the growing number of studies targeting pop-in occurrences in load-depth curves, almost all previous research relies on manual inspection or private scripts with undisclosed methods for the detection and quantification of pop-ins, creating a lack of easily accessible, reproducible event detection software. There is an urgent need for adaptable, open-source solutions that can be used "out of the box" by non-programmers and provide extensibility for power users as nanoindentation tools grow, spanning both traditional materials laboratories and emerging geophysical applications. To advance the next generation of automated pop-in analysis, researchers can submit new detection techniques, parameter settings, or visualization modules through our public `merrypopins` GitHub repository. We, therefore, welcome feature requests, bug reports, and community-contributed enhancements.

Using a variety of detection techniques ensures that `merrypopins` can detect pop-in events in many material systems and experimetal circumstances. `merrypopins` primary uses the Savitzky-Golay filter and Fourier-domain differentiation methods for pop-in detection. Savitzky-Golay's local polynomial smoothing maintains prominent curve characteristics while reducing high-frequency noise (Savitzky & Golay, 1964). Fourier spectral methods identify abrupt discontinuities with minimal parameterization (Cooley et al., 1969). Both strategies are computationally efficient, highly interpretable, and require only a few user-tunable parameters (window length, polynomial order, or frequency threshold), making them excellent for quick initial screening.

`merrypopins` also includes two other machine learning methods. Isolation Forest and convolutional autoencoders enable data-driven adaptation. Isolation Forest, an unsupervised ensemble-based statistical framework, can detect anomalies in multidimensional feature spaces without labeled instances (Liu et al. (2008)). This is especially useful when the pop-in magnitudes or frequencies are unknown beforehand. Convolutional autoencoders learn hierarchical feature representations directly from data, capturing subtle nonlinear patterns that classical approaches may overlook (Malhotra et al., 2016). However, they require more resources. These four techniques balance sensitivity, interpretability, and processing cost, allowing researchers to select and combine algorithms based on dataset size, noise characteristics, and analytic goals.

In addition to pop-in detection, `merrypopins` also includes a statistical analysis suite. This suite provides functions to automatically calculate stress-strain curves, calculate precursor statistics (i.e., are there events occuring prior to a pop-in such as yielding), and temporal statistics across pop-in events. These are both accessible in the library and in a no-code streamlit app.

The `merrypopins` library was developed using a tutorial-driven software development framework Aiken (2020). Instead of starting with predetermined architectural specs, this approach converts the scientist's process into a live, executable lesson (often a Jupyter notebook). Developers and researchers worked iteratively, with academics creating function stubs in a scientific narrative framework and developers implementing these functions based on real-world usage cases. This strategy ensures that scientific usability drives software design.

Code Availability

The `merrypopins` package can be installed via:

```
pip install `merrypopins`
```

Alternatively, the package can be found on github (https://github.com/SerpRateAI/merrypopins).

Contributions can be made by forking the repository and making a pull request.

The streamlit app is accessible via the streamlit website (https://merrypopins.streamlit.app/).

Acknowledgements

This project has received funding from the Norwegian Research Council (SerpRateAI, grant no. 334395) and is supported by EPOS-eNLarge funded by the Dutch Research Council (NWO) Roadmap for large-scale research infrastructure. We would like to thank Alissa Kotowski for fruitful conversations.

References

Aiken, J. M. (2020). Science coding vs software development. In *substack*. substack. https://mnky9800n.substack.com/p/science-coding-vs-software-development

Aiken, J. M., Jones, G., Yin, X., Abele, A. K., Woods, C., Westaway, R. M., & Bamber, J. L. (2025). 4DModeller: A spatio-temporal modelling package. *Journal of Open Source Software*, *10*(106), 7047. https://doi.org/10.21105/joss.07047

Bruker. (n.d.). *Hysitron TI 990 TriboIndenter brochure*. Bruker. https://www.bruker.com/en/meta/forms/bns-form-pages/brochures/ni/hysitron-ti-990.html

Cooley, J. W., Lewis, P. A. W., & Welch, P. D. (1969). The fast fourier transform and its applications. *IEEE Transactions on Education*, *12*(1), 27–34. https://doi.org/10.1109/TE.1969.4320436

Ispánovity, P. D., Ugi, D., Péterffy, G., Knapek, M., Kalácska, S., Tüzes, D., Dankházi, Z., Máthis, K., Chmelík, F., & Groma, I. (2022). Dislocation avalanches are like earthquakes on the micron scale. *Nature Communications*, *13*(1), 1975.

Kalidindi, S. R., & Pathak, S. (2008). Determination of the effective zero-point and the extraction of spherical nanoindentation stress-strain curves. *Acta Materialia*, *56*(14), 3523–3532. https://doi.org/10.1016/j.actamat.2008.03.036

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. *2008 Eighth IEEE International Conference on Data Mining*, 413–422. https://doi.org/10.1109/ICDM.2008.17

Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). *LSTM-based encoder-decoder for multi-sensor anomaly detection*. https://doi.org/10.48550/arXiv.1607.00148

Morris, J. R., Bei, H., Pharr, G. M., & George, E. P. (2011). Size effects and stochastic behavior of nanoindentation pop in. *Physical Review Letters*, *106*(1), 165502. https://doi.org/10.1103/PhysRevLett.106.165502

Oliver, W. C., & Pharr, G. M. (1992). An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *Journal of Materials Research*, *7*(6), 1564–1583. https://doi.org/10.1557/JMR.1992.1564

Oliver, W. C., & Pharr, G. M. (2004). Measurement of hardness and elastic modulus by instrumented indentation: Advances in understanding and refinements to methodology. *Journal of Materials Research*, *19*(1), 3–20. https://doi.org/10.1557/jmr.2004.19.1.3

Pathak, S., & Kalidindi, S. R. (2015). Spherical nanoindentation stress-strain curves. *Materials Science and Engineering R*, *91*(1), 1–36. https://doi.org/10.1016/j.mser.2015.02.001

Sato, Y., Shinzato, S., Ohmura, T., Hatano, T., & Ogata, S. (2020). Unique universal scaling in nanoindentation pop-ins. *Nature Communications*, *11*(1), 4177.

Savitzky, Abraham., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Shim, S., Bei, H., George, E. P., & Pharr, G. M. (2008). A different type of indentation size effect. *Scripta Materialia*, *59*(1), 1095–1098. https://doi.org/10.1016/j.scriptamat.2008.07.026

Woods, C., Hedges, L., Edsall, C., Brooks-Pollock, E., Parton-Fenton, C., McKinley, T. J., Keeling, M. J., & Danon, L. (2022). MetaWards: A flexible metapopulation framework for modelling disease spread. *Journal of Open Source Software*, *7*(70), 3914. https://doi.org/10.21105/joss.03914