Many-probes multi-object spatially-resolved analyses of galaxy clusters in the big data era

Fabio Castagna[1,2,*], *Stefano* Andreon[2], *Alberto* Trombetta[1], and *Marco* Landoni[3]

[1]Department of Pure and Applied Sciences, Computer Science Division, Insubria University, 21100 Varese, Italy

[2]INAF–Osservatorio Astronomico di Brera, via Brera 28, 20121 Milano, Italy

[3]INAF-Osservatorio Astronomico di Brera, via Emilio Bianchi 46, 23807 Merate, Italy

Abstract. The thermal Sunyaev-Zeldovich (SZ) effect and the X-ray emission offer separate and highly complementary probes of the thermodynamics of the intracluster medium, particularly on their radial dependence. We already released JoXSZ, the first publicly available code designed to jointly fit SZ and X-ray data coming from various instruments to derive the thermodynamic radial profiles of galaxy clusters, including mass. JoXSZ follows a fully Bayesian forward-modelling approach, adopts flexible parametrization for the thermodynamic profiles and includes many useful options that users can customize according to their needs. We are including shear measurement in our joint analysis, and moving from single-cluster to multi-cluster analyses, allowing to quantify the heterogeneity of thermodynamic properties within the cluster population. At the same time, we are creating a suitable framework that efficiently stores and optimally processes huge volumes of data being released by the current and new generation surveys.

1 Introduction

Galaxy clusters are the largest and most massive gravitationally bound objects in the Universe, and thus they offer a unique tracer of cosmic evolution [1]. The thermodynamic properties of a galaxy cluster can be probed through observations in the optical, X-ray, or microwave bands via the Sunyaev-Zeldovich (SZ) effect [2, 3], that we are combining into a single joint analysis.

We have already released JoXSZ [4, source code available on GitHub[1]], the first publicly available code designed to jointly fit SZ and X-ray data to derive the thermodynamic profiles of galaxy clusters.

We are currently working to include shear data as a third complementary component in the fit and to move from single-cluster to multi-cluster analyses, allowing to gather thermodynamic measurements on populations of clusters. Since it will involve large amount of observations coming from current and next generation facilities, we are creating a suitable framework that efficiently stores and optimally processes huge volumes of data.

---

[*]e-mail: fabio.castagna@inaf.it

[1]https://github.com/fcastagna/joxsz

2 `JoXSZ`

`JoXSZ` is a Python code that combines SZ and X-ray data into a single analysis following a fully Bayesian forward-modelling approach. `JoXSZ` accounts for beam smearing and data analysis transfer function, for the SZ calibration uncertainty and X-ray and SZ background level systematics. It adopts extremely flexible parametrization for the thermodynamic profiles and it employs a consistent temperature across the various parts of the code, allowing for differences between X-ray and SZ gas mass weighted temperatures when required by the user, and calculates the correct Poisson-Gauss expression for the joint likelihood.

    `JoXSZ` is built upon the SZ data fitting pipeline described in `PreProFit` [5, source code available on GitHub[2]] and an updated version of the X-ray data cube fitter `MBProj2`, originally developed by [6]. `JoXSZ` merges and extends these two processes into a unique joint and consistent model based on a Markov chain Monte Carlo (MCMC) fitting algorithm.

2.1 Inputs and Outputs

For each cluster, `JoXSZ` requires the user to specify the SZ surface brightness radial profile, and the X-ray data cube in the form of multiple energy band radial profiles. The SZ analysis requires the point spread function (PSF), the SZ transfer function (a correction for data affected by filtering), and the conversion factor from Compton parameter to the observed data unit with temperature dependence (see [5] for more details). In X-ray, by using a data cube we exploit the information about temperature contained in the X-ray spectra. The components required for processing X-ray data, *i.e.* the response matrix file (RMF) and the ancillary response file (ARF), must be provided by the user (see [4] for more details).

    Through MCMC computation, `JoXSZ` samples the joint and marginal posterior distribution of parameters and derives the radial profiles of main thermodynamic properties, *i.e.* pressure, electron density, temperature, entropy. Other thermodynamic profiles are computed as usual [6], including mass profile and gas fraction under the assumption of hydrostatic equilibrium.
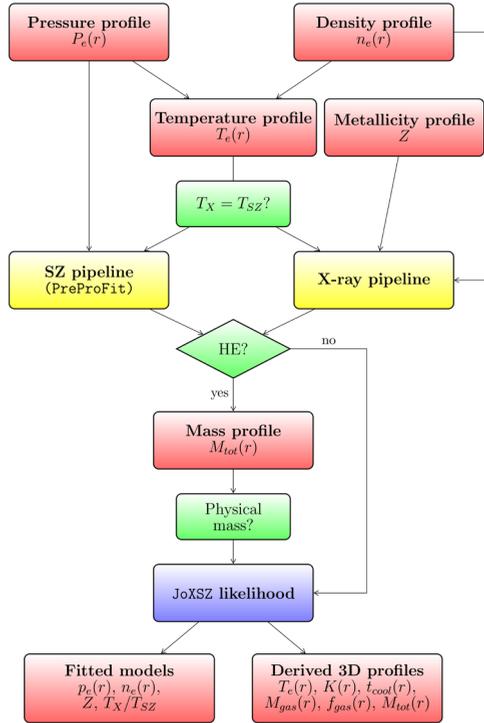
2.2 Program flow

Figure 1 depicts the block diagram describing the program flow of `JoXSZ`. In the SZ pipeline, outlined in Fig. 2, `JoXSZ` parametrizes the pressure profile, then projects it onto a two-dimensional map using forward Abel transform, convolves the map with the instrumental beam and the transfer function, and finally derives the surface brightness profile through appropriate conversion factors. If the beam file already includes the transfer function, the user has to specify it and `JoXSZ` will correctly handle it. Regarding the X-ray analysis, shown in Fig. 3, `JoXSZ` makes full use of the spatial-spectral X-ray data cube, profiting of the better angular resolution of X-ray data, when available, to measure thermodynamic profiles at small radii unprobed by lower resolution SZ data.

    `JoXSZ` is compatible with SZ data coming from different sources: NIKA, MUSTANG and SPT, whose provenence should be indicated because teams releasing the data adopts different standards. After specification, `JoXSZ` will process the data accordingly. `JoXSZ` has been already used for data coming from Chandra, NIKA and MUSTANG [4, 7], and is being used for data from XMM, XRT, SPT, and Planck [8]. `JoXSZ` is also able to read data in different formats, and to automatically handle measurement units throughout the code using `astropy.units` [9] upon correct specification from the user of the measurement units of input data.

---

[2]https://github.com/fcastagna/preprofit

Figure 1. Block diagram showing the program flow. Radial profiles are in red. Options are in green. Analysis pipelines are in yellow. Data enter in the blue box.

## 2.3 Thermodynamic profiles modelization

Users are allowed to formulate the pressure profile through either parametric model or binned profile: a generalized Navarro, Frenk & White (gNFW) model [10], a cubic spline model or a power-law interpolation model. Optionally, users can apply a prior constraint on the outer slope of the pressure profile. The multi-parameter (up to 10) model introduced by Vikhlinin [11] is used to parametrize the electronic density profile, and given the large number of possible parameters it should be flexible enough to satisfy the needs of most users.

The temperature profile is simultaneously influenced by the pressure constraints from SZ observations and the X-ray data cube, *i.e.* the X-ray spectra at various distances from the cluster center. JoXSZ allows users either to consider a unique temperature profile $T_{SZ} = T_X$, or to make a distinction between the gas mass weighted temperature T_{SZ} and the X-ray temperature T_X, introducing the multiplicative parameter $\log(T_X/T_{SZ})$.

JoXSZ assumes a flat metallicity profile, whose value can be fitted, or fixed, at user request.

2.4 Additional functionalities

Additional functionalities are implemented in the code, each user can choose whether to adopt them or not. By default, JoXSZ includes a calibration parameter accounting for uncertainties in the SZ measurements and a X-ray background scaling parameter that accounts for
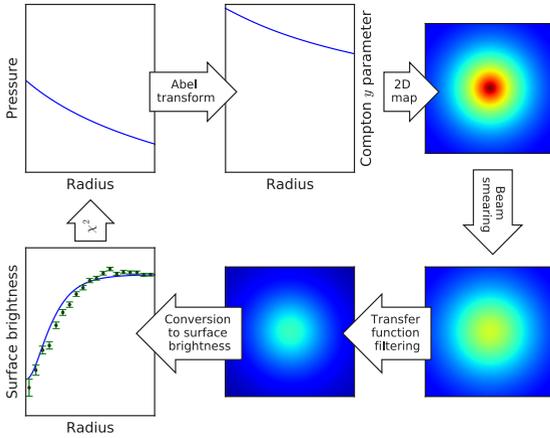
Figure 2. Graphical representation of the SZ single-cluster analysis pipeline within JoXSZ. The multi-object version loops over the clusters.
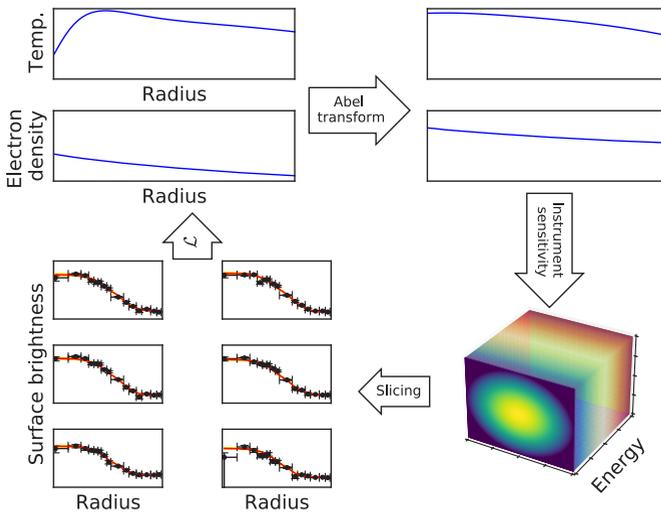
Figure 3. Graphical representation of the X-ray single-cluster analysis pipeline within JoXSZ. The multi-object version loops over the clusters.

differences in background level between the cluster and control field directions. In the SZ analysis, a pedestal parameter for the surface brightness can be added in the case of maps with non-zero level. Another option is to apply a prior constraint on the integrated Compton parameter. JoXSZ can be easily used for feasibility studies adopting simulated data, e.g. Gaussian approximations for the PSF and the transfer function.

If dealing with a large filtering image, JoXSZ gives the opportunity to neglect extremely small scales to increase the execution speed, but at the user's own risk.

To further save CPU time, users may ask to start the exploration of the parameter space not far from the likelihood maximum by specifying a (guessed) value for the cluster's characteristic radius, from which pressure parameters are derived assuming the universal profile defined by Arnaud et al. [12] as first approximation. This option is also available for a binned pressure model: JoXSZ automatically converts the parameters from the universal profile into the required binning values.

3 Current and future developments

To fully take advantage of the multi-probe potential of current and incoming surveys, such as Vera Rubin Observatory, Euclid, eROSITA, Roman Space Telescope, ACT, and SPT, we are currently adding shear measurements to our joint SZ and X-ray analysis. At the same time, we are paving the way to extend the analysis from single clusters at a time to multiple clusters at the same time, meaning that parameter estimation will be simultaneously conducted at both individual level and population of clusters level through a Bayesian hierarchical model. Such a tool will be of great interest for the astrophysical community, giving the opportunity to answer more complex questions [13], e.g. testing General Relativity [14] or assessing the heterogeneity of thermodynamic properties within the cluster population, such as the scatter in pressure or mass profiles.

Since the code of `JoXSZ` includes highly time-consuming operations, we are currently working on optimizing the computation, both reducing redundant operations and exploiting parallel computing techniques. In view of the extension from single to multi-object analyses, we need to develop typical features of Big Data Analytics, in order to transform the huge avalanche of data into accurate information. To do so, we are constructing an adequate infrastructure which is based on state-of-the-art providers for data storage (Amazon Web Services, e.g. see Landoni et al. [15]) and data processing (Apache Spark [16]). Frameworks like Apache Spark allow us to define very specific functionalities on datasets of large dimensions, and to define computation procedures that can be easily scaled on parallel and distributed architectures. Furthermore, they are able to integrate data analysis techniques based on deep learning, such as convolution neural networks, recently used to estimate cluster masses through X-ray data.

References

[1] G.M. Voit, Reviews of Modern Physics **77**, 207 (2005), `astro-ph/0410173`

[2] R.A. Sunyaev, Y.B. Zeldovich, Ap&SS **7**, 3 (1970)

[3] R.A. Sunyaev, Y.B. Zeldovich, Comments on Astrophysics and Space Physics **4**, 173 (1972)

[4] F. Castagna, S. Andreon, A&A **639**, A73 (2020), `2005.12315`

[5] F. Castagna, S. Andreon, A&A **632**, A22 (2019)

[6] J.S. Sanders, A.C. Fabian, H.R. Russell, S.A. Walker, Monthly Notices of the Royal Astronomical Society **474**, 1065 (2018), `1705.09299`

[7] S. Andreon, C. Romero, F. Castagna, A. Ragagnin, M. Devlin, S. Dicker, B. Mason, T. Mroczkowski, C. Sarazin, J. Sievers et al., MNRAS **505**, 5896 (2021), `2106.11327`

[8] S. Andreon, et al., in prep. (2021)

[9] Astropy Collaboration, T.P. Robitaille, E.J. Tollerud, P. Greenfield, M. Droettboom, E. Bray, T. Aldcroft, M. Davis, A. Ginsburg, A.M. Price-Whelan et al., A&A **558**, A33 (2013), `1307.6212`

[10] D. Nagai, A.V. Kravtsov, A. Vikhlinin, ApJ **668**, 1 (2007), `astro-ph/0703661`

[11] A. Vikhlinin, A. Kravtsov, W. Forman, C. Jones, M. Markevitch, S.S. Murray, L. Van Speybroeck, The Astrophysical Journal **640**, 691 (2006), `astro-ph/0507092`

[12] M. Arnaud, G.W. Pratt, R. Piffaretti, H. Böhringer, J.H. Croston, E. Pointecouteau, A&A **517**, A92 (2010), `0910.1234`

[13] E.D. Feigelson, R.S. de Souza, E.E.O. Ishida, G. Jogesh Babu, Annual Review of Statistics and Its Application **8**, 493 (2021), `2005.13025`

[14] V.F. Cardone, P. Karmakar, M. De Petris, R. Maoli, Phys. Rev. D **103**, 064065 (2021), `2008.00580`

[15] M. Landoni, P. Romano, S. Vercellone, J. Knödlseder, A. Bianco, F. Tavecchio, A. Corina, ApJS **240**, 32 (2019), `1901.00410`

[16] M. Zaharia, M. Chowdhury, M.J. Franklin, S. Shenker, I. Stoica et al., HotCloud **10**, 95 (2010)