Extremely Rapid X-ray Flares of TeV Blazars in the RXTE Era

S. F. Zhu1,2,3, Y. Q. Xue2,3, W. N. Brandt1,4,5, W. Cui6, and Y. J. Wang2,3

1 Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA; sxz89@psu.edu
2 CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei 230026, People’s Republic of China; xuey@ustc.edu.cn
3 School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, People’s Republic of China
4 Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
5 Department of Physics, 104 Davey Lab, The Pennsylvania State University, University Park, PA 16802, USA
6 Department of Physics and Astronomy, Purdue University, West Lafayette, IN 47907, USA

Received 2017 October 24; revised 2017 November 19; accepted 2017 December 1; published 2018 January 19

Abstract

Rapid flares from blazars in very high-energy (VHE) γ-rays challenge the common understanding of jets of active galactic nuclei (AGNs). The same population of ultra-relativistic electrons is often thought to be responsible for both X-ray and VHE emission. We thus systematically searched for X-ray flares at sub-hour timescales of TeV blazars in the entire Rossi X-ray Timing Explorer archival database. We found rapid flares from PKS 2005−489 and S5 0716+714, and a candidate rapid flare from 1ES 1101−232. In particular, the characteristic rise timescale of PKS 2005−489 is less than half a minute, which, to our knowledge, is the shortest among known AGN flares at any wavelengths. The timescales of these rapid flares indicate that the size of the central supermassive black hole is not a hard lower limit on the physical size of the emission region of the flare. PKS 2005−489 shows possible hard lags in its flare, which could be attributed to particle acceleration (injection); its flaring component has the hardest spectral energy distribution when it first appears. For all flares, the flaring components show similar hard spectra with parameter values of \( \Gamma = 1.7−1.9 \), and we estimate the magnetic field strength \( B \sim 0.1–1.0 \, \text{G} \) by assuming synchrotron cooling. These flares could be caused by inhomogeneity of the jets. Models that can only produce rapid γ-ray flares but little synchrotron activity are less favorable.

Key words: galaxies: active – galaxies: jets – gamma rays: galaxies – X-rays: galaxies

1. Introduction

Blazars, including BL Lac objects and flat-spectrum radio quasars (FSRQs), are a special class of radio-loud active galactic nuclei (AGNs) that have one of their relativistic jets pointing very close to our line of sight (e.g., Urry & Padovani 1995). FSRQs have luminous broad emission lines that are weak or absent in BL Lac objects. Due to Doppler boosting, the emission of a blazar is usually dominated by the jet whose spectral energy distribution (SED) shows two broad humps that smoothly extend from radio to γ-rays. The low-energy hump can extend from radio to soft X-rays. According to the frequency of the first hump, BL Lac objects are further divided into low-frequency-peaked BL Lac objects (LBLs; \( \nu_{\text{peak}} < 10^{14} \, \text{Hz} \)), intermediate-frequency peaked BL Lac objects (IBLs; \( 10^{14} < \nu_{\text{peak}} < 10^{15} \, \text{Hz} \)) and high-frequency peaked BL Lac objects (HBLs; \( \nu_{\text{peak}} > 10^{15} \, \text{Hz} \), e.g., Padovani & Giommi 1995; Abdo et al. 2010a). The high-energy hump extends from hard X-rays to γ-rays, even sometimes the very high-energy (VHE) TeV band. Such VHE blazars, typically HBLs, are called TeV blazars.

The low-energy hump is attributed to the synchrotron emission of highly relativistic electrons gyrating in a magnetic field in the jet. The origin of the high-energy hump, however, is still debated. A popular explanation is inverse-Compton emission from the same population of relativistic electrons that produce the synchrotron emission. The seed photons of the inverse-Compton scattering process could be local synchrotron photons (usually for BL Lacs) in the jet and/or external photons from the central engine (usually for FSRQs) or the cosmic microwave background. These models are thus called leptonic models. The hadronic models, on the other hand, attribute γ-ray emission to synchrotron emission of protons (Mücke & Protheroe 2001; Mücke et al. 2003; Fraija & Marinelli 2015) or proton-induced cascades (Mannheim 1998).

The observed photon flux of blazars varies significantly across the electromagnetic spectrum on timescales from minutes to years (e.g., Wagner & Witzel 1995; Ulrich et al. 1997). The origin of the variability is not well understood. Generally, the variability is noise-like (e.g., Kataoka et al. 2001; Chatterjee et al. 2012), similar to the variability of radio-quiet AGNs (e.g., Markowitz et al. 2003). However, blazars are also known to have bursts that show flare-like structures (e.g., Marscher et al. 2010), which may have recognizable patterns (Sasada et al. 2017). The outbursts can be explained by internal shocks of the jets when a new relativistic blob of plasma catches up with an old blob and accelerates particles to ultra-relativistic energies (e.g., Spada et al. 2001). Based on several similarities between the jet emission and corona-disk emission, the ultimate origin of the variability may still be accretion-rate fluctuations of the disk (McHardy 2008).

The shortest variability timescale is a crucial parameter because it serves as an independent constraint on the physical scale of the emission region (Tavecchio et al. 1998), which cannot be easily provided by other observational measurements. Blazars are usually most variable at frequencies just above the two SED humps (e.g., Ulrich et al. 1997; Madejski & Sikora 2016), which usually fall in the hard X-ray and TeV bands in the case of TeV blazars (e.g., Aleksic et al. 2015a; Baloković et al. 2016; Bartoli et al. 2016). In particular, an increasing number of TeV blazars show γ-ray flaring activity on timescales from several to a few tens of few minutes that are detected by ground-based Cherenkov telescopes, including both BL Lac objects (e.g., Gaidos et al. 1996; Aharonian et al. 2007; Albert et al. 2007;
et al. (2013) and FSRQs (Aleksić et al. 2011). The minute-scale variability in the TeV band (e.g., Aharonian et al. 2007) has strong implications for our understanding of AGN jets (Begelman et al. 2008). X-ray and TeV emission may be directly related to the same high-energy tail of the relativistic electron population. Indeed, the light curves of HBL-type TeV blazars in the X-ray and TeV bands are usually correlated (e.g., Aleksić et al. 2015b; Furniss et al. 2015b). Attempts to search for extremely rapid X-ray variability have been made (e.g., Cui 2004; Xue & Cui 2005; Paliya et al. 2015; Pnyal et al. 2015). The same source can have minute-scale variability in both the X-ray band and TeV band (e.g., Mrk501, Xue & Cui 2005; Albert et al. 2007). However, “orphan” TeV flares that have no X-ray counterparts are occasionally reported (e.g., Krawczynski et al. 2004; Blazejowski et al. 2005; Acciari et al. 2009; Fraija et al. 2015).

Rapid TeV variability has germinated various models to explain the small timescales. Most models involve some very compact regions moving in the rest frame of the jet. These compact regions could be “jets in a jet” that are either produced by magnetic reconnection processes in a Poynting flux-dominated jet (Giannios et al. 2009) or relativistic turbulence in the jet (Narayan & Piran 2012). The minijets-in-a-jet model can consistently produce the statistical properties of blazar flux (Biteau & Giebels 2012). There are also models involving a red giant star being stripped of its envelope by the jet (Barkov et al. 2012) and models involving beams of magneto-centrifugally accelerated electrons occasionally pointing toward us (Ghisellini et al. 2009).

Several well-studied TeV blazars show rich spectral behavior in X-rays, which may represent the general behavior of the synchrotron peak of all AGN jets. The X-ray spectra are usually curved (Massaro et al. 2004) and can only locally be fitted by a power law. The spectral variation with flux can be complex (Zhang et al. 2002; Cui 2004). Generally, the spectrum hardens when the flux increases (e.g., Gliozzi et al. 2006; Xue et al. 2006; Tramacere et al. 2009), but photon indices can saturate at higher fluxes (Xue & Cui 2005; Giebels et al. 2007). The synchrotron peak usually moves to higher frequencies with increasing flux during outbursts (e.g., Pian et al. 1998), but no correlation between the break energy and the flux exists when a broken power law is adopted to fit the X-ray spectra (Xue & Cui 2005; Giebels et al. 2007; Garson et al. 2010). A cooling break in the spectrum of emitting particles cannot explain these features (Wierzcholska & Wagner 2016), and some special particle acceleration processes may be involved (Madejski & Sikora 2016). There are also energy-dependent lags between the variations of different energy bands. In some flares, soft bands lag behind hard bands (e.g., Zhang et al. 2002), while lags in the opposite direction can also happen (e.g., Ravasio et al. 2004; Sato et al. 2008). Hysteresis in the HR (hardness ratio)–flux diagram is often used as a diagnostic of lags. Clockwise loops (e.g., Acciari et al. 2009; Kapanadze et al. 2016) in the HR–flux plane are a sign of soft lags while counterclockwise loops (e.g., Tramacere et al. 2009) are a sign of hard lags. The same source can exhibit both clockwise and counterclockwise loops; the observed patterns are further complicated by the superposition of flares at different timescales (Cui 2004). The above knowledge of TeV blazars in the X-ray regime comes from studies focusing on timescales of hours to weeks. We will extend this kind of analysis to much smaller timescales in this paper.
faint sources. Background data were then simulated using the appropriate model.\textsuperscript{10} We applied the GTIs to both observational data and simulated background and extracted light curves in initial 16 s bins from channels that correspond to \( \sim 2–20\) keV. The light curves of net count rates were calculated using \texttt{lcmath} in the HEASoft (v6.19) package. As only PCU2 among the five proportional counter units (PCUs) of PCA is almost always in operation, we extracted light curves from PCU2 for flaring event selection. We visually inspected every light curve to select events in individual observations that contain a complete or nearly complete sub-hour flaring profile. Specifically, we require the flare to have apparent rise and fall; we also require the existence of a plateau either before the rise and/or after the fall to assess the completeness of the flare and the level of the background component.\textsuperscript{11} Note that \textit{RXTE} is in low Earth orbit, and thus an uninterrupted light curve is usually less than \( \sim 50\) minutes due to the Earth’s occultation or passage through the South Atlantic Anomaly, etc., which limits the timescales of the events investigated. After identifying fast flaring events, we extracted light curves and spectra from all the PCUs available during that observation to achieve a higher signal-to-noise ratio (S/N) for further analysis.

From the complete \textit{RXTE}/PCA database, two new fast flaring events were found. They belong to PKS 2005–489 and S5 0716+714. The light curves of these two events are shown in Figure 1. We also plot an event of Mrk 501 (reported in Xue & Cui 2005). A fast-flaring candidate event of 1ES 1101–232 is reported in Appendix A, which has relatively low credibility because of limited S/N. These flaring observations generally lack simultaneous observations in other wavebands.

We checked for potential contamination by soft electron flares that were not screened out by the criteria “\textsc{Electron2} \& \textsc{Electron2}.L.0.1” in data cleaning. The contemporaneous Electron2s of each event were well below 0.1 and did not show any apparent electron flaring activity that might be responsible for the X-ray flares. The longitudes and latitudes of the satellite at the onsets of the X-ray flares did not cluster in the anomalous high background region (cf. Figure 7 of Xue & Cui 2005). Further support for the genuineness of the flaring events comes from their light curve and spectral features explored below. They behave like well-known X-ray flares of TeV blazars, only at much smaller timescales. In conclusion, we did not find any sign of contamination of soft electrons or any other known sources for any of the flaring events. However, the possibility of an unrelated X-ray transient in the field of view still cannot be ruled out entirely, as the PCA lacks the capacity of imaging.

3. Light Curve Fitting

Light curves in additional energy bands (see the second column of Table \textsuperscript{2} and Figures 2–4) are extracted according to the energy-channel conversion table.\textsuperscript{12} Thanks to the high data quality, variability at timescales down to the time resolution (16 s) of the light curves (see Figure 2) is seen. Because of the variety of data quality and gain epochs (Jahoda et al. 2006), we do not have uniform definitions for different bands for all three

\begin{align}
F(t) &= F_\text{c} + F_0(e^{\frac{t}{\tau_r}} + e^{\frac{t}{\tau_d}})^{-1},
\end{align}

where \( F_\text{c} \) represents the constant flux level underlying the flare, and \( \tau_r \) and \( \tau_d \) are the characteristic rising and decaying timescales\textsuperscript{14} of the flare. \( t_0 \) indicates the transition from rising

\textsuperscript{10} We adopted the faint background model file \texttt{pca_bkgd_cmfnair-tlf7_eMv20051128.mdl}.

\textsuperscript{11} We do not adopt quantitative criteria to select the events, automatically, because quantitative criteria are inescapable of subjective tweak and visual inspection is almost always necessary.

\textsuperscript{12} See the table at http://heasarc.gsfc.nasa.gov/docs/xte/e-c_table.html.

\textsuperscript{13} In addition to the frequently used Equation (1), some similar analytical expressions have been used to describe the flare profiles of blazars; see Albert et al. (2007), Giebels et al. (2007), Chatterjee et al. (2012).

\textsuperscript{14} The doubling and halving timescales are \( \tau_r \times \ln 2 \) and \( \tau_d \times \ln 2 \).
Table 2
Fitting Results for the Fast Flaring Events

| Band      | Energy (keV) | $F_i$ (c/s) | $F_0$ (c/s) | $t_0$ (s) | $\tau_i$ (s) | $\tau_d$ (s) | $t_F$ (s) | $F_F/F_c$ | $\xi$ | $\chi^2$/dof |
|-----------|--------------|-------------|-------------|-----------|---------------|--------------|-----------|-----------|------|-------------|
| PKS 2005–489 |
| Soft      | 1.94–5.47    | 2.08±0.02   | 3.81±0.28   | 303±8     | 23±5          | 160±16       | 343±5     | 2.3       | 0.75 | 0.957/166  |
| Medium    | 5.47–10.11   | 0.98±0.02   | 3.45±0.24   | 318±10    | 29±3          | 141±14       | 356±6     | 3.2       | 0.66 | 1.05/166   |
| Hard      | 10.11–20.30  | 0.33±0.02   | 2.43±0.33   | 324±34    | 29±18         | 105±25       | 353±12    | 5.3       | 0.56 | 0.895/166  |
| Full      | 1.94–20.30   | 3.4±0.04    | 9.49±0.47   | 311±6     | 25±4          | 143±11       | 348±6     | 2.8       | 0.70 | 1.18/166   |

| $S5$ 0716+714 |
| Soft      | 2.06–5.31    | 0.19±0.08   | 2.04±0.28   | 1344±111  | 222±27        | 825±77       | 1573±54   | 7.5       | 0.58 | 1.48/23    |
| Hard      | 5.31–10.11   | 0.00±0.08   | 1.67±0.21   | 1216±108  | 161±45        | 1588±1013    | 1551±1707 | >17       | 0.82 | 1.62/23    |
| Full      | 2.06–10.11   | 0.17±0.09   | 3.95±0.35   | 1258±53   | 178±29        | 1158±139     | 1547±18   | 17        | 0.73 | 1.15/159   |

| Mrk 501 first substructure |
| Soft      | 1.94–5.82    | 5.60±0.03   | 3.3±0.02    | 1918±32   | 99±27         | 110±75       | 1924±14   | 1.30      | 0.05 | 0.99/151   |
| Hard      | 5.82–20.30   | 5.5±0.05    | 2.4±0.38    | 2001±30   | 147±14        | 33±45        | 1961±30   | 1.38      | −0.64 | 1.09/151   |
| Full      | 1.94–20.30   | 10.96±0.06  | 6.9±1.1     | 1961±47   | 119±21        | 67±9         | 1936±18   | 1.33      | −0.28 | 1.23/151   |

| Mrk 501 second substructure |
| Soft      | ...          | ...         | 1.78±0.28   | 2206±25   | 40±15         | 160±71       | ...       | ...       | 0.60  | ...        |
| Hard      | ...          | ...         | 1.73±0.35   | 2315±64   | 30±79         | 302±225      | ...       | ...       | 0.82  | ...        |
| Full      | ...          | ...         | 3.7±0.6    | 2177±54   | 43±55         | 213±83       | ...       | ...       | 0.67  | ...        |

Notes.

a Note our fitting cannot constrain the hard-band quiescent flux level of $S5$ 0716+714, so the hard-band amplitude is unmeasurable. We thus set the full-band amplitude as the lower limit of that of the hard band.

b We fitted the model to the soft- and hard-band light curves in bins of 96 s, while the full-band light curve is in bins of 16 s.

to falling, and the count rate actually peaks at

$$t_p = t_0 + \frac{\tau_i \tau_d}{\tau_i + \tau_d} \ln \left( \frac{\tau_d}{\tau_i} \right),$$

which equals $t_0$ only when the flare is symmetrical ($\tau_i = \tau_d$). Therefore, we define the amplitude of the flare as $F_F/F_c$, i.e., the count rate at $t_p$ ($F_p$) over the constant level, instead of $F_0/F_c$.

The symmetry of a flare is described by

$$\xi = \frac{\tau_d - \tau_i}{\tau_d + \tau_i},$$

whose value is in the range of $[-1, 1]$. $\xi = -1 (=1)$ represents completely right (left) asymmetric profiles with a zero falling (rising) timescale; $\xi = 0$ indicates a symmetric flare.

The uncertainties of $t_p$ and $F_p$ have to be propagated from the errors of other parameters. We adopted an MCMC (Markov Chain Monte Carlo) algorithm to fit the light curves, which returns reliable probability intervals of timescales and amplitudes by sampling from their posterior distributions. We first performed weighted least-squares fitting using a numerical minimizer to obtain the best estimates of $F_c$, $F_0$, $t_0$, $\tau_i$, and $\tau_d$. Starting from these initial values, we took 1000 random walk steps in parameter space. The samples of $t_p$ and $F_p$ were calculated according to Equations (1) and (2). Note that we added a second flaring component to the model in the fitting of Mrk 501 (see the bottom panel of Figure 1).

Figure 2. Top three panels: The rapid X-ray flare of PKS 2005–489 in three energy bands in bins of 16 s. The red solid curves are weighted least-square models; the vertical blue dashed lines indicate $t_p$, whose uncertainties are shown as the shaded blue regions. Bottom two panels: Hardness ratios (hard band to soft band and medium band to soft band); the red solid curves are hardness ratios calculated from the weighted least-square models above. The data points are in bins of 32 s.

PKS 2005–489 1991/08/02 (MJD=50662.97097)
3.2. Correcting Light Curve Error Bars

The initial fits have reduced Chi-square values ($\chi^2 = \chi^2/dof$) in the range of 0.48–0.80, which indicates that the assigned error bars are larger than true statistical fluctuations. Indeed, the standard RXTE/PCA data reduction pipeline overestimates the light curve errors (Nandra et al. 2000). The error estimation of the net light curves is propagated from the error estimation of the observed light curves and the simulated background light curves, of which the latter is too smooth to be described by the assumed Poisson statistics. We decided to correct the error estimation using $\sigma_{\text{net}}^2 = \sigma_{\text{obs}}^2 + k^2\sigma_{\text{bkg}}^2$, where $0 \leq k^2 < 1$. The correction factor $k^2$ can be determined by forcing the excess variance\(^{15}\) of the quiescent parts in the top and bottom panels of Figure 1 to be zero. The resulting correction factors of different segments at different energy bands span from −0.13 to 0.71. We decided to fix $k^2 = 0$ as in Nandra et al. (2000). We have ignored errors on the background in the light curve analysis below, unless otherwise stated. The fitting results after correcting the error bars are tabulated in Table 2. Note that we still report the least-square results in Table 2 as the estimation of each parameter, but the 1σ intervals are derived from MCMC fitting. The reduced Chi-square values of most fits are around 1.

3.3. Light Curve Fitting Results

Every light curve shows some flare-like structure above a constant “background” flux level, which actually varies on longer timescales. From the flux levels of the constant components, the events occur when the sources are in relatively high states, but they do not always coincide with the periods with the highest flux levels. For example, the RXTE/PCA count rates of PKS 2005−489 can be 10 times the constant flux level here as found about one and a half years later (Perlman et al. 1999).

The variation amplitude is higher in harder bands, which suggests that the flaring component has a harder spectrum than the corresponding constant component. The mixed spectra become harder when the flux rises and the flaring component becomes more prominent. We discuss the spectral variability of PKS 2005−489 in detail in Section 3.4. We also fit the spectra of PKS 2005−489, S5 0716+714, and Mrk 501 and confirm that the flaring components have harder spectra in Section 4.

There is a trend of rising timescales being shorter and decaying timescales being longer at softer energies. This trend is obvious in PKS 2005−489 and the first flare substructure of Mrk 501 (see Table 2; see also Mrk 421 in Appendix B). This suggests that the variability is caused by electron acceleration and cooling (e.g., Fraija et al. 2017). As a consequence, the flares are more right asymmetric in harder bands (i.e., smaller $\xi$ values). S5 0716+714 and the second flare substructure of Mrk 501 do not follow the patterns, although the error bars prevent any solid conclusions. The difference of timescales between energy bands is not obvious for S5 0716+714, mainly because the decay of the flare was not completely sampled and $\tau_d$ cannot be constrained well. Moreover, only one PCU was operating during the observation, so the S/N is low.

The vertical lines in Figures 2 and 4 suggest hard lags in the variability of PKS 2005−489 and Mrk 501, which means that the variation of hard photons lags that of soft photons. The

\(^{15}\)Excess variance is the variance after subtracting the mean square error (e.g., Nandra et al. 1997; Vaughan et al. 2003).

\(^{16}\)We define hardness ratio as the count rate of hard band over that of soft band, $HR = \frac{H}{S}$. The errors are calculated as $\sigma_{HR} = \frac{1}{\pi} \left( \frac{2H}{H} + \frac{2S}{S} \right)$.
change of the relative fraction of the two components. In other words, if the flaring component has a different hardness ratio from that of the constant component, then even if neither of the spectra changes over time, the observed overall hardness ratio will still change due to the flux variation of the flaring component (see Sun et al. 2014; Ramolla et al. 2015). However, the hysteresis loop in panel (a) suggests the spectrum of the flaring component is intrinsically variable; otherwise, the track in the HR–flux plane while the flux is rising will be identical to the track while the flux is declining, instead of forming a loop. To discriminate the effects caused by the mixing of different components and the flare’s intrinsic spectral variability, we further subtract the constant component $F_c$ from the light curve of each band, where $F_c$ is from Table 2. We calculated the hardness ratios from the resulting flare-only light curves in panels (c), (d), and (e) and an HR–flux diagram in panel (b). The plots suggest that the flare emerges with a hard spectrum, softens gradually as the flux rises, then hardens near the highest flux, and finally softens as the flare fades away. Only a counterclockwise loop is apparent in the HR–flux plane for the flare-only light curve. Based on the evidence above, we know that the clockwise loop in panel (a) could be caused by the sudden emergence of a hard flaring component and the spectral variability soon follows the spectral variability of this flaring component due to its increasing dominance. The flaring component itself has complex spectral variability and the counterclockwise loop in panel (b) suggests the existence of a hard time lag, albeit being subject to large uncertainties. Note that the overall oblique “8” shape before subtracting the constant component in panel (a) is reminiscent of a flare of Mrk 421 that lasted ~60 ks and was reported by Garson et al. (2010) using Suzaku data. We reanalyzed the Suzaku data of Mrk 421 and performed the same spectral variability analysis as for PKS 2005–489 above in Appendix B. We reproduced...
the same results as for PKS 2005−489 in the flare of Mrk 421 and thereby strengthened our conclusions. The similar behavior of spectral variability, seen both in the extremely rapid flare of PKS 2005−489 and in the long-duration flare of Mrk 421, is reminiscent of the scale invariant nature of X-ray flares from TeV blazars (Cui 2004; Xue & Cui 2005).

We also calculated the CCF (Edelson & Krolik 1988; Welsh 1999) of the soft-band and hard-band flare-only light curves but did not find an obvious time lag. Any time lag in this fast flaring event of PKS 2005−489 could be intrinsically small compared with the time resolution of the observation (16 s). The bandpass of our data is narrow, spanning about one order of magnitude, so any energy-dependent lags may not be significant. The size of lags may also be positively correlated to the duration of the flares (Zhang et al. 2002), which means that flares of shorter duration have smaller lags. Note that in Table 2, the differences of $t_0$, $t_p$, and $\tau_t$ between different bands are small compared with relatively large error bars.

4. Spectral Fitting

Motivated by the two-component model in light curve fitting, we carried out spectral fitting in a similar way. By dividing the observation into flaring and quiescent phases according to the light curves, we separately extracted and jointly fitted the spectra of the two phases. The spectra were extracted from the top layers of the operating PCUs. The spectra and models are shown in Figure 6 in the EFE representation and the photon indices are tabulated in Table 3. We used channels that correspond to 3−20 keV for PKS 2005−489 and Mrk 501 and 3−10 keV for S5 0716+714, respectively. In addition to Galactic absorption, we used a power law and sum of two power laws in XSPEC (v12.9.0) to fit the spectra of the quiescent and flaring phases, respectively. We tied the power law in the quiescent phase to one of two power laws in the flaring phase, for both normalization and photon index (see the legends of Figure 6). We also calculated the X-ray luminosity of the constant component and the total energy of the flaring component in the X-ray band (namely average flare luminosity times the length of the defined flare phase) in Table 3. Note that the spectra should suffer from overestimation of error bars due to the same reason discussed in Section 3.2, but we are not able to correct the error estimation of the spectra as we did in the light curve fitting. One consequence of the overestimation of error bars is that the
parameters in the spectral fitting have large confidence intervals. Another consequence is that we might be able to fit the spectra with many models. For example, we can fit the spectra of the flaring phase, quiescent phase, and whole observation with simple power laws, and the Chi-square statistics can still be acceptable. We think that describing the spectra as a combination of different power laws is more physically appropriate and is consistent with the light curve fitting.

Note that due to the limited length of the observation, we have neither a complete flaring nor a quiescent phase of S5 0716+714. Therefore, we extracted the spectra of the two observations just before and after the flare, which are referred to as pre-flare and post-flare, respectively. The gaps between the observations are 3–4 days. The three observations are plotted in the bottom-right panel of Figure 6, where we fitted each of the pre-flare and post-flare light curves with a constant flux. The spectra are jointly fitted with a tied “constant” component photon index, but the normalizations of the “constant” component are free parameters in this case. In other words, the constant component varies on timescales of a few days. If we leave the normalization of the constant component underlying the flare (factor “A” in the legend of top-right panel) free, the value of “A” in the fitting results is negligible but its error bars are large. Additionally, we are not able to constrain the photon index of the flaring component ($\Gamma_2$) very well in this case. Thus, we fixed $A = 0.6D$, according to the light curve fitting results in the bottom-right panel of Figure 6. The choice of 0.6 or another reasonable value does not affect the result of $\Gamma_2 < \Gamma_1$.

The above fitting shows that for all three flares, the flaring components have harder photon indices than the corresponding constant components (see Table 3). The difference is not as apparent for Mrk 501 because of the relatively small variation amplitude (i.e., relatively weak flaring component). The photon index of the constant component ranges from 2.0 to 2.7, indicating that we are observing different declining parts of the synchrotron hump for different sources. The photon indices of the flaring components, on the other hand, lie in a narrower range around 1.8. Every estimated average flux (averaged over the length of the flare phase we defined) of the events is $\sim 10^{-11}$ erg cm$^{-2}$ s$^{-1}$ (see Table 3). Using the “efficiency limit” of compact sources (e.g., Fabian 1979; Brandt et al. 1999), $\Delta L / \Delta t \lesssim 2 \times 10^{41} \eta_{0.1}$ erg s$^{-1}$, the matter to energy conversion efficiency $\eta$ is greater than 1 for PKS 2005–489 and S5 0716+714, suggesting the existence of significant boosting of the emission (e.g., Remillard et al. 1991). Due to the small flare amplitude, the efficiency of Mrk 501 does not exceed 1.

5. Discussion

5.1. The Flaring and Constant Components

We interpret all the light curves that we analyzed in detail (including Mrk 421 in Appendix B) as a superposition of a constant component and a flaring component (e.g., Fraija et al. 2017). The two-component model depicts a simpler scenario than a single component undergoing an outburst. For the case of a single emitting region, the variation in the HR–flux plane is more complex. Furthermore, there must have been a sudden enhancement of some key physical quantity, and this quantity has to fall back later to its value preceding the burst.

5.2. X-Ray Radiation Process of S5 0716+714

S5 0716+714 is an IBL (Ackermann et al. 2011) with a synchrotron hump peaking at optical wavelengths (Anderhub et al. 2009). It has frequent intra-day variability and hysteresis loops in the color–magnitude plane in the optical band (e.g., Man et al. 2016). Praly et al. (2015) reported two rapid flares of S5 0716+714 in the X-ray band with low significance. The X-ray emission covered by the RXTE bandpass could be a mixture of both synchrotron and inverse-Compton emission (Wierzcholska & Siejkowski 2016). As such, S5 0716+714 will not necessarily behave the same way as HBLs in X-rays. Depending on the state (low or high) of the source when the observation was made, the dominant X-ray radiation process may change accordingly. The spectrum may steepen when S5 0716+714 brightens, considering that the synchrotron emission in the soft X-rays becomes increasingly important (Giommi et al. 1999). However, the synchrotron component itself hardens due to the shifting of the synchrotron peak (Ferrero et al. 2006; Zhang 2010). The steep photon index and the high flux levels (3–10 keV) of pre-flare and post-flare observations are consistent with synchrotron emission. The flaring component has a slightly inverted spectrum, but the even higher flux level and short timescales still support an origin as synchrotron emission. Therefore, the flare we observed from S5 0716+714 does not seem to have a different radiation process from HBLs.

5.3. Constraints on the Emission Region

The upper limit on the physical scale of the flaring region can be given by

$$R \approx \frac{\delta I_{\text{flare}}}{(1+z)} \approx 10^{14} \left( \frac{\delta}{30} \right) \left( \frac{I_{\text{flare}}}{100\text{ s}} \right) \text{cm},$$

(4)

Table 3

| Source          | $\Gamma_1$ | $\Gamma_2$ | $F_{\text{quiescent}}^a$ | $F_{\text{flare}}^a$ | $F_{\text{flare\-only}}^a$ | $\chi^2/\text{dof}$ | $L_X^b$ | $E^c$ |
|-----------------|------------|------------|--------------------------|------------------------|----------------------------|----------------------|---------|-------|
| PKS 2005–489   | 2.73 ± 0.05| 1.71 ± 0.18| 29.5                     | 42.4                   | 12.9                       | 0.88/38              | $3.86 \times 10^{44}$ | $2.08 \times 10^{47}$ |
| Mrk 501        | 2.01 ± 0.02| 1.86 ± 0.20| 115.7                    | 130.0                  | 14.3                       | 0.76/38              | $3.23 \times 10^{44}$ | $5.13 \times 10^{46}$  |
| S5 0716+714    | 2.48 ± 0.27| 1.89 ± 0.15| 5.7 (pre) 4.8 (post)     | 12.3                   | 12.2                       | 0.32/16              | $1.90 \times 10^{45}$ | $1.25 \times 10^{45}$  |

Notes.

a Fluxes are in units of $10^{-12}$ erg cm$^{-2}$ s$^{-1}$; the energy range is 3–20 keV for PKS 2005–489 and Mrk 501, but 3–10 keV for S5 0716+714.

b X-ray luminosity of the constant component luminosity in units of erg s$^{-1}$.

c Energy of the flaring component in units of erg.
where $\delta$ is the Doppler boosting factor, and $t_{\text{flare}}$ is the observed variability timescale. Because the flares are asymmetric, the size of the emitting region is reflected by the rising timescale $\tau_i$ (Zhang et al. 2002). Searching the literature, we find black hole mass estimates for three out of the four sources (except 1ES 1101–232). We plot the rising timescale in the galaxy’s frame against black hole mass in Figure 7. Note that the typical Doppler factor of BL Lac objects is 10–20 (20–30 for FSRQs, e.g., Hovatta et al. 2009). The size of the black hole is often thought to be a natural lower limit on the physical scale of the emission region. We also plot this lower limit as a function of black hole mass assuming different Doppler factors. The variability timescales should lie above the lines, if the size of the black hole is a hard lower limit. The rising timescale of PKS 2005–489 is, as far as we know, shorter than any variability timescale of AGNs at any wavelength ever reported (e.g., Remillard et al. 1991; Yaqoob et al. 1997; Gallo et al. 2004; Xue & Cui 2005; Aharonian et al. 2007; Albert et al. 2007; Aleksic et al. 2014; Kara et al. 2016). Figure 7 shows that the Doppler factor of PKS 2005–489 has to be larger than several hundred, which appears unrealistic, in order to support the idea that the black hole sets a lower limit on the physical size of the flaring region. So far, there is no correlation found between the observed minimum variability timescale and the black hole mass (Wagner 2008; Vovk & Babić 2015). The events of PKS 2005–489 and Mrk 421 (Appendix B) display remarkably similar spectral evolution, supporting the same process driving the spectral variability. However, the timescales of the two events are different by two orders of magnitude, reflecting an intrinsic difference between the two jets. Noticeably, the famous Mrk 421 has many more X-ray observations than PKS 2005–489; however, the variability timescale of Mrk 421 has never been found to be as short as the PKS 2005–489 event (Cui 2004; Paliya et al. 2015; Pryan et al. 2015). A lower limit on the variability timescale probably does exist, and probably is not set by the central supermassive black hole.

The synchrotron cooling time of emitting electrons is given by $t_{\text{cool}} \approx 6\tau_i m_e c/\gamma e^2 B^2$ (Rybicki & Lightman 1979). The observed photon energy at the synchrotron peak is given by $E_p = 8\delta h \nu \approx (3e\hbar/4\pi m_e c) \gamma e^{-2} B^2$ (Rybicki & Lightman 1979). Combining the above two equations, we have

$$t_{\text{cool}} = 3.04 \times 10^3 B^{-3/2} \gamma^{-1/2} E_p^{-1/2} \text{s},$$

where $E_p$ is in units of observed keV (Zhang et al. 2002). If we take PKS 2005–489 as an example, and adopt $t_{\text{cool}} = 143 \text{s}$, $E_p = 10\text{keV}$ and $\delta = 30$, we can have $B \approx 1.2\text{G}$. The estimated magnetic field of Mrk 501 is similar to that of PKS 2005–489, but the magnetic field of S5 0716+714 is four times weaker. Radio-loud AGNs are potential accelerators of cosmic rays. We estimated the maximum energy of protons, if they can be accelerated in the same region as the electrons, as $E_{\text{max}} = eB R \sim (2-9) \times 10^{16} \text{eV}$.

### 5.4. The Rarity of Extremely Rapid Flares of Blazars

It is surprising to find such rapid variability of PKS 2005–489, because it was not found to be variable on timescales less than a day (Perlman et al. 1999; Rector & Perlman 2003; H.E.S.S. Collaboration et al. 2010, 2011). Out of the ~160 PCA pointings on PKS 2005–489, only two have positive excess variance,\(^\dagger\) including the one shown in Figure 1(a).

Indeed, although the list of blazars that show extremely rapid flares is growing, these events are rare (e.g., Feigelson et al. 1986) and usually unexpected. The rate of occurrence of sub-hour flares is about once per 4 Ms in the RXTE/PCA database. Most of the sources have only one such event reported, either in $\gamma$-ray or X-ray (e.g., Guo et al. 2014; Hovatta et al. 2009; Aharonian et al. 2007; and this paper). These rapid flares do not seem to be the extremely short cases from a continuous distribution of the timescale of flares (e.g., Li et al. 2017; Sasada et al. 2017). It remains unknown whether such events exist in all wavebands, so we do not know the total energy output of the flares. The biggest challenge is to coordinate multiple instruments to target the same source simultaneously and hope rare unpredictable flaring events happen.

### 5.5. Particle Acceleration

Assuming the fast-cooling regime, the photon index of the flaring component ($\Gamma = 1.7–1.9$) indicates that the accelerated electrons have an effective energy spectral index ($p = 2\Gamma - 2$) in the range of 1.4–1.8, which can be achieved by relativistic magnetic reconnection (e.g., Guo et al. 2014). Indeed, the SED of PKS 2005–489 shows very low Compton dominance (H.E.S.S. Collaboration et al. 2010, 2011), which indicates high magnetization. The results of SED modeling (e.g., Anderhub et al. 2009; Abd et al. 2011; H.E.S.S. Collaboration et al. 2011; Aleksic et al. 2015a), which usually assume one single homogeneous emitting region, have low-strength magnetic fields ($B \sim 0.01–0.1\text{G}$) compared with the estimation of Section 5.3 and below equipartition. This may require that

\(^{\dagger}\)We do not correct the error bars of the light curves here, but the light curves are consistent with being flat under visual inspection.
either the jet is structured (Ghisellini et al. 2005) or only the regions that are responsible for the fast flares we observed have high magnetization.

The direction of the loops in the HR–flux plane is thought to be determined by the competition between the acceleration/ejection timescale, cooling timescale, and escape timescale (Kirk et al. 1998). In practice, it is actually determined by which part of the synchrotron spectrum we are observing. The distinctive pattern in the spectral variation of the flares of PKS 2005−489 and Mrk 421 is not predicted by time-dependent homogeneous one-zone models (e.g., Kirk et al. 1998; Chiaberge & Ghisellini 1999). The spectral variability pattern in Figures 5 and 10 could possibly be produced by time-dependent inhomogeneous blazar models (e.g., Böttcher & Dermer 2010).

6. Summary

We searched the entire RXTE archival database for rapid X-ray flares of TeV blazars that last less than one hour. We investigated the temporal and spectral properties of the fast flares discovered under a two-component assumption. Our analysis also includes an X-ray flare of Mrk 421 using Suzaku data that has similar spectral variability to that of PKS 2005−489 in Appendix B. Our main findings are as follows:

1. We discovered two new fast X-ray flares from PKS 2005−489 and S5 0716+714 and a candidate flare from 1ES 1101−232. The event of PKS 2005−489 shows, as far as we know, the most rapid variation of AGNs that has been observed at any wavelength. The extremely small timescale ($\tau_r < 30$ s) defies the size that corresponds to the light-crossing time of the supermassive black hole as a lower limit on the size of the flaring region.

2. The flares are usually superimposed on a constant/slowly varying component. The flaring component generally has a harder X-ray spectrum ($\Gamma = 1.7$−1.9) than the constant component ($\Gamma = 2.0$−2.7).

3. The higher data quality of the X-ray observations can provide more detail than $\gamma$-ray observations. The flare-only component shows a counterclockwise pattern in the HR–flux diagram, providing a sign of likely hard lags. This component also has the hardest spectrum right at its appearance.

RXTE ceased science operation in 2012 January, but X-ray observatories like Chandra, XMM-Newton, Swift, and NuSTAR are still accumulating exposures on TeV blazars. For example, XMM-Newton has 9 Ms of exposure on TeV blazars as of 2017 July. A natural follow-up work is to search for additional rapid X-ray flares from TeV blazars in these archival databases. With the increased sample, we may study more details of the spectral variability of the flares, thus shedding light on their origin and particle acceleration processes. We can also study the correlations between flares and the properties of the jets and central engine, as well as the reason for the small flaring timescales.

We thank the anonymous referee for his/her helpful comments that improved the paper. S.F.Z., Y.Q.X., and Y.J. W. acknowledge support from the 973 Program (2015CB857004), NSFC-11473026, NSFC-11421303, the CAS Frontier Science Key Research Program (QYZDJ-SSW-SLH006), and the Fundamental Research Funds for the Central Universities. S.F.Z. and W.N.B. acknowledge support from Chandra X-ray Center grant G04-15093X.

Figure 8. The candidate rapid X-ray flare of 1ES 1101−232. The ObsID is 95387-02-07-00, and only PCU2 was in operation during the observation. The gray points and blue squares are the data of 16 s bins and 96 s bins, respectively. The red dashed curve is the least-square model fitting to the 96 s-bin light curve.

Appendix A

A Candidate Fast Flare of 1ES 1101−232

We found one fast X-ray flaring event of 1ES 1101−232, which is an HBL at a relatively high redshift ($z = 0.186$) compared with other HBLs. The full-band (2−20 keV) light curve in 16 s bins (Figure 8) shows elevated flux in the second half of the observation. When shown in 96 s bins, the light curve clearly manifests an almost complete flare profile and can be fitted using the model of Equation (1). The amplitude ($L_{\text{p}}/L_0 = 1.93$) of this flare is modest, and the rising timescale ($\tau_r = 60$s) is almost as fast as PKS 2005−489. We only reported it as a candidate because the flux preceding the flare seems to be in continuous declining. The noisy data prevent us from performing data analysis in the same fashion as other sources (sub-band light curve fitting, joint spectral fitting, etc.).

Appendix B

Spectral Variability of Mrk 421

The flare-only component of PKS 2005−489 shows an interesting spectral variability pattern, but the noisy data prevent us from making strong conclusions. The oblique “8” pattern in Figure 5(a) is reminiscent of a flare of Mrk 421, which allows us to confirm our finding in another source using data from a different satellite.

We downloaded Suzaku data (ObsID 703043010) of Mrk 421 (Garson et al. 2010). We only used data from the X-ray Imaging Spectrometer (XIS) on board Suzaku. The data were reprocessed and screened using AEPIPELINE included in the Suzaku FTOOLS. In addition to the standard screening criteria, we also required the cut-off rigidity to be larger than 6 GV/c following Garson et al. (2010). We extracted light curves separately from the cleaned event files of XIS0 and XIS3 using XSELECT in initial 16 s bins. The source region has an inner radius of 35 pixels and outer radius of 408 pixels, while the background region is an annulus with an inner radius of 432 pixels and outer radius of 464 pixels. The light curves were extracted in two energy bands: 0.5−2 keV and 2−10 keV, which are referred to as the soft band and hard band, respectively. Light curves of the same energy band but
different detectors (XIS0 and XIS3) were then merged. We rebinned the light curves in a size of 5760 s in accordance with the orbital period of Suzaku. Finally, the background light curves were subtracted to obtain the estimation of net count rates from the source.

We clipped the light curves and kept only the segment of Flare 2 that is defined in Garson et al. (2010). We fitted the soft- and hard-band flares using a model that is analogous to Equation (1). Only the constant component underlying the flare was replaced by a slowly varying component using a linear function $F(t) = m + \text{slope} \times t$. This component is in long-term decline in both bands. The fitted models were shown in the top panel of Figure 9. The error bars of the light curves are exceedingly small, which renders the fitting statistically unacceptable and suggests ultimate incorrectness of the model due to the existence of sub-structures. However, variation of the smooth model curves in the figure match reasonably well with observational data.

The timescales are $\tau_s = 9630$ s and $\tau_d = 9261$ s in the soft band and $\tau_s = 14756$ s and $\tau_d = 6796$ s in the hard band, which are consistent with the pattern that hard band rises slowly and decays fast. The fitted long-term declining background was subtracted from the observed light curve of each band to leave out the flare-only component. We plotted hardness ratio variation with time and flux in the bottom panel of Figure 9 and the right panel of Figure 10. Figure 10 shows the same patterns as panels (a) and (b) of Figure 5. The effective photon indices ($\Gamma_{\text{eff}} \sim 1.84$) are also close to the photon indices of the flaring components that are listed in Table 3.
