SLAC-PUB-9679
astro-ph/0210214
March 2003

IMPLICATIONS OF VARIABILITY PATTERNS OBSERVED IN TEV BLAZARS ON THE STRUCTURE OF THE INNER JET

Chiharu Tanihata[1,2], Tadayuki Takahashi[1,2], Jun Kataoka[3], and Greg M. Madejski[4]

to appear in ApJ

ABSTRACT

The recent long look X–ray observations of TeV blazars have revealed many important new features concerning their time variability. In this paper, we suggest a physical interpretation for those features based on the framework of the internal and external shock scenarios. We present a simplified model applicable to TeV blazars, and investigate through simulations how each of the model parameters would affect to the observed light curve or spectrum. In particular, we show that the internal shock scenario naturally leads to all the observed variability properties including the structure function, but for it to be applicable, the fractional fluctuation of the initial bulk Lorentz factors must be small, $\sigma'_\Gamma \equiv \sigma_\Gamma/\Gamma_{\rm avrg} \ll 0.01$. This implies very low dynamical efficiency of the internal shock scenario. We also suggest that several observational quantities – such as the characteristic time scale, the relative amplitude of flares as compared to the steady ("offset") component, and the slope of the structure function – can be used to probe the inner jet. The results are applied to the TeV blazar Mrk 421, and this, within the context of the model, leads to the determination of several physical parameters: the ejection of a shell with average thickness of $\sim 10^{13}$ cm occurs on average every 10 minutes, and the shells collide $\sim 10^{17}$ cm away from the central source.

Subject headings: BL Lacertae objects: individual (Mrk 501, PKS 2155–304, Mrk 421) — galaxies: active — radiation mechanisms: non-thermal — X–rays: galaxies

1. INTRODUCTION

Blazars are active galactic nuclei exhibiting the most rapid and largest amplitude variability of all AGN. Historically, radio observations first revealed that the emission was luminous and rapidly variable. With that, assuming that the radio emission was due to synchrotron radiation, calculations of Compton up-scattering predicted much higher X–ray fluxes than the observed values unless the radio emission was relativistically beamed (Hoyle et al. 1966; Jones, O'Dell, & Stein 1974). This led to our current model for blazars where the entire electromagnetic emission arises in a relativistic jet pointing close to the line of sight (Blandford & Rees 1978). Subsequent radio observations using the Very Long Baseline Interferometry (VLBI) showed superluminal motion in many sources, which served as the direct evidence for the relativistic motion.

The broadband spectra of blazars consist of two peaks, one in the radio to optical–UV range (and in some cases, reaching to the X–ray band), and the other in the X–ray to γ–ray region. From the high polarization of the radio to optical emission, the lower energy peak is best interpreted as produced via the synchrotron process by relativistic electrons in the jet. The higher energy peak is believed to be due to Compton up-scattering by the same population of relativistic electrons. Several possibilities exist for the source of the seed photons; these can be the synchrotron photons internal to the jet (Jones, O'Dell, & Stein 1974; Ghisellini & Maraschi 1989), but also external, such as from the broad emission line clouds (Sikora, Begelman, & Rees 1994) or from the accretion disk (Dermer, Schlickeiser, & Mastichiadis 1992; Dermer & Schlickeiser 1993).

Blazars are commonly detected as γ–ray sources. A number of them with peak synchrotron output in the X–ray range also have been detected in the TeV range with ground-based Cherenkov arrays. These are the so-called "TeV blazars." In TeV blazars, X–rays provide the best means for studying variability properties: this is because X-ray flux is presumably produced by electrons that are accelerated to the highest energies (where the cooling time scales are most rapid), and thus the dilution by the non-varying components is the smallest.

Variability studies of blazars have entered a new stage after a number of continuous long-look X–ray observations conducted with the *ASCA* satellite. One such observation, of Mrk 421, conducted in 1998, showed for the first time that flaring is actually occurring on a daily basis, and that long-duration flares detected in previous observations were probably unresolved superpositions of multiple, more rapid flares (Takahashi et al. 2000). Such excellent data provided the new knowledge about the radiation processes, allowing an exploration of the actual dynamics of the particles that are accelerated in the jets.

In §2, we present the properties of X–ray variability observed from TeV blazars, and summarize the issues we need to explain. In §3, we consider the internal shock scenario, which involves shells propagating rapidly along the jet. Using this scenario, we simulate expected X-ray light curves, and study how various quantities that can be measured from observations depend on the input parameters to the model. In §4, we compare the results of simulations

[1] Institute of Space and Astronautical Science, 3-1-1 Yoshinodai, Sagamihara, 229-8510, Japan
[2] Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan
[3] Department of Physics, Tokyo Institute of Technology, Tokyo, 152-8551, Japan
[4] Stanford Linear Accelerator Center, Stanford, CA, 94309-4349, USA

Work supported in part by the Department of Energy contract DE-AC03-76SF00515.

to the data to determine whether the observed variability features can be reproduced, and if so, what are the implications on model parameters. We briefly consider the possibility of the external shock scenario in §5, and give a summary in §6.

2. VARIABILITY PROPERTIES OBSERVED IN TEV BLAZARS

One of the most surprising results from the long look observations of TeV blazars was the repeated occurrence of flares with a time scale of $\sim$ one day. This was first observed in the 7–day observation of Mrk 421 in 1998 (Takahashi et al. 2000), and was confirmed via the 10–day observations of both Mrk 501 and PKS 2155–304 in 2000 (Tanihata et al. 2001). The latter two sources were in a relatively low flux state compared to previous observations, which implied that a high state is not a requirement for rapid variability. In other words, this feature indicates that the time scale of the rise and decay of the flares are similar to the time scale of the repeating of the flares.

Another observational fact is that the X–ray flares always appear to lie on top of an underlying offset-like component (see, e.g., light curves in Urry et al. 1997; Tanihata et al. 2001; Zhang et al. 2002). At this point, we cannot distinguish whether this is due to superpositions (or pile-ups) of flares, or whether there is a separate, non-varying offset component, but the observations indicate that there is always some component such that a flare does not start from zero, but from a level comparable to the flare amplitude.

An important advantage of long continuous observations is that the variability can be treated statistically. Recent structure function analysis have showed that there is clearly a characteristic time scale $t_{\rm chr}$ of an order of a day for each of three TeV blazars, Mrk 421, Mrk 501, and PKS 2155–304, and also that the variability power at shorter time scales than $t_{\rm chr}$ is strongly suppressed (Kataoka et al. 2001; Tanihata et al. 2001). Together with the near-symmetry of the flares and also by comparing the various possible time scales, it was suggested that this $t_{\rm chr}$ is determined by some dynamical time scale, instead of energy dependent cooling or acceleration times (Kataoka 2000).

X–ray emission from TeV blazars shows spectral variability, where the common trend indicates a harder X–ray spectrum for higher intensity states. Furthermore, the improved spectral coverage allows a measurement of the exact location of the synchrotron peak. Perhaps the most striking case is Mrk 501, where the peak frequency shifted from below 1 keV to over 100 keV during the very high flux state (Pian et al. 1998; Tavecchio et al. 2001). Likewise, Fossati et al. (2000b) has shown using the *BeppoSAX* observations of Mrk 421 in 1997 and 1998 that the peak frequency shifted to the higher energy for higher intensity states; this is particularly apparent during the flare observed in 1998 April. This was also shown in the long look observation of Mrk 421 in 1998 using the combined *ASCA* and *EUVE* spectrum (Tanihata et al. 2002). Through a detailed analysis of the spectral evolution at the rise of the flare, we found several flares that start to appear from the higher energy, which strongly suggests that an appearance of a new harder (i.e., higher synchrotron peak frequency) emission component is associated with the generation of the flare.

Summarizing, the following features are the issues concerning the variability derived from observations, that we need to explain in considering any model: (1) Daily-flares (i.e. $t_{\rm cycle} \sim t_{\rm chr}$), (2) the offset component, (3) the structure function, and (4) the energy dependence.

3. THE INTERNAL SHOCK SCENARIO

Among many studies addressing the mechanism of particle acceleration in jets, acceleration via shocks appears to be the most viable. Such shocks can efficiently accelerate particles to very high energies (e.g. Longair 1994; Bell 1978; Drury 1983; Blandford & Eichler 1987; Jones & Ellison 1991), and since it is highly probable that shocks form inside jets, we consider it in more detail here.

In order to form a shock, there must be a large velocity difference between the colliding parcels of matter. The key idea of the internal shock model is a shock-in-jet scenario, where the central engine injects energy into the jet in a discontinuous manner, producing individual shells having slightly different bulk Lorentz factors and energies. If this occurs, there will be collisions by a faster shell catching up to a slower one, forming a shock. This internal shock scenario is among the most promising models to explain the emission of gamma–ray bursts (e.g. Sari & Piran 1995), although it was originally suggested in reference to AGN more than 20 years ago by Rees (1978). It has been recently suggested that this model could be successfully applied to blazars, as it can explain some of their basic properties such as the low efficiency (Ghisellini 2001; Spada et al. 2001; Sikora, Błażejowski, Begelman, & Moderski 2001).

3.1. *The Model*

In order to investigate whether the observed variability properties can be reproduced in the internal shock scenario, we developed a simulation code. Recently, Spada et al. (2001) has given a detailed calculation covering the formation, propagation, and collision of such shells. Included are hydrodynamic calculations to determine the structure of the shock fluid, and the full radiation spectrum is derived by summing up all the locally produced spectra from the electrons accelerated by the shocks. The model and simulations presented here are a much simplified version, to be compared specifically against the actual measurements of TeV blazars.

Due to the difference of the initial velocity of the shells ejected from the base of the jet, collisions occur when a faster shell catches up to a slower one. This is where the shock is formed, electrons are accelerated, and lose energy through radiation. Since we consider only the colliding shells, in our simulations we generate pairs of shells with one having a bulk Lorentz factor (BLF) of Γ_1 ejected from the base of the jet, and a following shell with a BLF of Γ_2 ejected after an initial separation distance of D_0. If Γ_2 is larger than Γ_1, the latter shell will catch up to the former one, and the two will merge into a single shell producing a shock. Each collision then generates radiation (called hereafter a "shot") for a duration of $t_{\rm shot}$. The time profile of each shot is assumed to have a symmetric linear rise and decay. The collisions are distributed randomly in time following a Poisson distribution, and superposition of

these individual shots results in the output light curve.

For simplicity, the rest mass of all shells is assumed to be the same, and the shell thickness l is assumed to be equal to the initial separation D_0. The average frequency of the collision is set to be consistent with the separation of the shells (i.e. $F_{\rm col} = c/2(D_0 + l)$), and only the first collision is considered. The initial BLFs are assumed to be distributed around an average value $\Gamma_{\rm avrg}$, following a Gaussian distribution with its width described by the sigma, σ_Γ. In this case, there are only three input parameters: two of them describe the distribution of the initial BLFs of the ejected shells ($\Gamma_{\rm avrg}$, σ_Γ), and one describes the initial separation of the two colliding shells (D_0). We also define the fractional width $\sigma'_\Gamma \equiv \sigma_\Gamma/\Gamma_{\rm avrg}$.

With the assumptions described above, using the momentum and energy conservation laws in an inelastic collision, the BLF of the merged shell is

$$\Gamma_{\rm m} = (\Gamma_1\Gamma_2)^{\frac{1}{2}}. \quad (1)$$

The newly generated total internal energy is

$$E_{\rm m} = Mc^2(\Gamma_1 - \Gamma_{\rm m}) + Mc^2(\Gamma_2 - \Gamma_m), \quad (2)$$

and thus the dynamical efficiency is given by

$$\eta = \frac{Mc^2(\Gamma_1 - \Gamma_{\rm m}) + Mc^2(\Gamma_2 - \Gamma_m)}{Mc^2\Gamma_1 + Mc^2\Gamma_2}. \quad (3)$$

We assume that all of the newly generated internal energy is converted to the random energy of the electrons.

Each collision will take place at distance

$$D \sim \frac{2\Gamma_1^2\Gamma_2^2}{\Gamma_2^2 - \Gamma_1^2} D_0 \quad (4)$$

from the core, where in making the above approximation, we used Γ_1^2, $\Gamma_2^2 \gg 1$. We assume that the jet is collimated into a cone with an opening angle $\theta \sim 1/\Gamma$. The radius of the shell at the location of the collision can thus be written as

$$R = D\tan\theta \sim \frac{D}{\Gamma}. \quad (5)$$

The time duration of each shot $t_{\rm shot}$ results from several competing time scales: the acceleration and cooling time scales, and the dynamic time scale, which includes the hydrodynamic time and the angular spreading time. It has been shown for TeV blazars that the cooling or acceleration times are significantly shorter than the observed rise and decay time scales of the X–ray and gamma–ray flares (Kataoka 2000; Tanihata et al. 2001), and thus the dynamical time scale most likely determines $t_{\rm shot}$, and also filters out any faster variability.

The hydrodynamic time scale is determined by the time that the shock takes to cross the shell. Using the conservation of mass, energy, and momentum at the shock, and the equality of pressure and velocity along the contact discontinuity, the Lorentz factors of the forward and reverse shocks $\Gamma_{\rm fs}$ and $\Gamma_{\rm rs}$ are given by (Kobayashi, Piran, & Sari 1997)

$$\Gamma_{\rm fs} \simeq \Gamma_{\rm m}\sqrt{(1 + \frac{2\Gamma_{\rm m}}{\Gamma_1})/(2 + \frac{\Gamma_{\rm m}}{\Gamma_1})} \quad (6)$$

$$\Gamma_{\rm rs} \simeq \Gamma_{\rm m}\sqrt{(1 + \frac{2\Gamma_{\rm m}}{\Gamma_2})/(2 + \frac{\Gamma_{\rm m}}{\Gamma_2})}. \quad (7)$$

We note that Equations (6) and (7) are for the case of relativistic shocks, where the adiabatic index is 4/3. For non-relativistic shocks, the adiabatic index should be 5/3, but since this does not make a large difference, we use this formula for all shocks in this paper. Following Kobayashi, Piran, & Sari (1997), we estimate the shock crossing time $t_{\rm crs}$ by the longer time scale of the two shocks to cross the shell; this happens to be the time which the reverse shock takes to cross the faster shell. Because the emitting region is moving with a relativistic speed towards the observer, the observed duration of the flare will be shortened by a factor of $(1 - \beta\cos\theta) \sim 1/\Gamma_{\rm m}^2$, and thus,

$$t_{\rm crs} = \frac{l}{c\ (\beta_2 - \beta_{\rm rs})\ \Gamma_{\rm m}^2} \sim \frac{2}{c\ \Gamma_{\rm m}^2}\left(\frac{1}{\Gamma_{\rm rs}^2} - \frac{1}{\Gamma_2^2}\right)^{-1} D_0 \quad (8)$$

where $\beta_2 c$ and $\beta_{\rm rs} c$ are respectively the velocity of the catching up shell and the reverse shock. The angular spreading time is given by

$$t_{\rm ang} = \frac{R}{c\ \Gamma_{\rm m}}. \quad (9)$$

The average time cycle of the collisions is determined from the frequency of the ejection of the shells, and thus

$$t_{\rm col} = \frac{4}{c}\ D_0. \quad (10)$$

For the $t_{\rm shot}$ in the simulations, we use the longer one of the $t_{\rm crs}$ and $t_{\rm ang}$ for each collision. As a result, since $t_{\rm crs}$ is always longer than $t_{\rm ang}$ for all cases, thus $t_{\rm shot}$=$t_{\rm crs}$.

3.2. *Results and the Dependence on Each Parameter*

We first consider the case of $\Gamma_{\rm avrg}$=10, σ'_Γ=0.05, and D_0=3×10^{13} cm. The distribution of the collision distances in this case is shown in Figure 1(a). Our simulations show that the collisions take place at distances ranging from $D \sim 10^{17}$ cm up to $D \sim 10^{20}$ cm. Figure 1(b) shows the amount of the newly generated internal energy for each collision $E_{\rm m}$ plotted against the collision distance D, and Figure 1(c) is the time scale of each generated flare $t_{\rm shot}$ plotted against D. It is apparent that $E_{\rm m}$ is larger and $t_{\rm shot}$ is shorter for collisions which occur at smaller D.

A portion of the simulated light curve for this case is shown in Figure 1(d). It can be seen that the overall light curve is characterized by repeating flares having time scales of $\sim 50 - 100$ ks, resembling the light curves observed from TeV blazars. This results from the fact that the shells which collide at smaller distances have larger $E_{\rm m}$ and shorter $t_{\rm shot}$; with this, the amplitude of the emission becomes much larger compared to the shots generated from collisions at larger distances. Accordingly, only the shots produced by collisions at the smallest distances will be apparent as flares in the observed light curve. The average frequency of the collision in this case is $F_{\rm col} = 0.24$ mHz, corresponding to one collision per ∼4 ks on average, while the number of flares which are observed is only about 6% of the total number of shots. The distribution of the resolved flares are shown as the shadowed area in Figure 1(a).

The calculated structure function for the simulated light curve is shown in Figure 1(e). A clear break, indicating a characteristic time scale $t_{\rm chr}$ is seen, and the slope at the shorter time scales is $\beta \sim$2. This indicates that the values of $t_{\rm shot}$ of observable flares are restricted to a rather narrow range, which determines $t_{\rm chr}$, and that there is very little variability power below $t_{\rm chr}$ – exactly what we have observed from TeV blazars.

Another remark regards the offset-like component. This is due to the other 94% of the collisions that generate the longer, smaller amplitude shots, which overlap each other, resulting in the observed offset. In the following, we show how each of the parameters affects on these results.

3.2.1. *Dependence on* $\Gamma_{\rm avrg}$

We first consider the effect of the change in the average BLF $\Gamma_{\rm avrg}$, while the relative width σ'_Γ, and D_0 are kept constant. When $\Gamma_{\rm avrg}$ increases by a factor of a, Equation 4 implies that the collisions will occur at a distance D which is greater, by a factor of a^2. On the other hand, Equation 8 shows that the the duration of each shot will be conserved, and thus only the amplitude will be different in the resulting light curve.

3.2.2. *Dependence on* D_0

We then consider the case where the initial separation of the colliding shells D_0 changes, while $\Gamma_{\rm avrg}$ and σ_Γ are kept constant. When D_0 is larger by a factor of a, Equations 8 and 10 imply that both $t_{\rm crs}(=t_{\rm shot})$ and $t_{\rm col}$ will be longer by a factor of a, and thus the simulated light curve will simply be a stretched-in-time version of the original one – the amount of relative offset components will be identical. The collision distance D will become larger by a factor of a.

3.2.3. *Dependence on* σ'_Γ

This turns out to be the most important parameter. Here, we fix $\Gamma_{\rm avrg}$=10 and D_0=3×10^{13} cm, and simulate light curves with different σ'_Γ. First, it can be shown from Equation 4 that as σ'_Γ becomes larger, the collisions start to take place at shorter distances from the core.

Since the characteristic time scale $t_{\rm chr}$ is always determined by the shots due to collisions which took place at the smallest distances, the time scales of the observable flares in the light curve become shorter when σ'_Γ is larger. This is shown in the simulated light curve for different σ'_Γ (σ'_Γ=0.001, 0.005, and 0.05) in Figure 2. Here, the flares appear more spiky as σ'_Γ becomes larger. Note that since all 3 simulations assume the same D_0, the number of collisions per unit time is the same for all 3 light curves. On the other hand, the number of visible flares in the light curves is clearly different. In order to see the difference in $t_{\rm chr}$, we calculated the structure function for the simulated light curves. This is shown in Figure 3. The break, indicating $t_{\rm chr}$, is clearly seen to shift to the shorter time scales as σ'_Γ becomes larger.

The differences in the spikiness in the light curve can be regarded as the differences in the relative amplitude of the flare and offset components. Importantly, what we have shown is that this relative amplitude, $r_{\rm fo}$, changes with the value of σ'_Γ. This indicates that, assuming that all of the offset component is generated by the internal shocks, we can compare the observed $r_{\rm fo}$ to the that derived through simulations. Indeed, the actual observations have clearly shown a larger offset component than, for instance is apparent in Figure 2(c), which suggests that σ'_Γ must be relatively small. We will quantify this in section 4.

3.2.4. *Effect of Fluctuation in l*

We have so far fixed the shell thickness l to be equal to the initial separation of the two shells D_0. Here at the end of this section, we consider the effect when l also fluctuates. We assume that l is distributed around an average D_0, following a Gaussian function with its width described by the relative sigma, σ'_l ($\equiv\sigma_l/D_0$). We also make an assumption that the mass of the ejected shells scales with the shell thickness (i.e. density of shell is constant).

The effect of the fluctuation is demonstrated in the plotted correlation of the duration of the radiation ($t_{\rm shot}$) and the dissipated energy ($E_{\rm m}$) for each collision in Figure 4. For the case where l is fixed (a), $t_{\rm shot}$ and $E_{\rm m}$ become simply respectively shorter and larger as $\Delta\Gamma$ ($\equiv\Gamma_2-\Gamma_1$) of the colliding shells increases, indicating the clear trend in the correlation. When l fluctuates, the spread becomes wider, as shown in Figures 4(b)(c).

This indicates that there will be increasing power in the faster variability time scales when σ'_l becomes larger. For a shell with smaller l, $E_{\rm m}$ will be reduced because of the smaller mass of the shell. However, since the shock crossing time $t_{\rm crs}$ ($=t_{\rm shot}$) also becomes shorter, the *amplitude* of the shot will be as large as that of a longer shot due to a larger l. This is demonstrated in the calculated structure function from the light curves simulated for different σ'_l, shown in Figure 5. While the break is seen to stay nearly at the same value, the slope at the faster time scale appears to flatten as σ'_l increases.

3.3. *Summary of the Simulations*

The simulation results can be summarized as follows:

1. Collisions of two shells which had the largest relative velocity, and accordingly collided at the shortest distances (D), are the shots which appear as the strongest observable flares in the light curve.

2. These shots determine the characteristic time scale ($t_{\rm chr}$) of the variability. There is very little variability power on time scales shorter than this $t_{\rm chr}$.

3. An offset component will arise from emission due to overlapping shots produced by collisions at larger distances. The relative amplitude of flare to offset ($r_{\rm fo}$) is a function of the initial width of the bulk Lorentz factor (σ'_Γ).

4. The dependences of each parameter are:

   - $\Gamma_{\rm avrg}$ determines $E_{\rm m}$: higher $E_{\rm m}$ yields higher $\Gamma_{\rm avrg}$.
   - D_0 determines the normalization of the time series: longer $t_{\rm chr}$ is a result of larger D_0.
   - σ'_Γ has an effect on $t_{\rm chr}$ and $r_{\rm fo}$: larger σ'_Γ results in smaller $t_{\rm chr}$ and larger $r_{\rm fo}$.
   - σ'_l has effect on the slope β of the structure function: β flattens as σ'_l increases.

4. APPLICATION OF THE INTERNAL SHOCK MODEL TO THE X-RAY DATA FOR TEV BLAZARS

4.1. *Light Curves*

In the previous section, we have shown that the internal shock scenario naturally predicts the main features of

blazar light curves described in §2, under the condition of $\sigma'_\Gamma \ll 0.01$. One prediction is that the typical time scale of the observed flares ($t_{\rm chr}$) always becomes similar to the time scale of the flare *cycles*. This is very much consistent with the actual observations, where the "day-scale" flares are observed "daily."

The next prediction concerns the features observed in the structure function analysis. The observations show that all TeV blazars show a break in the structure function, with $t_{\rm chr} \sim 1$ day. The slope below this $t_{\rm chr}$ is steep (>1), suggesting that very little variability power is below this $t_{\rm chr}$. We have shown that the structure functions calculated from the simulated light curves show the same features. In fact, this provides an explanation for the non-existence of shorter time scale variability – there are no collisions until a certain distance D from the central core.

In order to see the actual σ'_Γ dependence on $t_{\rm chr}$, we simulated light curves for a series of σ'_Γ for the case of $\Gamma_{\rm avrg}$=10, and D_0=3×10^{13} cm as in §3.2.3. We calculated the structure function for each simulated light curve, and assumed that it is well described by two power laws. We then fitted for the two slopes, and estimated the characteristic time scale as the point where the two slopes cross (as shown as the dotted lines in Figure 3). The derived values of $t_{\rm chr}$ plotted as a function of σ'_Γ are shown in Figure 6. It appears that $t_{\rm chr}$ becomes shorter with larger σ'_Γ, and the best-fit power-law gives a slope of ~ -0.9. Note that we have shown in §3.2.1 that the time scales do not depend on $\Gamma_{\rm avrg}$ if σ'_Γ is constant.

Another point regards the presence of the offset component. As discussed in §2, the observed flares in TeV blazars always appear to lie on top of an underlying offset component. The data do not tell us directly whether this offset is due to flares overlapping each other, or due to some steady emission component. What we have shown here is that considering the internal shock scenario, there will always be many overlapping flares which will appear as the offset component.

The amount of this offset component should also be useful in modeling of an actual observation, as this suggests that σ'_Γ can be estimated from an observed light curve if it has a sufficient length to estimate the relative amplitude of the flare and offset component. In an attempt to quantify the offset component, here we use the parameter $r_{\rm fo}$, describing the relative amplitude of the flare and offset component as follows. We first generate a histogram of the count rates which form a peak. Since the lower and higher end of the peak represent the minimum and maximum count rates in the light curve, these two can be considered as an indicator of the amplitude of the offset component, and the offset-plus-flare component. As there are fluctuations, we define the offset amplitude as the point where 10% of total counts is reached, and the offset-plus-flare amplitude as the point where 90% of the total counts is included. The amplitude of the flare component is estimated by subtracting the offset amplitude from the offset-plus-flare amplitude.

We calculate this flare-to-offset ratio $r_{\rm fo}$ for the same set of simulated light curves used in Figure 6. The result is shown in Figure 7, which shows that $r_{\rm fo}$ increases with increasing σ'_Γ. As we showed in §3, $r_{\rm fo}$ does not depend on $\Gamma_{\rm avrg}$ nor D_0, which means that σ'_Γ can be directly estimated from a $r_{\rm fo}$ given by an observed light curve.

Finally, we note that the $t_{\rm chr}$ plotted in the Figure 6 is for the case of D_0=3×10^{13} cm. Given that the time series scales linearly with D_0, the vertical axis in Figure 6 can be regarded as $t_{\rm chr}/D_{0,3\times10^{13}}$, where $D_{0,3\times10^{13}}$ denotes D_0 in units of 3×10^{13} cm. Accordingly, this indicates that D_0 can also be estimated if $r_{\rm fo}$ and $t_{\rm chr}$ can be measured. This is interesting, given that the initial separation is a value determined from the frequency of the emitted shells. Thus, D_0 should reflect the frequency of the activity of the central engine, that generates and ejects individual shells.

4.2. *Variation of the Synchrotron Peak Frequency*

The important observational result from the spectral analysis was that the peak synchrotron frequency increases during both high intensity states, and also within the daily flares. In this section, we discuss whether this can be interpreted within the same model. We start with formulating the peak energy, by generally following the prescriptions of Inoue & Takahara (1996) and Kirk, Rieger & Mastichiadis (1998).

The peak frequency of the synchrotron spectrum reflects the maximum energy of the accelerated electrons, which is determined from where the cooling and the acceleration time scale becomes comparable. The cooling time of the electrons due to synchrotron and inverse Compton emission can be written as (Rybicki & Lightman 1979),

$$\tau_{\rm cool}(\gamma) = \frac{3m_{\rm e}c}{4(U_{\rm B}+U_{\rm soft})\sigma_{\rm T}\gamma}, \tag{11}$$

where $U_{\rm B}$ and $U_{\rm soft}$ are the energy densities of the magnetic field and the soft photons to be upscattered in the Inverse Compton process, $m_{\rm e}$ the rest mass and and γ is the random Lorentz factor of the electron, and $\sigma_{\rm T}$ is the Thomson cross section.

The acceleration time is not as well understood as the cooling time. Perhaps the most promising theory suggests that it is determined from the time scale of the 1st order Fermi acceleration process operating in a shock. In this case, the acceleration time can be approximated by considering the mean free path $\lambda(\gamma)$ for the scattering of the electrons with the magnetic disturbances. Taking the mean free path to be proportional to the Larmor radius by introducing another parameter ξ ($\lambda \equiv \xi\frac{\gamma mc^2}{eB}$), the acceleration time is given by,

$$\tau_{\rm acc}(\gamma) = \frac{20\lambda(\gamma)c}{3v_{\rm s}^2} = \frac{20m_{\rm e}c^3}{3eB}\frac{\xi}{v_{\rm s}^2}\gamma, \tag{12}$$

where $v_{\rm s}$ is the shock velocity. By equating the radiative cooling and acceleration times in Equations 11 and 12, the maximum energy of the electrons is,

$$\gamma_{\rm max} = \frac{v_{\rm s}}{c}\left(\frac{9eB}{80(U_{\rm B}+U_{\rm soft})\sigma_{\rm T}\xi}\right)^{\frac{1}{2}}. \tag{13}$$

Furthermore, in the case of TeV blazars, the Compton cooling for the energy range of electron emitting X–rays is strongly suppressed by the Klein-Nishina cutoff, and thus $U_{\rm soft}/U_{\rm B} \ll 1$ (e.g. Kataoka 2000; Li & Kusunose 2000). Thus for this case, the observed maximum synchrotron frequency can be written as

$$\nu_{\rm max} = 1.2\times10^6 B\delta\gamma_{\rm max}^2 \tag{14}$$

$$\simeq 2.49\times10^{21}\frac{\delta_{10}}{\xi}\left(\frac{v_{\rm s}}{c}\right)^2 {\rm Hz}, \tag{15}$$

where δ_{10} denotes the beaming factor in units of 10. Note that the above implies that when synchrotron losses dominate (as is likely to be the case for TeV blazars), the observed peak frequency is independent on B.

We will first consider the differences in the synchrotron spectrum for the flare components and the underlying offset component. Above, we have shown that in the framework of the internal shock scenario, the observed flares can be regarded as being due to collisions which had the largest $\Delta\Gamma$ within the distribution, while the offset component results from emission due to fainter, overlapping flares produced by collisions that had smaller $\Delta\Gamma$.

Since the collisions which generate the observed flares occur at shorter distances D, the magnetic field B is expected to be higher. Accordingly, the radiative cooling times of the electrons are shorter there, preventing electrons from being accelerated to higher energies. On the other hand, Equation 12 suggests that the acceleration time will also be shorter, due to the larger B, and also due to larger shock velocity $v_{\rm s}$. Equation 15 shows that concerning the observed synchrotron peak *frequency*, the changes in B cancel out and only values of $v_{\rm s}$, δ, and ξ affect $\nu_{\rm max}$.

In Figure 8(a), we show the calculated shock velocity $v_{\rm s}$ for each collision (in units of c; $\beta_{\rm s}$=$v_{\rm s}/c$) as a function of D, for the case of $\Gamma_{\rm avrg}$=10, σ'_Γ=0.05, and D_0=3×10^{13} cm (similar to the ones used in Figure 1). It shows that $v_{\rm s}$ decreases as D increases, following a relation of $\beta_{\rm s}\propto D^{-1}$. Accordingly, if ξ were to be similar for all collisions, Equation 15 shows that $\nu_{\rm max}$ will be higher for collisions at smaller D. Since these are the shots that appear as flares, this is consistent with the actual observations, where the flare spectrum components show higher $\nu_{\rm max}$ than the offset component. The calculated $\nu_{\rm max}$, considering the Bohm limit (i.e. ξ=1) is shown in Figure 8(b). The dependence of $\nu_{\rm max}$ on D is $\nu_{\rm max}\propto D^{-2}$.

Finally, in addition to the spectral variation during the day-scale flares, the same trend regarding the relationship between the flux and synchrotron peak frequency is also observed in comparing different observations. The same discussions as above will hold for the case when σ'_Γ becomes larger. The shock velocity will become higher for the collisions which generate the observed flares, and accordingly the average $\nu_{\rm max}$ will be higher. Another way to generate a change in the average $\nu_{\rm max}$ is the change in the $\Gamma_{\rm avrg}$. If $\Gamma_{\rm avrg}$ increases, it can be shown from equation 15 that $\nu_{\rm max}$ will also increase.

4.3. *The Case of Mrk 421*

Here we consider the observed light curve of Mrk 421 during the *ASCA* long look in 1998 April. In modeling the observed light curve, we wish to reproduce the observables determined from these data such as the characteristic time scale $t_{\rm chr}$ and the flare-to-offset ratio $r_{\rm fo}$ (as defined in § 4.1).

The first step is to estimate σ'_Γ from the observed $r_{\rm fo}$. The normalized light curve (in counts $\rm s^{-1}$) and the calculated histograms of the count rate are shown in Figure 9. The calculated $r_{\rm fo}$=0.7 is compared with the derived relation of σ'_Γ and $r_{\rm fo}$ from the simulations shown in Figure 7, which gives an estimate of $\sigma'_\Gamma\sim$0.001. Then, from the observed characteristic time scale $t_{\rm chr}\sim$ 40 ks (Tanihata et al. 2001), D_0 can be estimated from Figure 6 as

$$D_0\sim\frac{40}{120}\times3\times10^{13}\sim1\times10^{13}{\rm cm}. \quad (16)$$

From Equation 10, this would give $t_{\rm col}\sim$ 1300 seconds, which indicates that the shells are ejected from the central engine on average every ∼10 minutes.

On the other hand, we also remark here that the $r_{\rm fo}$ derived in Figure 7 is for the variations of the dissipated energy. Concerning the $r_{\rm fo}$ for a particular energy, this will have an energy dependence, which requires a proper treatment of synchrotron formulae, and will also depend on the energy dependence of the efficiency of the energy that is to be converted to radiation. This might be further investigated by studying the difference of $r_{\rm fo}$ between different observations. In particular, a similar analysis as above, applied to the *BeppoSAX* observation of Mrk 421 in 2000 May, shows a larger value of $r_{\rm fo}$=1.7, whereas the value is somewhat similar to the *ASCA* observations for the *BeppoSAX* observations in 2000 April. This is shown in Figure 10 and 11. An examination of the light curve during 2000 May reveals that flares with shorter time scales are dominant in the light curve. These differences in the characteristic time scales can be seen clearly in the structure functions calculated for each observation, shown in Figure 12. The break in the structure function is at ∼10 ks for the 2000 May data, whereas it is at ∼30–40 ks for the other two. While this might be only a coincidence, the 3 observations show a trend that $r_{\rm fo}$ is larger when $t_{\rm chr}$ is shorter: this is actually the trend expected from our model – when σ'_Γ is larger, $r_{\rm fo}$ is larger and $t_{\rm chr}$ is shorter.

4.4. *The Efficiency*

One concern is that the dynamical efficiency η of the collisions is rather small. For instance, for the parameters derived above for Mrk 421, η given by Equation 3 is $\sim10^{-6}$ for the collisions that occur at the shortest D. For the more distant collisions, this becomes even smaller.

It has been claimed that the jet cannot radiate all of its energy in the sub-parsec region considered here, since a substantial power must be transported to the kpc-scale radio lobes. With this, the efficiency must be low. However, this seem too low – as this means that only 10^{-6} of the total energy is radiated at the base of the jet, which indicates that the jet is 10^6 times energetic than is observed via the blazar phenomenon. On the other hand, since there is no measure of the energetics of the entire jet, such low radiative efficiency near the core cannot be simply ruled out. Nevertheless, as it does seem rather low, we discuss how this could be increased.

The very low efficiency mainly results from the fact that we intend to reproduce the offset component, which reduces the σ'_Γ, and thus reduces the velocity difference of the two colliding shells. Thus, the easiest way to increase the efficiency would be to have larger $\Gamma_2-\Gamma_1$, but this would make the flare time scales shorter than the flare cycles, which conflicts with the observed flares that occur repeatedly (daily-flares).

Part of the low efficiency also is due to the fact that we are assuming a Gaussian distribution for the initial BLF, which emphasizes the effect. The origin of the modulation of the Lorentz factor is not known; it can be due to any physical condition of the central engine that is not stable,

such as instabilities in the innermost parts of an accretion disk or magnetic eruptions in the corona (e.g., Sikora & Madejski 2000). Thus one way to increase the efficiency would be to assume a broader distribution, such as a flat distribution within a certain range (e.g., Spada et al. 2001; Kobayashi, Piran, & Sari 1997), which would increase the number of efficient collisions. According to our simulations, this was shown to increase the efficiency by roughly one order of magnitude. We can consider even more extreme distributions where the values of the BLFs are concentrated at the low and high end of the distribution. At most, this could improve the efficiency by another order of magnitude, but not any more, since $\Gamma_2 - \Gamma_1$ is always small. If this is still too inefficient, this would indicate that there is a problem assuming random distribution of BLFs for the shells in the internal shock scenario. In this case, we would probably need to consider another possible origin of the offset component. For instance, if there is emission from significantly larger distances, such as by reconfinement shocks or by shocks due to collisions with inhomogeneities that have much larger radial extensions, this may produce a low-amplitude relatively steady component. For this case, the efficiency can be much larger. On the other hand, we emphasize here that the internal shock scenario alone reproduces successfully many of the observed features such as the daily-flares, the existence of the characteristic time scale $t_{\rm chr}$ and the non-existence of variability power below $t_{\rm chr}$.

5. THE EXTERNAL SHOCK SCENARIO

In the previous sections we have discussed how the internal shock model successfully reproduces the observed variability of the both the fluxes and spectra of blazars. While the this model works rather nicely naturally explaining the observed properties, we briefly consider the alternative model, the external shock scenario.

In the internal shock scenario, a shock is generated when a faster shell catches up to a slower shell. This is where electrons are accelerated to relativistic energies and radiate through synchrotron and inverse Compton emission. The alternative scenario is that these shells do not collide with each other but run into either an external material or external field, where the shock emerges (e.g. Dermer et al. 1999). This is somewhat similar to the acceleration mechanism which is usually considered in supernova remnants or the afterglows of gamma-ray bursts. In these two examples, the external material is provided by the interstellar medium, while in the case of blazars, the promising candidates for the external medium are broad line clouds and/or intercloud material. We will discuss below how this may work in blazars; below, we will call such external material "clumps."

Here, we assume a shell ejected with a Lorentz factor Γ, which runs into a broad line cloud "clumps" (or any other external material) at a distance D. In similarity to the case of the internal shock scenario, given that the radiation cooling time is much faster than the observed variability time scales, the observed variability time scales are most likely determined by the dynamical time scale. There are two time scales which may affect the dynamical time scale: the time for the shock to cross the region, and the angular spreading time. Here, in contrast to the internal shock scenario, the shock can be considered as relativistic.

For this case, the time for the shock to cross the shell is given by

$$t_{\rm crs} = \frac{l}{2c\Gamma^2}, \tag{17}$$

where l is the shell thickness. The angular spreading time is similar to the case of the internal shock model, and thus

$$t_{\rm ang} = \frac{R}{c\Gamma} = \frac{D}{c\Gamma^2}, \tag{18}$$

where R is the radius of the shell at distance D.

If the shock crossing time $t_{\rm crs}$ were to determine the observed time scale of the flares, which is of the order of a day, Equation 17 would suggest that the thickness of the shell must be as large as $l \sim 3 \times 10^{17}$ cm. This means that the central engine must be continuously ejecting material for 100 days to produce a single shell – which is in severe conflict with the observed rate of flares which occur daily. Thus, in contrast to the case of the internal shocks, the angular spreading time $t_{\rm ang}$ should be the dominating time scale in the case of external shocks. For this case, to obtain flares with time scales of $\sim$ 1 day, the location of the shock is calculated to be at $D \sim 3 \times 10^{17} \Gamma_{10}^2$ cm, where Γ_{10}=$\Gamma/10$.

Because the emission is beamed, the observed time scale of $\sim$ 1 day reflects a $\sim$ 10 day emission in the jet co-moving frame. The daily occurrence of flares indicates that there must be at least 10 shells radiating at the same time, which requires at least 10 separate "clumps." The external fields must also be restricted in a rather limited range, as to keep the flare time scales similar. Figure 13 shows a simulated light curve for the case of Γ=15, and D=5.4 $\times$ 10^{17} cm (calculated for $t_{\rm shot} = t_{\rm ang}$ = 80 ks). Here, the average flare cycle is set to 5 ks, showing that it is possible for the offset component to be produced only if the flare cycle is high (i.e. substantial number of "clumps" of the external material available for collisions with the jet at the same time).

We also remark that for the case of the external shock model, the collisions are likely to occur farther away from the central source than for the internal shock model. A concern arises to whether the broad line clouds are still present at such distances. From observations of the time delay between the continuum and line variations in AGNs, the broad line regions are suggested to extend to distances around 10^{16-18} cm, although there are no measurements for blazars. If this is of the same order for blazars, the distance calculated above (for Γ=15, $t_{\rm shot}$=1 day) is actually at the right location, but may be close to the limit. Summarizing, we do not claim that the discussion above rules out the external shock scenario, but we remark that more fine-tuning of the parameters is necessary as compared with the internal shock model.

6. SUMMARY

The high-quality light curves and spectra obtained via the recent long look observations of TeV blazars provided the first opportunity to use the variability as a new tool to study the structure of jets in blazars in more detail than was previously possible. Our approach is unique as it investigates not only a single flare, but considers a series of

flares resulting from well-sampled, long-duration time series. We summarized the observed variability properties and suggested a physical interpretation to explain these features based on the framework of the internal shock scenario.

We presented a simplified model applied specifically for TeV blazars, and investigated through simulations how each of the parameters would affect the observed light curve or spectrum. In particular, we showed that the internal shock scenario naturally accounts for the observed variability properties, but it requires a condition that the fluctuation of the initial bulk Lorentz factors are small, $\sigma'_\Gamma \ll 0.01$. Remarkably, this explains both features observed in the flux but also in the spectral variability – the repeatedly occurring flares, the offset component, the structure function, and the shift of the synchrotron peak frequency. We also showed that several useful observational quantities can be used to probe the physical parameters of the inner jet: the characteristic time scale, the flare-to-offset ratio, and the slope of the structure function. We applied this model to the *ASCA* X–ray light curves of the TeV blazar Mrk 421, which allows us to determine several physical parameters of the jet such as the frequency of ejection of shells (on average one shell every 10 minutes), the average thickness of the shell ($\sim 10^{13}$ cm), and the location of their collisions (typically $\sim 10^{17}$ cm away from the central source). We also briefly commented on the external shock scenario, and claimed that this scenario is viable, but requires rather detailed fine-tuning to the parameters.

We thank Marek Sikora and Fumio Takahara for valuable discussions on this manuscript. We would also like to thank the anonymous referee for constructive comments to improve this paper. Support for this work was provided by the Fellowship of Japan Society for Promotion of Science for Young Scientists, and by NASA via Chandra grant no. GO0-1038A from SAO to Stanford University.

REFERENCES

Bell, A. R. 1978, MNRAS, 182, 147
Blandford, R. & Eichler, D. 1987, Phys. Rep., 154, 1
Blandford, R. D. & Rees, M. J. 1978, Phys. Scr, 17, 265
Dermer, C. D., Schlickeiser, R., & Mastichiadis, A. 1992, A&A, 256, L27
Dermer, C. D., & Schlickeiser, R. 1993, ApJ, 416, 458
Drury, L. O'C. 1983, Rep. Prog. Phys., 46, 973
Fossati, G. et al. 2000, ApJ, 541, 153
Fossati, G. et al. 2000, ApJ, 541, 166
Ghisellini, G., & Maraschi, L. 1989, ApJ, 340, 181
Ghisellini, G. 2001, ASP Conf. Ser. 227: Blazar Demographics and Physics, 85
Hoyle, F., Burbidge, G. R., & Sargent, W. L. W. 1966, Nature, 209, 709
Inoue, S., & Takahara, F. 1996, ApJ, 463, 555
Jones, F. C. & Ellison, D. C. 1991, Space Science Reviews, 58, 259
Jones, T. W., O'Dell, S. L., & Stein, W. A. 1974, ApJ, 188, 353
Kataoka, J., Ph.D Thesis, University of Tokyo
Kataoka, J. et al. 2001, ApJ, 560, 659
Kirk, J. G., Rieger, F. M., & Mastichiadis, A. 1998, A&A, 333, 452
Kobayashi, S., Piran, T., & Sari, R. 1997, ApJ, 490, 92
Longair, M. S. 1994, Cambridge: Cambridge University Press, —c1994, 2nd ed.,
Pian, E., et al. 1998, ApJ, 492, L17
Rees, M. J. 1978, MNRAS, 184, 61P
Rybicki, G. B., & Lightman, A. P. 1979, Radiative Processes in Astrophysics (New York: Wiley)
Sari, R. & Piran, T. 1995, ApJ, 455, L143
Sikora, M., Begelman, M. C., &Rees, M. J. 1994, ApJ, 421, 153
Sikora, M., & Madejski, G. 2000, ApJ, 534, 109
Sikora, M., Błażejowski, M., Begelman, M. C., & Moderski, R. 2001, ApJ, 561, 1154
Spada, M., Ghisellini, G., Lazzati, D., & Celotti, A. 2001, MNRAS, 325, 1559
Takahashi, T., et al. 2000, ApJ, 542 L105
Tanihata, C. et al. 2001, ApJ, 563, 569
Tanihata, C. et al. in preparation
Tavecchio, F. et al. 2001, ApJ, 554, 725
Urry, C. M. et al. 1997, ApJ, 486, 799
Zhang, Y. H., et al. 2002, ApJ, 572, 762

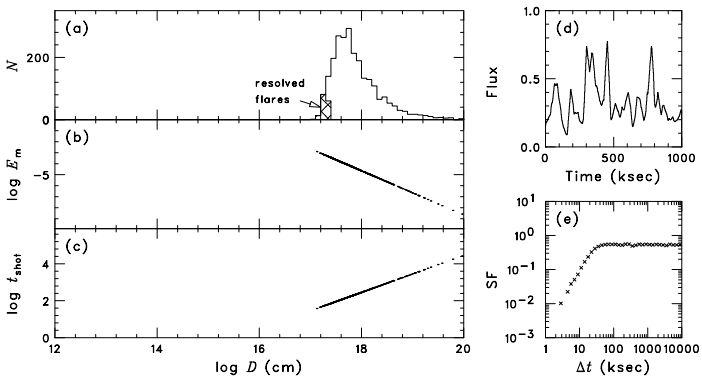

FIG. 1.— The results of a simulation of the internal shock model for the case of $\Gamma_{\rm avrg}$=10, σ'_Γ=0.005, and D_0=3×10^{13} cm. (a) The histogram of the collision distances D. (b) The dissipated energy and the (c) the generated time scale of the flare at the collision, showed as a function of D. (d) The resulting simulated light curve, and (e) the structure function calculated from the light curve. It is apparent that the simulated light curve shows repeating resolved flares, and that the structure function clearly shows a characteristic time scale.

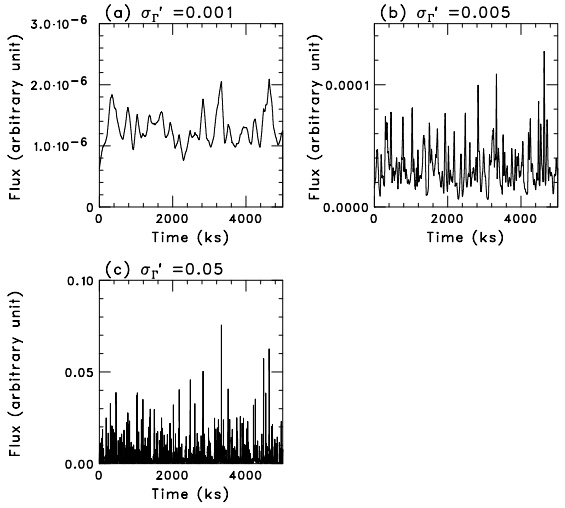

FIG. 2.— The simulated light curves for the case of σ'_Γ= (a) 0.001, (b) 0.005, and (c) 0.05. $\Gamma_{\rm avrg}$ is fixed to 10, and D_0 is fixed to 3×10^{13} cm. It is shown that the relative amplitude of the flare to offset component increases as σ'_Γ becomes larger.

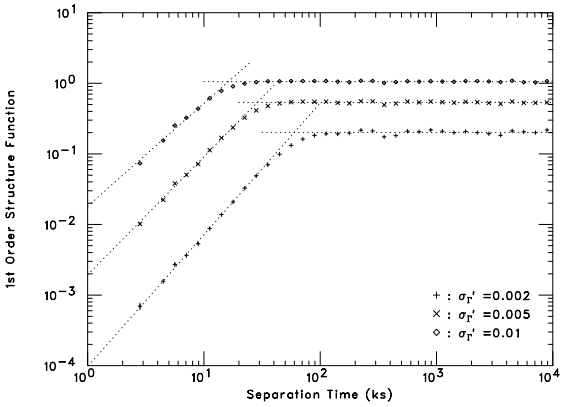

FIG. 3.— The structure function calculated from the simulated light curves for the case of σ'_Γ=0.002, 0.005, and 0.01. $\Gamma_{\rm avrg}$ is fixed to 10, and D_0 is fixed to 3×10^{13} cm. The location of the break, indicating the characteristic time scale, shifts to the shorter time scale as σ'_Γ increases.

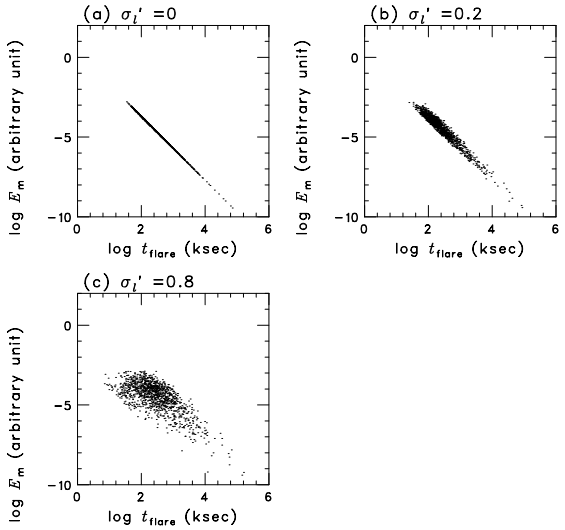

FIG. 4.— The correlation of the flare time scale and the energy dissipated in each collision for the case of σ'_l= (a) 0.0, (b) 0.2, (c) 0.8. The other parameters are fixed to $\Gamma_{\rm avrg}$=10, σ'_Γ=0.005, and D_0=3×10^{13} cm.

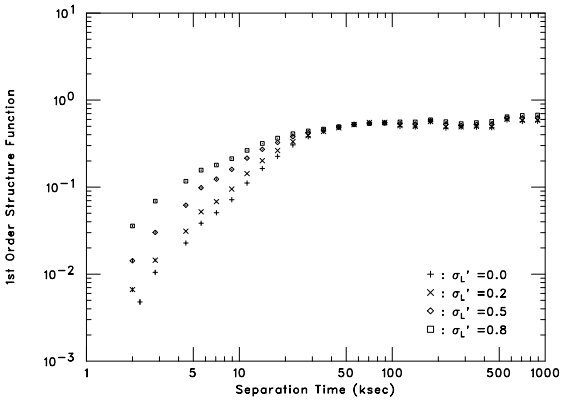

FIG. 5.— The structure function calculated from the simulated light curves for the case of σ_l'=0.0, 0.2, 0.5, 0.8. The other parameters are fixed to $\Gamma_{\rm avrg}$=10, σ_Γ'=0.005, and D_0=3 × 10^{13} cm. The slope of the derived structure function flattens as σ_l' increases.

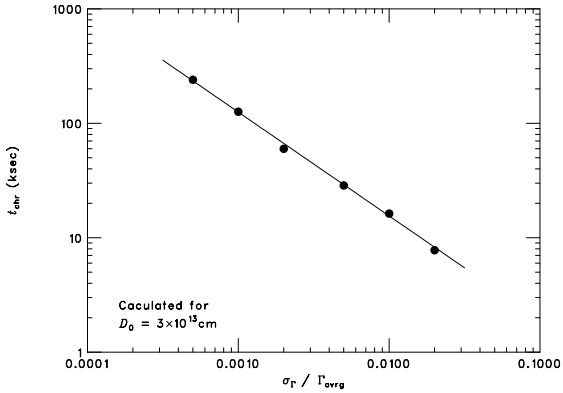

FIG. 6.— The derived characteristic times scales of the variability plotted as a function of the initial width of the BLF distribution, σ_Γ'. The case of D_0=3 × 10^{13} is shown.

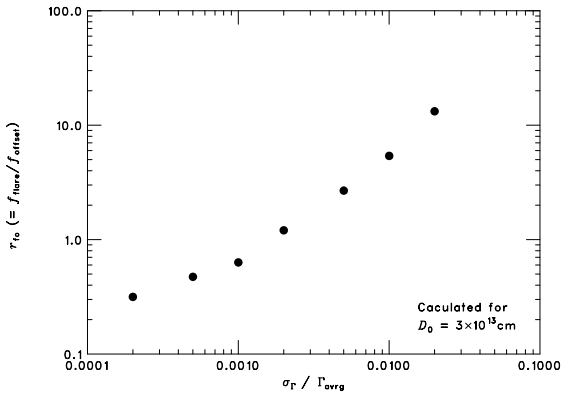

FIG. 7.— The ratio of the flare to offset component amplitude, plotted as a function of the initial width of the BLF distribution, σ_Γ'.

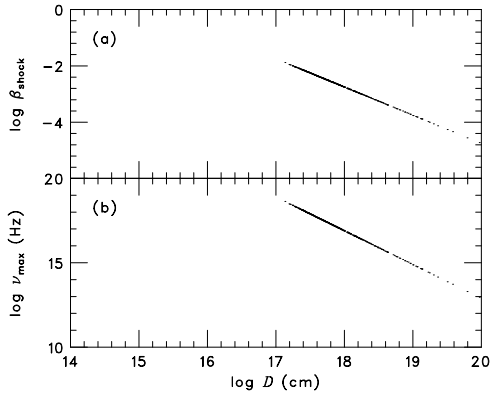

FIG. 8.— The relative velocity of the reverse shock (a) and the resulting synchrotron peak frequency (b) calculated for the case of $\Gamma_{\rm avrg}$=10, σ'_Γ=0.005, and D_0=3×10^{13} cm. The Bohm limit is assumed.

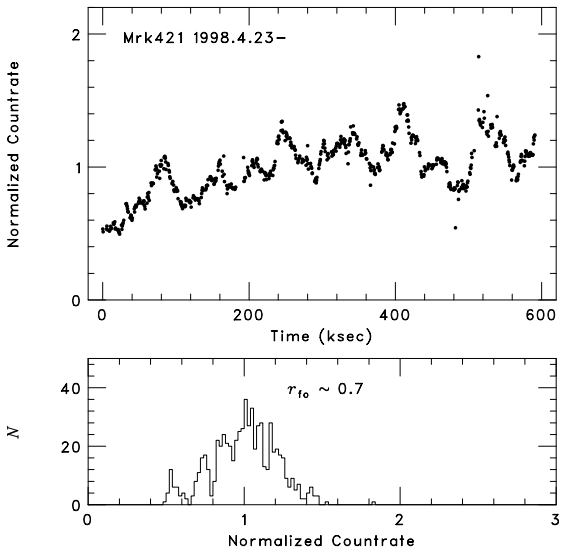

FIG. 9.— The observed normalized 0.6–7.5 keV count rate light curve and the histogram of the count rate for Mrk 421 during the *ASCA* long look in 1998 April. The bottom panel shows the number histogram of the count rate, where the lower and higher end of the peak can be considered as an indicator of the amplitude of the offset component, and the offset-plus-flare component. The derived flare-to-offset ratio defined in §4.1 is calculated to be $r_{\rm fo}$ ∼0.7 for this case.

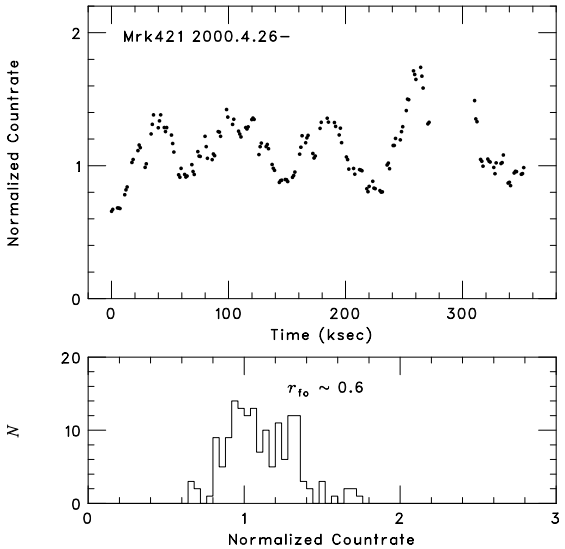

FIG. 10.— The observed normalized count rate light curve and the histogram of the count rate for the *BeppoSAX* observation of Mrk 421 during 2000 April. The light curve was generated by integrating over the MECS band. The bottom panel shows the number histogram of the count rate, where $r_{\rm fo}$ ∼0.6 is derived.

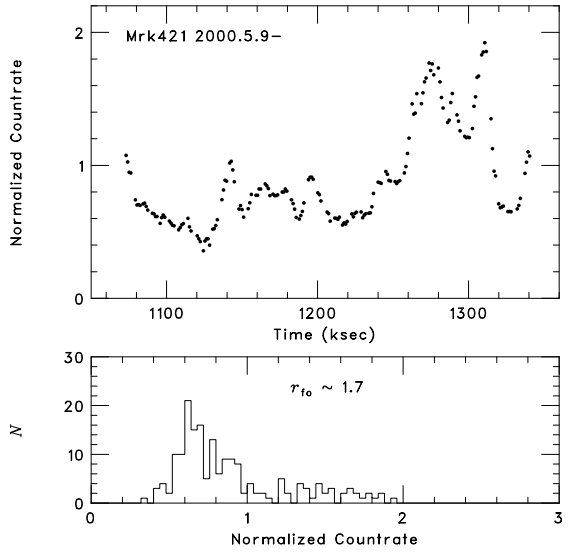

FIG. 11.— The observed normalized count rate light curve and the histogram of the count rate for the *BeppoSAX* observation of Mrk 421 during 2000 May. The light curve was generated by integrating over the MECS band. The bottom panel shows the number histogram of the count rate, where $r_{\rm fo}$ ∼1.7 is derived.

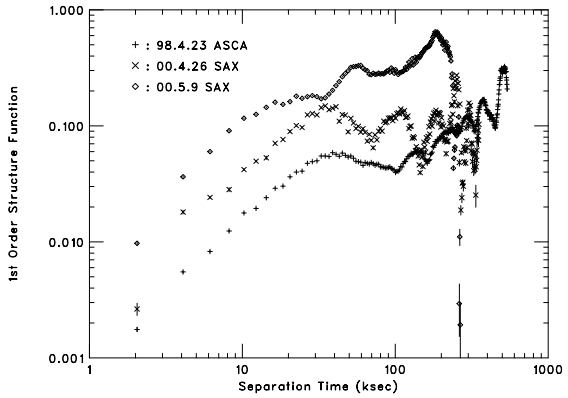

FIG. 12.— The structure function calculated for the observed light curve of Mrk 421 for the 3 observations shown in Figures 9-11.

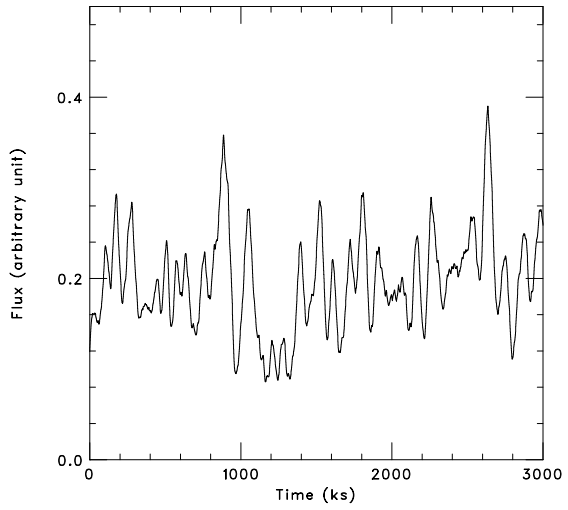

FIG. 13.— The simulated light curve considering the external shock scenario, assuming parameters Γ=15 and D=5.4 × 10^{17}.