Detectability of exoplanet transits with Athena’s WFI instrument. Testing for white and correlated noise.

Stefania Carpano\textsuperscript{a}, Jörn Wilms\textsuperscript{b}, and Arne Rau\textsuperscript{a}

\textsuperscript{a}MPE, Giessenbachstr. 1, 85748 Garching, Germany
\textsuperscript{b}Remeis-Observatory & ECAP, Sternwartstr. 7, 96049 Bamberg, Germany

ABSTRACT

One of the science goals of the Athena mission is to detect and characterise, in the X-ray domain, transits of hot Jupiter-like planets orbiting their parent stars. To date, the only candidate for this kind of studies is HD 189733b, a Jupiter-size planet in a 2 d orbit, for which a transit depth of 6-8\% has been observed accumulating several Chandra and XMM-Newton observations.

We simulate in this work realistic light curves of exoplanet transits using the Athena end-to-end simulator, SIXTE, and derive the expected signal-to-noise ratios (SNR) for different instrument configurations and planetary system parameters. We first produce flat light curves for the currently existing WFI instrument designs and for different source fluxes to extract the expected (white noise) standard deviation. Next, moderate levels of correlated noise and transits of different depths are added to the light curves.

As expected, for pure white noise the SNR is proportional to the square root of the flux, to the light curve bin size and to the number of co-added transits, and by definition proportional to the transit depth. When correlated noise starts to be significant, rebinning the data will only slightly increase the SNR, depending on the noise characteristics. Considering only white noise, a transit observed in a source like HD 189733, that has a flux around $5 \times 10^{-13}$ erg s\textsuperscript{-1} cm\textsuperscript{-2} and a transit depth of about 5\% can be detected with a SNR $> 3$ in a unique transit. With correlated noise, several transits might be necessary.

We also simulate trapezoidal shaped transits and try to recover the ingress/egress times after addition of noise. The relative error on the fitted ingress times is below 10\% for most of the light curves with SNR $> 1$.

Keywords: Athena, WFI instrument, exoplanets, simulations, performance analysis

1. INTRODUCTION

“Hot” Jupiter-sized planets, i.e. in a close orbit, are believed to interact with their parent stars inducing high energy emission during these processes. The transiting planet would produce an eclipse in the X-ray light curve with a depth and duration linked to its high-energy absorption radius, which might be larger than the optical one. To date, the only exoplanet for which X-ray transits have been detected is HD 189733b. It is a hot Jupiter of 1.138 M\textsubscript{Jup} orbiting a K-type star in a 2.22 d orbit, published by Poppenhaeger et al. \cite{1} who revealed eclipses with a transit depth of 6--8\% by accumulating 7 Chandra and XMM-Newton observations.

Due to the larger effective area of Athena compared to current X-ray observatories, transits from hot Jupiters should be recorded with a higher SNR. The Athena supporting Science document on “Solar system and exoplanets” \cite{2} describes the proposed strategy to study and characterise the star-planet interactions at X-ray wavelengths.

Here, we present realistic X-ray light curve simulations of transiting Jupiter like planets for the Athena Wide Field Imager \cite{3 and 4}, using the SIXTE end-to-end simulator \cite{5}. First we describe the performed simulations, and then present the expected SNR for different instrument configurations and planetary system parameters. We also simulate transits having a trapezoidal shape and try to recover the ingress/egress times in noisy light curves.

Further author information: (Send correspondence to S. Carpano)
S. Carpano.: E-mail: scarpano@mpe.mpg.de, Telephone: +49 89 300003522
Table 1. Count rate (Rate) in count s⁻¹, number of valid patterns (Nvalid) and pileup valid pattern (NPvalid), expressed in %, as a function of source flux (in the 0.1–15 keV band) for the WFI large detector in window mode (w16), and for two mirror outer radii and with or without external light blocking filter.

### 2. LIGHT CURVE SIMULATIONS

#### 2.1 Flat light curves produced by SIXTE

Simulations of flat light curves, without transits, are performed using SIXTE (Simulation of X-ray telescopes), version 2.0.7. The software and instrument files related to Athena can be downloaded at the website [http://www.sternwarte.uni-erlangen.de/research/sixte/](http://www.sternwarte.uni-erlangen.de/research/sixte/). The first step consists in creating SIMPUT (Simulation Imput) Files, using the `simputfile` command, that will be used to run the simulator. For this, an input spectral model, in a format compatible with XSPEC [6] has to be provided. To be realistic but without going into details, we use the model derived for HD 189733 by Poppenhaeger et al. [1]: an `apec+apec` model with $kT_1 = 0.25$ keV, $N_1 = 1.81 \times 10^{-4}$ and $kT_2 = 0.62$ keV, $N_2 = 6.6 \times 10^{-5}$. We left default values for the abundances.

Besides the spectral model, a theoretical input light curve is necessary for the simulator, and is created as flat (being 1 everywhere). The next step consists of generating the event files using the SIXTE simulator. These contain information about the number of single, double, triple, quadruple and invalid patterns that are expected to occur statistically. Observations of bright sources, i.e. > 1 mCrab, with the large detectors of the WFI in full frame operation are affected by pile-up [7]. We therefore decided to simulate light curves using configuration files of the WFI’s large detector in a window mode (`depfet_b1l_ff_w16.xml`) which reads 16X512 pixels with a readout time of 2.5 µs per line (so 40 µs per frame)*. As the design of Athena telescope is not finalised, we use the WFI response files for two different mirror outer radii, the nominal mirror assembly (NMA, 1469 mm) and the small mirror assembly (SMA, 1190 mm). In addition, we simulate observations with and without the external optical/UV light blocking filter of the WFI. The version of the response files used for the analysis are `v20150504`.

We also run the SIXTE simulator for different flux values going from $1 \times 10^{-10}$ to $5 \times 10^{-13}$ erg cm⁻² s⁻¹ (in the 0.1–15 keV energy band), and using a short exposure time of 10000 s. These dummy flat light curves are only used to extract the standard deviations, which increase as the square root of the flux. Transits with different depths are added in a further stage.

In Table 1, we display the mean count rate (Rate) expressed in count s⁻¹, the number of valid patterns (Nvalid) and pileup valid patterns (NPvalid), expressed in %, extracted from the different event files using the

---

*We note, that the final operational modes of the WFI are not defined yet and that the adapted mode is only one possible option.
Comparing with count rates from Dauser et al. [7], we have a much higher difference in the values between the “with” and “without” filter configurations. This is because the spectrum is here much softer, a thermal model instead of a Crab-like power-law spectrum, and thus more sensitive to the responses at lower energies. The number of valid patterns using the large detector in window mode (w16) is more than 95% for all fluxes and the number of pile-up valid events is very small ($\lesssim 1\%$).

In the rest of the paper, we report results only for the WFI instrument with the NMA, which has the largest effective area. On the other hand, the configuration “with” filter is used since the observed stars will likely be bright in the visible band (e.g., HD189733 is 7.65 mag [8]). Results for the other WFI instrument configurations (without filter and SMA) can be scaled according to the flux changes reported in Table 1. The X-IFU instrument has a smaller effective area and is therefore not simulated in this work.

2.2 White versus correlated noise

The output of the SIXTE simulations provides light curves containing Gaussian white noise only, caused by the uncertainty in the number of photons getting through the mirrors and collected by the detectors, including as well the expected background noise. A larger effective area will increase the SNR. The white noise has also the property that SNR enhances with the bin size of the light curve (i.e. in the context of exoplanets with the duration of the transit). With correlated noise, things get more complicated. Pont et al. [9] describes in detail, the influence of red/correlated noise in the planetary transit detection. Depending on the characteristics of this noise, resampling the data might not be so helpful anymore.

2.3 Simulating light curves with $1/f^\gamma$ correlated noise

In correlated noise time series, the value of a data point at time $t$ is correlated to some extend with the preceding data points. For an autoregressive model of order $n$ AR($n$), that is used in this work, the value $x$ at time $t$ is defined as:

$$x(t) = c + \sum_{i=1}^{n} a_i x(t-i) + w(t)$$

where $c$ is a constant, $a_i$ are the autoregressive coefficients and $w(t)$ is the (uncorrelated) white noise at time $t$.

In the context of this work, the values of the AR coefficients may vary from source to source, and possibly from transit to transit. For the simulations performed here, light curves with $1/f^\gamma$ noise are produced, i.e. with a Power Spectral Density described by a power law of slope $\gamma$ (varying from 0 to 2). Without moving to the frequency domain, a $1/f^\gamma$ noise can be produced using the formulas (115) and (116) from Kasdin [10]:

$$x(t) = -a_1 x(t-1) - a_2 x(t-2) - a_3 x(t-3) - ... + w(t)$$

where

$$a_0 = 1$$

$$a_i = \left( i - 1 - \frac{\gamma}{2} \right) \frac{a_{i-1}}{i}$$

The $a_i$ coefficient are generated until $i = 63$ as is it done in MATLAB for the dsp.ColoredNoise routine. $w(t)$ is the white noise light curve, with the standard deviation being extracted from the SIXTE flat light curves, proportional to the square root of the source flux. When $\gamma = 0$, only the white noise is describing the process.

2.4 Adding transits

Once the correlated noise light curves of 50000s are produced and normalised to 1, single transits of a fixed duration of 5000s and different depths (5%, 10%, 15% and 20%) are added. The shape of the transit is assumed to be square.
2.5 Example of simulated light curves

Figures 1 and 2 show examples of light curves simulated with white noise ($\gamma = 0$) only and with correlated noise ($\gamma = 1$), respectively. The time series are shown for fluxes ranging from $1 \times 10^{-10}$ to $5 \times 10^{-13}$ erg cm$^{-2}$ s$^{-1}$, from left to right, and for transit depths ranging from 20% to 5% from top to bottom. The source rate is normalised to 1 and light curves are binned with a size of 100 s. The noiseless light curves are overplotted in red.

Figure 1. Simulated white noise light curves with normalised count rate, for the WFI instrument with NMA and with external light blocking filter. Fluxes decrease from left to right and transit depths decrease from top to bottom. The red lines show the noiseless light curves.
3. SIGNAL-TO-NOISE RATIO

The signal-to-noise ratio is calculated as the ratio between the transit depth \(d\) and the standard deviation of the simulated light curve \(\sigma\) (not including the transit). For a single transit:

\[
\text{SNR}_{1\text{tr}} = \frac{d}{\sigma}
\]  

(5)

Because light curves of transits observed at different epochs are uncorrelated, co-adding the data from different transits should increase the SNR by a factor \(\sqrt{N_{\text{tr}}}\), where \(N_{\text{tr}}\) is the number of co-added transits.

Figure 3 shows the SNR derived from the set of simulated light curves, as a function of the source flux. Results for the different transit depths are shown in red (5%), green (10%), cyan (15%), and blue (20%), for data binned at 1 and 100 s (left and right columns, respectively). SNR for light curves with white noise (\(\gamma = 0\)), moderate correlated noise (\(\gamma = 0.5\)), and stronger correlated noise (\(\gamma = 1.0\)) are shown in the graphics of the top, middle and bottom, respectively.

The numbers indicated at the left of the graphs correspond to the SNR for a flux of \(1 \times 10^{-11}\) erg cm\(^{-2}\) s\(^{-1}\), for the different transit depths. The SNR still increases as the square root of the source flux which means that the standard deviation of the light curves, with the noise defined in 2.3, is proportional to the standard deviation derived from the white noise. Binning the data with a larger bin size is clearly less useful for stronger correlated noise. For \(\gamma = 1\), the SNR only increases by a factor of 1.7 when increasing the bin size by a factor of 100, while for pure white noise we gain a factor of 10.
4. RECOVERING THE TRANSIT INGRESS/EGRESS TIMES

In the above simulations, the shape of the transit was assumed to be square which means that the transit ingress or egress times are negligible with respect to the total duration of the eclipse. In this section we define a more realistic trapezoidal transit shape as shown in Fig. 4, where $t_a$ and $t_d$ are the start and end of the eclipse, respectively, $t_b$ marks the end of the transit ingress and $t_c$ the start of the transit egress. Furthermore we make the following assumption:

- $t_a$ and $t_d$, and hence as well the total duration of the eclipse, are approximatively known,
- $|t_b - t_a| = |t_d - t_c| = t_{ing} = t_{egr}$, i.e. the ingress and egress times ($t_{ing}$ and $t_{egr}$) are the same and hence the transit is symmetric.

In the simulations performed in this section, we assume $t_a$ and $t_d$ to be constant and allow only $t_{ing}$ to be variable with values of 200s, 500s, 750s and 1000s for a total eclipse duration of 5000s. Noise is then added to these theoretical light curves as it is done previously with sources fluxes having 6 different values, the transit depths 4 values, the bin size 2 values and the $\gamma$ exponent 3 values, so that a total set of 576 different light curves are simulated.
For each of them a trapezoidal transit is fitted using the `curve_fit` function of Python ([http://docs.scipy.org/doc/scipy-0.15.1/reference/generated/scipy.optimize.curve_fit.html](http://docs.scipy.org/doc/scipy-0.15.1/reference/generated/scipy.optimize.curve_fit.html)). The theoretical values of $t_a$ and $t_d$ are given as initial guess parameters. To check the performance of the fit we plot the relative error of the transit ingress time, defined here below, as a function of the SNR:

$$\left|\frac{t_{\text{ing}1} - t_{\text{ing}2}}{t_{\text{ing}2}}\right|$$

where the subscripts 1 and 2 refer to the fitted and theoretical values, respectively.

Results of the fit performance are shown in Fig. 5 where the relative error (in %) of the time ingress is plotted as a function of the SNR for the fits with finite or small error bars (top), and with infinite or very large error bars (bottom). From the graph, we can conclude that with a SNR > 1 most of the time ingress values can be recovered with an error below 10%. This result can eventually be improved by applying some denoising filter to the light curves before performing the fit, but this will not be tested here.
5. CONCLUSIONS

*Athena* will certainly contribute in more detections and characterisation of planetary transits of Hot Jupiter-like planets in the X-ray band, whereas these are currently performed mainly in optical and infrared bands. Currently only one source, HD 189733, has transits observed in the X-rays with a SNR $> 3$.

For a particular source characterised by its specific flux and transit depth, the SNR can only increase by binning the data to the length of the transit duration and by accumulating data from several transits. As an example, for a source similar to HD 189733, with an X-ray flux of $5 \times 10^{-13} \text{erg cm}^{-2} \text{s}^{-1}$, a transit depth of 5% with a duration of 2000 s could be detected with the *Athena* WFI with a SNR $> 3$ in a single transit, in the case of pure white noise. It is still difficult to quantify the number of transits required to detect such a small depth in the presence of correlated noise, this can be few as it can be many, depending on the noise characteristics. We also note that several techniques exist in the literature to filter out correlated noise induced by stellar variability, in order to increase the detectability of planetary transits. It is out of the scope of this work to test them, but these could be apply on a later stage, on real observations.

Furthermore, we mention that a database containing some X-ray properties of exoplanets (like stellar spectra, light curves and luminosities) can be found on the webpage [http://sdc.cab.inta-csic.es/xexoplanets/](http://sdc.cab.inta-csic.es/xexoplanets/) and more details are given in the related paper from Sanz-Forcada et al. [11]. The X-ray fluxes of the currently known exoplanet systems (not all of them transiting), are provided in the Table 3 of the manuscript, where values...
and upper-limits have been calculated for 54 stars using XMM-Newton and Chandra data. Only 5 over the 54 sources have fluxes between $5 \times 10^{-13}$ and $1 \times 10^{-10} \text{ erg cm}^{-2} \text{ s}^{-1}$ (although they are computed in the 0.12–2.48 keV band but should not differ too much to the ones extracted from our broader 0.1–15 keV band). All other sources are much fainter in the X-ray wavelengths.

As a last point, in the case that X-ray bright transiting exoplanetary systems are found in the coming years, it might possible to perform differential spectroscopy to derive the expected spectra absorbed by the planet atmosphere. In that case, the X-IFU instrument, that has a much higher spectral resolution, could be used for this purpose.

REFERENCES