XMM-Newton **Calibration Technical Note**

FITTING X-RAY SPECTRA ANALYSIS: χ^2 vs C-statistics

A.M. Pérez

July 27, 2020

1 Introduction

The purpose of this document is test C-stat and χ^2 statistics, to conclude which is the best way to fit X-ray spectra (at high energies, E > 3 keV).

Typically, X-ray spectral fitting has been done using χ^2 statistics, using the background-subtracted spectra, obtained from the observed spectra and the corresponding background file. However, X–ray spectra of statistical sources are often characterized by relatively low number of counts per spectral bin and it is a common practice to rebin the spectra to obtain a relatively high number of counts per bin (that implies loss of spectra details). On the other hand, χ^2 implies that the statistics is Poissonian. To properly use the Poissonian statistics, the application of χ^2 should be fitting source plus background spectra.

$$\chi^2 = \sum_{i=1}^{n} \frac{(N_i - s_i)}{\sigma^2} \tag{1}$$

C-statistics is a much better statistics to be applied in spectra with bins with a low number of counts. Then, this method can be used to fit simultaneously the source+background spectra and the corresponding background. An advantage of this approach is that this statistics can be used to obtain uncertainty ranges on the parameters of the model. However, C-stat has not a similar value to χ^2 that could be interpreted as a goodness of fit indicator. The C-stat implementation in XSPEC [1] is a modification of the original one proposed by [2], and is expressed as:

$$C = 2\sum_{i=1}^{n} s_i - N_i + N_i \ln(\mathrm{N_i/s_i}) \tag{2}$$

For a spectrum with high number of counts per bin, $C \to \chi^2$, but if the number of counts is small, C will be significantly smaller than the number of bins. [4] has developed a simple numerical approximation to evaluate the goodness of fit using C-statistics. These formulas has implemented in python scripts.

For this work, we will use XMM–EPIC (PN and MOS) source+background spectra of several sources (GRB111209A,IRAS13224-3809,LBQS1228+1116,NGC5408X-1,OJ287). In general, the model used to fit the source (in the range 2-12keV) is a power-law, except for NGC7172.

The software used for this analysis is pyXspec.

2 Background Parameters

The X-ray background is very complex. In general, the spectral components in the 0.7-12keV are:

- the emission from the Galaxy Halo (HALO)
- the cosmic ray induced continuum (NXB)
- the cosmic X-ray background (CXB)
- the local hot bubble
- the quiescent soft protons (QSP)
- the fluorescence lines

Figure 1 shows an example of fitting in a region of Abell2142 with all background components. Note that this fitting includes the source. In this work, we are interested in the energy range 2-12keV. Then, the emission of Galactic Halo and the local hot bubble are negligible.

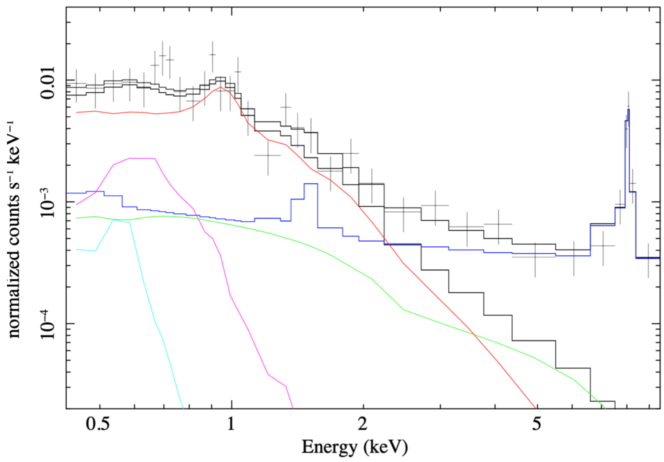

Figure 1: EPIC-PN spectra and fit for a region in Abell2142. The solid lines show the various components used for the fitting procedure: the source (red), the NXB (blue), the CXB (green), the Galactic halo emission (magenta), and the local hot bubble (cyan). Credits: [3]

The problem lies with the high number of free parameters to set and the low S/N of background spectra. To solve the first issue, we follow the same procedure that previous works [3, 5], fixing several parameters (see below). On the other hand, we use a high S/N background spectra, obtained from a larger area on the detector. We have tried different combinations of the single components present in the background model and finally, the best fit obtained for the XMM-EPIC background has three components:

- an absorbed power–law (CXB), with fixed Γ_{CXB}=1.46
- a single power–law (NXB), with fixed Γ_{NXB}=0.24
- several gaussians to fit fluorescence lines (see Table A2 in [5]) plus a Cu escape peak at 6.3 keV produced by Cu-Kα)

The result of this best fit for M1 and PN spectra is shown in Figure 2. We will use these parameters to fit with C-statistics the spectra (see Section 4), fixing the exponents of power-laws models, energy and width of fluorescence lines, leaving free all normalization parameters.

3 χ^2 fitting

The fit of the EPIC-XMM spectra of the sources has been done fitting simultaneously PN and MOS spectra. For the sources, the 2–12keV is well defined with a power-law single model, excepting NGC7172 spectrum, that is a composed model of an absorbed power law plus a gaussian (Fe Kα emission line). Figure 3 shows the results of χ^2 fitting for all sources. The spectra used in this method are binned to have a high enough level of counts per bin. The error in parameters obtained for all sources fitted with a power-law single model are in Table 1, using *error* command. For NGC7172, the parameters (and errors) are in Table 2. Note that,

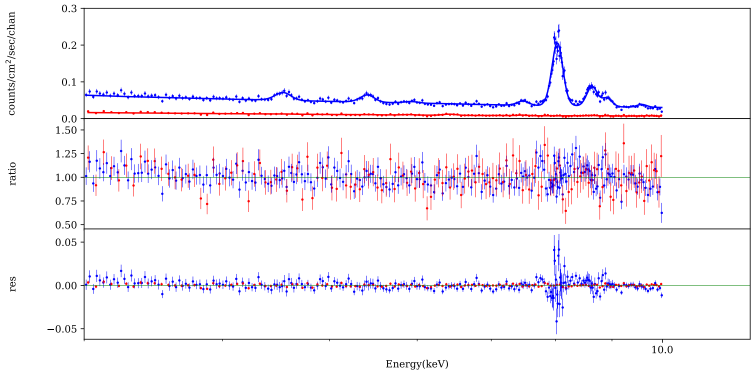

Figure 2: XMM-EPIC fit of background spectra. PN: Blue line, M1: Red line

for fitting, it has been forced that some parameters (Γ in power-laws, Line Energy and σ, nH) are the same for PN and MOS spectra, leaving free normalizations. The errors are well defined, and lower than 7% for the Γ parameter. Note that, in χ^2 fitting, we always fit subtracting the background, because chi^2 statistics does not work with spectra with a low number of counts.

Source	Γ	Γ_{err}(90%)	NormPN	NormPN$_{err}$(90%)	NormMOS	NormMOS$_{err}$(90%)
GRB111209A	1.83	-0.12/0.12	2.08e-4	-3.65e-5/4.43e-5	2.13e-4	-3.72e-5/4.52e-5
IRAS13224-3809	2.10	-0.09/0.09	2.77e-4	-3.58e-5/4.12e-5	2.97e-4	-3.73e-5/4.29e-5
LBQS1228+1116[1]	1.71	-0.09/0.09	2.70e-4	-3.54e-5/4.07e-5	3.06e-4	-3.88e-5/4.45e-5
LBQS1228+1116[2]	1.72	-0.06/0.06	3.01e-4	-2.94e-5/3.24e-5	3.18e-4	-3.04-5/3.36e-5
NGC5408X-1[1]	3.09	-0.14/0.15	7.61e-4	-1.43e-4/1.79e-4	8.50e-4	-1.59e-4/2.01e-4
NGC5408X-1[2]	2.95	-0.21/0.22	6.80e-4	-1.83e-4/2.56e-4	7.31e-4	-1.90e-4/2.65e-4
OJ287	1.66	-0.09/0.09	3.91e-4	-5.02e-5/5.74e-5	4.15e-4	-5.13e-5/5.86e-5

Table 1: Model parameter results with χ^2 fitting method

To analyze the goodness of fit, we vary Γ and normalization covering a wide range of values (using *steppar* command), and obtain for each pair, a value of χ^2. Results are plotted in Figure 4. The good fits, that correspond to lowest values of $f(\chi^2) = exp(-\chi^2)$ are concentrated in a small range for both parameters.

4 C-statistics fitting

To use C-stat fitting, as we mentioned before, we have implemented a numerical approximation to obtain a C parameter that it can be used for testing the goodness of fit. With this method, we fit separately MOS and PN spectra, each one with its appropriate background. Then, the model selected is composed by a power-law plus the three components of background defined in section 2. Figure 5 shows several fits for sources using this statistics. The parameters obtained are in Table 3.

We try to obtain errors in parameters and goodness of fit in the same way that with χ^2 statistics (using *error* and *steppar* commands), but it does not work. Next step is made simulations to perform an analysis of the goodness of fit using C-statistics.

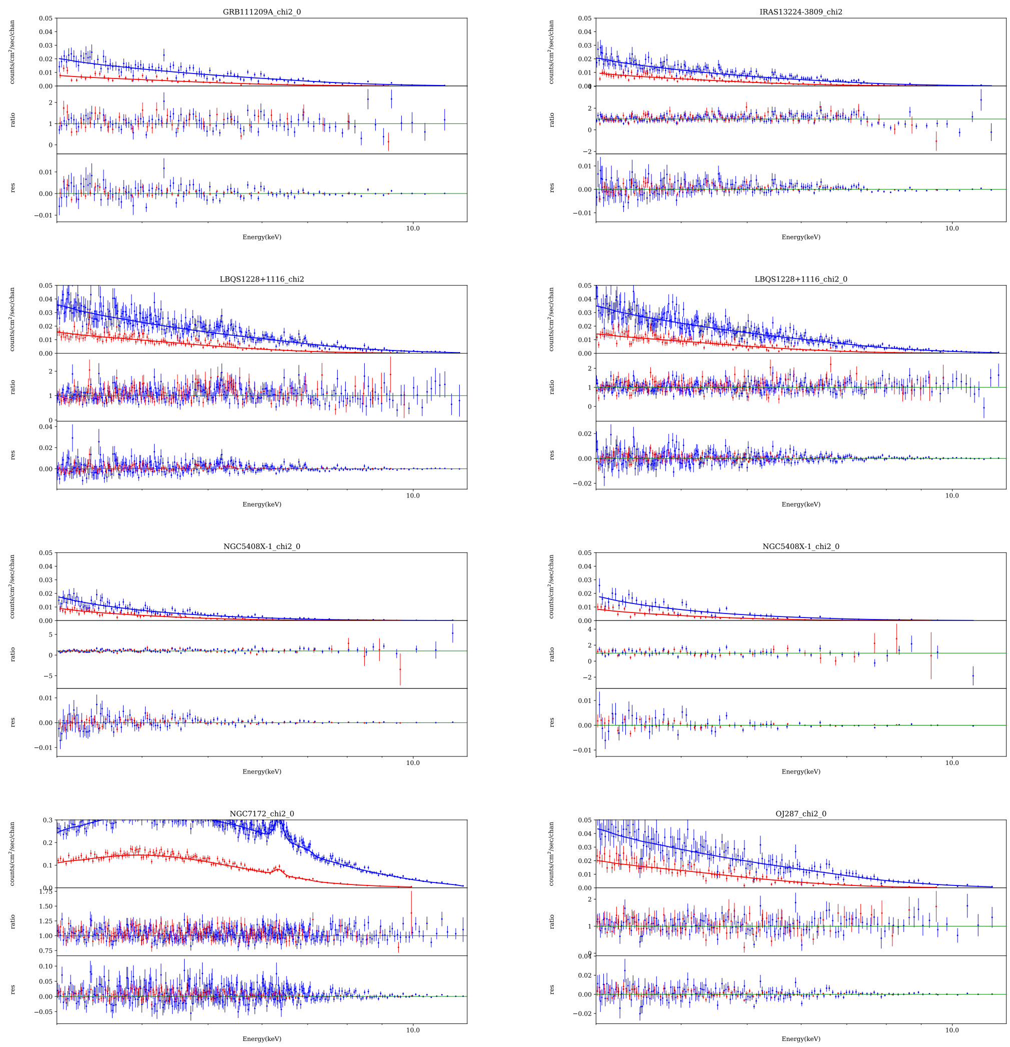

Figure 3: Fitting for all sources with a power–law model using χ^2 statistics. PN (blue points data; blue line model) and MOS (red points data; red line model) are plotted.

5 Simulations

For each spectra, we build a sample of simulated spectra (5000), based on the correspondent observational one. Then, using the same models and restrictions that we describe in previous sections, we repeat the fitting procedure for these simulated spectra. We apply both statistics, χ^2 and C.

First, we study the distributions of χ^2 and C parameters. In all cases, we can accept that the models fitted are suitable, as shown in Figure 7. In this plot, we present several examples (for NGC5408-X and LBQS1228+116 sources) of χ^2 and C distributions.

Finally, we analyze the variation of Γ and normalization parameters in the simulations. As we shown in Figure 8, the ranges covered in simulations are smaller than those obtained with *steppar* (for χ^2 statistics, see Figure 6), although both density plots are compatible. For C-statistics, the range is wider (mainly in the normalization parameter), but not sufficient to obtain conclusive results. In both cases, the best fitting values are concentrated in a small range of Γ and normalization values.

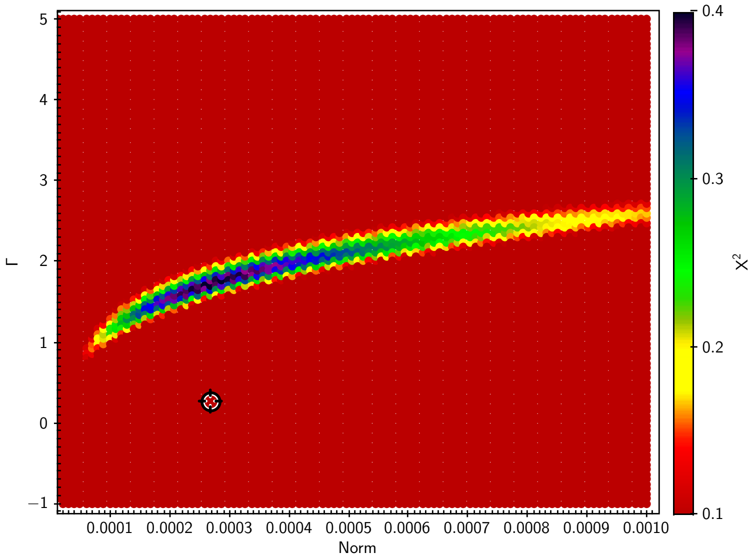

Figure 4: Γ vs normalization with χ^2 fitting density plot (for source LBQS1228+116). The best values are in blue

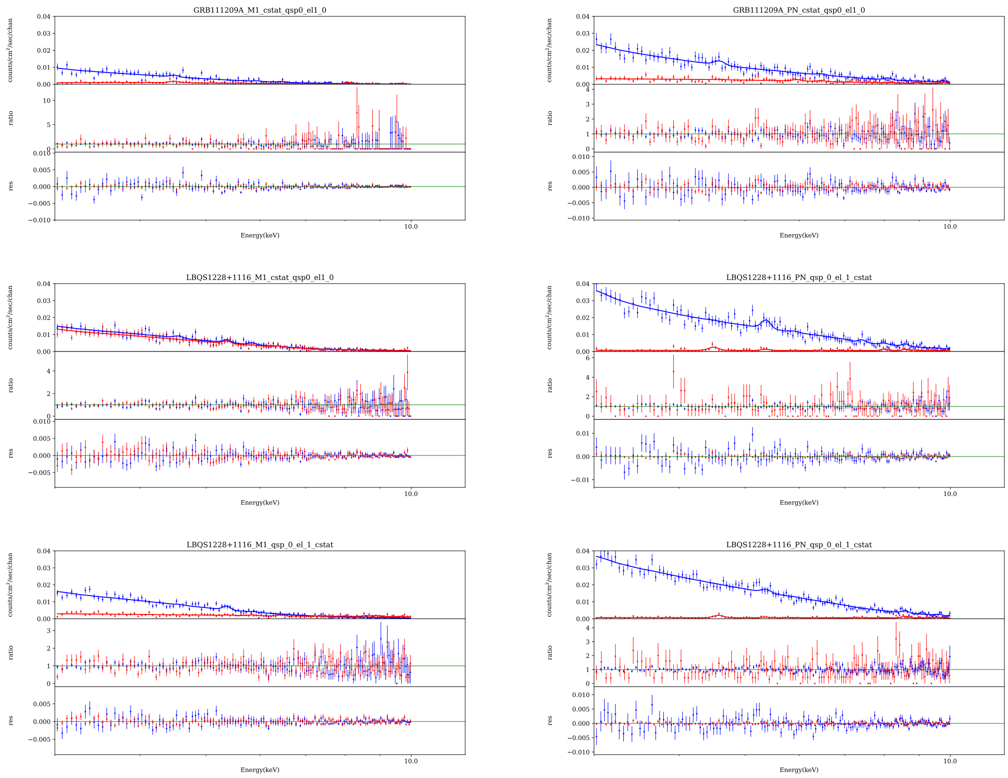

Figure 5: Fitting for several sources with a power–law model using C-statistics. Left: MOS (blue points data; blue line model) and background (red points data; red line model) spectra. Right: The same, for PN spectra

6 Conclusions

- The background in the range 2–12keV is well characterized by an absorbed power–law (CXB), with fixed Γ_{CXB}=1.46, a single power–law (NXB), with fixed Γ_{NXB}=0.24

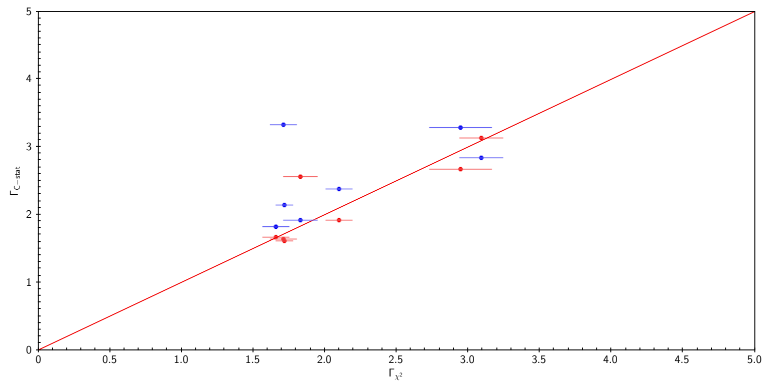

Figure 6: Γ obtained with χ^2 fitting vs Γ from C-stat fitting. Red: PN, Blue: MOS

and several gaussians to fit fluorescence lines plus a Cu escape peak at 6.3 keV produced by Cu-Kα).

- The results obtained with χ^2 and C-stat are, in general, compatible, but not identical.
- Both methods are stable, as shown our simulations. The χ^2 behavior is better than C-statistic one, but it should be noted that the spectra used for this statistics are binned, and then there are details that can be lost.
- Errors obtained with χ^2 statistics are quite small.

To obtain more conclusive results, it would be required:

- Fitting with C-stat MOS and PN spectra (and its correspondent background) simultaneously, to define better the parameters (Γ of the power-law) of the model
- Implementing a method to quantity the parameter errors in C-stat fitting, in the same way that with χ^2 statistics.

7 References

[1] Arnaud, K.A., 1996, ASPC 101,17A

[2] Cahs, W., 1979, ApJ 228, 939

[3] Eckert, D., et al., 2014, A&A 570, 119

[4] Kaastra, J.S., 2017, A&A 605, A51

[5] Leccardi, A. & Molendi, S., 2008, A&A, 486, 359

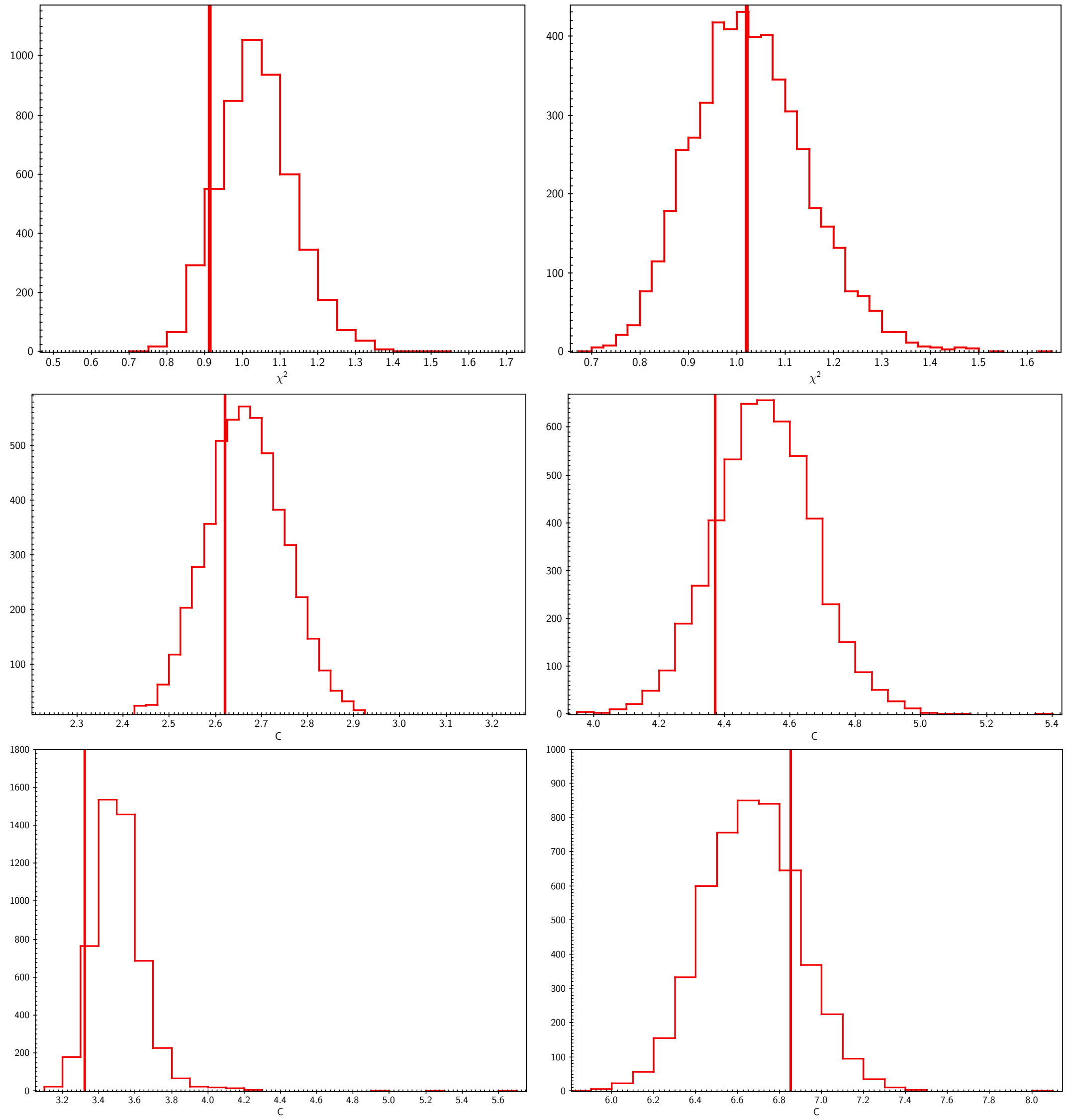

Figure 7: Distribution of χ^2 and C for NGC5408-X (left) and LBQS1228+116 (right) obtained from simulations. Red vertical line is the value correspondent to the observational spectrum fitting. From up to bottom: χ^2, C for MOS1 and C for PN spectra.

nH1	$nH1_{err}$(90%)	LineE	$LineE_{err}$(90%)	σ	σ_{err}(90%)	nH2	$nH2_{err}$(90%)	Γ	Γ_{err}(90%)	N1	$N1_{err}$(90%)	N2	$N2_{err}$(90%)
0.261	-0.26/14.36	6.407	-0.018/0.019	7.01e-2	-1.78e-2/1.90e-2	13.98	-13.79/0.79	1.66	-0.05/0.05	2.96e-5	-5.22e-5/1.17e-5	7.91e-5	-7.68e-4/8.63e-4
										3.82e-5	-9.75e-6/1.86e-5	8.47e-5	-8.12e-4/9.12e-4

Table 2: Model parameter results with χ^2 fitting method for NGC7172

Source	Γ_{MOS}	NormMOS	Γ_{PN}	NormPN
GRB111209A	2.56	2.46e-4	1.92	1.82e-4
IRAS13224-3809	1.92	6.49e-5	2.38	3.02e-4
LBQS1228+1116[1]	1.64	3.80e-5	3.32	1.46e-3
LBQS1228+1116[2]	1.61	2.49e-4	2.14	2.73e-4
NGC5408X-1[1]	3.13	9.38e-4	2.83	5.72e-4
NGC5408X-1[2]	2.67	5.34e-4	3.28	8.94e-4
OJ287	1.66	3.91e-4	1.82	4 9.37e-5

Table 3: Model parameter results with C-stat fitting method

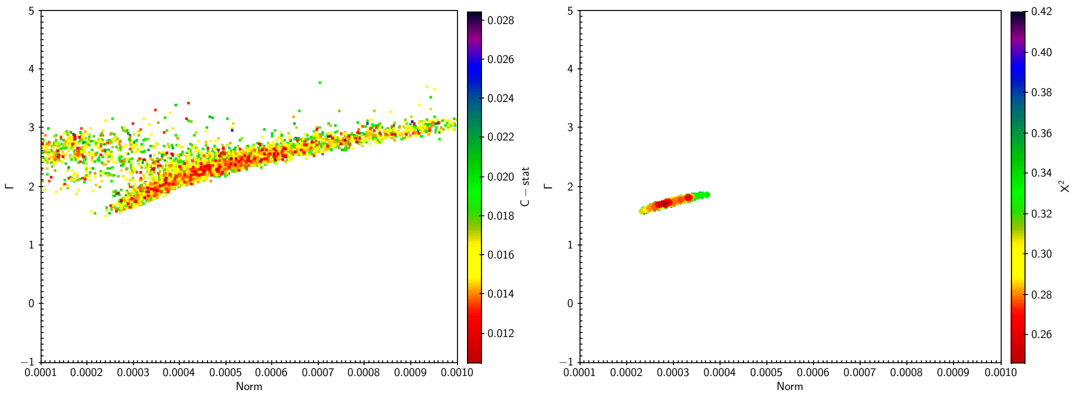

Figure 8: Γ vs normalization with χ^2 fitting density plot for simulations (for source LBQS1228+116). The best values are in red.