

# The EXOD search for faint transients in *XMM-Newton* observations: Method and discovery of four extragalactic Type I X-ray bursters

I. Pastor-Marazuela<sup>1,2,3,4</sup>, N. A. Webb<sup>3,4</sup>, D. T. Wojtowicz<sup>3,4,5</sup>, and J. van Leeuwen<sup>2,1</sup>

<sup>1</sup> Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands  
e-mail: [i.pastormarazuela@uva.nl](mailto:i.pastormarazuela@uva.nl)

<sup>2</sup> ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

<sup>3</sup> Université de Toulouse, UPS-OMP, IRAP, Toulouse, France

<sup>4</sup> CNRS, IRAP, 9 Av. colonel Roche, BP 44346, 31028 Toulouse CEDEX 4, France

<sup>5</sup> IRIT, Université Toulouse III – Paul Sabatier, 118 route de Narbonne, 31062 Toulouse CEDEX 9, France

Received 8 October 2019 / Accepted 16 May 2020

## ABSTRACT

**Context.** The observations carried out with *XMM-Newton* have produced a very extensive X-ray source catalogue in which the standard pipeline determines the variability of sufficiently bright sources through  $\chi^2$  and fractional variability tests. Faint sources, however, are not automatically checked for variability, and this means that faint, short timescale transients are overlooked. From dedicated X-ray searches, as well as optical and radio archive searches, we know that some such dim sources can still be identified with high confidence.

**Aims.** Our goal is to find new faint, fast transients in *XMM-Newton* EPIC-pn observations. To that end we created the EPIC-pn *XMM-Newton* outburst detector (EXOD) algorithm, which we run on the EPIC-pn full-frame data available in the 3XMM-DR8 catalogue.

**Methods.** In EXOD, we computed the variability of the whole field of view by first binning in time the counts detected in each pixel of the detector. We next computed the difference between the median and maximal number of counts in each time bin and pixel to detect variability. We applied EXOD to 5751 observations in the full frame mode and compared the variability of the detected sources to the standard  $\chi^2$  and Kolmogorov–Smirnov (KS) variability tests.

**Results.** The algorithm is able to detect periodic and aperiodic variability, with both short and long flares. Of the sources detected by EXOD, 60–95% are also shown to be variable by the standard  $\chi^2$  and KS tests. EXOD computes the variability over the entire field of view faster than the light curve generation takes for all the individual sources. We detect a total of 2961 X-ray variable sources. After removing the spurious detections, we obtain a net number of 2536 variable sources. Of these we investigate the nature of 35 sources with no previously confirmed classification. Amongst the new sources, we find stellar flares and AGNs, in addition to four extragalactic type I X-ray bursters that double the known neutron-star population in M31.

**Conclusions.** This algorithm is a powerful tool for the prompt detection of interesting variable sources in *XMM-Newton* observations. EXOD also detects fast transients that other variability tests would classify as non-variable due to their short duration and low number of counts. This is of increasing importance for the multi-messenger detection of transient sources. Finally, EXOD allows us to identify the nature of compact objects through their variability and to detect rare compact objects. We demonstrate this through the discovery of four extragalactic neutron-star low-mass X-ray binaries, doubling the number of known neutron stars in M31.

**Key words.** methods: data analysis – X-rays: general – X-rays: binaries – X-rays: bursts – stars: flare

## 1. Introduction

*XMM-Newton* has been inspecting the X-ray sky since its launch in 1999. Its three on-board telescopes, each with a geometric effective area of  $\sim 1550 \text{ cm}^2$  at 1.5 keV, are combined to provide the largest total collecting area of any X-ray telescope launched (Jansen et al. 2001). This has allowed for the compilation of a very large catalogue of X-ray detections. We used the eighth data release of the *XMM-Newton* catalogue, 3XMM-DR8, which contains 775 153 detections (Rosen et al. 2016)<sup>1</sup>.

The large majority of these detections are of steady, unvarying sources. Compared to this static sky, the dynamic sky remains relatively unexplored. Many violent and rare variable phenomena such as tidal disruption events and X-ray bursters can be observed, which can be used for studying (astro)physics,

including strong gravity and thermonuclear explosions. The X-ray regime is one part of multi-messenger time-domain astronomy, which focuses on this dynamical sky. The detection of cosmic rays and neutrinos (e.g. Hirata et al. 1987), of extragalactic bursts in radio (Lorimer et al. 2007), type Ia supernovae in the optical (e.g. Riess et al. 1998), and of double neutron star mergers in gravitational waves (Abbott et al. 2017) show that transient surveys are of vital importance to understand the physics behind the explosive universe.

*XMM-Newton* was not built as a transient detector. However, the high sensitivity and high time resolution of its European Photon Imaging Cameras (EPIC) allow it to record fast X-ray transients. Nevertheless, the detection of these transients and potential follow-up with other instruments depends on the sensitivity and performance of the variability detection pipeline. The pipeline processing for the 3XMM catalogue includes such a variability study, for sources with total EPIC counts (including the detectors pn, MOS1 and MOS2) exceeding 100.

<sup>1</sup> 3XMM-DR8 website: [http://xmssc.irap.omp.eu/Catalogue/3XMM-DR8/3XMM\\_DR8.html](http://xmssc.irap.omp.eu/Catalogue/3XMM-DR8/3XMM_DR8.html)

The variability of these brightest sources is studied through two complementary tests; the first is a simple  $\chi^2$  test where the time series are fitted to a constant model. The time bin width is the lowest integer multiple of 10 s for which the average number of counts is  $\geq 18$  counts  $\text{bin}^{-1}$ . Sources with a  $\chi^2$  probability of constancy  $P(\chi^2) \leq 10^{-5}$  are flagged as variable (Watson et al. 2009). The second method is the study of the fractional variability amplitude,  $F_{\text{var}}$ , as described in Rosen et al. (2016) and references therein.  $F_{\text{var}}$  is given by the square root of the excess variance normalised by the mean count rate. This provides the scale of the variability. Although  $F_{\text{var}}$  is given in the catalogue, it is not used to flag sources as variable.

The  $\chi^2$  statistic can be applied to binned data sets where the observed number of counts deviates from expectation approximately following a Gaussian distribution. Cash (1979) showed that when the number of counts per bin falls below  $\sim 10$ – $20$ , the  $\chi^2$  statistic becomes inaccurate. This justifies the requirements for the light curve generation in *XMM-Newton*'s pipeline.

There are, however, some drawbacks to this technique. Faint, variable sources can easily go unnoticed, since their time series are not generated. Additionally, sources that are variable on timescales shorter than the duration of the time bins will not be flagged as variable, since one single data point with a very high count rate is likely to be disregarded by the  $\chi^2$  variability test.

A number of objects can display variability within a single *XMM-Newton* observation. These can be distinguished in multiple ways. In a sample containing stars, AGN and different compact object systems, the compact objects show a higher X-ray flux variation factor (Lin et al. 2012). Furthermore, the features in the light curves are usually specific to a particular type of object. Thus, a number of source classes have previously been detected and identified:

Type I X-ray bursts are outbursts lasting from seconds to minutes, emitted by low mass X-ray binaries (LMXB) with an accreting neutron star (NS; e.g. Parikh et al. 2013). One subclass of High mass X-ray binaries (HMXB) called supergiant fast X-ray transients (SFXT) are HMXB with a neutron star accreting the winds of its companion. These sources can present multiple X-ray flares that are visible for a few hours (see Sidoli 2013, for a review). Cool late-type stars can show flares, due to magnetic reconnection. These flares can last from minutes to hours and show an exponential rise and decay (e.g. Pye et al. 2015; Favata et al. 2005; Imanishi et al. 2003). Cataclysmic variables (CVs) show modulated light curves with periods between  $\sim 0.2$  and  $\sim 1.5$  h (e.g. Kuulkers et al. 2003; Southworth et al. 2011).

While some sources from the list above have been previously detected with *XMM-Newton*, a more sensitive algorithm could detect a higher number; improving population statistics, and increasing the chance of detection.

We also aim to detect rarer transients such as short gamma ray bursts (SGRBs), the electromagnetic counterpart to NS mergers, as well as long gamma ray bursts (LGRBs), believed to be emitted from the jets formed during the death of massive stars in core-collapse supernovae. Whilst GRBs generally have spectra that peak in the gamma-rays, the spectra are broadband and can also be observed in X-rays. For distant GRBs, the peak luminosity will be redshifted into the *XMM-Newton* energy range. If a GRB with a small angle between the jet and the line of sight is within the field of view (FoV) of an *XMM-Newton* observation, it could be detected as a very faint, short transient. A handful of distant LGRBs has been detected by the *Swift* mission (Salvaterra 2015, and references therein).

Short X-ray bursts may accompany fast radio bursts (FRBs) – the rare, extragalactic, millisecond radio transients. Such a con-

nection is implied by the behaviour of Galactic magnetar SGR 1935+2154. That source is capable of emitting short FRB-like radio bursts with fluences as high as a MJy ms (Bochenek et al. 2020; Scholz 2020). These radio bursts are accompanied by bright, short X-ray bursts (Mereghetti et al. 2020; Tavani et al. 2020; Zhang et al. 2020; Ridnaia et al. 2020). Extragalactic FRBs may thus potentially also be accompanied by a transient X-ray counterpart (e.g. Kaspi & Beloborodov 2017; Metzger et al. 2019). Previous follow-up efforts and archival data searches so far have not found convincing X-ray counterparts to extragalactic FRBs (Scholz et al. 2016; Bhandari et al. 2018; DeLaunay et al. 2016). More observations, more sensitive instruments, or perhaps more sensitive algorithms applied to existing instruments could offer the step needed to detect a counterpart or to set more stringent limits on their existence.

Variability is thus a feature that can be used as a diagnosis tool to shed light upon the nature of a source and potentially allow the identification of new and rare compact objects. Hence, there is a need for new algorithms that are computationally cheaper than the generation of individual light curves that can be applied to faint sources.

Previous attempts to detect faint transients in *XMM-Newton* data include the EXTraS project<sup>2</sup>, which searched for transients and periodic variability (De Luca et al. 2015; Novara et al. 2020). The transient detection method employed there combines a source search on short time intervals with an imaging technique. The time intervals are defined through a modified Bayesian block algorithm (Scargle 1998; Scargle et al. 2013) that finds count-rate changes in the EPIC observations and runs a source detection algorithm on the time-resolved images.

In this paper, we present EXOD, a new algorithm that aims to detect short, faint transients by computing the variability of the whole FoV, then by detecting variable sources through an imaging technique. This can be used to promptly spot interesting sources that vary on timescales as short as 3 s. Due to the difference in the methodologies between EXTraS and EXOD (see Sect. 3), particularly in the way the time windows are defined and the stage at which the source detection algorithm is run, we expect to find different sources.

The remainder of this paper is organised as follows: Sect. 2 explains the data we used and how it was reduced. Section 3 details the EXOD algorithm and the variability detection technique as well as the comparison with other variability tests. Section 4 discusses the results and reports the detection of new variable sources. In Sect. 5, we discuss our work and potential applications and we give our conclusions in Sect. 6.

## 2. Data

### 2.1. Observation mode

We analysed the set of EPIC-pn observations contained in the 3XMM-DR8 catalogue performed in the full frame mode. The full frame mode was chosen for the following reasons: (1) The FoV covers a large surface of  $27.2' \times 26.2'$  and thus contains a higher number of sources than other EPIC-pn observation modes. (2) A time resolution of 73.3 ms allows the study of short timescale variability. (3) EPIC-pn receives about twice the number of photons that fall on EPIC-MOS (Strüder et al. 2001a; Turner et al. 2001). These properties are key to finding new variable, faint sources. We analysed a total number of 5751 observations with these characteristics.

<sup>2</sup> <http://www.extras-fp7.eu/index.php>

## 2.2. Filtering observations

We filtered the EPIC FITS pn IMAGING mode event list (PIEVLI) files using the Science Analysis System (SAS) version 16.1.0<sup>3</sup> with the following criteria:

We limited the energy range to 0.5–12 keV. We filtered out the 0.2–0.5 keV energy range to avoid the spurious detections caused by low-energy noise (Watson et al. 2009).

In order to ignore high background count rates from soft proton flares we removed periods in which the 10–12 keV count rate exceeds a certain threshold. We tested the *XMM-Newton* Science Analysis System (SAS) task `bkgoprtrrate` for this. It computes the rate maximising the signal to noise ratio for a given source. However, we found that the rate computed by this function is not optimal to detect faint variable sources. Although this rate increases the number of counts per source, the sources present an extrinsic variability due to the proton flares.

Figure 1 is an example, showing the background count rate in observation 0112370801. One can see that the `bkgoprtrrate` optimal background rate (green solid line) is too high to filter out several background flares. Choosing this rate would thus imply keeping a contaminated observation, affecting the variability detection. The count rate recommended by the SAS for filtering EPIC-pn observations, 0.4 cts s<sup>-1</sup> (pink dotted line) filters all the high background periods of this observation. However, this value could filter too many photons in the case of an observation with a higher low background rate or a source showing an extremely bright outburst. We thus chose a threshold of 0.5 cts s<sup>-1</sup> (orange dashed line). This threshold filters most of the background flares that are also filtered with 0.4 cts s<sup>-1</sup>. The steepness of the light curve at the beginning and the end of the background flares implies that the difference of the good time intervals between count rates of 0.4 and 0.5 will be small. The threshold of 0.5 cts s<sup>-1</sup> was adopted after testing it on ~700 observations. Whilst a slightly refined value may be better for some observations, this value was found to be appropriate for the large majority of the observations studied. We used the extra precaution of removing all the time windows (see Sect. 3.1.1 for definition) where a background flare starts or ends to avoid contamination.

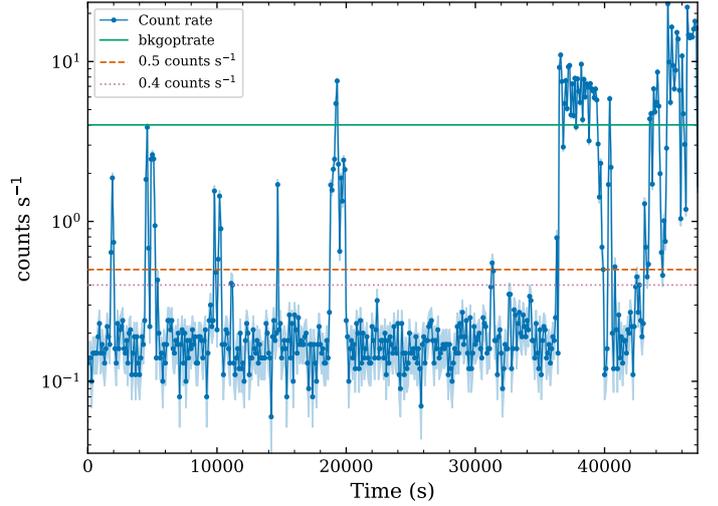
We used the SAS task `tabgtigen` with a count rate of 0.5 cts s<sup>-1</sup> and time bins of 100 s to generate the good time interval (GTI) file and filtered the bad time intervals from the observations. Additionally, we applied the standard filters `#XMMEA_EP` and `PATTERN<=4`.

## 3. Method

### 3.1. The algorithm

Since the light curve generation is necessary for the variability computation of the sources in *XMM-Newton*'s pipeline, only the variability of the brightest sources is studied. In order to detect fast, faint transients coming from fainter sources within *XMM-Newton* observations, the whole FoV should be explored.

We developed EXOD, the EPIC-pn *XMM-Newton* Outburst Detector, an algorithm that detects fast and faint transients by computing the variability of every single EPIC-pn pixel from each observation, as we present below. The code is publicly available online<sup>4</sup>.



**Fig. 1.** Count rate of observation 0112370801 between 10–12 keV as a function of time since the beginning of the observation. The vertical axis is in logarithmic scale and we use a time binning of 100 s. Green solid line: optimal rate as computed by `bkgoprtrrate`, in this case ~4 cts s<sup>-1</sup>. Orange dashed line: 0.5 cts s<sup>-1</sup>. Pink dotted line: 0.4 cts s<sup>-1</sup>.

The algorithm computes the variability as explained in the following section using the previously generated filtered events file and GTI file. The required input parameters used for the variability detection can be found in Table 1.

#### 3.1.1. Variability computation

The variability computation can be divided in the steps depicted in Fig. 2 and explained next: (1) First, we extract the time and pixel of arrival of every event detected during the observation from the events file. (2) Secondly, photons detected in a 3 × 3 square around each pixel are added together to increase the signal-to-noise ratio and to reduce stochastic variability. (3) Next we bin the events per pixel into time windows ( $TW$ ), whose duration we give as an input. (4) We then extract the data from the GTI file. (5) We compute the good time ratio of each time window, that is, the time belonging to the GTI of the considered time window divided by the duration of the time window. (6) Next we divide the number of counts per time window by the good time ratio. Only those time windows with the good time ratio above a critical value,  $r_{GT}$ , will be considered. This will normalise the number of photons that have been detected during a time window that has been shortened by the bad time periods. We note that the  $r_{GT}$  default value is 1 (see Table 1). This removes every time window that has been partially or totally contaminated by background flares, avoiding detecting these flares as variable sources. (7) Finally, the variability  $\mathcal{V}$  of each pixel is given by:

$$\mathcal{V} = \begin{cases} \max(C_{\max} - \tilde{C}, |C_{\min} - \tilde{C}|) / \tilde{C} & \text{if } \tilde{C} \neq 0 \\ C_{\max} & \text{if } \tilde{C} = 0 \end{cases} \quad (1)$$

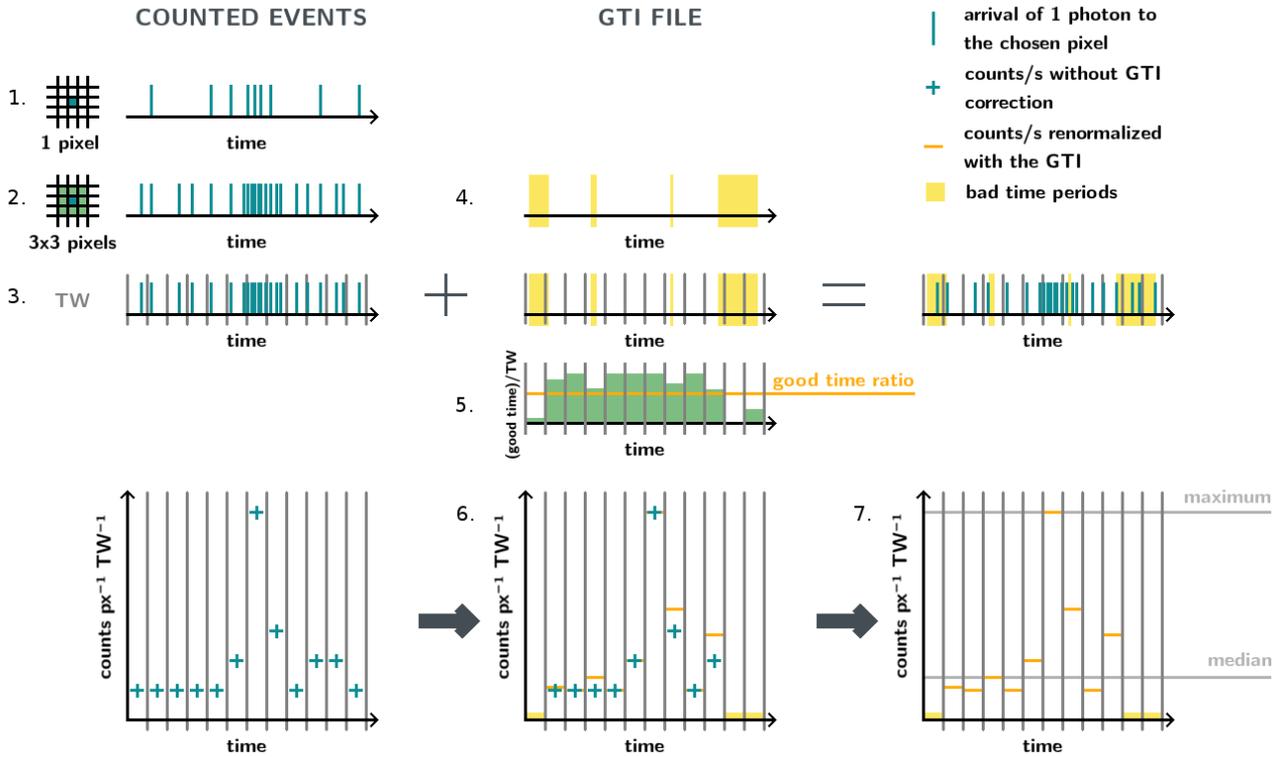
where  $C_{\max}$  and  $C_{\min}$  are, respectively, the maximal and minimal number of counts per time window for that pixel and  $\tilde{C}$  is the median number of counts over the time windows for the pixel. The expression  $C_{\max} - \tilde{C}$  targets sources presenting outbursts, while  $|C_{\min} - \tilde{C}|$  points to those sources with a period of lower flux. Considering the maximum between the two allows the detection of a wider variety of phenomena. The division by

<sup>3</sup> [https://xmm-tools.cosmos.esa.int/external/xmm\\_user\\_support/documentation/sas\\_usg/USG/](https://xmm-tools.cosmos.esa.int/external/xmm_user_support/documentation/sas_usg/USG/)

<sup>4</sup> <https://github.com/InesPM/EXOD>

**Table 1.** Input parameters of the variability computation.

Symbol	Parameter	Accepted values	Default value	Function
$DL$	Detection level	$\mathcal{R}^+$	10	Level above which an area is considered as being variable. A lower value allows the detection of more sources, but a higher proportion of the detected sources will be non-variable.
$TW$	Time window	$\mathcal{R}^+$	100 s	Duration of the time windows. It is also the minimal timescale of the variability that can be explored.
$b$	Box size	[3,5,...,63]	5	Length in pixels of the detection box, limited by the size of the CCDs. It determines the extension of the variable area that we want to detect.
$r_{GT}$	Good time ratio	[0;1]	1.0	Critical (good time)/ $TW$ above which the time window will be taken into account. Choosing 1.0 will remove all time windows that have been partially or entirely contaminated by soft proton flares.



**Fig. 2.** Diagram of the variability computation with the different stages of the algorithm. The vertical blue lines indicate the arrival time of the photons during the observation. The vertical gray lines represent the time bins. The blue crosses indicate the counts per time window per pixel without filtering for the GTI. The yellow shaded regions indicate the time periods contaminated by soft proton flares. The green rectangles show the good time over the duration of the time window for each time window. The orange horizontal lines show the counts per time window per pixel after renormalising with the GTI, and the yellow horizontal lines show the time windows that will not be taken into account because they are below the good time ratio. Finally, the gray horizontal lines indicate the maximum  $C_{\max}$  and the median  $\bar{C}$  counts/s that will be used to compute the variability.

the median  $\bar{C}$  gives the variability relative to the flux. The algorithmic notation is given in Appendix C.

### 3.1.2. Variable source detection

The variable areas are detected with the sliding box technique, where a box of size  $|b|^2$  pixels will move through all the pixels of the observation, with  $b$  representing an odd number.

This procedure has been broadly used in X-ray observatories, for instance the *Einstein* observatory (Gioia et al. 1990),

ROSAT<sup>5</sup>, and currently in *Chandra* (Calderwood et al. 2001) and *XMM-Newton* (Watson et al. 2009).

We first compute the median variability value  $\bar{V}$  (Eq. (1)) for all detector pixels. We next calculate the sum of the variability in a box of size  $|b|^2$  to obtain  $V_{\text{box}}$ . The central position of the box is then shifted to the contiguous pixel and the process is repeated until  $V_{\text{box}}$  has been calculated for every available position within the detector (based on Watson et al. 2009). The variability of a

<sup>5</sup> <https://heasarc.gsfc.nasa.gov/wgacat/>

**Table 2.** Parameters used for the variability computation, with good time ratio  $r_{GT} = 1.0$ .

$TW$ (s)	3			10			30			100		
$DL$	5	6	7	6	7	8	7	8	9	8	9	10
$b$ (pixels)	3						5					

**Notes.** For each time window ( $TW$ ), the three respective detection levels ( $DL$ ) were used, whereas the two given box sizes  $b$  were used for all the  $TW$ - $DL$  combinations.

box  $\mathcal{V}_{\text{box}}$  centered on the pixels  $x,y$  is given by the following equation:

$$\mathcal{V}_{\text{box}(x,y)} = \sum_{i=x-(b-1)/2}^{x+(b-1)/2} \sum_{j=y-(b-1)/2}^{y+(b-1)/2} \mathcal{V}_{i,j}, \quad (2)$$

Where  $b$  is the length of the box in raw pixel units.  $\mathcal{V}_{\text{box}}$  will not be computed at the border of the CCDs, since the size of the box would be smaller than  $|b|^2$ .

If for a certain box  $\mathcal{V}_{\text{box}}$  is above a chosen threshold, we consider the pixels contained in the box as variable. The value of this threshold is given by the following expression:

$$\mathcal{V}_{\text{box}} > DL \times |b|^2 \times \tilde{\mathcal{V}}, \quad (3)$$

where  $DL$  is the detection level. When two consecutive boxes are variable, the pixels of both boxes are joined into a single variable area.

The variable sources are located at the centre of the variable areas. The position of the (X,Y) coordinates in raw pixels is the mean value of the position of the pixels belonging to the variable areas. The raw pixel position and the number of the CCD of the variable sources are given as an output, as well as the distance from the centre to the outermost pixel belonging to the variable area, in pixels. We do not consider sources detected in CCD = 4, RAWX = 12; CCD = 5, RAWX = 12; CCD = 10, RAWX = 28 since these pixels are known to be damaged (Strüder et al. 2001b) and produce a high number of spurious detections.

### 3.2. Detection parameters

We applied the algorithm to the observations defined in Sect. 2 with different time windows, detection levels, and box sizes. We thus empirically determined the optimal parameters for finding most variable sources, while minimising spurious detections. The different values of the time windows check for variability on different timescales. We chose  $TW = 3, 10, 30, 100$  s to optimise time variability with computation time.

We performed these variability tests with the parameters given in Table 2 and  $r_{GT} = 1.0$  on a subset of 2284 observations. These are the observations that were included in the incremental releases of the 3XMM-DR7 and 3XMM-DR8 catalogs. We used the optimal parameters obtained to analyse the light curves of the sources detected in the remaining observations.

### 3.3. Comparison with other variability tests

In order to determine the robustness of the algorithm, we needed a comparison with the existing variability tests. To do so, we determined the fraction of false positives as well as the fraction of false negatives.

To measure the fraction of false negatives, we performed a positional cross-match of the EXOD sources with the 3XMM-DR8 sources flagged as variable. For this we used the catalogue manipulation tool TOPCAT (Taylor 2005).

To measure the fraction of false positives, we first recreated for reference the standard variability measure of the EXOD-detected sources. We generated the light curve of the detected sources with the SAS function `evselect` and subsequently computed the probability of constancy with the  $\chi^2$  and Kolmogorov-Smirnov tests,  $P(\chi^2)$  and  $P(KS)$  respectively, with the FTOOLS XRONOS function `lcstats`. Comparisons with other test are found in Pastor-Marazuela (2018)<sup>6</sup>. We used the HEASOFT version 6.22.1 (Blackburn 1995).

The  $\chi^2$  test can only be applied to data with a Gaussian distribution, which occurs when the number of counts per bin is  $\geq 18$ . As we generated light curves regardless of the number of counts, not every light curve satisfies this criterion. We thus additionally apply the KS test, which makes no assumptions about the binning of the data and can be safely applied to sources with low counts<sup>7</sup>.

If  $P(\chi^2) \leq 10^{-4}$  or  $P(KS) \leq 10^{-4}$ , we consider the source to be variable.  $P(\chi^2) \leq 10^{-4}$  gives a  $>4\sigma$  confidence level on the variability detection with the  $\chi^2$  test. We define the robustness of the algorithm as the ratio of sources detected with EXOD that are variable according to at least one of the variability tests  $\chi^2$  or KS.

The light curves were extracted automatically by giving the position in raw pixel units of the detected variable source, then using a set of SAS tasks described below that optimise the position and radius of the source. We used the  $TW$  duration as the bin size of the light curves. For these tasks, the SAS summary file and the EPIC FITS global background time series (FBK-TSR) are needed additionally to the filtered observation and the GTI file.

Since we noted that extended sources, bright sources, or out-of-time (OOT) events can trigger spurious detections of variability within an observation, we only applied the automatic light curve generation when the number of detections per observation was less than six.

We determined this six-source threshold empirically, by noting that less than  $\sim 2.6\%$  of the observations have more than five variability detections, and that these are usually spurious detections. However, we visually inspected the variability plots (such as the one shown in Fig. B.1) of the fields of view with more than five detections in order to avoid missing interesting candidates.

**Source region.** From the raw pixel and CCD of the detected variable sources, we can determine the position in sky pixel space with the SAS task `ecoordconv`. The position and the radius of the source are corrected with the function `eregionanalyse` with `backval=0.1`. This function returns a circle containing 90% of the energy of the source.

It is known that the shape of the point spread function (PSF) in the EPIC-pn detector depends on its angle with respect to the centre of the detector (Read et al. 2011). Whereas a centered (on-axis) source is well approximated by a circular region, a non-centered (off-axis) source is elongated due to off-axis aberration (astigmatism), giving a larger, energy-dependent radius. Although elliptical regions increase the signal to noise

<sup>6</sup> Searching for fast transients in *XMM-Newton* data. Master thesis: <https://www.uva.nl/en/profile/p/a/i.pastormarazuela/i.pastor-marazuela.html>

<sup>7</sup> Kirkman, T.W. (1996) Statistics to Use. <http://www.physics.csbsju.edu/stats/>

ratio of the light curves, we chose circular extraction regions since this option is a sufficiently good approximation, especially when dealing with faint sources for which only a small number of pixels will be above the background level. We also find it to be more reliable in such an extensive study to automatically extract the counts of all the sources and find nearby background extraction regions. The selection of elliptical regions becomes of greater importance when performing spectral analyses.

**Background region.** To subtract the background from the source light curve, we use background regions determined with the SAS task `ebkgreg`. This task searches the optimal background region following only geometrical criteria. To avoid selecting a region containing sources, we extract the background from a filtered events file where we have removed the sources. We obtain the positions of all the sources of the observation from the FBKTSR file. We choose the same radius as for the source region.

### 3.4. M31

There are some fields observed by *XMM-Newton* where we expect to detect more than five variable sources. In this work we have decided to focus on one specific field, M31. Because of its crowded field, M31 required analysis that was more manual than described above.

M31 is located at  $\sim 0.78$  Mpc<sup>8</sup>, and is the nearest spiral galaxy to the Milky Way. With an angular size of  $\sim 200' \times 80'$ , EPIC-pn's FoV covers its central region. This makes the study of M31 very interesting, since one observation contains a high number of extragalactic sources.

There is a total of 48 *XMM-Newton* EPIC-pn observations of M31 that have not been strongly polluted by high background flares. The high density of sources required selecting the extraction regions manually. In most observations, the number of detected sources exceeds the limit of five. We have thus made an exception for this limit for M31, where we expect to detect a higher number of variable sources.

### 3.5. Fast radio bursts

We know EXOD is sensitive to bursts lasting up to 100s (see, Sect. 4.3). There are now both theories and observations linking FRBs with magnetars; and the latter are known to show X-ray flares lasting up to a few seconds. FRB theories such as those put forward by Beloborodov (2017) and Metzger et al. (2017) require young magnetars as the power source. Meanwhile, observations of SGR 1935+2154 show that Galactic magnetars can emit energetic flashes that show as both FRB-like radio bursts (Bochenek et al. 2020) and X-ray bursts (Zhang et al. 2020), simultaneously.

We thus tested whether EXOD could again find variability missed by standard tools. We reanalysed observations 0792382801 and 0792382901 of FRB 121102, the first known repeating FRB (Scholz et al. 2017). FRB 121102 was highly active in radio during these *XMM-Newton* observations (Chatterjee et al. 2017; Law et al. 2017), but Scholz et al. (2017) found no transient X-ray counterparts. Using our same opti-

mal parameters defined in Sect. 3.2, we also found no sign of variability.

The energy upper limit estimates that Scholz et al. (2017) set on the X-ray emission of  $10^{45} - 10^{47}$  erg at a distance of 972 Mpc (Tendulkar et al. 2017), remains several orders of magnitude lower than that of GRBs. The detection of an FRB afterglow with *XMM-Newton* could be possible only if its host galaxy was located at a lower redshift. Following up on non-repeating low-DM FRBs, such as FRB 110214 at  $169 \text{ pc cm}^{-3}$  (Petroff et al. 2019), could help to achieve a detection, especially if possible near real-time (cf. Maan & van Leeuwen 2017).

## 4. Results

### 4.1. Detected sources

In this section, we present the sources EXOD detected in our set of observations. This includes the detected variable sources for which the light curves were automatically generated (less than six detections per observation) as well as those with more than five detections that we determined as non-spurious through a visual inspection.

There were 221 observations with more than five variability detections, with only 20 containing non-spurious detections. Among these non-spurious detections, we generated the light curve of 83 sources. We include these sources in the following analyses.

#### 4.1.1. Variability

In Fig. 3, we show the results obtained after applying EXOD as described in the previous sections. By looking at the test parameters, one can see that the number of detected sources increases with decreasing detection level  $DL$ . Decreasing  $DL$  also increases the fraction of non-variable sources according to the  $\chi^2$  and KS tests, although this is accompanied by an increase of the net number of variable sources. In most cases a  $3 \times 3$  pixel box results in a slightly higher fraction of variable sources than a  $5 \times 5$  box.

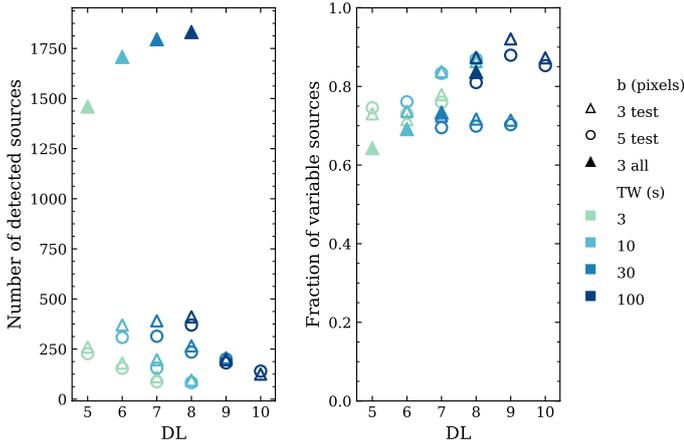
In order to detect the highest possible number of previously unknown variable sources, we chose a low detection level with a box size of  $3 \times 3$  pixels to analyse the full data set, accepting the accompanying increase in false positives. These final optimal parameters are given in Table 3, together with the number of sources detected among the 5751 observations. The results are plotted in Fig. 3 as filled symbols, with the same color code as the test observations.

From these results, we can determine that the robustness of the algorithm as defined in Sect. 3.3 is between  $\sim 60$  and  $95\%$  and depends on the chosen parameters. For the optimal parameters,  $\sim 84\%$  of the sources are deemed to be variable using the  $\chi^2$  and KS tests for a time window  $TW = 100$  s. This percentage decreases when  $TW$  decreases, and reaches  $\sim 64\%$  for  $TW = 3$  s. EXOD is thus more robust for longer time windows.

The remaining fraction of sources are usually detected in short or highly polluted observations, notably when the net exposure time is below 5000 s, and can be clearly identified. We thus advise visually inspecting the results when using EXOD for such short exposure observations.

Although we perform a large number of independent trials (especially the number of TWs, in four trials  $\times$  number of detection boxes), statistical fluctuations of sources that are not intrinsically variable are not a significant cause of false positives, given the high minimum detection levels we require.

<sup>8</sup> M31 properties obtained from NASA/IPAC Extragalactic Database (NED): <https://ned.ipac.caltech.edu/byname?objname=M31>. The given distance is an average of the published distances.



**Fig. 3.** Properties of the sources detected by EXOD. Test observations with the detection parameters listed in Table 2: empty markers. All observations with the parameters in Table 3: filled markers. *Left*: number of detected sources as a function of  $DL$ . *Right*: fraction of sources confirmed to be variable through the  $\chi^2$  and KS tests among the detected ones as a function of  $DL$ . The colors correspond to different  $TW$ : light green for  $TW = 3$  s, light blue for  $TW = 10$  s, medium blue for  $TW = 30$  s and dark blue for  $TW = 100$  s. The shapes represent the box size: triangles for  $b = 3 \times 3$  pixels and circles for  $b = 5 \times 5$  pixels.

When applying the optimal parameters to the whole set of observations, we find the fraction of non-variable sources to be higher than that in the test observations. We are investigating the cause for this but we suspect the later date of the DR7 and DR8 observations to play a role, as the average duration of the later observations has increased. Given the longer duration, the number of flares that can be present in an observation increases. In a shorter observation, the lower number of counts may give rise to a higher number of spurious detections due to the aforementioned outliers of the Poissonian fluctuations.

The sum of the sources detected with the optimal parameters is 6649. Since some of the sources are detected with different time windows, we performed an auto-correlation on the position of the sources with TOPCAT. We searched for internal matches in a radius of  $15''$  around the location of each source. This radius corresponds to the Half Energy Width (HEW) of EPIC-pn at 1.5 keV (Strüder et al. 2001a). After removing the duplicated selections, we obtain a grouped list with a total number of 2961 sources detected among the 5751 observations.

The *XMM-Newton* pipeline finds slightly more variable sources because it searches longer timescales than EXOD. Sources varying more slowly than our longest timescale of 100 s might not be detected.

Although EXOD may miss some variables found by the standard pipeline, it does find variable sources this pipeline missed. We investigate these in Sect. 4.3.

#### 4.1.2. Cross-correlation with other catalogues

To determine the nature of the EXOD-detected sources, we queried them by position in the SIMBAD database<sup>9</sup> (Wenger et al. 2000), using *astroquery*. This returns the name, location and source type of the nearest catalogued source. We allowed a  $15''$  radius around the detection coordinates. We then placed the sources in one of the categories listed in Table 4. These include six physical classes, one group of known extended sources, and

**Table 3.** Optimal parameters and detected sources.

Property	Value			
$TW$ (s)	3	10	30	100
$DL$	5	6	7	8
$b$ (pixels)	3	3	3	3
Detected sources	1459	1707	1795	1831
Variable sources (%)	64.3	69.2	73.4	83.7
Sum	6792	Total	2961	

**Notes.** Parameters used to generate the light curves of all the observations, number of sources detected and percentage of variable sources according to the  $\chi^2$  and KS tests.  $r_{GT} = 1.0$ . Bottom line: sum of the detected sources and total number of sources after auto-correlation.

two categories of unknowns: the “no identification” class for previously detected sources whose type is not known; and the “without counterpart” class where no detection has previously been made at all. If two sources were identified with the same SIMBAD entry, only one is counted. For that reason, the number of classified sources in Table 2 (2961) differs from the total number of sources in Table 4 (2907).

In a preliminary cross-correlation with SIMBAD, we found  $\sim 1200$  classified as “without counterpart”. A visual inspection of these sources revealed that most of these belonged to one of the following categories: (1) Out of Time events, (2) detections in the PSF of a bright source, (3) detections in an extended source and (4) hot pixels. We thus cross-matched these with the 3XMM-DR8 catalogue with TOPCAT in a  $15''$  radius around the detection position and the radius of the sources given in the catalogue to take into account their extension. The sources matching a 3XMM-DR8 source were added to the “no identification” (point-like sources) or the “extended sources” class ( $>6''$  radius). This reduced the number of “without counterpart” to  $\sim 400$ . A visual inspection of the remaining “without counterpart” sources showed that these were spurious detections.

Given our search criteria, we could have expected to find new sources through the variability search if they were faint and short enough to be drowned in the background noise and would thus not be included in the regular XMM catalogue. Whilst no completely new sources were identified here, further modifications to the detection parameters could still reveal such objects.

In Fig. 4, we show the result of the cross-correlation with SIMBAD for the four sets of parameters as thin bars. The cross-correlation of the grouped sources with SIMBAD and 3XMM-DR8 is shown as thick bars with a solid black contour. The cross-correlation only with SIMBAD is plotted as dashed black contours for comparison.

If we remove the number of spurious detections from our set of detected sources, we obtain a net count of 2,536 variable sources detected with EXOD.

We can compare this number to the number of variable sources in the catalogue. 3XMM-DR8 contains 775 153 sources. 438 342 of these sources were detected in observations that we have analysed. 102 498 of these sources have generated time series. Finally, 3418 sources are catalogued as variable, compared to the net 2536 sources in this work.

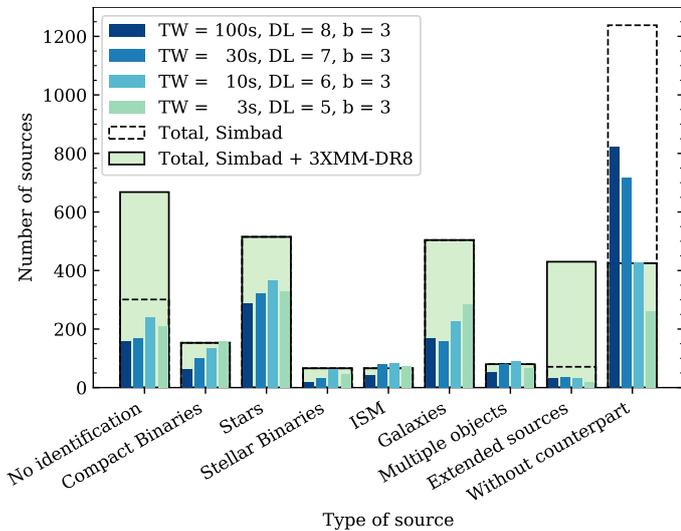
However, the number of sources that are classified as variable in 3XMM-DR8 and EXOD simultaneously is 777. This corresponds to a false negative rate of 77.3%. The high rate of false negatives was expected, since we are targeting a very specific kind of short-term variability, and sources varying on longer timescales are not detected. Nevertheless, the light curve

<sup>9</sup> SIMBAD: <http://simbad.u-strasbg.fr/simbad/>

**Table 4.** Source category classification.

Category	Description	Number of sources
Compact binaries	Cataclysmic variables, X-ray binaries, ULXs, novae.	153
Stars	Stars, pulsars	515
Stellar binaries	Binary stellar systems	66
ISM	Interstellar medium	66
Galaxies	Galaxies, AGNs, QSOs	504
Multiple objects	Galaxy clusters, groups of galaxies, stellar clusters or associations.	80
Extended sources	The angular distance $d$ to the catalogued source satisfies $10'' \leq d \leq 15''$ and it belongs to the category ISM, galaxies or multiple objects.	430
No identification	Previously known source of unknown type.	668
Without counterpart	No association has been found within $15''$ .	425

**Notes.** Naming of the categories used in Fig. 4, with a description of the objects included in each category and the number of sources detected for each.



**Fig. 4.** Classification of the sources detected by EXOD in source types. The thin bars show the SIMBAD association of the sources detected, from darker to lighter, with 100 s, 30 s, 10 s and 3 s respectively. The thick bars with a black dashed contour show the SIMBAD association of the grouped sources. The thick light green bars with a solid black contour show the SIMBAD association with the 3XMM-DR8 correction.

of 688 (27.1%) of the EXOD variable sources were not generated in 3XMM-DR8, so there was no previous available information about their variability. Lastly, Table 3 shows that EXOD detects a high number of variable sources according to the  $\chi^2$  or the KS tests, indicating that we are sensitive to a variability that the tests applied in *XMM-Newton*'s pipeline are not always adapted to detect.

#### 4.1.3. Detected variability

Although EXOD was specifically designed to detect faint, short outbursts, the inspection of the detected sources showed that it is able to detect a large diversity of variable phenomena. In Fig. 5 we show a selection of six light curves from different variable sources. These light curves were extracted automatically. We also plot  $\tilde{C}$  and  $C_{\max}$  to compare the light curves to the value of the variability obtained with EXOD.

These sources differ in their variability timescales (short and long flares), periodicity or aperiodicity, and speed of decline and

rise. They also span a range of physical classes. In Fig. 5, we present the following sources: in (1), an ultra compact binary showing periodic oscillations whose maximal values seem to increase during the observation (Steehns et al. 2006). In (2), a high proper motion star showing a flare during the observation (Kirkpatrick et al. 2016). In (3), we can see flares from a supergiant fast X-ray transient, aperiodic with varying maxima (Sidoli et al. 2012). In (4), an eclipsing binary of beta Lyr type (Bruch 2017). In (5), we see a magnetic cataclysmic variable whose periodic variability is due to variation in the opacity of the accretion curtain that comes into view as the system rotates (Mason 1985; Fuchs et al. 2016). In (6), an LMXB presents first an eclipse during which the flux decreases, then a type-I X-ray burst (in't Zand et al. 2019). We note that we refiltered the last observation since the  $0.5 \text{ cts s}^{-1}$  tabgtigen threshold cut a part of the flare.

#### 4.2. Computational performance of the algorithm

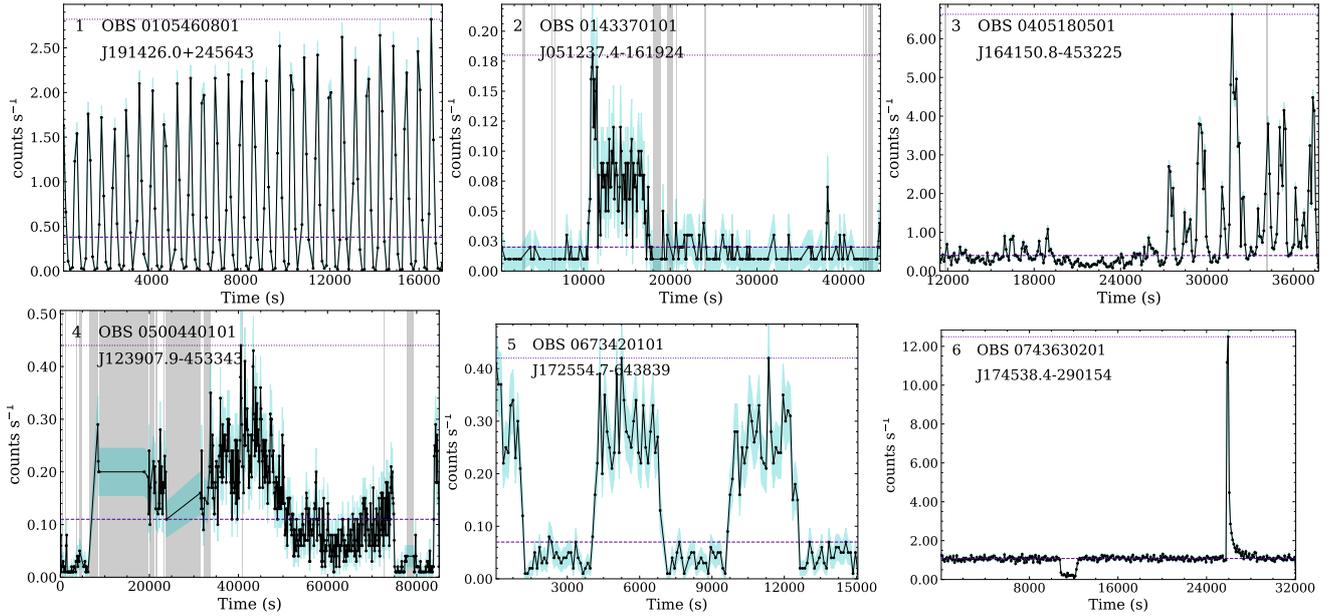
We measured the absolute computation time  $t_{\text{comp}}$  taken by EXOD to analyse each observation. Additionally, we measured the computation time for the generation of a single light curve for each detected variable source. Experiments were conducted on a virtual machine with 38 cores at 2.3 GHz and 320 GB of RAM. We fitted these values to a power law with the expression given in Eq. (4) as a function of the total duration of the GTI,  $t_{\text{gti}}$ . The result is plotted in Fig. 6,

$$t_{\text{comp}} = a + b \times t_{\text{gti}}^c, \quad (4)$$

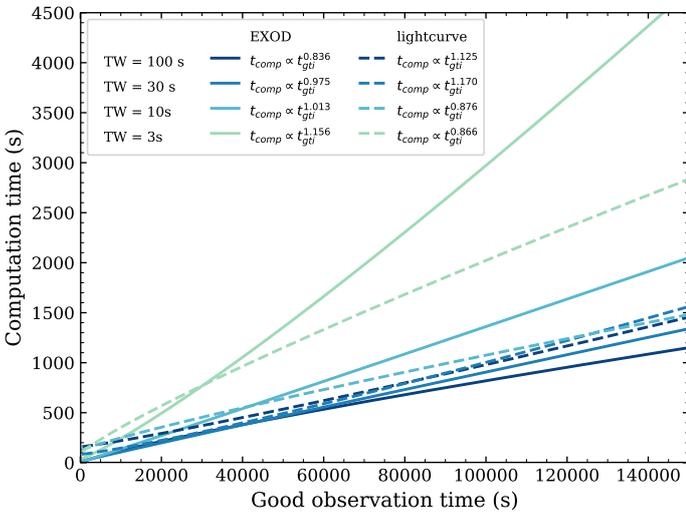
where  $a$ ,  $b$ , and  $c$  are the parameters determined by the fit. We computed these parameters by considering the computation time for all of the analysed observations with time windows of 3 s, 10 s, 30 s and 100 s. These values are shown in Table 5.

The computation time of the algorithm includes the variability computation and the source detection time, whereas the computation time of the light curve includes some SAS and FTOOLS tasks that are required for the light curve generation: cifbuild, eregionanalyse, ebkgreg, evselect, epiclccorr, lcstats and lcurve.

The EXOD computation time increases for shorter  $TW$ , as expected given the increased number of data points. The light curve computation time is similar for 10, 30 and 100 s  $TW$ , and increases for  $TW = 3$  s. For 100 and 30 s  $TW$ , EXOD can compute the variability of the whole observation before even a single light curve is generated. While for a shorter  $TW$ , the computation is slower than generating a single light curve, EXOD is



**Fig. 5.** Example of light curves of sources with different types of variability detected by EXOD. The light curves are plotted in black with cyan shaded regions representing the  $1\sigma$  error bars. In each plot we give the OBSID where it was detected and the 3XMM name of the source. The dashed purple line represents  $\bar{C}$ , the median number of counts. The dotted purple line represents  $C_{\max}$ , the maximal number of counts. The gray vertical shaded regions represent the bad time intervals. (1) RX J1914.4+2456, ultra compact binary showing periodic oscillations (Steehns et al. 2006). (2) WISEA J051237.57-161925.2, high proper motion star showing a stellar flare (Kirkpatrick et al. 2016). (3) IGR J16418-4532, flares from a supergiant fast X-ray transient (Sidoli et al. 2012). (4) V\* V1129 Cen, eclipsing binary of beta Lyr type (Bruch 2017). (5) LSQ J172554.8-643839, magnetic cataclysmic binary showing periodic eclipses (Fuchs et al. 2016). (6) AX J1745.6-2901, LMXB showing a type I X-ray burst (in’t Zand et al. 2019).



**Fig. 6.** Computation time as a function of the observation time that has not been polluted by background flares (good time) fitted to a power law. The solid lines indicate the EXOD detector computation time. The dashed lines indicate the light curve computation time for one source. Blue (dark) is the computation time for a time window of 100 s, green (medium) for 30 s, orange (medium) for 10 s and pink (light) for 3s.

faster than light curve generation when more than one source is present in the FoV.

#### 4.3. Previously steady sources found to be variable

Since no completely new sources were found among the “unknown source” class (Sect. 4.1.2), we analysed the “no

**Table 5.** Computation time fit parameters.

TW(s)	EXOD			Light curve		
	$a$	$b$	$c$	$a$	$b$	$c$
100	0.151	0.054	0.836	158.204	0.002	1.125
30	12.614	0.012	0.975	83.171	0.001	1.170
10	5.210	0.012	1.013	121.960	0.040	0.876
3	36.822	0.005	1.156	89.109	0.091	0.866

**Notes.** Parameters of the fit to the computation time for the expression (4) for EXOD and the light curve generation.

identification” sources in more detail. We collected the sources not flagged as variable by *XMM-Newton*’s pipeline and visually inspected these. We removed the sources misidentified in the automatic procedure or affected by high background rates. Since the outburst selection was based on a visual inspection, we checked the  $P(\chi^2)$  and P(KS) of the pn light curves. When these probabilities of constancy were  $>10^{-4}$ , we also computed them for the MOS1 and MOS2 light curves, and only kept the sources for which the  $P(\chi^2)$  and P(KS) for pn, MOS1 or MOS2 were  $<10^{-4}$ . We ended up with a subset of 26 sources, whose light curves are presented in Fig. A.1. Among the M 31 observations for which we manually selected the light curve extraction regions, we found a further 9 sources previously detected but not flagged as variable, plotted in Fig. A.2. The properties of these sources are summarised in Table 6.

All the new transients were detected with  $TW = 100$  s. Some of these transients were also detected by shorter  $TW$ , but no new sources were detected only with shorter time windows. If the duration of the transient is  $\lesssim 100$  s, most of the photons will fall within the same time bin if  $TW = 100$  s, increasing the signal to

**Table 6.** New variable EXOD sources.

ID	Name	OBSID	RA (J2000)	Dec (J2000)	Burst dur. (s)	Var.	Comments
1	J173046.7+521846	0021750201	17:30:46.7	+52:18:46		PN	AGN candidate.
2	J160502.1+430401	0025740101	16:05:02.1	+43:04:01		PN	Short observation.
3	J191515.1+044348	0075140501	19:15:15.1	+04:43:48		PN	No SIMBAD object associated.
4	J083941.3+192901	0101440401	08:39:41.3	+19:29:01	3000	PN	Star in cluster.
5	J182929.4-092530	0135744801	18:29:29.4	-09:25:30		PN	Stellar flare. Off-axis detection.
6	J015709.1+373739	0149780101	01:57:09.1	+37:37:39	800	PN	Stellar flare.
7	J070509.8-112940	0201390201	07:05:09.8	-11:29:40		PN	Short observation. Star.
8	J174610.8-290021	0202670701	17:46:10.8	-29:00:21	100	PN	Type I X-ray burst candidate.
9	J010909.2+132337	0203280301	01:09:09.2	+13:23:37		PN	AGN candidate.
10	J015727.2-004041	0303110101	01:57:27.2	-00:40:41	200	PN	Short observation, off-axis detection.
11	J092927.6+504810	0556210401	09:29:27.6	+50:48:10		PN	Quasar candidate.
12	J233504.9-534751	0604870332	23:35:04.9	-53:47:51		PN	AGN.
13	J183658.4-072119	0606420101	18:36:58.4	-07:21:19	3000	PN	Close to CCD gap. Star.
14	J090150.9-015815	0655340159	09:01:50.9	-01:58:15	600	PN	Short observation. Blue ctp.
15	J081907.9-384302	0655650201	08:19:07.9	-38:43:02	2500	PN	IR-optical ctp. Star.
16	J221448.2+002707	0673000136	22:14:48.2	+00:27:07	400	MOS	QSO.
17	J090629.9-000911	0725300150	09:06:29.9	-00:09:11		MOS	Short observation. Star.
18	J090335.4+013224	0725300157	09:03:35.4	+01:32:24		PN	Short observation. Seyfert 1.
19	J011552.2-003058	0747400134	01:15:52.2	-00:30:58	300	PN	No SIMBAD object associated.
20	J012517.1-001829	0747410134	01:25:17.1	-00:18:29	300	PN	QSO.
21	J012351.6+000831	0747410144	01:23:51.6	+00:08:31		MOS	Short observation, highly contaminated.
22	J014150.5+000754	0747430146	01:41:50.5	+00:07:54	400	PN	Optical-UV ctp.
23	J144506.2+685817	0763640601	14:45:06.2	+68:58:17	200	PN	Contaminated observation. Blue ctp.
24	J175131.6-401533	0763700301	17:51:31.6	-40:15:33	1000	PN	var_flag=True. Optical ctp. Star.
25	J113407.5+005223	0770380401	11:34:07.5	+00:52:23	200	PN	No SIMBAD ctp.
26	J022133.7-042346	0785101401	02:21:33.7	-04:23:46	1000	PN	NIR-blue ctp. Star.
M 31							
1	J004307.5+412019	0109270101	00:43:07.5	+41:20:19	400	PN	Type I X-ray burst. NS-LMXB.
2	J004242.5+411657	0405320701	00:42:42.5	+41:16:57		PN	
3	J004215.6+411720	0650560201	00:42:15.6	+41:17:20	300	PN	Type I X-ray burst. NS-LMXB.
4	J004252.4+411648	0650560501	00:42:52.4	+41:16:48		PN	
5	J004209.5+411745	0674210201	00:42:09.5	+41:17:45		PN	
6	J004210.9+411248	0674210201	00:42:10.9	+41:12:48	500	PN	Type I X-ray burst. NS-LMXB.
7	J004215.1+411234	0674210301	00:42:15.1	+41:12:34		PN	
8	J004212.1+411758	0727960401	00:42:12.1	+41:17:58	500	MOS	Type-I X-ray burst. NS-LMXB.
9	J004231.2+411938	0727960401	00:42:31.2	+41:19:38		PN	

**Notes.** Newly variable sources detected by EXOD with  $TW = 100$  s. Column 1 gives the source ID in this paper. Column 2 the name of the source in 3XMM-DR8. Column 3 the observation ID in which the source was detected, Cols. 4 and 5 the RA and Dec respectively. Column 6 the duration of the burst if there is one, Col. 7 in which instrument the light curve is variable (PN or MOS), and Col. 8 some comments, including the presence of a counterpart (ctp.). The top 30 are the general survey, the bottom 9 are in the direction of M 31.

noise ratio. With a lower  $TW$ , the number of photons per time bin is lower, and thus has lower chances to reach the detection level. We found some interesting candidates that had been detected with  $TW = 3$  s, but a closer look at these sources showed that they were damaged pixels. In particular, the pixel  $RAWX = 1$ ,  $RAWY = 72$ ,  $CCD = 8$  and the surrounding pixels were detected in  $\sim 20$  observations as what looked like very short, bright outbursts. In some cases, these had been included in 3XMM-DR8 as real detections. The problem with the damaged pixels being detected as sources is addressed in 4XMM-DR9 (Webb et al. 2020), where there will be a flag to indicate that the source is likely to be spurious if it falls on or close to one of these pixels. Most of the sources detected with EXOD that turned out to have a probability of constancy  $> 10^{-4}$  have in common that the net exposure time is less than 5000 s, that as mentioned in Sect. 4.1.1, cannot be ruled out as being due to statistical fluctuations.

While discussing the totality of the newly identified variable sources is beyond the scope of the paper, we analyse in more detail some particularly interesting candidates that show clear outbursts, some of them surprisingly undetected by *XMM-Newton*'s pipeline. In some cases this was done through the identification of a counterpart. In other cases, we extracted the spectra of the source and carried out spectral fitting using XSPEC version 12.9.1p. We note however, that the low number of counts preclude us from performing a detailed spectral analysis, thus we use simple models like `powerlaw` and `bbody` to characterise the emission. We compute fluxes using the pseudo model `cflux` in XSPEC. For the four transients found in M 31 showing outbursts, we refine our analysis in extracting the spectra separately for the burst and the persistent emission by visually selecting the times for each period. We fit each spectra separately and compute the persistent luminosity  $L_{\text{pers}}$ , peak luminosity  $L_{\text{peak}}$  and luminosity ratio ( $L_{\text{peak}}/L_{\text{pers}}$ ) to study the properties of the emission in each

case, in order to put constraints on their nature. Again, due to the low number of counts, we rely on WEBPIMMS<sup>10</sup> to compute this luminosity ratio.  $L_{\text{pers}}$  is estimated using the model of the persistent emission and the median count rate of the light curve while  $L_{\text{peak}}$  is computed with the burst model and the peak count rate of the light curve. Since the distance to M31 is known, we assume that the sources are located at the same distance, 0.78 Mpc, to obtain the luminosity. Below we give a detailed analysis and discuss the nature of the selected sources. The detailed results and the plots of the spectral fitting can be found in Appendix D.

*J173046.7+521846, OBS 0021750201, Source 1.* This source, 2XMM J173046.8+521847 in SIMBAD, is classified as an active galactic nucleus (AGN) candidate (Lin et al. 2012). It shows an aperiodic variability within the duration of the observation (~3500 s). Long term X-ray variability (months/years) is common in AGNs, but short term variability (hours/days) has only been observed in Seyfert types 1.8 and 1.9 (e.g. Hernandez-Garcia et al. 2017 and references therein). We thus propose that this source belongs to one of those classes.

*J083941.3+192901, OBS 0101440401, Source 4.* Although the automatic SIMBAD query classified this source in the “no identification” group, we later found that the source is classified as a star in a cluster with a manual query of the SIMBAD database (Hambly et al. 1995). This source, Cl\* NGC 2632 HSHJ 283 in SIMBAD, shows an outburst lasting ~3000 s, with a main burst lasting ~500 s. The flare was not detected by the XMM-Newton pipeline. Since we have not checked the totality of the source associations, we suspect there might be additional misidentifications like this one amongst the detected sources.

*J015709.1+373739, OBS 0149780101, Source 6.* We detected this ~800 s long outburst in the direction of the open cluster NGC 752, located at a distance of ~430 pc (Daniel et al. 1994; Giardino et al. 2008). We identify a potential blue counterpart at a distance of 0.92'' with a magnitude of 22.83 in the B<sub>j</sub> photographic band (NBXA027004, Lasker et al. 2008). The counterpart is point like; if it is a star in this open cluster, the X-ray transient could be a stellar flare that would correspond to a luminosity of  $1.6 \times 10^{30}$  erg s<sup>-1</sup>.

While our paper was under review, Alp & Larsson (2020) also reported a discovery of this source that they refer to as XT 030206. The authors identify the counterpart as a starburst galaxy and infer a photometric redshift of  $z = 1.17$  by fitting the SED of the optical spectrum. They conclude that the most likely explanation for this X-ray transient is a supernova shock breakout (SBO). As the SED fitting makes the analysis of Alp & Larsson (2020) more thorough we defer to their interpretation of the transient as an SBO. A new X-ray outburst of the source in the future would strengthen the case for the stellar flare.

*J174610.8-290021, OBS 0202670701, Source 8.* Although this observation is highly contaminated by soft proton flares, we detected a variable source with an outburst lasting ~100 s. Surprisingly, the light curve of this source was generated by XMM-Newton’s pipeline, and the burst is visible<sup>11</sup>. However, the source was not classified as variable by the  $\chi^2$  test, probably due to the short duration of the outburst. This source,

CXOU J174610.8-290019 in SIMBAD, is classified as an “X-ray source”. It is a perfect example of why EXOD is better adapted to detecting short transients than other variability tests.

This source is highly variable. Its flux decreased by more than a factor 10 between 2000 and 2004 and then was never detected again by Swift or XMM-Newton, even though these satellites returned to the field on numerous occasions as the source is towards the Galactic center. It has a very absorbed spectrum, typical of sources in the Galactic center,  $n_{\text{H}} \sim 1.6 \times 10^{23}$  cm<sup>-2</sup>, indicating a distance of ~8 kpc.

The luminosity in the burst, assuming a distance of 8 kpc, reaches  $\sim 10^{34}$  erg s<sup>-1</sup>. Its non-flare spectrum shows indications of the presence of an emission line at ~6.66 keV, that we identify as an iron line or a cyclotron resonance line (see Appendix D). The number of counts is too low to draw firm conclusions on the nature of this source, but it is most likely that it is an accreting neutron star in an X-ray binary. Whether it is an LMXB or an HMXB is not clear, since the search for an optical counterpart is complicated due to its location in a crowded region and the addition of the high extinction from the observed photoelectric absorption. A 100 s flare would be uncommon amongst HMXBs, but not unexpected due to the high variability that these sources present (see e.g. Chaty 2011, and references therein). A type I X-ray burst from an NS-LMXB, given the timing and spectral properties of the flare, could be an explanation.

*J183658.4-072119, OBS 0606420101, Source 13.* The outburst of this source lasts ~3000 s. It is located in a crowded region in the galactic plane and presents numerous optical/NIR counterparts within a 6'' radius. We thus conclude that this source is most likely a star.

*J081907.9-384302, OBS 0655650201, Source 15.* This source shows an outburst that lasts ~300 s, but it is preceded and followed by higher flux periods before going back to its quiescent state. Including these periods, the outburst lasts ~2500 s. This type of variability is also observed in stellar flares.

*J175131.6-401533, OBS 0763700301, Source 24.* This outburst, lasting 1000 s, is classified as a variable source in 3XMM-DR8, but it has no SIMBAD object associated. The source is however in a crowded region of the Galactic plane, indicating that this is most likely a star.

*J113407.5+005223, OBS 0770380401, Source 25.* This puzzling source has a burst that lasts only ~200 s, expected for type I X-ray burst, with some structure appearing in the burst. The burst spectrum appears to be fairly soft, although this is based on the joint fit of the persistent and burst emission due to the low number of counts.

Alp & Larsson (2020) presented also this source (under name XT 151219) as an SBO while our paper was under review. They associate the transient with a host galaxy located at redshift  $z = 0.62$ . Although the presence of four galaxies within a 15'' radius from the position of the X-ray source<sup>12</sup> casts some doubts on the identification of the host galaxy, Alp & Larsson (2020) find the flux and spectral properties of the source to be in good agreement with blue supergiant SBO predictions. In general, we too argue that this is likely an extragalactic source because it is located well outside the galactic plane and there is no clear bright counterpart other than background galaxies.

*J022133.7-042346, OBS 0785101401, Source 26.* This source shows a ~1000 s outburst, with a linear rise and an exponential decay. The source has a red-NIR counterpart at a distance

<sup>10</sup> WEBPIMMS: <https://heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3pimms/w3pimms.pl>

<sup>11</sup> Automatically generated light curve of Source 8: <http://xmm-catalog.irap.omp.eu/detection/102026707010084>

<sup>12</sup> NED query: <https://ned.ipac.caltech.edu/>

of  $\sim 125$  pc (Gaia Collaboration 2016, 2018; Bailer-Jones et al. 2018), indicating that it is a star.

*J004307.5+412019, OBS 0109270101, Source M31-1.* This outburst lasts  $\sim 400$  s. The source, [ZGV2011] 23 in SIMBAD, had been previously classified as a low-mass X-ray binary (Zhang et al. 2011), but the nature of the accretor was unknown. We extracted the spectrum of the burst and the persistent emission separately, and fitted each one with an absorbed black body (tbabs\*tbody in XSPEC) and with an absorbed power law (tbabs\*pow). The low number of photons does not allow for a clear preference for one of these models, although a power law gives a slightly better fit for both burst and persistent emission.

From persistent to burst emission, the resulting black body temperature increases from  $\sim 0.2$  to  $\sim 0.3$  keV, whereas the power law index  $\Gamma$  goes from  $\sim 3.3$  to  $\sim 1.9$  (see Appendix D), indicating an apparent spectral hardening during the burst. From the power law indices and a count rate going from  $0.01 \text{ cts s}^{-1}$  to  $0.14 \text{ cts s}^{-1}$ , we computed a  $L_{\text{peak}}/L_{\text{pers}} \sim 30$ . A persistent flux of  $\sim 0.3 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$  scales to a luminosity of  $\sim 2 \times 10^{36} \text{ erg s}^{-1}$  at the distance of M31, 0.78 Mpc. This corresponds to a burst luminosity of  $\sim 6 \times 10^{37} \text{ erg s}^{-1}$ , around 30% of the Eddington luminosity for an NS ( $\sim 1.8 \times 10^{38} \text{ erg s}^{-1}$ ). The presence of such a burst and the luminosity rise indicate that this is likely a type I X-ray burst and would, in this case, identify the accretor as an NS. Such a source shows the power of EXOD since it allows one to identify the nature of a compact object.

*J004215.6+411720, OBS 0650560201, Source M31-3.* This outburst lasts  $\sim 300$  s. The source is classified as an X-ray binary candidate (Lin et al. 2012). The spectral fitting seems to indicate a hardening of the source during the burst, with  $\Gamma$  going from  $\sim 1.6$  in the persistent emission to  $\sim 0.7$  during the burst, although still consistent within the error bars. These values are nevertheless hard for type I X-ray bursts, making it difficult to identify the nature of the compact object in the system. By using the aforementioned power law indices and the count rates from the persistent ( $0.02 \text{ cts s}^{-1}$ ) and burst ( $0.13 \text{ cts s}^{-1}$ ) emission, we calculate  $L_{\text{peak}}/L_{\text{pers}} \sim 16$ . The presence of such a burst makes the type I X-ray burster the most likely scenario, although with the present data we cannot robustly confirm its nature.

*J004210.9+411248, OBS 0674210201, Source M31-6.* The outburst lasts  $\sim 200$  s, and it is followed by a higher flux period before going back to the quiescent state. Including this, the outburst lasts  $\sim 500$  s. This type of outburst is consistent with a type I X-ray burst. The source, XMMM31 J004211.0+411247 in SIMBAD, where it is catalogued as an “X-ray source”, has been detected in 44 *XMM-Newton* observations. It is so faint that its spectrum and time series have only been extracted in 3 out of 44 observations. The observation where we detected its variability was not included<sup>13</sup>, and the source is thus not catalogued as variable.

We fitted the spectrum with an absorbed black body (tbabs\*tbody) and an absorbed power law (tbabs\*pow). The fitted power law gives  $\Gamma \sim 2.4$  for the persistent emission and  $\Gamma \sim 1.1$  at peak. However, there are only 30 available counts during the flare and the source was additionally located next to a chip gap in the pn detector, reducing the number of available photons for the spectral fitting. We obtain  $L_{\text{peak}}/L_{\text{pers}} \sim 40$  for a count rate going from  $0.01 \text{ cts s}^{-1}$  to  $0.13 \text{ cts s}^{-1}$ . The flare spectrum as well as the duration of the burst are all comprised within

the typical values of type I X-ray bursts, and it makes this source an NS-LMXB candidate.

*J004212.1+411758, OBS 0727960401, Source M31-8.* This outburst lasts  $\sim 500$  s. Previously identified as an LMXB named [ZGV2011] 27 (Zhang et al. 2011), it has been detected in 47 *XMM-Newton* observations. In order to identify the nature of the accretor, we fitted its persistent and burst spectrum with an absorbed black body (tbabs\*tbody) and with an absorbed power law (tbabs\*pow). This source was also located near a chip gap in the pn detector, considerably reducing the number of photons used for the spectral fitting, with only 20 available photons during the burst, and thus not enough to detect spectral shape deviations between the persistent and burst emission. However, the presence of a short burst reinforces the previous identification with an LMXB and makes an accreting NS the most likely explanation.

## 5. Discussion

### 5.1. Effectiveness and speed

We present a new algorithm that can be used as a tool to promptly detect interesting sources variable on short timescales in *XMM-Newton*’s EPIC-pn observations. EXOD is computationally inexpensive when compared to other variability tests, like  $\chi^2$  and KS. This makes EXOD faster than generating the light curve of all the sources in one observation.

The combination of the parameters  $TW = 100$  s,  $DL = 8$ ,  $r_{\text{GT}} = 1$  and  $b = 3$  pixels detects a higher number of variable sources according to the  $\chi^2$  and KS variability tests with a lower proportion of non-variable sources. We chose a low  $DL$  in order to detect a higher number of new variable sources, accepting the consequent increase in false positives. The variability of the whole observation with these parameters is computed faster than the generation of a single light curve with the SAS. This makes it optimal to find new variable sources. However, a visual inspection of the detected sources becomes necessary since the amount of detections in bright or extended sources is higher than with higher  $DL$ s, and the  $TW = 100$  s can lead to a worse localisation than with other  $TW$ .

With the optimal parameters applied to different time windows, the percentage of variable sources according to the  $\chi^2$  and KS tests varies between 64.3% and 83.7%. However, only 22.7% of the sources classified as variable in 3XMM-DR8 are detected. The high rate of false negatives is partially compensated by the fact that we get variability information from faint sources for which the light curves were not previously automatically generated (27.1% of the EXOD sources).

### 5.2. Contamination

The source detection with the sliding box technique gives the same problems when applied here as when applied to the source detection in *XMM-Newton*’s pipeline. It is a simple technique that usually gives good results, except in complicated cases where extended sources, bright sources, crowded regions or OOT events are present, since the spatial background can vary rapidly and give rise to spurious variability detections. These problems have been listed in Watson et al. (2009), and such sources have been flagged as spurious in the 3XMM-DR8 catalogue thanks to visual screening. Fortunately, it is easy to distinguish real sources from false detections with a visual inspection.

We are confident that the variability computation algorithm gives the desired results. Nevertheless, the number of spurious detections ( $\sim 1/8$  of the total) remains quite large due to the

<sup>13</sup> Source M31-6: <http://xmm-catalog.irap.omp.eu/source/201125704010111>

source detection procedure. A more sophisticated source detection algorithm could get rid of this problem, but it would probably be more computationally expensive. A comparison of different source detection procedures for *XMM-Newton* can be found in [Valtchanov et al. \(2001\)](#).

We have noted an additional issue for bright sources: the spikes of the PSF are instrumental features with considerable stochastic variability, and they give rise to spurious variable detections around sources with a high number of counts. In most of these cases, the centre of the PSF is detected as a low variability region. This can be visualised in [Fig. B.1](#).

### 5.3. Discoveries

EXOD computes the variability of the whole field of view, which uncovers the variability of faint sources with  $<100$  counts. It also allows the identification of variable sources that had been classified as non variable by *XMM-Newton*'s pipeline due to a non-optimal time binning or to the short duration of the outburst, helping to unravel the nature of these sources.

Applying EXOD to 5751 EPIC-pn observations in the full frame mode with different parameters led to the detection of a total of 2961 sources. Of these, 2536 were previously known sources in 3XMM-DR8 or SIMBAD, and we consider the remaining ones to be spurious detections. These variable sources belong to a wide variety of categories, including stellar flares, cataclysmic variables, type I X-ray bursts, supergiant fast X-ray transients, supernova shock breakouts, AGNs and more.

Finally we discuss 35 sources previously known, but not known to be variable among those classified as “no identification” with SIMBAD and not classified as variable in the 3XMM-DR8 catalogue. Some of the sources that were detected in observations with a net exposure time lower than 5000 s are not classified as variable by other variability tests. We thus recommend the EXOD user to be cautious when analysing such short observations.

Variability is a prominent feature that can be used for classification purposes. Earlier works aiming to identify unidentified sources in the *XMM-Newton* catalogue applied a position cross-match to catalogs at other wavelengths. Other known parameters, such as optical to X-ray flux ratio, are used afterwards to discriminate between source classes ([Pineau et al. 2011](#); [Lin et al. 2012](#)). More recent works have applied machine learning algorithms where timing parameters are the major classifying feature ([Lo et al. 2014](#); [Farrell et al. 2015](#)). Around 27.1% of the EXOD detected sources did not have a previously generated light curve and thus no variability classification. Additionally, many EXOD detections that were bright enough for a light curve generation were classified as non-variable by *XMM-Newton*'s pipeline, where the light curve showed a clearly visible outburst. EXOD thus provides an additional component that can be used to improve the classification of faint sources.

### 5.4. M31

We identified four transients in M31 that are likely to be neutron-star low mass X-ray binaries given their variability and spectra. These double the population of known NSs in M31.

Previous neutron star searches in M31 have used various techniques. [Pietsch & Haberl \(2005\)](#) found two NSs in LMXBs from *XMM-Newton* data, through their bursting nature. [Esposito et al. \(2016\)](#) and [Zolotukhin et al. \(2017\)](#) found a 1.2 s X-ray pulsar in M31, while [Rodríguez Castillo et al. \(2018\)](#) detected a 3 s X-ray pulsar.

Radio pulsar searches have also been carried out. [Rubio-Herrera et al. \(2013\)](#) used the Westerbork telescope at 328 MHz to detect six bursts at a similar dispersion measure (DM), suggesting a neutron star emitting as a rotating radio transient, but the periodicity searches have been unsuccessful, and the source has not been confirmed. Deep radio-pulsar searches with LOFAR at 150 MHz have been reported by [Mikhailov \(2018\)](#) and [van Leeuwen et al. \(2020\)](#), without detections.

That adds up to a total number of known neutron stars in M31 of only four. In this galaxy, more massive than the Milky Way, many more than these four must clearly exist.

Given the increased number of *XMM-Newton* observations of M31 since the detections of e.g. [Pietsch & Haberl \(2005\)](#), and the sensitivity of EXOD to outburst detections demonstrated in [Sect. 4.3](#), additional type I X-ray bursters could be expected to be found through EXOD.

And, indeed, our four new type I X-ray burster candidates J004307.5+412019, J004215.6+411720, J004210.9+411248, and J004212.1+4111758 significantly increase the limited available population of known extragalactic neutron stars.

### 5.5. Further potential for discovery

EXOD has the potential to discover further examples similar to the tidal disruption event recently discovered with unexplained quasi periodic eruptions as it declines to quiescence ([Miniutti et al. 2019](#)). Whilst the data showing the quasi periodic eruptions was not in the 3XMM-DR8 catalogue, another similar source was included, RX J1301.9+2747 ([Giustini et al. 2020](#)). This source was not found during the initial study as the observation was taken in the Extended prime full window mode, a mode that was not included in our study. However, the source was detected highly significantly by a run of EXOD on this data. It was identified as a galaxy and therefore there may be more of these rare and unexplained objects in that category. Identifying them should provide data that will help us understand the nature of the quasi periodic eruptions. Our goal is to apply EXOD to *XMM-Newton* observations performed in the Extended Full Frame mode and the Large Window mode in the future, where we expect to find a high number of unknown transients.

The two sources referred as source 6 and 25 in this paper that were independently detected by [Alp & Larsson \(2020\)](#) and identified as SBOs while this paper was under review, represent exceptional examples of EXOD's extragalactic transient discovery potential.

It is very likely that there are other variable sources that have an inaccurate identification following the automatic SIMBAD query. Some could also have been eliminated by the  $0.5 \text{ cts s}^{-1}$  GTI rate threshold. Finding these would require manual perusal of the associations, and a visual inspection of all the detected sources. This will be pursued in future work.

EXOD can further be adapted to other existing X-ray observatories with similar detector properties, such as *Chandra*, *NuSTAR*, *Swift*, or *eROSITA*. It could also be applied to future missions like *Athena*.

## 6. Conclusions

In this work, we present EXOD, a new algorithm able to detect sources that vary within the duration of an *XMM-Newton* observation. It applies an imaging technique proven in optical and radio transient surveys. Its main strength is to detect the variability of faint sources for which no light curve has been generated by *XMM-Newton*'s automatic pipeline.

We tested its performance by applying it to 5751 observations and subsequently implementing the  $\chi^2$  and Kolmogorov–Smirnov variability tests on the detected sources. With this technique, we were able to find a net count of 2536 variable sources such as stellar flares, type-I X-ray bursts, supergiant X-ray transients, cataclysmic variables, AGNs, and QSOs.

Thirty-five of the sources detected in *XMM-Newton* archival data were unknown transients<sup>14</sup>. In spite of the low number of counts of these sources, usually <100, we performed spectral fitting to get a hint on their nature, and we searched for archival counterparts at other wavelengths.

We find that four of these sources are extragalactic type I X-ray bursters, located in M 31. Other sources, in spite of having analysed their spectra and looked for counterparts at other wavelengths, continue to be unidentified since the number of photons of these sources remains too low to draw any firmer conclusions.

EXOD is a computationally inexpensive algorithm, making it particularly advantageous to rapidly find sources of interest in *XMM-Newton* observations and to provide additional information suitable to identify the nature of a transient source. It can be applied to future *XMM-Newton* observations, and we are looking to later adapt it to other X-ray observatories with similar detection techniques. This could yield discoveries of new transients and ease their multi-wavelength, even multi-messenger follow-up studies.

*Acknowledgements.* IPM thanks A. Gúrpide and the anonymous referee for their input and discussions. IPM, NAW & DTW are grateful to the CNES for financing this study. IPM further acknowledges funding from the Netherlands Research School for Astronomy (NOVA5-NW3-10.3.5.14). JvL received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 617199 (“ALERT”), and from Vici research programme “ARGO” with project number 639.043.815, financed by the Netherlands Organisation for Scientific Research (NWO). This research has made use of data obtained from the 3XMM *XMM-Newton* serendipitous source catalogue compiled by the 10 institutes of the *XMM-Newton* Survey Science Centre selected by ESA; of the SIMBAD database, operated at CDS, Strasbourg, France; and of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

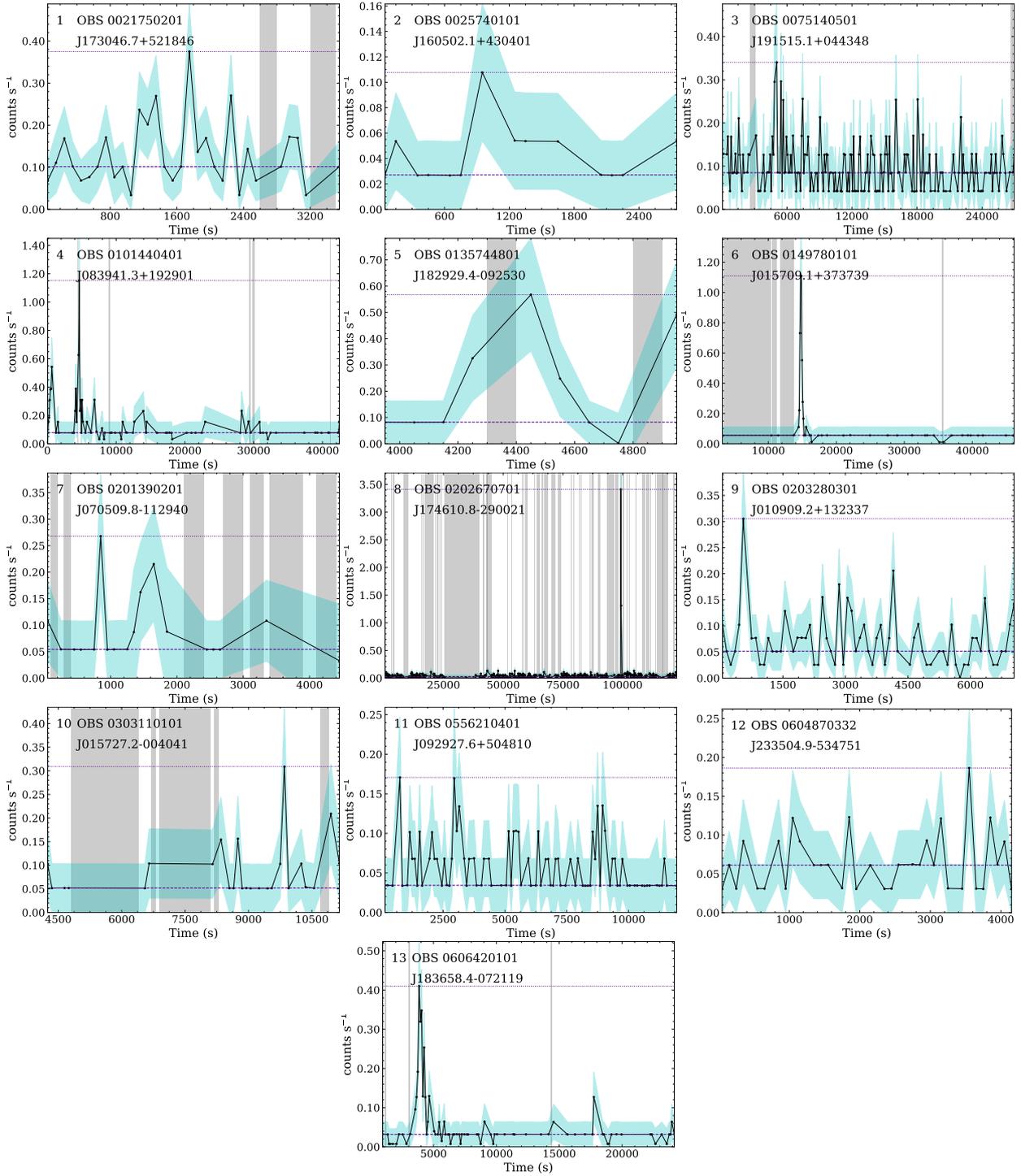
## References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, *ApJ*, **848**, L13
- Alp, D., & Larsson, J. 2020, *ApJ*, **896**, 39
- Bailer-Jones, C. A. L., Rybizki, J., Fousneau, M., Mantelet, G., & Andrae, R. 2018, *AJ*, **156**, 58
- Beloborodov, A. M. 2017, *ApJ*, **843**, L26
- Bhandari, S., Keane, E. F., Barr, E. D., et al. 2018, *MNRAS*, **475**, 1427
- Blackburn, J. K. 1995, *FTOOLS*, **77**, 367
- Bochenek, C., Kulkarni, S., Ravi, V., et al. 2020, *ATel*, **13684**
- Bruch, A. 2017, *New Astron.*, **57**, 51
- Calderwood, T., Dobrzycki, A., Jessop, H., & Harris, D. 2001, *ASP Conf. Ser.*, **238**, 443
- Cash, W. 1979, *ApJ*, **228**, 939
- Chatterjee, S., Law, C. J., Wharton, R. S., et al. 2017, *Nature*, **541**, 58
- Chaty, S. 2011, *ASP Conf. Proc.*, **447**, 29
- Daniel, S. A., Latham, D. W., Mathieu, R. D., & Twarog, B. A. 1994, *PASP*, **106**, 281
- DeLaunay, J. J., Fox, D. B., Murase, K., et al. 2016, *ApJ*, **832**, L1
- De Luca, A., Salvaterra, R., Tiengo, A., et al. 2015, *Proceedings of Swift: 10 Years of Discovery - PoS(SWIFT 10)*, 135
- Esposito, P., Israel, G. L., Belfiore, A., et al. 2016, *MNRAS*, **457**, L5
- Farrell, S. A., Murphy, T., & Lo, K. K. 2015, *ApJ*, **813**, 28
- Favata, F., Flaccomio, E., Reale, F., et al. 2005, *ApJS*, **160**, 469
- Fuchs, J. T., Dunlap, B. H., Dennihy, E., et al. 2016, *MNRAS*, **462**, 2382
- Gaia Collaboration (Prusti, T., et al.) 2016, *A&A*, **595**, A1
- Gaia Collaboration (Brown, A. G. A., et al.) 2018, *A&A*, **616**, A1
- Giardino, G., Pillitteri, I., Favata, F., & Micela, G. 2008, *A&A*, **490**, 113
- Gioia, I. M., Maccacaro, T., Schild, R. E., et al. 1990, *ApJS*, **72**, 567
- Giustini, M., Miniutti, G., & Saxton, R. D. 2020, *A&A*, **636**, L2
- Hambly, N., Steele, I., Hawkins, M., & Jameson, R. 1995, *A&AS*, **109**, 29
- Hernandez-Garcia, L., Masegosa, J., Gonzalez-Martin, O., et al. 2017, *A&A*, **602**, A65
- Hirata, K., Kajita, T., Koshihara, M., et al. 1987, *Phys. Rev. Lett.*, **58**, 1490
- Imanishi, K., Nakajima, H., Tsujimoto, M., Koyama, K., & Tsuboi, Y. 2003, *PASJ*, **55**, 653
- in't Zand, J. J. M., Kries, M. J. W., Palmer, D. M., & Degenaar, N. 2019, *A&A*, **621**, A53
- Jansen, F., Lumb, D., Altieri, B., et al. 2001, *A&A*, **365**, L1
- Kaspi, V. M., & Beloborodov, A. M. 2017, *ARA&A*, **55**, 261
- Kirkpatrick, J. D., Kellogg, K., Schneider, A. C., et al. 2016, *ApJS*, **224**, 36
- Kuulkers, E., Norton, A., Schwöpe, A., & Warner, B. 2003, in *Compact Stellar X-ray Sources*, eds. W. Lewin, & M. van der Klis (Cambridge, UK: Cambridge University Press), *Cambridge Astrophys. Ser.*, **421**
- Lasker, B. M., Lattanzi, M. G., McLean, B. J., et al. 2008, *AJ*, **136**, 735
- Law, C. J., Abruzzo, M. W., Bassa, C. G., et al. 2017, *ApJ*, **850**, 76
- Lin, D., Webb, N. A., & Barret, D. 2012, *ApJ*, **756**, 27
- Lo, K. K., Farrell, S., Murphy, T., & Gaensler, B. M. 2014, *ApJ*, **786**, 20
- Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, *Science*, **318**, 777
- Maan, Y., & van Leeuwen, J. 2017, *IEEE Proc. URSI GASS*.
- Mason, K. 1985, *Space Sci. Rev.*, **40**, 99
- Mereghetti, S., Savchenko, V., Gotz, D., et al. 2020, *ATel*, **13685**
- Metzger, B. D., Berger, E., & Margalit, B. 2017, *ApJ*, **841**, 14
- Metzger, B. D., Margalit, B., & Sironi, L. 2019, *MNRAS*, **485**, 4091
- Mikhailov, K. V. 2018, Ph.D. Thesis, Universiteit van Amsterdam, The Netherlands
- Miniutti, G., Saxton, R. D., Giustini, M., et al. 2019, *Nature*, **573**, 381
- Novara, G., Esposito, P., Tiengo, A., et al. 2020, *ApJ*, **898**, 37
- Parikh, A., José, J., Sala, G., & Iliadis, C. 2013, *Prog. Part. Nucl. Phys.*, **69**, 225
- Pastor-Marazuela, I. 2018, Master's Thesis, Université de Toulouse, France
- Petroff, E., Oostrum, L. C., Stappers, B. W., et al. 2019, *MNRAS*, **482**, 3109
- Pietsch, W., & Haberl, F. 2005, *A&A*, **430**, L45
- Pineau, F.-X., Motch, C., Carrera, F., et al. 2011, *A&A*, **527**, A126
- Pye, J. P., Rosen, S., Fyfe, D., & Schroeder, A. C. 2015, *A&A*, **581**, A28
- Read, A. M., Rosen, S. R., Saxton, R. D., & Ramirez, J. 2011, *A&A*, **534**, A34
- Ridnaia, A., Golenetskii, S., Aptekar, R., et al. 2020, *ATel*, **13688**
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, **116**, 1009
- Rodríguez Castillo, G. A., Israel, G. L., Esposito, P., et al. 2018, *ApJ*, **861**, L26
- Rosen, S. R., Webb, N. A., Watson, M. G., et al. 2016, *A&A*, **590**, A1
- Rubio-Herrera, E., Stappers, B. W., Hessels, J. W. T., & Braun, R. 2013, *MNRAS*, **428**, 2857
- Salvaterra, R. 2015, *J. High Energy Astrophys.*, **7**, 35
- Scargle, J. D. 1998, *ApJ*, **504**, 405
- Scargle, J. D., Norris, J. P., Jackson, B., & Chiang, J. 2013, *ApJ*, **764**, 167
- Scholz, P., & CHIME/FRB Collaboration 2020, *ATel*, **13681**
- Scholz, P., Spitler, L. G., Hessels, J. W. T., et al. 2016, *ApJ*, **833**, 177
- Scholz, P., Bogdanov, S., Hessels, J. W. T., et al. 2017, *ApJ*, **846**, 80
- Sidoli, L. 2013, ArXiv e-prints [arXiv:1301.7574]
- Sidoli, L., Mereghetti, S., Sguera, V., & Pizzolato, F. 2012, *MNRAS*, **420**, 554
- Southworth, J., Gänsicke, B. T., & Breedt, E. 2011, *Proc. Int. Astron. Union*, **7**, 123
- Steehls, D., Marsh, T. R., Barros, S. C. C., et al. 2006, *ApJ*, **649**, 382
- Strüder, L., Briel, U., Dennerl, K., et al. 2001a, *A&A*, **365**, L18
- Strüder, L., Aschenbach, B., Bräuninger, H., et al. 2001b, *A&A*, **375**, L5
- Tavani, M., Ursi, A., Verrecchia, F., et al. 2020, *ATel*, **13686**
- Taylor, M. B. 2005, *TOPCAT & STIL*, **347**, 29
- Tendulkar, S. P., Bassa, C. G., Cordes, J. M., et al. 2017, *ApJ*, **834**, L7
- Turner, M. J. L., Abbey, A., Arnaud, M., et al. 2001, *A&A*, **365**, L27
- Valtchanov, I., Pierre, M., & Gastaud, R. 2001, *A&A*, **370**, 689
- van Leeuwen, J., Mikhailov, K., Keane, E., et al. 2020, *A&A*, **634**, A3
- Watson, M. G., Schröder, A. C., Fyfe, D., et al. 2009, *A&A*, **493**, 339
- Webb, N. A., Coriat, M., Traulsen, I., et al. 2020, *A&A*, in press, **10.1051/0004-6361/201937353**
- Wenger, M., Ochsenein, F., Egret, D., et al. 2000, *A&AS*, **143**, 9
- Zhang, S.-N., Tuo, Y.-L., Xiong, S.-L., & Li, C.-K. 2020, *ATel*, **13687**
- Zhang, Z., Gilfanov, M., Voss, R., et al. 2011, *A&A*, **533**, A33
- Zolotukhin, I. Y., Bachetti, M., Sartore, N., Chilingarian, I. V., & Webb, N. A. 2017, *ApJ*, **839**, 125

<sup>14</sup> Although two were identified with an independent transient search algorithm while this paper was under review.

## Appendix A: Additional figures

Figures A.1, and A.2 show the lightcurves of the sources listed in Table 6 with 100 s time binning.



**Fig. A.1.** Light curves of sources detected with EXOD, present in 3XMM-DR8, but not classified as variable by the automatic pipeline. The light curves are plotted in black with cyan shaded regions representing the  $1\sigma$  error bars. The dashed purple line represents  $\bar{C}$ , the median number of counts. The dotted purple line represents  $C_{\max}$ , the maximal number of counts. The gray vertical shaded regions represent the bad time intervals. Each plot gives the source ID, the OBSID in which it was detected and the 3XMM name of the source.

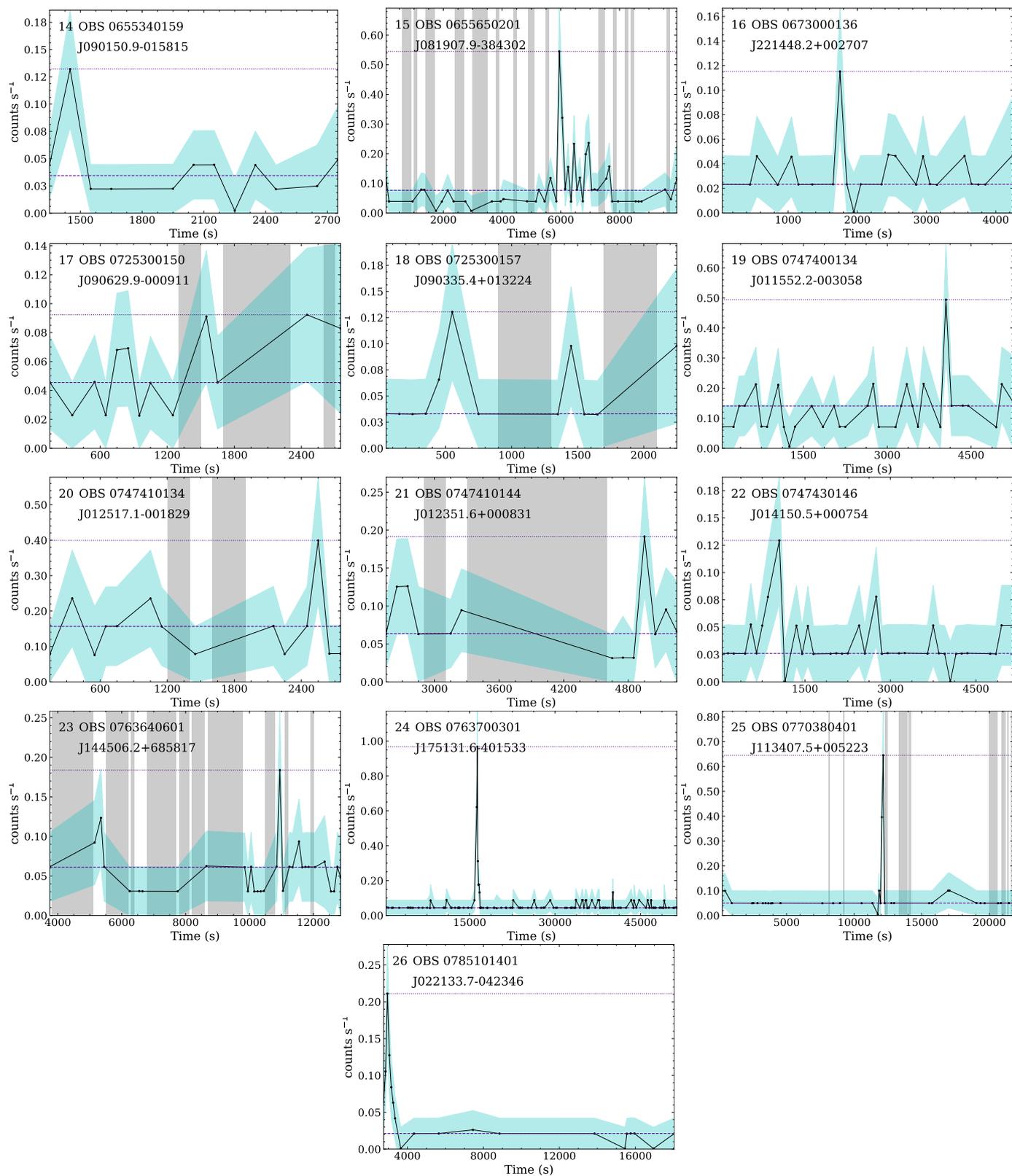


Fig. A.1. continued.

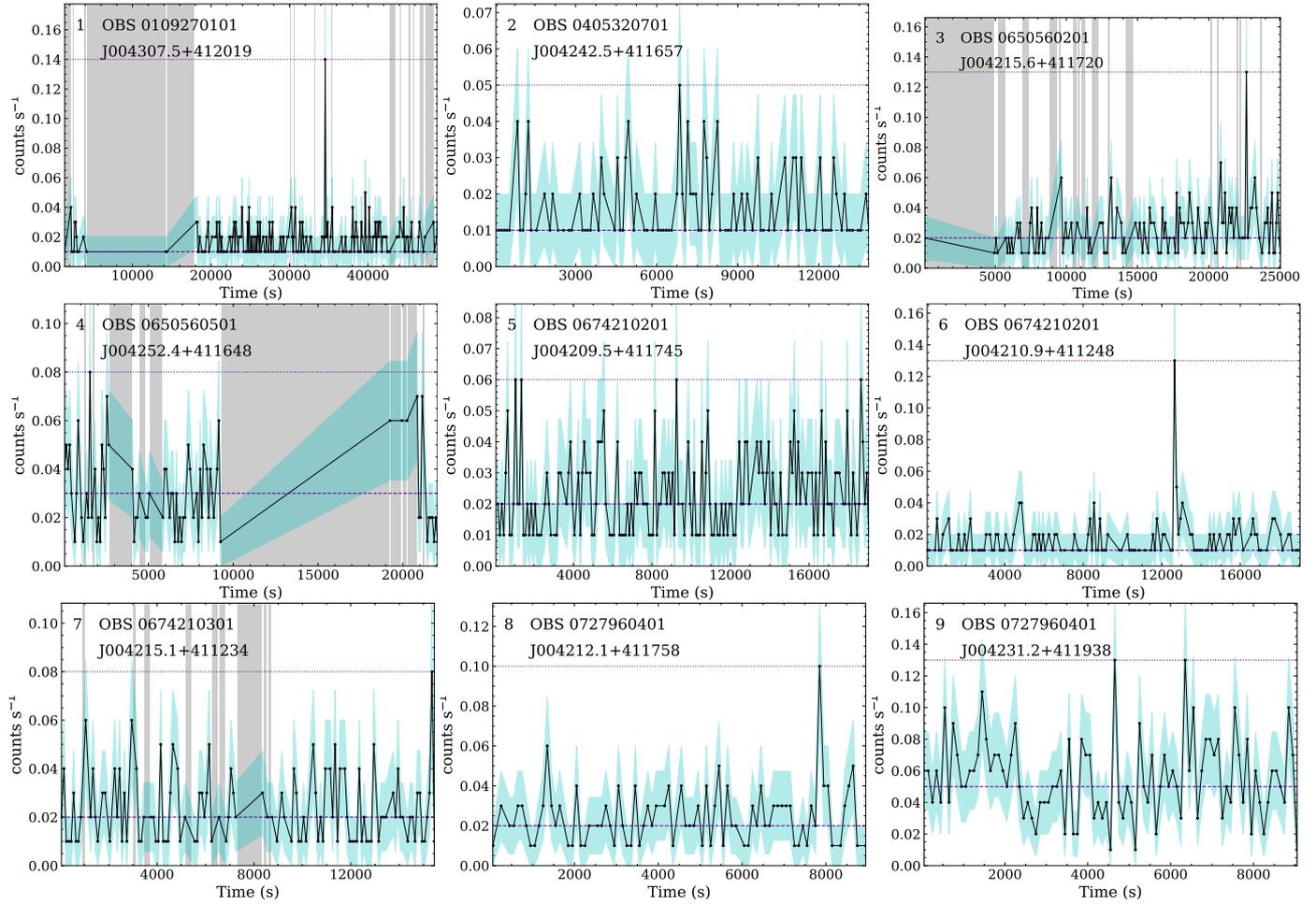


Fig. A.2. Same as Fig. A.1, but with sources found in M 31.

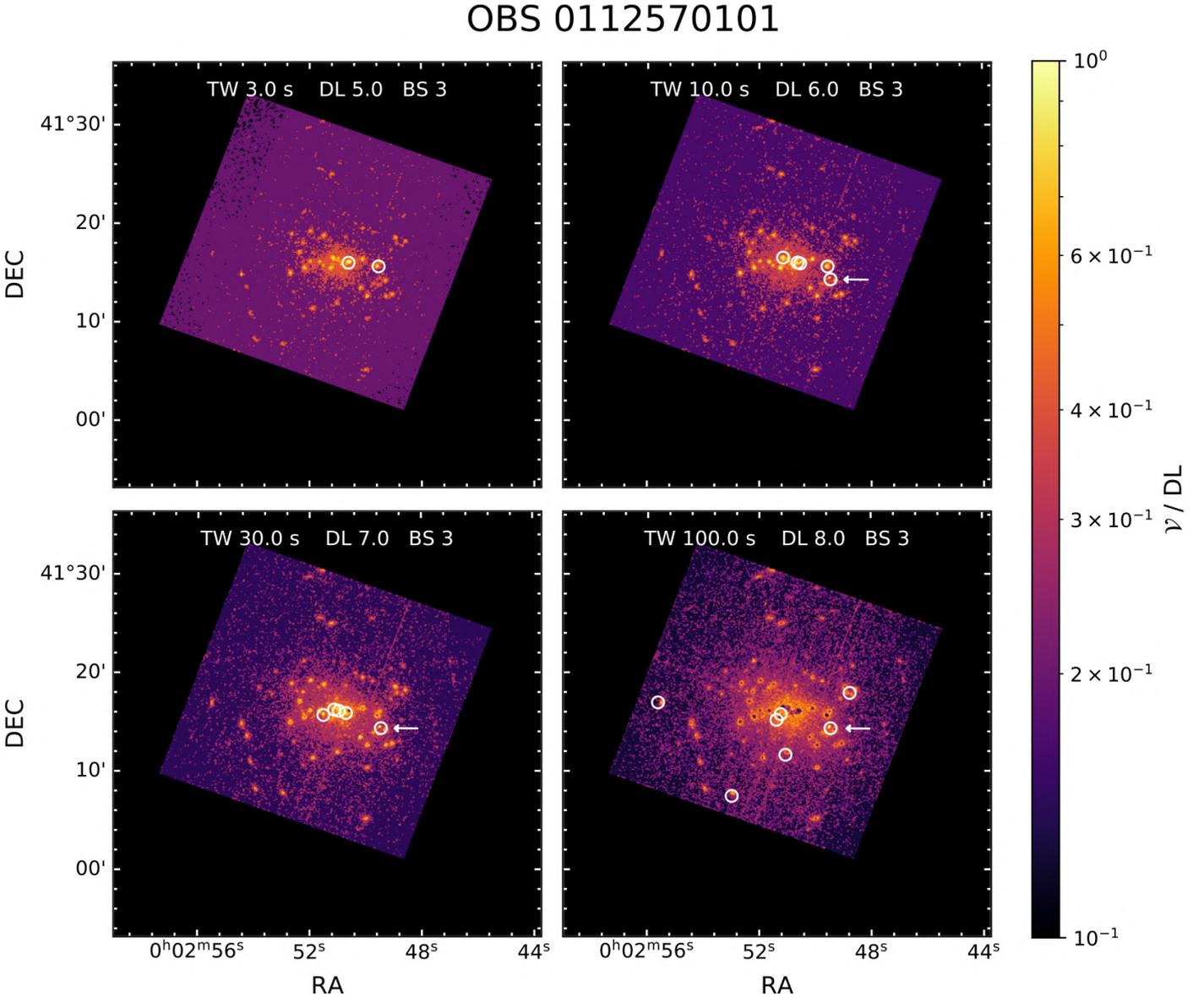
## Appendix B: EXOD output

In Fig. B.1, we present an example of one of the M31 observations analysed with the four sets of parameters given in Table 3 to illustrate the output of EXOD and how the chosen time window affects the value of the variability.

In this particular observation, [Pietsch & Haberl \(2005\)](#) reported the detection of a type I X-ray burst in M31. Other

sources are detected in the same field of view as being variable, amounting up to seven with  $TW = 100$  s.

We can see that the sources detected with different  $TW$  can change. A remarkable fact is that some of the brightest sources that are variable on short timescales (3, 10 s) are seen as a low-variability centre embedded in a variable surrounding area at  $TW = 100$  s.



**Fig. B.1.** Variability of observation 0112570101 of M31's central region. Computed with  $TW = 3$  s (*top left*), 10 s (*top right*), 30 s (*bottom left*) and 100 s (*bottom right*). The colorbar is the same for all the plots and represents the variability divided by the detection level used for each time window, from less variable ( $V/DL = 0.1$ , darker) to more variable ( $V/DL = 1$ , lighter). The detected variable sources are marked by white circles. One of the sources that is detected with  $TW = 100$  s,  $TW = 30$  s and  $TW = 10$  s in this observation was reported in [Pietsch & Haberl \(2005\)](#) as a type I X-ray burst, and is it marked here with white arrows.

**Appendix C: Algorithm**

Algorithm 1 shows the variability computation explained in Sect. 3.1.1 with an algorithmic notation.

**Algorithm 1: EXOD variability computation**


---

**Data:** The events list  $\mathcal{E} = \{ \langle x, y, t \rangle \}$ , the good time intervals  $GTI = \{ \langle t_n, t_m \rangle \}$ , the list of TWs  $T = \{ \langle t_n, t_{n+TW} \rangle \}$  and the good time ratio  $r_{GT}$

**Result:** The variability per pixel, in cts s<sup>-1</sup>

**begin**

```

/* Event count per pixel per time window */
C ← (0)64, 254, |T|;
foreach x ∈ [0..64] do
  foreach y ∈ [0..254] do
    foreach tw ∈ T do
      Cx, y, tw ← |{ ∪i=x-1x+1 ∪j=y-1y+1 ∪t ∈ tw ei, j, t ∈ E }|;
/* Good time ratio computation and TWs filtering */
tw ← T0;
foreach g ∈ GTI do
  /* Deletion of the TWs outside GTIs */
  while tw ∩ g = ∅ do
    if rGT ≠ 0 then remove tw in T, C and  $\tilde{C}$  tw ← next TW ∈ T;
  /* Application of rGT for each TW into a GTI */
  while tw ∩ g ≠ ∅ do
    rGT(tw) ←  $\frac{|tw \cap g|}{|tw|}$ ;
    if rGT(tw) ≥ rGT then
      C-, -, tw ← (xi, j, tw × rGT(tw) : ∀ i ∈ [0..64] ∀ j ∈ [0..254]);
      tw ← next TW ∈ T;
/* Variability computation */
V ← (0)64, 254;
foreach x ∈ [0..64] do
  foreach y ∈ [0..254] do
     $\tilde{C}$  ← median(Cx, y, tw : ∀ tw ∈ T);
    Cmax ← max(Cx, y, tw : ∀ tw ∈ T);
    Cmin ← min(Cx, y, tw : ∀ tw ∈ T);
    if  $\tilde{C}$  ≠ 0 then
      Vx, y ← max(Cmax -  $\tilde{C}$ , |Cmin -  $\tilde{C}$ |);
    else
      Vx, y ← Cmax;
return V;

```

---

## Appendix D: Spectral fitting

In this section, we give the detailed results of the spectral fittings performed with Xspec. The fitted spectra are shown in Fig. D.1, and Table D.1 shows the C-stat fit of the sources to an absorbed black body or an absorbed power law model. The spectra were generated with the SAS task `evselect` using standard filters as recommended. The spectra were rebinned to have at least five counts per spectral bin.

J174610–290021, without the photons emitted during the flare in observation 0202670701, was fitted with an absorbed black body with a Gaussian emission line (`tbabs*(bbody+gaus)`). The fitted parameters include the following:

$$n_{\text{H}} = 1.08^{+1.87}_{-1.08} \times 10^{22} \text{ atom cm}^{-2}$$

$$kT = 2.12^{+1.33}_{-0.78} \text{ keV}$$

$$\text{Gaussian energy} = 6.66^{+0.17}_{-0.13} \text{ keV}$$

$$\text{Gaussian } \sigma = 0.28^{+0.22}_{-0.12}$$

$$\text{Flux} = 1.175 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$$

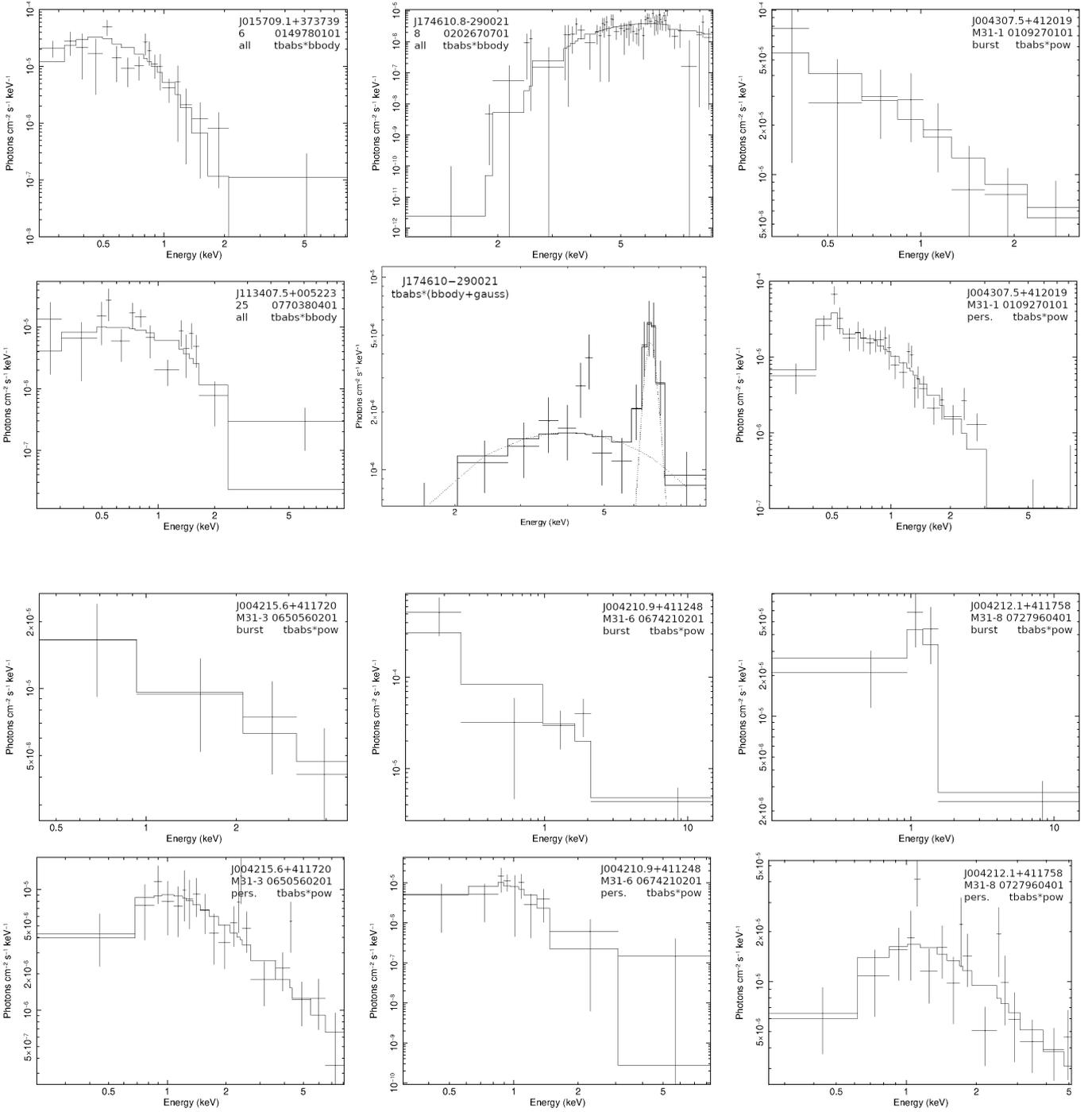
$$\text{C-stat} = 0.8 \text{ (9 d.o.f.)}$$

We should note that the  $n_{\text{H}}$  value obtained by fitting the spectrum during the flare is an order of magnitude smaller than the value of the persistent emission. The large errors are due to the small number of data points. Both values are compatible at  $3\sigma$ .

**Table D.1.** Results from fitting different models to the pn spectra.

(1) Object	(2) Time selection	(3) $n_{\text{H}}$ ( $\times 10^{22}$ at. cm $^{-2}$ )	(4) kT (keV)	(5) $\Gamma$	(6) C-stat (d.o.f.)	(7) $F_{\text{abs}}$ ( $10^{-13}$ erg cm $^{-2}$ s $^{-1}$ )	(8) $F_{\text{unabs}}$ ( $10^{-13}$ erg cm $^{-2}$ s $^{-1}$ )
6 J015709.1+373739	A	<0.12	$0.16 \pm^{0.03}_{0.03}$	...	19.79 (19)	$0.16 \pm^{0.03}_{0.03}$	$0.26 \pm^{0.19}_{0.05}$
		<1.14	...	$2.95 \pm^{7.05}_{0.76}$	14.83 (19)	$0.20 \pm^{0.12}_{0.09}$	$0.41 \pm^{4208}_{0.19}$
8 J174610.8–290021	A	$22.71 \pm^{44.14}_{10.52}$	$1.92 \pm^{2.90}_{1.37}$	...	88.88 (90)	$1.81 \pm^{0.36}_{0.33}$	$4.10 \pm^{3.88}_{1.34}$
		$30.38 \pm^{57.93}_{15.97}$	...	$1.74 \pm^{3.28}_{0.72}$	88.69 (90)	$1.86 \pm^{0.36}_{0.33}$	$9.17 \pm^{47.8}_{5.22}$
25 J113407.5+005223	A	< 0.18	$0.30 \pm^{0.07}_{0.08}$	...	36.30 (20)	$0.15 \pm^{0.04}_{0.04}$	$0.17 \pm^{0.04}_{0.04}$
		$0.31 \pm^{0.39}_{0.23}$	...	$3.55 \pm^{6.04}_{2.12}$	35.56 (20)	$1.89 \pm^{75.8}_{1.51}$	$0.18 \pm^{0.12}_{0.06}$
M 31-1 J004307.5+412019	B	<0.50	$0.32 \pm^{0.18}_{0.17}$	...	1.72 (2)	$0.83 \pm^{0.54}_{0.39}$	$1.15 \pm^{1.90}_{0.42}$
		<0.12	...	$1.91 \pm^{1.02}_{1.04}$	0.68 (2)	$1.91 \pm^{4.45}_{1.09}$	$2.68 \pm^{4.07}_{1.02}$
	P	<0.16	$0.22 \pm^{0.03}_{0.03}$	...	73.23 (51)	$0.23 \pm^{0.03}_{0.03}$	$0.39 \pm^{0.11}_{0.05}$
M 31-3 J004215.6+411720	B	<0.39	$0.68 \pm^{0.56}_{0.23}$	...	2.72 (3)	$1.33 \pm^{1.36}_{0.63}$	$1.47 \pm^{1.32}_{0.66}$
		<0.54	...	$0.74 \pm^{1.67}_{0.67}$	0.48 (3)	$3.55 \pm^{3.86}_{2.31}$	$3.60 \pm^{3.85}_{1.94}$
	P	<0.19	$0.65 \pm^{0.20}_{-0.13}$	...	50.42 (50)	$0.56 \pm^{0.21}_{0.14}$	$0.61 \pm^{0.21}_{0.13}$
M 31-6 J004210.9+411248	B	<3.86	$0.71 \pm^{0.36}_{0.45}$	...	7.23 (2)	$3.22 \pm^{2.29}_{1.00}$	$3.50 \pm^{2.76}_{1.49}$
		<3.55	...	$1.15 \pm^{0.42}_{0.47}$	5.87 (2)	$5.18 \pm^{3.34}_{2.56}$	$5.18 \pm^{3.34}_{2.56}$
	P	0.12 <sup>(*)</sup>	$0.23 \pm^{0.09}_{0.06}$	...	18.59 (23)	$0.13 \pm^{0.05}_{0.04}$	$0.23 \pm^{0.08}_{0.07}$
M 31-8 J004212.1+411758	B	0.12 <sup>(*)</sup>	$0.43 \pm^{0.17}_{0.12}$	...	0.28 (2)	$1.75 \pm^{1.04}_{0.69}$	$2.15 \pm^{1.06}_{0.76}$
		0.32 <sup>(*)</sup>	...	$2.34 \pm^{0.72}_{0.64}$	1.18 (2)	$2.51 \pm^{1.84}_{1.10}$	$6.15 \pm^{4.27}_{2.11}$
	P	< 0.19	$0.73 \pm^{0.11}_{0.11}$	...	31.21 (23)	$1.37 \pm^{0.38}_{0.30}$	$1.48 \pm^{0.39}_{0.30}$
		$0.32 \pm^{0.22}_{0.17}$	...	$1.55 \pm^{0.62}_{0.54}$	24.69 (23)	$2.27 \pm^{0.95}_{0.69}$	$3.06 \pm^{0.99}_{0.59}$

**Notes.** Column (1) gives the name of the object in this paper and in 3XMM, (2) time selection of the fit refers to all photons of the observation (A), the burst (B) or persistent (P) emission. Column (3) gives the interstellar absorption; Cols. (4) and (5) give the black-body temperature (kT) or the power law index ( $\Gamma$ ) respectively; Col. (6) the goodness of fit measured using the C-statistic and the number of degrees of freedom; Cols. (7) and (8) give an estimate of the absorbed ( $F_{\text{abs}}$ ) and unabsorbed ( $F_{\text{unabs}}$ ) fluxes in the 0.2–10.0 keV band, respectively. All the errors are given for 90% confidence for one interesting parameter. For the fluxes, the errors are at 68% confidence. <sup>(\*)</sup> Values frozen for the fit.



**Fig. D.1.** EPIC-pn fitted spectra in the 0.2–10 keV band. Beginning by first pair of rows and then second pair of rows, *from left to right*: (1) J015709.1+373739 fitted with a `tbabs*bbbody` model. (2) J113407.5+005223 fitted with a `tbabs*bbbody` model. (3) and (4) respectively: J174610.8–290021 fitted with a `tbabs*bbbody` model and persistent emission fitted to a `tbabs*(bbbody + gauss)` model. (5) and (6) J004307.5+412019 burst and persistent emission fitted with a `tbabs*pow` model. (7) and (8) J004215.6+411720 burst and persistent emission fitted with a `tbabs*pow` model. (9) and (10) J004210.9+411248 burst and persistent emission fitted with a `tbabs*pow` model. (11) and (12) J004212.1+411758 burst and persistent emission fitted with a `tbabs*pow` model.