THE ASTROPHYSICAL JOURNAL, 541:153–165, 2000 September 20


# X-RAY EMISSION OF MARKARIAN 421: NEW CLUES FROM ITS SPECTRAL EVOLUTION. I. TEMPORAL ANALYSIS

G. FOSSATI
Center for Astrophysics and Space Sciences, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0424; gfossati@ucsd.edu

A. CELOTTI, M. CHIABERGE, AND Y. H. ZHANG
International School for Advanced Studies, via Beirut 2-4, 34014 Trieste, Italy

L. CHIAPPETTI
Istituto di Fisica Cosmica G. Occhialini, via Bassini 15, 20133 Milano, Italy

G. GHISELLINI,[1] L. MARASCHI, AND F. TAVECCHIO
Osservatorio Astronomico di Brera, via Brera 28, 20121 Milano, Italy

E. PIAN
ITeSRE/CNR, via Gobetti 101, 40129 Bologna, Italy

AND

A. TREVES
Università dell'Insubria, via Lucini 3, 22100 Como, Italy
*Received 2000 February 9; accepted 2000 April 21*

## ABSTRACT

Mrk 421 was repeatedly observed with *BeppoSAX* in 1997 and 1998. This is the first of two papers where we present the results of a thorough temporal and spectral analysis of all the data available to us, focusing in particular on the flare of 1998 April, which was simultaneously observed also at TeV energies. Here we focus on the time analysis, while the spectral analysis and physical interpretation are presented in the companion paper. The detailed study of the flare in different energy bands reveals very important new results: (1) hard photons lag the soft ones by 2–3 ks—a behavior opposite to what is normally found in high-energy peak BL Lac X-ray spectra; (2) the flare light curve is symmetric in the softest X-ray band, while it becomes increasingly asymmetric at higher energies, with the decay being progressively slower than the rise; and (3) the flux decay of the flare can be intrinsically achromatic if a stationary underlying emission component is present. The temporal and spectral information obtained challenge the simplest models currently adopted for the (synchrotron) emission and most importantly provide clues on the particle acceleration process.

*Subject headings:* BL Lacertae objects: general — BL Lacertae objects: individual (Markarian 421) — galaxies: active — X-rays: galaxies — X-rays: general

## 1. INTRODUCTION

Blazars are radio-loud active galactic nuclei (AGNs) characterized by strong variability, large and variable polarization, and high luminosity. Radio spectra smoothly join the infrared-optical-UV ones. These properties are successfully interpreted in terms of synchrotron radiation produced in relativistic jets and beamed into our direction because of plasma moving relativistically close to the line of sight (e.g., Urry & Padovani 1995). Many blazars are also strong and variable sources of GeV $\gamma$-rays, and in a few objects the spectrum extends up to TeV energies. The hard X- to $\gamma$-radiation forms a separate spectral component, with the luminosity peak located in the MeV–TeV range.

The emission up to X-rays is thought to be due to synchrotron radiation from high-energy electrons in the jet, while it is likely that $\gamma$-rays derive from the same electrons via inverse Compton (IC) scattering of soft (IR-UV) photons—synchrotron or ambient soft photons (e.g., Sikora, Begelman, & Rees 1994; Ghisellini et al. 1998).

The contributions of these two mechanisms characterize the average blazar spectral energy distribution (SED), which typically shows two broad peaks in a $\nu F_\nu$ representation (e.g., von Montigny et al. 1995; Sambruna, Maraschi, & Urry 1996; Fossati et al. 1998b): the energies at which the peaks occur and their relative intensity provide a powerful diagnostic tool to investigate the properties of the emitting plasma, such as electron energies and magnetic field (e.g., Ghisellini et al. 1998). Moreover, variability studies, both of single band and of simultaneous multifrequency data, constitute the most effective means to constrain the emission mechanisms at work in these sources as well as the geometry and modality of the energy dissipation. The quality and amount of X-ray data on the brightest sources start to allow us to perform a thorough temporal analysis as function of energy and determine the spectral evolution with good temporal resolution.

In X-ray–bright BL Lac objects (HBL, from high-energy peak BL Lac objects; Padovani & Giommi 1995), the synchrotron maximum (usually) occurs in the soft X-ray band, and the inverse Compton emission extends in some cases to the TeV band where—thanks to ground-based Cerenkov telescopes—four sources have been detected up to now: Mrk 421 (Punch et al. 1992), Mrk 501 (Quinn et al. 1996), 1ES 2344+514 (Catanese et al. 1998), and PKS 2155–304 (Chadwick et al. 1999). If the interpretation of the SED properties in terms of synchrotron and IC radiation is

[1] Osservatorio Astronomico di Brera, via Bianchi 46, 22055 Merate (Lecco), Italy.

correct, a correlation between the X-ray and TeV emission is expected.

Mrk 421 ($z = 0.031$) is the brightest BL Lac object at X-ray and UV wavelengths and the first extragalactic source discovered at TeV energies, where dramatic variability has been observed with doubling times as short as 15 minutes (Gaidos et al. 1996). As such it was repeatedly observed with X-ray satellites, including *BeppoSAX*. Remarkable X-ray variability correlated with strong activity at TeV energies has been found on different occasions (Macomb et al. 1995, 1996; Takahashi et al. 1996; Fossati et al. 1998a; Maraschi et al. 1999). In particular, the 1998 *BeppoSAX* data presented here were simultaneous with a large TeV flare detected by the Whipple Observatory (Maraschi et al. 1999).

This paper is the first of two that present the results of a uniform, detailed spectral and temporal analysis of *BeppoSAX* observations of Mrk 421 performed during 1997 and 1998. Here we focus on the data reduction and the timing analysis and also discuss the results on the spectral variability derived from the different properties of the flux variations in different energy bands.

The paper is organized as follows. We briefly summarize the characteristics of *BeppoSAX* (§ 2) and introduce the observations studied (§ 3). We then address the temporal analysis of the variability, considering several energy bands and comparing the light-curve features by means of a few simple estimators for the 1997 and 1998 observations (§ 4). The remarkable flare observed in 1998 is the object of a further deeper analysis, reported in § 5, focused on timescales and time lags. Section 6 contains a summary of the results of the temporal analysis, preparing the ground for the comprehensive discussion presented in Paper II (Fossati et al. 2000). There they are considered together with the results of the spectral analysis and thus used to constrain a scenario able to interpret the complex spectral and temporal findings.

## 2. *BEPPOSAX* OVERVIEW

For an exhaustive description of the Italian/Dutch *BeppoSAX* mission, we refer to Boella et al. (1997) and references therein. The narrow field co-aligned instrumentation (NFI) on *BeppoSAX* consists of a low-energy concentrator spectrometer (LECS), three medium-energy concentrator spectrometers (MECS), a high-pressure gas scintillation proportional counter (HPGSPC), and a phoswich detector system (PDS). The LECS and MECS have imaging capabilities in the 0.1–10 keV and 1.3–10 keV energy bands, respectively, with energy resolution of 8% at 6 keV. At the same energy, the angular resolution is about 1.′2 (half-power radius). In the overlapping energy range the MECS effective area (150 cm$^2$) is $\sim$3 times that of the LECS. Furthermore, the exposure time for the LECS is limited by stronger operational constraints to avoid UV light contamination through the entrance window (LECS instrument is operated during Earth dark time only). The HPGSPC covers the range 4–120 keV, and the PDS the range 13–300 keV. In the overlapping interval the HPGSPC has a better energy resolution than the PDS, but it is less sensitive than both PDS and MECS. Therefore, HPGSPC data will not be discussed in this paper.

The present analysis is based on the SAXDAS linearized event files for the LECS and the three MECS experiments, together with appropriate background event files, as produced at the *BeppoSAX* Science Data Center (revs. 0.2, 1.1, and 2.0). The PDS data reduction was performed using the XAS software (Chiappetti & Dal Fiume 1997) according to the procedure described in Chiappetti et al. (1999).

## 3. OBSERVATIONS

Mrk 421 was observed by *BeppoSAX* in the springs of 1997 and 1998. The journal of observations is given in Table 1.

### 3.1. *1997*

The 1997 observation comprised several pointings spanning the interval between April 29 and May 7. MECS data are not available for May 7 because of the failure of the detector unit 1 on May 6. In this paper we will not consider the LECS data of this last day of the 1997 campaign because, unfortunately, LECS data alone do not provide useful spectral and variability information. The *net* exposure time (excluding the May 7 data) was $\sim$52 and $\sim$117 ks for LECS and MECS, respectively, while the on-source time coverage computed as the sum of the observations between each $T_{\rm start}$ and $T_{\rm stop}$ is about 58 hr.

TABLE 1

JOURNAL OF OBSERVATIONS

| DATE (UTC) | | NET EXPOSURE TIME (ks) | | | |
|---|---|---|---|---|---|
| Start | Stop | LECS | MECS | PDS[a] | *BeppoSAX* ARCHIVE NUMBER |
| 1997/04/29:04:02 | 1997/04/29:14:42 | 11.6 | 21.8 | 18.4 | 50032001 |
| 1997/04/30:03:19 | 1997/04/30:14:42 | 11.4 | 24.1 | 23.0 | 50032002 |
| 1997/05/01:03:17 | 1997/05/01:14:42 | 11.2 | 23.8 | 22.8 | 50032003 |
| 1997/05/02:04:10 | 1997/05/02:09:41 | 4.4 | 11.4 | 11.0 | 50016002 |
| 1997/05/03:03:24 | 1997/05/03:09:41 | 4.3 | 11.7 | 11.2 | 50016003 |
| 1997/05/04:03:25 | 1997/05/04:09:45 | 4.9 | 12.2 | 11.6 | 50016004 |
| 1997/05/05:03:32 | 1997/05/05:09:45 | 5.0 | 11.9 | 11.4 | 50016005 |
| 1997/05/07:04:47 | 1997/05/07:10:27 | 6.0 | ... | 13.1 | 50016006[b] |
| 1998/04/21:01:52 | 1998/04/22:03:13 | 23.6 | 29.6 | 42.2 | 50686002 |
| 1998/04/23:00:27 | 1998/04/24:06:37 | 27.2 | 34.7 | 50.4 | 50686001 |

[a] Net exposure time over half of the total area because of the rocking mode in which the instrument is operated.

[b] On 1997 May 7 MECS 1 had a fatal failure. Only LECS and PDS data are available for that day, and they are not discussed in the paper.

Results on the first half of the 1997 campaign (April 29–May 1) have been presented by Guainazzi et al. (1999), where the details and motivation of the observations are given. We reanalyzed those data along with the new ones, applying the same techniques, in order to obtain a homogeneous set of results, necessary for a direct comparison.

### 3.2. *1998*

In 1998 *BeppoSAX* observed Mrk 421 as part of a long monitoring campaign involving *BeppoSAX*, *ASCA* (Takahashi, Madejski, & Kubo 1999), *RXTE* (Madejski et al. 2000, in preparation), and coverage from ground-based TeV observatories (Maraschi et al. 1999). The *BeppoSAX* observation comprised two distinct long pointings, started respectively on April 21 and 23. The total *net* exposure time for MECS telescopes has been 64 ks, and the LECS one adds up to 50 ks. Hereafter, we will refer to the two 1998 pointings simply as April 21 and April 23, leaving out the year.

The actual on-source time was 55.5 hr. Unfortunately, celestial and satellite constraints in late 1998 April were such that in each orbit there are two intervals during which the reconstructed attitude is undefined. One occurs when the source is occulted by the dark Earth, but the second, longer one occurs during source visibility periods. Thus, there are about 19 minutes per orbit when the NFIs are pointing at the source, but there is a gap in attitude reconstruction despite the high Earth elevation angle. During these intervals the source position is drifting along the $\pm X$ detector axis, from a position in the center to about 12′ off, passing the strongback window support. These intervals are excluded by the LECS/MECS event file generation. Using the XAS software it is possible to accumulate MECS images also during such intervals, showing that the source is actually drifting. The drift causes a strong energy-dependent modulation, clearly visible in light curves accumulated with XAS and which would be very difficult to recover, as required to build a response matrix integrated from different position dependent matrices weighted according to the time spent in each position. We then did not try to recover these data for the analysis.

In the case of PDS the source remains inside the collimator flat-top during the drift, and therefore we used for the accumulation all events taken when the Earth elevation angle was greater than 3° (inclusive of attitude gaps).

## 4. TEMPORAL ANALYSIS

Light curves have been accumulated for different energy bands. The input photon lists were extracted from the full data set selecting events in a circular region centered on the position of the point source. The extraction radii are 8′ and 6′ for LECS and MECS, respectively, chosen to be large enough to collect most of the photons in the whole energy range:[2] the used values ensure that less than 5% of the photons are missed, at all energies.

For MECS data we considered the merged photon list of all the available MECS units, which were 3 for the 1997 observations and only 2 (namely MECS units 2 and 3) in 1998.

[2] For bright and soft sources—this matters especially because the LECS point-spread function (PSF) gets rapidly broad below the carbon edge, i.e., $E \lesssim 0.3$ keV—it is recommended to select quite large values.

The expected background contribution is less than a fraction of a percent (typically of the order of 0.2%–0.5%), and therefore the light curves have not been background subtracted.

In order to maximize the information on the spectral/temporal behavior, the choice of energy bands shall be such that they are as independent as possible. This should take into account the instrumental efficiencies and, at a certain level, the details of the spectral shape. The main trade-off is the number of photons in each band, which should be large enough to obtain statistically meaningful results. Taking into account the main features of the LECS and MECS effective areas (e.g., the carbon edge at 0.29 keV which provides with independent energy bands below and above this energy) and the steepness of the Mrk 421 spectrum (above a few keV the available photons quickly become insufficient), we determined the following bands: 0.1–0.5 keV and 0.5–2 keV for LECS data and 2–3 and 4–6 keV for MECS data. Their respective "barycentric" energies[3] are 0.26, 1.16, 2.35, and 4.76 keV, respectively, thus providing a factor $\sim$20 leverage, useful to test the energy dependence of the variability characteristics.

### 4.1. *1997*

The four resulting light curves for 1997 are presented in Figure 1. The source shows a high degree of flux variability, with possibly a major flare between the third and fourth pointings. The vertical scale is logarithmic, allowing a direct comparison of the amplitude of variations at different energies. As anticipated we are not going to consider further the

[3] These have been computed *directly from the count spectra*, and thus they are not model dependent. Moreover, despite the observed spectral variability their values do not change more than a few eV.

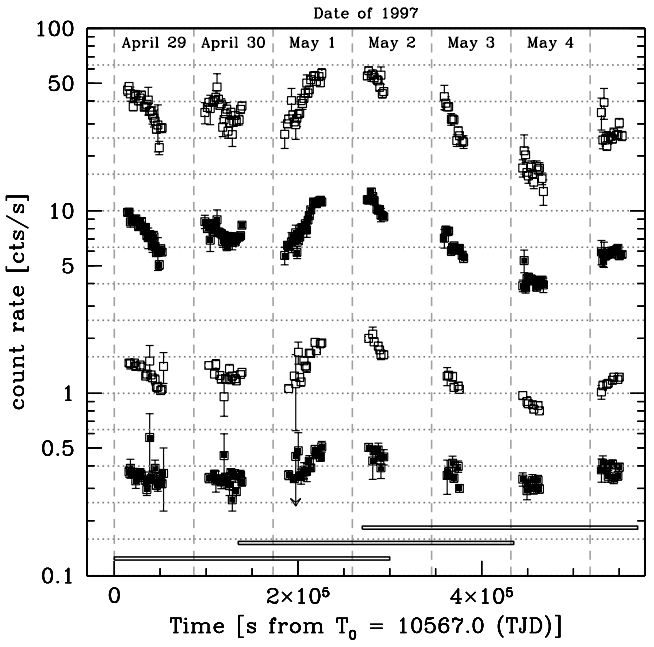


FIG. 1.—Light curves in four energy bands for the 1997 observations. They are shown in order of decreasing energy from top to bottom: 4–6 keV (multiplied by a factor of 100), 2–3 keV (multiplied by a factor of 10), 0.5–2 keV, and 0.1–0.5 keV. The binning time is 1500 s. The reference time TJD = 10,567.0 corresponds to 1997 April 29:00:00 UT. The horizontal bars shown at the bottom of the plot mark the time intervals considered for the analysis of the variability properties in § 4.3: subsets 1 to 3 from left to right.

LECS data of May 7, however for the sake of completeness we report that the count rate was of $0.37 \pm 0.01$ and $1.06 \pm 0.04$ counts $s^{-1}$ for the 0.1–0.5 and 0.5–2.0 keV energy bands, respectively, comparable to the level measured on May 5.

The comparison of data with the overlaid reference grid shows that the amplitude of variability is larger at higher energies, as commonly observed in blazars at energies above the SED peak (e.g., Ulrich, Maraschi, & Urry 1997). We will discuss this issue more quantitatively in § 4.3.

### 4.2. *1998*

The 1998 light curves are shown in Figure 2 (see the caption for details about the axis scales). The overall appearance is somewhat different, being dominated by a single isolated flare at the beginning of the campaign. The variability amplitude is similar to the 1997 one in each corresponding energy band (see also § 4.3).

The most striking and important result of the campaign is that in correspondence with the X-ray flare of April 21 a sharp TeV flare was detected by the Whipple Cerenkov Telescope, with amplitude of a factor 4 and a halving time of about 2 hr (Maraschi et al. 1999). In Figure 3 the Whipple TeV ($E \geq 2$ TeV) light curve is shown together with the *BeppoSAX* ones: LECS (0.1–0.5 keV), MECS (4–6 keV), and PDS (12–26 keV) instruments. The peaks in the 0.1–0.5 keV, 4–6 keV, and 2 TeV light curves are simultaneous *within 1 hr*. Note however that the TeV variation appears to be both larger and faster than the X-ray one. The 12–26 keV light curve shows a broader and slightly delayed peak with respect to lower energy X-rays. However, the very limited statistics does not allow us to better quantify this event. A more detailed account on this particular result is presented and discussed by Maraschi et al. (1999).

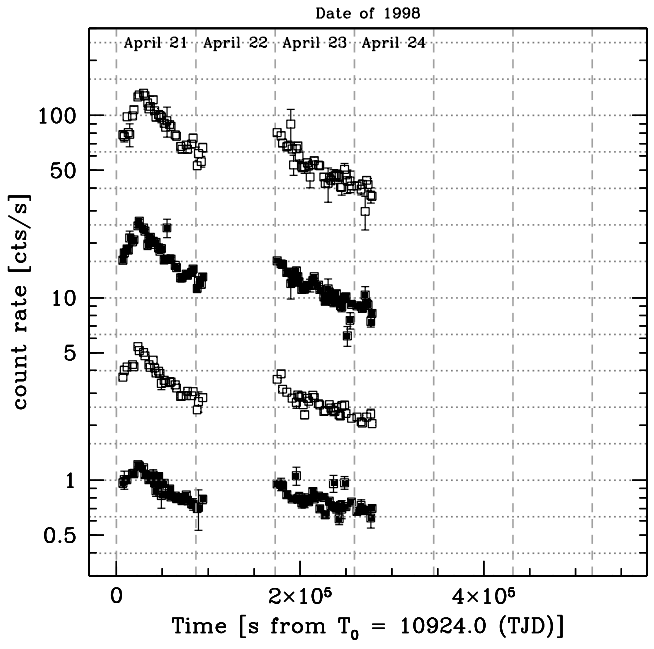


FIG. 2.—Light curves in four energy bands for the 1998 observations. They are shown in order of decreasing energy from top to bottom: 4–6 keV (multiplied by a factor of 100), 2–3 keV (multiplied by a factor of 10), 0.5–2 keV, and 0.1–0.5 keV. The binning time is 1500 s. The reference time TJD = 10,924.0 corresponds to 1998 April 21:00:00 UT. In order to allow an easier comparison with the 1997 light curves, the $X$ scale has been expanded to the same length of that of Fig. 1, and also the $Y$-axis dynamic range is the same.

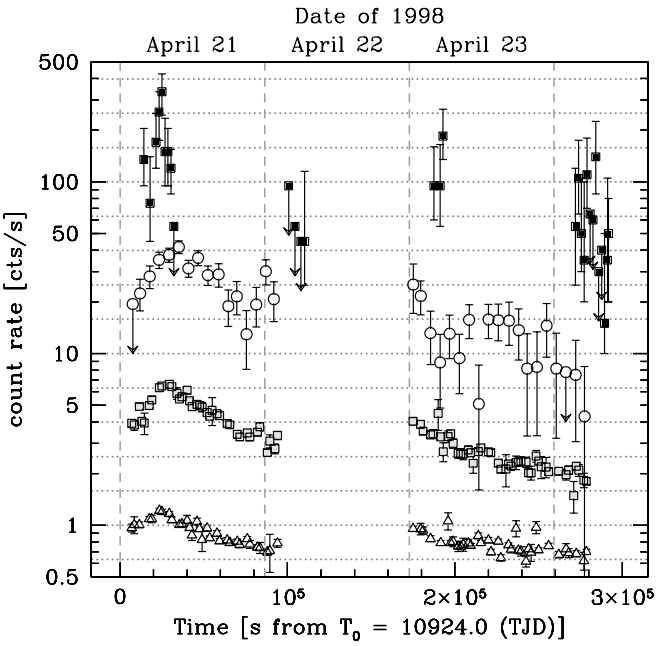


FIG. 3.—Light curves at TeV and X-ray energies during the 1998 campaign. They are shown in order of decreasing energy from top to bottom: Whipple $E \geq 2$ TeV (multiplied by a factor of 500), PDS 12–26 keV (one point per orbit, multiplied by a factor of 50), MECS 4–6 keV (multiplied by a factor 5), and LECS 0.1–0.5 keV (both with 1500 s bins). The count rate units are counts $s^{-1}$ for *BeppoSAX* data and counts $minute^{-1}$ for Whipple data. The thicker dashed grid marks the days, while the thinner ones have a 20 ks spacing. TJD = 10,924.0 corresponds to 1998 April 21:00:00 UT.

It is also worth noticing that the positive (flux) offset of the beginning of April 23 with respect to the end of the April 21 suggests the presence of a second flare occurring between the two pointings. This seems to be also confirmed by the *RXTE* all-sky monitor (ASM) light curve, shown in Figure 4 together with the MECS one in the (2–10 keV) nominal working energy range of the ASM. Single ASM ~90 s dwells have been rebinned in 14,400 s (i.e., 4 hr) intervals,

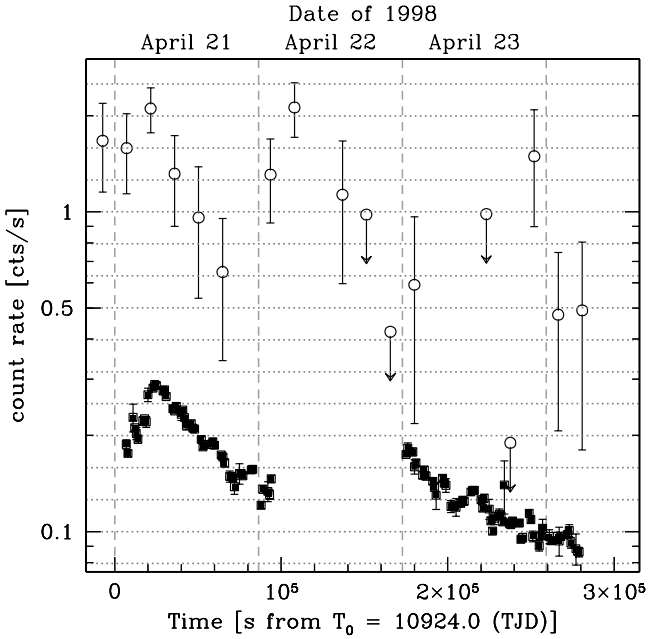


FIG. 4.—*RXTE* all-sky monitor light curve (*top*) compared with the MECS in the same (2–10 keV) band, scaled by a factor of 1/20.

weighting the contribution of each single dwell on its effective exposure time and on the quoted error. The MECS light curve bin is 1000 s. Indeed, the ASM light curve appears to unveil the presence of a second flare occurring in between the two *BeppoSAX* pointings, with a brightness level similar to that of the detected one.

On the other hand the few Whipple data points at the time of this putative second X-ray flare [at $T = (100–120) \times 10^3$ s; see Fig. 3] do not indicate any (major) TeV activity: the count rate measured by the Whipple Telescope was in fact significantly lower than that measured simultaneously with the first X-ray flare.

### 4.3. *Comparison of 1997 and 1998 Variability Characteristics*

We can characterize the variability in different X-ray bands by two commonly used estimators: the fractional *rms* variability parameter $F_{\rm var}$ and the *minimum "doubling/halving time"* $T_{\rm short}$ (for definitions and details see the Appendix and Zhang et al. 1999).

#### 4.3.1. *Fractional rms Variability*

For each band (0.1–0.5, 0.5–2, 2–3, and 4–6 keV), we computed $F_{\rm var}$ for the whole 1997 and 1998 data sets that have the same on-source coverage (i.e., $\sim 55$ hr).

Moreover, for the 1997 we considered three (partially overlapping) subsets of data, each spanning a time interval comparable to the length of the whole 1998 observation, i.e., $\sim 300$ ks. The intervals cover the first 300 ks, a middle section, and the last 300 ks, as marked on Figure 1.

For each data set $F_{\rm var}$ is computed for six different light-curve binning intervals, namely, 200, 500, 1000, 1500, 2000, and 2500 s, and then the weighted average[4] of these values is computed. The results are listed in Table 2 and shown in Figure 5.

If we treat all the data sets as independent and momentarily *do not* consider *subset 1* of the 1997 light curve (Fig. 5, *open triangles*) on which we will comment later, the following results seem to appear:

[4] To the $F_{\rm var}$ values obtained for each different time binning, we assigned as weight $\equiv 1/\sigma^2$, where $\sigma$ is their respective 1 $\sigma$ uncertainty.

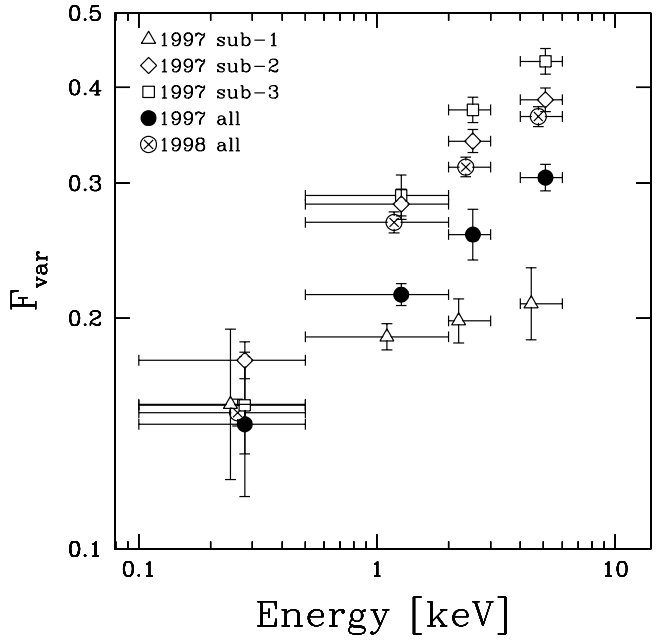


FIG. 5.—Plot of the rms variability parameter, $F_{\rm var}$, vs. energy. A small energy shift has been applied for the sake of clarity.

1. In terms of rms variability the lower energy flux is less variable than the higher energy one, as already pointed out in § 4.1. This holds independent of both the length of the time interval and the state of the source.

2. There is no relation between the brightness level and the variability amplitude since although in 1998 the source was at least a factor of 2 brighter than in 1997 the corresponding $F_{\rm var}$ are not larger.

What is different for subset 1 of the 1997 light curve?

First, two caveats on $F_{\rm var}$. This quantity is basically only sensitive to the average excursion around the mean flux and does not carry any information about either duty cycle or the actual flux excursion. Therefore, in order for $F_{\rm var}$ to meaningfully represent the variability, the "extrema" of the

TABLE 2

BASIC VARIABILITY ESTIMATORS

| Parameter | LECS 0.1–0.5 | LECS 0.5–2 | MECS 2–3 | MECS 4–6 |
|---|---|---|---|---|
| Energy (keV) ..... | 0.26 | 1.18 | 2.36 | 4.76 |
| rms Variability Estimator $F_{\rm var}$ | | | | |
| 1998 all ..... | $0.151 \pm 0.006$ | $0.267 \pm 0.009$ | $0.315 \pm 0.009$ | $0.366 \pm 0.011$ |
| 1997 all ..... | $0.145 \pm 0.035$ | $0.215 \pm 0.007$ | $0.257 \pm 0.020$ | $0.305 \pm 0.012$ |
| 1997 subset 1..... | $0.154 \pm 0.039$ | $0.189 \pm 0.008$ | $0.198 \pm 0.014$ | $0.209 \pm 0.024$ |
| 1997 subset 2..... | $0.176 \pm 0.010$ | $0.282 \pm 0.013$ | $0.340 \pm 0.012$ | $0.385 \pm 0.014$ |
| 1997 subset 3..... | $0.154 \pm 0.024$ | $0.289 \pm 0.018$ | $0.374 \pm 0.015$ | $0.432 \pm 0.017$ |
| Shortest Timescale $T_{\rm short}$ (ks) | | | | |
| 1997: | | | | |
| All data ..... | $80.6 \pm 46.6$ | $28.6 \pm 6.4$ | $25.5 \pm 2.4$ | $25.5 \pm 2.0$ |
| Halving ..... | $167.4 \pm 31.0$ | $62.9 \pm 4.4$ | $31.0 \pm 3.1$ | $32.6 \pm 8.0$ |
| Doubling..... | $81.7 \pm 62.1$ | $28.2 \pm 6.6$ | $27.3 \pm 4.1$ | $28.7 \pm 3.4$ |
| 1998: | | | | |
| All data ..... | $63.8 \pm 12.1$ | $28.1 \pm 4.2$ | $27.3 \pm 2.4$ | $21.4 \pm 1.5$ |
| Halving ..... | $63.8 \pm 12.1$ | $35.7 \pm 4.1$ | $36.9 \pm 8.7$ | $34.5 \pm 2.4$ |
| Doubling..... | ...[a] | $30.4 \pm 12.1$ | $28.4 \pm 3.3$ | $23.6 \pm 1.9$ |

[a] It has not been possible to derive a reliable estimate for this case due to insufficient data.

source variability have to be sampled. This is clearly not the case for the first 1997 sub–light curve (see Fig. 1). Furthermore, $F_{\rm var}$ is estimated from the comparison of the variances of the light curve and the measurements, but while the latter ones probably obey a Gaussian distribution and are much smaller than the former ones, the "probability" for the source to show a significant deviation from the average (i.e., the count rate histogram distribution) is in most cases highly non-Gaussian. For a typical "flaring" light curve this has a "core" at low count rates, resulting from the long(er) time spent by the source in between flares and a very extended tail to high(er) rates. The implicit assumption of Gaussian variability can thus affect the estimate of $F_{\rm var}$ (as for large amplitude and low duty cycle variability).

Thus, $F_{\rm var}$ is a particularly poor indicator especially for observations with a window function like the one of 1997, for which the source temporal covering fraction is of the same order of the single outburst, making likely to miss the flare peak, and the total span of the observation does not significantly exceed the frequency of peaks. We therefore believe this affects the results relative to the first subsample of 1997.

In order to obtain a variability estimator more sensitive to the dynamic range toward higher count rates (flaring), we also considered a modified definition of it, $F_{\rm var,med}$: to study only the width (amplitude) of the brighter tail, only data above the *median* count rate are considered to compute variance and expected variance. Note that the modified definition is still affected by the problem of "representativity" of all the source states. Since the median is always smaller than the average (for all energies, and for all trial time binnings), the values of $F_{\rm var,med}$ are larger than $F_{\rm var}$, but in every other respect the results are qualitatively unchanged.

#### 4.3.2. *Doubling/Halving Timescale*

Similarly, the "*minimum halving/doubling*" timescale has been computed for each energy band taking the average of values for the different input light curve binning (but discarding the 200 s one because too strongly subject to spurious results). As in Zhang et al. (1999) we rejected a $T_{{\rm short},ij}$ if the fractional uncertainty was larger than 20%.

In order to minimize the contamination by isolated data points, we filtered the light curves excluding data lying at more than 3 $\sigma$ from the average computed over the six nearest neighbors (three on each side), and for the 500 s binning light curves we did not consider pairs between data closer than three positions along the time series. Finally, instead of taking the single absolute minimum value of $T_{{\rm short},ij}$ (over all possible $i$, $j$ pairs), $T_{\rm short}$ is determined as the average over the five lowest values.

The resulting $T_{\rm short}$ are listed at the bottom of Table 2 and shown in Figure 6*a*. The main findings are as follows: (1) there is not significant difference between 1997 and 1998 observations, at any energy; (2) there is weak evidence that *the softest X-ray band* exhibits slower variability; while (3) the timescales *for the three higher* energy bands are indistinguishable, and there is no sign of a trend with energy. We also distinguish *doubling* and *halving* timescales. Unfortunately, it has not been possible to obtain an estimate of the *doubling* time for the lowest energy band for 1998 data because there are only a few, uncertain, data during the rise of the flare. We find that (see Table 2 and Figs. 6*b* and 6*c*) (1)

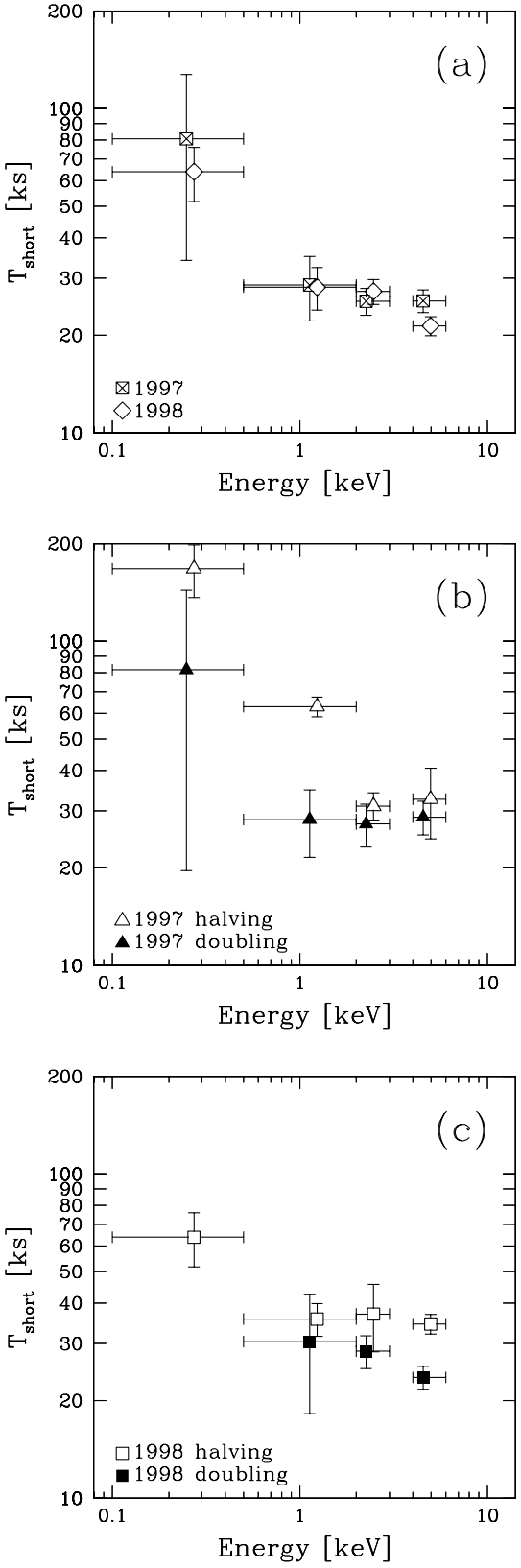


FIG. 6.—(*a*) Comparison of global $T_{\rm short}$ for 1997 (*crossed squares*) and 1998 (*diamonds*), (*b*) comparison of doubling (*filled triangles*) and halving (*open triangles*) timescales for 1997, and (*c*) comparison of doubling (*filled squares*) and halving (*open squares*) timescales for 1998 (see text).

again the softest X-ray band shows longer (halving) timescales and (2) there is marginal evidence that the *doubling* timescale is shorter than the *halving* one (i.e., the rise is faster than the fall). This latter result seems to hold for both 1997 and 1998 light curves and for each single trial binning time, and for each energy (except in five cases out of 35).[5] However, it is strongly weakened by the averaging over the several time binnings because $T_{\rm short}$ slightly increases with the duration of the time bins, thus yielding a larger uncertainty on the average.

## 5. 1998 DETAILED TEMPORAL ANALYSIS

We performed a more detailed analysis of the characteristics of the April 21 flare, in particular, to determine any significant different behavior among different energy bands. There are three main goals: (1) to measure the flare exponential (or power law) decay timescales (or power-law index), (2) to evidence possible differences in the flare rise and decay timescales, and (3) to look for time lags among the different energies.

We considered the energy ranges already adopted in § 4.

### 5.1. *Decay Timescales*

A first estimate of the decay timescales has been presented by Maraschi et al. (1999), where a simple exponential law has been fit to the decaying phase of the light curves and a dependence of the *e*-folding timescale on the energy band has been found. Here we analyze the issue of determining the timescale in more detail, and in particular we consider the requirement for the presence of an underlying steady emission.

In particular, the decay timescales for different energy bands have been estimated by fitting the postflare light curves with an exponential decay superimposed to a constant (quasi-steady) flux eventually constrained by the fit (we immediately stress that an additional nonzero steady emission is necessary in order to obtain a meaningful fit with an exponential decay). In one set of fits the underlying contribution is set to zero. More precisely, we performed the fits with three possible conditions for the level of the steady contribution: unconstrained, constrained, and zero.

The analytical expression of the function used in the fits is

$$F(t) = F_{\rm steady} + F_{\rm flaring}\, e^{-(t-T_{\rm ref})/\tau} = F_{\rm steady}(1 + \mathcal{R} e^{-(t-T_{\rm ref})/\tau}) .$$

The model parameters are then the decay timescale $\tau$, the absolute value of $F_{\rm steady}$, and the ratio $\mathcal{R}$ between the flaring and the steady component taken at a reference time $T_{\rm ref}$, which ideally should correspond to the peak of the flare.

The "reference time" $T_{\rm ref}$ for the April 21 flare was $T_{\rm ref} = 25 \times 10^3$ s ($T = 0 = 10{,}924.0$ TJD), and $T_{\rm ref} = 170 \times 10^3$ s for April 23. Note that the value of $\mathcal{R}$ obtained from the fit to the second data set is only a lower limit on the actual flare amplitude since we do not have an estimate of the time at which the peak of this putative second flare occurred.

In any case, as our focus is actually on the properties of the first, well-defined outburst, we used the parameters obtained from the April 23 data only to constrain the contribution from the steady component. Furthermore, it turns out that the determination of the timescale $\tau$ is not affected by the uncertainty on its contribution (see Table 3 and Fig. 8).

We adopted a 1000 s binning for the 0.5–2 keV and 2–3 keV data, which have better statistics, while the binning time for the 0.1–0.5 keV and 4–6 keV light curves is 2000 and 1500 s, respectively. The analysis has been performed in two stages, for each energy band:

1. Fits were made to the April 21 and 23 data sets independently, yielding values of $\tau$, $F_{\rm steady}$, and $\mathcal{R}$. As an example, in Figure 7 the 2–3 keV light curve is shown together with its best-fit models for April 21 and 23.
2. A fit was made of the April 21 postflare phase constraining the level of the underlying steady component to be below the level of April 23 (the actual constraint is the upper bracket for a two-parameter 90% confidence interval). We assume that the underlying emission varies on a timescale longer than that spanned by the observations.

As anticipated we also modeled the decay of the April 21 flare setting to zero the offset.

The results are summarized in Tables 3 and 4. The best-fit $\tau$ for all three cases are plotted in Figure 8 as a function of energy.

The 0.1–0.5 keV timescale is not affected by the constraint on the baseline flux, and the 4–6 keV confidence interval only suffers a minor cut (smaller error bar), while both the 0.5–2 and 2–3 keV parameters are significantly changed because the best-fit unconstrained level of the steady contribution is significantly higher than allowed by the April 23 data (see, for instance, Fig. 7).

The main findings are as follows:

1. The decay timescales depend critically on the presence/absence of a contribution by a nonvariable com-

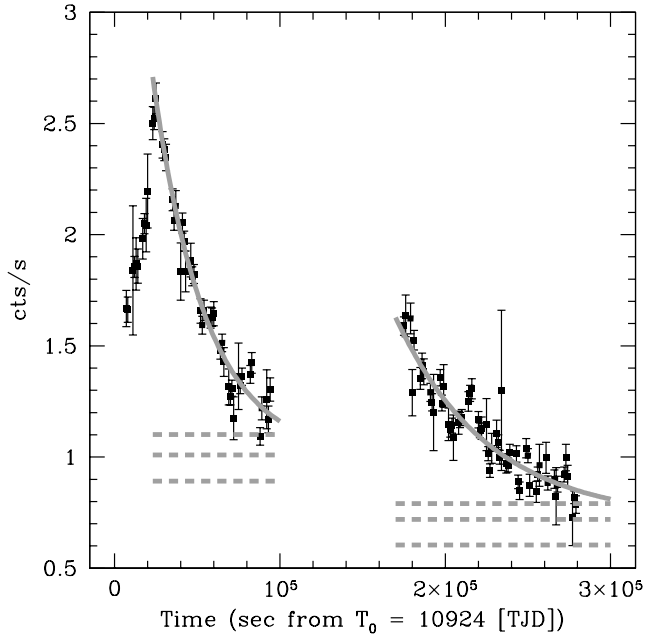


FIG. 7.—Example of the fit of the light curve with a constant + exponential decay. The 2–3 keV data are shown, with a 1000 s time bin, together with the best-fit model (solid line through the data points) and the best-fit level of the steady component (three dashed lines corresponding to the best-fit value and the 90% confidence interval for two interesting parameters).

[5] We stress here that although the individual timescales are statistically consistent with each other the behaviors appear to be systematic at all energies.

TABLE 3

PARAMETERS OF FIT TO LIGHT CURVES

| Parameter | LECS 0.1–0.5 | LECS 0.5–2 | MECS 2–3 | MECS 4–6 |
|---|---|---|---|---|
| **Constant + Exponential Decay** | | | | |
| April 21, zero baseline | | | | |
| $\tau$[a] (ks) | $125.3^{+8.6}_{-5.8}$ | $91.8^{+1.9}_{-1.5}$ | $83.8^{+1.8}_{-2.2}$ | $74.4^{+3.1}_{-2.3}$ |
| April 21, unconstrained | | | | |
| $\mathscr{R}$[b] | $0.77^{+0.24}_{-0.13}$ | $1.38^{+0.19}_{-0.15}$ | $1.58^{+0.22}_{-0.16}$ | $2.97^{+3.01}_{-0.77}$ |
| $F_{\rm steady}$ (counts s$^{-1}$) | $0.68^{+0.05}_{-0.08}$ | $2.23^{+0.12}_{-0.18}$ | $1.01^{+0.07}_{-0.09}$ | $0.39^{+0.09}_{-0.12}$ |
| $\tau$[a] (ks) | $29.9^{+11.4}_{-6.2}$ | $34.4^{+4.5}_{-3.2}$ | $32.1^{+4.3}_{-4.0}$ | $40.7^{+11.1}_{-7.4}$ |
| April 21, constrained | | | | |
| $\mathscr{R}$[b] | $0.77^{+0.24}_{-0.13}$ | $1.66^{+0.11}_{-0.04}$ | $2.19^{+0.16}_{-0.04}$ | $2.97^{+3.01}_{-0.49}$ |
| $F_{\rm steady}$[c] (counts s$^{-1}$) | $0.67_{-0.07}$ | $1.94_{-0.07}$ | $0.79_{-0.03}$ | $0.39_{-0.12}$ |
| | [<0.72] | [<1.94] | [<0.79] | [<0.39] |
| $\tau$[a] (ks) | $30.7^{+10.6}_{-6.9}$ | $41.7^{+1.8}_{-1.2}$ | $43.0^{+2.0}_{-1.7}$ | $41.2^{+10.6}_{-1.9}$ |
| April 23, unconstrained | | | | |
| $\mathscr{R}$[b] | $0.38^{+0.09}_{-0.07}$ | $0.95^{+0.27}_{-0.13}$ | $1.25^{+0.24}_{-0.16}$ | $1.35^{+0.17}_{-0.11}$ |
| $F_{\rm steady}$ (counts s$^{-1}$) | $0.70^{+0.02}_{-0.04}$ | $1.77^{+0.16}_{-0.24}$ | $0.72^{+0.06}_{-0.08}$ | $0.36^{+0.03}_{-0.03}$ |
| $\tau$[a] (ks) | $29.0^{+22.0}_{-10.3}$ | $66.6^{+21.1}_{-13.3}$ | $55.6^{+14.0}_{-9.1}$ | $39.2^{+8.6}_{-6.0}$ |
| **Constant + Power-Law Decay** | | | | |
| $\mathscr{R}$[b] | >1.715 | >15.15 | >23.04 | >39.37 |
| $F_{\rm steady}$[c] (counts s$^{-1}$) | <0.600 | <0.619 | <0.489 | <0.175 |
| $\eta$[d] | $0.54^{+0.37}_{-0.18}$ | $0.54^{+0.05}_{-0.02}$ | $0.62^{+0.16}_{-0.02}$ | $0.70^{+0.09}_{-0.03}$ |

NOTES.—In the *unconstrained* case the flux level of the steady component is left free to adjust during the fit. In the *constrained* case its flux level is required to be less or equal to the upper value of the 90% confidence interval obtained by the fit of the April 23 light curve in the corresponding energy band. The values reported within square brackets represent the imposed upper limits on $F_{\rm steady}$ (we do not quote the upper error value because in all four cases the upper boundary of the confidence interval simply coincides with the imposed upper limit). Quoted errors are for a 1 $\sigma$ confidence interval for one or two interesting parameters (i.e., $\Delta\chi^2 = 1.0$ or 2.3) for the case *without* and *with* the baseline, respectively.

We do not list the $\chi^2$ values because they are not really meaningful estimates of the goodness of the fit. This happens because the errors on each bin of the light curves are very small and there is significant variability on very short timescales, which of course is not accounted for by the simple exponential decay and eventually it has the effect of giving an artificially high $\chi^2$ value. For the fits of April 21, the values of reduced $\chi^2$ are between 1.6 and 4.1 for the cases including the baseline (for both exponential and power-law decays) and between 2.6 and 6.5 for the pure exponential (zero baseline) case. The number of degrees of freedom (for April 21 data sets) is between 20 and 35, depending on the energy band.

[a] Timescale of the exponentially decaying flaring component.

[b] Ratio between the flaring and steady components at $T_{\rm ref}$ (see text).

[c] The 90% confidence upper limit, for $\Delta\chi^2 = 4.61$, for the level of the steady component.

[d] Power-law decay spectral index.

ponent. This is true not only for the values themselves (for the cases *without* baseline are between a factor of 2 and 4 longer) but also for the relationship between the timescales and energy. In fact, further results are as follows:

*a*) In the cases with baseline, the timescales range between 30 and $45 \times 10^3$ s and *do not* show a clear (if any) relationship with the energy, rather suggesting an achromatic postflare evolution. In fact, according to a $\chi^2$ test, the values of $\tau$ for the four energies are consistent with coming from the same distribution (the so-called null hypothesis).

*b*) On the contrary, in the case of pure exponential decay the timescales follow a weak inverse relationship with the energy, $\tau \propto E^{-0.18\pm0.02}$ (both when considering $\tau$ or $E$ as the dependent variable). The $\chi^2$ probability in favor of the "null hypothesis" is in this case negligible.

2. The amplitude of the variability (as measured by $\mathscr{R}$) positively correlates with the energy, as reported in many cases for HBLs (e.g., Sambruna et al. 1994; Ulrich et al. 1997).

#### 5.1.1. *Is There a Steady Component?*

Variability timescales and their energy dependence carry precious information on the physics of the source: the puzzling dependence of the results on the hypothesis of the presence of a steady component requires further analysis.

Here we only discuss a simple test of the robustness of the results of the previous section that, however, provides useful hints. This consists in determining $\tau$ for different choices of the end time ($T_{\rm end}$) of the interval over which the fit is performed. The underlying idea is that if the shape of the decay is close to the one we are assuming (exponential with/without an offset) one can expect that the model parameters do not change much as a function of $T_{\rm end}$. On the contrary, if the chosen model provides only a poor description of the actual decay characteristics, the parameters will likely take different values depending on $T_{\rm end}$, unless of course we postulate a process whose timescale changes with time.

For each of the three cases (constrained, unconstrained, zero baseline), we tried five different choices of the end time, namely, $T_{\rm end} = 60$, 70, 80, 90, and $100 \times 10^3$ s (from $T_0$), the

TABLE 4

TIMESCALES FOR DIFFERENT EXPONENTIAL DECAY FITS

| Parameter | LECS 0.1–0.5 | LECS 0.5–2 | MECS 2–3 | MECS 4–6 |
|---|---|---|---|---|
| | | April 21, Zero Baseline | | |
| $T_{end} = 100 \times 10^3$..... | $125.3^{+8.6}_{-5.8}$ | $91.8^{+1.9}_{-1.5}$ | $83.8^{+1.8}_{-2.2}$ | $74.4^{+3.1}_{-2.3}$ |
| $T_{end} = 90 \times 10^3$ ..... | $120.3^{+8.2}_{-5.7}$ | $89.0^{+2.0}_{-1.5}$ | $80.3^{+2.4}_{-1.7}$ | $72.0^{+3.7}_{-2.0}$ |
| $T_{end} = 80 \times 10^3$ ..... | $116.2^{+9.0}_{-6.1}$ | $84.9^{+2.0}_{-1.7}$ | $73.8^{+2.6}_{-1.8}$ | $68.5^{+3.8}_{-2.4}$ |
| $T_{end} = 70 \times 10^3$ ..... | $106.2^{+7.8}_{-5.9}$ | $80.7^{+2.0}_{-1.9}$ | $70.8^{+2.5}_{-2.1}$ | $68.5^{+3.4}_{-3.0}$ |
| $T_{end} = 60 \times 10^3$ ..... | $95.3^{+9.5}_{-6.6}$ | $76.0^{+2.7}_{-2.1}$ | $70.0^{+3.6}_{-2.8}$ | $70.3^{+5.9}_{-4.1}$ |
| $\mathscr{P}$[a] ..... | $4 \times 10^{-2}$ | $6 \times 10^{-8}$ | $1 \times 10^{-5}$ | ... |
| | | April 21, Unconstrained Baseline | | |
| $T_{end} = 100 \times 10^3$..... | $29.9^{+11.4}_{-6.2}$ | $34.4^{+4.5}_{-3.2}$ | $32.1^{+4.3}_{-4.0}$ | $40.7^{+11.1}_{-7.4}$ |
| $T_{end} = 90 \times 10^3$ ..... | $33.8^{+14.8}_{-8.4}$ | $37.7^{+5.5}_{-4.1}$ | $34.0^{+5.6}_{-4.7}$ | $46.3^{+16.5}_{-10.4}$ |
| $T_{end} = 80 \times 10^3$ ..... | $26.9^{+10.9}_{-7.1}$ | $34.5^{+5.6}_{-4.8}$ | $38.6^{+10.4}_{-7.5}$ | $53.6^{+19.6}_{-16.9}$ |
| $T_{end} = 70 \times 10^3$ ..... | $32.3^{+24.0}_{-11.1}$ | $34.6^{+9.7}_{-5.7}$ | $45.6^{+26.5}_{-12.5}$ | $61.7^{+12.0}_{-26.3}$ |
| $T_{end} = 60 \times 10^3$ ..... | $32.6^{+70.4}_{-14.3}$ | $26.3^{+8.0}_{-5.0}$ | $24.2^{+11.2}_{-5.9}$ | $18.9^{+14.4}_{-6.2}$ |
| $\mathscr{P}$[a] ..... | ... | ... | ... | ... |
| | | April 21, Constrained Baseline | | |
| $T_{end} = 100 \times 10^3$..... | $30.7^{+10.6}_{-6.9}$ | $41.7^{+1.8}_{-1.2}$ | $43.0^{+2.0}_{-1.7}$ | $41.2^{+10.6}_{-1.9}$ |
| $T_{end} = 90 \times 10^3$ ..... | $33.3^{+15.1}_{-8.0}$ | $41.2^{+3.4}_{-1.3}$ | $42.0^{+3.0}_{-1.5}$ | $46.2^{+16.6}_{-7.2}$ |
| $T_{end} = 80 \times 10^3$ ..... | $26.6^{+11.3}_{-3.4}$ | $40.8^{+3.3}_{-1.2}$ | $40.6^{+8.9}_{-1.7}$ | $53.1^{+19.4}_{-14.3}$ |
| $T_{end} = 70 \times 10^3$ ..... | $31.9^{+24.4}_{-8.8}$ | $40.5^{+6.1}_{-1.4}$ | $45.5^{+26.2}_{-7.1}$ | $63.5^{+9.3}_{-24.4}$ |
| $T_{end} = 60 \times 10^3$ ..... | $32.7^{+69.3}_{-9.4}$ | $40.2^{+5.0}_{-1.8}$ | $41.5^{+9.6}_{-2.4}$ | $44.3^{+23.0}_{-3.9}$ |
| $\mathscr{P}$[a] ..... | ... | ... | ... | ... |

NOTES.—The decay of the April 21 flare is modeled with an exponential function over a nonvariable flux contribution. The listed numbers are the characteristic $\tau$ values (expressed in ks) of the exponential decay, as a function of energy and $T_{end}$. Quoted errors are for a 1 $\sigma$ confidence interval for one or two interesting parameters for the case *without* and *with* the baseline, respectively. The timescales for the $T_{end} = 100$ ks case are the same as Table 3.

[a] Probability that the values of the timescale $\tau$ obtained for the different choices of $T_{end}$ are drawn from a unique parent distribution ($\chi^2$ with respect to the weighted average) for those cases where there is a significant rejection of the "null hypothesis."

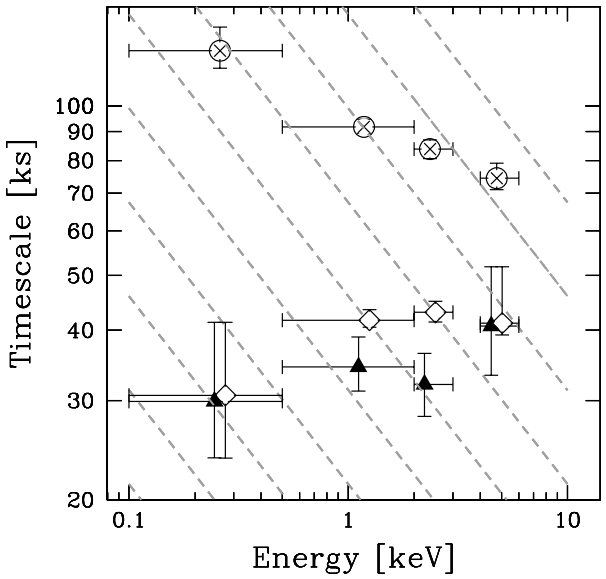


FIG. 8.—Decay timescales $\tau$ plotted vs. the barycentric energy of the data from which they have been derived. The different symbols are crossed circles for the zero baseline, filled triangles for the unconstrained baseline, and diamonds for the constrained baseline case. The dashed lines show the loci of the relation $\tau \sim E^{-1/2}$ that holds if pure radiative cooling is responsible for the spectral variability during the burst decay.

last one corresponding to the entire duration of the data coverage on April 21.

To evaluate the likelihood that the five values of $\tau$ obtained for these choices of $T_{end}$ are drawn from the same (normal) distribution, we considered the probability $\mathscr{P}$ of the $\chi^2$ computed with respect to their weighted[6] average.[7]

In Table 4 we summarize the results obtained under each of the different hypotheses. There are a few points to note:

1. In the *zero baseline* hypothesis, the timescales obtained changing $T_{end}$ are *not consistent* with being several different realizations from a unique distribution. Furthermore, the clear trend of (decreasing) $\tau$ with (decreasing) $T_{end}$ supports the possibility that the variation of $\tau$ is significant.

2. On the contrary, in both cases *with* baseline the values of $\tau$ for the 0.1–0.5, 0.5–2, and 2–3 keV bands are very close for all values of $T_{end}$. For the 4–6 keV band the scatter is somewhat larger, likely due to the fact that at this energy there is the largest amplitude of variability, i.e., the smallest contribution of the putative baseline. The same kind of

[6] To each measure we assigned as $\sigma$ (and in turn as weight $\equiv 1/\sigma^2$), the mean between its plus and minus uncertainties.

[7] It should be noted that the weighted average has the property of minimizing the $\chi^2$ for a data set, and thus the resulting values of probability that the data do belong to a unique distribution are upper limits.

effect of lowered sensitivity to the details of a putative baseline is responsible for the good result of the 4–6 keV band in the case *without* baseline.

3. Even in the case *without* baseline, the values of $\tau$ for the various energy bands become very similar for $T_{\rm end} = 60 \times 10^3$ s. Therefore, even if the failure of the test by the pure exponential decay model could be interpreted as the signature of a mechanism whose characteristic timescale changes with time (unlike a true exponential decay), there is anyway an indication that in the earlier stages of the decay the evolution is not energy dependent.

On the basis of this simple consistency check we can therefore infer that *if the decay were exponential* there is evidence for the presence of an underlying nonvariable component. Thus, the postflare phase does not show any dependence on energy, and the timescale is of the order of 30–40 ks.

### 5.2. *Power-Law Decay*

As an alternative description of the decay of the April 21 flare, we consider a power law, where $F_{\rm flaring} \propto (T/T_{\rm ref})^{-\eta}$, again leaving the possibility of an offset from a steady component. The results are summarized in Table 3 and Figure 9.

A fundamental difference with respect to the exponential decay case is that an additional steady component is never required: the power-law fits yield only (not very stringent) upper limits on $F_{\rm steady}$, and only for the 0.1–0.5 keV band is the formal best-fit value for $F_{\rm steady}$ different from zero ($F_{\rm steady} = 0.26$ counts s$^{-1}$).

Again, there is only a weak relationship of the decay properties with energy, with a marginal indication that the flux decrease is faster at higher energies. The values for the four energy bands are consistent with a common (weighted average) value of $\eta = 0.583$, while for the subset comprising only the three higher energy data sets the probability for the "null hypothesis" is reduced to about 0.07.

### 5.3. *Rise versus Fall*

A further interesting question is whether the rise and the decay phases of the flare are characterized by the same timescale (i.e., if the flare is symmetric or asymmetric) and, in particular, how these properties might correlate with energy.

The observational coverage of the rise is not good enough to apply the same direct technique used for the decay (§ 5.1), to constrain the timescale of an exponential rise, with all the parameters left free. Thus, the flux level of the steady component is fixed at the best-fit value obtained for the decay or set to 0.

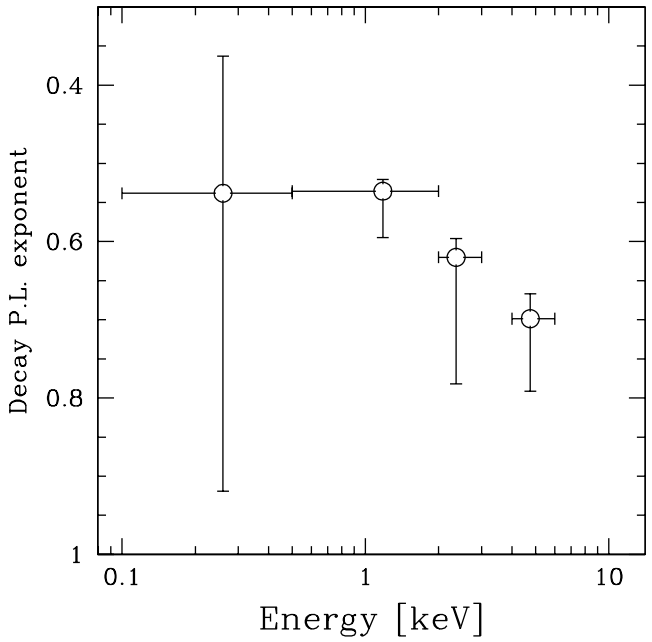


FIG. 9.—Index of the power-law decay plotted vs. energy. Only the value for the April 21 postpeak decay is plotted.

We are mainly interested in the comparison of the properties of the outburst just before and after the top of the flare (see discussion in Paper II), and the interval of about 17–18 ks preceding the flare is long enough for this purpose.

We thus focused on the "tip" of the flare, considering data on the decay side only up to the time when the flux reaches approximately the level at which the observation starts. The duration of the postpeak intervals needed to decrease the flux to the preflare level are of about 15–20, 25, 30–35, and 40 ks for the four energy bands. Since the uncertainties on the decay time of the light curves are not larger than 5 ks, the differences in duration and the trend (longer time at higher energy) are real.

The differences between rise and fall timescales, estimated with fits to the light curves, are shown in Table 5 and plotted in Figure 10. The flare is symmetric for the softest energy band, while it is definitely asymmetric in the other three bands, with an asymmetry that appears to increase systematically with energy: at higher energies the rise phase

TABLE 5

DIFFERENCE BETWEEN RISE AND FALL TIMESCALES

| Case[a] | LECS 0.1–0.5 | LECS 0.5–2 | MECS 2–3 | MECS 4–6 | $\mathcal{P}$[b] |
|---|---|---|---|---|---|
| $F_{\rm steady} = 0$ | $2.9^{+33.4}_{-12.3}$ | $26.3^{+15.0}_{-3.4}$ | $33.1^{+11.2}_{-2.5}$ | $38.6^{+4.8}_{-3.4}$ | ... |
| $F_{\rm steady}$ best-fit unconstrained | $1.2^{+6.9}_{-4.3}$ | $13.7^{+4.3}_{-1.9}$ | $14.8^{+7.9}_{-1.3}$ | $23.6^{+8.5}_{-1.9}$ | $3 \times 10^{-2}$ |
| $F_{\rm steady}$ best-fit constrained | $1.3^{+6.2}_{-4.3}$ | $16.1^{+6.3}_{-2.1}$ | $19.8^{+10.9}_{-1.6}$ | $24.6^{+9.1}_{-2.2}$ | $1 \times 10^{-2}$ |

NOTES.—The values of $\Delta T = T_{\rm fall} - T_{\rm rise}$ have been computed for two different binning of the light curves, 500 and 1000 s, and the average is shown (in ks). Confidence intervals are 1 $\sigma$ for one interesting parameter because the baseline was not free to vary. The upper error bars (giving the largest possible value of $\Delta T$) derive from the upper boundary of the $T_{\rm fall}$ confidence interval and the lower boundary of the $T_{\rm rise}$ confidence interval. The lower error bars (giving the lowest possible value of $\Delta T$) derive from the lower boundary of the $T_{\rm fall}$ confidence interval and the upper boundary of the $T_{\rm rise}$ confidence interval.

[a] In each case the level of the steady component was kept fixed at the best-fit value for each energy band, as listed in Table 3.

[b] Probability that the four values of $\Delta T$ are drawn from the same parent distribution, computed using the $\chi^2$ of the data with respect to their weighted average.

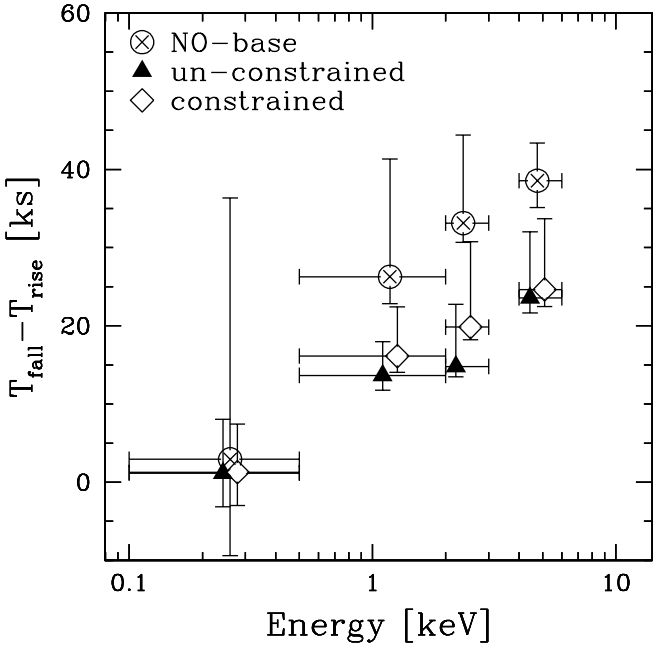


FIG. 10.—Difference between the postpeak decay and the prepeak rise timescales, plotted vs. energy. The different symbols are crossed circles for the $F_{steady} = 0$ case, filled triangles for the $F_{steady}$ set at the best-fit value of the unconstrained case, and diamonds for the $F_{steady}$ set at the best-fit value obtained for the April 23 data (see Table 3).

is increasingly faster than the decay (the significance is reported in Table 5).

It is worth stressing that these numbers indicate only relative changes in the timescales before and after the peak of the flare. On the other hand, since we found that the decay timescales are very similar once the baseline flux is taken into account (§ 5.1), this asymmetry probably means that the rise times get shorter with increasing energy.

### 5.4. *Time Lag*

We performed a detailed cross-correlation analysis using two different techniques suited to unevenly sampled time series: the discrete correlation function (DCF) (Edelson & Krolik 1988) and the modified mean deviation (MMD) (Hufnagel & Bregman 1992). Moreover, Monte Carlo simulations, taking into account "flux randomization" (FR) and "random subset selection" (RSS) of the data series (see Peterson et al. 1998), were used to statistically determine the significance of the time lags between different X-ray bands obtained with the DCF and MMD. We refer for the relevant details of such analysis to Zhang et al. (1999).

We binned the light curves over 300 s in the 0.1–1.5 and 3.5–10 keV bands (whose effective barycentric energies are $E \simeq 0.8$ and $E \simeq 5$ keV), and a trial time step of 720 s is adopted for both the DCF and MMD. DCF amplitude versus time lag is plotted in Figure 11. A negative value means that variations in the 3.5–10 keV band light curve lag those occurring in the 0.1–1.5 keV one (i.e., hard lag). The best Gaussian fits for both DCF and MMD result in negative time lags of $-2.8 \pm 0.2$ (DCF) and $-2.1 \pm 0.3$ (MMD) ks, indicating that the medium-energy X-ray photons lag the soft ones.

The cross-correlation peak distribution (CCPD) obtained from the FR/RSS Monte Carlo simulations is shown in Figure 12 for both the DCF and the MMD methods. The average lags resulting from the CCPD are $-2.7^{+1.9}_{-1.2}$ ks for

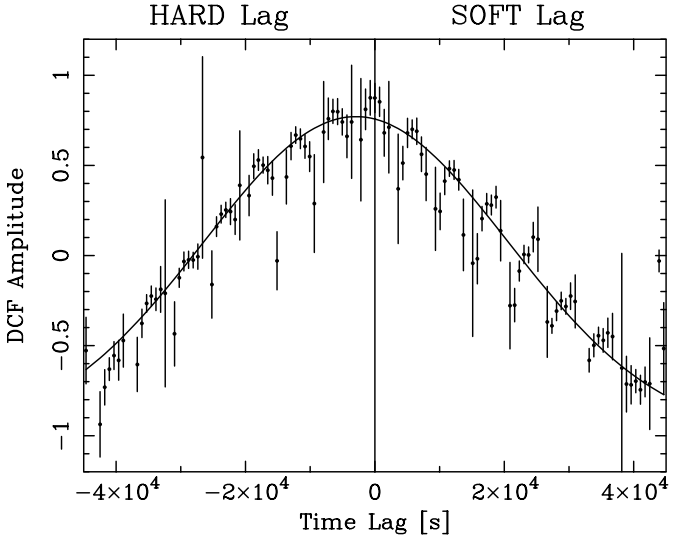


FIG. 11.—Result of the cross-correlation DCF method applied to the 0.1–1.5 and 3.5–10 keV light curves. Negative values correspond to the 3.5–10 keV band lagging the 0.1–1.5 keV one. The continuous curve superimposed to the data is the best fit with a constant + Gaussian model. The actual fit has been performed only to the "core" of the data, i.e., taking into account the data in the interval ranging from −30,000 to 30,000 s.

DCF and $-2.3^{+1.2}_{-0.7}$ ks for MMD (1 $\sigma$ confidence intervals, two-sided with respect to the average), confirming the significance of the above results with high (>90%) confidence.

The total integral probabilities for a negative lag (that would be the actual measure of the confidence of the "discovery" of the hard lag) are ≃95.0% (DCF) and ≃98.7% (MMD).

## 6. SUMMARY OF VARIABILITY PROPERTIES

We presented a comprehensive temporal analysis of the flux variability characteristics in several energy bands of

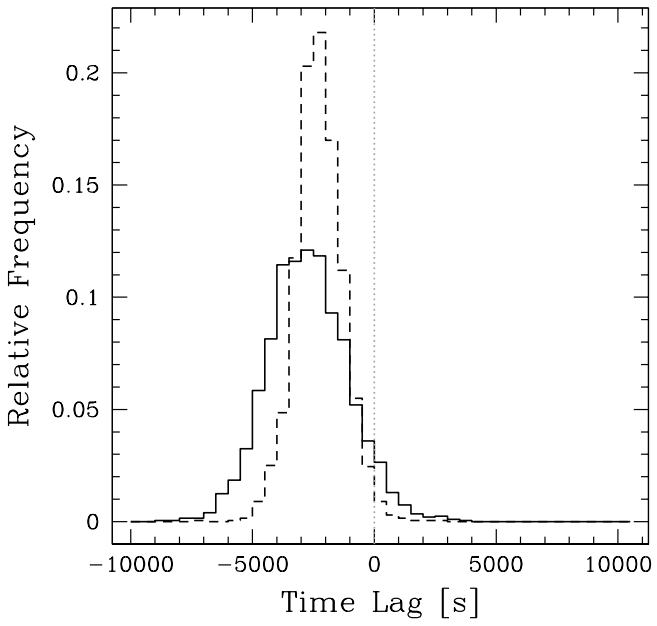


FIG. 12.—Cross-correlation peak distributions for the FR/RSS simulations. The solid line is for the analysis with DCF technique and the dashed for the case of MMD. Negative values correspond to the 3.5–10 keV band lagging the 0.1–1.5 keV one.

*BeppoSAX* observations of Mrk 421. The primary results of this study are the following:

1. The fractional *rms* variability is higher at higher energies.
2. The fractional *rms* variability does not change with the brightness state of the source.
3. The *minimum halving/doubling time* is longer for the softest energy band.
4. The *minimum halving/doubling time* at a given energy does not change with the brightness state of the source.
5. There is a hint that the *doubling* timescale is shorter than the *halving* timescale.

Findings 1 and 2, on $F_{\rm var}$, confirm the well-known (although not fully explained) phenomenology of HBLs (see Ulrich et al. 1997 and references therein).

The results on the variability timescales (findings 3, 4, 5) show good agreement with those found with the more thorough analysis of the 1998 flare, which can be summarized as follows:

6. The flare decay is consistent with being achromatic, both if modeled as an exponential decay with an additional contribution from a steady component and in the case of a power-law decay.
7. The results on timescales are at odds with the simplest possibility of interpreting the decay phase, i.e., that it is driven by the radiative cooling of the emitting electrons. This would give rise, in the simplest case, to a dependence of the timescale on energy, $\tau \sim E^{-1/2}$. The tracks corresponding to this relationship are overlaid to the data in Figure 8, and it is clear that they cannot be reconciled.
8. The harder X-ray photons lag the soft ones, with a delay of the order of few ks. This finding is *opposite* to what is commonly found in the best monitored HBL where the soft X-rays lag the hard ones (e.g., Takahashi et al. 1996 for Mrk 421; Urry et al. 1993, Zhang et al. 1999, and Kataoka et al. 2000 for PKS 2155−304; Kohmura et al. 1994 for H0323+022). The latter behavior is usually interpreted in terms of cooling of the synchrotron-emitting particles.
9. Possible asymmetry in the rise/decay phases: the flare seems to be symmetric at the energies corresponding (roughly) to the peak of the synchrotron component (see the spectral analysis in Paper II), while it might have a faster rise at higher energies. This could be connected to the observed hard lag.
10. Finally, the characteristics of the April 21 flare suggest the presence of a quasi-stationary emission contribution, which seems to be dominated by a highly variable peaked spectrum.

## 7. CONCLUSIONS

*BeppoSAX* has observed Mrk 421 in 1997 and 1998. We analyzed and interpreted the combined spectral and temporal evolution in the X-ray range. During these observations the source has shown a large variety of behaviors, providing us a great wealth of information, but at the same time revealing a richer than expected phenomenology.

In this paper we have presented the first part of the analysis, focused on the study of the variability properties.

The fact that the higher energy band lags the softer one (with a delay of the order of 2–3 ks) and the energy dependence of the shape of the light curve during the flare (with faster flare rise time at higher energies) provide strong constraints on any possible time-dependent particle acceleration mechanism. In particular, if we are indeed observing the first direct signature of the ongoing particle acceleration, progressively "pumping" electrons from lower to higher energies, the measure of the delay between the peaks of the light curves at the corresponding emitted frequencies would provide a tight constraint on the timescale of the acceleration process.

The decomposition of the observed spectrum into two components (a quasi-stationary one and a peaked, highly variable one) might allow us to determine the nature and modality of the energy dissipation in relativistic jets.

In Paper II we complement these findings with those of the time-resolved spectral analysis and develop a scenario to interpret the complex spectral and temporal phenomenology.

We are grateful to the *BeppoSAX* Science Data Center (SDC) for their invaluable work and for providing standardized product data archive, and to the *Rossi*XTE ASM team. We thank Gianpiero Tagliaferri and Paola Grandi for their contribution to our successful *BeppoSAX* program and for useful comments and the anonymous referee for useful suggestions that have improved the clarity of the paper. A. C., M. C., and Y. H. Z. acknowledge the Italian MURST for financial support. This research was supported in part by the National Science Foundation under grant PHY 94-07194 (A. C.). Finally, G. F. thanks Cecilia Clementi for providing tireless stimulus.

## APPENDIX

## DEFINITIONS OF $F_{\rm var}$ AND $T_{\rm short}$

The fractional rms variability amplitude is a useful parameter to characterize the variability in unevenly sampled light curves. It is defined as the square root of the so-called excess variance (e.g., Nandra et al. 1997). This parameter, also known as the true variance (Done et al. 1992), is computed by taking the difference between the variance of the overall light curve and the variance due to measurement error, normalized by dividing by the average squared flux (count rate).

We consider a data set $F_i(t_i)$ $(i = 1, N)$, with an uncertainty $\sigma_i$ assigned to each point.

The fractional rms variability parameter is then defined as

$$F_{\rm var} \equiv \frac{(\sigma_F^2 - \Delta_F^2)^{1/2}}{\langle F \rangle}, \tag{A1}$$

where

$$\sigma_F^2 \equiv \frac{1}{N-1} \sum_{i=1}^{N} (F_i - \langle F \rangle)^2 , \tag{A2}$$

$$\Delta_F^2 \equiv \frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 . \tag{A3}$$

The definition of the *shortest variability timescale* is the following:

$$T_{\rm short} \equiv \min_{\forall i,j} (T_{{\rm short},ij}) , \tag{A4}$$

$$T_{{\rm short},ij} \equiv \left| \frac{F_{ij} \Delta T_{ij}}{\Delta F_{ij}} \right| , \tag{A5}$$

where $F_{ij} \equiv (F_i + F_j)/2$, $\Delta F_{ij} \equiv F_i - F_j$, and $\Delta T_{ij} \equiv T_i - T_j$.

## REFERENCES


Boella, G., Butler, R. C., Perola, G. C., Piro, L., Scarsi, L., & Bleeker, J. A. M. 1997, A&AS, 122, 299
Catanese, M., et al. 1998, ApJ, 501, 616
Chadwick, P. M., et al. 1999, ApJ, 513, 161
Chiappetti, L., & Dal Fiume, D. 1997, in Proc. Fifth Workshop, Data Analysis in Astronomy, ed. V. Di Gesú et al. (Singapore: World Scientific), 101
Chiappetti, L., et al. 1999, ApJ, 521, 552
Done, C., Madejski, G. M., Mushotzky, R. F., Turner, T. J., Koyama, K., & Kunieda, K. 1992, ApJ, 400, 138
Edelson, R. A., & Krolik, J. H. 1988, ApJ, 333, 646
Fossati, G., et al. 1998a, in The Active X-Ray Sky, ed. L. Scarsi, H. Bradt, P. Giommi, & F. Fiore (Amsterdam: Elsevier), 423
Fossati, G., et al. 2000, ApJ, 541 (Paper II)
Fossati, G., Maraschi, L., Celotti, A., Comastri, A., & Ghisellini, G. 1998b, MNRAS, 299, 433
Gaidos, J. A., et al. 1996, Nature, 383, 319
Ghisellini, G., Celotti, A., Fossati, G., Maraschi, L., & Comastri, A. 1998, MNRAS, 301, 451
Guainazzi, M., et al. 1999, A&A, 342, 124
Hufnagel, B. R., & Bregman, J. N. 1992, ApJ, 386, 473
Kataoka, J., Takahashi, T., Makino, F., Inoue, S., Madejski, G. M., Tashiro, M., Urry, C. M., & Kubo, H. 2000, ApJ, 528, 243
Kohmura, Y., Makishima, K., Tashiro, M., Ohashi, T., & Urry, C. M. 1994, PASJ, 46, 131
Macomb, D. J., et al. 1995, ApJ, 449, L99
———. 1996, ApJ, 459, L111 (erratum)
Maraschi, L., et al. 1999, ApJ, 526, L81
Nandra, K., George, I. M., Mushotzky, R. F., Turner, T. J., & Yaqoob, T. 1997, ApJ, 476, 70
Padovani, P., & Giommi, P. 1995, ApJ, 444, 567
Peterson, B. M., Wanders, I., Horne, K., Collier, S., Alexander, T., Kaspi, S., & Maoz, D. 1998, PASP, 110, 660
Punch, M., et al. 1992, Nature, 358, 477
Quinn, J., et al. 1996, ApJ, 456, L83
Sambruna, R. M., Barr, P., Giommi, P., Maraschi, L., Tagliaferri, G., & Treves, A. 1994, ApJ, 434, 468
Sambruna, R. M., Maraschi, L., & Urry, C. M. 1996, ApJ, 463, 444
Sikora, M., Begelman, M. C., & Rees, M. J. 1994, ApJ, 421, 153
Takahashi, T., et al. 1996, ApJ, 470, L89
Takahashi, T., Madejski, G., & Kubo, H. 1999, Astropart. Phys., 11, 177
Ulrich, M.-H., Maraschi, L., & Urry, C. M. 1997, ARA&A, 35, 445
Urry, C. M., et al. 1993, ApJ, 411, 614
Urry, C. M., & Padovani, P. 1995, PASP, 107, 803
von Montigny, C., et al. 1995, ApJ, 440, 525
Zhang, Y. H., et al. 1999, ApJ, 527, 719