Detecting stellar flares in photometric data using hidden Markov models

J. Arturo Esquivel [ID],[1, 2] Yunyi Shen [ID],[3] Vianey Leos-Barajas [ID],[1, 2, 4]
Gwendolyn Eadie [ID],[1, 2, 5] Joshua S. Speagle [ID],[1, 2, 5] Radu V Craiu [ID],[1, 2] Amber Medina [ID],[6]
and James R. A. Davenport [ID][7]

[1]*Department of Statistical Sciences, University of Toronto, Toronto, ON, Canada*
[2]*Data Sciences Institute, University of Toronto, Toronto, ON, Canada*
[3]*Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA*
[4]*School of the Environment, University of Toronto, Toronto, ON, Canada*
[5]*David A. Dunlap Department of Astronomy & Astrophysics, University of Toronto, Toronto, ON, Canada*
[6]*Department of Astronomy, University of Texas-Austin, Austin, TX, USA*
[7]*Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA*

ABSTRACT

We present a *hidden Markov model* (HMM) for discovering stellar flares in light curve data of stars. HMMs provide a framework to model time series data that are non-stationary; they allow for systems to be in different states at different times and consider the probabilities that describe the switching dynamics between states. In the context of stellar flares discovery, we exploit the HMM framework by allowing the light curve of a star to be in one of three states at any given time step: *Quiet*, *Firing*, or *Decaying*. This three-state HMM formulation is designed to enable straightforward identification of stellar flares, their duration, and associated uncertainty. This is crucial for estimating the flare's energy, and is useful for studies of stellar flare energy distributions. We combine our HMM with a *celerite* model that accounts for quasi-periodic stellar oscillations. Through an injection recovery experiment, we demonstrate and evaluate the ability of our method to detect and characterize flares in stellar time series. We also show that the proposed HMM flags fainter and lower energy flares more easily than traditional sigma-clipping methods. Lastly, we visually demonstrate that simultaneously conducting *detrending* and flare detection can mitigate biased estimations arising in multi-stage modelling approaches. Thus, this method paves a new way to calculating stellar flare energy. We conclude with an example application to one star observed by TESS, showing how the HMM compares with sigma-clipping when using real data.

Keywords: TESS, M-dwarfs, Time-series, stellar flares

Corresponding author: J. Arturo Esquivel
a.esquivel@mail.utoronto.ca

1. INTRODUCTION

Almost all stars in the universe with convection surface produce *stellar flares* — bursts of energy emitted from the star that are thought to be caused by magnetic reconnection (see, e.g., Forbes 1991; Donati & Landstreet 2009). Properly estimating the energy distribution of flares as a function of a stars' mass, age, and other characteristics is fundamental for understanding (i) the evolution of stellar magnetic fields, (ii) stellar rotation and mass-loss rates, and (iii) the highly energetic radiation environment to which planets orbiting these stars are subjected. Past studies have found relationships between flare energies and decay times, frequency of flares and stellar rotation rates, and flare duration and peak luminosity (e.g., see Davenport 2016; Medina et al. 2020; Chang et al. 2015).

Flares are detected in the time series data of a star's brightness measurements — a sudden, sharp increase in brightness followed by a slower decay usually indicates a stellar flare. However, detecting stellar flares is complicated by the fact that most stars also exhibit small, quasi-periodic oscillations in their brightness over time.

It is common practice to identify flares in time series data using non-parametric models, before assuming anything about their time series signature or shape (Chang et al. 2015). Current methods to detect stellar flares in time series data rely on multi-stage data processing and "sigma-clipping"; after the stationary and quasi-periodic part of the time series is modeled and removed (called *detrending*), points lying outside a pre-defined confidence interval are highlighted as potential flares (e.g. Davenport 2016; Hawley et al. 2014; Osten et al. 2012; Walkowicz et al. 2011; Yang et al. 2017; Günther et al. 2020). Medina et al. (2020) also note that in the 3-sigma approach to flare detection and flare energy estimation, the largest source of uncertainty comes from defining the end of the flare. The sigma-clipping approach may also struggle

to identify compound flares, although change-point detection does not suffer from this problem (Chang et al. 2015).

After the flares' locations in the time series are detected, a template or model for flare shape (e.g., Davenport et al. 2014) is often used to estimate flare parameters. The detrending process is often done using a flexible model, such as Gaussian processes (e.g. *celerite* Foreman-Mackey et al. 2017). However, detrending methods are typically done before the flare detection step, and this pre-processing may absorb lower-energy flares in the time series data, and bias the energy profile. Therefore, it would be beneficial to model the trend of the time series and the flares simultaneously.

Other methods such as change-point detection have also been explored to identify potential flares (Chang et al. 2015). While this approach detects the most energetic flares, it can struggle to detect the medium- to low-energetic flares that are part of the flare energy distribution. The *stella* software, which uses a convolutional neural network (CNN) to find flares in TESS data, is an efficient tool for finding flares, but still relies on a probability threshold for flare detection (Feinstein et al. 2020).

In this work, we introduce a new approach to stellar flare detection using *hidden Markov models (HMMs)*. HMMs are flexible time series models that are popular in many domains, including ecology (McClintock et al. 2020; Adam et al. 2019; Leos-Barajas et al. 2017), health (Williams et al. 2020) and sports (Ötting et al. 2021). The advantage of using HMMs in the context of stellar flare detection is that they are more likely to detect medium- to low energy flares than traditional sigma-clipping methods.

HMMs provide a way to model different states underlying a time series, with a probability associated to the transitions between them (Zucchini et al. 2017). This is a very natural scheme to approach the detection of stellar flares — a

star may be in a "quiet", flare-"firing", or flare-"decaying" state. Thus, if an HMM is fit to the light curve of a star, every point in that time series can be estimated to come from one of these three states. This allows one to discover both the firing and decaying phases of each flare. The decaying state is particularly helpful, as it helps characterize the end of a flare as the star transitions back to the quiet state.

In addition to using an HMM, we simultaneously fit *celerite* to model the quasi-periodic trend in the star's light curve. We show that this simultaneous fitting of *celerite* and the HMM not only removes the need for iterative fitting when searching for flares, but also improves the *celerite* fit overall. That is, *celerite* does not as easily absorb small flares, nor the decaying portion of larger flares. For our entire analysis, we adopt a Bayesian approach.

This paper presents our HMM for stellar flares detection, shows its merits, and applies it to a stellar light curve measured from TESS. Our paper is organized as follows: In Section 2, we describe the data that motivated this study. In Section 3, we thoroughly describe our method; we begin with a quick overview of *celerite* (Section 3.1), followed by an introduction to HMMs (Section 3.2); then we proceed to describe our observational model (Section 3.3), our particular HMM (Section 3.4), our model fitting techniques (Section 3.5), how we identify and characterize flares (Section 3.6), and the injection recovery experiment performed (Section 3.7). The results of our injection recovery experiment and the application of our HMM to TESS data are presented in Section 4. We conclude with a discussion and a summary of future research directions in Sections 5 and 6.

2. DATA

To test and demonstrate our HMM approach, we use M dwarf TIC 031381302 two-minute cadence data measured by TESS. We use this star's Pre-Search Data Conditioning Simple Aperture Photometry (PDCSAP) light curve for both our injected flare tests (Section 3.7) and for a case study demonstration detecting real flares (Section 4.2). This star was chosen because it has long portions of the time series undergoing quiescent oscillations, in which simulated flares could be injected to test our method. At the same time, there are parts of the time series for this star that have known flare events — these portions of the light curve are used to demonstrate that our HMM can recover the same flares as other methods.

3. METHODS

To identify flares in a stellar brightness time series, we simultaneously model the quasi-periodic changes of the star and the star's flares. For the former, we use *celerite* (Foreman-Mackey et al. 2017) and for the latter we use the HMM described in this paper. Readers familiar with *celerite* may want to skip ahead to Section 3.2. Those readers familiar with HMMs may want to skip to our particular setup in Section 3.4.

For quick reference, a list of our mathematical notation is shown in Table 1.

3.1. *Celerite and detrending*

To account for the star's quasi-periodic changes as well as the mean brightness, we use *celerite*, a physically-motivated Gaussian Process (GP) widely used to model the trend of stellar light curves (Foreman-Mackey et al. 2017). In principle, one could use any kernel provided in *celerite*, but in this study we use the rotation kernel, the same one used in Medina et al. (2020), which consists of a sum of two simple harmonic oscillators.

For the GP, let μ, $\mathcal{K}$, and $\boldsymbol{f}$ denote the mean function, the kernel parameter, and the (latent) trend, respectively. The value of $\boldsymbol{f}$ at time t is denoted f_t. For a single star, the observed light curve Y_t is modeled as the sum of the trend f_t and a flaring channel Z_t,

Table 1. Notation used in this paper.

	Meaning
Y_t	observed brightness of star at time t
$\boldsymbol{f}$	quasi-periodic trend of the time series (modeled with *celerite)*
f_t	*celerite*-modeled trend at time t
Z_t	flaring channel (time series without trend)
$\boldsymbol{\mu}$	mean flux in Quiet state
$\mathcal{K}$	kernel for *celerite*
σ^2	variance of measurement noise
S_t	state of time series at time t
Q	Quiet state
F	Firing state
D	Decay state
$\log(\lambda)$	log average flux increase during firing state
$\mathrm{logit}(r)$	logit of decay rate
$p_{Q|Q}, p_{Q|F}$	transition probabilities from Quiet state
$p_{F|F}, p_{F|D}$	transition probabilities from Firing state
$p_{D|Q}, p_{D|F}, p_{D|D}$	transition probabilities from Decay state

$$\boldsymbol{f}|\mathcal{K}, \mu \sim \mathrm{Celerite}(\mu, \mathcal{K})$$
$$Y_t = f_t + Z_t. \tag{1}$$

We use priors for the parameters of μ and $\mathcal{K}$ recommended in Foreman-Mackey et al. (2017). The HMM for the flaring channel Z_t is described in Section 3.4.

3.2. *An introduction to HMMs*

In its basic form, an HMM is a doubly stochastic process composed of an observable, state-dependent, process $\{Z_t\}_{t=1}^T$ and a state process $\{S_t\}_{t=1}^T$. At each point in time, t, the time series is assumed to be in one of N possible states, (i.e., the time series follows a state process $S_1, S_2, ..., S_t$). The states are taken to be discrete latent variables generated from a first-order Markov chain that evolves over time according to an $N \times N$ transition probability matrix with entries $p_{i,j} = \mathrm{Pr}(S_t = j|S_{t-1} = i)$, for $i, j \in \{1, \ldots, N\}$. The state at time $t = 1$ is

taken to be generated according to an initial state distribution.

The observations are modeled assuming they are emitted from a set of state-dependent distributions, i.e.,

$$g_n(Z_t) = g(Z_t|S_t = n) \text{ for } n \in \{1, \ldots, N\}. \tag{2}$$

For example, if the HMM is a three-state model ($N = 3$), then there would be three different distributions g, which describe a different data generating process conditioned on each state n. In practice, the number of states N and the distributions g are defined using scientific domain knowledge. The parameters of these distributions may be estimated within the model or fixed. The $N \times N$ matrix describing the probability of transition between states is also estimated in practice.

When fitting an HMM to real time series data, one can use the estimated model to obtain the state sequence $\hat{s}_1, \ldots, \hat{s}_T$; the most likely, under the assumed model to underlie the observations. This is known as state decoding and can be effi-

ciently carried out through the Viterbi algorithm (Viterbi 1967; see also Forney 1973 for a detailed description).

The latter uses the estimated state-dependent distributions $\hat{g}_n$ to compute the probability density $\hat{g}_n(z_t)$ of each observation when a specific state $n \in \{1, \ldots, N\}$ is active. The algorithm combines these probability densities with the estimated transition probabilities to recursively determine the most likely sequence leading to each possible state n, at each time $t \in \{1, \ldots, T\}$. This is done through recursively computing the quantities

$$\xi_{t,n} = \left(\max_i(\xi_{t-1,i}\hat{p}_{i,n})\right)\hat{g}_n(z_t)$$

for $t = 2, \ldots, T$, initialized with $\xi_{1,n} = \widehat{\Pr}(S_1 = n)\hat{g}_n(z_1)$. i.e., $\xi_{t,n}$ corresponds to the likelihood of the state sequence from time 1 to $t - 1$ most likely to lead to state n being the one active at time t, given the observations between 1 and t. The most likely full state sequence is then determined by recursively maximizing over these likelihoods, starting with

$$\hat{s}_T = \arg\max_{n=1,\ldots,N} \xi_{T,n},$$

and setting

$$\hat{s}_t = \arg\max_{n=1,\ldots,N} \xi_{t,n}\hat{p}_{n,\hat{s}_{t+1}},$$

for $t = T - 1, T - 2, \ldots, 1$.

Additional references and some foundational papers about HMMs are Baum et al. (1970), Rabiner (1989), and Zucchini et al. (2017).

3.3. Observational Model

In our application, we assume an observational model for the random variable Y_t given state S_t:

$$Y_t|S_t = f_t + Z_t|S_t, \qquad (3)$$

where f_t is the celerite-modeled trend at time t (note that it does not depend on the underlying state), and Z_t is the *flaring channel*. The latter's

distribution depends on whether the star is in a quiet, firing, or decaying state, and is described next.

3.4. A Quiet-Firing-Decay HMM for Flare Events

We propose a three-state hidden Markov model (Hamilton 1990) for modeling flare events in the detrended light curve. Each point in the time series can result from one of three (hidden, unobservable) states: Quiet, Firing, or Decay (denoted as Q, F and D respectively). The Q state is used to model the time series when the star is not in any flare event, while the F and D states are used to model the increasing and decreasing phase of a flare.

Recall the probabilities p_{S_{t-1},S_t} of switching between states at each step in the time series (akin to a Markov chain); they are conditional probabilities of the form $pS_{t-1}|S_t$, where S_{t-1} is the previous state; for example, $p_{F|Q}$ denotes the probability of transitioning to state F given that the star is in state Q. The interpretation given to each of these transitions is illustrated in Table 2.

The transition from Q to F accounts for the firing rate of flares from the star's quiet state. Once the star is in the flaring state, the transition from F to F accounts for the increasing phase of a flare, while the transition from F to D accounts for the decay of the flare. When the time series is in the decay state, transiting from D to F corresponds to a compound flare. Note that we forbid the transition from Q to D (i.e., when the star is quiet, it will not suddenly "decay") and from F to Q (i.e., when the star is flaring, the flare will not spontaneously disappear). All other transitions are allowed, and are each modeled with parameters that account for different physical characteristics.

We model the flaring channel Z_t at time t given the states $S_t \in \{Q, F, D\}$ and the previous step Z_{t-1} as follows:

Table 2. Interpretation of transitions between states

| | | to | | |
		Q	F	D
from	Q	remain quiet	start firing	*forbidden*
	F	*forbidden*	increased firing	start decaying
	D	return to quiet	start compound flare	decaying continues

$$
\begin{aligned}
Z_t|(S_t = Q, Z_{t-1}, \sigma^2) &\sim \mathcal{N}(0, \sigma^2), \\
Z_t|(S_t = F, Z_{t-1}, \lambda, \sigma^2) &\sim \mathcal{N}(Z_{t-1}, \sigma^2) + \\
&\quad \text{Exp}(\lambda), \\
Z_t|(S_t = D, Z_{t-1}, r, \sigma^2) &\sim \mathcal{N}(rZ_{t-1}, \sigma^2).
\end{aligned}
\tag{4}
$$

The distributions above are the g_n distributions mentioned in Section 3.2 and eq. (2). Note that the value of the time series at time t is always dependent on the current state of the star S_t and on the value of the time series in the previous step (Z_{t-1}) when in states F or D. The Q state is modeled as an independent normal random variable with variance σ^2. When formulating the distributions in (4) we assume that: when quiet, the flaring channel will just be measurement noise on top of the quiescent trend; when firing, the flaring channel will be around the previous channel plus an independent flux increase exponentially distributed; and when decaying, the flaring channel will be centered around a scaled value of the previous channel.

Recall that the model for the observed light curve y_t is the combination of the trend and the flaring channel; that is, *celerite* and the HMM are fit simultaneously, such that $\boldsymbol{f}$ is a latent variable (Figure 1). We call this combined model *celeriteQFD*. A summary of the HMM parameters, their prior distributions, and hyperparameter values can be found in Table 3. invGamma is the inverse Gamma distribution and the transition probabilities (e.g., $p_{Q|F}, p_{F|D}$, etc.) use Dirichlet prior distributions. This setting is used for all injection-recovery and real data examples.
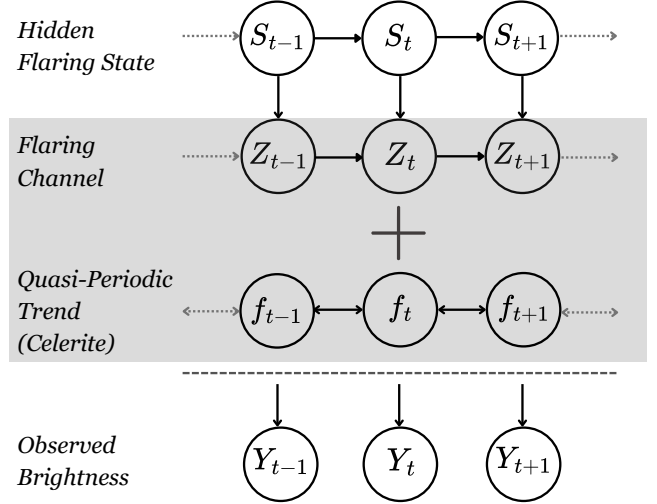

Figure 1. Graphical representation of our complete model for observed brightness decomposed into its various parts. The Quasi-Periodic Trend reflects the average brightness of the star when it is not flaring. The Flaring Channel represents the extra brightness due to the state of the star. The Hidden Flaring State represents the (unobserved) state of the star (Q,F, or D).

3.5. *Computation, Model Fitting, and State Decoding*

For computational purposes, we split each time series $\{y_t\}_{t=1}^T$ into smaller chunks of 2000 time steps. We subtract the overall mean from the light curve to center it around zero. Our model was implemented in stan (Carpenter et al. 2017), while the C++ code for *celerite* was adopted from the Python library EXO-PLANET. The posterior distributions of model parameters and derived quantities are sampled using a dynamic Hamiltonian Monte Carlo algorithm (Hoffman & Gelman 2014). States are decoded using the Viterbi algorithm as described in Section 3.2

Table 3. Prior setting of the QFD part of the model

Parameter	Meaning	Prior distribution	Hyperparameter used
μ	Mean flux at quiet	$N(\mu_0, \sigma_0^2)$	$\mu_0 = 0, \sigma_0^2 = 100\sigma^2$
σ^2	variance of measurement noise	$\text{invGamma}(\alpha, \beta)$	$\alpha = .01, \beta = .01$
$\log(\lambda)$	log average flux increasing during firing	$N(\mu_\lambda, \sigma_\lambda^2)$	$\mu_\lambda = 0, \sigma_\lambda^2 = 1e3$
$\text{logit}(r)$	logit of decay rate	$N(\mu_r, \sigma_r^2)$	$\mu_r = 0, \sigma_r^2 = 1e3$
$p_{Q|Q}, p_{Q|F}$	transition probabilities from Quiet state	$\text{Dir}(\alpha_Q)$	$\alpha_Q = (1, 0.1)$
$p_{F|F}, p_{F|D}$	transition probabilities from Firing state	$\text{Dir}(\alpha_F)$	$\alpha_F = (1, 1)$
$p_{D|Q}, p_{D|F}, p_{D|D}$	transition probabilities from Decay state	$\text{Dir}(\alpha_D)$	$\alpha_D = (1, 0.1, 1)$

(see also Forney 1973). The `stan` implementation as well as the injection recovery tests can be found in the first author's github repository Esquivel-Arturo/celeriteQFD.

To obtain samples from the posterior distributions of the model parameters, we use two MCMC chains and sample 1000 (2000 in total) posterior samples, after discarding the first 1000 samples obtained during the warm-up period. For each joint posterior sample of the parameters, we use the Viterbi algorithm to uncover the most likely state sequence that could have generated the data. In this manner we are able to propagate the uncertainty around our parameters to produce 2000 most likely state sequences and capture the variability in state decoding results. Thus, for every point in the time series we have a "decoding distribution" (see Figure 2 top panel) of the states (Q,F,D). For each point in the time series, we estimate the state of the star to be the one that appears most frequently across the 2000 decodings of that time step.

Each *celeriteQFD* implementation on 2000 time steps took between 1 and 4 hours to run using two cores of a M1 MacBook Pro with 16 GB of memory.

3.6. *Identifying and Characterizing Flares*

To identify flares after fitting *celeriteQFD*, we consider, all consecutive points decoded in a non-Q state (i.e., the rise F and fall D of the flare) are used to define the duration of a flare. A flare is considered over once the time series returns to state Q. In other words, the duration of a flare is defined as the time elapsed from when the star enters the flaring state (Q to F) to when the star re-enters the quiet state (D to Q). For example, in Figure 2 the flare was estimated to commence with the peak red point and end right before time 1331.30.

This method also allows us to find compound flares (e.g., the time series could be in the decaying phase of a flare, and then start firing again). As the state decoding of a compound flare will have multiple peaks we skip any peaks identified in a flare's duration when searching for the next flare. We can also quantify uncertainty around the Viterbi "decoding distributions", e.g., Figure 3 shows two flares detected very close to each other. From the decoding proportions we can see 10% of the Viterbi state decodings point to a compound flare instead of two separate flares.

3.7. *Injection Recovery Experiment*

To test the ability of our HMM to detect stellar flares, we perform an injection recovery experiment: we inject simulated flares into a real stellar time series, apply our HMM algorithm to detect the simulated flares, and compare our results to the ground truth and those of using a sigma-clipping approach.

We use the mean-centered flux time series data from one star as the base time series for our injection recovery experiment, and randomly inject
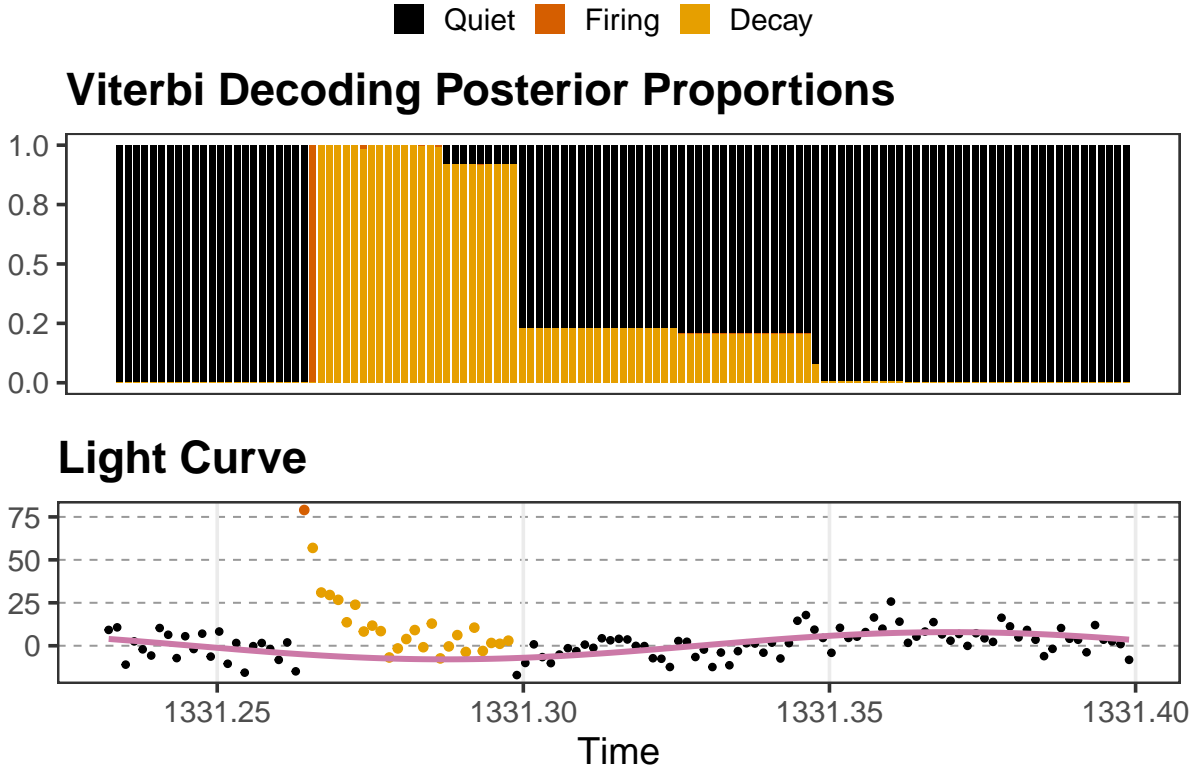
Viterbi Decoding Posterior Proportions

Light Curve

Figure 2. Detected flare example from implementing *celeriteQFD* on the real time series of TIC 031381302 *Top panel:* State "decoding distributions" across all Viterbi sequences per observation. *Second panel:* shows the fit of *celeriteQFD*, the estimated trend (purple curve) and assigned state to each point (black points are state Q, red points F, and orange points are D).

Kepler flares following the procedure outlined in Davenport et al. (2014). We use TIC 031381302 (from day 1325.292 to day 1327.377, $n = 1501$). This time segment was chosen to avoid already-discovered natural flares.

We inject five flares, both small and large, at randomly chosen points in the time series. The time scale of our Kepler-like flares, $t_{1/2}$, is proportional to its peak flux. The peak fluxes of the injections are *i.i.d.* Pareto (i.e., they follow a truncated power-law distribution):

$$p(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}, \tag{5}$$

where x_m and α are the distribution parameters.

We use different parameter values for $(t_{1/2}, x_m, \alpha, \delta, \beta)$ to simulate small and large flares. We use $(5 \times 10^{-5}, 10, 1, 30, 150)$ for small flares, and $(5 \times 10^{-5}, 50, 1, 0, 300)$ for large flares.

For each set of parameter values, we separately simulate and inject five flares into the time series, and perform our analysis of flare recovery. We repeat the procedure 100 times for each parameter scheme. Although we do not explicitly study our method's ability to recover compound flares, we do allow the simulated flares to overlap in time and form compound flares in the base time series.

3.8. Flare Detection Evaluation

For each simulated time series, we run *celeriteQFD* and obtain a Bayesian estimate of the state of the star (Q,F, or D) at every time step (see Section 3.5 and Figure 2). Once we have the estimated states for all points in a time series, flares are identified as described in Section 3.6.

To evaluate the accuracy of flare detection using our HMM framework (Section 3.2), we com-

pare the number of detected flares to the ground truth (i.e., flares injected). For each fitted model (100 per method and flare scheme), we calculate the true positive, false positive, and false negative rates of detection, as well as the sensitivity and the positive predictive value (PPV):

$$sensitivity = \frac{TFD}{TFT}, \qquad (6)$$

$$PPV = \frac{TFD}{FD}, \qquad (7)$$

where TFD is the number of true flares detected; TFT the total number of true flares; and FD the total number of flares detected. We also assess flare detection in terms of the full duration of the flaring processes, i.e. we compute the per observation sensitivity and PPV:

$$sensitivity = \frac{TFD_o}{TFT_o}, \qquad (8)$$

$$PPV = \frac{TFD_o}{FD_o}, \qquad (9)$$

where TFD_o is the number of observations part of a true flare correctly identified; TFT_o is the total number of observations that are part of a true flare; and FD_o is the total number of observations identified to be part of a detected flare. Note that sensitivity and PPV ((6) and (7)) should each be 1 if we perfectly identify all true flares. Similarly, the per observation metrics ((8) and (9)) should be 1 if the entire duration of every flare is correctly identified.

We compare our results to those obtained with the more commonplace sigma-clipping a–$b\sigma$ rule, where a is the number of consecutive points in the time series that are $b\sigma$ away from the mean flux μ of the detrended time series (i.e., the rule outlined in Chang et al. 2015 and used in e.g. Medina et al. 2020; Ilin et al. 2019). In particular, we compare against using a 1–3σ approach and when fitting *celerite* for detrending in the this

method, we use the same priors and kernel as with *celeriteQFD*.

4. RESULTS

4.1. *Injection recovery*

An example of one injection recovery experiment under the small flare scheme is shown in Figure 4. The flux was centered, i.e. it is the raw flux data from which we subtracted the grand mean. It shows the estimated trend and state sequence from our HMM (second panel), along with the ground truth (top panel) and with the result of using a $1 - 3\sigma$ approach (third panel). We use $1 - 3\sigma$ because it is more sensitive than the usual $3 - 3\sigma$ and so it is more likely for it to detect small injections. In this example it can be seen how, through state D, the HMM directly identifies a larger proportion of flaring events than sigma-clipping. Moreover, note that *celerite* can absorb part of the flares into the estimated trend of the time series, and so reduce the chance that a flare is detected by sigma-clipping rules.

The results of the injection recovery experiments are presented in Figure 5. The box plots show a comparison of our method's detection performance with that of $1 - 3\sigma$ clipping for small and large flares schemes. They show the sensitivity and PPV distributions across the 100 simulations of each setting. e.g., the first (light-blue) box on the top-left panel corresponds to the distribution of the per flare sensitivity (see (6)) across all 100 sigma-clipping models fitted to the 100 injections of five small flares simulated.

From part (a) of Figure 5, it can be seen that both methods perform almost ideally when it comes to detecting the occurrence of large and small flaring events. Flare detection sensitivity, i.e., the probability of correctly detecting a flare was computed to be 1 in almost all the 100 simulations. They also achieve a similar performance in terms of the PPV (see (7)); i.e., the probability of an identified flare occurrence
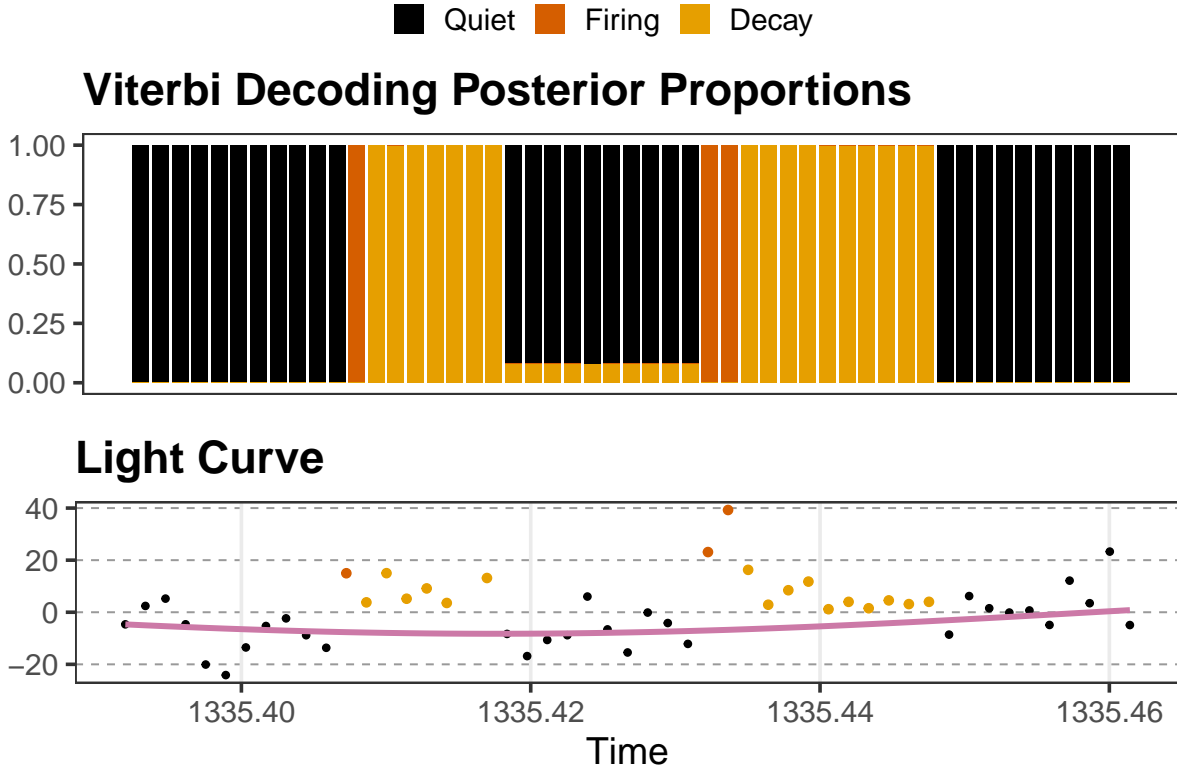
Figure 3. Example of two flares detected from implementing *celeriteQFD* on the real time series of TIC 031381302. *Top panel:* State "decoding distributions" across all Viterbi sequences per observation. *Second panel:* shows the fit of *celeriteQFD*, the estimated trend (purple curve) and assigned state to each point (black points are state Q, red points F, and orange points are D).

being correct. However, under the large flares scheme, *celeriteQFD* often performed slightly better, achieving PPVs of 1 (it was 0.8 almost always with sigma-clipping). This indicates that our method is less susceptible to producing false detections.

These results suggest our proposed HMM framework is at least as good as sigma-clipping for flaring events detection tasks. But the meaningful difference of our method consists of its ability to directly provide an estimate of the full duration of flaring events. Figure 5 part (b) contains the distributions of the per observation performance metrics (see (8) and (9)). i.e., the metrics are computed using all points identified as part of a flare by the methods, as compared with all the points that truly belong to injected flares. The plots show that both methods very rarely flagged observations outside of a real flare

(PPVs are concentrated very close to 1). Also one can clearly see the difference between methods when it comes to spotting all light points that are part of a flare; *celeriteQFD* consistently identified more than 50% of the observations forming part of a small flare and close to 70% for large flares.

Sigma-clipping alone is never used for a full characterization of a flare, which is usually done through further data modeling steps (see Chang et al. (2015) for example). Still, this experiment demonstrates the capacity our method has to describe the entire duration of detected flares without the need for extra steps.

Another crucial difference of our method is that it simultaneously carries out detrending and flare detection. By considering the light curve observations to be a combination of the long-term trend and a flaring channel (see (3)),
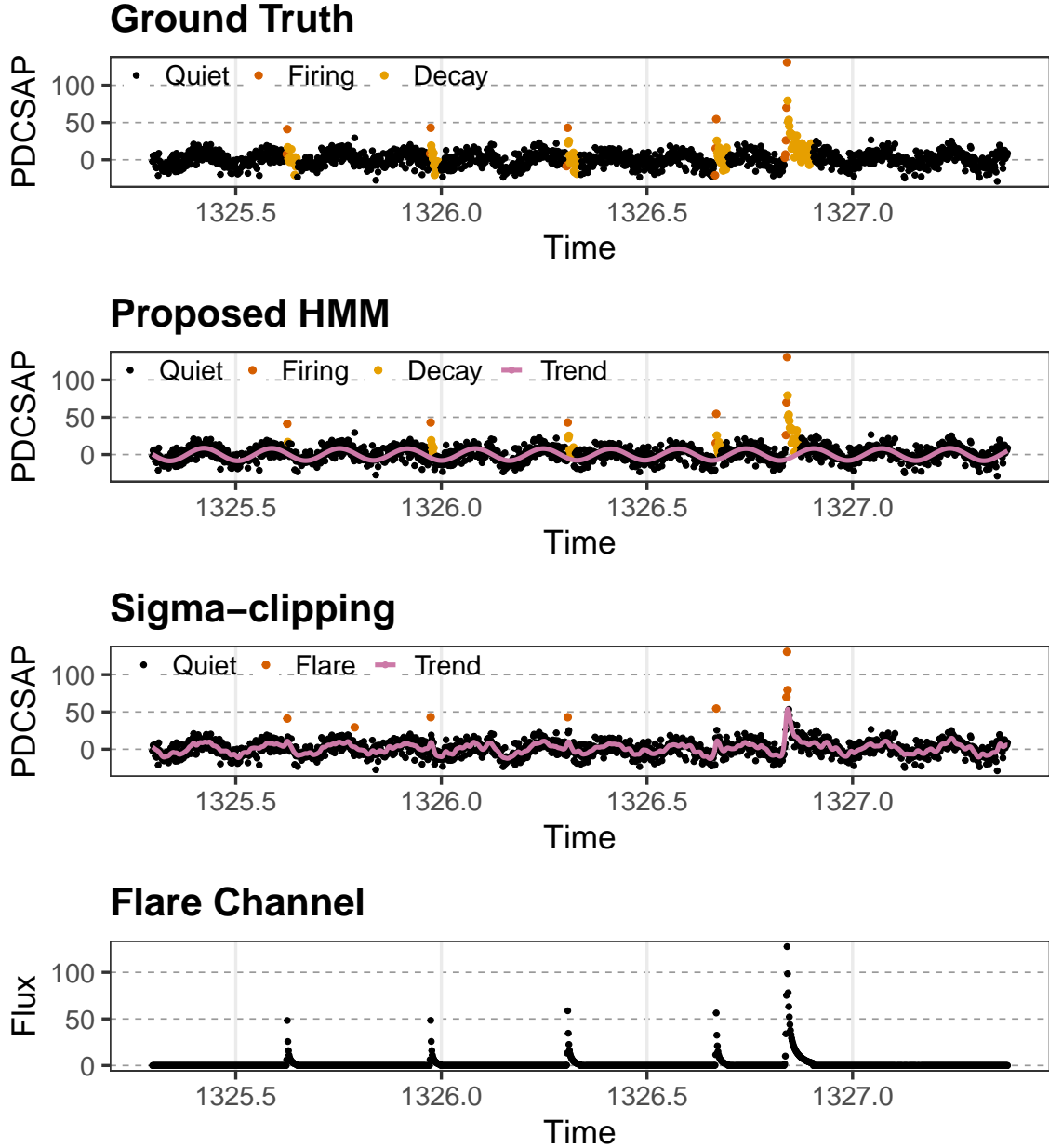
Ground Truth

Proposed HMM

Sigma–clipping

Flare Channel

Figure 4. An example of one injection recovery simulation in the small flare scheme. *Top panel:* ground truth — the real time series for TIC 031381302 with five simulated flares injected (their real states colour-coded). *Second panel:* shows the fit of our proposed algorithm that simultaneously models the trend with *celerite* (purple curve) and assigns states to each point in the time series (black points are state Q, red points F, and orange points are D). *Third panel:* the sigma-clipping approach that uses *celerite* alone to model the trend (purple curve), with outliers beyond 3σ (red points) used to identify flares. *Bottom panel:* the flare channel that was injected into the time series.

both components are modeled accounting for the effects of the other part on the observational process. i.e., the posterior distributions of the

celerite parameters contain information on the HMM parameters and vice-versa.

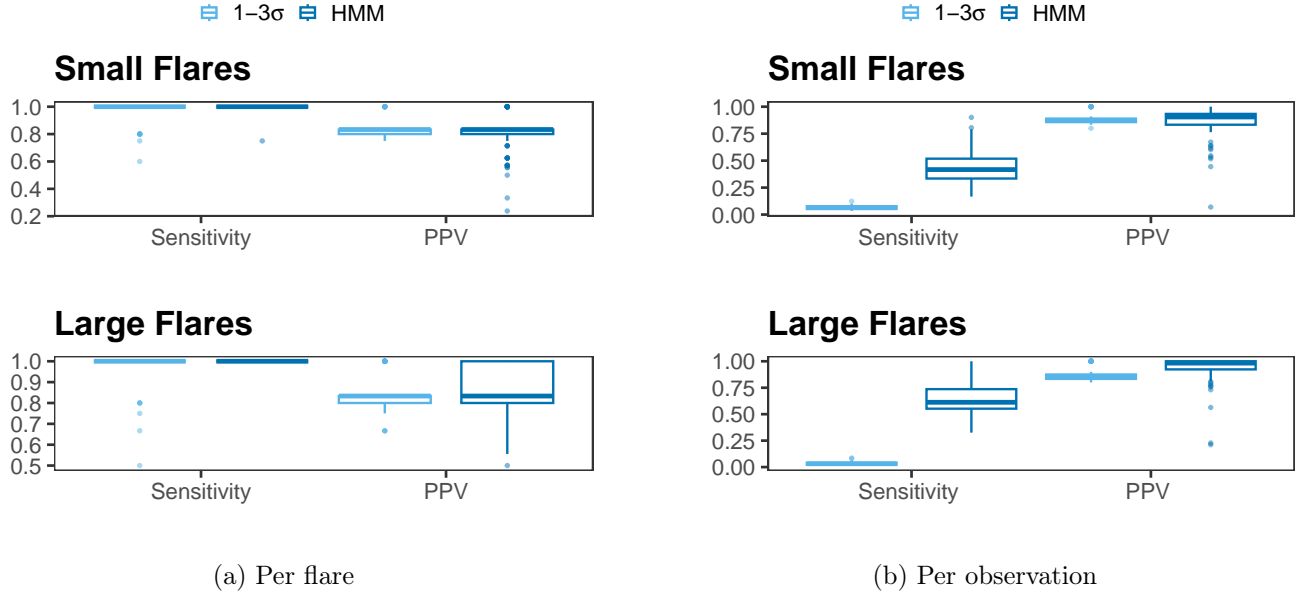Figure 6 illustrates this by showing a zoomed-in flare detected by both methods from the exam-

(a) Per flare

(b) Per observation

Figure 5. Flare recovery sensitivity and positive predictive value (PPV) distributions across 100 small and large flares injections, using the 1-3σ rule and *celeriteQFD* (HMM). *(a)* shows results from detecting flare occurrences (equations (6) and (7) are used). *(b)* shows results on a per-observation basis (equations (8) and (9) are used).

ple in Figure 4. The first thing to note, is that, overall, the *celerite*-estimated trend is considerably less affected by light curve variability when fitted simultaneously with the HMM. Moreover, in the zoomed flare, note that *celerite*, when used alone, can absorb part of the flares into the trend. Since the data points correspond to an injected flare, we know the increased flux of those observations is not part of the long-term pattern of the light curve and that detrending should ideally ignore it. *celeriteQFD*, is not only able to estimate the trend unaffected by the increased flux, it is also capable of identifying many of the observations as part of the decaying phase of a flare. This result is critical, suggesting that better detrending can be achieved, further leading to less biased estimates of the flares' energies.

4.2. *Case Study: Photometric Data from the TESS Mission*

As a demonstration of real flares detection, we apply our HMM method to a large portion of the TIC 031381302 light curve. The time series was sliced into pieces of 2000 time steps to

more efficiently fit the models and conduct flare detection. Trace plots of the MCMC sub-chains corresponding to the parameters were produced and inspected without any indication of lack of convergence.

A large portion (from day 1325.292 to day 1353.177) of TIC 031381302 PDCSAP mean-centered flux, along with the resulting fit and state decoding, are shown in Figure 7. During the period observed, a total of 11 flares were detected, with an average duration of 10.3 observations (approximately 0.01545 days or 1334.88 seconds). The estimated (using the posterior median) transition probability ($pQ|F$) was 0.00170, with a 95% credible interval of $(0.00007, 0.00820)$. i.e., the estimated probability that this star starts firing at any particular time, given that it was quiet in the previous time step, is 0.17%.

A smaller portion (from day 1333.627 to day 1336.404) of both models fitted to the TIC 031381302 are shown in Figure 8. The right-hand side shows a zoom-in into a portion containing what both methods identified as a flare
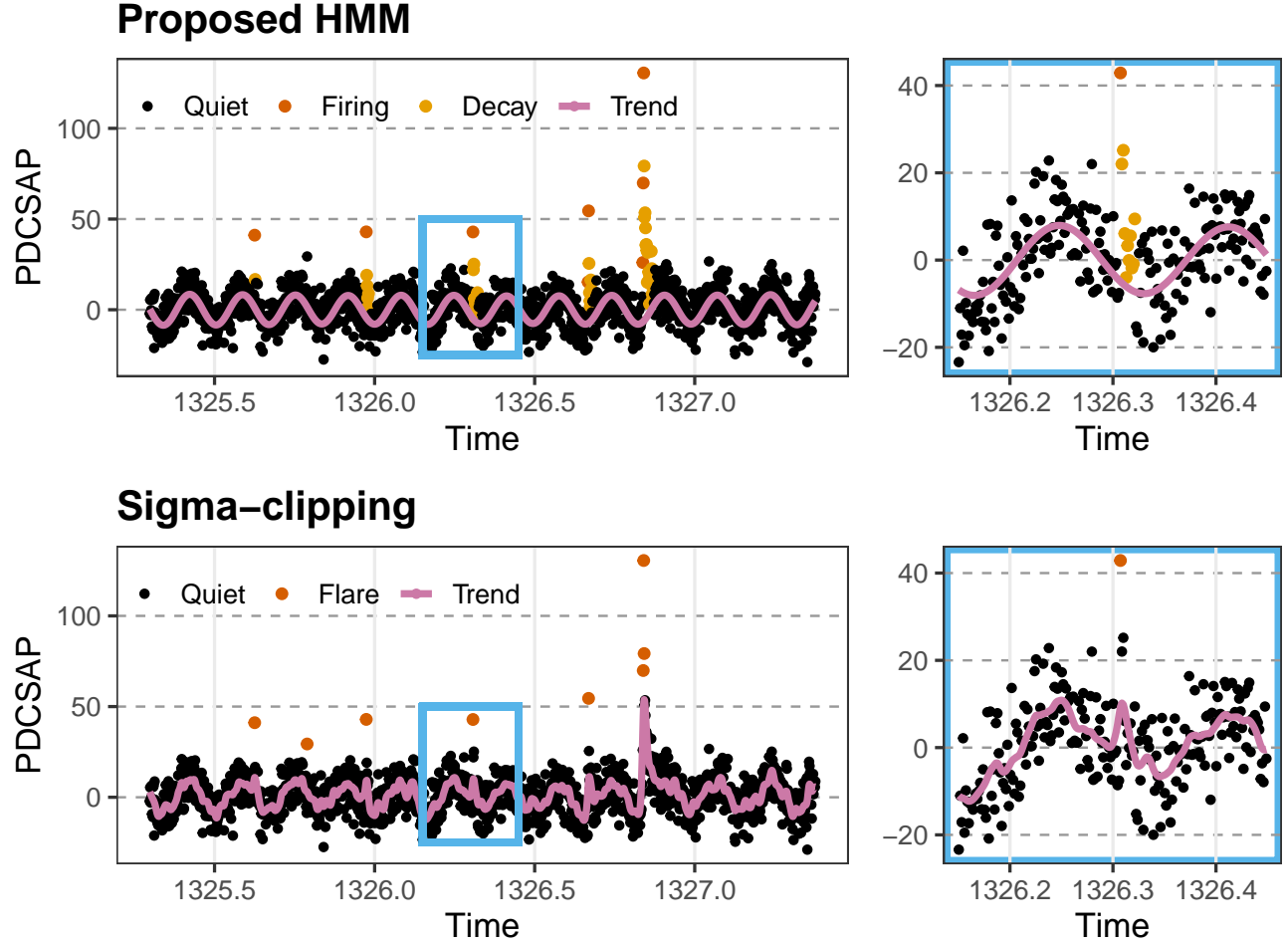
Figure 6. An example comparison of *celeriteQFD* and standard sigma-clipping for identifying the injected flares to part of the TIC 031381302 light curve. *Top row: celeriteQFD*, that simultaneously models the trend with *celerite* and assigns states to each point in the time series. *Bottom row:* the sigma-clipping approach that uses *celerite* alone to model the trend, with outliers beyond 3σ used to identify flares. The right-hand column shows a zoomed-in portion of one of the flares identified in both methods.

event. Note the similarity with many aspects of the flare shown in the right-hand side of Figure 6. The trend modeled using *celerite* only (bottom row) gets distorted, absorbing observations of higher brightness. Given that this points follow two observations of peak brightness, it is rather likely that at least some of them correspond to the decaying phase of a flare. This is precisely the type of case in which flare energies could be underestimated. Also note that using our approach, the estimated trend remained unaltered in this window and the HMM identified multiple

observations to be in decaying state (the same way as in the synthetic case).

By comparing the simulation results, where the ground truth is know, with real data results, it seems *celeriteQFD* can in fact better model the long-term trend of a light curve. Additionally, these results indicate the model is capable of directly identify light points conforming the decaying phase of a real flare, determining the duration of flaring events. Moreover, through the "decoding distributions" it directly provides a way to quantify the uncertainty about the estimated durations of the flares. e.g, in Figure 2
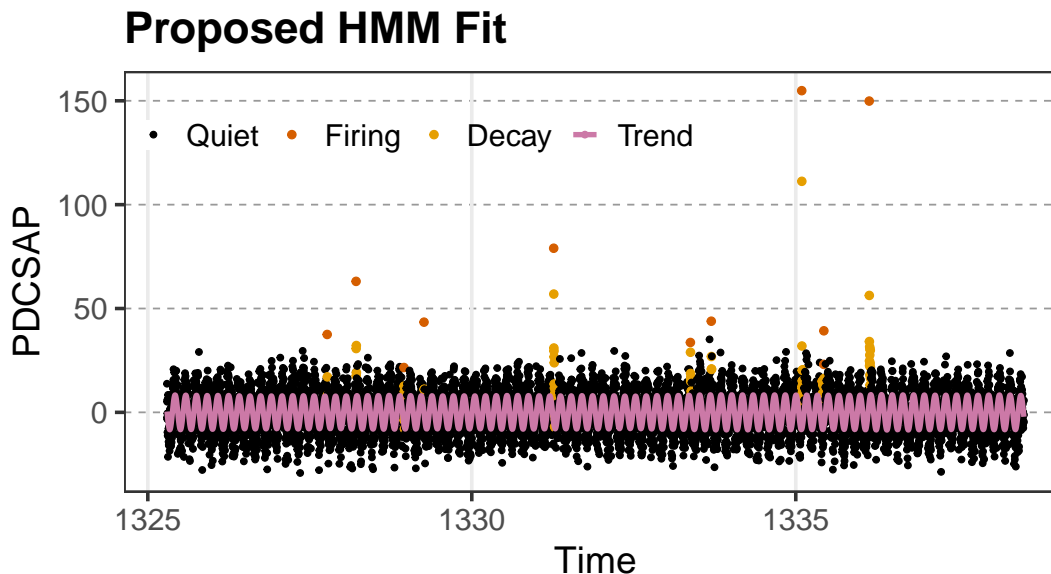
Proposed HMM Fit

Figure 7. TIC 031381302 mean centered light curve along with the fit of *celeriteQFD* that simultaneously models the trend with *celerite* and assigns states to each point in the time series.

top panel it can be seen that almost 20% of the Viterbi sequences estimated the decaying phase of the flare detected extends up until time 1331.35. Similarly, it can be seen that around 10% of them identified the flare to end 9 time steps earlier than estimated using the majority state. This uncertainty can be easily propagated into the final goal of producing energy distributions, potentially leading to more reliable and comprehensive distributions.

5. DISCUSSION

As shown in this work, an advantage of having the HMM as a flaring model is that we can identify the whole course of a flare, via states assigned to each time point. We no longer need to cross-correlate the time series with stellar flare templates, and thus remove a step in the analysis process. The HMM approach also allows us to detect compound flares more easily. Moreover, these can be used to give a probabilistic sense to the duration of the flares, making it possible to produce more comprehensive distributions of the energies of the flares.

The inclusion of a state associated with decaying and the capability to simultaneously perform

detrending and state decoding constitute a relevant benefit of our proposed method. Through our injection recovery experiment and analysis of a real star, we have shown that *celeriteQFD* can produce better and more stable estimates of the long-term trend of a light curve than *celerite* alone. The agreement between the results obtained for synthetic cases (where the ground truth is known) and real data cases indicates that biased estimation of the trend is indeed an issue that can arise when detrending is done prior to flare detection. It also provides some reassurance that our model is better equipped to handle the problem and prevent such bias. Further, we have shown that our HMM method can detect flares of lower energy that might be missed by other methods, even sensitive methods such as $1 - 3\sigma$ clipping. Thus, the HMM approach to flare detection could be well-suited for detecting flares in more "inactive" G-type stars.

It is worth mentioning that our flare recovery experiments are in no way exhaustive of flare morphologies present in stars. While it is true that stellar flare templates are more physically motivated than *celeriteQFD*, and that *celerite-*
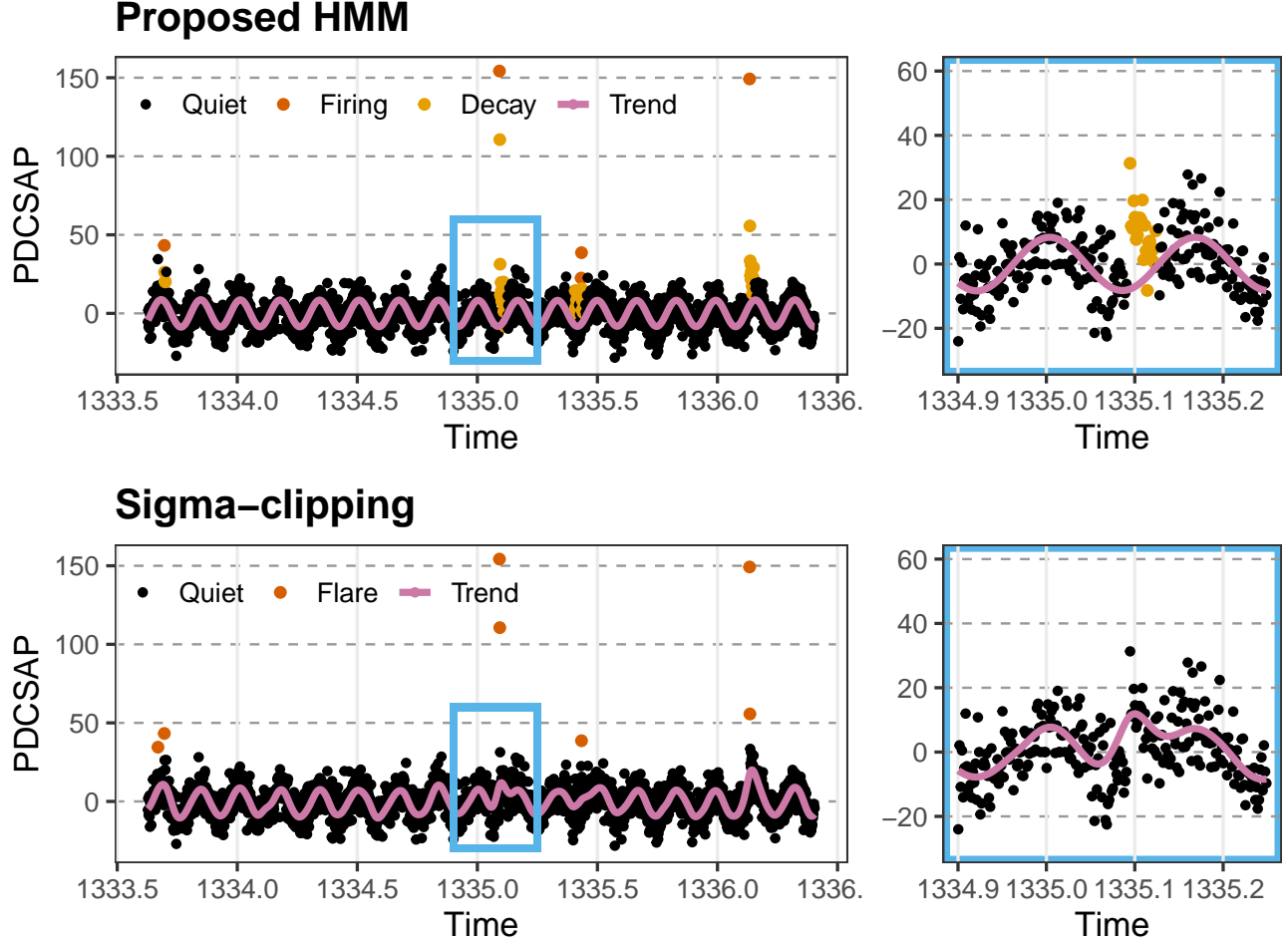
Figure 8. An example comparison of our HMM approach and standard sigma-clipping for identifying flares, using real data part of the TIC 031381302 light curve. *Top row: celeriteQFD*, that simultaneously models the trend with *celerite* and assigns states to each point in the time series. *Bottom row:* the sigma-clipping approach that uses *celerite* alone to model the trend, with outliers beyond 3σ used to identify flares. The right-hand column shows a zoomed-in portion of one of the flares identified in both methods.

QFD is not a generative model, HMMs still classify states very well even when the generative process is not specified correctly Ruiz-Suarez et al. (2022). Also, note that we do not make any strong assumptions about the nature of flare events. The few restrictions on the matter are made through the state-dependent distributions, which parameters are estimated. This provides flexibility for the model to produce different results in different contexts. We intend to explore this further, and conduct energy recovery experiments on different types of simulated flares,

assessing our method's performance and contrasting it with that of other methods.

Considering the ultimate goal of estimating flares' energies and their distributions, this method paves a new way to calculating stellar flare energy. *celeriteQFD* could prove be generally applicable and quantitatively reliable. We plan to continue this work by using our model on other light curves from different stars measured by TESS.

6. CONCLUSIONS & FUTURE WORK

In this paper, we have introduced a hidden Markov model for discovering stellar flares in

time series data of M-dwarf stars. Our approach has some notable advantages over previous approaches.

First, our method simultaneously fits a *celerite* model for the quiescent state of the star and a three-state flaring model through a hidden Markov process. With this approach, we not only obtain a better estimate of the quiescient state of the star, but also eliminate the need for sigma-clipping approaches and iterative fitting of time series data. Moreover, with the combined approach of *celerite* and the HMM, *celerite* does not absorb early or late parts of the flares. Concurrently, the HMM better identifies the whole course of the flare, and can also identify compound flares easily.

Second, through our flare injection recovery experiment, we find that our HMM method for flare detection achieves the same or better sensitivity and positive predictive value compared to sigma clipping (Figure 5).

Third, our method enables a coherent path for uncertainty quantification. Rather than providing a single most-likely flare state sequence for the time series, we obtain a posterior distribution of most-likely flare state sequences by propagating the uncertainty from our parameter estimates. This allows us to capture the variability and uncertainty on the duration of each flare.

While our approach has significant advantages, one potential disadvantage is computation time. Currently, it can take on the order of a couple of hours to run the HMM model on a single star's time series as measured by TESS, on modest resources (see Section 3.5). We are currently exploring approaches to overcome this challenge, as the advantages of our method seem to outweigh this minor (and surmountable) drawback.

Another potential criticism of our approach is that we have not developed the HMM to realistically simulate stellar time series with flares. That is, in this work, we are measuring the capacity of our proposed HMM plus the *celerite* model to identify flares well, but we are not evaluating the model's generative properties. This is something that could be improved upon and explored in the future.

Overall, this work is a promising initial step and proof-of-concept in developing a robust flare detection algorithm that does not rely on sigma-clipping or iterative approaches. There many avenues that we plan to explore in future work:

- We aim to speed up the computation time so that we can apply our HMM approach to a large sample of M-dwarf stars observed by TESS and recover stellar flares.

- In a follow-up paper, we will estimate both the flare energy distribution and flare frequency distribution (FFD) through the posterior distribution of the Viterbi state-decoded sequences. Through this approach, we will be able to propagate the uncertainties in the duration of the flare from the state sequences to the energy of the flare in a coherent way. This, combined with the improved sensitivity and PPV of our method in detecting small and large flares, should produce better estimates and increase our confidence regarding the FFD of M-dwarf stars.

- Further work may explore incorporating multivariate data (i.e., time series data across multiple bands) through a hierarchical Bayesian paradigm.

- Ultimately, it could be fruitful to design a hierarchical model that includes the FFD parameters at the population level. In this way, many M-dwarf stars could be fit with our HMM approach simultaneously, and both their individual parameters and the population-level parameters of the FFD would be modeled in a coherent way.

- After detecting stellar flares in a star, and obtaining the state-decoded sequences that quantify the uncertainty in the duration of each flare for that star, one could still use stellar flare templates to model the flares and obtain parameter estimates of interest. A study comparing this approach with standard approaches in the literature could provide further insight into the benefits (or not) of using HMMs in this framework.

- The HMM technique presented here could also be further developed for time series data measured across multiple bands, which is now becoming more commonplace (Joseph et al. 2024, e.g.,). This would allow a HMM analysis of stellar flares measured by future data sets (e.g., CubeSat Poyatos et al. 2023).

To our knowledge, this paper is one of the first applications of HMMs to an astronomy problem, and the first to do so for stellar flare detection. This statistical method has promise not only for stellar flare detection but also for other areas in astrophysics with time series data, such as gamma ray bursts, fast radio bursts, and quasars. Our hope is that this paper is useful as a starting-off point for the astronomical community to use this method in both stellar flare detection and other areas of astronomy.

AUTHOR CONTRIBUTIONS

J.A.E and Y.S. contributed equally to this project. Y.S. wrote the original code, performed most of the analysis, and wrote parts of the paper. J.A.E. subsequently modified the code, checked and performed additional analyses, wrote all code and analysis to produce all of the figures (except Figure 1), and contributed substantially to writing Sections 3, 4, and 5.

V.L.B. provided expertise on HMMs and co-supervised J.A.E. and Y.S.. V.L.B also wrote parts of Section 3.2 and made Figure 1. V.L.B. was also a co-I on the DSI grant (see Acknowledgements) that funded this project.

G.M.E. was the PI on the DSI grant that funded this work. G.M.E. also co-supervised J.A.E. and Y.S., and provided expertise on Bayesian anaylsis and astrostatistics. G.M.E. edited and contributed to all sections of the paper, and contributed substantially to Section 6.

J.S. co-supervised J.A.E. and Y.S.. J.S. contributed to the data handling and interpreation of results, and provided comments on the manuscript.

R.V.C. was a co-I on the DSI grant that funded this work. R.V.C. provided expertise on Bayesian analysis and inference, and provided comments on the manuscript.

A.M. co-supervised Y.S., and contributed substantial expertise on M-dwarfs. A.M. also helped write parts of Section 1.

J.R.A. D. provided expertise on M-dwarfs and stellar flares, provided scientific comments and feedback, and contributed to the editing of the final paper.

REFERENCES

Adam, T., Griffiths, C. A., Leos-Barajas, V., et al. 2019, Methods in Ecology and Evolution, 10, 1536, doi: https://doi.org/10.1111/2041-210X.13241

Baum, L. E., Petrie, T., Soules, G., & Weiss, N. 1970, The annals of mathematical statistics, 41, 164

Carpenter, B., Gelman, A., Hoffman, M. D., et al. 2017, Grantee Submission, 76, 1

Chang, S. W., Byun, Y. I., & Hartman, J. D. 2015, The Astrophysical Journal, 814, 35, doi: 10.1088/0004-637X/814/1/35

Davenport, J. R., Hawley, S. L., Hebb, L., et al. 2014, Astrophysical Journal, 797, doi: 10.1088/0004-637X/797/2/122

Davenport, J. R. A. 2016, The Astrophysical Journal, 829, 23, doi: 10.3847/0004-637X/829/1/23

Donati, J.-F., & Landstreet, J. 2009, Annual Review of Astronomy and Astrophysics, 47, 333

Feinstein, A. D., Montet, B. T., Ansdell, M., et al. 2020, AJ, 160, 219, doi: 10.3847/1538-3881/abac0a

Forbes, T. 1991, Geophysical & Astrophysical Fluid Dynamics, 62, 15

Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, The Astronomical Journal, 154, 220, doi: 10.3847/1538-3881/aa9332

Forney, G. D. 1973, Proceedings of the IEEE, 61, 268

Günther, M. N., Zhan, Z., Seager, S., et al. 2020, AJ, 159, 60, doi: 10.3847/1538-3881/ab5d3a

Hamilton, J. D. 1990, Journal of econometrics, 45, 39

Hawley, S. L., Davenport, J. R. A., Kowalski, A. F., et al. 2014, The Astrophysical Journal, 797, 121, doi: 10.1088/0004-637X/797/2/121

Hoffman, M. D., & Gelman, A. 2014, Journal of Machine Learning Research, 15, 1593. http://jmlr.org/papers/v15/hoffman14a.html

Ilin, E., Schmidt, S. J., Davenport, J. R., & Strassmeier, K. G. 2019, Astronomy & Astrophysics, 622, A133

Joseph, W., Stelzer, B., Magaudda, E., & Martínez, T. V. 2024, arXiv preprint arXiv:2401.17287

Leos-Barajas, V., Gangloff, E. J., Adam, T., et al. 2017, Journal of Agricultural, Biological and Environmental Statistics, 22, 232

McClintock, B. T., Langrock, R., Gimenez, O., et al. 2020, Ecol. Lett., 23, 1878

Medina, A. A., Winters, J. G., Irwin, J. M., & Charbonneau, D. 2020, The Astrophysical Journal, 905, 107, doi: 10.3847/1538-4357/abc686

Osten, R. A., Kowalski, A., Sahu, K., & Hawley, S. L. 2012, The Astrophysical Journal, 754, 4, doi: 10.1088/0004-637X/754/1/4

Ötting, M., Langrock, R., & Maruotti, A. 2021, AStA Adv. Stat. Anal., 1

Poyatos, J., Fors, O., & Cama, J. M. G. 2023, arXiv preprint arXiv:2302.12566

Rabiner, L. R. 1989, Proceedings of the IEEE, 77, 257

Ruiz-Suarez, S., Leos-Barajas, V., & Morales, J. M. 2022, JABES, 27, 339

Viterbi, A. J. 1967, IEEE Transactions on Information Theory, 13, 260

Walkowicz, L. M., Basri, G., Batalha, N., et al. 2011, The Astronomical Journal, 141, 50, doi: 10.1088/0004-6256/141/2/50

Williams, J. P., Storlie, C. B., Therneau, T. M., Jr, C. R. J., & Hannig, J. 2020, J. Amer. Statist. Assoc., 115, 16

Yang, H., Liu, J., Gao, Q., et al. 2017, The Astrophysical Journal, 849, 36, doi: 10.3847/1538-4357/aa8ea2

Zucchini, W., MacDonald, I. L., & Langrock, R. 2017, Hidden Markov models for time series: an introduction using R (CRC press)