Bayesian data analysis of atomic spectra

M. Trassinelli[1] *

[1]Institut des NanoSciences de Paris, CNRS, Sorbonne Université, 75005 Paris, France

Synopsis Bayesian statistical methods are widely used in different physics domains, such as cosmology and particle physics, but rarely in atomic physics. Here, we present some applications that are commonly encountered in the analysis of atomic spectra like the determination of the number of lines present in the spectrum or the most adapted line profile. To solve these issues, we have developed a specific new analysis program named `Nested_fit`, based on Bayesian statistical methods.

In some case of spectra analysis, the standard probability methods reach their limitations for accurately study experimental data sets. The major restrictions are encountered when prior knowledge have to be included from past experiments or strong boundaries (like the positive value of a particle mass), or when probability value has to be assigned to different models instead of choosing one of them to fit a given experimental data set. These issues often occur in several fields of physics, including atomic and molecular physics, while Bayesian statistical methods are not so commonly used.

The Bayes' theorem and their lemmas permit to include prior knowledge in analysis of new data and to assign probabilities to different models describing a given spectrum in a very straightforward way. `Nested_fit` code [1] is based on these methods (more precisely on the nested sampling algorithm [2]) and is presented here. For a given data set and chosen model, the analysis program provides (i) the *Bayesian evidence* for the comparison of different hypotheses/models and (ii) the parameter probability and correlation distributions.

To give concrete illustrations, we consider a series of examples: the study of the presence or not of non-resolved satellite peaks in a high-resolution X-ray spectra of pionic atoms [3] or in a photoemission spectra of gold nanodots [4], and the analysis of low-statistics spectra, in a high-resolution X-ray spectrum of He-like uranium [5] or in a photoemission spectra of carbon nanodots [6]. In cases where the number of components cannot be clearly identified like for the He-like uranium spectrum (Fig. 1 inset), we show that the main component can anyway be determined. The probability distribution of a main line position can in fact be obtained by the distributions from the single models weighted by the probability assigned to the different models. This particular example shows the potential of Bayesian statistics compared to standard model selection criteria, as the likelihood-ratio test and Akaike information criterion, that can indicate the favor of a model with respect to another, but cannot handle critical cases where a clear favor is absent.

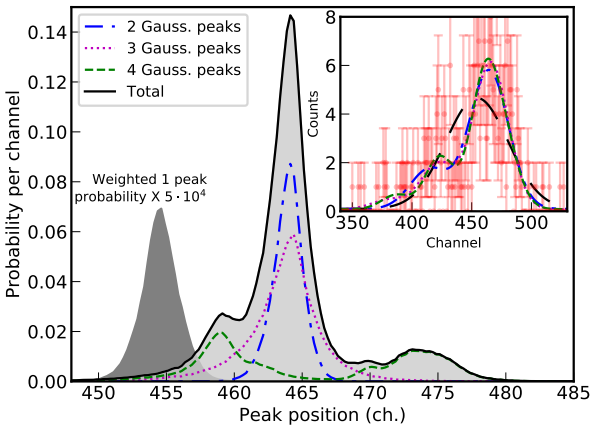

Figure 1. Probability distribution of the main peak position from the weighted average of different model [1]. In the inset, the corresponding fits are presented with the data.

References

[1] Trassinelli M 2017 *Nucl. Instrum. Methods B* **408** 301
[2] Skilling J 2006 *Bayesian Anal.* **1** 833
[3] Trassinelli M, Anagnostopoulos D F *et al.* 2016 *Phys. Lett. B* **759** 583
[4] Lévy A, De Anda Villa M *et al.* submitted to *Langmuir*
[5] Trassinelli M, Kumar A, Beyer H F *et al.* 2009 *Eur. Phys. Lett.* **87** 63001
[6] Papagiannouli I, Patanen M, Blanchet V *et al.* 2018 *J. Phys. Chem. C* **122** 14889

*E-mail: martino.trassinelli@insp.jussieu.fr