

# CROSS CALIBRATION FOR XMM-NEWTON



Fernando Bayon Rodriguez  
Mathematics Faculty  
Complutense University of Madrid  
[fbayon@sciops.esa.int](mailto:fbayon@sciops.esa.int)  
Tutors: M.Kirsch, M. Stuhlinger

# **Project: automated cross calibration fitting tool**

XMM-SOC has created an automated cross calibration tool to fit the X-ray data from XMM-Newton

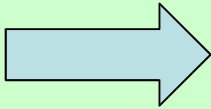
In order to “check in” new observations manual steps must be made such as  
region definition  
and  
finding the best fit models

# Throughout the presentation we will find out:

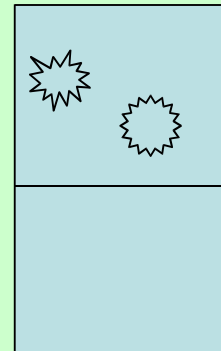
- Why pileup appears and how it can ruin our spectra
- How the fit concept appears and why it is helpful
- What mathematical and statistical tools and knowledge we use to do the fitting
- How we use these to know if our fit is a good one

# EPIC Pile-up

- It's the arrival of more than one photon in a pixel or adjacent pixels before it's read out  
2 photons or 1 with high energy?



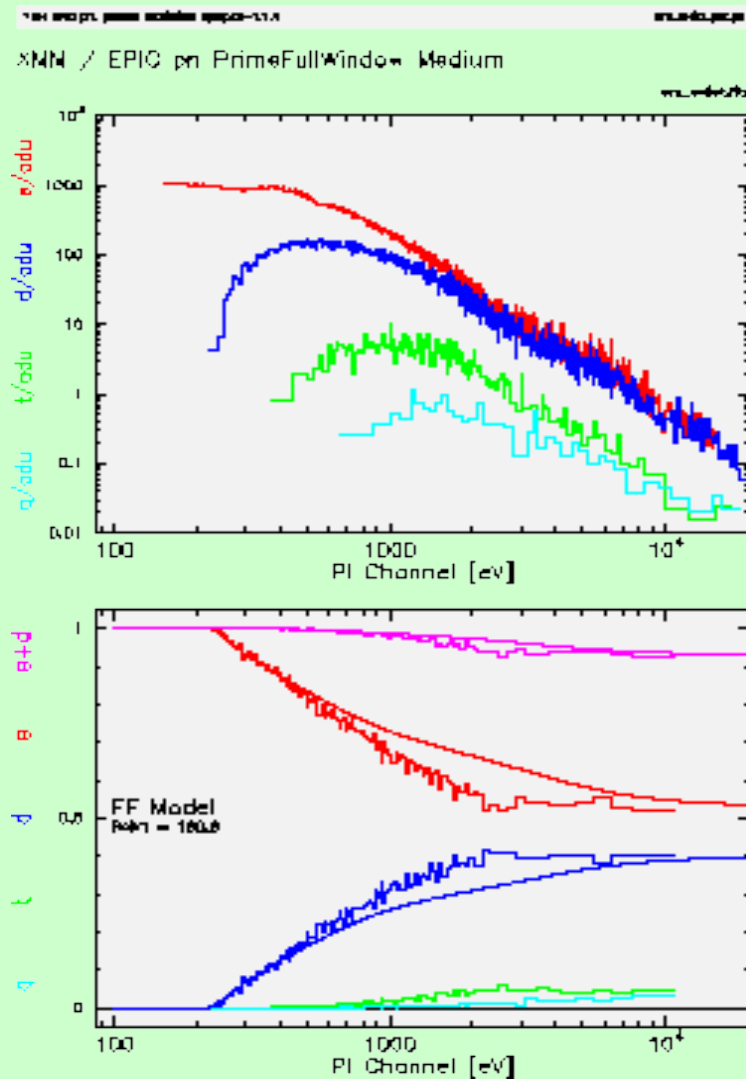
- Technology can not detect it
- Affects psf and the spectra



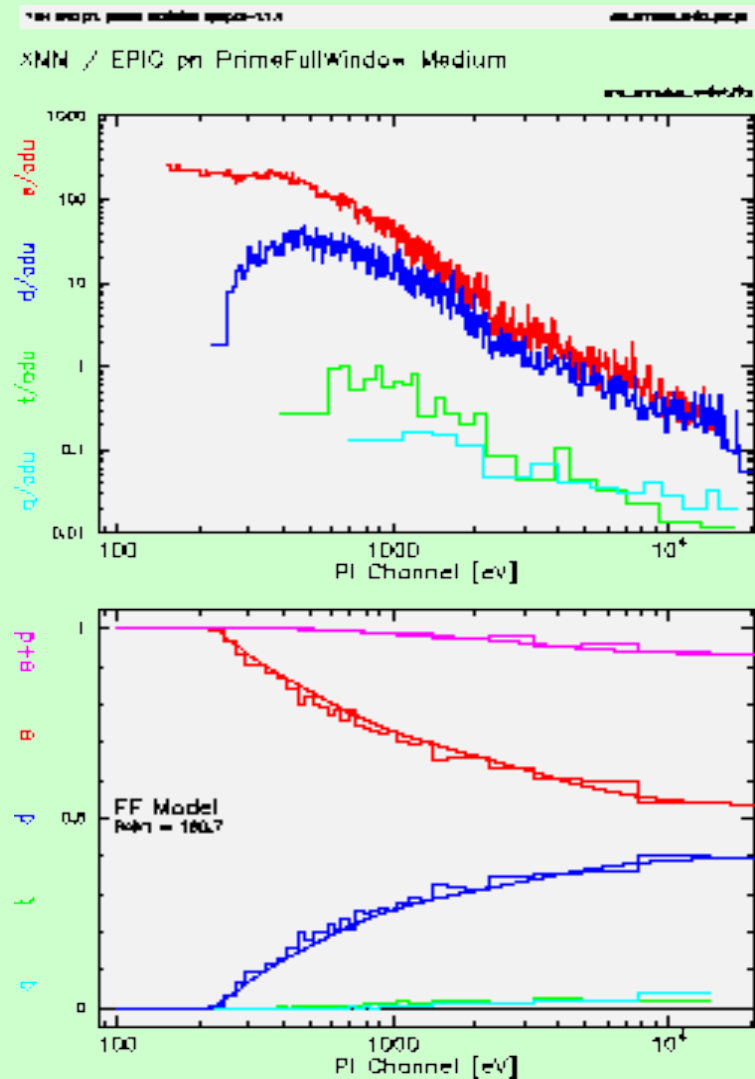
# Solving pile-up

- SAS epatplot gives relation of single and double events
- Compare with models
- Exclude regions with heavy pileup

# Source with pile up



# Clean of pile up



XMM-Newton does not give us the actual spectrum but photon counts from each instrument channel. Both are related as follows:

$$C(I) = \int_0^{\infty} f(E) R(I, E) dE$$

Known data:

$C(I)$ : photon counts

$R(I, E)$ : instrumental response

- Variable:

$f(E)$  = the actual spectrum

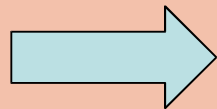
Our goal is to clear out  $f(E)$  as a function of all the rest known variables from the relation:

$$C(I) = \int_0^{\infty} f(E) R(I, E) dE$$



# The problem

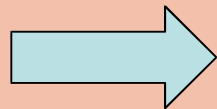
- Very tough restrictions (that usually do not occur) are needed to invert the equations. This makes inversions to be **non-unique** and **unstable** to small changes.



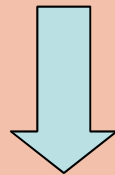
inversion **impossible!**

# The problem

- Very tough restrictions (that usually do not occur) are needed to invert the equations. This makes inversions to be **non-unique** and **unstable** to small changes.



inversion **impossible!**



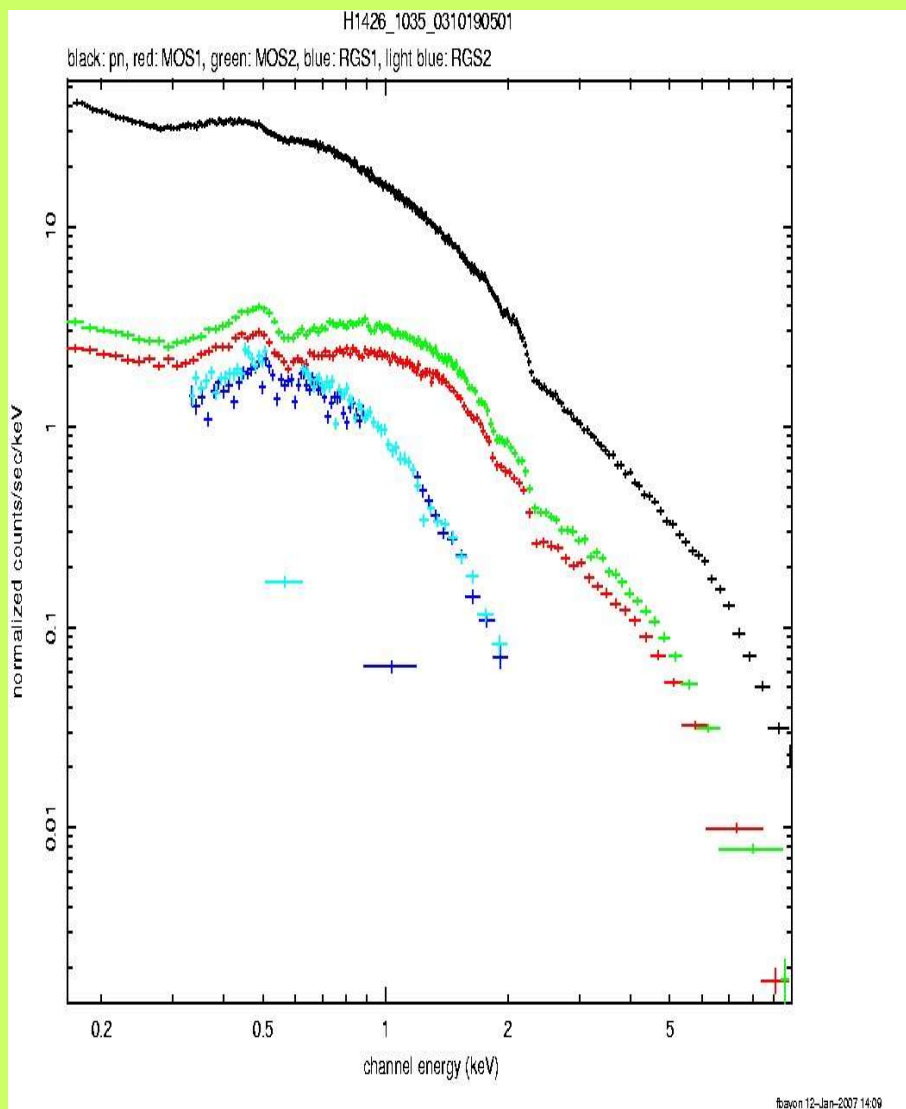
fit concept is needed

# The solution: fitting

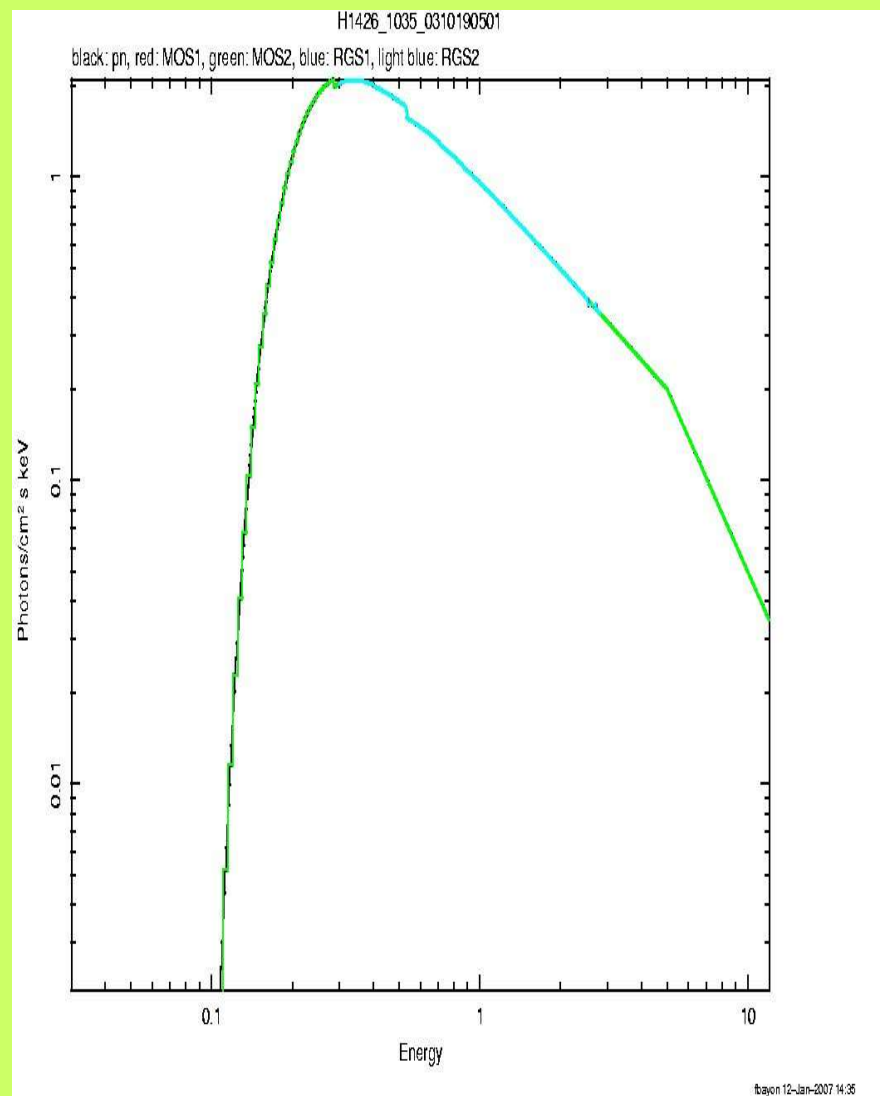
- The general idea is to find a simple model with few parameters that models well our data.

**First:** we choose a model spectrum in terms of few parameters  $F(E, p_1, p_2, \dots, p_n)$

# Observed spectrum C(I) for obs ID 0310190501



# Model Spectrum Phabs\*Brokenpowerlaw



**Then:** calculate the predicted count spectrum for our model in each channel.

$$C_m(I) = \int_0^{\infty} f(E, p_1, p_2 \dots p_n) R(I, E) dE$$

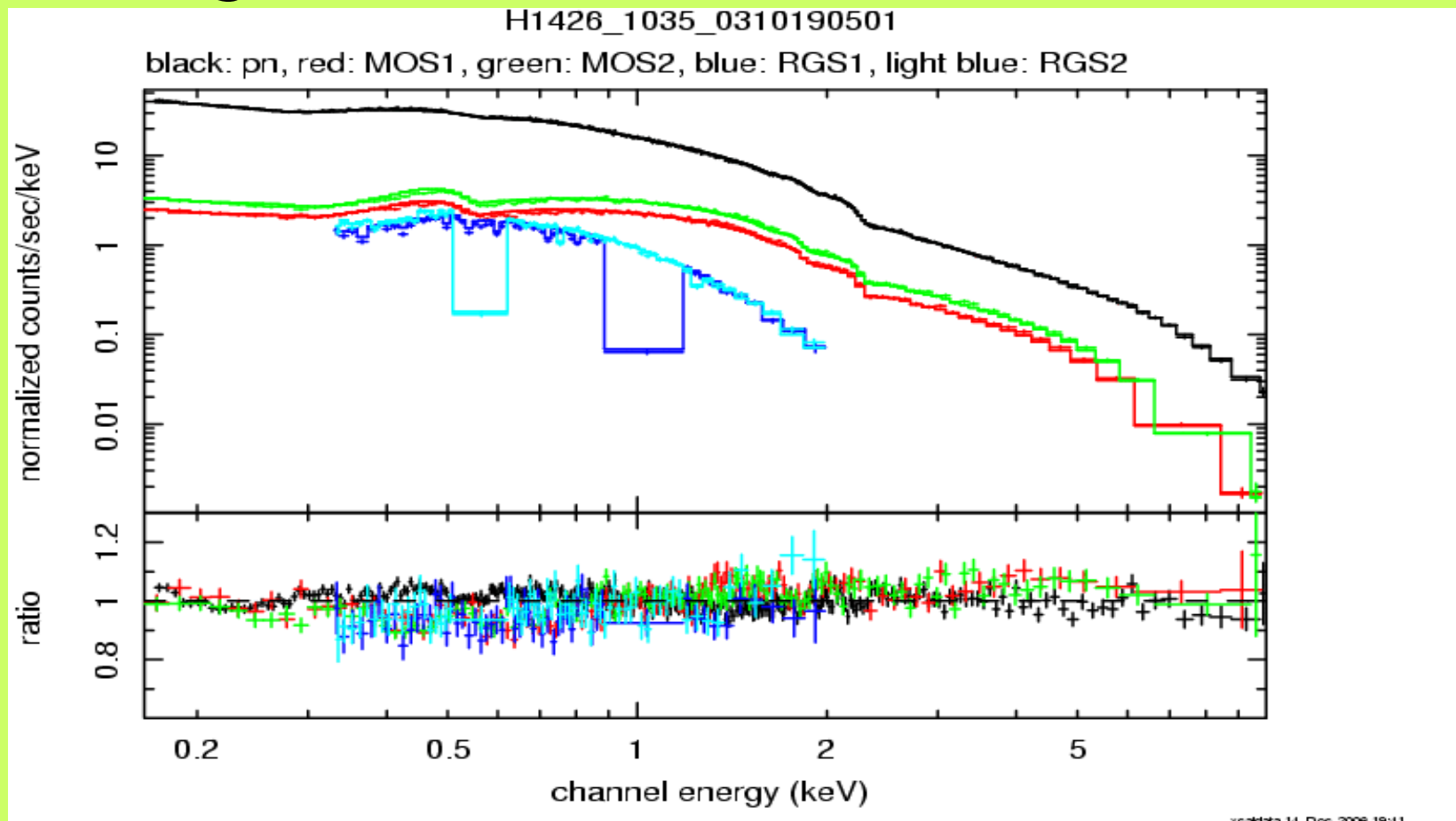
And **compare** this value to the observed one with the Chi-Squared statistic to test if we have chosen a good model

$$\frac{\sum_N (C(I) - C_m(I))^2}{\sigma(I)^2}$$

**Now:** vary the parameters  $p_1, p_2 \dots p_n$  to find the best fit

# To obtain:

- Fitting for 0310190501



# Comparison of model and data

- How do we know if the model we have chosen really ‘fits’ our source data? To study this we will consider the Chi-Squared statistic.

In general:

$$\sum_N \frac{(y_i - f(x_i))^2}{\sigma_i^2}$$

# In general:

$N$  = the number of data points

$\sigma_i$  = error made in iteration  $i$ , usually  
estimated by the square root of the number  
of iterations for  $i$

$y_i$  = model variable

$f(x_i)$  = actual value



# In our particular case:

$N$  = number of channels

$\sigma_i$  = error for channel  $i$

$C(i)$  = number of counts observed

$C_m(i)$  = folded model spectrum

So for our particular case the chi square is now:

$$\chi^2 = \sum_N \frac{(C(i) - C_m(i))^2}{\sigma_i^2}$$

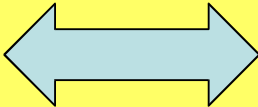
- If we have chosen a good model, the difference between the observed counts and the expected model counts in each channel should be about the same as the error expected for this channel.

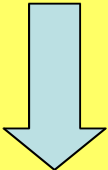
$C(i) - C_m(i) = \sigma(i)$  and if this happens it implies

$$\frac{C(i) - C_m(i)}{\sigma(i)} = 1 \quad \Rightarrow \quad \frac{(C(i) - C_m(i))^2}{\sigma(i)^2} = 1$$

We have this equality for all N channels. If we remember how we had defined the chi-squared we arrive to:

$$\chi^2 = \sum_N \frac{(c(i) - c_m(i))^2}{\sigma(i)^2} = N$$

In Xspec  $\chi^2 = N$    $\chi^2 = N - p = \nu$

p small compared to N 

degrees of freedom

So in general to see how good a model is we will look for:

The reduced  
Chi Squared

$$\frac{\chi^2}{\nu} = 1$$

# If chi-square test fails

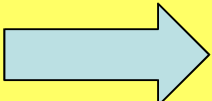
- If  $\frac{\chi^2}{\nu} > 1 \quad \Rightarrow \quad \frac{C(i) - C_m(i)}{\sigma(i)} > 1$

2 cases:

$C(i) - C_m(i)$  too big compared to  $\sigma(i)$

$\Rightarrow$  wrong model

OR  
 $\Rightarrow$  errors have been underestimated

- If  $\frac{X^2}{v} < 1$    $\frac{C(i) - C_m(i)}{\sigma(i)} < 1$

  $\sigma(i)$  too big

 errors have been overestimated

# My part of the project

- Defining pile-up free regions for observations in the archive: aprox 300 regions defined
- Finding the best fit model and parameters for the spectra

Many left in the archive