The Journal of Open Source Software

quickBayes: An analytical approach to Bayesian loglikelihoods

Anthony Lim[1]

1 Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Campus, Didcot, Oxfordshire, OX11 0QX

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Øystein Sørensen

Reviewers:

- @JonasMoss
- @mattpitkin
- @prashjet

Submitted: 27 November 2023
Published: unpublished

Summary

It is common in science to have multiple hypotheses that could describe your data. Each one of these hypothesis provides a mathematical description of the data and gives unique physical insight into the data. However, it is not always obvious which hypothesis is the correct one and to demonstrate this figure 1 show Quasi Elastic Neutron Scattering (QENS) data fitted with one and two Lorentzian peaks. The quickBayes package is designed to make it easier for users to determine which hypothesis is correct given their data. At present there are example workflows for:

- Determining the number of Lorentzians in quasielastic neutron scattering data.
- Determining the number of exponential decays in MuSR data.

The quickBayes package comes with an API that can be easily be extended by users to calculate the most likely hypothesis given their data.

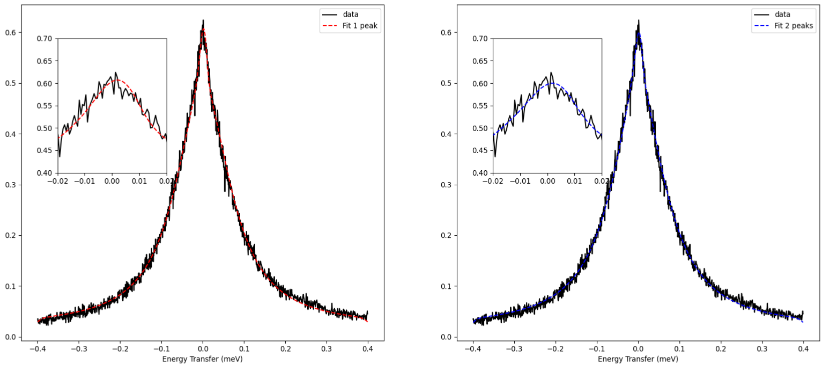

Figure 1: Two plots of the same raw data from a Quasi Elastic Neutron Scattering (QENS) experiment, showing the fits for one and two peaks (it also includes a linear background and an elastic peak). The inserts show zoomed in images of the peak centre. From the loglikelihood calculation the most likely number of peaks is two.

Statement of need

The quickBayes package started as a replacement for the quasielasticbayes package (Sivia & Howells, n.d.), which was used for:

- Determining the most likely number of Lorentzians in QENS data.

- Calculating the loglikelihood of a stretched exponential in QENS data.
- Calculating a contour of the probability of a stretched exponential as a function of beta and sigma (full width half maxima).

The quasielasticbayes package is highly successful within the field of QENS (Sivia et al., 1992). During the development of quickBayes it became clear that the fundamental concept would benefit other areas of research such as MuSR and was written to be easily extendable into new domains. The new code is written in Python with clear sections for fitting functions, fitting engines and workflows. The workflows are designed to use the fitting functions and engines as building blocks, allowing the workflow to focus on the specific steps of the data analysis (e.g. splines, rebinning). The code has been developed to make it easy to read and understand, with comprehensive automated testing of the functionality. These changes make the new code more maintainable and reliable, while ensuring it is easily extendable by the user community.

Traditionally the calculation of the loglikelihood is not analytic and needs to be done with Bayesian inference, using a computationally expensive model such as Markov chain Monte Carlo (von Toussaint, 2011). The calculation of the loglikelihood in both quasielasticbayes and quickBayes uses a few assumptions to simplify the calculation so it can be evaluated analytically (Sivia & Skilling, 2006). These simple assumptions reduce the computational cost significantly allowing users to quickly determine the most likely hypothesis given their data. In the QENS community it is difficult to determine if the data contains one or two Lorentzians, the quasielasticbayes and now quickBayes packages provide a way to quickly calculate which is more likely.

The quasielasticbayes package was written in Fortran and had become almost impossible to maintain due to poor code quality and a lack of unit testing. Furthermore, compiling the package only worked in very specific cases due to floats being frequently converted between real and complex variables when passed between functions. Therefore, a new code was required that captured the key functionality of the quasielasticbayes package and led to the development of quickBayes.

Acknowledgments

I would like to thank Jari Fowkes for his helpful comments and insights. Thank you to the main tester of the package Spencer Howells. I would also like to thank the reviewers for qucikBayes; Martyn Gigg, Silke Schomann, Robert Applin and Jess Farmer.

References

Sivia, D. S., & Howells, W. S. (n.d.). *Quasielasticbayes version 0.2.0*.

Sivia, D. S., J., C. C., Howells, W. S., & König, S. (1992). Bayesian analysis of quasielastic neutron scattering data. *Physica B: Condensed Matter*, *182*(4), 341–348. https://doi.org/10.1016/0921-4526(92)90036-R

Sivia, D. S., & Skilling, J. (2006). *Data analysis a bayesian tutorial* (Second). Oxford University Press. ISBN: 9780198568322

von Toussaint, U. (2011). Bayesian inference in physics. *Rev. Mod. Phys.*, *83*, 943–999. https://doi.org/10.1103/RevModPhys.83.943