Publications of the Astronomical Society of the Pacific, 130:104504 (12pp), 2018 October https://doi.org/10.1088/1538-3873/aada6f

NFD: Toward Real-time Mining of Short-timescale Gravitational Microlensing Events

Jiaming Qiu[1], Yankui Sun[1,3], Chao Wu[2,3], Zhihui Du[1], and Jianyan Wei[2]
[1] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, People's Republic of China; syk@mail.tsinghua.edu.cn
[2] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, People's Republic of China; cwu@bao.ac.cn
Received 2018 February 28; accepted 2018 August 14; published 2018 September 14

Abstract

To search short-timescale microlensing (ML) events ($T_E < 1$ *day*) from high-cadence, wide-field survey in real time, we present an algorithm called NFD (normalized feature deviation) to monitor all the observed light curves and to alert abnormal deviation in the data stream of light curves. The NFD algorithm framework consists of three main modules: feature stream extraction, statistical analysis of the feature stream, and discrimination. This algorithm is designed to be effective for ML event search from variable baselines, as well as constant baselines. To thoroughly evaluate the performance of our algorithm, a large-scale simulative light curve data set of variable stars based on the physical model of gravitational microlensing effect was built, yielding 139,968 valid sample light curves in total; then, a more challenging simulative data set for evaluation under GWAC-like environment was generated, including 12,960 constant sample light curves and 3240 variable sample light curves. Sample precision rate (SPR) and average detecting position (ADP) as two metrics are proposed, where SPR measures the accuracy of our algorithm, and ADP reflects alarm sensitivity. Golang implementations are done to evaluate the computational efficiency of our algorithm. Results of simulative experiments show that our algorithm is fast (176,000 stars in 3 seconds), accurate (SPR is 89% on constant baselines, and 73% on variable baselines in the second data set), and sensitive (in average, it can detect all kinds of ML event before the peak on both constant and variable baselines), making it fairly competent for the real-time mining mission of short-timescale gravitational microlensing events.

Key words: gravitational lensing: micro

Online material: color figures

1. Introduction

Gravitational microlensing (ML) can play an important role in detecting dark objects in a broad mass range from black hole to planets (Paczyński 1986). Microlensing events are very rare (Alcock et al. 2000) because the alignment required is so precise and difficult to predict. Modern observation and search technologies can monitor tens of millions of potential source stars, and search more than 2000 events every year (Hamadache et al. 2006; Udalski et al. 1994 & Udalski et al. 2015; Wyrzykowski et al. 2015).

Current ML surveys tend to take high-cadence observations to search short-timescale events, which provides a new way to detect low-mass, free-floating planets. In 2011, Sumi et al. (2011) first reported the evidence of free-floating planetary mass objects is related to short-timescale microlensing events ($T_E < 2$ day). Mroz et al. (2017b) discovered a shorter timescale Neptune-mass, free-floating planet candidate. Some observations also show the existence of Earth-mass and super-Earth-mass free-floating planets (Mroz et al. 2017a) in short-timescale microlensing events. That is, the characteristic timescale of microlensing events depends on the mass of the lens (Mroz et al. 2017a). To search short-timescale ML, a wide field of view (FOV) and a high cadence are required for the instrument and observation strategy.

The ground-based, wide-angle camera array (GWAC) is a part of the follow-up telescope network of the SVOM mission (see Wei et al. 2016), which is comprised of 36 cameras, each with an 18 cm diameter optical refractor telescope and 4k × 4k e2v back illuminated CCD detector. Every pixel corresponds to a 11.17 scale, thus the whole pixel mosaic gives a FOV of 12.7 × 12.7 degrees. The total FOV will cover ∼5000 deg^2, which allows the GWAC to monitor millions of stars brighter than 16.0 mag (V band, GWAC with unfiltered exposure) at the same time. The primary scientific goal of GWAC is to observe prompt, optical emissions of gamma-ray bursts (GRBs) by monitoring the sky area of ECLAIRs (A GRB detector of the SVOM mission payload; see Schanne 2008). According to the observation strategy of SVOM, each camera of GWAC will monitor 1 ∼ 2 fields with a 15 second (exposure time = 10 sec, readout time = 5 sec) cadence for one observation

[3] Author to whom any correspondence should be addressed.

night. The set of 36 cameras of the GWAC is expected to be completed at the end of 2019, and will be mounted provisionally at Xinglong observatory.[4] To improve its efficiency from sky background brightness, it is planning to ship to other ideal observatories of Xinjiang or Tibet in China. Besides the primary scientific goal observation of the prompt emission, GWAC will also be a good instrument in searching ML of timescales less than one day, benefitting from its high cadence and wide field of view.

We propose to search short-timescale ML from GWAC through a huge volume of light curve data in real time. It can trigger follow-up observations to identify the short-timescale events after they are discovered. It is required to solve two problems: (i) the process and management of a large volume of light curve data in real time, and (ii) to search the short-timescale ML candidates in real time. For problem (i), Wan et al. (2016) did pathfinding work on real-time processing and managing of a large volume of GWAC light curves by using the MonetDB platform, which we are pushing ahead through engineering done by our collaboration computer team. This paper will focus on solving short-timescale ML events by searching millions of light curves with the baselines of constant or variability.

Udalski et al. (1994) first developed real-time search system, called Early Warning System (EWS), which has succeeded in operation in OGLE project for more than 20 years (Udalski 2003; Udalski et al. 2015). For search from the variable baselines, Wyrzykowski et al. (2006) have first succeeded in systematic search for ML events from the variable baselines in OGLE-III data. Recently, an algorithm called D-ARIMA that dynamically adjusts the parameters of ARIMA for real-time anomaly detection on light curves was proposed (Feng et al. 2017), but their experiments are conducted on limited number of light curves, without speed test under GWAC-like hardware environments. We also find previous work about real-time anomaly detection of timeseries. In 2016, Ahmad and Purdy proposed doing real-time anomaly detection using Hierarchical Temporal Memory (HTM; Ahmad & Purdy 2016), but this algorithm is too time consuming and memory consuming for large-scale, real-time data, and it is not tested on light curves.

In this paper, we investigate the problem of a real-time search for short-timescale gravitational ML events from a huge number of light curves, and propose an algorithm framework, called normalized feature deviation (NFD), which is motivated by the idea of HTM algorithm. Thorough experiments are conducted on our generated large-scale simulative data sets to evaluate the performance of NFD.

The organization of this paper is as follows. In Section 2, we present insight and detailed descriptions of the NFD algorithm framework. Section 3 introduces the large-scale simulative data sets that we created for evaluation, and shows experimental results. Section 4 is the conclusion of our work.

[4] http://www.xinglong-naoc.org/html/en/

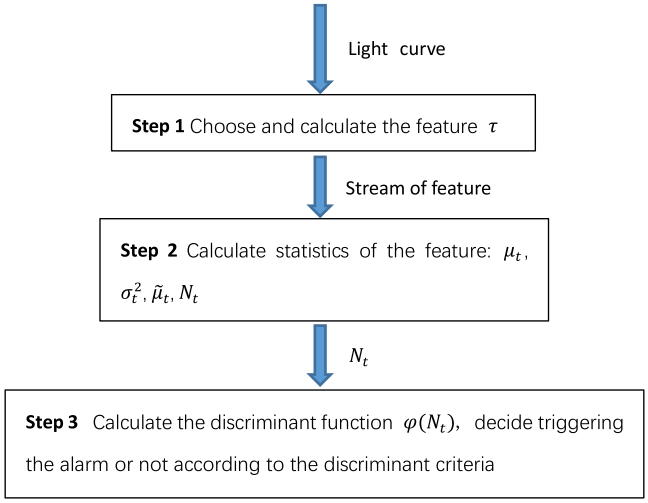

Figure 1. Pipeline of NFD algorithm framework.
(A color version of this figure is available in the online journal.)

2. NFD Algorithm Framework

The NFD algorithm framework, as its name indicates, focuses on the deviation of a specific pre-chosen feature of the light curve, which is normalized. We measure this deviation by calculating the difference between the set of long-term features covering the past state of the light curve and the set of short-term features representing the current state of the light curve. If the deviation is too large, we think something anomalous is happening (e.g., gravitational microlensing effect). The idea of the NFD has been inspired by previous works (Ahmad & Purdy 2016), but we are making a few more efforts to generalize it so that we get an algorithm framework comprised of several components: features, discriminant functions, and discriminant criteria. Each of these components can be changed to meet various time/space constraints, so that the algorithm is highly modularized.

The flowchart of NFD is shown in Figure 1. It consists of the following three steps:

Step 1. Choosing an appropriate feature of the light curve, and calculate it if needed.

The feature τ is the basic calculation unit of our algorithm framework, and there will be a value of it at each sample time, the same as the brightness. For simplicity and efficiency, we can use brightness itself as the feature, however, statistics of the light curve, like the derivative of the brightness with regard to the time, or outputs from complicated sequential models (e.g., recurrent neural network, long short-term memory (Hochreiter & Schmidhuber 1997) and hierarchical temporal memory (Ahmad & Purdy 2016)) can also be used as features. These

kinds of features might be more powerful for measuring deviations, but they will bring extra computations.

Step 2. Calculating the statistics of feature.

We define the history window as the most recent L feature values, and the decision window as the most recent S feature values of the light curve, where L and S are hand-picked parameters whose values should be related to the timescale of target anomalies, as well as to the sampling interval, while also satisfying $L \gg S$. For example, in our experiments where the sampling interval is 15 seconds, we set the value of L to be from 8000 to 16,000, spanning from 1 day to 3 days, and we set the value of S to be from 15 to 45, spanning from about 4 minutes to 10 minutes.

In this step, we calculate the mean, μ_t, and the variance, σ_t^2, of the features in the history window, as well as the mean of the features $\tilde{\mu}_t$ in the decision window. Therefore, the normalized variable N_t can be calculated, which represents the normalized observed value of the random variable $\tilde{\mu}_t$ from the distribution $N(\mu_t, \sigma_t^2)$:

$$\mu_t = \frac{\sum_{i=0}^{L-1} \tau_{t-i}}{L}, \tag{1}$$

$$\sigma_t^2 = \frac{\sum_{i=0}^{L-1} (\tau_{t-i} - \mu_t)^2}{L-1}, \tag{2}$$

$$\tilde{\mu}_t = \frac{\sum_{i=0}^{S-1} \tau_{t-i}}{S}, \tag{3}$$

$$N_t = \frac{\tilde{\mu}_t - \mu_t}{\sigma_t}. \tag{4}$$

Step 3. Making a decision according to the discriminant function and discriminant criteria.

Calculate the discriminant function, $\varphi(N_t)$, and then use pre-chosen discriminant criteria to decide whether to trigger the alarm signal for anomaly event. In this paper, we use the Q-function as a discriminant function to measure the deviation of normalized variable. Q-function is the tail probability of the standard normal distribution, defined as

$$Q(x) = \frac{1}{2} erf\left(\frac{x}{\sqrt{2}}\right) = \frac{1}{2} erfc\left(\frac{x}{\sqrt{2}}\right), \tag{5}$$

where

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt. \tag{6}$$

In other words, $Q(x)$ measures the probability that the observed value of a normal Gaussian variable is larger than the mean of it by x standard variations. To decide whether to trigger the alarm, we set a probability threshold, ε, and compare $Q(x)$ to ε unilaterally or bilaterally, depending to the type of feature. If we use the predict error or the anomaly score produced by sequential predictive models such as LSTM and HTM, the discriminant criteria should be unilateral (i.e., triggering the alarm when $Q(x) < \varepsilon$); however, if original brightness is used as feature, the discriminant criteria should be bilateral (i.e., triggering alarm when $Q(x) < \varepsilon$ or $Q(x) > 1 - \varepsilon$), as it should be considered anomalous when the value of brightness is too large or too small.

The proposed NFD algorithm framework has several advantages. First, it is simple and fast, especially when the original value of the timeseries is used as its feature. Well-known prediction-based timeseries models like ARIMA (Box & Pierce 1970), random forests (Breiman 2001), and logistic regression either need explicit modeling of timeseries under unrealistic hypotheses or complicated training or parameter-estimating process, while the NFD algorithm does not need any "training" step or does not rely on carefully chosen model parameters; according to the large-scale stress test in Section 3.5.2, NFD can simultaneously process 176,000 light curves while the sampling interval is 3 seconds. Second, it is robust to occasional noise point effect. As it utilizes a time window that usually lasts from minutes to hours, occasional noise points will be smoothed, avoiding false triggers. Last but not least, other algorithms can be easy embedded into the NFD algorithm if we consider their outputs as features.

3. Experiments

In this section, we conduct thorough experiments to evaluate the performance of our algorithm. Our experiments are mainly conducted on a Windows 10 machine with Intel Core i7-7700k 4.20 GHz CPU and 32 GB RAM, and the experiment codes are written in Python. We use the original brightness as features for NFD in all the experiments for simplicity, but it can easily be extended to use other features.

3.1. Simulated Light Curve Data Set

As in formula (2) in Wyrzykowski et al. (2006), the microlensing model can be expressed as

$$f(t) = f_{s0} * A(t) + f_b, \tag{7}$$

where f_{s0} is the flux of baseline. It is denoted as $f_s + f_{var}$ in Wyrzykowski et al. (2006). According to the observation strategy of GWAC, we can ignore the impact of blending stars in our data set, as GWAC will avoid a crowded field. Hence, f_b is assumed as 0 in our paper. After considering noise affect, the above formula can be expressed in a magnitude unit following our simulation:

$$\begin{aligned} I_{obs}(t) &= I_{s0} + I_A + N(\sigma^2) \\ &= -2.5 \log_{10}(f_{s0} * A(t)) + N(\sigma^2), \end{aligned} \tag{8}$$

where $I_{obs}(t)$ denotes the simulated observation brightness of ML event, I_{s0} is the baseline brightness, I_A is the magnification of the microlensing, and $N(\sigma^2)$ is noise. The unit of all the

terms is in magnitude. To simulate the value of three terms, we will use following three steps.

Step 1. Microlensing magnification I_A.

We employ the simplest case of single-point mass lens given in the standard form (Paczyński 1986):

$$A(t) = \frac{[u(t)]^2 + 2}{u(t)\sqrt{[u(t)]^2 + 4}} \tag{9}$$

$$u(t) = u_0 + \left| \frac{t - t_0}{T_E} \right|, \tag{10}$$

where u_0 is the minimum impact parameter in units of the Einstein radius at the moment of crossing time of Einstein radius at t_0, and t_E. Combined with formula (8), we can simulate various microlensing magnification with basic parameters. We use I_{A0} to denote $I_A(t_0)$, which means maximum magnification.

To evaluate the performance of detection time position of our algorithm, we need define the microlensing start and end times. Based on our experience and data set property, we define the beginning of the event as the first data point y_i in the generated gravitational microlensing signal that satisfies $|I_{A\_i} - I_{A\_i+1}| \geqslant \delta$ mag , and the end of the event as the last point y_j that satisfies $|I_{A\_j} - I_{A\_j+1}| \geqslant \delta$ mag, while δ represents a small mag value used to distinguish the core component of the signal from the long flat tail (we use $\delta = 0.01$ in our experiments, while other values might also be applicable). We use time T' to denote the duration from start time to end time defined above.

Step 2. Baseline I_{s0} simulation

The baseline brightness is simulated with the model

$$I_{s0}(T, m_A, \varphi, t, I_0) = m_A \sin\left(\frac{2\pi t}{T} + \varphi\right) + I_0 \tag{11}$$

Where I_{s0} consists of two components: variable brightness, and constant brightness. The variable brightness is used with simple sinusoidal shape. m_A is amplitude of the variable brightness, T is period, and φ is phase. The constant part has no effect to our algorithm performance, except for its related noise in practice. Therefore, we assume $I_0 = 0$ in our paper.

To meet the convention term of signal processing, source brightness lensed part in light curve (time length is T') will be called the signal part, while the other part is defined by the background (no lensing effect) in our paper.

In our simulation, we will generate a background light curve with a time span of $3 + T'$ days as shown in Figure 2.

Step 3. Combining microlensing magnification (I_A) and baseline (I_{s0}) with observation noise ($N(\sigma^2)$).

We first combine the microlensing magnification (I_A) and baseline (I_{s0}) by three steps: (1) zero-pad the microlensing magnification at the right to expand its duration from T' days to $3 + T'$ days; (2) element-wisely add I_A and I_{s0}; and (3) horizontally reverse the combined signal. Visualized

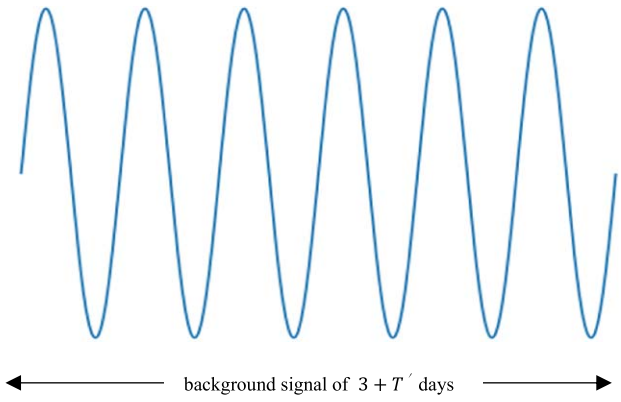

Figure 2. Background signal I_{s0}.
(A color version of this figure is available in the online journal.)

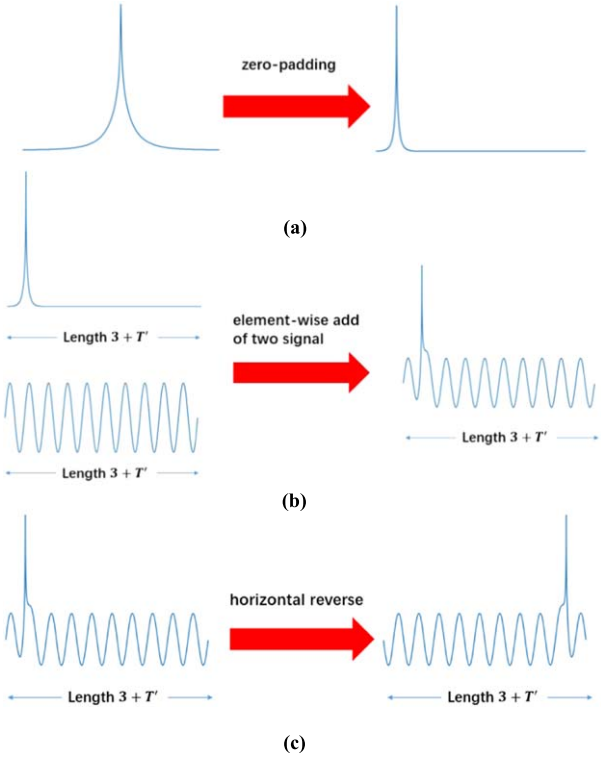

Figure 3. Process of combining two signals.
(A color version of this figure is available in the online journal.)

combining procedure is shown in Figure 3. The reason that we zero-pad the microlensing signal at the right of it (instead of at the left of it) and do extra reversing step after addition is that while the phase at the left-end of the background signal is decided given φ, the phase at the right-end can still change while T' changes. By the pipeline we design, we assure that the microlensing signal aligns to the phase of the sine curve which

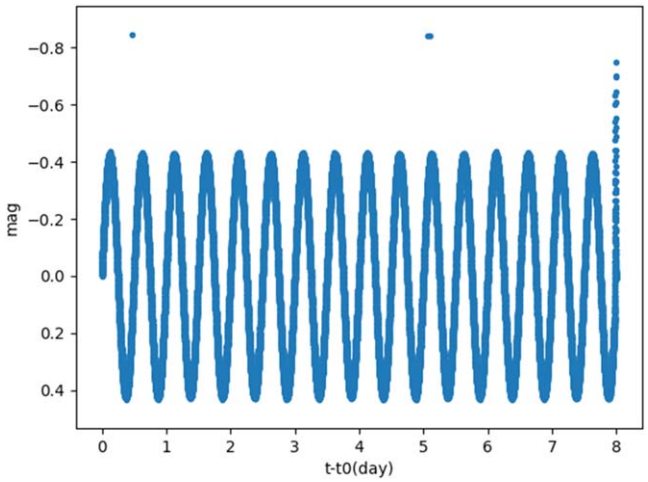

Figure 4. Sample light curve with $\sigma = 0.005$, $T' = \frac{1}{24}day$, $T = 0.5$day, $I_{A0} = 0.84$, $\frac{I_{A0}}{m_A} = 2$, $\varphi = 0$.
(A color version of this figure is available in the online journal.)

is decided by sampling from a range of given values in step 2, regardless of the value of T'.

After combining I_A and I_{s0}, we add observation noise $N(\sigma^2)$ on the combined signal, yielding the simulated light curve,

$$I_{obs}(t) = I_{s0} + I_A + N(\sigma^2). \tag{12}$$

As the noise property is unknown, in this paper we assume some values of it in all the experiments based on our experience. In addition to the basic noise, we randomly inject outlier points, whose magnitudes are $I_{obs}(t_0)$, into the combined signal to test the robustness of the algorithm to occasional outlier points.

As in Equation (12), a sample light curve is determined by the following seven parameters: T, m_A, φ, t_0, t_E, u_0, σ^2. For the purpose of evaluation, we define $\frac{I_{A0}}{m_A}$ as the relative saliency of the gravitational microlensing signal with regard to the background signal. As I_{A0} is determined by u_0, adjusting the value of u_0 and $\frac{I_{A0}}{m_A}$ is equivalent to adjusting the value of u_0 and m_A. Noticing that t_0 is determined by T', and T' is determined by t_E, we only need to directly determine the value of the following six parameters: u_0, T', $\frac{I_{A0}}{m_A}$, T, φ, σ. A sample light curve with $\sigma = 0.005$, $T' = \frac{1}{24}day$, $T = 0.5day$, $I_{A0} = 0.84$, $\frac{I_{A0}}{m_A} = 2$, $\varphi = 0$ is shown in Figure 4.

3.2. Evaluation Metrics

To simulate the real-time processing environment, the algorithm will be made to process a sample light curve point by point according to the time order. We set the starting point of evaluation at a unified proportion of the full length of each sample. For example, if we set the starting proportion to be 30%, then for each sample light curve, the algorithm will only need to give its result for the last 70% points of it, and the first 30% data points are provided as historical data for training or other usages. For all of the experiments in this paper, we set the starting proportion to be 50%. Assuming that a gravitational microlensing event happens at the time range $[t_1, t_2]$, we define that if the algorithm triggers an alert at moment, t', where $t_1 \leqslant t' \leqslant t_2$, then this trigger is a true-positive trigger, otherwise it is a false-positive trigger. If there is a true-positive trigger in a sample light curve, then we stop further testing on it because we have successfully mined the only event in it, so there might be multiple false-positive trigger, but at most one true-positive trigger will happen on a light curve. We choose to measure the performance of our algorithm using sample as the basic granularity, and we propose two evaluation metrics. The first metric, which we call sample precision rate (SPR), is defined as follows:

$$SPR = \frac{n_{correct}}{n_{all}}, \tag{13}$$

where $n_{correct}$ represents the number of samples on which the algorithm gives no false trigger and exactly one true trigger (we call these samples correct samples), and n_{all} represents the total number of samples.

The second metric, which we call average detecting position (ADP), is only evaluated on correct samples. It is defined as follows:

$$ADP = \frac{1}{n}\sum_{k=1}^{n} \frac{t_{(k)} - t_{0(k)}}{T'_{(k)}}, \tag{14}$$

where n represents the number of correct samples, and $t_{(k)}$ represents the time of the first true trigger on sample k, while $t_{0(k)}$ and $T'_{(k)}$ represent t_0 and T' of the injected ML signal of sample k respectively. Because t_0 locates at the center of the duration of ML signal, the average detection position ranges from -50% to 50%.

3.3. Selection of Alarming Threshold

As mentioned in Section 2, we need to decide the value of the probability threshold ε for alarm triggering, which is somehow sensitive to the pattern of data. To do so, we construct a data set on which we conduct experiments to find out best value of ε. This data set is denoted as General-Mini in Table 1, with 11,664 light curves in total, and we further filter out samples with unreasonably large background signal magnitude, i.e., $m_A > 1.5$, yielding 5616 valid samples in total. We set $L = 8000$ (because GWAC samples data every 15 seconds, a window of 8000 data points spans about two days) and $S = 15$ (We find the relations between the metrics and ε are similar for different value of S, however the overall results are best when $S = 15$). Table 2 shows the results of the experiment.

Table 1
Parameters of Data Sets in Section 3.3 and Section 3.4

Data Set	u_0	T'(days)	σ	T(days)	$\frac{I_{A0}}{m_A}$	φ
General-Mini	[0.01, 0.1, 0.9]	$\left[\frac{1}{24}, \frac{3}{24}, ..., \frac{21}{24}, \frac{23}{24}\right]$	[0.005, 0.01, 0.05, 0.1]	[0.2, 0.3, ..., 0.9, 0.1]	$\left[\frac{1}{5.0}, \frac{1}{4.0}, \frac{1}{3.0}, \frac{1}{2.0}, 1.0, 2.0, 3.0, 4.0, 5.0\right]$	0
General-Large	[0.001, 0.251, 0.501, 0.750, 1]	$\left[\frac{1}{24}, \frac{3}{24}, ..., \frac{21}{24}, \frac{23}{24}\right]$	[0.005, 0.171, 0.337,0.503,0.668,0.834, 1, 1.5, 2]	[0.2, 0.3, ..., 0.9, 0.1]	$\left[\frac{1}{5.0}, \frac{1}{4.0}, \frac{1}{3.0}, \frac{1}{2.0}, 1.0, 2.0, 3.0, 4.0, 5.0\right]$	$\left[0, \frac{1\pi}{3}, \frac{2\pi}{3}, \frac{3\pi}{3}, \frac{4\pi}{3}, \frac{5\pi}{3}\right]$

Table 2
Evaluation Results with Different ε on General-Mini

ε	SPR (%)	ADP (%)
0.0425	69.2	−6.9
0.0375	**69.6**	**−6.2**
0.0325	69.6	−5.6
0.0275	69.0	−5.1
0.0225	68.4	−4.6
0.0175	66.8	−4.2

Note. The bolded numbers indicate the best value.

From Table 2, we can conclude that, using SPR as the most important metric, the algorithm has best performance when $\varepsilon = 0.0375$.

3.4. Experimental Results on Large-scale Simulative Data Set

In this section, we construct a larger data set and conduct experiments on this data set to explore the relations between evaluation metrics and data characteristics. This data set is denoted as General-Large in Table 1, yielding 262,440 light curves in total, which is much more than that of General-Mini, and we consider different values of φ in this data set. Also, we filter out samples whose m_A is larger than 1.5, yielding 139,968 valid samples in total. We let $L = 8000$ and $\varepsilon = \varepsilon_{best} = 0.0375$, as done before. However, in this section we set $S = 45$ when applying NFD because we find that for very large data sets, larger value of S leads to more stable results. We show the overall results and relations between two metrics and four most important parameters (T', σ, $\frac{I_{A0}}{m_A}$ and φ) in the following pages.

• **SPR**

On General-Large, SPR of our algorithm is 67.2%. The relations between SPR and four chosen parameters are shown in Figure 5. Specifically, in Figure 5(a), the relations between SPR and $\ln \frac{I_{A0}}{m_A}$ as well as σ are shown simultaneously. In Figures 5(b) and (c), SPR under different values of φ and T' are shown.

As we can see from Figure 5(a), SPR is getting higher with $\frac{I_{A0}}{m_A}$ getting larger and σ getting smaller, which is consistent to our expectation and common sense since the more salient it is, and the less noisy it is, the easier it can be detected. SPR is higher than 0.9 in average when $\ln \frac{I_{A0}}{m_A}$ is not less than 1, which in our data set means that $\frac{I_{A0}}{m_A}$ is not less than 3 (see Figure 6), and σ is less than 0.15, but it drops down to ∼0.6 when $\frac{I_{A0}}{m_A}$ equals 2, and ∼0.4 when $\frac{I_{A0}}{m_A}$ equals 1. In Figure 5(b), we find that SPR achieves its best when $\varphi = \pi$ (see Figure 7), because when $\varphi = \pi$, the gravitational microlensing signal is overlaid on the peak of the background signal, making it more outstanding. In Figure 5(c), we find that SPR is getting higher with T' getting longer, which is similar to that of Figure 5(a), however, an interesting phenomenon in Figure 5(c) is that the points in it are separated into five layers. After empirical analysis, we conjecture that this phenomenon is caused by the "cutting" of gravitational microlensing signals in the first step of sample generating process. Because there are five different values of u_0 for the gravitational microlensing signal, the ascending speed of them is different, causing different proportion of them are cut. Since the value of u_0 determines the remaining length of the signal, the number of layers in Figure 5(c) equals the number of values of u_0 in our data set which is five.

• **ADP**

On General-Large, ADP of our algorithm is −7.0%. From Figure 8 we see the relations between ADP and four chosen parameters. The results are presented in the same way as that of SPR. As the figures show, the relations between the ADP and the four parameters are similar to that of SPR, and for most values of the four parameters, the average detection position of our algorithm is before the peak of the gravitational microlensing, which shows the sensitivity of our algorithm. However, we can see that in Figure 8(c), the points are separated into approximately two layers which is similar to that in Figure 5(c) with only the difference of number of layers. We empirically find that the changing of the ADP with regard to u_0 is far less sensitive than that of SPR, so the layer-separating effect is less obvious in Figure 8(c).

3.5. Experiments in GWAC-like Simulative Environment

In this section, we focus on evaluating our algorithm in simulative environment that is close to the actual runtime

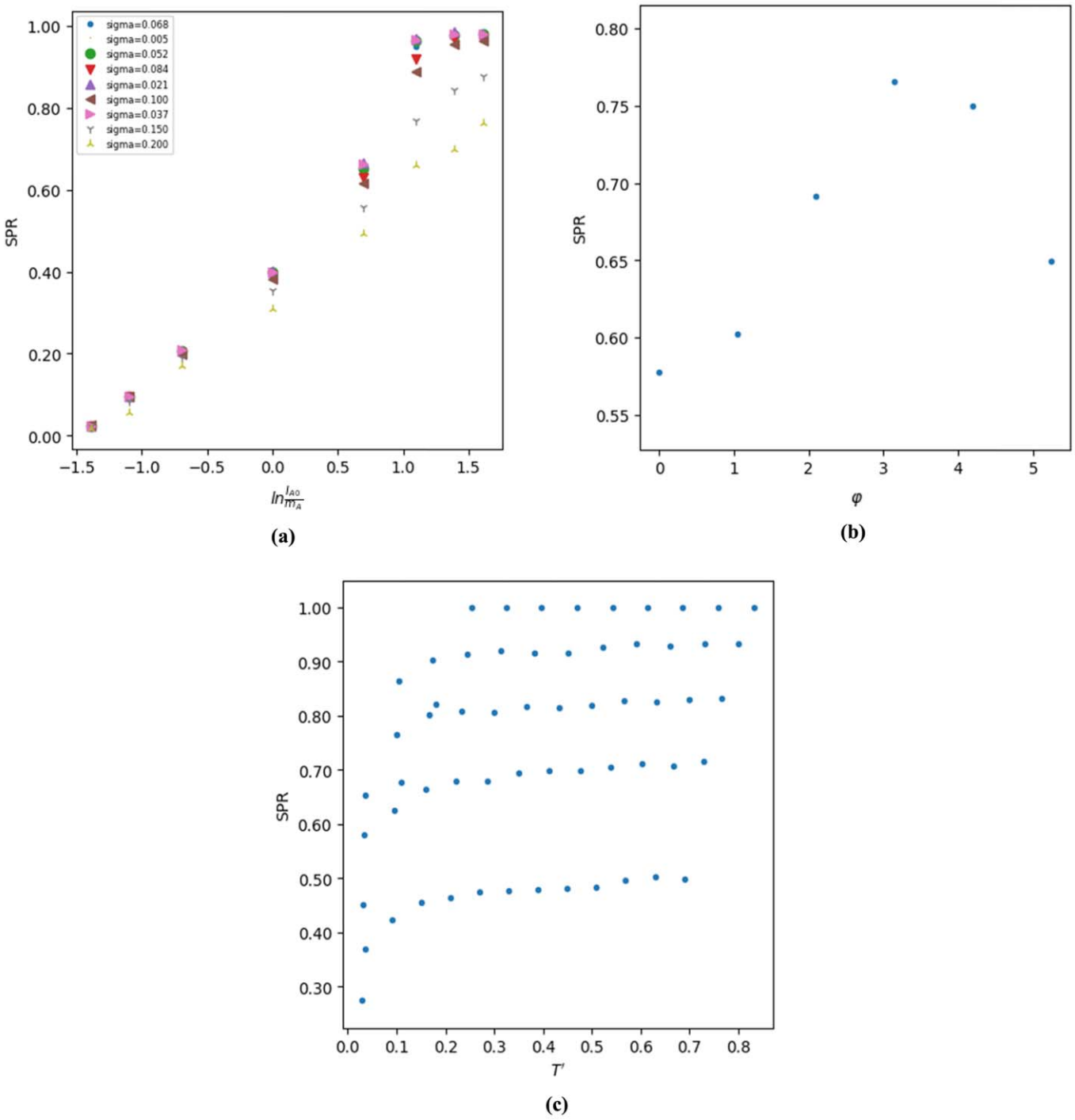

Figure 5. Relations between SPR and four parameters on General-Large.
(A color version of this figure is available in the online journal.)

environment of GWAC. The evaluation consists of two parts: the first part includes experiments on elaborately simulated data sets, including constant and variable stars, which are based on the observation parameters of GWAC for accuracy and sensitivity evaluation, and large-scale stress tests using random data generator for speed and robustness evaluation.

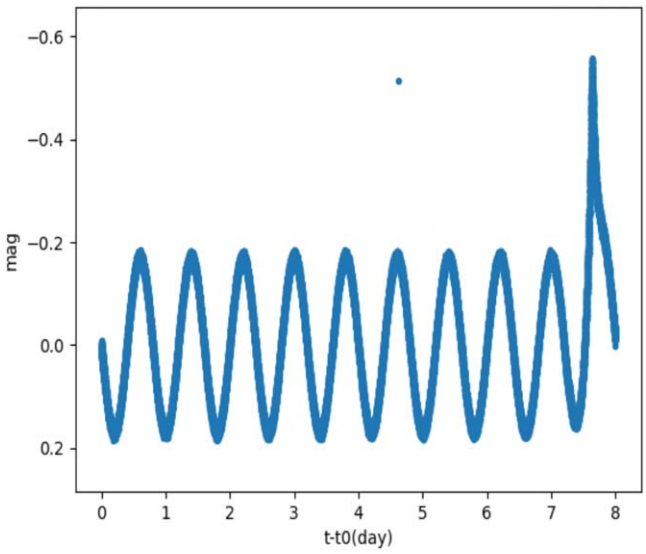

Figure 6. Sample light curves with $\frac{I_{A0}}{m_A} = 3$.
(A color version of this figure is available in the online journal.)

3.5.1. Accuracy and Sensitivity Evaluation

To evaluate the accuracy and sensitivity of our algorithm for GWAC observation missions using simulative data sets, the observation parameters of GWAC and the observation process itself must be carefully considered. The median noise level σ of GWAC is 0.06, so we consider a reasonable range around it while injecting noise into the generated light curves. Another important characteristic of GWAC observation routine is that the observation is not held for all day long. Normally, GWAC will be launched at night only, and the exact duration of observation might differ with regard to external conditions like weather, which means that the light curves that we have finally attained are discontinuous timeseries. To evaluate the performance of our algorithm under this scenario, we build another relatively smaller simulative data set. This data set includes two subsets: constant baselines with 12,960 light curves in total and variable baselines with 3240 light curves in total. These two subsets are denoted as GWAC-Con and GWAC-Var in Table 3, with the difference comparing to Table 1 that we fix m_A at specific values for constant baseline and variable baseline individually instead of controlling $\frac{I_{A0}}{m_A}$, in order to match the magnitude of baselines under GWAC-like scenario.

In GWAC-Var, for each parameter composition, we first generate *continuous* samples according to the setting in Section 3.1. Then, we generate the corresponding *discontinuous* samples by segmenting the original sample in parts that are about eight hours long and randomly change the phase of each part (some generated samples are shown in Figures 9 and 10). We also lengthen these samples to 192 hours

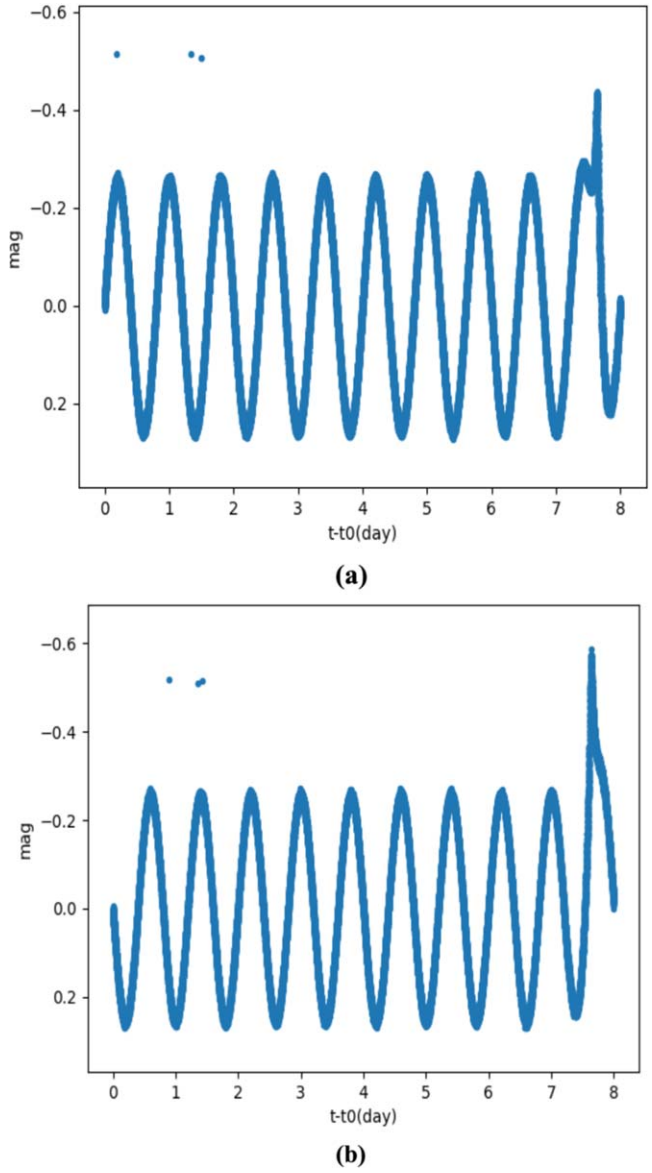

Figure 7. Sample light curves with $\varphi = 0$ (Figure (a)) and $\varphi = \pi$ (Figure (b)).
(A color version of this figure is available in the online journal.)

(eight days for continuous samples and 24 days for discontinuous samples, including the duration of gravitational microlensing signal) because we conjecture in advance that longer history window might be needed due to the discontinuity. We apply approximate optimal parameter search for history window length, decision window length, and trigger threshold for best performance, and yield approximately optimal result with $L = 16000$, $S = 45$ and $\varepsilon = 0.0075$.[5]

[5] Accurate grid search can be applied to gain best parameter settings, but it is extremely computationally costly, and is not necessary for our research. Therefore, we choose to apply approximate search in this paper.

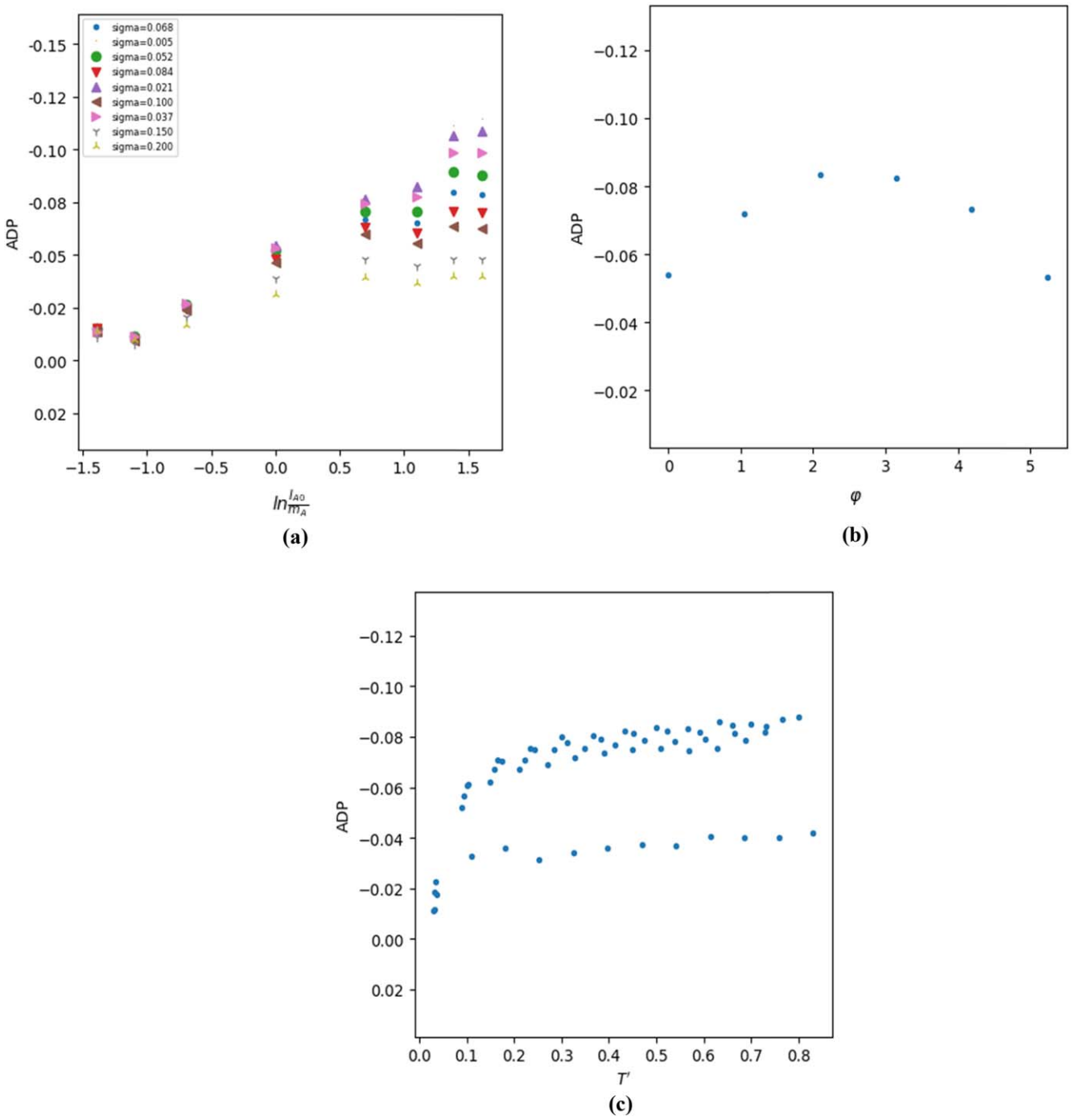

Figure 8. Relations between ADP and four parameters on General-Large (b). (A color version of this figure is available in the online journal.)

Results of our algorithm on GWAC-Con and GWAC-Var are shown in Tables 4–Table 7.

According to Table 4, the SPR of our algorithm on GWAC-Con is 89.0%, and the ADP is −7.2%, which means that under the constrain of absolutely no false positive, our algorithm can recall 89% of all ML signals, and we can tell the existence of ML signal before the peak by 7.2% of the total lasting time of the ML signal in average. SPR and ADP under different I_{A0} are shown in Table 5 (note that we show the relationship between I_{A0} and the metrics instead of $\frac{I_{A0}}{m_A}$ and the metrics since $m_A = 0$),

Table 3
Parameters of Data Sets in Section 3.5

Data Set	u_0	T′(days)	σ	T(days)	m_A	φ
GWAC-Con	[0.001, 0.251, 0.501, 0.750, 1]	$\left[\frac{1}{24}, \frac{3}{24}, ..., \frac{21}{24}, \frac{23}{24}\right]$	[0.01, 0.06, 0.1, 0.15]	[0.2, 0.3, ..., 0.9, 0.1]	0	$\left[0, \frac{1\pi}{3}, \frac{2\pi}{3}, \frac{3\pi}{3}, \frac{4\pi}{3}, \frac{5\pi}{3}\right]$
GWAC-Var	[0.001, 0.251, 0.501, 0.750, 1]	$\left[\frac{1}{24}, \frac{3}{24}, ..., \frac{21}{24}, \frac{23}{24}\right]$	0.06	[0.2, 0.3, ..., 0.9, 0.1]	0.25	$\left[0, \frac{1\pi}{3}, \frac{2\pi}{3}, \frac{3\pi}{3}, \frac{4\pi}{3}, \frac{5\pi}{3}\right]$

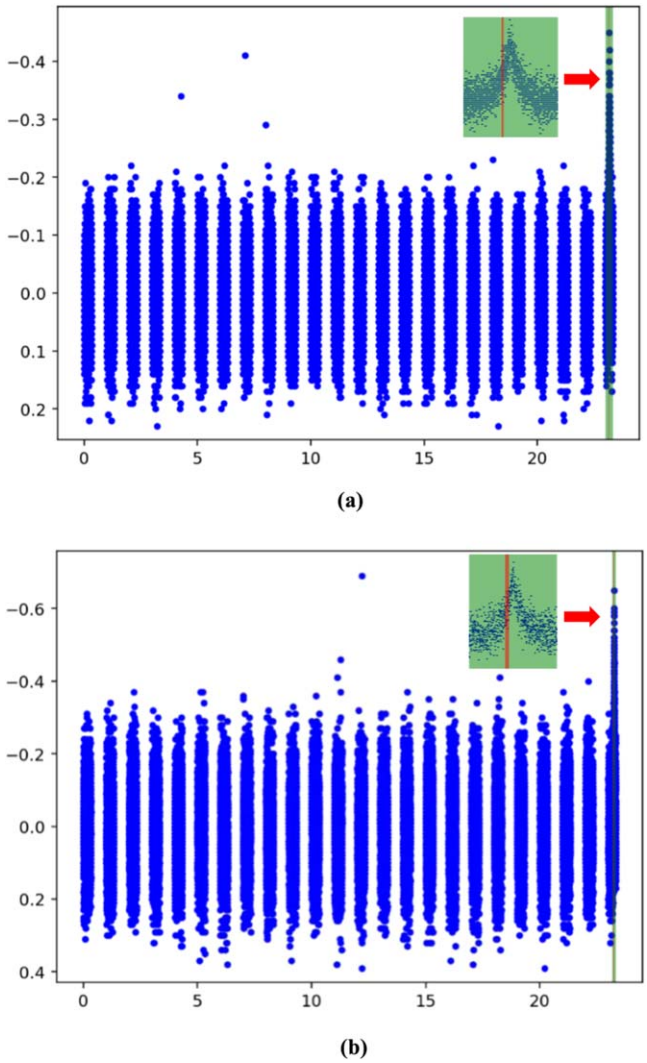

Figure 9. Some generated constant baselines and visualized results on them. The green rectangle area means the ground truth gravitational event section, while the red area means the triggering section given by our algorithm. Panels (a) and (b) are both correct samples. The σ in panel (a) is 0.06, which is the typical noise level of GWAC, while σ is a little bit larger in panel (b), with the value of 0.1.

(A color version of this figure is available in the online journal.)

and some visualized results are shown in Figure 9. From Figure 9 we can see that our algorithm did not trigger false alarms on manually injected outliers under different noise levels ($\sigma = 0.06$ in Figure 9(a), and 0.1 in Figure 9(b)), proving its robustness. It might be noticed that when the amplitude of ML event is extremely large, SPR might suffer decline, and the reason of this phenomenon is that those outlier points with extremely large value may have affect the statistics, though they might not cause false trigger at once. However, outlier points with extremely large values are rare cases in the practice; besides, the accuracy of our algorithm is still higher than 80%, which is acceptable in practice under the constrain of zero false positive.

According to Table 6, SPR of our algorithm is 73.5% on continuous samples and 73.2% on discontinuous samples on GWAC-Var, and ADP has increased from −6.3% to 3.0%. Although ADP has increased while discontinuity is introduced, the gap of SPR is trivial. From this comparison, we find that the correctness of our algorithm is robust to discontinuities. SPR and ADP under different $\frac{I_{A0}}{m_A}$ are shown in Table 7, and visualized detection results are shown in Figure 10.

3.5.2. *Large-scale Stress Test*

Each GWAC camera will take one image every 15 seconds, containing about 176,000 stars, and the photos will be sent to a high-performance cluster through dedicated networks. According to the 15s sampling frequency of GWAC, the algorithm must finish all the calculating in less than 10s (about 5s for the preprocessing step) to satisfy the requirement of real-time processing. All of the computation will be held on the cluster, including source extraction, cross match, and light curve analysis, and the data produced by one camera will be processed by one corresponding high-performance machine. In this experimental environment, we use a single machine with 32-core, 2.2 GHz CPU, 256 GB memory, and a 50T SATA hard disk. A data generator that directly produces a data batch consisting of 176,000 data points, each of which is the newest brightness value of its corresponding star every 3 seconds, is used to simulate one camera. Even in such a far stricter time constraint, the Golang implementation of our algorithm can

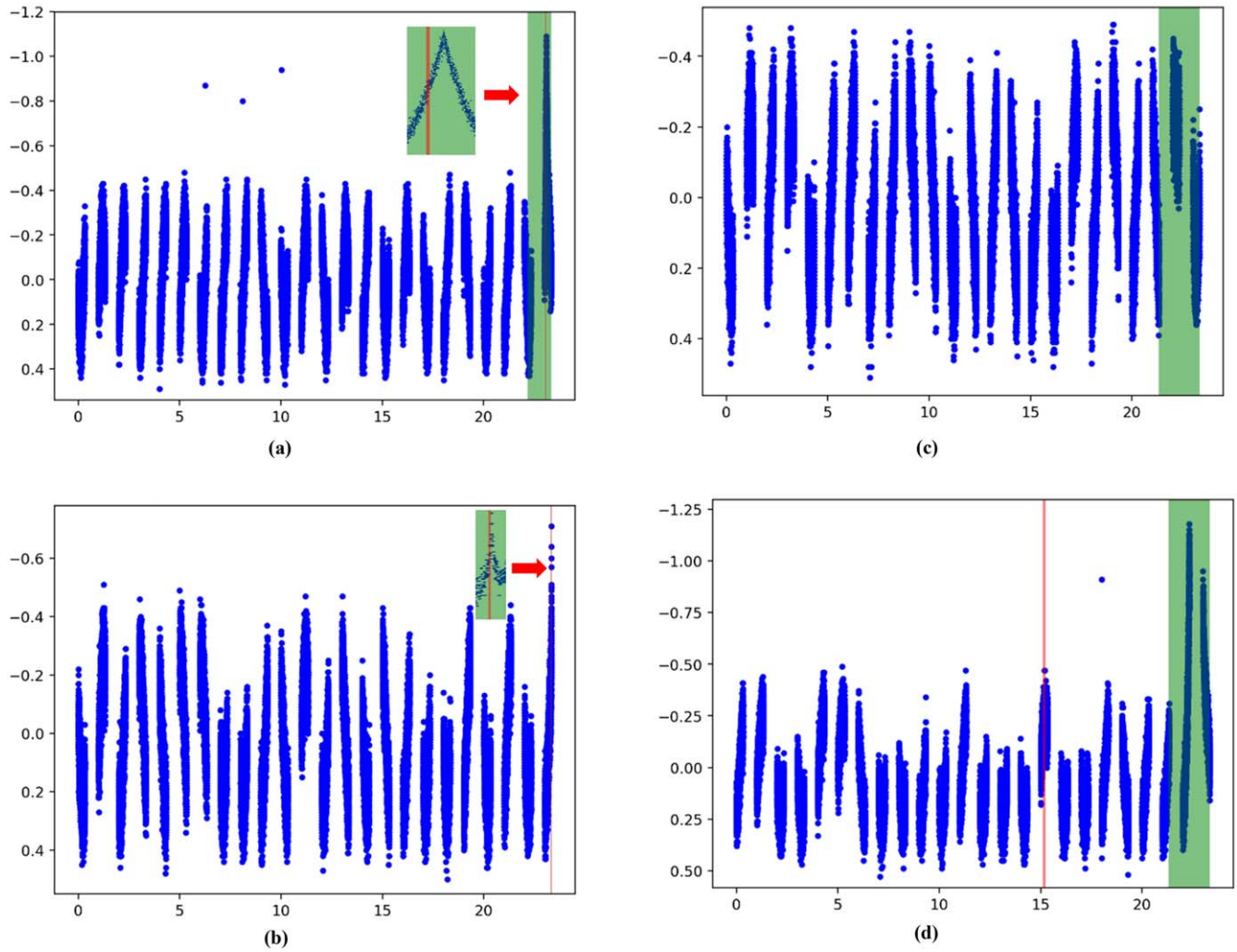

Figure 10. Some generated variable baselines and results of running our algorithm on them. The green rectangle area means the ground truth gravitational event section, while the red area means the triggering given by our algorithm. We can see that panels (a) and (b) are correct samples, while panels (c) and (d) is not (our algorithm did not give any trigger in panel (c), and gives wrong triggers on panel (d)).
(A color version of this figure is available in the online journal.)

Table 4
Overall Result on GWAC-Con

SPR (%)	ADP (%)
89.0	−7.2

Table 5
Evaluation Results with Different I_{A0} on GWAC-Con

I_{A0}	SPR (%)	ADP (%)
0.32	72.8	−3.3
0.51	97.9	−2.2
0.84	99.9	−6.8
1.52	93.1	−10.7
7.49	81.3	−12.8

finish analyzing all the data in time with visualization module attached, which proves the efficiency of our algorithm.

4. Conclusion

To search ML events, especially short-timescale events, from light curves, we present a simple and flexible algorithm framework, called normalized feature deviation (NFD), to monitor/process data stream of light curves. This algorithm has following advantages in searching ML events: (i) Fast processing: the processing time is less than 3 seconds under the

Table 6
Overall Result of Continuous/discontinuous Baselines on GWAC-Var

Continuity	SPR (%)	ADP (%)
Continuous	73.5	−6.3
Discontinuous	73.2	+3.0

Table 7
Evaluation Results of Discontinuous Baselines with Different $\frac{I_{A0}}{m_A}$ on GWAC-Var

$\frac{I_{A0}}{m_A}$	SPR (%)	ADP (%)
1.28	26.7	+12.9
2.04	53.7	+9.0
3.36	91.5	+6.4
6.08	97.5	−0.0
29.96	96.8	−3.3

pressure of 176,000 light curve streams at parallel. (ii) Smooth occasional noise point effect: it will avoid false triggers from occasional noise points with high deviation. (iii) Short delay: a large part of ML events can be searched before maximum. It makes possible follow-up observations after trigger. Our test on large-scale simulated variable baseline light curve data sets (139,968 valid samples, including both continuous and discontinuous data streams) indicates that our method is feasible on challenging data, while experiments on GWAC specified simulated data sets (12,960 constant stars and 3240 variable stars) prove that it is capable of dealing with all kinds of light curves likely observed by GWAC (a ground instrument of SVOM mission). As GWAC will generate high-cadence light curves (1point/15 second), it is possible to find ML with short timescale from hours to one day. Our real-time method will be suitable for ML search from GWAC data.

As no domain-specific knowledge is introduced into the NFD algorithm, it has the potential to be applied for other kinds of astronomical events. And the large-scale simulated light data sets are of great significance for studies and comparisons of mining algorithms. In the future, we plan to investigate more features and discriminant functions, conducting experiments on them, as well as testing the actual performance of our algorithm when GWAC is officially operated.

The authors thank Prof. R. Street and Dr. W. Zhu for their helpful discussion and suggestions, and the anonymous referee for the many insightful comments and suggestions that greatly improved the paper. This work was supported by the National key Research Program of China “Scientific Big Data Management System” (No. 2016YFB1000600), and the National Natural Science Foundation of China (No. U1431108, No. 61671272).

References

Ahmad, S., & Purdy, S. 2016, arXiv:1607.02480
Alcock, C., Allsman, R. A., Alves, D. R., et al. 2000, ApJ, 541, 734
Box, G., & Pierce, D. 1970, J. Am. Stat. Assoc., 65, 1509
Breiman, L. 2001, Mach. Learn., 45, 5
Feng, T., Du, Z., Sun, Y., et al. 2017, in 6th 2017 IEEE International Congress on Big Data (Honolulu, HI), 224
Hamadache, C., Guillou, L. L., Tisserand, P., et al. 2006, Astron. Astrophys., 454, 185
Hochreiter, S., & Schmidhuber, J. 1997, Neural Comput., 9, 1735
Mroz, P., Ryu, Y. H., Skowron, J., et al. 2017b, arXiv:1712.01042
Mroz, P., Udalski, A., Poleski, R., et al. 2017a, Natur, 548, 173
Paczyński, B. 1986, ApJ, 304, 1
Schanne, S. 2008, arXiv:0807.0738
Sumi, T., Kamiya, K., Bennett, D. P., et al. 2011, Natur, 473, 349
Udalski, A. 2003, AcA, 53, 291
Udalski, A., Szymanski, M., & Kaluzny, J. 1994, AcA, 44, 227
Udalski, A., Szymański, M. K., & Szymański, G. 2015, AcA, 65, 1
Wan, M., Wu, C., Wang, J., et al. 2016, PASP, 128, 114501
Wyrzykowski, Ł., Rynkiewicz, A. E., Skowron, J., et al. 2015, ApJS, 216, 12
Wyrzykowski, Ł., Udalski, A., Mao, S., et al. 2006, AcA, 56, 145
Wei, J., Cordier, B., Antier, S., et al. 2016, arXiv:1610.06892