Separating States in High-Energy Astronomical Sources Using Hidden Markov Models

Robert Zimmerman[1,2] David A. van Dyk[2] Vinay L. Kashyap[3] Aneta Siemiginowska[3]

[1]University of Toronto [2]Imperial College London [3]Harvard & Smithsonian

Abstract

We present a new method to distinguish between different states (e.g., quiescent and flaring) in astronomical sources with count data. The method models the underlying physical process as latent variables following a continuous-space Markov chain. For the underlying state process, we consider several continuous-space hidden Markov models of varying complexity, under which we can infer the source's underlying physical state at any given time. We then dichotomize these predictions using a finite mixture model to produce binary classifications. Applying this method to X-ray data from the active dMe flare star EV Lac to distinguish quiescent from flaring states, we find that a first-order autoregressive process efficiently separates the two states: flaring occurs over 30–40% of the observations and is characterized by higher temperatures and emission measures, and we can also identify a well-defined persistent quiescent state.

Introduction

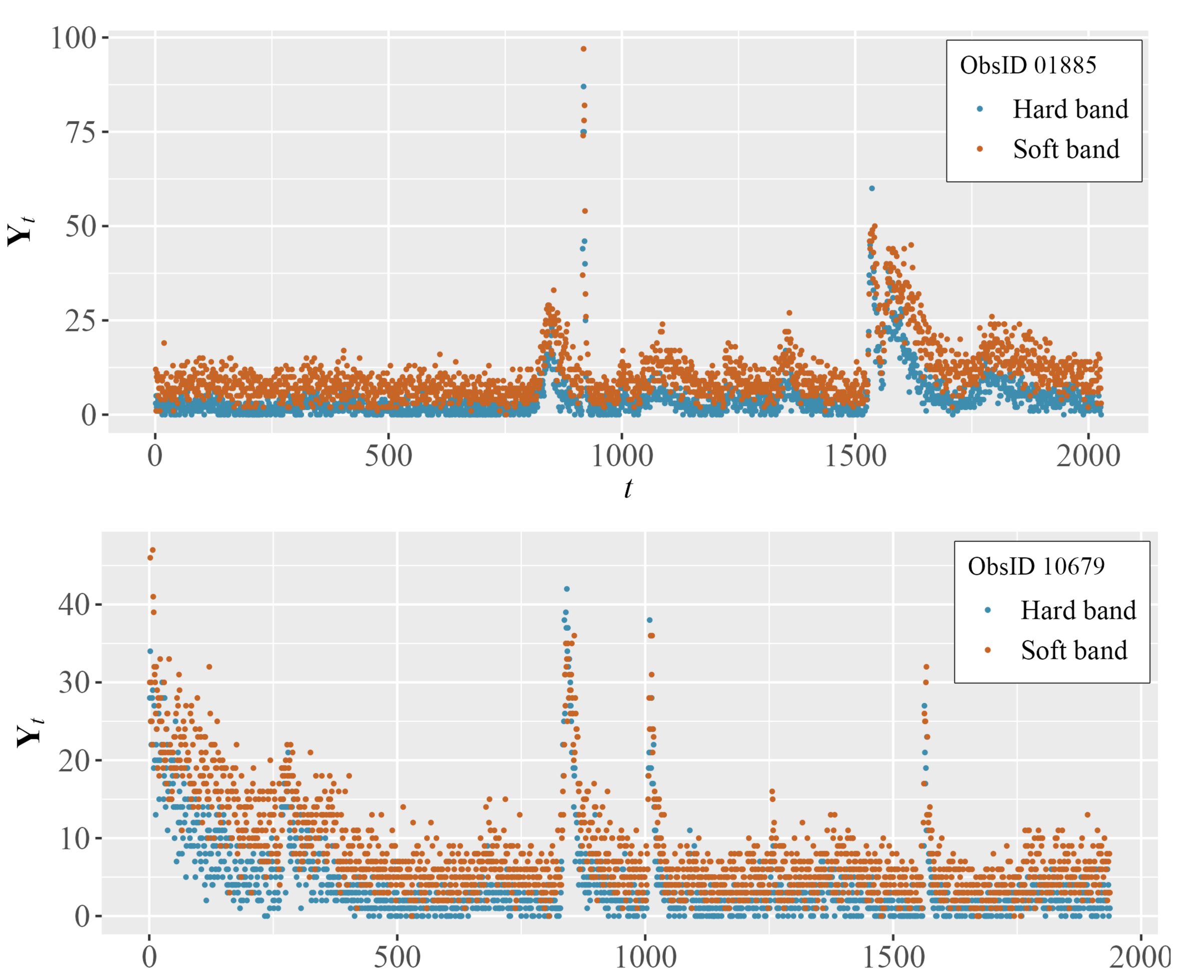

Figure 1: Two light curves of EV Lac (ObsID 01855 from December 2001 and ObsID 10679 from March 2009), each split into hard (1.5–8.0 keV) and soft (0.3–1.5 keV) bands

- Certain stars produce sporadic short-duration flares from their coronae
- We want to understand the proportion of time these stars spend in flaring versus quiescent states
- For nearby stars like the sun, we have much directly-observed "continuous" information (e.g., [1])
- For distant stars emitting X-rays, we have only light curves computed from the time and energy of photons recorded by X-ray telescopes like Chandra
- Most previous work applied ad-hoc rules of black-box/model-free learning methods to estimate flaring states; the best guess for EV Lac is 39% of time spent flaring [2]
- Our new approach is to model the flaring and quiescent states in two stages using hidden Markov models:

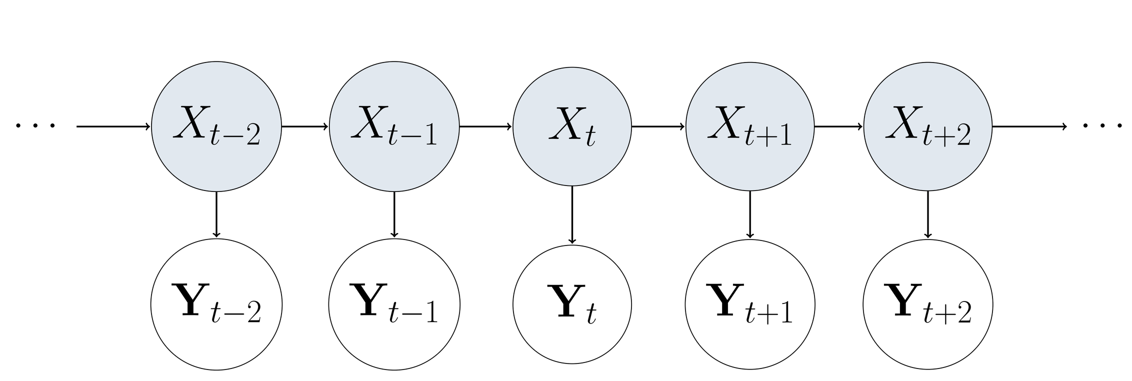

Figure 2: A graphical model representing the standard discrete-time HMM dependence structure

Stage 1: HMMs for Flaring Sources

- A (discrete-time) *hidden Markov model (HMM)* assumes that a latent Markov process $X_{1:T} = (X_1, X_2, \dots, X_T)$ generates observed data $\mathbf{Y}_{1:T} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_T)$ such that $\mathbf{Y}_s$ and $\mathbf{Y}_t$ are conditionally independent given $X_{1:T}$
- When the state space $\mathcal{X}$ is finite — say $\mathcal{X} = \{1, 2, \dots, K\}$ — the HMM is a *discrete-space* HMM, with initial probabilities $\mathbb{P}(X_1 = i)$, transition probabilities $\mathbb{P}(X_t = j \mid X_{t-1} = i)$, and state-dependent mass functions $\mathbb{P}(\mathbf{Y}_t = \mathbf{y}_t \mid X_t = i)$ for $i, j \in \mathcal{X}$
- We can calculate and maximize the likelihood efficiently and then predict the hidden states at each time t — however, the conditional independence assumption fails for our data:

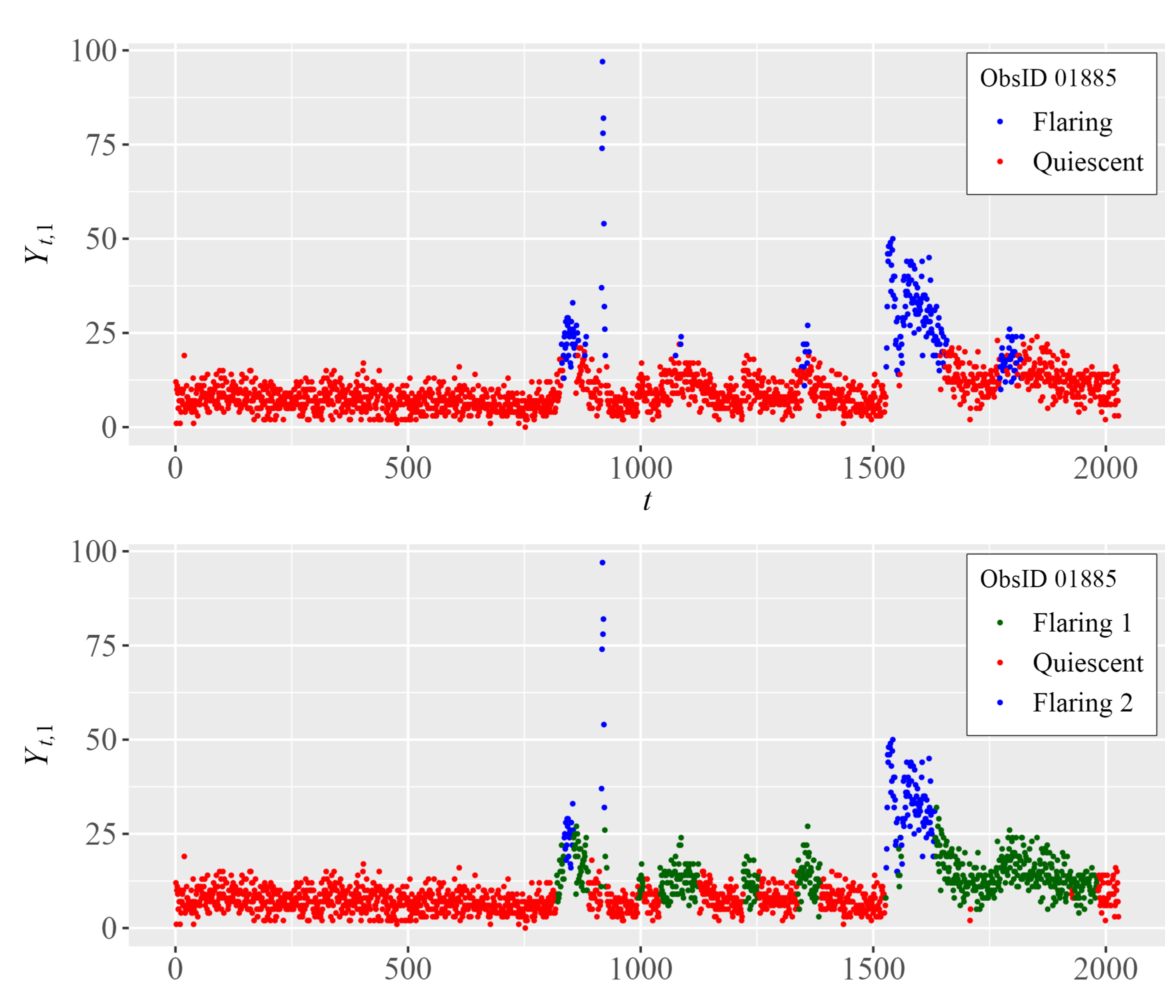

Figure 3: December 2001 soft-band light curve colored with classifications based on 2- and 3-state HMMs

- In contrast, for *continuous-space* HMMs, the underlying Markov chain takes values in $\mathcal{X} = \mathbb{R}^d$, which allows for smooth transitions between quiescence to flaring
- While the likelihood is now intractable, we can approximate it using a discrete-space likelihood
- We consider several variants of the model and select a VAR(1) process on a line:

$$\mathbf{Y}_t \mid X_t \sim \mathrm{Pois}(w \cdot \beta_1 \cdot e^{X_t}) \cdot \mathrm{Pois}(w \cdot \beta_2 \cdot e^{\sigma_2 X_t/\sigma_1}),$$
$$X_t = \phi X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$$

Stage 2: Classification Into Flaring and Quiescent Intervals

- Once the Stage 1 HMM is fit, we make posterior state predictions for each state X_t as $\hat{X}_t = \operatorname*{argmax}_{x_t \in \mathcal{X}} \mathbb{P}_{\hat{\boldsymbol{\theta}}}(X_t = x_t \mid \mathbf{Y}_{1:T} = \mathbf{y}_{1:T})$
- We then view the predictions $\hat{X}_1, \dots, \hat{X}_T$ as fresh "data" and approximate their distribution by a 2-component mixture: $\alpha \cdot F_1 + (1 - \alpha) \cdot F_2$
- Assuming that the distribution F_2 corresponds to "flaring", $(1 - \alpha)$ is the overall proportion of time spent in this state
- **If we see a clear sustained quiescent period** $[t_1, t_q]$, we use $\hat{X}_{t_1:t_q}$ as training data for a KDE for the quiescent mixture component
- We approximate the flaring mixture component with a step function, and fit the entire mixture using a custom-designed EM algorithm
- This yields $100\% \cdot (1 - \hat{\alpha}) \approx 45\%$ for the September 2001 EV Lac data, agreeing with [2]
- **If we don't see such a quiescent period**, we use a 3-component normal mixture which is easily fit with an EM algorithm
- Applying this method to the March 2009 EV Lac data yields $100\% \cdot \hat{\alpha}_3 \approx 27\%$, also agreeing with [2]

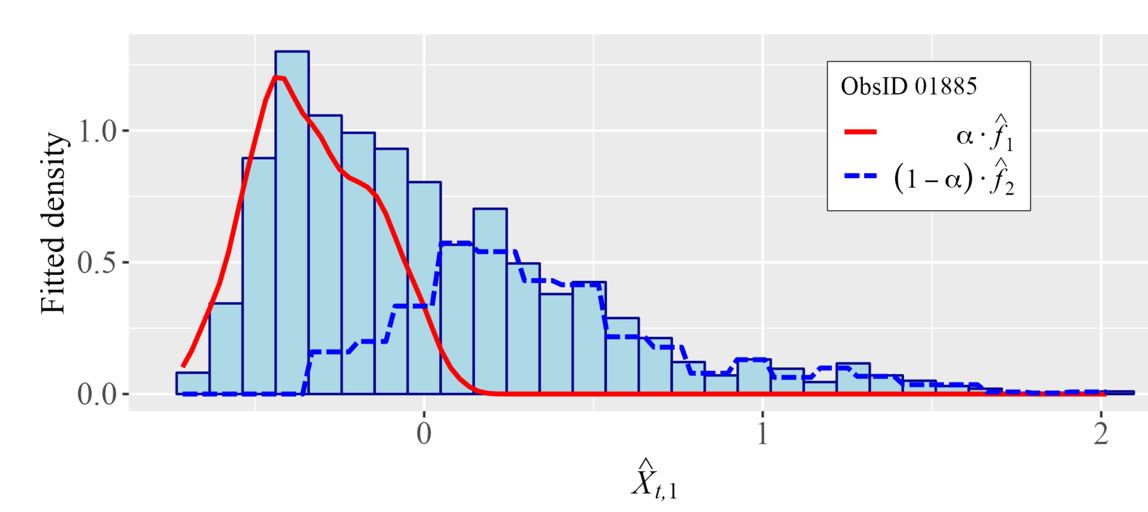

Figure 4: Fitted component densities overlaid on a histogram of the predicted states $\hat{X}_1, \dots \hat{X}_T$ for September 2001

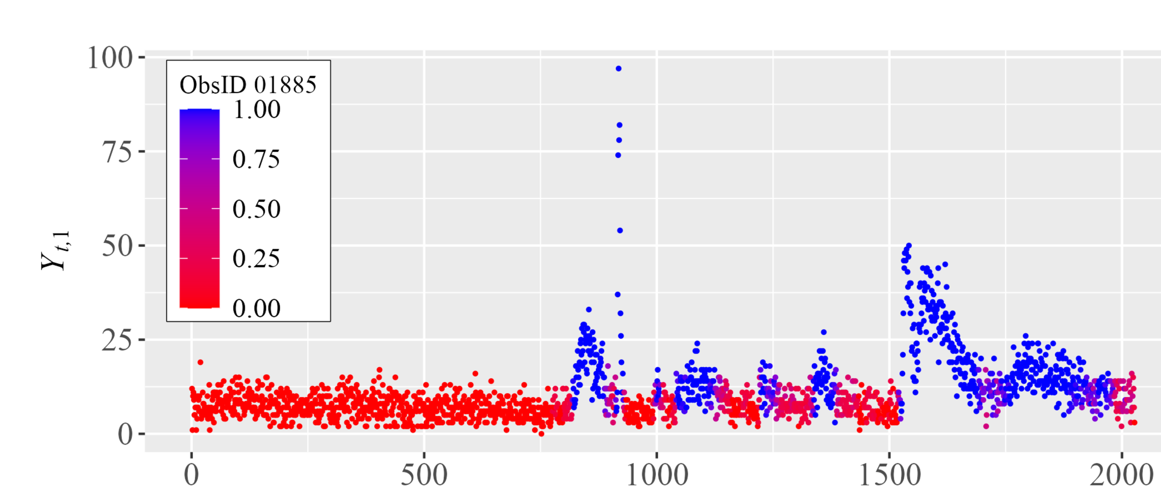

Figure 5: Posterior flaring state probabilities used to color the soft-band counts for September 2001

References

[1] Stanislavsky A. et al., 2020, *J. Atmos. Sol.-Terr. Phys.*, 208, 105407

[2] Huenemoerder D.P. et al., 2010, *ApJ*, 723, 1558

Statistical Sciences UNIVERSITY OF TORONTO

IMPERIAL

CENTER FOR ASTROPHYSICS HARVARD & SMITHSONIAN