XMM-Newton Technical Note

XMM-SOC-CAL-TN-0233

# On the Rate and Energy-Dependent PHA Correction for EPIC-pn Fast Modes: Column by Column Analysis

Simone Migliari and Michael Smith

25 May 2022

## 1 Introduction

In EPIC-PN fast mode observations, there is evidence for a rate-dependent shift in PHA-channels of the collected photons: above a certain threshold, the higher the total count-rate the larger the shift to higher PHA-channels. Furthermore, within a single observation, the higher the spectral energy-channel, the larger the shift. The assumed origin of the rate-dependent shift is a higher charge transfer efficiency due to an overall increase in shifted charge. Hence, rather than simply the count rate, the most useful parameter to describe the effect is the rate of shifted electrons.

In Migliari & Smith (2019) and Migliari et al. (2020), we analysed all the archive EPIC-PN fast mode observations available at the time to calibrate their energy shifts as a function of shifted electron rate. One of the key choices in the analysis was to calculate the rate and extract the spectra from an area of the CCD of 21 columns, centered on the brightest. From the selected area, for each observation, we extracted two parameters: 1) the number of shifted electrons of the CCD caused by the impinging photons, and 2) the energy peak of a Gaussian fit to a proxy of the reference line edges. This way, we could analyse a plot with rate of shifted electrons vs reference energy for each observation. Given that, in absence of a rate-dependent energy shifts, the reference energy should not vary, we calculated the observed discrepancy to a constant energy line and created CCFs to correct for it, given the shifted electron rate of the observation.

The current CCF files (EPN_CTI_0053.CCF and EPN_CTI_0054) correct a rate dependent effect in the EPIC-pn Timing and Burst Modes that affects the energy-scale precision, and are used by default since SAS 20. This rate-dependent correction, called RDPHA, is calculated in the PHA space (see Guainazzi et al. 2013, 2014). The RDPHA is a third correction to be applied for energy-scale accuracy of the EPIC-pn timing and burst modes, the other two being X-ray Loading and the special gain correction (Guainazzi et al. 2014).

In the analysis described above, we deliberately lose information on the specific columns and on their possible different responses. In this work, we repeat the analysis column by

column for each EPIC-PN fast mode observation available. The goal is to assess the possible role of the column uniqueness in the energy calibration.

## 2 Sample and Data Analysis

The rationale of the RDPHA correction applied to Timing and Burst Mode observations is described in Guainazzi et al. 2013 and 2014. Herewith, following the definitions and nomenclature in Migliari & Smith (2019) and Migliari et al. (2020), we summarize the scope and methods and list the calibration procedure used for this analysis:

In order to assess the energy precision at low energy, we analysed the energy-channel location of the Si K-edge at $\sim 1.8$ keV and Au M-edge at $\sim 2.2$ keV, where the EPIC-pn has the strongest gradients of the effective area. Moreover, to assess the high energy range and investigate if there is any energy dependence in the energy shift, we also analyse the energy-channel location of the Au L-edge at $\sim 11.9$ keV. In order to simplify the fit of the edges and convert it to simple Gaussian fits in the channel-spectrum, Guainazzi et al. (2013, 2014) defined an empirical Color Ratio (hereafter, CR) that turns out to be very sensitive to changes in the effective area. The CR is defined as:

$$CR(PHA) \equiv \frac{|C(PHA) - C(PHA - Shift)|}{C(PHA) + C(PHA - Shift)} \quad (1)$$

where $C(PHA)$ are the counts in a particular PHA channel and $Shift$ was empirically determined to specifically emphasise those edges: $Shift = 25$ for the low-energy edges at $\sim 1.8$ and $\sim 2.2$ keV, and $Shift = 150$ for the high-energy edge at $\sim 11.9$ keV. Indeed, the CR forms sharp peaks at the Si K-edge, Au M-edge and at the Au L-edge. These peaks can be robustly fitted with a phenomenological model comprised of simple Gaussians plus a power law for the continuum.

The calibration procedure was the following.

1) We collected 450 timing-mode and 95 burst-mode EPIC-pn observations in the XMM-Newton science archive.

2) We determined if a source is variable with the file keyword=EXVAR. We excluded observations if $10 \times \sqrt{|EXVAR|/AVERAGE} > 1$.

3) We filtered out extended sources and sources with suspected pileup.

4) We filtered out the observations with no offset map and applied the XRL correction to the remaining sample.

5) We selected regions in the CCD corresponding to single columns for each of the observations of the sample.

6) We estimated the CR as a function of the PHA channel for each extraction region for the whole sample and fitted the spectrum with a model of Gaussians plus a power-law.

7) We applied a second filter to the sample based on the goodness of fit of the CR: we kept the observations for which the estimated best-fit values of the PHA peaks have a significance of more than $5\sigma$.

8) We collected the best-fit values of the PHA Gaussian peaks and plot them against the number $N_e$ of shifted electrons per pixel per second. $N_e$ is an instrumental proxy of the total count rate, calculated as:

$$N_e = \frac{\sum_{i=1}^{N_{ph}} E_i}{N_{pxl} \times T_{exp} \times 3.6} \tag{2}$$

where $E_i$ is the energy of the i-th photon (estimated using ADU=5 eV), $N_{ph}$ is the number of detected photons, $N_{pxl}$ is the number of pixels of the column where each spectrum was extracted, $T_{exp}$ is the exposure time and the factor 3.6 (in eV) is the energy required to produce an electron-hole pair (see also Guainazzi et al. 2014b).

9) For the remaining sample, we estimated the RDPHA correction by fitting the three edges correlations independently, using power-law functions:

$$\begin{array}{lll} PHA = A_1 & \text{for} & N_e < N_e^{low} \\ PHA = A_2 + A_3 Log(N_e) & \text{for} & N_e^{low} \leq N_e \leq N_e^{high} \\ PHA = A_4 & \text{for} & N_e > N_e^{high} \end{array}$$

where $N_e^{low}$ and $N_e^{high}$ are respectively the lowest and highest $N_e$ in the sample. $A_1$ and $A_4$ are determined by the condition of continuity with $PHA = A_2 + A_3 Log(N_e)$ at $N_e^{low}$ and $N_e^{high}$.

10) The goal, *in case of calibration improvement*, was to create a CCF file containing the values:

$$\begin{array}{lll} \Delta PHA = 0 & \text{for} & N_e < N_e^{low} \\ \Delta PHA = PHA(N_e) - PHA(N_e^{low}) & \text{for} & N_e^{low} \leq N_e \leq N_e^{high} \\ \Delta PHA = PHA(N_e^{high}) - PHA(N_e^{low}) & \text{for} & N_e > N_e^{high} \end{array}$$

where the correction algorithm in this case would depend on the column number, and would correct the RDPHA effect for each column independently within an observation.

For comparison with the plots that we will analyse in this work, we show in Fig. 1 and Fig. 2 the current best calibration sample for Timing and Burst Modes, as described in Migliari & Smith 2019 and Migliari et al. 2020.

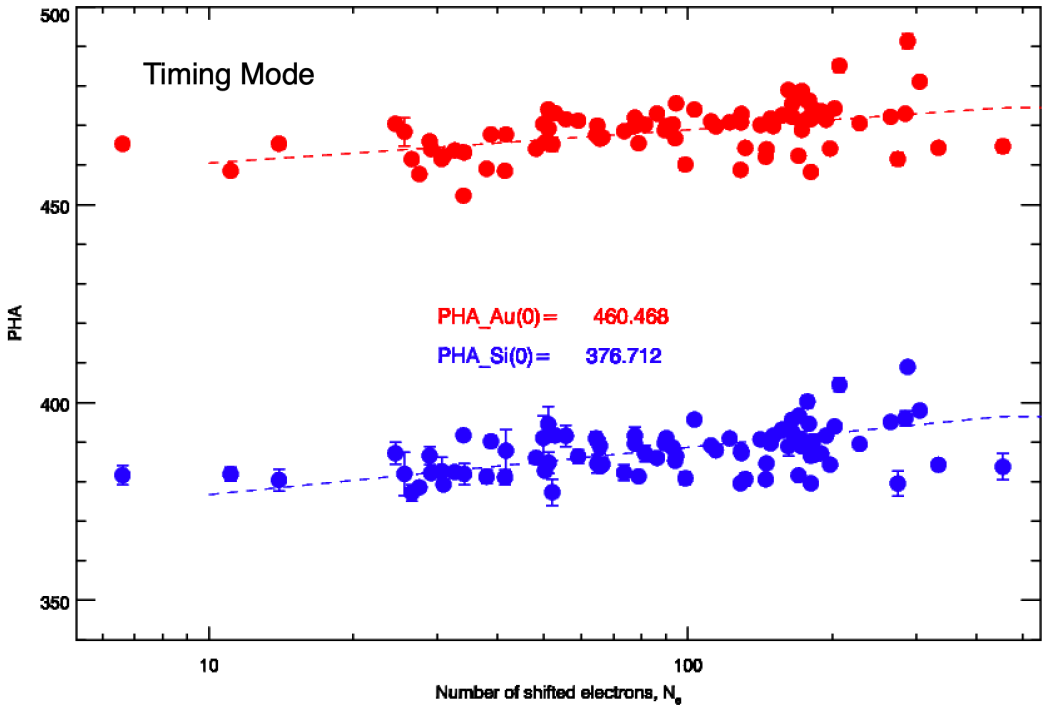


Figure 1: Number of shifted electrons vs PHA channel for EPIC-PN timing mode observations (From Migliari & Smith 2019)

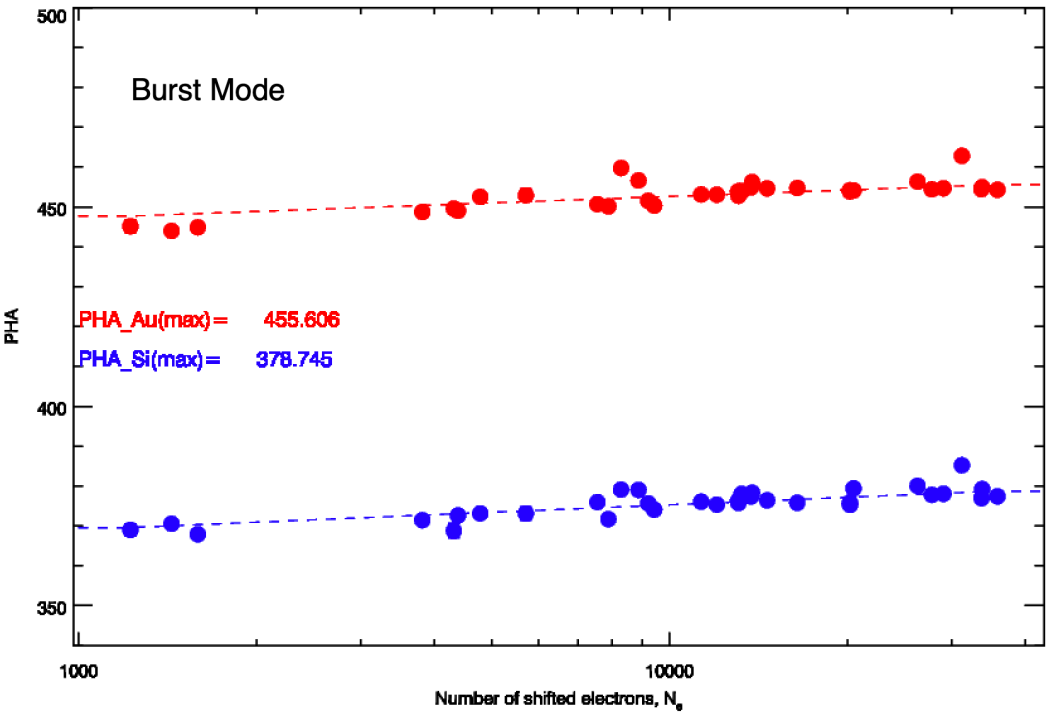


Figure 2: Number of shifted electrons vs PHA channel for EPIC-PN burst mode observations (From Migliari et al. 2020)

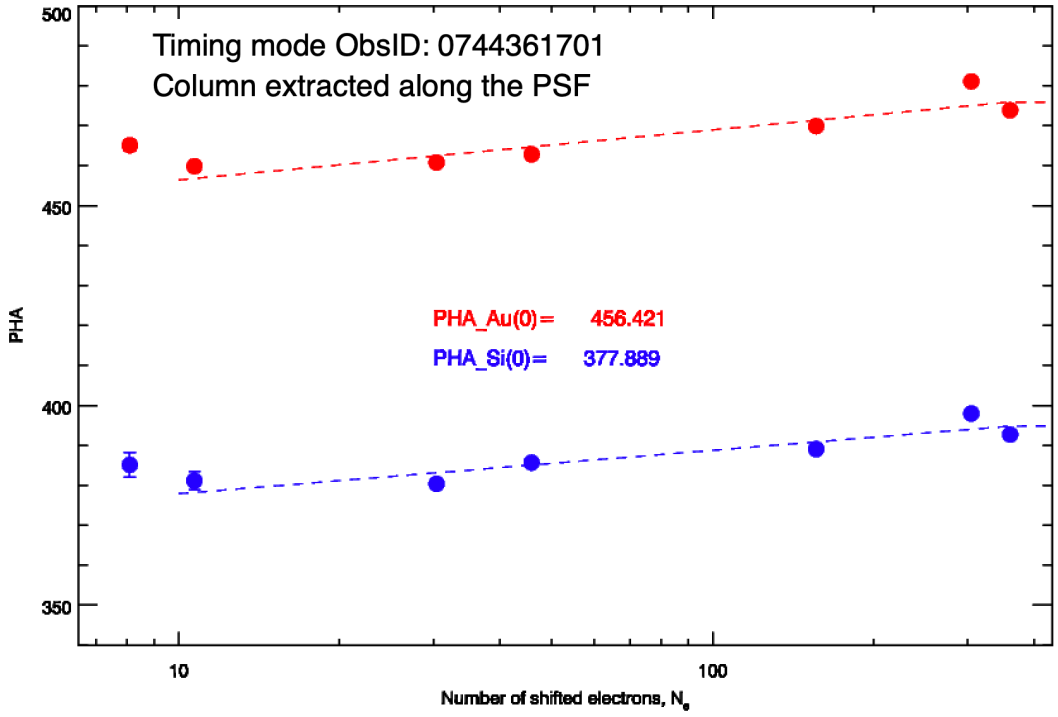


Figure 3: Ne-PHA correlations of the Timing Mode observation 0744361701, where each point comes from the extraction of a spectrum from one column.

## 3 Assessing Calibration Along the PSF Profile

First of all, we assessed the RDPHA effect along the PSF of single observations. By comparing the RDPHA effects observed using the 21-column averaged regions with that of single column with different count-rates along the PSF of a single observations, we wanted to check if effects due to the column number alone (and not the difference in observation set-ups) could be visible. We took the brightest column as central PSF reference and extracted the spectra of every other column to sample the PSF shape, symmetrically accumulating columns at the same distance to the brightest to improve statistics. In Fig. 3, we show the Ne-PHA plot of the Timing Mode observation 0744361701, where each point comes from the extraction of a spectrum from one column. Comparing with the current 21-column averaged RDPHA effect (Fig. 1), we see that both slope and normalization are consistent, *modulo* statistical scatter. Given that a similar rate-dependent effect is observed within a single observation and averaging a large number of observations with different setups and observational conditions, e.g. like temperature of the CCD, etc., this indicates that such parameters are not dominating the RDPHA effect.

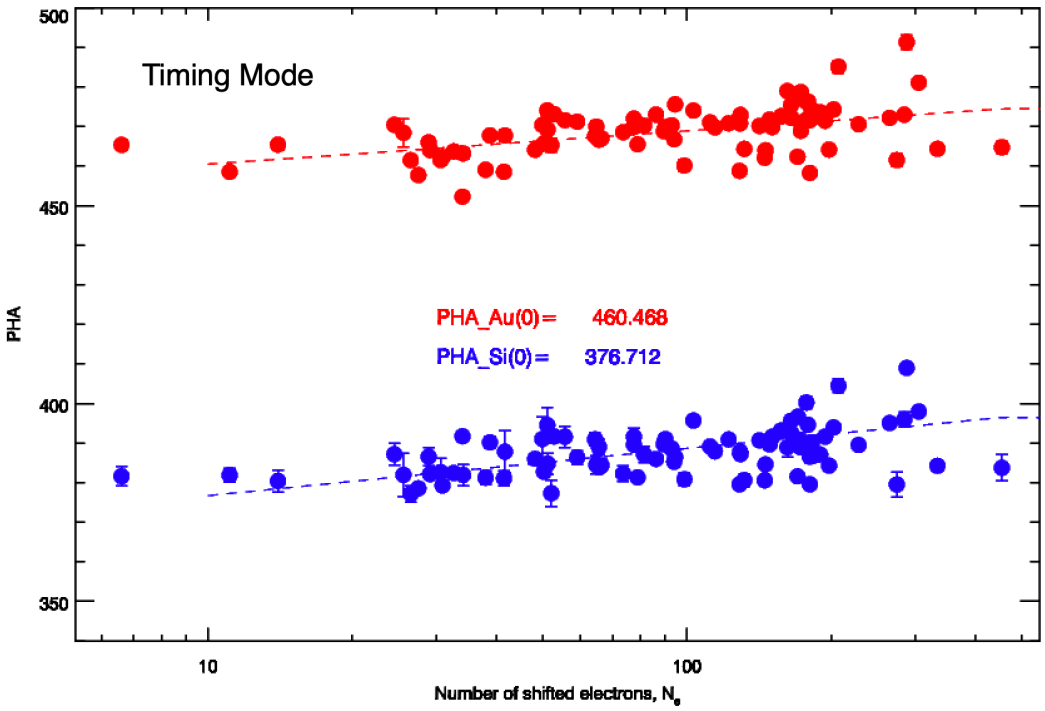


Figure 4: Ne-PHA correlation of the brightest column for the whole observation sample.

# 4 Analysing Brightest Column Only

We analysed only the brightest column for the whole observation sample. The resulting Ne-PHA plot is shown in Fig. 4 for Timing mode observations. The analysis results using only the brightest columns are unsurprisingly very similar (*modulo* statistics) to those using the 21-column averaged centered in the brightest column. Indeed, because of the steepness of the PSF, most of the photons fall in the brightest column.

## 4.1 Calibration as a Function of Column Number

As a step forward, we filtered the results by column number. In Fig. 5, we show the plot of columns 36, 37 and 38, i.e., the most common brightest columns in the sample, for which we have the best statistics. We observe a similar trend for 36 and 37, that is more or less constant, albeit with different normalization in energy shifts, 36<37, but a different trend for the PHA of 38 that increases as a function of Ne. If we compare with Fig. 4 (one column overall), it looks like the normalization difference between 36 and 37, their scatter at high Ne, and the increasing correlation of 38, is what creates an overall increasing correlation, even though only column 38 seems to have a significant positive slope.

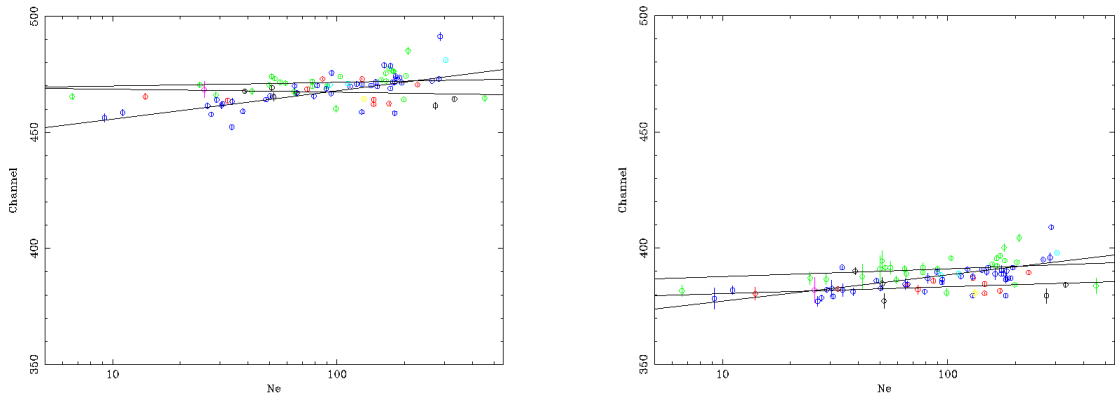


Figure 5: Ne-PHA correlations for columns 35 (open circles), 36, (filled circles), 37, (filled triangles), 38 (filled stars), 39 (open stars), 41 (open diamonds), 46 (open squares). Left panel: power law fit of the columns 36 (Red), 37 (Green) and 38 (Blue) for the Au M-edge. Right panel: power law fit of columns 36 (Red), 37 (Green) and 38 (Blue) for the Si K-edge.

### 4.2 Calibration as a Function of Filter

Using the brightest column selection as in Section 4, we plot in Fig. 6 the observations taken with different filters: Thin, Medium, Thick. The goal is to see if different filters have a significant effect on the RDPHA calibration. We note that most of the observations are taken either with Thin or Medium filters. Thick filters are more rare. There is no obvious difference between the three samples, apart that the slope seems to increase from Thick (consistent with being constant) to Medium to Thin. However, the statistics is too low, especially with Thick Filter observations, to draw any solid conclusion.

### 4.3 Calibration as a Function of Column Number and Filter

In Fig. 7 and 8, we show the fits of the three columns, 36, 37 and 38, for each filter. Due to statistics, fits have been performed only for observations with thin and medium filters. We can see that the trends are similar, so within the same column the fit does not significantly change if the filter changes from Thin to Medium.

## 5 Analysing Column by Column: Accumulating by Column Number

Finally, we extracted approximately 9000 spectra out of the timing mode observations and 1900 out of the burst-mode observations, that is, column-by-column spectra of each observation of the sample for 20 columns along the PSF. Of course, we expect the statistics to play an important role: the farther from the brightest column, the less counts and less

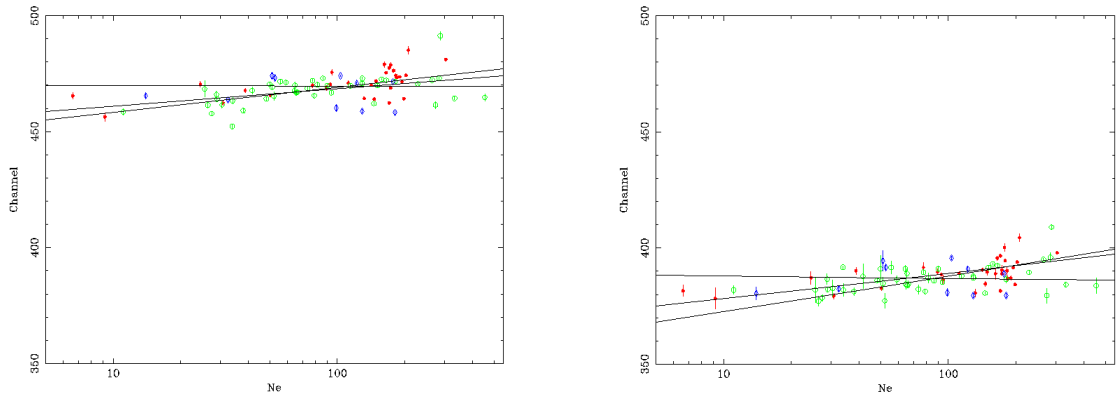


Figure 6: Fits of columns 35 to 46 (markers as in Fig 5) filtered by filter thickness: Thick (Blue), Medium (Green) Thin (Red). Left panel: power law fit for the Au M-edge. Right panel: power law fit for the Si K-edge. The lines show the trends of Thin and Medium filter observations are very similar. The slopes decreases from Thin to Medium to Thick filters, however, no conclusion can be drawn from the Thick filter sample because of statistics.

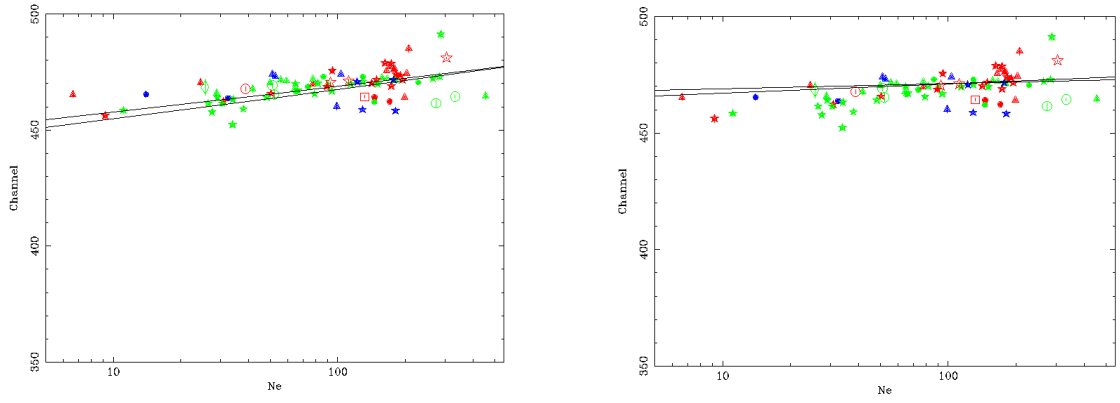


Figure 7: Correlations of the Au M-edge for the column 35 to 46 (markers as in Fig 5) grouped by filter: Thin (Red), Medium (Green), Thick (Blue). Left Panel: The solid lines represents the fits of column 38-only for thin and medium filters. Right Panel: The solid lines represents the fits of column 37-only for thin and medium filters.

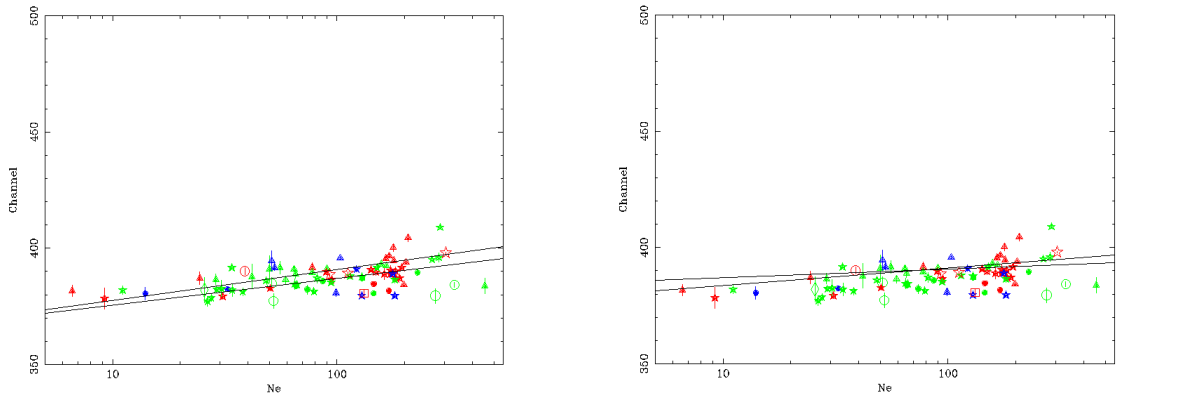


Figure 8: Correlations of the Si K-edge for the column 35 to 46 (markers as in Fig 5) grouped by filter: Thin (Red), Medium (Green), Thick (Blue). Left Panel: The solid lines represents the fits of column 38-only for thin and medium filters. Right Panel: The solid lines represents the fits of column 37-only for thin and medium filters.

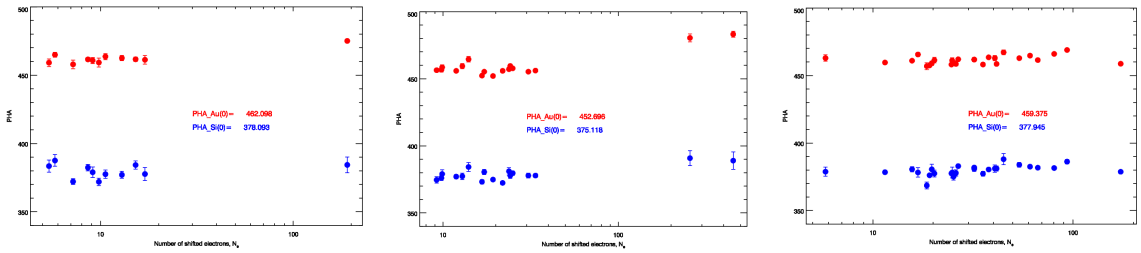


Figure 9: Burst Mode correlations for column 30, 33 and 34 (left to right) showing that additional two uncertainties are introduced: lower statistics, and a different reference baseline for each column.

statistics to fit the energy edges.

Fig.9 show examples of burst-mode correlations for two of the edges. It is clear how the column by column analysis introduces two uncertainties. The first is of course the diminished statistics in terms of number of sampled points. The second is the uncertainty on the reference baseline for the edge energy shifts. The latter is an additional free parameter with respect to our previous analysis averaging columns, that needs to be estimated. For example, in the burst mode observations (the same reasoning applies to timing mode), if we take the PHA channel of reference of the Si peak, the reference channel depends on the data fit of each column and it is in the range 371-383, that means a range of 12 channels. For comparison, the current CCF correction maximum shift for burst mode is approximately 9-10 channels, comparable to the error introduced by the reference parameter uncertainty.

## 6 Results and (No) Updates

We conclude that the column by column analysis with the current samples cannot be used to improve significantly, if at all, the calibration algorithms used in the current CCFs for fast mode EPIC pn observations.

## 7 References


Guainazzi M. 2013, XMM-CCF-REL-0299,
"Coefficients of the Rate-Dependent PHA (RDPHA) correction based on the derivative spectra"
(available at: `http://xmm2.esac.esa.int/docs/documents/CAL-SRN-0299-1-1.ps.gz`)

Guainazzi M., et al. 2014b, XMM-SOC-CAL-TN-0083,
"Spectral calibration accuracy in EPIC-pn fast modes"
(available at: `http://xmm2.esac.esa.int/docs/documents/CAL-TN-0083.pdf`)

Guainazzi M. 2014a, XMM-CCF-REL-0312,
"RDPHA calibration in the Fe line regime for EPIC-pn Timing Mode"
available at: `http://xmm2.esac.esa.int/docs/documents/CAL-SRN-0312-1-1.pdf`

Migliari s., Smith M., 2019, Rate and Energy-Dependent PHA Correction for EPIC-pn Timing Mode, available at: `https://xmmweb.esac.esa.int/docs/documents/CAL-SRN-0369-0-0.pdf`

Migliari S, Valtchanov I., Smith M., Schartel N., 2020, EPIC-pn Energy Scale: Rate and Energy-Dependent PHA Correction for Burst Mode and Long-Term CTI Updates, available at: `https://xmmweb.esac.esa.int/docs/documents/CAL-SRN-0376-1-1.pdf`