



Universiteit
Leiden
The Netherlands

Hunting dark matter with X-rays

Franse, J.

Citation

Franse, J. (2016, December 20). *Hunting dark matter with X-rays. Casimir PhD Series*. Retrieved from <https://hdl.handle.net/1887/45082>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/45082>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/45082> holds various files of this Leiden University dissertation.

Author: Franse, J.

Title: Hunting dark matter with X-rays

Issue Date: 2016-12-20

5 | CORRELATION METHOD FOR WEAK LINE SEARCHES

5.1 Introduction

Spectral modeling of X-ray observations is a well-established field with well-established techniques and tools. Current X-ray detectors have energy resolutions of order 100 eV, however, which can cause fitting degeneracies between various features, such as between various emission lines, or between the continuum and emission lines. This paper introduces a proof-of-concept to distinguish spectral components by their scaling characteristics between objects and environments.

A particular example, and the nominal science case on which we focus the development, is the possibility that the Dark Matter may decay with a monochromatic X-ray emission line as a result. This kind of Dark Matter has been well described, and covers a wide range of particles. For example Essig et al. (2013) describes a number of these along with current bounds, and more can be found in e.g. Abazajian et al. (2001); El Aisati et al. (2014); Frandsen et al. (2014); Iakubovskiy (2014).

Currently, a candidate signal for decaying Dark Matter is being scrutinized thoroughly. It was discovered in spectra of galaxies and clusters of galaxies (Boyarsky et al., 2014a; Bulbul et al., 2014a), and subsequently studied in many other objects (e.g., Boyarsky et al., 2015; Jeltema & Profumo, 2015; Urban et al., 2015; Malyshev et al., 2014; Ruchayskiy et al., 2015; Jeltema & Profumo, 2016; Bulbul et al., 2016a; Franse et al., 2016, and references therein), although the final verdict remains due. The main approach to establishing whether this interpretation is correct, is to compare the strength of this signal between objects, and seeing that the signal scales as the total Dark Matter mass inside the field-of-view of the telescope over the distance to the object squared. However, the signal strength has to be measured from spectral fits. In single objects, the low statistical significance means that the flux has a large uncertainty and may be susceptible to spectral model degeneracies. The statistics may be increased by stacking a number of objects together and fitting to the resulting spectrum (Bulbul et al., 2014a). This has a number of advantages, among which that instrumental effects smear out among the spectrum (as long as the stacked sample has a large enough spread in redshifts). However, the spectrum will be a superposition of varying environments, so that it becomes more difficult to relate

model parameters to each other by simple physical relationships. In addition, a detection in a stack precludes the possibility of comparing signal strength between the members of the stack, although comparing the signal strength between different stacks is still possible.

Here we present a method to combine the statistical power of a stacked spectrum with the ability to constrain the origin of a signal based on the scaling between the single objects. Even though it would be impossible to measure the signal strength from each object individually, we combine the limited number of X-ray events and the knowledge of the mass and redshift of each object into a single estimator that reflects not only the strength of a Dark Matter decay signal, but also whether or not its origin is actually Dark Matter decay.

Scaling up this method to use as much archival data as possible, and introducing more advanced simulations and minimization approaches along with the required computing power, it is fully expected that this method can provide competitive limits on the Dark Matter decay. With the loss of the *Hitomi* mission, our approach may be ideally suited to testing the Dark Matter interpretation of the 3.5 keV signal in the future.

As a bonus, the method does not require background subtraction and can be made insensitive to detector and mass calibration systematic effects under certain circumstances. This method is also easily generalized and applied to any object or type of signal.

In Section 5.2 we will lay the mathematical groundwork for performing the correlation, while Section 5.3 details how to use this formalism to perform a search for a Dark Matter decay signal. We describe the data used and the practical implementation of the method in Section 5.4 and discuss the error estimation in Section 5.5. Section 5.6 contains the central results of our work, where a set of simulated signals is added to the real dataset, and we report the success rate of recovering the simulated signals using our method. Our conclusions are summarized in Section 5.7, and in Section 5.8 we discuss the performance of our method and possible improvements. The appendices contain derivations of various equations presented throughout the main text, the results of a few null tests, and a note on the practical difference between a correlation and a weighted average and instructions on how to incorporate weighting in the correlation.

5.2 Formalism

The formalism is adapted from Refregier et al. (1997), where the correlation w can be expressed as

$$w = \frac{\langle NI \rangle}{\langle N \rangle \langle I \rangle} - 1. \quad (5.1)$$

Here, quantities N and I are being correlated over some field that is divided into cells, and the averages indicated by $\langle \dots \rangle$ are performed over these cells. In this work we are interested in possible Dark Matter signals, and therefore correlate the detected X-rays (represented as I) with a catalog of galaxy groups that serve as a proxy for Dark Matter mass overdensities. These are represented by N , but N can contain any (combination of) physical characteristics of the galaxy groups. We call this the correlator throughout this work. Of course, N can represent any other kind of object if so desired. The basic

expressions for these are simply

$$\langle NI \rangle = \frac{1}{N_C} \sum_{i=1}^{N_C} I_i N_i \quad (5.2)$$

$$\langle I \rangle = \frac{1}{N_C} \sum_{i=1}^{N_C} I_i \quad (5.3)$$

$$\langle N \rangle = \frac{1}{N_C} \sum_{i=1}^{N_C} N_i. \quad (5.4)$$

with N_C the number of cells in the field. If the two quantities are uncorrelated, their averages become independent so that $\langle NI \rangle = \langle N \rangle \langle I \rangle$ and $w = 0$. For a correlation, w becomes larger than 0, and smaller than 0 when an anti-correlation exists. Below, we will extend this formalism in various ways, modifying the above expressions

As noted, the correlator N can be chosen to represent any kind of physical property or object. In this work we are interested in galaxy groups as tracers of the Dark Matter density field, and will use various (combinations of) galaxy group properties as the correlator.

The central methodology in this work is to perform the correlation in bins of spectral energy and therefore in redshift-space. Essentially, it is possible to distinguish the various components of the raw X-ray data by their spectral behaviour relative to the redshifts of the physical objects that are being correlated with.

Since our interest is in the X-rays emitted by the galaxy groups, we work in the rest-frames of each of those groups. When computing $\langle NI \rangle$, whenever a cell i contains a galaxy group k , the entire X-ray spectrum of that cell I_i is blueshifted back to the group's redshift z_k , represented as $\zeta(I_i, z_k)$ or ζ_{ik} for short. It is possible for a cell to contain multiple groups at different redshifts. In that case, this blueshifting is performed once for each group. The expression becomes

$$\langle NI \rangle = \frac{1}{N_C} \sum_{i=1}^{N_C} \sum_{k=1}^{N_g(i)} \zeta(I_i, z_k) N_k. \quad (5.5)$$

with $N_g(i)$ the number of groups in cell i .

The X-rays coming from the galaxy groups will indeed be redshifted (which is what the ζ_{ik} corrects for), but all backgrounds and foregrounds that are present in I_i will also be shifted by ζ_{ik} . This shift is random with respect to the fore-/backgrounds, a fact that can be used to separate the galaxy group emission from those backgrounds. In order to do that correctly, $\langle I \rangle$ should in a sense describe the average expected result of performing these shifts on the backgrounds. In other words, $\langle I \rangle$ is the average X-ray spectrum of the entire field, averaged in turn over all the shifts performed while calculating $\langle NI \rangle$.

In addition, the correlator N_k used in $\langle NI \rangle$ is effectively giving random weights to the backgrounds in each cell that contains a group. Analogously to the blueshifting described above, therefore, the computation of $\langle I \rangle$ must also include the distribution of the values of the correlator N_k in the group catalog. Specifically, during the calculation of $\langle NI \rangle$, the redshifts and correlator values come in pairs (that is to say, the redshift of

one group always goes along with the correlator value of the same group), so all in all, to calculate $\langle I \rangle$, the distribution of (z_k, N_k) must be taken into account. This gives

$$\langle I \rangle = \frac{1}{N_C} \sum_{i=1}^{N_C} \frac{1}{\sum_l^{N_g} N_l} \sum_{k=1}^{N_g} \zeta(I_i, z_k) N_k \quad (5.6)$$

where N_g is the total number of groups in the catalog, the factor $\sum_l^{N_g} N_l$ accounts for the proper normalization.

Note that without the blueshifting, it would not be required to write $\langle I \rangle$ explicitly in terms of (sums involving) N_k , as the sum over the groups k would be independent of the sum over the cells i , and the expression would be mathematically equivalent to Equation 5.3. Due to the inclusion of ζ_{ik} , the two sums mix and we get Equation 5.6.

5.2.1 (De)composition of the Correlation

If the true X-ray emission is known in terms of its components, we can predict what the correlation should be for various choices of the correlator. We will actually use $w\langle I \rangle$ as this turns out to be the most useful choice. Below we give a few examples of this, with derivations given in Appendix 5.A.1. With the expressions below in hand, we shall see that it is possible to invert the process and obtain an estimate of the true X-ray emission per component.

If the true emission is given by an intrinsic luminosity that is the same for each group (L_{grp}), the observed flux for each cell would be $I_i = L_{grp}/(4\pi D_{lum,i}^2)$. Using as correlator then $1/(4\pi D_{lum}^2)$, the correlation will result in

$$w\langle I \rangle = L_{grp} \frac{\sum_k 1/D_k^4}{\sum_k 1/D_k^2} \left(1 - \frac{1}{N_C} \right) \quad (5.7)$$

where the sums run over all groups, and for convenience we write $4\pi D_{lum,k}^2 = D_k^2$. Given the catalog and the known distances, it is possible to solve for L_{grp} .

A more general case would be when the true emission is given by an arbitrary number of components, each of which scales with a different power β of the group mass M like so

$$I_i = \sum_{\beta} L_{\beta} M_i^{\beta} / D_i^2. \quad (5.8)$$

If the correlator used is then M^{γ}/D^2 (for an arbitrary γ), the resulting correlation is

$$\begin{aligned} w_{\gamma}\langle I_{\gamma} \rangle &= \sum_{\beta} L_{\beta} \frac{\sum M^{\beta+\gamma}/D^4}{\sum M^{\gamma}/D^2} \left(1 - \frac{1}{N_C} \right) \\ &\equiv \sum_{\beta} L_{\beta} A_{\beta\gamma} \end{aligned} \quad (5.9)$$

where the sums run over all objects k in the catalog and the k subscripts on all powers of M and D have been omitted for readability.

In this case, the masses and distances of the groups are known, the correlation $w_{\gamma}\langle I_{\gamma} \rangle$ is measured, but there are an arbitrary number of unknown L_{β} 's (if at least the β 's are

known, otherwise there are twice as many unknowns). It is still possible to solve for these unknowns if the correlation is repeated for at least as many different values of γ as there are unknowns. In that case, Equation 5.9 becomes a system of equations that can be solved analytically if all the components can indeed be accurately described by Equation 5.7 and the correlations is noise-free.

Of course, this is not the most general case imaginable, but the principle holds for any model as long as it scales with known (for some value of known) properties of the objects in the catalog. This principle also holds in the presence of foregrounds and backgrounds, as these correlate out and do not contribute to $w\langle I \rangle$ on average.

This is the central principle that we will use throughout this paper to search for a Dark Matter decay signal. Of course the true emission of the galaxy groups may not be perfectly described like in Equation 5.7 with a minimal set of components, nor will the measurement of the correlation be noise free, so Section 5.3 will detail how to put the principle into practice.

5.3 Methodology

In this section we will describe how to use the formalism from Section 5.2 to search for Dark Matter decay-like signals in practice.

After choosing a suitable field, and having obtained the X-ray data and galaxy group catalog appropriate for that field, a correlation can be computed given some correlator. The choice of correlator should be related to the signal that is being looked for. In the case of Dark Matter decay, the expected behaviour is M^γ/D^2 with $\gamma = 1$, so that is a good starting point.

In principle we could then compute some more correlations, for varying values of γ for example. Then we could assume some model for the intrinsic correlated emission (like Equation 5.7), and continue to solve the system of equations 5.9 using this set of correlations (having different correlators each), and a choice of β 's. Regarding the search for Dark Matter decay, then at least one of those β 's should be 1, and the objective is to find a significant line-like contribution of the $\beta = 1$ component. This approach will be discussed in Section 5.8.1, but turns out to be not very efficient to solve directly.

As we shall see in Section 5.6, even quite weak signals (in terms of signal-to-noise) can sometimes be seen by eye in the correlation spectrum (the value of w as a function of restframe energy). Therefore we shall use a quite naive estimator which will prove to work remarkably well (also in Section 5.6) and takes advantage of the fact that we are looking for a spectral line feature (although a generalized methodology will be described later).

This simple estimator to test for the presence and strength of a decay line is to take for a particular energy bin the difference between that bin's w value, minus the average of the neighboring bins, by way of continuum subtraction. We will refer to this estimator as Δw . In fact, to make use of Equation 5.9, we should actually use $\Delta(w_\gamma \langle I_\gamma \rangle)$, which we will write as Δ_γ for short.

Some arbitrariness is present in the choice on the width of the energy bin. Considering the spectral resolution of the instrument, also the choice of which bins to use for the 'signal region' and which bins to use for the continuum estimate is somewhat arbitrary. If

we call the energy bin under consideration the ‘central bin’; the central bin together with potential extra bins that are considered to contain signal the ‘signal bins’; the bins that are used to estimate the continuum ‘side bins’; and potentially some bins between the signal bins and the side bins that may be ignored as ‘skipped bins’; then the estimator can be written as

$$\begin{aligned} \Delta_\gamma &= \Delta(w_\gamma \langle I_\gamma \rangle) \\ &= \sum_b^{\text{signal bins}} w_\gamma \langle I_\gamma \rangle(E_b) - \frac{N_{\text{signal bins}}}{N_{\text{side bins}}} \sum_s^{\text{side bins}} w_\gamma \langle I_\gamma \rangle(E_s) \end{aligned} \quad (5.10)$$

We will indicate the choice of which bins are used schematically for example like so

$$sskxXxkss$$

representing the bins in energy space, with X the central bin, x the remaining signal bins, k skipped bins and s the side bins used for continuum subtraction. In words, we will say this scheme uses 1 extra signal bin, 1 skipped bin and 2 side bins (since the scheme is symmetric around the central bin, we refer only to one side).

A significant non-zero and positive value for Δ_γ is an indicator for a line-like emission feature. Below, we will describe how to use this estimator Δ_γ to search for Dark Matter decay-like signals. Four different approaches will be described, in order of increasing complexity. The efficacy of each will be tested in Section 5.6.

- It is already possible to use $\Delta_{\gamma=1}$ to obtain $L_{\beta=1}$ if one assumes that the excess is caused by a Dark Matter decay-like signal, and solves Equation 5.9 like so; $L_1 = \Delta_1/A_{11}$. However, this carries no intrinsic information about the origin of the excess.
- One way to test whether a Dark Matter decay may be the cause of non-zero and positive Δ_γ could be to take the set of computed correlations $\{\Delta_\gamma\}$ (for a set of varying γ 's), and solve the resulting system of equations for $L_{\beta=1}$, assuming a single Dark Matter decay component.

$$\begin{aligned} \Delta_1 &= L_1 A_{11} \\ \Delta_2 &= L_1 A_{12} \\ &\vdots \\ \Delta_n &= L_1 A_{1n} \end{aligned} \quad (5.11)$$

Because real data contains noise, and because it is not *a priori* known whether a single $\beta = 1$ component is the correct interpretation, this system of equations probably does not have an perfect analytical solution. Instead, one should solve this system using some fitting or optimization procedure. This difference between solving the system of equations using the the set of γ 's, and only using $\gamma = 1$, is that one may judge whether the Dark Matter decay is a consistent interpretation (ie., that assuming a single $\beta = 1$ component is correct) by requiring that the residuals from the fitting are small.

- Instead of just requiring small residuals from a given assumption (single $\beta = 1$ component), it is more convincing to let the optimization procedure find the best-fit β for the obtained $\{\Delta_\gamma\}$. So, still only a single component is assumed, but the model now is two-parameter (β and L_β) and the system of equations to be solved is

$$\begin{aligned}\Delta_1 &= L_\beta A_{\beta 1} \\ \Delta_2 &= L_\beta A_{\beta 2} \\ &\vdots \\ \Delta_n &= L_\beta A_{\beta n}\end{aligned}\tag{5.12}$$

- Lastly, it is of course possible to include more than a single component to describe $\{\Delta_\gamma\}$. The value of β for each component can be picked *a priori*, or be taken as a free parameter. The number of different correlators used to solve this system should be at least equal to the number of free parameters.

Having decided upon which of the above approaches to employ, one obtains the ‘decomposition’ of the correlation (best-fit solution to Equation 5.9), one for each energy bin. An example of such a ‘spectrum’ is given in Figure 5.1, for a simulated signal (see Section 5.6). Such a figure can then be used to determine whether a Dark Matter decay signal is present in the data. The requirements for that are; 1) a significant positive value for the $L_{\beta=1}$ component, 2) a best-fit β close to 1 (if β was a free parameter), 3) a high goodness-of-fit and 4) a line-like shape in the ‘decomposition spectrum’ (with regards to the instrument resolution).

5.4 Data and Implementation

To test and explore this method, we utilize the publicly available data and catalogs of the COSMOS survey (Scoville et al., 2007). This field of roughly two square degrees has been covered in most wavelength regimes, among with deep *XMM-Newton* exposures, and multiple catalogs of Dark Matter tracers are available. The COSMOS field is relatively empty of very bright objects in any wavelength and galactic foregrounds are low and reasonably homogeneous.

The main Dark Matter tracer population that we use is the X-ray selected galaxy group catalog by George et al. (2011). It contains 183 groups between redshifts of 0 and 1, which have been detected by a spatial wavelet analysis on the X-ray mosaic and confirmed by optical galaxy overdensities. This we refer to as the G11-sample. Another catalog is also available, which is the optically selected group catalog by Knobel et al. (2012). It enables us to check for systematics and in the case of a possible line candidate test the robustness of that detection. This catalog uses spectroscopic redshift measurements and a friend-of-friends algorithm to detect galaxy groups. It contains around 1500 groups of which almost 200 have 5 or more spectroscopic members. We shall refer to this catalog as the ‘20k’ sample, since it is based on the zCOSMOS spectroscopic survey (Lilly et al., 2007) of nearly 20,000 redshifts. The zCOSMOS field is slightly smaller than the full COSMOS field.

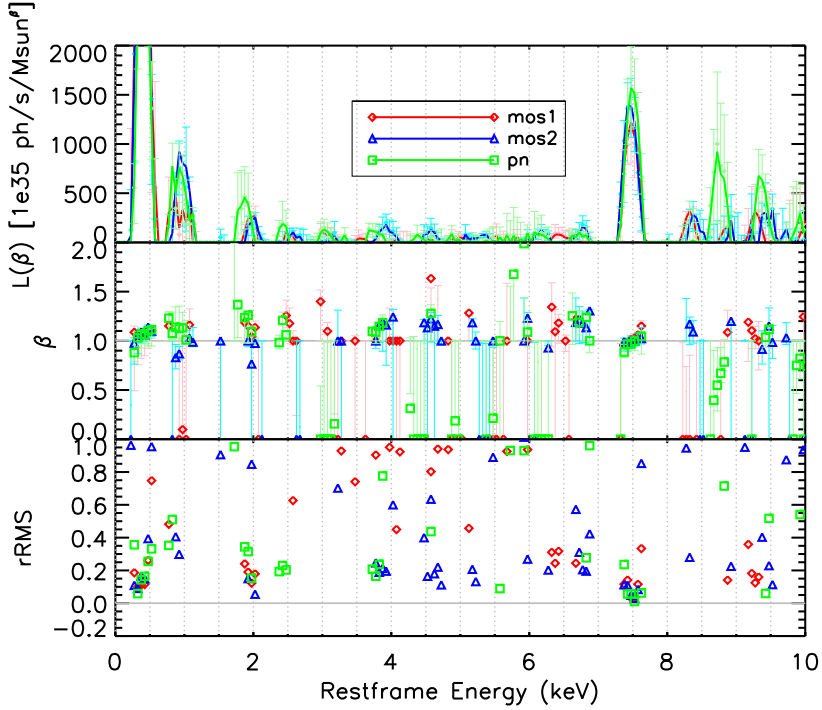


Figure 5.1: How to determine if there's a Dark Matter decay line in the data? Example plot of simulated data showing the last step in the methodology (see text, Sections 5.2 and 5.3). All four requirements are met: 1) a significant positive measured value for L_β (top panel) reflecting the intrinsic luminosity of the correlated excess emission; 2) a best-fit value of β consistent with 1 (middle panel), reflecting that the correlated excess emission scales with mass to the power $\beta = 1$, as is required of Dark Matter decay; 3) a sufficiently low relative RMS deviation (bottom panel) for this combination of best-fit values of L_β and β ; 4) a line-like profile in the top panel, since Dark Matter decay is a monochromatic signal. Therefore it can be concluded that a Dark Matter decay signal is present at 7.5 keV in this simulated data. This 'decomposition spectrum' is for the real COSMOS data (Section 5.4) and a simulated injected signal (Section 5.6). The simulated signal is indeed at 7.5 keV and is a very strong signal in order to demonstrate the method. Note that the 'significant' peak below 0.5 keV is an artifact of the estimator Δ_γ due to edge effects (the edge of the spectrum). Different colors and symbols indicate data from the three different cameras (mos1, mos2 and pn instruments represented in red and by diamonds, in blue and by triangles, and in green and by squares respectively). The error bars in the top and middle panels indicate 68% bootstrap intervals (Section 5.5).

We use all available *XMM-Newton* data of the COSMOS field (Hasinger et al., 2006) that is available in the science archives. The standard *SAS* pipelines were used to reduce the data and filter the data for flares using the `espfilt` procedure. Additional flare filtering was performed based on in-FoV and out-FoV high energy count rates (De Luca & Molendi, 2004), with observations having a ratio over 1.3 being excluded. The whole COSMOS field was then divided into a square grid with 75'' spacing, resulting in 5256 cells. This cell size was chosen to be somewhat larger than the *XMM-Newton* PSF. Using the *SAS* procedures `rmfgen` and `arfgen` the effective area curves for each exposure of each cell were generated. The different exposures of each cell were then combined by weighting by the cleaned exposure time, effective area and usable solid angle (not all pixels are usable). The data from the three instruments (MOS1, MOS2 and PN) and the correlation products were kept separate throughout this work unless otherwise noted. The main computational time sink in the entire method is the generation of `rmf` and `arf` files for each exposure of each cell. This is the reason for using a flexible and multi-functional square grid.

Obs ID	Exposure (ks)			Obs ID	Exposure (ks)		
	MOS1	MOS2	PN		MOS1	MOS2	PN
0203360101	28.9	29.0	24.3	0302350701	17.8	18.2	14.7
0203360201	15.1	15.1	10.6	0302350801	18.9	18.6	15.1
0203360301	30.1	30.3	24.2	0302350901	7.5	8.1	3.4
0203360401	27.2	26.9	20.2	0302351001	36.1	37.7	28.2
0203360501	26.0	26.3	19.9	0302351101	15.1	15.5	11.5
0203360601	23.1	22.1	18.4	0302351201	15.0	14.8	12.1
0203360701	32.1	33.0	27.0	0302351301	18.6	18.5	15.3
0203360801			14.1	0302351401	16.2	17.2	8.7
0203360901	20.7	21.4	16.2	0302351501	14.1	13.9	9.5
0203361001	14.0	13.6	10.6	0302351601	28.9	29.4	21.9
0203361101	19.6	20.5	7.9	0302351701	19.0	18.3	14.3
0203361201	25.3	25.0	21.6	0302351801	17.3	17.4	14.1
0203361301	25.1	25.0	21.6	0302351901	11.4	10.5	7.5
0203361401	30.5	29.8	26.1	0302352001	4.8	4.8	3.5
0203361501	23.2	23.4	13.3	0302352201	7.6	7.8	3.9
0203361701	30.0	29.8	25.1	0302352301	4.4	5.4	3.0
0203361801	26.4	26.4	22.9	0302352401	18.8	18.4	15.0
0203361901	22.9	23.2	19.5	0302352501	22.5	22.5	19.1
0203362001	8.1	7.5	4.6	0302353001	2.7	2.7	2.1
0203362101	59.0	58.2	51.2	0302353101	17.5	17.8	13.6
0203362201	28.3	29.2	13.4	0302353201	11.3	11.0	7.6
0302350101	12.9	14.0	11.5	0302353301	12.2	12.4	8.7
0302350201	13.6	14.4	8.9	0302353401	8.8	9.5	5.1
0302350401	7.3	7.7		0501170101	32.1	31.5	26.5
0302350501	18.4	18.6	15.1	0501170201	29.8	30.2	22.0
0302350601	17.0	16.4					
Total					99.3	99.9	75.5

Table 5.1: Cleaned exposure times of the used *XMM-COSMOS* observations. Entries without exposure listed were discarded for having a Fin-Fout ratio too high (see text).

During development, it was found that the outer regions of the full COSMOS field are not usable for the correlation. As mentioned, the 20k-sample of galaxy groups was only based on the zCOSMOS field. However, the G11-sample suffers from the reduced sensitivity at the outsides of the COSMOS field, caused by the vignetting effect of *XMM-Newton*. Even in the mosaic of all *XMM-Newton* exposures, the outermost regions of the COSMOS field are only covered by the edges of the detector, which have lower effective area. Including the full COSMOS field when correlating using the G11 sample caused residual systematic effects for that reason. Throughout the rest of this work, we shall consider only the zCOSMOS field, whose boundaries coincide well with the region of lower sensitivity of the COSMOS X-ray mosaic.

For solving the system of equations described in Section 5.3, we employ the IDL minimization routine `TNMIN` with a relative least-squares implementation. This procedure is general enough to be able to solve the multi-component decompositions and allows limiting the L_β parameters to only positive values.

5.5 Error Estimation

Although it is possible to simply propagate in the usual way the Poisson noise through the entire calculations of $\langle NI \rangle$, $\langle I \rangle$, and w or $w\langle I \rangle$, it is likely that upon scaling up this method the Poissonian noise becomes subdominant to for example chance alignments of sources and features in the astrophysical or instrumental background. More importantly maybe, is that the w_γ (and Δ_γ) and their errors will be correlated between different values for γ . Also the energy bins within a correlation spectrum will be correlated up to a point due to the blueshifting procedure (and due to the instrumental spectral resolution).

For the purposes then of the (estimators) of the decomposition, the propagated Poissonian noise is insufficient. The most efficient way to emulate all possible sources of uncertainty at the same time is to perform a bootstrapping resampling of the dataset. We split the zCOSMOS field into 9 evenly sized subfields. For each bootstrap iteration then, a new realization of the zCOSMOS field is obtained by drawing randomly with replacement 9 subfields from the original 9. The entire algorithm is then applied to this new ‘survey’.

However, by resampling the field, the group catalog for each realisation will also be different. The values of w and $w\langle I \rangle$ depend on the group catalog, so that the spread in values of w from the bootstrap does not represent an uncertainty in the ‘true’ value of w since there is no one true value for w among the bootstrap realizations. The decomposition in terms of L_β does not depend on the group catalog in principle, as L_β is an intrinsic properties of the physical process. However, a resampling of the field can seriously affect the significance of any correlated signal, if the resampled catalog contains many fewer, less massive and/or higher redshift groups for example. Therefore, during bootstrapping we only accept resamplings of the field if the resulting group catalog is within 0.02 dex (about 5%) of the original group catalog in terms of $\sum_k M_k/D_k^2$, which serves as a proxy for the expected total received signal photons.

This constrained bootstrap provides an uncertainty interval for the decomposition, and is now also suited for obtaining uncertainties on w and $w\langle I \rangle$. The requirement threshold on the similarity of bootstrapped group catalogs is somewhat arbitrary. Larger (more lenient) thresholds will broaden the uncertainty interval (overestimating the error), while

stricter requirements risk limiting the bootstrap process so it is no longer representative of the actual errors, leading to underestimation of the errors. The threshold of 0.02 dex mentioned above has been established by trial-and-error, it having been observed that after applying an initial lenient threshold, only the lower end of the uncertainty interval was affected, as would be expected. Stronger thresholds would however also start to reduce the size of the uncertainty interval at the high end. In addition, at 0.02 dex, only 20% of the resamplings comply with the threshold. Stronger thresholding would therefore likely inhibit the proper working of the bootstrap. In Section 5.6 we compare the the bootstrap interval to the scatter in simulated signal recovery, and find that the bootstrap interval is larger than the scatter. This may indicate that the bootstrap errors are still overestimating the errors due to other effects. In order to improve error estimation a suite of full end-to-end simulations would be required, possibly supported by cosmological hydrodynamic simulations and mock catalogs.

5.6 Recovery of Injected Decay Signals

To actually test the ability of this method to detect Dark Matter decay-like emission signals without the need to rely on spectral modeling, we inject into the real COSMOS dataset several simulated Dark Matter decay signals. For this purpose, we utilize the `fakeit` procedure, which creates a simulated event list given such information as the telescope response and exposure, and input parameters for the simulated signal. In this particular case, we seek to relate the simulated Dark Matter decay to actual Dark Matter particle models, to place the performance into context. We compare to the sterile neutrino model as discussed in e.g. Dodelson & Widrow (1994); Abazajian et al. (2001); Boyarsky et al. (2009c), which for a given particle mass m_{DM} and mixing angle $\sin^2(2\theta)$ generates a flux of

$$F_{DM} = 3.9 \cdot 10^{-7} \frac{ph}{cm^2s} \left[\frac{\sin^2(2\theta)}{10^{-8}} \right] \left[\frac{m_{DM}}{5keV} \right]^4 \times \left[\frac{D_L(z=0.1)}{D_L(z)} \right]^2 \left[\frac{M}{10^{13}M_\odot} \right] \quad (5.13)$$

for an object at redshift z with mass M .

We then inject a simulated redshifted gaussian line into the data at each cell that contains an object from the catalog, at $E = m_{DM}/2(z+1)$ with a normalization as given by the above equation and according to the mass and luminosity distance of the object. The generated event list is concatenated to the original cell's event list, and the correlation can be computed again.

We simulated a number of decay signals at various values for m_{DM} and $\sin^2(2\theta)$, for the G11 sample of galaxy groups. Because the injected signal is known and we can compare the pre-injection and post-injection correlation, we prove in Figure 5.2 that Equation 5.9 is correct. Here, we plot the input luminosity per unit mass against the luminosity derived from the difference between pre- and post-injection $w(\gamma=1)$ and equation 5.9. The correspondence is one-to-one up to the poissonian noise from the simulation.

For a few of these we show w as a function of restframe energy in Figure 5.3. Clearly, the only difference between these cases is the addition of a line at a particular energy. The

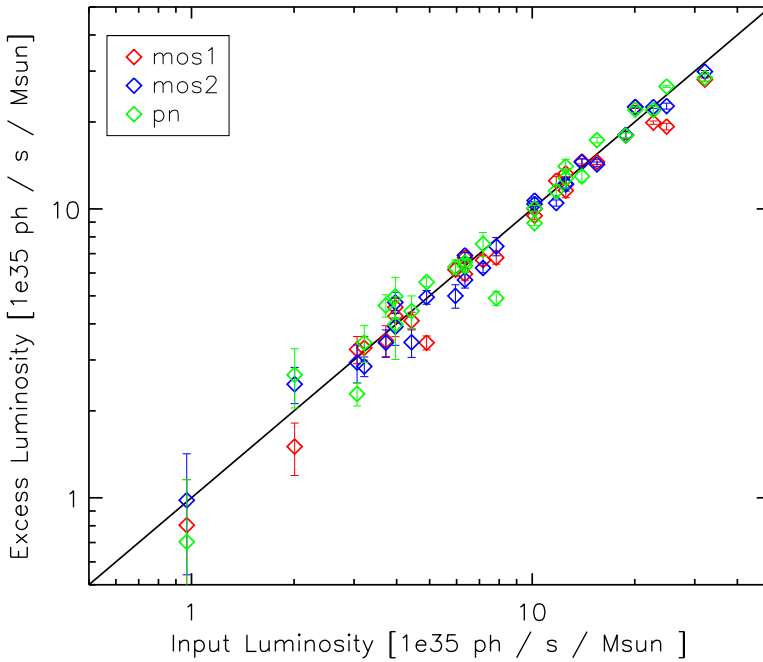


Figure 5.2: Practical proof of the validity of Equation 5.9 and therefore the formalism presented in this work. The input intrinsic Dark Matter decay luminosity (in $10^{35} \text{ ph s}^{-1} M_{\odot}^{-1}$) to the simulation compared to the *change* in correlation pre- and post-injection as expressed through Equation 5.9. Note that this is not a realistic measurement (see Section 5.6 for that), but rather the response of the correlation w to the injection of a fake signal. Excellent agreement confirms in practice the behaviour of the formalism as described in Section 5.2, while the scatter is due to the randomized nature of the `fakeit` simulation's photon distributions. Colors indicate the 3 different detectors, while the error bars in this case indicate the Poisson noise on the added signal.

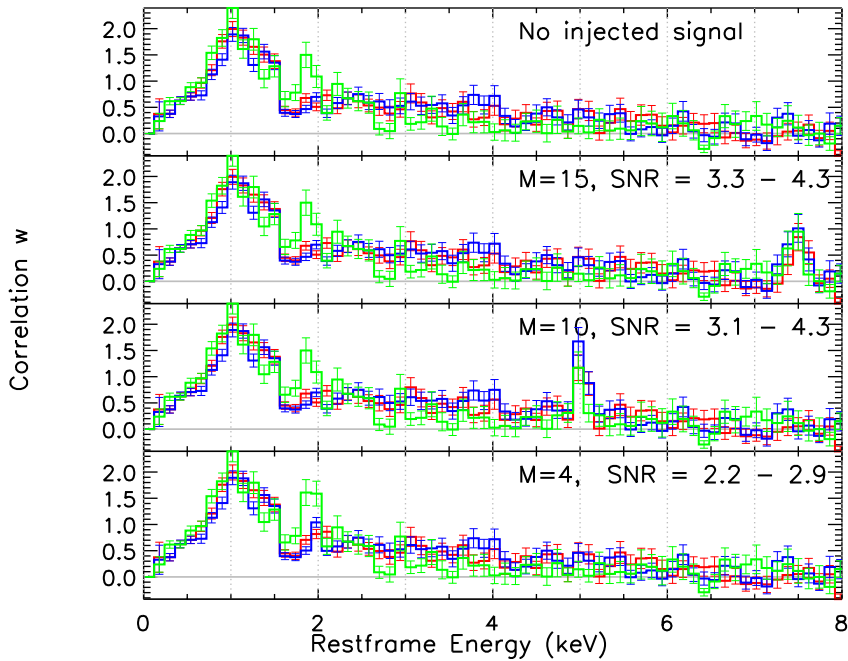


Figure 5.3: Example of the effect of a few simulated signals. Correlation as a function of energy for the nominal zCOSMOS field, using the G11 sample, correlating with M/D_{lum}^2 using energy bins of 120 eV, in the top panel. Other panels are the same, but with a simulated line added for a particle with the indicated mass (in keV, so the signal is at an energy half that mass). The SNR indicated is the total amount of raw simulated photons divided by the square root of the raw total amount of photons at the energy of injection. The range on the SNR indicated is spanned by the different cameras. Red, blue and green lines refer to MOS1, MOS2, and PN cameras respectively. Error bars are Poissonian propagated errors for indicative purposes only (see Section 5.5).

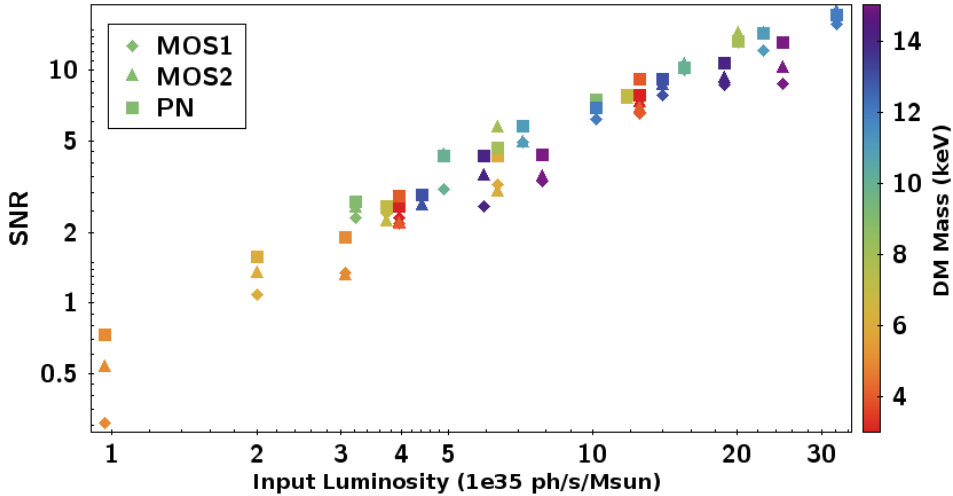


Figure 5.4: The set of simulated signals used in Section 5.6, by input luminosity versus the signal-to-noise ratio in terms in total injected signal photons to the square root of the total photons. The three cameras are shown separately as the SNR differs between them. Points are coloured by the Dark Matter particle mass used for that simulation, which is related to the injection restframe energy. Most of the scatter is due to different background counts between cameras and between restframe energies.

figure also indicates the signal-to-noise ratio of the injected signal, expressed as the total amount of simulated line photons over the square root of the total amount of photons in this energy bin as determined from a traditional stacking of the cells that contain groups (taking the proper blueshifting to restframe into account). It can be seen that even low SNR injected signals can be picked out by eye in some cases from the correlation.

The SNR for all injected signals is shown in Figure 5.4, separately for each instrument, and as a function of the input luminosity. Our sample has three injected signals below an SNR of 2, four signals between an SNR of 2 and 3, and SNR of 3 and higher starts at input luminosities of about $5 \cdot 10^{35}$ ph/s/ M_{\odot} . We draw attention to the fact that such weak (low SNR) signals would be challenging to detect in traditional stacking analyses. For reference, the simulated signal contains on average 1 photon per group at an SNR of around 6 (there are 180 groups in the G11 sample up to a redshift of 1.0).

Below we will show the efficiency at which the injected simulated signals are recovered using a variety of estimators as described in Section 5.3. For the remainder of this section, all results shown are for the following choices for the computation of the correlation and the estimator. We limit the field to that of the zCOSMOS sample, and use the full G11 X-ray selected catalog of galaxy groups except for those that fall outside the zCOSMOS field. In the computation of the correlation we only use events detected at an energy greater than 200 eV and less than 10 keV, and use energy bins (in restframe) of 50 eV. The correlator used is M^{γ}/D^2 with $\{\gamma\} = \{0, 0.25, 0.75, 1.0, 1.25, 1.75, 2.25\}$, i.e., we obtain a set of 7 correlations for different values of γ . The estimator Δ_{γ} is computed as

per Section 5.3, using 2 extra signal bins, 2 skipped bins and 3 side bins.

The best results are obtained using the approach where the set of Δ_γ are fit to a single component with a free β . The results are shown in Figure 5.5. The majority of all injected signals is recovered between 80 and 100% of the injected signal. High input luminosities are more likely to be correctly recovered. Essential is the fact that almost all best-fit β 's are found within 10% of the correct value of 1. It is confirmed that the correct value for β is equally well recovered for simulated injected signals that scale as M^2 (not shown).

The efficiency of the other three estimators mentioned in Section 5.3 is shown for reference in Figure 5.6. The most naive estimator, to simply use $\Delta_{\gamma=1}$ and solve Equation 5.9 with a single $\beta = 1$ component, returns a distribution that is centered on a ratio of measured-to-input of 1, with quite a broad spread. Fitting to the set of correlators, like in Figure 5.5, but with β fixed to 1 gives a distribution similar to Figure 5.5 but with more outliers. Fitting the set of correlators to multiple components (5 in this case) does not work well at all. This will be discussed further in Section 5.8.

As mentioned in Section 5.5, bootstrap errors are employed to estimate uncertainties on the best-fit values of L_β (and β). These errors were already shown in the left panel of Figure 5.5 on the individual simulated signals, for 100 bootstrap realizations. We check the validity of this approach by comparing the distribution of recovery ratios of the injected signals to the distribution of recovery ratios in the bootstrap realizations with the same injected signals. To be clear, the bootstrap only resamples the field, it does not simulate additional realizations on the injected signal. As can be seen in Figure 5.7, the bootstrap distribution for best-fit β resembles closely that of the nominal injected signals in Figure 5.5. As was described before, the bootstrap resampling also changes the group catalog so that some realizations may contain much less significant signal than others. A less significant signal will be more difficult to recover, so that the bootstrap intervals would overestimate the uncertainty at the low-recovery end. Therefore, as mentioned, the bootstraps are performed in such a way that only realizations of the group catalog are used that are close in expected signal significance to the nominal group catalog. For a too lenient filtering, then, the error bars are overestimated. However a too strict filtering would defeat the purpose of the bootstrap method in the first place. We determined from trial-and-error that a requirement of similarity of 0.02 dex in terms of the ratio of the value of $\sum_k M_k/D_k^2$ is justified, but not stronger. As can be seen in Figure 5.7, the distribution of the bootstrap realizations is however broader and peaks at slightly lower ratios than the nominal distribution in Figure 5.5. Although the number of simulated signals is not very large, this may indicate that the bootstrap errors overestimate the true uncertainty interval, despite the filtering of the bootstrap realizations. The source of this is currently unknown, but we will continue to use the bootstrap estimation of the errors as a conservative measure.

The choice of which bins to use for the calculation of Δ_γ (Equation 5.10) is of course of impact on the results presented above. The choice presented was taken for providing a distribution that peaked close to 1 (based on trial-and-error). Taking fewer signal bins or fewer skipped bins typically reduces the recovery ratio, but therefore also reduced the probability of over-estimation. Performing the correlation with smaller bins allows more freedom to refine this choice, but also increases computation times.

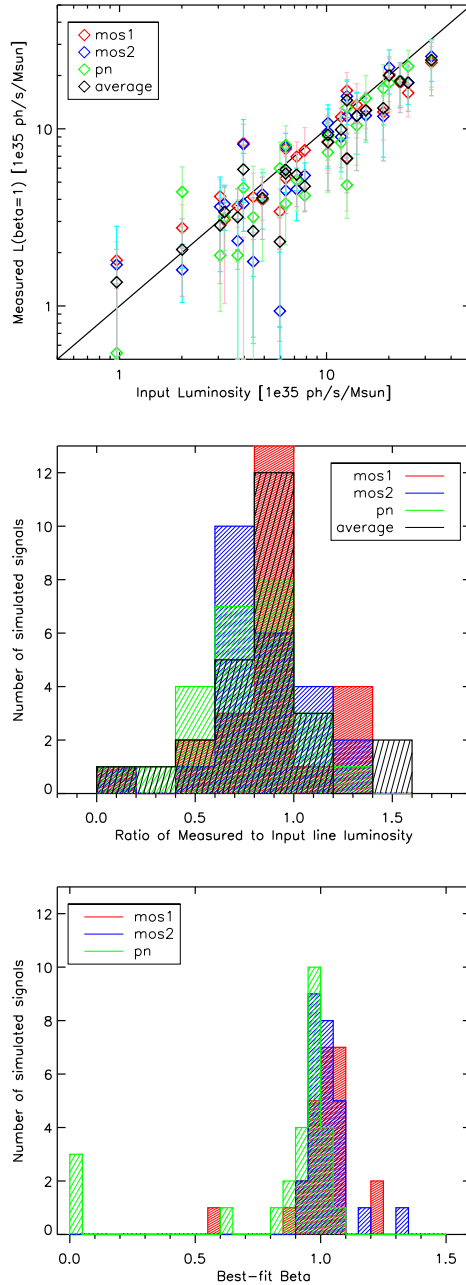


Figure 5.5: Recovery of injected decay signals using Δ_{γ} with M^{γ}/D^2 and $\{\gamma\} = \{0, 0.25, 0.75, 1.0, 1.25, 1.75, 2.25\}$ to fit a single component with L_{β} and β free parameters. In all panels, colors indicate the different detectors, and black is the weighted average of the detectors per injection. *Top:* for each individual injected signal the input luminosity and recovered L_{β} . Error bars indicate the distribution of 100 bootstrap realizations. *Middle:* histogram of the 26 injected signals in terms of the ratio of recovered to input signal. *Bottom:* histogram of the 26 injected signals in terms of the best-fit β .

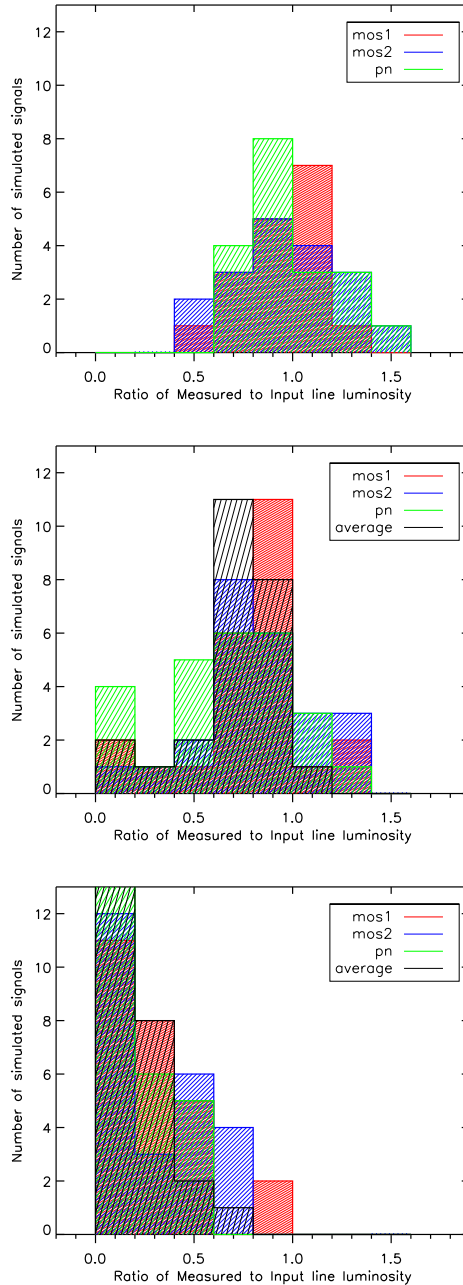


Figure 5.6: Recovery of injected decay signals using three different methods from Section 5.3. Shown as histograms of the 26 injected signals in terms of the ratio of recovered to input signal. In all panels, colors indicate the different detectors, and black is the weighted average of the detectors per injection. *Top:* directly solving Equation 5.9 for a single $\gamma = 1$ and a single $\beta = 1$ components. *Middle:* fitting Δ_γ with M^γ/D^2 and $\{\gamma\} = \{0, 0.25, 0.75, 1.0, 1.25, 1.75, 2.25\}$ to a single component with fixed $\beta = 1$. *Bottom:* fitting Δ_γ with M^γ/D^2 and $\{\gamma\} = \{0, 0.25, 0.75, 1.0, 1.25, 1.75, 2.25\}$ to five different components with $\{\beta\} = \{0.1, 0.5, 1.0, 1.5, 2.0\}$.

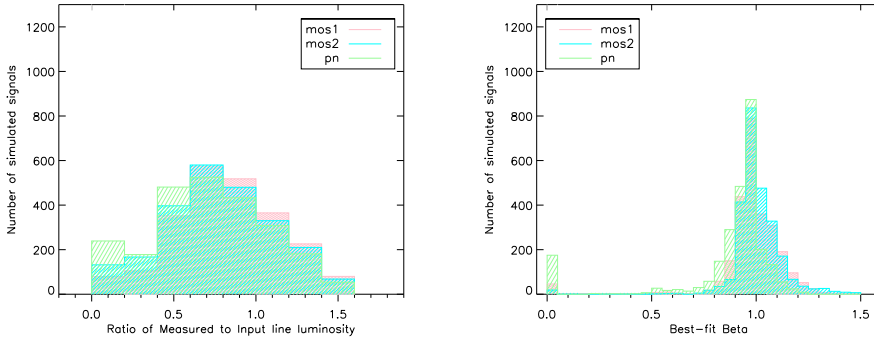


Figure 5.7: Same as Figure 5.5, but for all 100 bootstrap realizations of the 26 injected signals.

5.7 Conclusion and Summary

We have presented a new method to search for Dark Matter decay signals that does not rely on spectral modeling and does not require dedicated background subtraction. Both of these properties have been shown to be exhibited by our proof-of-concept implementation of the method on the COSMOS data using injected simulated Dark Matter decay signals.

The method relies on a spatial correlation and a correlation in redshift- and energy-space at the same time to automatically remove any and all backgrounds and foregrounds. The other main advantage is that by computing the correlation using different properties as the correlator, the behaviour of any correlated emission can be studied as a function of those properties. It is also possible to decompose the correlated emission into different components that behave differently. This method requires large datasets, but makes it possible to combine the advantages of a traditional stacking analysis with those of a population study (Sections 5.2 and 5.3).

We have performed a number of simulations of fake decay signals that have been injected into the real data, and subsequently attempted to recover this signal using our method. Our most effective estimator recovers the majority of the simulated signals to within 20 to 30% of the input luminosity, depending on the signal strength, and confirms its Dark Matter-like behavior (Section 5.6). It does so reliably for signals with a signal-to-noise ratio (raw photon count-based) as low as 3, without the need for any background subtraction or spectral modeling.

We expect that with more advanced fitting methods and considerable computing resources additional advantages of this method can be unlocked, such as independence of calibration uncertainties of both the instrument and the group properties (Section 5.8).

The sensitivity of the COSMOS dataset as a training set is not sufficient to reach the current state-of-the-art reported Dark Matter decay sensitivity, but the method is particularly well equipped to be scaled up and take advantage of all available archival data. As our simulations show that even very weak signals in terms of photon-count-based SNR have a good probability to be detected by our method, it is fully expected that it is possible to reach or surpass those sensitivities, while at the same time providing a robust tool for determining a potential signal's physical origin (Section 5.8).

5.8 Discussion

5.8.1 Decomposition Reliability

The estimator Δ_γ introduced in Section 5.3 requires a certain minimum of spectral information to detect line-like excesses. However, for weaker signals in regimes crowded with other emission lines (or other non-monotonic features) this approach may not work very well. In such a case it would be necessary to either take into account modeling of the full spectrum, or disregard any spectral information altogether. The latter being the main objective of the concept in this work. In this case, then, the Δ_γ estimator is skipped and the full set of $w_\gamma \langle I_\gamma \rangle$ is decomposed into a number of components as per Equation 5.9. Although in the current implementation presented in this work, the performance is sub-par, this approach offers dramatic advantages in the form of calibration independence once the reliability has been increased.

The results presented in Section 5.6 focused on decomposing the estimator Δ_γ into a single component with one or two free parameters. In the bottom panel of Figure 5.6, some results of a decomposition of $\{\Delta_\gamma\}$ into 5 components with a single free parameters each show that this approach did not work well in this case. Foregoing the estimator and employing the decomposition of the entire correlated signal delivers better results however. For the same set of correlators, the same fake injected signals and the same components (ie., components with the same β 's), the results for the L_β component are shown in Figure 5.8. This is an improvement compared to the same decomposition but only relying on the estimator Δ_γ , but compares badly to the single-component estimators.

The main problem for the multi-component minimization procedure is likely that the coefficients $A_{\beta\gamma}$ from Equation 5.9 behave monotonically with both β and γ , which may incur some degeneracies if the number of components is too large. Additionally, it may not be clear what the ideal choice of β 's should be for the decomposition. Although the $\beta = 1$ choice is clearly physically motivated, the usual astrophysical signals in a narrow energy bin may not necessarily scale as any power of the group mass, but rather as some more complicated function. The set of 5 values used above was chosen solely in order to provide enough opportunity for any signal that is not Dark Matter decay to ‘choose’ components with $\beta \neq 1$.

Possibilities for improvement then might be found along the following lines. Firstly, the decomposition may be performed with an intermediate number of components, such as a single $\beta = 1$ component and one or two components with a free (within some limits) β . Secondly, the correlators nor the components chosen necessarily all have to be different powers of the same property. It should be possible to combine, for example, mass and temperature correlators, as long as the additive components of the decomposition are physically independent processes. Thirdly, the minimization procedure would likely profit, especially for more decompositions with more than 1 component, of a more advanced implementation, such as a Markov-chain Monte-Carlo method to more fully explore any degeneracies.

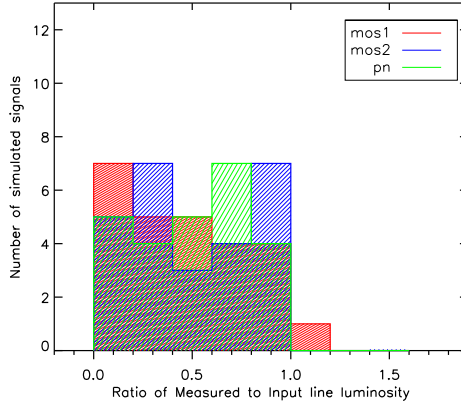


Figure 5.8: Recovery of injected decay signals using the full correlated signal $w_\gamma \langle I_\gamma \rangle$ with M^γ/D^2 and $\{\gamma\} = \{0, 0.25, 0.75, 1.0, 1.25, 1.75, 2.25\}$ to fit 5 components with $\{\beta\} = \{0.1, 0.5, 1.0, 1.5, 2.0\}$. The recovery rate leaves something to be desired, but is already better than those of the bottom panel of Figure 5.6. Possible improvements to this approach are discussed in Section 5.8.1.

5.8.2 Independence of Calibration Uncertainties

As alluded to earlier, the decomposition of the full correlated signal $w_\gamma \langle I_\gamma \rangle$ can be used to obtain an estimator that is independent of telescope and mass calibration. To be explicit, if the instrumental (mis-)calibration can be described as,

$$I = mI' + a$$

with I the measured X-rays flux, I' the true intrinsic X-ray flux, m a multiplicative calibration factor and a an additive calibration component; and if the mass calibration suffers only from a multiplicative bias so that $M = bM'$ with M the measured (reported) mass, M' the true mass and b the multiplicative bias; then the following estimator is insensitive to all of a , m and b :

$$\begin{aligned} Q_{\beta\gamma} &\equiv \frac{L_\beta A_{\beta\gamma}}{w_\gamma \langle I_\gamma \rangle} \\ &= \frac{L'_\beta A'_{\beta\gamma}}{w'_\gamma \langle I_\gamma \rangle'} = Q'_{\beta\gamma} \end{aligned} \quad (5.14)$$

where all primed quantities represent the true values of the measured un-primed quantities. The derivation can be found in Appendix 5.A.2. The estimator $Q_{\beta\gamma}$ rather intuitively represents the fraction of the total correlated emission (the denominator) sourced by the component with exponent β . It also has the convenient property that

$$Q_\gamma \equiv \sum_{\{\beta\}} Q_{\beta\gamma} = 1 \quad (5.15)$$

by definition if the decomposition is perfect and noise-free.

The estimator effectively sacrifices any knowledge about the intrinsic value of any L_β (which is why Q still depends on γ) to get rid of the dependence on the calibration. The obvious advantage of this is that multiple instruments can be combined in order to increase sensitivity without introducing additional systematics related to calibration. We expect that it is also possible using this formalism to combine not just different instruments but also different fields and surveys without losing the benefits described above. As long as the type of object being correlated with is the same across surveys (ie, described by the same physics and thus by the same decomposition), the L_β measured from the decomposition should be the same for each survey performed with the same telescope (assuming the telescope calibration does not change between surveys, and that the mass calibration b is also the same for each survey). In that case, the L_β 's from different surveys but the same instrument can be combined by some weighted average, or one may perform the fitting procedure to all surveys simultaneously, keeping the values of L_β tied between surveys (the coefficients $A_{\beta\gamma}$ will change between surveys). To obtain $Q_{\beta\gamma}$, a single survey has to be chosen as an arbitrary reference point. After all, the L_β are the same for each survey, but $A_{\beta\gamma}$, w_γ and $\langle I_\gamma \rangle$ are not. As long as the same survey is chosen as the reference point for each instrument, the Q 's of the individual instruments can now also be combined to obtain a final single value for $Q_{\beta\gamma}$.

In principle, $Q_{\beta\gamma}$ can also be constructed for the estimator Δ_γ , without the loss of calibration independence. For single-component estimators however, the actual value of $Q_{\beta\gamma}$ does not provide a lot of insight, except that any deviation from the value of 1 is equal to the residual in the fitting process. The advantage would be found though in the ability to reduce the error bars by being able to combine results from different instruments without having to worry about calibration issues.

Since the full decomposition (without using the estimator Δ_γ) requires more advanced minimization or fitting implementations to deliver reliable results (see Section 5.8.1), and because in simulating the injected Dark Matter decay signals the available telescope response files are assumed to be correct, practical proof of the usefulness of this estimator Q will have to be postponed until future work that is able to employ a full suite of end-to-end simulations of the data.

5.8.3 Flexibility and Robustness

One advantage of this method is that when a candidate signal is found, it is possible to perform many tests by merely changing the parameters of the correlation and decomposition. A real signal should be present in all circumstances albeit with different significance. Jack-kniving is the first obvious possibility; the dataset can be split by redshift, by various properties of the objects in the catalog like mass or size, or by sky location. The process of bootstrapping already covers some of these variations in a way. Secondly, the object catalog can be changed. This may help root out possible selection bias. For example, if the X-ray coverage is not of uniform depth, the objects in an X-ray-selected catalog could be correlated with the exposure depth (see Section 5.4). Thirdly, even though the instrumental backgrounds cancel out very well on average (see Section 5.B), chance alignments of spectral and spatial inhomogeneities with catalog objects may occur. It is possible to check for robustness here by simply masking known instrumental lines in the raw data (observed frame). Because of the correlation being performed in redshift-space (object

restframe), any real signal will persist even though the significance will be lower over some energy range.

5.8.4 Sensitivity

With regards to the sensitivity to Dark Matter decay signals, this dataset is limited mostly by the total number of (expected) signal photons, as can be seen from the simulations in Section 5.6. Increasing the expected number of photons can be achieved in three ways; increased exposure depth, a larger number of objects in the group catalog, or a larger field. With increased photon statistics, it also becomes increasingly important to reduce the amount of noise in the correlation caused by chance alignments. Because the formalism depends on backgrounds (and other effects not related to the group catalog) canceling out on average, and because the field under consideration (or any field) is finite, some chance alignments between objects and some spatial or spectral inhomogeneities in the backgrounds are expected, inducing fluctuations in the correlation (see Section 5.B). Both higher object counts and larger fields (more cells) will improve the fidelity of correlation in this respect and thus reduce the (bootstrap) error bars. Note that the objects should have an appreciable range in redshifts for the correlation in redshift-space to work properly.

The simulated signals in this work are stronger than what is already currently ruled out from previous studies. The 3.5 keV signal (Boyersky et al., 2014a; Bulbul et al., 2014a), for example, would have an L_β of the order $0.1 \text{ ph s}^{-1} M_\odot^{-1}$, an order of magnitude weaker than the weakest signal we've simulated here.

5.A Derivations

5.A.1 Correlation (De)composition

In this Appendix, Equation 5.9 is derived. We will write N_k for the correlator (which can represent any combination of properties of the galaxy group k). For readability, we will assume the X-rays consists of a single monochromatic Dark Matter decay component, and a homogeneous backgrounds component. We will generalize this later. The detected X-ray flux in a cell i at observed energy E' that contains a galaxy group at redshift z_i is then

$$I_i(E') = \delta [E'(1 + z_i) - E_0] LM_i/D_i^2 + I_{bg,i}(E') \quad (5.16)$$

with L the intrinsic luminosity per unit mass of the component, E_0 the restframe energy of this monochromatic component, M_i the group's mass and D_i^2 short for $4\pi D_{lum,i}^2$ representing the luminosity distance to the group. The δ -function makes sure the monochromatic component only contributes to $I_i(E')$ if $E' = E_0/(1 + z_i)$.

We then put this into the equations for the correlation, where we will sometimes write a sum over N_{cwg} meaning all cells that contain a group, and we will use the shorthand

$\sum_k N_k$ to mean the sum over all groups in the catalog.

$$\frac{\langle NI(E) \rangle}{\langle N \rangle} = \frac{1}{N_C} \frac{N_C}{\sum_k N_k} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)} [I_i(E' = E/(1+z_k)) N_k] \quad (5.17)$$

$$= \frac{1}{\sum_k N_k} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)} \left(\delta[(1+z_i)E/(1+z_k) - E_0] LM_k N_k / D_k^2 \right. \\ \left. + I_{bg,i}(E/(1+z_k)) N_k \right) \quad (5.18)$$

Note that $\sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)}$ is just the same as the sum over all groups, so that $z_i = z_k$ in this summation;

$$\frac{\langle NI(E_0) \rangle}{\langle N \rangle} = \frac{L}{\sum_k N_k} \sum_k^{N_g} N_k M_k / D_k^2 + \frac{1}{\sum_k N_k} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)} I_{bg,i}(E_0/(1+z_k)) N_k \quad (5.19)$$

Repeat for $\langle I \rangle$

$$\langle I(E) \rangle = \frac{1}{N_C \sum_k N_k} \sum_i^{N_C} \sum_k^{N_g} \left[I_i(E' = E/(1+z_k)) N_k \right] \quad (5.20)$$

$$= \frac{1}{N_C \sum_k N_k} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g} \left(\delta[(1+z_i)E/(1+z_k) - E_0] LM_i N_k / D_i^2 \right) \\ + \frac{1}{N_C \sum_k N_k} \sum_i^{N_C} \sum_k^{N_g} I_{bg,i}(E/(1+z_k)) N_k \quad (5.21)$$

Notice that the first double-summation runs over all cells *with a group* and all groups in the catalog, instead of *all* cells and all groups. Therefore, in this first double-sum, z_i and z_k are not always the same. At $E = E_0$, the δ -function however does enforce $z_i = z_k$, so that

$$\langle I(E_0) \rangle = \frac{L}{N_C \sum_k N_k} \sum_k^{N_g} N_k M_k / D_k^2 + \frac{1}{N_C \sum_k N_k} \sum_i^{N_C} \sum_k^{N_g} I_{bg,i}(E_0/(1+z_k)) N_k \quad (5.22)$$

Now we note that the parts from both $\langle NI \rangle / \langle N \rangle$ and $\langle I \rangle$ that contain I_{bg} and are on average equal if the background and the groups are uncorrelated (so that one might replace every $I_{bg,i}$ with I_{bg}). This leads to

$$w \langle I \rangle (E_0) = L \frac{\sum_k N_k M_k / D_k^2}{\sum_k N_k} \left(1 - \frac{1}{N_C} \right) \quad (5.23)$$

Since the number of cells is typically very large (more than 3000 in this work), the contribution of the correlated components to $\langle I \rangle$ is typically very small, ie. $(1 - 1/N_C) \simeq 1$.

This can be generalized by replacing the scaling M_k/D_k^2 with any other desired scaling (which we will represent with S_k).

If the component is constant in energy instead of monochromatic, the result would instead be

$$w\langle I \rangle = L \left(\frac{\sum_k N_k M_k / D_k^2}{\sum_k N_k} - \frac{\sum_k M_k / D_k^2}{N_C} \right) \quad (5.24)$$

and independent of energy. Again due to the factor $1/N_C$, in practice the change between monochromatic and constant in energy is about 1% at most for the dataset used in this work.

For completeness, we must note that for a single monochromatic component at E_0 the relevant part of $\langle NI(E_j) \rangle$ for any other $E_j \neq E_0$ will be zero, but the part of $\langle I(E_j) \rangle$ containing $L\delta((1+z_i)E_j/(1+z_k) - E_0)$ will not, as there may be pairs of (z_i, z_k) that solve $(1+z_i)E_j/(1+z_k) = E_0$. This reduces the correlation at every $E_j \neq E_0$, but only marginally due to the factor $1/N_C$.

Any arbitrary physical scenario can be expressed as the sum of many δ components at different energies E_s with their scalings S_s :

$$I_i = \sum_s L_s S_{s,i} + I_{bg,i} \quad (5.25)$$

The contributions to $\langle NI \rangle$ and $\langle I \rangle$ can then also simply be summed so that finally the generalized expression is

$$w\langle I \rangle = \sum_s L_s \frac{\sum_k N_k S_{s,k}}{\sum_k N_k} \left(1 - \frac{1}{N_C} \right), \quad (5.26)$$

neglecting as before the smearing of each δ -function as part of $\langle I \rangle$. This approximation may not work well if one is interested in components whose intrinsic luminosity (in terms of $L_s \sum_k N_k S_{s,k} / \sum_k N_k$) is of the order N_C smaller than the dominant component, setting a fundamental sensitivity limit for a given number of cells (ie, size of the field).

5.A.2 Calibration Independence

Here we will prove that $Q_{\beta\gamma}$ (Equation 5.14) is indeed independent of telescope and mass calibrations. The derivation is very similar to that of the previous Section. In this case however, we will suppress the explicit writing of energy dependence for now for readability.

Let I'_X be the true intrinsic X-ray emission from catalog objects in particular cell, I'_r any physical X-rays not sourced by and therefore not correlated with catalog objects, m the multiplicative calibration of the instrument and a the additive calibration of the instrument and the instrumental background. Let M' be an object's true mass, and b the multiplicative bias on that mass, so that $M = bM'$. Let L'_β be the true intrinsic luminosity per unit (mass $^\beta$ /distance 2) for the X-ray emission component process that scales as (mass $^\beta$). Then the measured flux I_i in cell i will be

$$I_i = m(I'_X + I'_r) + a \quad (5.27)$$

$$= m(I'_r + \sum_{\{\beta\}} L'_\beta M_i'^\beta / D_i^2) + a \quad (5.28)$$

where the sum over the set of components $\{\beta\}$ runs over all physical processes at work (which we do not need to specify here).

Writing N_{cwg} for the number of cells that contain a group, $N_g(i)$ for the numbers of groups in cell i , then the measured quantities relate to the true quantities in the following way;

$$\langle N \rangle(\gamma) = \frac{1}{N_C} \sum_k^{N_g} b^\gamma M_k'^\gamma \quad (5.29)$$

$$= b^\gamma \langle N \rangle'(\gamma) \quad (5.30)$$

$$\langle NI \rangle(\gamma) = \frac{1}{N_C} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)} \left(m I_{r,i}' + a + \sum_{\{\beta\}} m L'_\beta M_i'^\beta / D_i^2 \right) (b^\gamma M_k'^\gamma) \quad (5.31)$$

$$= \frac{mb^\gamma}{N_C} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)} \left(\sum_{\{\beta\}} L'_\beta M_i'^\beta M_k'^\gamma / D_i^2 + I_{r,i}' M_k'^\gamma \right) + \frac{1}{N_C} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)} ab^\gamma M_k'^\gamma \quad (5.32)$$

$$= mb^\gamma \langle NI \rangle'(\gamma) + \frac{1}{N_C} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)} ab^\gamma M_k'^\gamma \quad (5.33)$$

$$\langle I \rangle(\gamma) = \frac{1}{N_C} \left(\sum_l^{N_g} b^\gamma M_l'^\gamma \right)^{-1} \left(\sum_i^{N_{\text{cwg}}} \sum_k^{N_g} m \sum_{\{\beta\}} L'_\beta M_i'^\beta b^\gamma M_k'^\gamma + \sum_i^{N_C} \sum_k^{N_g} (I_{r,i}' m + a) b^\gamma M_k'^\gamma \right) \quad (5.34)$$

$$= m \langle I \rangle'(\gamma) + \frac{1}{N_C} \left(\sum_l^{N_g} M_l'^\gamma \right)^{-1} \sum_i^{N_C} \sum_k^{N_g} a M_k'^\gamma \quad (5.35)$$

The uncorrelated emission I_r is still part of $\langle I \rangle$ and $\langle NI \rangle$, both in the primed and unprimed quantities, but cancels out in w and $w \langle I \rangle$. Notice that in Equation 5.33 we can simplify

$$\frac{1}{N_C} \sum_i^{N_{\text{cwg}}} \sum_k^{N_g(i)} ab^\gamma M_k'^\gamma = \frac{1}{N_C} \sum_k^{N_g} ab^\gamma M_k'^\gamma, \quad (5.36)$$

and in Equation 5.35

$$\frac{1}{N_C} \left(\sum_l^{N_g} M_l'^\gamma \right)^{-1} \sum_i^{N_C} \sum_k^{N_g} a M_k'^\gamma = \left(\sum_l^{N_g} M_l'^\gamma \right)^{-1} \sum_k^{N_g} a M_k'^\gamma, \quad (5.37)$$

so that those parts cancel out below;

$$w(\gamma)\langle I \rangle(\gamma) = \frac{\langle NI \rangle(\gamma)}{\langle N \rangle(\gamma)} - \langle I \rangle(\gamma) \quad (5.38)$$

$$= m \left(\frac{\langle NI \rangle'(\gamma)}{\langle N \rangle'(\gamma)} - \langle I \rangle'(\gamma) \right) = mw'(\gamma)\langle I \rangle'(\gamma) \quad (5.39)$$

The decomposition of the different correlating components as described in Section 5.2 is now affected as follows. Recall that in Equation 5.9

$$A_{\beta\gamma} = \frac{\sum_k M_k^{\beta+\gamma}/D_k^2}{\sum_k M_k^\gamma} \left(1 - \frac{1}{N_C} \right) \quad (5.40)$$

so that $A_{\beta\gamma} = b^\beta A'_{\beta\gamma}$. All of the above means that the solutions to Equation 5.9 will be

$$L_\beta = \frac{m}{b^\beta} L'_\beta \quad (5.41)$$

We can then construct the following quantity that reflects which fraction of the correlated emission is explained by the component with exponent β

$$Q_{\beta\gamma} \equiv \frac{L_\beta A_{\beta\gamma}}{w(\gamma)\langle I \rangle(\gamma)} \quad (5.42)$$

$$= \frac{L'_\beta A'_{\beta\gamma}}{w'(\gamma)\langle I \rangle'(\gamma)}. \quad (5.43)$$

This quantity is completely independent of detector calibration and of multiplicative mass bias, as shown.

The energy dependence of m and a were not explicitly taken into account above for readability. Comparing with the previous section, it will be clear that this does not change the independence of $Q_{\beta\gamma}$. The energy dependence of a functions exactly like $I_{bg,i}$ in the previous section, which canceled out completely since it is an uncorrelated additive component. Regarding $m = m(E') = m(E/(1+z_k)) \equiv m_k$, one may see that if this factor is introduced in the equations of the previous section, it becomes part of the summation like so

$$\frac{1}{\sum_k N_k} \sum_k^{N_g} m_k N_k M_k / D_k^2 \quad (5.44)$$

it is clear that the calibration of the instrument at some energy in the observed frame, $m(E')$, is uncorrelated with any properties of the group k . On average then, the preceding expression can be written as

$$\frac{1}{\sum_k N_k} \sum_k^{N_g} m_k N_k M_k / D_k^2 = \left(\frac{\sum_k m_k}{N_g} \right) \frac{\sum_k N_k M_k / D_k^2}{\sum_k N_k} \quad (5.45)$$

so that $\sum_k m_k / N_g$ can be associated with the factor m in Equations 5.33, 5.35, 5.39 and 5.41. All of which is to say that the intuitively expected

$$\frac{\langle NI \rangle}{\langle N \rangle \langle I \rangle} = \frac{\langle NmI' \rangle}{\langle N \rangle \langle mI' \rangle} = \frac{\langle m \rangle \langle NI' \rangle}{\langle N \rangle \langle m \rangle \langle I' \rangle}. \quad (5.46)$$

is confirmed to hold also when performing the correlation in redshift space.

5.B Null Tests

One of the main features of this method is the fact that it automatically finds the correlated emission, and does not require separate background subtraction or modeling. This is possible because of the large field and object catalog, and because of the correlation in redshift space. To show that this feature works as claimed, we perform the following null test.

We first create a instrumental background-only facsimile version of the COSMOS field. This is achieved by using the available *XMM-Newton* closed-filter calibration observations. These observations are performed, as indicated by the name, while the filter wheel is in the closed position, blocking any physical emission, leaving only the instrumental backgrounds. To create a mosaic that resembles the actual instrumental background in the COSMOS observations, the closed filter dataset is cut into parts with roughly similar numbers of event counts as the real COSMOS exposures. Each part is then made a piece of the mosaic by using the `skycast` procedure to give all events sky coordinates.

The resulting mosaic is then used instead of the real data to compute the correlation as described above. The same catalogs are used, but in addition the correlation is performed on a number of randomized versions of the catalogs. So for each iteration, all objects in the catalog are assigned to a random cell. This means that all catalog properties remain the same, such as the redshift distribution.

We show the result of 40 of such randomized catalogs of the G11-sample, using as correlator M/D_{lum}^2 , in Figure 5.9. The average of all the random catalogs is shown in red, and is very close to 0 for all energies, as is expected. The variation is shown in the red error bars, being the 68% intervals. This variation indicates the potential impact of chance alignments of objects in the catalog and spatial and spectral features in the instrumental backgrounds.

An additional on-the-spot null test can be included with every correlation. If the observed X-rays used for the correlation include events up to high energies, so that the telescope's effective area at these energies is essentially negligible, then the high-energy tail of the correlation spectrum should not include any correlated signal (up to some random fluctuations).

5.C Weighted Correlations

Although in a sense the performing of a correlation could be described as a weighted average, strictly speaking the equations for $\langle NI \rangle$ and $\langle I \rangle$ (Equations 5.5 and 5.6) would be different if they both employed weighting by some cell's property in the form of

$$X = \frac{\sum_i x_i w_i}{\sum_i w_i}. \quad (5.47)$$

The difference with correlating is that weighting is used when there are certain properties of the cell that are not related to the correlating property of the groups *nor* to the correlating property of the cell, but which for some reason should increase or decrease how important this single cell is to the total.

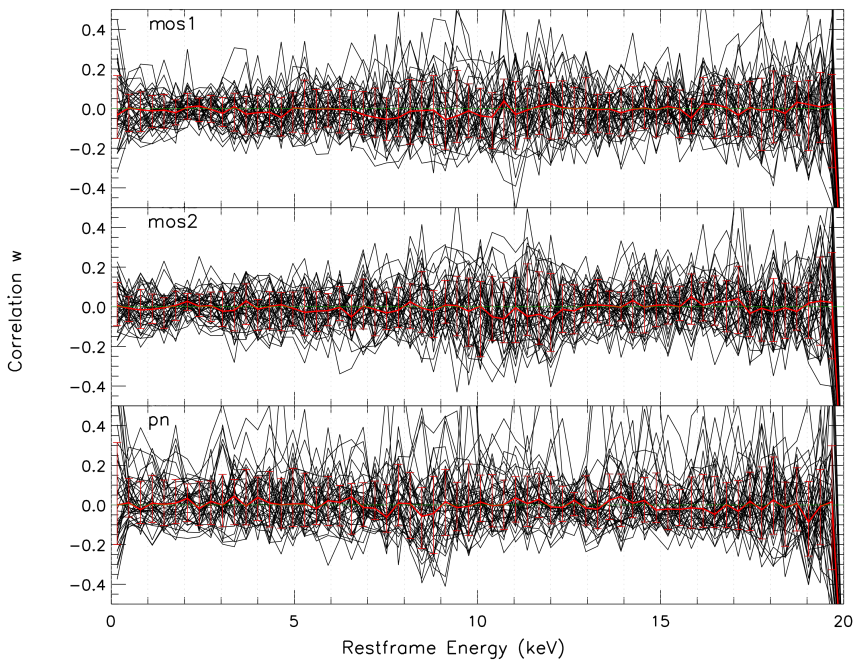


Figure 5.9: Results of the correlation for 40 randomized versions of the G11-sample. Median and 68% interval shown in red. Correlator used is M/D_{lum}^2 , with energy bins of 320 eV.

We will here show for completeness how a weighting can be incorporated in the formalism, although one should beware that including a weight with residual relations between the quantities being correlated will induce systematic offsets.

Let us first rewrite $\langle I \rangle$

$$\langle I \rangle = \frac{1}{N_C} \sum_{i=1}^{N_C} \frac{1}{\sum_{l=1}^{N_g} N_l} \sum_{k=1}^{N_g} \zeta_{ik} N_k \quad (5.48)$$

$$= \frac{N_g}{\sum_{l=1}^{N_g} N_l} \frac{\sum_{i=1}^{N_C} \sum_{k=1}^{N_g} \zeta_{ik} N_k}{\sum_{i=1}^{N_C} \sum_{k=1}^{N_g} 1} \quad (5.49)$$

For $\langle NI \rangle$, we first observe that

$$\sum_{i=1}^{N_C} \frac{1}{N_C(i, \theta)} \sum_{j=1}^{N_C(i, \theta)} \sum_{k=1}^{N_g(j)} 1 = N_g \quad (5.50)$$

so that

$$\frac{\frac{1}{N_C} \sum_{i=1}^{N_C} \frac{1}{N_C(i, \theta)} \sum_{j=1}^{N_C(i, \theta)} \sum_{k=1}^{N_g(j)} \gamma(I_i, z_k) N_k}{\frac{1}{N_C} \sum_{i=1}^{N_C} \frac{1}{N_C(i, \theta)} \sum_{j=1}^{N_C(i, \theta)} \sum_{k=1}^{N_g(j)} 1} = \langle NI \rangle \frac{N_C}{N_g} \quad (5.51)$$

and we can write

$$\langle NI \rangle = \frac{N_g \sum_{i=1}^{N_C} \frac{1}{N_C(i, \theta)} \sum_{j=1}^{N_C(i, \theta)} \sum_{k=1}^{N_g(j)} \gamma(I_i, z_k) N_k}{\sum_{i=1}^{N_C} \frac{1}{N_C(i, \theta)} \sum_{j=1}^{N_C(i, \theta)} \sum_{k=1}^{N_g(j)} 1} \quad (5.52)$$

Now when adding the weights a_i the equations become

$$\langle NI \rangle = \frac{N_g \sum_{i=1}^{N_C} \frac{1}{N_C(i, \theta)} \sum_{j=1}^{N_C(i, \theta)} \sum_{k=1}^{N_g(j)} \gamma(I_i, z_k) N_k a_i}{\sum_{i=1}^{N_C} \frac{1}{N_C(i, \theta)} \sum_{j=1}^{N_C(i, \theta)} \sum_{k=1}^{N_g(j)} a_i} \quad (5.53)$$

$$\langle I \rangle = \frac{N_g \sum_{i=1}^{N_C} \sum_{k=1}^{N_g} \zeta_{ik} N_k a_i}{\sum_{l=1}^{N_g} N_l \sum_{i=1}^{N_C} \sum_{k=1}^{N_g} a_i} \quad (5.54)$$

Initial tests of a weighting by each cell's product of exposure time, effective area and usable solid angle yielded systematic positive and increasing offsets at the high energy tail of the correlation spectrum (where it should be zero), for both the G11 sample and the 20k sample. We speculate this may be caused by telescope calibration uncertainty. Further application and study of any weights during the correlation was not pursued.