Bayesian Approach to Find a Long-Term Trend in Erratic Polarization Variations Observed in Blazars

Makoto UEMURA,1 Koji S. KAWABATA,1 Mahito SASADA,2 Yuki IKEJIRI,2 Kiyoshi SAKIMOTO,2
Ryosuke ITOH,2 Masayuki YAMANAKA,2 Takashi OHSGU,1 Shuji SATO,3 and Masaru KINO3

1Hiroshima Astrophysical Science Center, Hiroshima University,
1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526
uemuran@hiroshima-u.ac.jp
2Department of Physical Science, Hiroshima University,
1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526
3Department of Physics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602

(Received 2009 October 9; accepted 2009 November 16)

Abstract

We developed a method of separating a long-term trend from observed temporal variations of polarization in blazars using a Bayesian approach. The temporal variation of the polarization vector is apparently erratic in most blazars, while several objects have occasionally exhibited systematic variations, for example an increase of the polarization degree associated with a flare of the total flux. We assume that the observed polarization vector is a superposition of two distinct components: a long-term trend and a short-term variation component responsible for short flares. Our Bayesian model estimates a long-term trend that satisfies the condition that the total flux correlates with the polarized flux of the short-term component. We demonstrate that assumed long-term polarization components are successfully separated by the Bayesian model for artificial data. We applied this method to the three photopolarimetric data of OJ 287, S5 0716+714, and S2 0109+224. Simple and systematic long-term trends were obtained in OJ 287 and S2 0109+224, while no such trend was confirmed in S5 0716+714. We propose that the apparently erratic variations of polarization in OJ 287 and S2 0109+224 are due to the presence of the long-term polarization component. The behavior of polarization in S5 0716+714 during the period of observation implies the presence of a number of polarization components showing variations on a quite short time-scale.

Key words: galaxies: active — galaxies: nuclei

1. Introduction

Blazars are believed to be active galactic nuclei (AGN) with relativistic jets pointing toward us (e.g., Blandford & Rees 1978). Doppler-boosted nonthermal emission from the jet dominates from radio to γ-rays in blazars. The radio–optical emission is dominated by synchrotron emission, and hence highly polarized. Polarimetric observations have been extensively performed, since they provide a clue to the structure of the magnetic field in the jet (e.g., Angel & Stockman 1980; Mead et al. 1990). Another important feature of blazars is a violent variability in a wide range of time-scales from a few minutes to years (e.g., Hufnagel & Bregman 1992). Since the polarization degree (PD) and angle (PA) also change with time, dense and long-term polarimetric observations are essential to understanding the variation mechanism of blazars and the magnetic field structure in the jet.

Moore et al. (1982) conducted an intensive observation campaign for BL Lac with multilongitude optical observatories over a one-week period, and found that the polarization changes trace out a random walk in the Stokes Q–U plane. In general, erratic variations have been observed in polarization of most blazars (e.g., Angel & Stockman 1980). On the other hand, systematic variations in PD and PA were also occasionally detected in blazars, for example, flares associated with the increase of PD (Smith et al. 1986; Tosti et al. 1998; Efimov & Shakovskoy 1998) and the color variation correlated with the variation in PD (Cellone et al. 2007). Marscher et al. (2008) have recently reported a smooth rotation of the polarization vector in BL Lac, which indicates a propagation of the emitting region through a helical magnetic field in the jet. Thus, erratic and systematic variations of polarization have been both observed in blazars.

The erratic variation of the polarization vector can be predicted by a scenario that a number of independent sources with randomly oriented and strong polarizations are blinking (Moore et al. 1982; Jones et al. 1985; Impey et al. 1989). It is, however, difficult to explain the systematic variations by attributing it only to the random variation. Alternatively, a number of systematic variations can be apparently overlooked in PD and PA, if short-term polarization variations are superimposed on a long-term polarization trend. A schematic example is shown in figure 1. The upper and lower panels show the light curves of the total and polarized fluxes and the temporal variation of the polarization vector in the Stokes Q–U plane, respectively. In this example, the polarization vector is assumed to be a superposition of two components; one shows a short-term variation associated with a flare of the total flux, and the other is a long-term trend. In the case where there is no long-term trend, the polarization flux (PF) correlates with the total flux, independent of the PA of the short-term flare. In the case where the long-term trend has a significant PF,
component could be considered to be a long-term trend if it

as shown in the lower panel of figure 1, however, the PF can have no positive correlation with the light curve. If the PA of the short-term flare is random and observations are sparse, the variation of polarization can be apparently erratic.

The presence of the long-term component has been revealed by several observations. The short-term variability of blazars is characterized by a combination of shots (Hufnagel & Bregman 1992). A typical time-scale of the shot was estimated to be ~1 d (Takahashi et al. 2000; Kataoka et al. 2001). In addition to these short-term variations, it is well known that blazars show variations on a time-scale ranging from months to years. It is possible that the long- and short-term variation components have different polarization features. Marscher et al. (2002) reported on VLBI radio observations of blazars, which show two or more components in polarization maps. If the polarization components in the radio range can also be observed in the optical range, an observed optical polarization vector should be a composition of those distinct radio polarization components, since the angular resolution of optical observations is much lower than that of VLBI radio images. Hagen-Thorn et al. (2002) reported on a long polarimetric monitoring of BL Lac in 1969–1991, and found the existence of a preferred direction of the polarization vector. They propose a model that the polarization of BL Lac has two kinds of sources: a stationary component and a large number of randomly polarized components (also, see Blinov & Hagen-Thorn 2009). This stationary component could be considered to be a long-term trend if it varies gradually.

The two-component model shown in figure 1 is one of the simplest among models with multiple polarization components. It deserves further investigation if such a simple model can extract a systematic trend from erratic variations of polarization in blazars. Here, we present an exploratory Bayesian approach to separating a long-term trend from observed temporal variations of polarization in blazars. A description of the model and its method is shown in the next section. We apply the model to the observed photopolarimetric data of blazars and discuss the implications of the results in section 3. In section 4, we discuss the validity of the model and results. Finally, we summarize our findings in section 5.

2. Bayesian Model for the Separation of Long- and Short-Term Variation Components in Polarization

2.1. Model Description

We assume the following three conditions for the observed flux and polarization in blazars: (i) An observed polarization vector is a superposition of two distinct polarization vectors of long- and short-term variation components. (ii) The short-term component is responsible for observed variations of the total flux; in other words, the PF of the short-term component completely correlates with the light curve of the total flux. (iii) Temporal variations of the flux from the long-term component is small. Condition (ii) is based on the fact that several flares of blazars are associated with the increase of PF, as mentioned in section 1. In our picture, all flares are associated with PF flares although each variation is often hidden in the observed PF because of the presence of the stronger, long-term component. Condition (iii) is established just to simplify the calculation. It is difficult to separate the short- and long-term components simultaneously in both the light curve and the polarization vector in our model.

Since Stokes parameters are additive, condition (i) can be expressed with Stokes parameters Q and U as follows:

\[
\begin{align*}
Q_{\text{obs}} &= Q_L + Q_S, \\
U_{\text{obs}} &= U_L + U_S,
\end{align*}
\]

where \((Q_{\text{obs}}, U_{\text{obs}})\), \((Q_L, U_L)\), and \((Q_S, U_S)\) denote the linear polarization parameters of the observed, long-, and short-term components, respectively. The short-term component \((Q_S, U_S)\) is determined when \((Q_L, U_L)\) is given for \((Q_{\text{obs}}, U_{\text{obs}})\). PF, PD, and PA are calculated as \(\sqrt{(Q^2 + U^2)}\), PF/I, and \((1/2)\arctan(U/Q)\), respectively.

Bayesian statistics provide a method of estimating a posterior probability density function (PDF) of model parameters, from a likelihood function defined by the model and the data and from a prior PDF of the model parameters. We have developed a Bayesian model to estimate a time-series of Q and U of the long-term trend, that is, y = \{QLi, ULi\} (i = 0–N). We use N sets of photopolarimetric observations, x = \{Q_{\text{obs},i}, U_{\text{obs},i}, I_{\text{obs},i}\}, where \(I_{\text{obs}}\) denotes the observed total flux. According to the Bayesian theorem, the posterior distribution of y is calculated as

\[
P(y|x) = \frac{L(y|x)\pi(y)}{C},
\]

where \(L(y|x)\) is the likelihood of the model, and \(\pi(y)\) is the prior distribution of the parameters.
where \( C \) is a constant for normalization of PDF. Here, \( L(y|x) \) and \( \pi(y) \) are the likelihood function and the prior distribution of \( y \), respectively.

The likelihood function, \( L(y|x) \), is defined based on our assumed conditions (ii) and (iii). The PF of the short-term component can be calculated as \( \text{PF}_S = \sqrt{Q^2_S + U^2_S} \). We normalized \( I_{\text{obs}} \) and \( \text{PF}_S \) using the “Z-score” transformation, defined as \( a' = (a - \bar{a})/\sigma_a \), where \( a' \), \( \bar{a} \), and \( \sigma_a \) are the normalized- and original-parameter values, respectively, and its standard deviation, respectively. This normalization procedure reduces the uncertainty of a possible contribution of the long-term component to the total flux if the contribution is time-independent, as assumed on condition (iii). The likelihood function is, then, defined with these normalized parameters, \( I'_{\text{obs}} \) and \( \text{PF}'_S \), as follows:

\[
L(y|x) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2_{PF,j}}} \exp \left[-\frac{(I'_{\text{obs},j} - \text{PF}'_S)\sqrt{2}}{2\sigma^2_{PF,j}} \right].
\]

where \( \sigma_{PF'} \) is the error for \( \text{PF}'_S \) and approximately given by the error of the observed PF. We neglected the error of \( I_{\text{obs}} \) because photometric errors of the total fluxes are much smaller than the errors of polarization parameters. \( L(y|x) \) reaches the maximum when the profile of the light curve is the same as that of \( \text{PF}_S \).

The prior distribution, \( \pi(y) \), plays a role in controlling the time-scale of the long-term trend in our model. Our model requires that the time-scale of variations in \( (Q_L, U_L) \) is longer than that in \( (Q_S, U_S) \). The long-term variation definitely draws a smoother path than the short-term variation does. A smooth curve can be described by the condition that the sequence of the first difference of \( \{ Q_L \} \) and \( \{ U_L \} \) should follow a standard normal distribution. We define the prior distribution as follows:

\[
\pi(y) = \pi(\{ Q_L \})\pi(\{ U_L \}),
\]

\[
\pi(\{ Q_L \}) = \prod_i \frac{1}{\sqrt{2\pi w^2}} \exp \left[-\frac{(Q_{L,i} - Q_{L,i-1})^2}{2w^2} \right],
\]

\[
\pi(\{ U_L \}) = \prod_i \frac{1}{\sqrt{2\pi w^2}} \exp \left[-\frac{(U_{L,i} - U_{L,i-1})^2}{2w^2} \right].
\]

Here, \( \pi(y) \) has no physical meaning in our model. It only controls the smoothness of the long-term trend by using a hyperparameter, \( w \). The result obtained from this model evidently depends on \( w \). We determine an appropriate \( w \) using “empirical Bayesian” approach, by setting up the criterion for obtained results, which are discussed in subsections 2.3 and 4.1. The hyperparameter, \( w \), has the same dimension as \( Q \) and \( U \). Since \( w \) corresponds to a standard deviation of the normal distribution as can be seen in equations (5) and (6), it is useful to use \( w \) normalized by a typical observational error of \( (Q, U) \), \( \alpha \):

\[
\alpha = \frac{\sqrt{\sum_{i=1}^{N}(Q^2_{\text{err},i} + U^2_{\text{err},i})}}{N}.
\]

Thus, we can calculate \( L(y|x) \) and \( \pi(y) \) for \( y \) using the observations, \( x \). The posterior probability, \( P(y|x) \), was then estimated based on equation (2). The estimation was performed with the Markov chain Monte Carlo (MCMC) method with the Metropolis algorithm (Metropolis 1953). The calculation procedure of MCMC is as follows: At the \( n \)-th step of MCMC, we obtain \( p_n = L_n \sigma_n \) for \( y_n \). We then randomly move the step to another point in the parameter space of \( y \) by adding random values drawn from a standard normal distribution which has a dispersion chosen to efficiently sample the likelihood surface. As a result, we obtain \( p_{n+1} \). We count the \((n + 1)\)-th step if either \( p_{n+1} > p_n \) or \( p_{n+1}/p_n \) is larger than a uniform random number between 0 and 1; otherwise we reject it and retake the \((n + 1)\)-th step. After discarding the first \( 10^4 \) steps, we sample every 100 steps until the number of the sample becomes \( 10^5 \). This procedure forms a set of \( y \), and we obtained ten sets of \( y \) with different initial values of \( y \). We confirmed that no significant difference was seen in each set of \( y \). We merged them, and finally obtained the MCMC sample of \( y \). The median and 68.3% confidence intervals for \( y \) were determined from this combined sample.

2.2. Demonstration with Artificial Data

In this section, we demonstrate how our Bayesian model works using artificial data sets. We generated five sets of artificial data of temporal \((Q, U)\) variations and light curves. Each data set consists of 100 sets of \( Q, U \), and \( I \) in a time-series from \( t = 0 \) to 99. The duration and amplitude of the short-term flare were fixed in the light curve and PF as \( \Delta t = 10 \), \( \Delta I = 10 \), and \( \Delta \text{PF}_S = 10 \). Errors of \( Q \) and \( U \) were set to be 1.0. We used simple linear functions for brightening and fading phases of the flare. Figure 1 shows an example of the flare. Results of our analysis are independent of the flux out of the flare since the light curve is normalized in our procedure, as mentioned above. The short-term flares were randomly generated in the time and PA. The frequency of the flare was controlled by a parameter, \( \alpha \), which is a probability that a flare occurs at each time. Overlaps of flares were allowed. Figures 2 and 3 show respectively the variations in the \( Q-U \) plane and the light curves of the five sets of artificial data, labeled as (a), (b), (c), (d), and (e). We assumed sine-curves for the long-term trends, as depicted in the left panels of figure 2. The middle panels show \((Q_{\text{obs}}, U_{\text{obs}}) = (Q_L + Q_S, U_L + U_S)\).

We applied our Bayesian model to those artificial data. The hyperparameter, \( w \), was fixed to \( w/\sigma = 0.20 \) in all cases. The estimated long-term trends were shown in the right panels of figure 2, indicated by the filled circles. A typical error of \((Q_L, U_L)\) is also shown in each right panel. The temporal variations of \( \text{PF}_S \) are shown in the right panels in figure 3.

Data (a) was generated by a low flare frequency of \( \alpha = 0.02 \). The second and fourth flares apparently have larger amplitudes than the others in the light curve in figure 3, because two consecutive flares are incidentally overlapped. A sign of the assumed long-term trend can be seen in the middle panel of figure 2 because of the low flare frequency. The observed PF apparently shows negative correlations with the light curve in the flares of \( t < 30 \), while positive correlations are seen in \( t > 30 \). This is because the long-term trend has a significant polarization, and the polarization vectors of the early flares are

\[ \text{Observed PF} \]

\[ \text{Calculation PF} \]

\[ \text{Correlation PF} \]
Fig. 2. Artificial polarization data and estimated long-term trends in the $Q$–$U$ plane. Five sets of the data were used, as labeled (a), (b), (c), (d), and (e) (for the details of the data, see the text). The left, middle, and right panels show the assumed long-term trends, the short-term variations superimposed on the long-term trends, and the estimated long-term trends, respectively. The paths of the $(Q, U)$ variation of the data are also shown in the right panels with the dashed lines. In data (e), no long-term trend was assumed. The Bayesian estimations of the long-term trend were performed with $w/\sigma = 0.20$ in all cases.
Fig. 3. Light curves of the total and polarized fluxes of the five sets of the artificial data labeled as (a), (b), (c), (d), and (e) are indicated by the open and filled circles, respectively. In the right panels, the estimated long-term trends were subtracted from the polarized flux.

incidentally directed toward the origin of the $Q-U$ plane. As seen in figure 2, the Bayesian model well reproduced the assumed long-term trend. The corrected PF$_S$ has a definite correlation with the light curve throughout this entire period, as can be seen in figure 3.

Data (b) was generated from data (a) by random sampling. We “observed” data (a) with an observational probability of 50%. This data is used in order to evaluate the Bayesian model for data taken at irregular intervals, like typical ground-based observations of blazars. This experiment is important because the prior distribution in our model assumes data observed at regular intervals. The estimated PF$_S$ still has a definite correlation with the light curve. The estimated long-term trend is somewhat smoother than the assumed one, while it successfully reproduces general features of the assumed trend.

Data (c) is an example for data with a low signal-to-noise ratio (S/N). We generated $Q$ and $U$ of data (c) adding a Gaussian noise to $Q$ and $U$ of data (a) with a standard deviation of 2.3. This corresponds to $S/N = 3.0$ for each polarization flare having an amplitude of $\Delta$PF$_S = 10$. The estimated long-term component still shows a general trend of increasing $U$, while the “S”-shaped profile is weak. The correlation between the light curve and PF is more definitely improved even in the case where this long-term trend is subtracted, as can be seen in figure 3. This is probably because the substructure of the assumed long-term trend is negligible compared with the flares having a large amplitude in this case. We confirmed that the estimation of the long-term trend completely failed if the standard deviation of the Gaussian noise was 3.5 ($S/N \sim 2$). In the case of a standard deviation of 1.8 ($S/N \sim 4$), the long-term trend was well reproduced. Hence, our model is validated if polarization flares are detected with $S/N \geq 3$ in the case of data (a). In general, the allowable $S/N$ highly depends on the data itself. It is possible that a higher $S/N$ is required for the data with a higher flare frequency.

Data (d) shows more complicated temporal variations. It was generated by a high flare frequency of $\alpha = 0.30$, and
a more prominent long-term trend. Almost no characteristic
of the long-term trend is seen in the middle panel of data (c) in
figure 2, and the path in the \( Q-U \) plane is apparently erratic,
as generally observed in blazars. Nevertheless, the long-term
trend was successfully reproduced by the Bayesian model, as
can be seen in the right panel of figure 2. While there is no
significant correlation between the observed PF and the light
curve, we calculate that the correlation coefficient between PFs
and the light curve is \( 0.84^{+0.05}_{-0.07} \). This demonstrates that
the Bayesian method can be a powerful tool for analysis of appar-
ently erratic variations in polarization. We note that differences
between PFs and the light curve are larger than those in the
cases of data (a) and data (b). These indicate that the Bayesian
model possibly fails to reproduce the long-term trend in the
case of a very high flare frequency.

In data (e), there is no association between the light curve
and the polarization parameters. The light curve was gener-
ated in the same manner as in the other cases with \( \alpha = 0.30 \).
The variation of the polarization vector was set to be a purely
random walk in the \( Q-U \) plane without a long-term trend.
The Bayesian model found a solution of the long-term trend
even in this case, as shown in the right panel of data (e) in
figure 2. On the other hand, there is no significant correlation
between the estimated PFs and the light curve. This suggests
that the model cannot extract any meaningful long-term trend
from data (e).

2.3. Evaluation of the Result Obtained from the Bayesian
Model

In the preceding subsection, we demonstrated that the
Bayesian model successfully reproduced the assumed long-
term trends in the \( Q-U \) plane for the artificial data. It can,
however, generate a false long-term trend even from the data in
which no long-term trend was actually present. It is problematic
whether we can apply this model to real observations in
which we have no information about the presence of the long-
term trend. Hence, we need to adopt the criterion for judging
whether the results obtained from the Bayesian model are valid
or not. We also need to determine a more desirable range of
the hyperparameter, \( w \), because the result depends on \( w \).
In general, a hyperparameter in an empirical Bayesian model
can be estimated by maximizing the marginal likelihood. This
standard method is, however, insufficient for our model, as
discussed in subsection 4.1. Here, we adopt a specific crite-
ron for our model.

First, the correlation of PFs with the light curve should be
significantly improved compared with the correlation between
the observed PF and the light curve. The significance of the
difference in correlation coefficient is evaluated by a \( Z \)-test.
By using two correlation coefficients, \( r_i (i = 1 \text{ and } 2) \),
the test statistic, \( Z \), which has a standard normal distribution,
is defined as follows:

\[
Z_i = \frac{Z_1 - Z_2}{\sigma},
\]

\[ Z_i = \frac{1}{2} \log \left( \frac{1 + r_i}{1 - r_i} \right), \]

\[
\sigma = \sqrt{\frac{1}{N_i - 3}},
\]

where \( N_i \) is the number of samples. We can conclude that
the two correlation coefficients are significantly different with
a > 95% confidence level when \(|Z| > 1.96 \). In our Bayesian
model, a smaller \( w \) yields a smoother long-term trend, and
hence it results in a less improvement in the correlation coeffi-
cients between PFs and the light curve. As a result, a too-
small \( w \) is discarded by this procedure. We calculated that
\(|Z| \) in the cases of artificial data (a), (b), (c), (d), and (e)
in the preceding subsection are 23.74, 17.41, 9.71, 6.86, and
1.35, respectively. Thus, we conclude that the long-term trend
estimated for data (e) is a false result.

Second, we need to check whether an estimated long-term trend
actually has a longer time-scale variation than the short-
term trend has. As mentioned in subsection 2.1, a long-term
component is expected to draw a smooth path in the \( Q-U \)
plane, and we adopted the prior distribution with \( w \). In other
words, the hyperparameter, \( w \), is related to the variation time-
scale of the long-term trend in our model. A path of the long-
term trend becomes more complicated and erratic in the case of
a larger \( w \) (for example, see figure 5 in subsection 3.2). With
an extremely large \( w \), the variation amplitudes and time-scales
of \( Q_t \) and \( U_t \) can be comparable to those of \( Q_S \) and \( U_S \). Thus,
\( w \) should be restricted to be small enough to assure us that the
estimated \( Q_t \) and \( U_t \) actually exhibit “long”-term variations.
We evaluate it using a travel distance in the \( Q-U \) plane of
both long- and short-term components. The travel distance, \( d \),
defined for a set of \( \{Q_i, U_i\} \) is as follows:

\[
d = \sum_{i=1}^{N-1} \sqrt{(Q_{i+1} - Q_i)^2 + (U_{i+1} - U_i)^2}.
\]

In the present work, the estimated \( Q_t \) and \( U_t \) are acceptable
in the case where the ratio between \( d \) for long- and short-
term components, \( d_t \) and \( d_s \), satisfies the following condition:

\[
R = \frac{d_t}{d_s} < 0.10.
\]

We call \( R \) the “distance ratio”. In conjunction with the first
criterion of the correlation coefficient, we can restrict \( w \) to
a narrow range.

Results obtained from the Bayesian model are quantitatively
evaluated on the above criterion. It is, however, not unlikely
that a false long-term trend would be extracted from observ-
eations even if the result satisfies the above criterion. Hence,
results from the Bayesian model should be evaluated carefully.
In addition to the above criterion, we can check the validity of
results with their qualitative features. We expect that a long-
term trend exhibits not a random walk but a systematic motion
in the \( Q-U \) plane. A long-term trend depicting a very com-
licated path in the \( Q-U \) plane may merely indicate that there
is no systematic long-term trend in the observed \( (Q,U) \). We
should also suspect a result that is highly sensitive to a small
change in \( w \).
Fig. 4. Temporal variations of the observed and estimated parameters. The left, middle, and right panels show the results of OJ 287, S5 0716+714, and S2 0109+224, respectively, as indicated at the top of the panels. From top to bottom, the panels show (a) the observed light curve, (b) the observed polarization degree, (c) the observed polarized flux, (d) the observed polarization angle, (e) the estimated polarization flux of the short-term component, (f) the estimated polarization angle of the short-term component, (g) the estimated polarization flux of the long-term component, and (h) the estimated polarization angle of the long-term component. The observational errors and 1σ confidence levels of the estimated parameters are also indicated, while most of them are comparable with the symbol size.

3. Results: Application for Observed Photopolarimetric Data of Blazars

3.1. Data Description

We applied the Bayesian model to observed polarimetric data of blazars. The data were obtained with TRISPEC attached to the “Kanata” 1.5 m telescope at Higashi-Hiroshima Observatory during the campaign for our photopolarimetric observation of blazars in 2008–2009 (Watanabe et al. 2005; Uemura et al. 2008). A full description of the observation and the data reduction will be published in a forthcoming paper. In this paper, we focus on how our model works for the real data.

We used the V-band photopolarimetric data of OJ 287, S5 0716+714, and S2 0109+224. The individual observations of these three objects were performed on 79 nights during the period from 2008 October to 2009 May, 118 nights from 2008 August to 2009 May, and 56 nights from 2008 July to 2009 February, respectively. The light curves, PD, PF, and PA are shown in panels (a), (b), (c), and (d) in figure 4, respectively. Variations in the $Q–U$ plane are shown in figure 5.

OJ 287 exhibited short-term flares superimposed on a gradual fading trend in the light curve on a time-scale of days. In contrast, the PD remained high at 10%–20% during the period of observation, occasionally showing flares reaching a maximum of ~30%. The object almost stayed in the fourth quadrant in the $Q–U$ plane. This behavior implies the presence of two polarization components, that is, one for a long-term trend pointing to the fourth quadrant, and the other for short-term flares.

Short-term variations in S5 0716+714 had similar features to those in OJ 287, while the behavior of PD is totally different. The PD of S5 0716+714 rapidly fluctuated between 0% and 15%. No systematic PD variation correlating with the total flux can be seen, except for a possible negative correlation during a flare around Modified Julian Day (MJD) 54750. The temporal variation in the $Q–U$ plane is very erratic.
S2 0109+224 experienced a flare strongly associated with an increase of PD around MJD 54810–MJD 54820. Except for this flare, there can be no clear correlation between the light curve and PD. The PA apparently shows a systematic long-term trend: a gradually decreasing trend until MJD 54800, followed by an increasing trend, while short-term variations occasionally disturb the trend. This behavior in PA indicates the presence of long- and short-term polarization components, as in PA of OJ 287.

3.2. Bayesian Analysis of the Polarization Data

The Bayesian estimations of the long-term trend for these three objects were performed with $w/\sigma = 0.10$, 0.20, 0.30, 0.40, and 0.50. The different cases of $w/\sigma$ were used in order to check the dependency of the result on $w$. Figure 5 shows the estimated long-term trend in the $Q$–$U$ plane indicated by the filled circles. The top, middle, and bottom panels show the results with $w/\sigma = 0.10$, 0.20, and 0.50, respectively.

The long-term trend draws a more complicated path in the $Q$–$U$ plane with a larger $w/\sigma$, as expected. The test statistics, $|Z|$, and distance ratio, $R$, defined in subsection 2.3 for those results are summarized in table 1. Only the cases of $w/\sigma = 0.20$ for OJ 287 and S2 0109+224 are acceptable in our criterion: $|Z| > 1.96$ and $R < 0.10$. Temporal variations of PF and PA of the short-term components are shown in panels (e) and (f) of figure 4, respectively. Those of the long-term trends are in panels (g) and (h). Those panels show the results with $w/\sigma = 0.20$.

The long-term trend in S5 0716+714 draws a rather erratic path in the $Q$–$U$ plane, even with $w/\sigma = 0.10$. The two-component model is probably inadequate to explain the observed $Q$–$U$ variation in S5 0716+714. The $Q$–$U$ variation in S5 0716+714 may be caused by the number of polarization components having a variation time-scale of $< 1$ d (Moore et al. 1982; Jones et al. 1985). In this case, the time resolution of our observation could be too low to detect systematic
Bayesian Separation of Polarization in Blazars

Table 1. Test statistics, |Z|, and distance ratio, R, in the cases of OJ 287, S5 0716+714, and S2 0109+224.

| w/σ | OJ 287  | S5 0716+714 | S2 0109+224 |
|------|---------|------------|------------|
| 0.10 | 1.08, 0.04 | 0.33, 0.03 | 1.66, 0.04 |
| 0.20 | 2.85, 0.09 | 0.40, 0.06 | 2.27, 0.07 |
| 0.30 | 3.97, 0.14 | 1.39, 0.09 | 2.26, 0.10 |
| 0.40 | 4.66, 0.18 | 2.25, 0.13 | 2.75, 0.11 |
| 0.50 | 4.78, 0.21 | 2.38, 0.15 | 2.83, 0.16 |

variations of polarization. Rapid variations having a time-scale of less than an hour actually have been reported in S5 0716+714 (Stalin et al. 2006; Sasada et al. 2008).

As can be seen from figures 4 and 5, the PA of the long-term trend in OJ 287 gradually increases until MJD 54920, and then decreases with time. In other words, the polarization vector apparently oscillates in a narrow range of PA of 150°−170° smoothly during the period of observation. This feature of the long-term trend can be seen both in the cases of w/σ = 0.10 and 0.20, which indicates that it is stable. The PF of the long-term trend shows variations at a small amplitude of a factor of ~2. The long-term trend estimated in OJ 287 would, thus, be an ideal example for our two-component model.

In the case of S2 0109+224, the PF of the long-term trend first decreased until ~MJD 54710, and then the object stagnated near the origin of the Q–U plane during the next forty days. After that, the PF increased in the opposite direction of the early trend in the Q–U plane, as shown in figure 5. This component turned toward the decrease again at last. This feature can be commonly seen in all cases of w/σ, while substructures become prominent in the case of w/σ = 0.50. This behavior of the long-term trend may be interpreted as two distinct components: one which kept decaying until ~MJD 54710, and the other which became prominent later. We note that the PF of the long-term trend shows a sevenfold variation. Since our two-component model assumes no flux variation for the long-term trend, the large variation amplitude of PF might refute this assumption. This problem can be settled only when the long-term trend displays a major variation in PD with a constant total flux.

The most notable feature in the estimated short-term component is nonuniform in the distributions of PA in S2 0109+224 and possibly also OJ 287. Figure 6 shows the histograms of the PA of the short-term components. The figure also includes the distribution in S5 0716+714 just as a reference. We tested nonuniformity in the distributions using the Kolmogorov–Smirnov (KS) test. As a result, we calculated that KS probabilities in the cases of OJ 287, S5 0716+714, and S2 0109+224 were 0.09, 0.05, and < 0.01, respectively. Therefore, we confirmed that the nonuniformity is statistically significant in S2 0109+224 with a > 95% confidence level (or KS probability < 0.05). This indicates that the PA of the short-term component is not completely random, and hence implies that the source of the short-term variation is preferentially localized on an area where the direction of the magnetic field is fixed.

Finally, we note that a small w was actually rejected in all cases of the three objects because no significant improvement was achieved in the correlation between the total and polarized fluxes, as shown in table 1. This means that extremely simple long-term trends are insufficient for our model. The most simple long-term component would be considered as an average of Q and U during the period of observation. If the correlation would be improved with the differential polarization flux from the average (Q, U), the Bayesian model was not required. However, we confirmed that the correlation is not significantly improved with those differential polarization fluxes from the averages; |Z| for OJ 287, S5 0716+714, and S2 0109+224 are 1.22, 0.40, and 1.27, respectively. Our Bayesian model, therefore, has an advantage in extracting long-term trends from polarization variations, compared with the simple correlation with the average (Q, U).

4. Discussion

4.1. Estimation of the Hyperparameter, w, by Maximizing Marginal Likelihood

As mentioned in subsection 2.3, we establish the criterion to determine an appropriate value of the hyperparameter, w,
in our Bayesian model. In general, a hyperparameter of an empirical Bayesian model can be estimated by maximizing the marginal likelihood. The marginal likelihood, $M$, is equivalent to the constant, $C$, in equation (2), defined as

$$M(w) = \frac{L(y|x)\pi(x|w)}{P(y|x)}. \quad (14)$$

We calculated $M(w)$ based on the method described in Chib and Jeliazkov (2001). The model parameter, $y$, can be arbitrary since $M(w)$ is independent of $y$. The numerator of equation (14) was calculated from the best estimated parameters of $y$. The denominator can be estimated by the Monte Carlo integration drawing a sample from $P(y|x)$, which was obtained through our Bayesian analysis. The results are summarized in table 2.

$M(w)$ takes the maximum with $w/\sigma = 2.00\text{--}3.00$ for all cases of the three objects. A long-term component estimated by such a large $w/\sigma$ definitely draws a quite complicated path in the $Q-U$ plane, as can be seen from figure 5. As a result, it is not acceptable in our criterion about the distance ratio, $R$.

The ratio of two $M(w)$ is called the “Bayes factor” (BF), which we used to evaluate the models. We calculated that the BF for OJ 287, S5 0716+714, and S2 0109+224 are $\sim1.6$, $\sim1.6$, and $\sim2.3$ with the maximum of $M$ and $M(0.10)$, respectively. These BF are too small to conclude that the model with $w/\sigma = 2.00\text{--}3.00$ is significantly better than that with $w/\sigma = 0.10$ (Kass & Raftery 1995).

Thus, the maximization of $M(w)$ is not suitable for the determination of $w$ in our model. This is mainly because the prior distribution is not the real distribution for $Q_L$ and $U_L$. Another form of the prior distribution may be more suitable for the case of polarization variations in blazars. The practical criterion defined in subsection 2.3 is sufficient for our exploratory model in this paper.

4.2. Application in the $(Q/I, U/I)$ Plane

Our Bayesian model can find a long-term trend in the $Q-U$ plane. It can also apply to the $(q,u) = (Q/I, U/I)$ plane, while rigid limitations are imposed on the application. In our two-component model, the observed $(q,u)$ is described as

$$q = \frac{Q_L + Q_S}{I_{\text{total}}}, \quad u = \frac{U_L + U_S}{I_{\text{total}}} \quad (15)$$

The two polarization components, namely, $(q_L, u_L) = (Q_L/I_L, U_L/I_L)$ and $(q_S, u_S) = (Q_S/I_S, U_S/I_S)$, can be separated in the $q-u$ plane when $q_L$ and $u_L$ are much smaller than $q_S$ and $u_S$ (e.g., Messinger et al. 1997). If the PD of the long-term trend is not negligible, the polarization components cannot be separated without estimating $I_L$ and $I_S$.

In the case where the PD of the long-term trend is large, the application of the Bayesian model to the $q-u$ plane means an estimate of $(Q_L/I_{\text{total}}, U_L/I_{\text{total}})$ that leads to a positive correlation between the light curve and the modified PD calculated from $(Q_S/I_{\text{total}}, U_S/I_{\text{total}})$. Since $I_L$ is assumed to be time-independent in our model, the temporal behavior of the estimated $(Q_S/I_{\text{total}}, U_S/I_{\text{total}})$ is probably analogous to that of $(Q_S/I_S, U_S/I_S)$. In this sense, the Bayesian analysis in the $q-u$ plane is not completely nonsense, and can still provide meaningful results. A result in the $q-u$ plane would be noteworthy, particularly if the long-term trend in the $q-u$ plane draws a path analogous to that in the $Q-U$ plane.

Figure 7 is the same as figure 5, but for the $q-u$ plane. We found that the long-term trend estimated with $w/\sigma = 0.20$ is quite analogous to that in figure 5 in the case of S2 0109+224. The long-term trend in OJ 287 also exhibits a feature common to that in the $Q-U$ plane in terms of the oscillation of the polarization vector within a narrow range of PA. The similarity between the long-term trends in the $Q-U$ and $q-u$ planes implies that the short-term flares were associated with the increases of PD in OJ 287 and S2 0109+224. The short-term flares may originate from a region where the direction of the magnetic field is more aligned than the whole emitting region.

4.3. Future Studies

We showed that the polarization vectors in OJ 287 and S2 0109+224 can be separated into two components: the long-term trend and the short-term variation component. The result, however, provides no clear evidence for the existence of the two components in these objects. Our Bayesian model just gives us one of possible views for temporal variations of polarization in blazars. This view deserves to be discussed, since simple and systematic variations can be extracted from apparently erratic variations. On the other hand, further investigations are needed to confirm the two-component view.

One of the most direct ways to obtain evidence for the two-component scenario would be VLBI radio observations. Marscher et al. (2002) reported on the presence of multiple polarization components even within radio cores in blazars. VSOP-2 is a space VLBI project, which will resolve subparsec regions around the central blackhole of AGN (Hirabayashi et al. 2004). VSOP-2 will reveal the detailed structure of the polarization components in the radio core. A part of the radio polarization components might exhibit temporal variations synchronized with the optical polarization components resolved by our Bayesian model. If this is the case, our Bayesian model, for the first time, provides a tool for investigating the temporal evolution of each polarization component observed both in the optical and radio regimes.

Our model assumes that the PF of the short-term component correlates with the light curve of the total flux [condition (ii) in subsection 2.1]. In other words, we assume that the total flux exhibits a flare whenever a short-term polarization flare occurs. This assumption gives the condition on which the application of our Bayesian model is validated. Our model definitely yields a false result if it applies to, for example, short-term...
polarization flares without significant variations of the total flux. If such variations are typical in a blazar, our model is not validated for it. Our assumption should, hence, be evaluated on a case-by-case basis for objects. In our picture, a short-term flare with a larger amplitude is expected to have a close correlation with polarization because the specific polarization feature of a smaller flare would be readily diluted by other small flares. Hence, the polarization behaviors of short and large flares would provide the key to validating our assumption.

Marscher et al. (2008) reported on a rotation of the optical polarization vector in BL Lac and interpreted it as a shift of the emitting region through a helical magnetic field in the jet. Such a rotation event can have been missed if the observed polarization has a long-term trend as considered in our model. Our Bayesian model would provide a tool to search for rotation events of the polarization vector both in the observed, long- and short-term components. No systematic rotation event can be seen in the long- and short-term components in the three objects, as shown in subsection 3.2.

5. Summary

We developed a Bayesian model to extract a long-term trend from apparently erratic variations in polarization of blazars. Our Bayesian model successfully resolved the long- and short-term components in the artificial data. We applied this model to the photopolarimetric data of OJ 287, S5 0716+716, and S2 0109+224. Simple and systematic long-term trends were obtained in OJ 287 and S2 0109+224, while no meaningful trend was identified in S5 0716+716. We propose that all short-term flares were associated with the increase of the polarized flux in OJ 287 and S2 0109+224. The polarization variation in S5 0716+716 may be explained by a random variation caused by a number of polarization components having a quite short time-scale of variations.

We acknowledge useful discussions with Taichi Kato and Takuya Yamashita. This work was partly supported by a Grand-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan (19740104).
References

Angel, J. R. P., & Stockman, H. S. 1980, ARA&A, 18, 321
Blandford, R. D., & Rees, M. J. 1978, in Pittsburgh Conf. on BL Lac
Objects, ed. A. M. Wolfe (Pittsburgh: University of Pittsburgh
Press), 328
Blinov, D. A., & Hagen-Thorn, V. A. 2009, A&A, 503, 103
Cellone, S. A., Romero, G. E., Combi, J. A., & Martí, J. 2007,
MNRAS, 381, L60
Chib, S., & Jeliazkov, I. 2001, J. Am. Statistical Assoc., 96, 270
Efimov, Yu. S., & Shakhovskoy, N. M. 1998, BLAZAR Data, 1, 3
Hagen-Thorn, V. A., Larionova, E. G., Jorstad, S. G., Björnsson, C.-I.,
& Larionov, V. M. 2002, A&A, 385, 55
Hirabayashi, H., et al. 2004, Proc. SPIE, 5487, 1646
Hufnagel, B. R., & Bregman, J. N. 1992, ApJ, 386, 473
Impey, C. D., Malkan, M. A., & Tapia, S. 1989, ApJ, 347, 96
Jones, T. W., Rudnick, L., Aller, H. D., Aller, M. F., Hodge, P. E., &
Fiedler, R. L. 1985, ApJ, 290, 627
Kass, R. E., & Raftery, A. E. 1995, J. Am. Statistical Assoc., 90, 773
Kataoka, J., et al. 2001, ApJ, 560, 659
Marscher, A. P., et al. 2008, Nature, 452, 966
Marscher, A. P., Jorstad, S. G., Mattox, J. R., & Wehrle, A. E. 2002,
ApJ, 577, 85
Mead, A. R. G., Ballard, K. R., Brand, P. W. J. L., Hough, J. H.,
Brindle, C., & Bailey, J. A. 1990, A&A, 383, 183
Messinger, D. W., Whittet, D. C. B., & Roberge, W. G. 1997, ApJ,
487, 314
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H.,
& Teller, E. 1953, J. Chem. Phys., 21, 1087
Moore, R. L., et al. 1982, ApJ, 260, 415
Sasada, M., et al. 2008, PASJ, 60, L37
Smith, P. S., Balonek, T. J., Heckert, P. A., & Elston, R. 1986, ApJ,
305, 484
Stalin, C. S., Gopal-Krishna, Sagar, R., Wiita, P. J., Mohan, V., &
Pandey, A. K. 2006, MNRAS, 366, 1337
Takahashi, T., et al. 2000, ApJ, 542, L105
Tosti, G., et al. 1998, A&A, 339, 41
Uemura, M., et al. 2008, in Blazar Variability across the
Electromagnetic Spectrum (Trieste: Scuola Internazionale
Superiore di Studi Avanzati), 70
Watanabe, M., et al. 2005, PASP, 117, 870