MNRAS **000**, 1–15 (2018) Preprint 10 June 2019 Compiled using MNRAS LATEX style file v3.0

# Searching for fast extragalactic X-ray transients in *Chandra* surveys

G. Yang (杨光),[1,2]★ W. N. Brandt,[1,2,3] S. F. Zhu (朱世甫),[1,2] F. E. Bauer,[4,5,6] B. Luo (罗斌),[7] Y. Q. Xue (薛永泉),[8,9] and X. C. Zheng (郑学琛)[10]

[1]*Department of Astronomy and Astrophysics, 525 Davey Lab, The Pennsylvania State University, University Park, PA 16802, USA*
[2]*Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*
[3]*Department of Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA*
[4]*Instituto de Astrofísica and Centro de Astroingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Casilla 306, Santiago, Chile*
[5]*Millennium Institute of Astrophysics (MAS), Nuncio Monseñor Sótero Sanz 100, Providencia, Santiago, Chile*
[6]*Space Science Institute, 4750 Walnut Street, Suite 205, Boulder, Colorado 80301, USA*
[7]*School of Astronomy & Space Science, Nanjing University, Nanjing 210093, China*
[8]*CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei 2*
[9]*School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, China*
[10]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*



Accepted XXX. Received YYY; in original form ZZZ

**ABSTRACT**

Recent works have discovered two fast (≈ 10 ks) extragalactic X-ray transients in the *Chandra* Deep Field-South (CDF-S XT1 and XT2). These findings suggest that a large population of similar extragalactic transients might exist in archival X-ray observations. We develop a method that can effectively detect such transients in a single *Chandra* exposure, and systematically apply it to *Chandra* surveys of CDF-S, CDF-N, DEEP2, UDS, COSMOS, and E-CDF-S, totaling 19 Ms of exposure. We find 13 transient candidates, including CDF-S XT1 and XT2. With the aid of available excellent multiwavelength observations, we identify the physical nature of all these candidates. Aside from CDF-S XT1 and XT2, the other 11 sources are all stellar objects, and all of them have $z$-band magnitudes brighter than 20. We estimate an event rate of $59^{+77}_{-38}$ evt yr$^{-1}$deg$^{-2}$ for CDF-S XT-like transients with 0.5–7 keV peak fluxes $\log F_{\rm peak} \gtrsim -12.6$ (erg cm$^{-2}$ s$^{-1}$). This event rate translates to $\approx 15^{+20}_{-10}$ transients existing among *Chandra* archival observations at Galactic latitudes $|b| > 20°$, which can be probed in future work. Future missions such as *Athena* and the *Einstein Probe* with large grasps (effective area × field of view) are needed to discover a large sample (∼ thousands) of fast extragalactic X-ray transients.

**Key words:** X-rays: bursts – X-rays: general – X-rays: galaxies – X-rays: stars – Stars: activity – Methods: data analysis

## 1 INTRODUCTION

X-ray observations can provide uniquely insightful views of many astronomical phenomena such as accretion and mergers of compact objects (e.g. Brandt & Alexander 2015; Pooley et al. 2018). The X-ray sky is variable. Main-sequence stars (especially dwarfs) have strong flares powered by magnetic reconnection, generally lasting minutes to hours (e.g. Haisch et al. 1991; Güdel & Nazé 2009). X-ray binaries have various variability behaviors such as pulsations, bursts, and quasi-periodic oscillations (e.g. van der Klis 1989; Belloni & Stella 2014). Active galactic nuclei (AGNs) typically have red-noise X-ray variability, with characteristic amplitudes being $\lesssim$ 0.5 dex on timescales from ∼ an hour to ∼ 10 years (e.g. Markowitz et al. 2003a,b; Yang et al. 2016; Paolillo et al. 2017; Zheng et al. 2017). However, some relatively rare AGN and related phenomena, e.g. tidal disruption events, changing-look AGNs, and narrow-line Seyfert 1s, can have larger X-ray variability amplitudes (e.g. Komossa 2015; Kara et al. 2016; Ricci et al. 2016; Gallo 2018).

Recently, a new type of X-ray variability phenomenon has been revealed in the form of two relatively faint X-ray transients found in the *Chandra* observations of the *Chandra* Deep Field-South (CDF-S XT1 and XT2; Bauer et al. 2017;

★ E-mail: gyang206265@gmail.com (GY)

Xue et al. 2019). Both transients are fast ($T_{90} \approx 10$ ks,[1] observed-frame). Their origins are found to be extragalactic, with optical/near-infrared (NIR) counterparts at $z \approx 2.1$ (photometric redshift) and $z = 0.74$ (spectroscopic redshift), respectively. Both transients have $\gtrsim 100$ counts detected, corresponding to enormous amounts of energy release ($\gtrsim 10^{48}$ erg, assuming isotropic emission). Due to the lack of simultaneous multiwavelength observations and the small sample size of transients, the physical origins are not well determined with some possibilities being off-axis gamma-ray bursts, tidal-disruption events, mergers of neutron stars, and shock-breakout events. In this paper, we regard CDF-S XT1 and XT2 as the same "type" of transients considering their observational similarities in flux, timespan, and extragalactic origin, although their physical causes might be different.

Given the short timescales ($T_{90} \approx 10$ ks) and large numbers of counts ($\gtrsim 100$) for CDF-S XT1 and XT2, such transients should be easy to detect in any $\gtrsim 10$ ks *Chandra* exposure. The two transients are both detected in a small survey area ($\approx 480$ arcmin$^2$) and relatively short timespan (2014 October and 2015 March), indicating that a large population of X-ray transients might exist. Bauer et al. (2017) performed a preliminary transient search in the *Chandra* source catalog (CSC; Evans et al. 2010), which compiled *Chandra* observations before 2010 August 10. They did not find transients similar to the CDF-S transients. However, this CSC search is not conclusive, because the CSC is not dedicated to discovering fast transients and thus potential transients might be missed or poorly/incorrectly characterized. Also, many CSC sources have only a single short *Chandra* visit, making it difficult to ascertain the transient and quiescent levels. The CSC sources also generally lack deep optical/NIR observations, preventing further studies of the physical nature of potential transients.

To mitigate the above issues, in this work, we search for similar transients in *Chandra* archival observations of X-ray surveys. We develop a method to identify CDF-S XT-like transients in a single *Chandra* exposure, which is applicable to any *Chandra* imaging observation. In the surveys, most X-ray sources have been visited by two or more *Chandra* exposures, allowing us to inspect transients with multi-epoch X-ray data and study their quiescent behaviors. Deep multiwavelength data are critical in clarifying the physical origins of X-ray transients. CDF-S XT1 and XT2 have optical/NIR counterparts with $V \gtrsim 25$ mag and $H \gtrsim 24$ mag (Bauer et al. 2017; Xue et al. 2019), well beyond the detection limit of wide-field surveys such as SDSS (York et al. 2000) and UKIDSS (Lawrence et al. 2007). Glennie et al. (2015) discovered an X-ray transient in one *Chandra* archival observation, but were not able to clarify its physical origin due to the lack of deep multiwavelength data. Our selected X-ray surveys are accompanied by deep multiwavelength observations, allowing identifications of optical/NIR counterparts for the selected transients.

The main aim of this paper is to search for fast extragalactic X-ray transients that are similar to CDF-S XT1 and XT2 rather than general X-ray transients (although our search is effective for a fairly wide range of transients; see Appendix A). The structure of this paper is organized as follows. We detail our X-ray transient-selection algorithm and assess its efficiency with simulations in §2. We describe our X-ray data, selection of transient candidates, and optical/NIR counterparts in §3. We estimate the event rate of CDF-S XT-like transients based on our results and discuss the prospect of future missions in §4. We summarize our results in §5.

Throughout this paper, we assume a cosmology with $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. Quoted uncertainties are at the $1\sigma$ (68%) confidence level, unless otherwise stated. Quoted optical/infrared magnitudes are AB magnitudes.

## 2 METHODOLOGY

In §2.1, we detail our algorithm for transient-candidate searching, which is designed to find CDF-S XT-like events within individual *Chandra* exposures. Our algorithm is simple and fast, and can be easily implemented for any individual *Chandra* observations. We perform intensive Monte Carlo simulations and assess the sensitivity of our algorithm in §2.2.

### 2.1 Algorithm for Transient-Candidate Selection

Our algorithm works on an unbinned *Chandra* light curve, i.e. an array of photon arrival times of a source, for which the background has been estimated. Below, we denote $N_{\rm tot}$ ($N_{\rm bkg}$) as the number of total (background) counts for the light curve. We require that the source lies within an off-axis angle of $8'$, following previous *Chandra* studies (e.g. Vito et al. 2016; Yang et al. 2016). This is because *Chandra*'s performance (as measured by, e.g. effective area and PSF size) drops significantly beyond $8'$. Additionally, we require that the light-curve length is shorter than 50 ks to avoid large numbers of accumulated background counts in long exposures. Exposures longer than 50 ks are chopped into a few parts to meet this requirement (§3). In §2.2.3, we show that our algorithm reaches a uniform sensitivity for nearly all observations shorter than 50 ks. Note that the choice of 50 ks is somewhat subjective; the flux limit and the results of our transient search (§2.2.3 and §3.3) do not change significantly if we adjust this value between $\approx 16$ ks and $\approx 100$ ks. Choosing a value below $\approx 16$ ks could chop some observations into $\lesssim 8$ ks parts, which are ineffective in our selection of XT-like transients (see §2.2). Choosing a value above $\approx 100$ ks could leave some long observations unchopped, which have relatively high accumulated background, affecting transient detection.[2]

Our algorithm first calculates $N_1$ and $N_2$, defined as the numbers of counts at $t = (t_{\rm s}, t_{\rm m})$ and $t = (t_{\rm m}, t_{\rm e})$, respectively, where $t_{\rm s}$ and $t_{\rm e}$ are the times when the exposure starts and ends, respectively, and $t_{\rm m} = (t_{\rm s} + t_{\rm e})/2$, i.e. the midpoint of exposure time. Since typical *Chandra* observations are continuous and uninterrupted by background flares ($\approx 1\%$ of

[1] $T_{90}$ is defined as the time interval between the arrival times of the 5%-th photon and the 95%-th photon.

[2] If we do not chop the observations, the actually detected extragalactic transients among our data will be the same, although there will be four more stellar flares detected (§3).

exposure time), our two-part division of the exposure is legitimate.

We select a source in an observation as a transient candidate if it satisfies all of the following criteria (Method 1):

(A) $N_{\rm tot}$ is larger than the $5\sigma$ Poisson upper limit of $N_{\rm bkg}$;
(B) $N_1$ and $N_2$ are statistically different at a $>4\sigma$ significance level;
(C) $N_1 > 5\times N_2$ or $N_2 > 5\times N_1$.

Criterion A filters out faint sources that have low signal-to-noise ratios (S/N), and thus boosts the speed of the selection process. This criterion is also helpful in avoiding false detections caused by rare background flares, since flares can dominate the detected counts for faint sources. Criterion B selects sources that have significantly different count rates in the first-half and second-half exposures. Technically, we realize Criterion B with the $E$-test (Krishnamoorthy & Thomson 2004). The $E$-test can test if two Poisson variables ($N_1$ and $N_2$ in our case) are drawn from the same distribution, and simultaneously considers the statistical fluctuations of both variables. Criteria A and B are based on statistical significance, and they select high-S/N sources with significant variability. However, these criteria are not sufficient, since they cannot rule out AGNs which have stochastic variability. To deal with this AGN issue, we also add Criterion C, which requires that the flux-variation amplitude is large. Since the characteristic AGN variability amplitudes (on timescales from $\sim$ an hour to $\sim$ 10 years) are a factor of $\lesssim 3$ (§1), we choose the amplitude threshold as a factor of 5 to cleanly rule out AGN variability. We note that the choice of amplitude threshold is empirical: a low value could not remove AGNs effectively; a high value could miss potential transients. We have tested adjusting the threshold slightly (e.g., by a factor of 3 or 4 instead of 5), and the number of extragalactic transients we found in §3 does not change.

Method 1 is not efficient in selecting transients that happen at $t \approx t_{\rm m}$, because these transients will have similar $N_1$ and $N_2$. To overcome this defect, we also select transients with the following method. We denote $N_1'$ as the number of counts at $t = (t_{\rm bgn}, t_{\rm q1})$ plus that at $t = (t_{\rm q3}, t_{\rm end})$, where $t_{\rm q1}$ and $t_{\rm q3}$ are the first and third quartiles of the observation time, and $N_2'$ as the number of counts at $t = (t_{\rm q1}, t_{\rm q3})$. We also select a source as a transient candidate, if it satisfies (Method 2)

(A′) $N_{\rm tot}$ is larger than the $5\sigma$ Poisson upper limit of $N_{\rm bkg}$;
(B′) $N_1'$ and $N_2'$ are statistically different at a $>4\sigma$ significance level;
(C′) $N_1' > 5\times N_2'$ or $N_2' > 5\times N_1'$.

In §2.2, we prove the necessity of adopting both Method 1 and Method 2 for transient selection.

## 2.2 Efficiency of the Selection Algorithm

In this Section, we assess the efficiency of our transient-selection algorithm (§3.2) with Monte Carlo simulations. In §2.2.1, we detail our simulation configurations. In §2.2.2, we define a "gauge" to measure the efficiency of our algorithm. In §2.2.3, we present our simulation results, i.e. the performance of our algorithm.

### 2.2.1 Simulation Configurations

The simulations are based on a fiducial light-curve model. Since our main goal is to search for fast extragalactic transients analogous to CDF-S XT1 and XT2, we adopt a light-curve model similar to the best-fit models of these two transients (Bauer et al. 2017; Xue et al. 2019). The light-curve shape in the model is described by

$$\mathrm{cntR(t)} \propto \begin{cases} 0, & t < 0 \\ t, & 0 \le t < t_1 \\ t^{\alpha_1}, & t_1 \le t < t_2 \\ t^{\alpha_2}, & t \ge t_2 \end{cases} \quad (1)$$

where cntR is the count rate in units of counts $\mathrm{s}^{-1}$. Here, we follow the convention that the transient starts at $t = 0$. For $t$ between 0 and $t_1$, the cntR rises to the peak value. This time interval is very short ($\lesssim 100$ s for both CDF-S XT1 and XT2), and thus the exact functional form is not important. Here, we adopt a basic form of a linear rise and set $t_1 = 50$ s. For $t$ between $t_1$ and $t_2$, the light curve is roughly in a plateau with an index of $\alpha_1 = -0.1$. This plateau only exists for XT2 (2.3 ks) but not for XT1, and we adopt $t_2 = t_1 + 1$ ks. For $t > t_2$, the adopted cntR is a power-law decline with an index of $\alpha_2 = -2$, which is between those of XT1 ($-1.5$) and XT2 ($-2.2$). We adopt a power-law spectral shape with photon index of $\Gamma = 1.6$ for the model, which is consistent with those measured for both XT1 and XT2. We note that changing the model parameters slightly (e.g. changing $t_1$ to 100 s and $\Gamma$ to 2.0) does not significantly affect our simulation results. In Appendix A, we also perform simulations for some other types of transients that are significantly different from the CDF-S XTs, although these transients are not the main focus of this work; these simulations show that our algorithm can identify transients with timescales $\lesssim$ exposure time while the details of the light-curve shapes do not affect the sensitivity significantly. We plot the adopted light-curve model in Fig. 1. The $T_{90}$ for this light-curve setting is 9.4 ks, similar to those of XT1 and XT2. This similarity is expected, because our model in Eq. 1 is based on the light-curve shapes of XT1 and XT2. Under the fiducial-model configuration, the conversion between peak flux and total net counts is

$$N_{\rm net} \approx 1.6\times10^{14} F_{\rm peak}\ (\mathrm{erg\ cm^{-2}\ s^{-1}}). \quad (2)$$

The conversion factor is calculated with PIMMS, assuming a typical off-axis angle of $5'$ when accounting for vignetting (i.e. the drop of photon-collecting area toward large off-axis angle; see Appendix B for other off-axis angles).[3]

Background noise is also needed for the simulations. Here, background includes both detector background and sky X-ray background for 0.5–7 keV. The background-extraction region is an annulus centered at the X-ray source (see §3.1 for details). The background level rises as a function of off-axis angle. In the simulations, we assume a background of $5.6\times10^{-5}$ cnt $\mathrm{s}^{-1}$, which is the typical background level at an off-axis angle of $5'$ (see Appendix B for other off-axis angles). The adopted background is also approximately the median value for all X-ray sources in our studied surveys.

[3] See http://cxc.harvard.edu/toolkit/pimms.jsp for PIMMS; see http://cxc.harvard.edu/proposer/POG/html/chap4.html for vignetting.

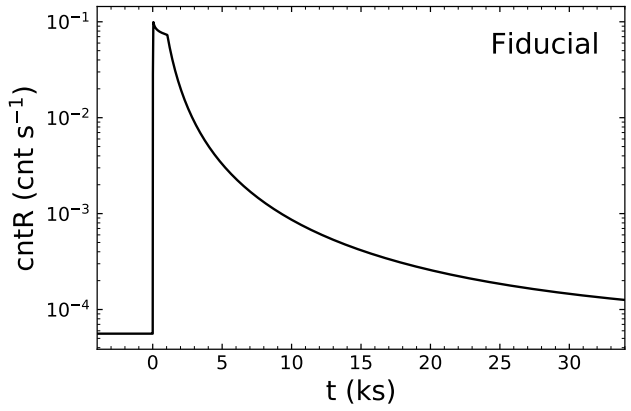


**Figure 1.** The fiducial light-curve model adopted in our simulations. The light-curve shape is similar to those of CDF-S XT1 and XT2. The time ($x$-axis) zero point is chosen such that the transient starts at $t = 0$. The plot is generated with peak flux $\log F_{\rm peak} = -12.0$ (cgs), for display purposes only. In the simulations, we test different $F_{\rm peak}$ values (§2.2.3).

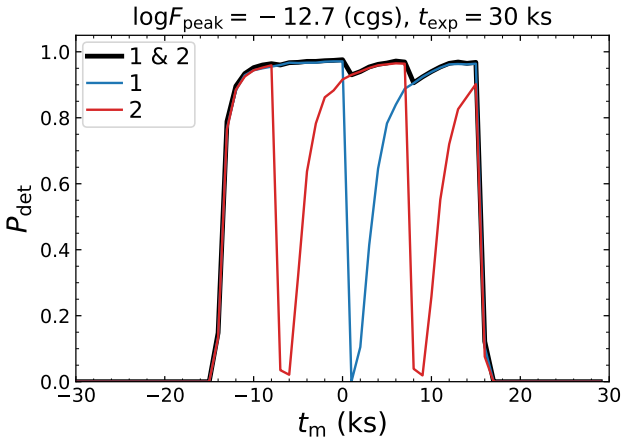


**Figure 2.** $P_{\rm det}$ as a function of exposure midpoint (see §3.2). The black curve represents the results of both Method 1 and Method 2; the blue and red curves represent the results of Method 1 and Method 2, respectively. $P_{\rm det}$ is calculated based on simulations (§2.2). The time ($x$-axis) zero point is chosen such that the transient starts at $t = 0$, and thus $t_{\rm m}$ means the relative time between the exposure midpoint and the transient start time. As labelled, different panels are for different net counts and exposure times. There are some significant drops in the curves of Method 1 and Method 2, which are related to our transient-detection algorithm. For example, when the transients happens at $t \approx t_{\rm m}$, the efficiency of Method 1 is low (see §2.1).

This background level only corresponds to ≈ 3 background counts for a 50 ks light curve, which is the longest light curve analyzed (see §2.1).

### 2.2.2 *Efficiency Gauge*

For a given set of $F_{\rm peak}$ and $t_{\rm exp}$ (exposure time), we can estimate the probability of transient detection ($P_{\rm det}$) as a function of $t_{\rm m}$ (observation midpoint; §2.1) with the simulation procedures described below. Since the transient starts at $t = 0$ (§2.2.1), $t_{\rm m}$ actually means the relative time between the exposure midpoint and the transient start time.

First, we simulate light curves in the time interval of $t = (-t_{\rm exp},\ t_{\rm exp})$. We divide $t = (-t_{\rm exp},\ t_{\rm exp})$ into small bins with $\Delta t = 5$ s. We then calculate the expected total counts in each bin. Using these values, we generate the counts in each bin with a Poisson distribution, which gives a simulated light curve. We repeat the procedures and generate 1,000 light curves. We apply both Method 1 and Method 2 (§3.2) for these light curves and calculate the fraction of successful detections. We adopt this fraction as the detection probability ($P_{\rm det}$).

Fig. 2 displays an example of $P_{\rm det}$ vs. $t_{\rm m}$ for $\log F_{\rm peak} = -12.7$ (cgs) and $t_{\rm exp} = 30$ ks. Besides showing the $P_{\rm det}$ when using both Method 1 and Method 2 (see §2.1), Fig. 2 also displays the $P_{\rm det}$ when using Method 1 and Method 2 separately. Note that $P_{\rm det}$ drops significantly for some $t_{\rm m}$ values when using Method 1 and Method 2 separately. However, such drops are greatly alleviated when using both Methods, indicating the necessity of our combined method strategy.

From Fig. 2, $P_{\rm det}$ (using both Methods) is not constant for different $t_{\rm m}$. This $P_{\rm det}$ variation makes it difficult to use $P_{\rm det}$ as a direct measure of algorithm performance as a function of $F_{\rm peak}$ and $t_{\rm exp}$. Therefore, we define an "effective" detection probability ($P_{\rm eff}$) averaged over different $t_{\rm m}$ as a gauge to measure the efficiency, i.e.

$$P_{\rm eff} = \frac{\int_{-\infty}^{+\infty} P_{\rm det}(t_{\rm m}) dt_{\rm m}}{t_{\rm exp}}. \quad (3)$$

From this definition, $P_{\rm eff}$ ranges from 0 to ≈ 1 for a given set of $F_{\rm peak}$ and $t_{\rm exp}$,[4] with higher values indicating higher average detection efficiency.

### 2.2.3 *Simulation Results*

We calculate $P_{\rm eff}$ for different $t_{\rm exp}$ and $F_{\rm peak}$ and show the results in Fig. 3. As expected, $P_{\rm eff}$ rises toward high $F_{\rm peak}$ at a given $t_{\rm exp}$, because brighter sources have higher S/N. We choose $\log F_{\rm peak} \approx -12.6$ (cgs) as our detection limit, above which $P_{\rm eff} \approx 1$ for a wide range of $t_{\rm exp} = 8$–50 ks. Note that this flux limit is much lower than the peak fluxes of CDF-S XT1 and XT2 (see Table 2). The estimated flux limit is mainly used to estimate the event rate in §4, we note that there are still non-zero probabilities to detect transients below this limit (see Fig. 3). Here, we remind readers that $t_{\rm exp} = 50$ ks is the maximum exposure time accepted by our algorithm (see §2.1). We note that the simulation results above are calculated from our fiducial model which is similar to CDF-S XTs (§2.2.1; see Appendix A for some other transient models), since our main purpose is to find CDF-S XT-like transients. The simulation results are based on the instrumental response and background at a typical off-axis

[4] $P_{\rm eff}$ might be slightly greater than unity, because a transient may be detected even when it is partially covered by the observations.

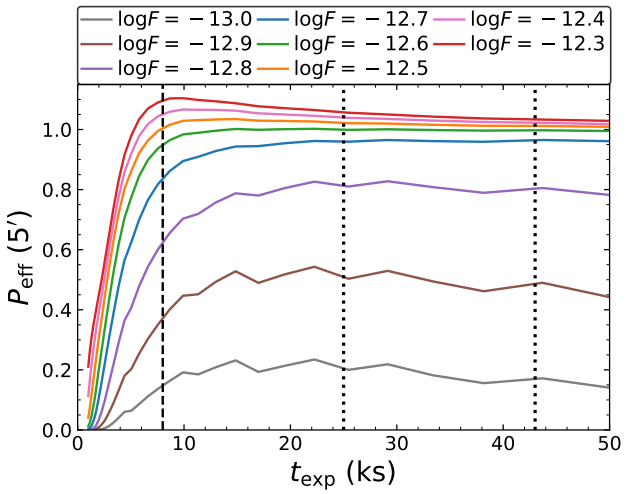


**Figure 3.** $P_{\rm eff}$ as a function of $t_{\rm exp}$ for a typical off-axis angle of 5′ (see §2.2.1). Different colors indicate different $F_{\rm peak}$ as labelled (cgs units). The black vertical dashed line marks $t_{\rm exp}$ = 8 ks. The black vertical dotted lines indicate the 20%–80% percentile range of the exposure times of our data (§3.2). Above our selected flux limit ($\log F_{\rm peak} = -12.6$; §2.2.3) $P_{\rm eff}$ rises to ≈ 1 for $t_{\rm exp} \gtrsim 8$ ks.

angle of 5′ (§2.2.1), and we present the results at other off-axis angles in Appendix B.

Below $t_{\rm exp}$ = 8 ks, $P_{\rm eff}$ drops significantly at a given $F_{\rm peak}$ (see Fig. 3). This is because, when the exposure time becomes significantly shorter than the transient timescale, the observed light curve will be similar to a normal variable source, and thus may not be selected by our algorithm. In our estimation of event rate (§4), we do not include observations that are shorter than 8 ks, although we do not discard these observations in our transient search (§3). Only a negligible fraction of observation time (≈ 0.1%; see §4) in our analyzed X-ray data is from < 8 ks exposures.

## 3 DATA AND ANALYSES

The scope of this paper is to search for CDF-S XT-like extragalactic transients. Utilizing the methodology detailed in §2, we first perform an initial search for transient candidates in the X-ray survey data (§3.1 and §3.2). Since stellar objects can have strong X-ray flares that might be selected by our algorithm, we need to exclude stars from our selected transient candidates. We perform this task with the high-quality multiwavelength data available for the surveys (§3.3).

### 3.1 X-ray Data and Processing

In this work, we analyze the *Chandra* survey data from the CDF-S, CDF-N, DEEP2, UDS, COSMOS, and E-CDF-S regions. The survey properties are summarized in Table 1. DEEP2 includes the full field of EGS (DEEP2-1) and three other fields (DEEP2-2, DEEP2-3, and DEEP2-4) with shallower (≈ 10 ks) exposures. The total exposure time of these surveys is 19 Ms. All the surveys are at high Galactic latitude ($|b| \gtrsim 40°$), matching our main interest of searching for extragalactic transients. Also, these surveys have deep multiwavelength coverage, allowing us to study the physical origins of the transients (§1 and §3.3).

We download all the *Chandra* data products of observations related to the surveys, and run the CHANDRA_REPRO script in CIAO 4.10.[5] The CHANDRA_REPRO script performs standard cleaning and calibration processes,[6] and yields a clean event file for each observation. Based on the data products of CHANDRA_REPRO, we produce the exposure map for each observation using the CIAO script FLUXIMAGE. The exposure maps denote the "effective" exposure times for different positions in the field of view, and instrumental factors such as bad pixels and vignetting are taken into account.

For each event file, we extract the 0.5–7 keV photons of each X-ray source presented in the X-ray catalogs (Table 1). Since the *Chandra* background is extremely low, any sources with $\gtrsim$ 10 net counts should be detected by the X-ray surveys. This level of counts is much lower than that of our transient-selection sensitivity (see below), and thus we should not miss any transients due to their absence in the X-ray catalogs. The total events are extracted from an aperture of $1.5 \times R_{90}$, where $R_{90}$ is the radius encircling 90% of the X-ray counts. We adopt $R_{90}$ as a function of off-axis angle from Table A1 of Vito et al. (2016). From simulations with the CIAO script SIMULATE_PSF, we find that this aperture size ($1.5 \times R_{90}$) encircles nearly all ($\gtrsim$ 98%) X-ray counts regardless of off-axis angle. The background events are extracted from an annulus with inner and outer radii of $1.5 \times R_{90}$ and $1.5 \times R_{90} + 20$ pixels. The background area is 9 times larger than the source area for a typical source at an off-axis angle of 5′. If the background region covers a nearby X-ray source, we mask the source (also with radius of $1.5 \times R_{90}$), and do not include the masked area when estimating the background. We note that changing the source and background extraction regions slightly will not affect our qualitative results. We estimate the background counts in the source region ($N_{\rm bkg}$; §2.1) by scaling the counts in the background region by a factor. Here, the scaling factor is the sum of the exposure-map values in the source area divided by that in the background area.

### 3.2 Selection of Transient Candidates

We apply the algorithm in §2.1 to the light curves extracted in §3.1. We note that the transient selection is only applied to sources with off-axis angle of < 8′ to avoid the low-quality X-ray data beyond 8′ (§2.1). If a light curve is longer than 50 ks (the maximum length accepted by our algorithm; §2.1), we chop it into several continuous parts with each having the same $t_{\rm exp}$ shorter than (or equal to) 50 ks (§2). For example, for a 80 ks exposure, we divide it into two parts each having $t_{\rm exp}$ = 40 ks. We then perform transient selection for each chopped light curve independently. After this observation-chopping process, we have 610 exposures with a median $t_{\rm exp}$

[5] CHANDRA_REPRO cannot process observation 1431 (CDF-S), which consists of two separate exposures. For this observation, we use the data products from Luo et al. (2017), who split the observation into two continuous exposures. We perform transient searching for these two exposures independently (§3.2), but do not find transient candidates in the two exposures.
[6] http://cxc.harvard.edu/ciao/ahelp/chandra_repro.html

**Table 1.** Properties of X-ray Surveys Analyzed in this Work

| Survey (1) | Area (2) | Total Exp. (3) | Obs. Num. (4) | Src. Num. (5) | Reference (6) |
|---|---|---|---|---|---|
| CDF-S | 0.13 | 6.9 | 101 | 1008 | Luo et al. (2017) |
| CDF-N | 0.12 | 2.0 | 20 | 683 | Xue et al. (2016) |
| DEEP2 | 3.28 | 3.7 | 139 | 2976 | Goulding et al. (2012); Nandra et al. (2015) |
| UDS | 0.33 | 1.2 | 25 | 868 | Kocevski et al. (2018); Suh et al. in prep. |
| COSMOS | 2.20 | 4.5 | 117 | 4016 | Civano et al. (2016); Marchesi et al. (2016) |
| E-CDF-S | 0.31 | 1.0 | 9 | 1003 | Xue et al. (2016) |
| All | 6.38 | 19.3 | 411 | 10554 | – |

NOTE. — (1) X-ray survey name. (2) Survey area in deg$^2$. (3) Total exposure time in Ms. (4) Number of *Chandra* observations (before chopping; see §3.2). (5) X-ray source number. (6) References where the survey details and source catalog are presented. Additional information about the CDF-S, CDF-N, and E-CDF-S can be found in Xue (2017).

of 30 ks and a 20%–80% percentile range of 25–43 ks. We show the $t_{\rm exp}$ distribution of these 610 exposures in Fig. 4.

For Method 1 (2), Criterion A (A′) selects a total of 9379 (9379) events in the 610 exposures analyzed. Among these events, Criterion B (B′) further selects 31 (24) events. Finally, Criterion C (C′) picks out 11 (5) events as the events selected by Method 1 (2). For the events filtered out by Criterion C (C′), ≈ 70% of them are stellar flares, identified with the methods detailed in §3.3; the other ≈ 30% have extragalactic origins. We have examined the light curves of these extragalactic sources and found all of them have significant non-zero quiescent fluxes, and thus they are likely AGNs rather than extragalactic transients. This result demonstrates the capability of Criterion C (C′) in removing AGN variability (§2.1). We merge the events selected by Method 1 and Method 2, leading to a sample of 13 unique transient candidates. Among these 13 candidates, 8 and 2 are uniquely selected by Method 1 and Method 2, respectively, indicating the importance of using both Methods (see §2).

We visually inspect the background light curves of these transient candidates, and do not find significant flares. We have checked the X-ray images of the transients in both sky and detector coordinates. For each source, the events are concentrated and extended in sky and detector coordinates, respectively. This indicates that the transient candidates are physical X-ray sources rather than hot pixels, because hot pixels will lead to extended (concentrated) patterns in the sky (detector) coordinates caused by *Chandra* dithering.

The X-ray properties of the 13 transient candidates are listed in Table 2. ID1 and ID2 are CDF-S XT1 and XT2, respectively. Their successful selection indicates that our method of transient searching is effective for selecting CDF-S XT-like transients (§2.2.3). For each transient candidate, we calculate the hardness ratio for the observation where the transient is identified. Here, hardness ratio is defined as $(H - S)/(H + S)$, where $H$ and $S$ are hard-band (2–7 keV) and soft-band (0.5–2 keV) net counts, respectively. The $1\sigma$ uncertainty is calculated with BEHR, a Bayesian code for hardness ratio estimation (Park et al. 2006). The results are listed in Table. 2. In Fig 5, we show the distribution of hardness ratios. The spectral shapes of XT1 and XT2 are harder than for other transient candidates.

In Fig. 6 (left), we show the light curves of the transient candidates during the observation when the transient happens. The light curves are derived from the X-ray events extracted in §3.2, and are binned by 5-count intervals. The data points in these light curves indicate total count rates, including contributions from the source and background. The estimated average background count rate is marked as the dashed line in each panel of Fig. 6 (left). The durations of XT1 and XT2 tend to be shorter than for other transient candidates (Fig. 6 left). The $T_{90}$ values of XT1 and XT2 are $5.0^{+4.2}_{-0.3}$ ks and $11.1^{+0.4}_{-0.6}$ ks, respectively (see Bauer et al. 2017 and Xue et al. 2019 for details). We do not derive $T_{90}$ for other sources, because $T_{90}$ cannot be derived for many transients that extend beyond the *Chandra* exposures (e.g. ID3 and ID9 in Fig. 6 left). Also, unlike XT1 and XT2, many of the other transient candidates have non-zero fluxes in the quiescent states, and thus their $T_{90}$ calculation requires careful subtraction of the quiescent fluxes, which is beyond the scope of this work.

We plot the long-term light curves in Fig. 6 (right), where each *Chandra* observation is represented by a data point. These data points indicate net count rates, which are background-subtracted. As expected, the transient observation generally has a count rate much higher than other observations. However, unlike the CDF-S XT1 and XT2 events, most of the other transient candidates have detectable signals in some of the non-transient observations. Also, CDF-S XT1 and XT2 tend to have higher hardness ratios than the rest of the selected transient candidates (Fig. 5). These differences indicate that most of the new transient candidates are physically distinct from CDF-S XT1 and XT2 (see §3.3).

### 3.3 Optical/NIR Counterparts

We have compiled the likelihood counterpart matching results from the survey catalogs (Table 1). All the transient candidates have optical/NIR counterparts. The counterpart properties are presented in Table 3. We also match the counterparts with the *Gaia* catalog (Gaia Collaboration et al. 2018) using a 1″ matching radius, and mark the sources with non-zero parallax and/or proper motion as "star" in Table 3.

We show the optical/IR image cutouts in Fig. 7. From

**Table 2.** X-ray Properties of Transient Candidates

| ID (1) | Survey (2) | RA (3) | DEC (4) | Pos. Unc. (5) | Obs. ID (6) | Off. Ang. (7) | HR (8) | $\log F_{\rm peak}$ (9) | Method (10) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | CDF-S | 53.16156 | −27.85934 | 0.32″ | 16454 | 4.3′ | $-0.13^{+0.09}_{-0.10}$ | −11.41 | 1,2 |
| 2 | CDF-S | 53.07648 | −27.87339 | 0.31″ | 16453 | 4.1′ | $-0.32^{+0.08}_{-0.09}$ | −12.18 | 1,2 |
| 3 | CDF-N | 189.02046 | 62.33728 | 0.20″ | 957 | 6.6′ | $-0.54^{+0.08}_{-0.12}$ | −12.59 | 1 |
| 4 | CDF-N | 189.10587 | 62.23467 | 0.10″ | 3389 | 3.2′ | $-0.82^{+0.05}_{-0.06}$ | −12.82 | 1 |
| 5 | DEEP2 | 215.07414 | 53.10650 | 0.36″ | 9875 | 6.7′ | $-0.72^{+0.08}_{-0.12}$ | −12.53 | 1 |
| 6 | DEEP2 | 214.96015 | 52.74344 | 0.26″ | 9456 | 6.6′ | $-0.63^{+0.12}_{-0.14}$ | −12.97 | 1 |
| 7 | DEEP2 | 214.61007 | 52.54347 | 0.20″ | 9735 | 4.8′ | $-0.83^{+0.04}_{-0.17}$ | −12.75 | 1,2 |
| 8 | DEEP2 | 214.66798 | 52.66658 | 0.11″ | 5849 | 3.0′ | $-0.77^{+0.08}_{-0.10}$ | −13.21 | 2 |
| 9 | DEEP2 | 252.12761 | 34.96337 | 0.53″ | 8636 | 7.5′ | $-0.62^{+0.09}_{-0.12}$ | −12.46 | 1 |
| 10 | UDS | 34.48317 | −5.09118 | 0.96″ | 17305 | 0.7′ | $-0.53^{+0.15}_{-0.18}$ | −13.14 | 1 |
| 11 | COSMOS | 149.75403 | 2.14188 | 0.30″ | 8021 | 4.0′ | $-0.80^{+0.09}_{-0.14}$ | −13.38 | 1 |
| 12 | COSMOS | 149.82641 | 2.71812 | 0.30″ | 15214 | 5.9′ | $-0.58^{+0.07}_{-0.09}$ | −12.59 | 1 |
| 13 | COSMOS | 149.99794 | 2.77972 | 0.90″ | 15211 | 6.5′ | $-0.69^{+0.12}_{-0.14}$ | −12.66 | 2 |

NOTE. — (1) Transient-candidate ID in this work. (2) X-ray survey name. (3), (4), and (5) X-ray source position and positional error from the corresponding survey catalog. The positional error is taken from the survey catalog, and is calculated based on all observations that cover the source (not only the observation in Column 6). For example, ID6 has a lower positional uncertainty than ID15, because the former has more total net counts than the latter (≈ 100 vs. ≈ 25). (6) *Chandra* ID of the observation where the transient is identified. (7) Off-axis angle of the transient in the observation. (8) Hardness ratio based on the observation in Column 6. The uncertainties are at the $1\sigma$ level and are calculated with BEHR (§3.2). (9) Logarithmic 0.5–7 keV peak flux converted from the peak count rate in Fig. 6 with the method in §2.2.1. (10) The Method(s) responsible for identifying the transient candidate.

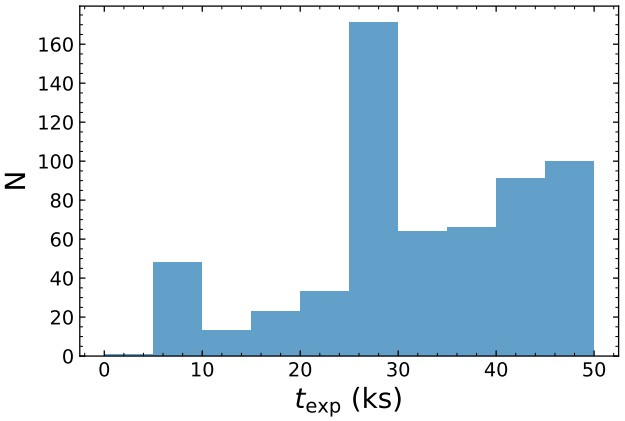


**Figure 4.** The $t_{\rm exp}$ distribution of the 610 exposures (after exposure chopping; §3.2) analyzed in this work. There is a peak at the $t_{\rm exp}$ = 25–30 ks bin, because the original observation set has many (74) exposures of $t_{\rm exp}$ = 50–60 ks, and these exposures are chopped to exposures of $t_{\rm exp}$ = 25–30 ks.

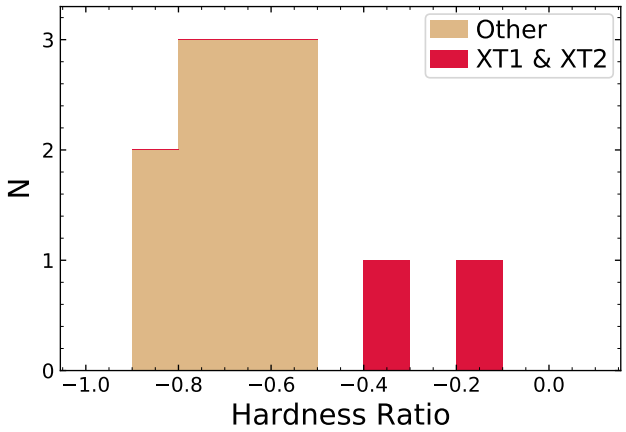


**Figure 5.** Hardness-ratio distribution of our selected transient candidates. CDF-S XT1 and XT2 are highlighted with the red color. The spectral shapes of XT1 and XT2 are harder than most of the other transients.

Fig. 7, the optical positions[7] are within (or marginally outside, i.e. ID6 and ID7) the $3\sigma$ X-ray positional errors, indicating that the X-ray and optical/NIR positions are generally consistent with each other. For ID6 and ID7, in the image cutouts nearby the X-ray positions, there appear to be no other optical/NIR sources except the counterparts, and thus the counterparts are likely the same physical objects as the X-ray sources.

From Table 3, ID1 and ID2 (namely CDF-S XT1 and XT2) are likely of extragalactic origin and have already been discussed in detail (Bauer et al. 2017; Xue et al. 2019). The other transients are relatively bright ($\rm mag_z < 20$), and all of them are reliably identified as stellar objects by optical/NIR spectroscopy and/or *Gaia*. Therefore, all the new transient candidates (aside from CDF-S XT1 and XT2) are stellar flares. These stellar objects have different variability properties, e.g. some have significant non-zero fluxes detected in the non-bursting observations (e.g. ID3 and ID4; see Fig. 6) while others do not (e.g. ID5 and ID9). However, since the

[7] The positional errors of the optical/NIR counterparts are not provided in the corresponding catalogs. Estimating the optical/NIR positional errors requires addressing factors such as CCD saturation and seeing (for ground-based telescopes), which are beyond the scope of this work.

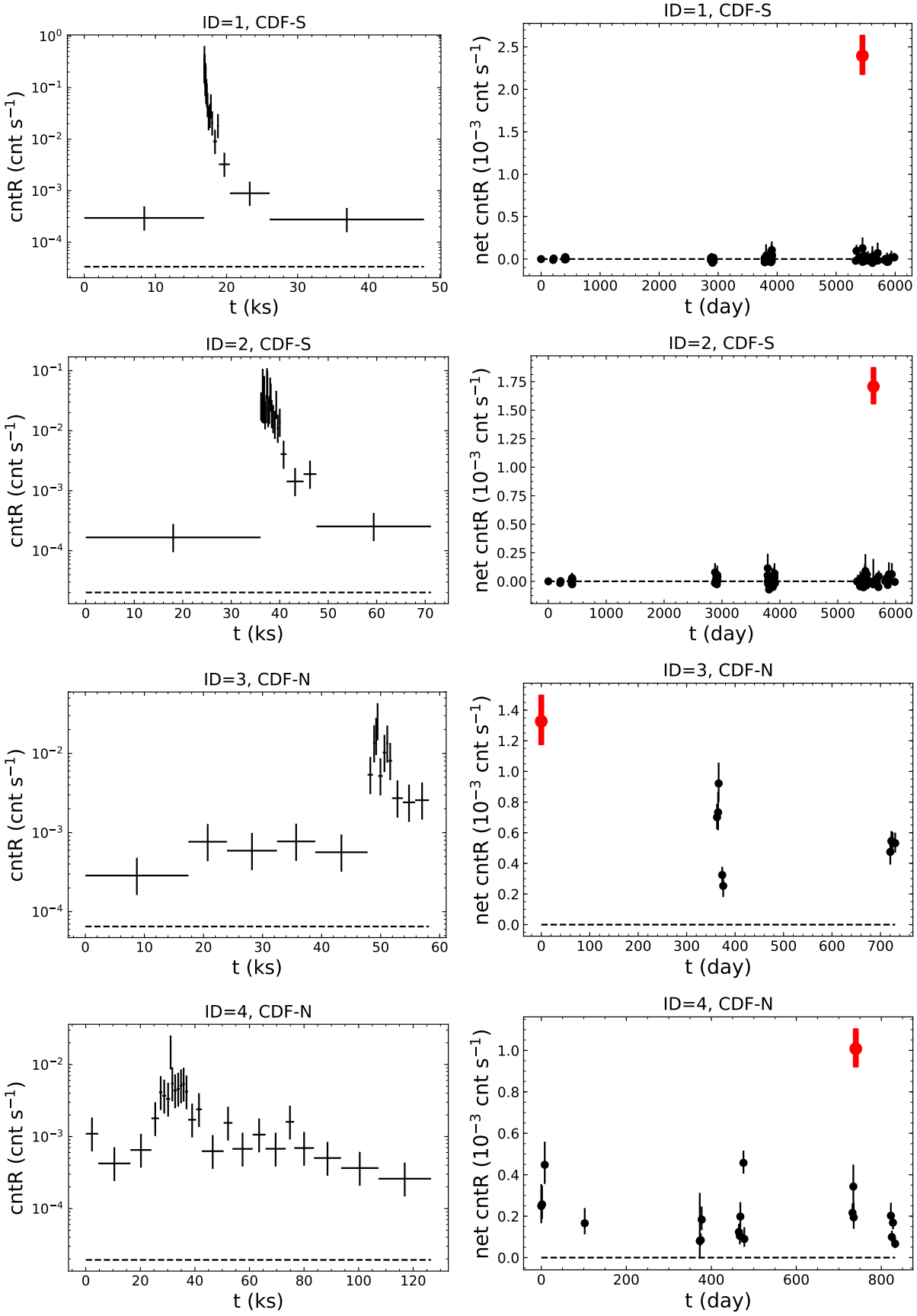


**Figure 6.** Light curves for each transient candidate. The left panels are light curves for the observation with the transient, with each bin including 5 counts. The horizontal dashed lines indicate the estimated average background count rates. The right panels are long-term light curves with each data point representing a *Chandra* observation. The transient observation is highlighted in red color. The horizontal dashed lines indicate a net count rate of zero.

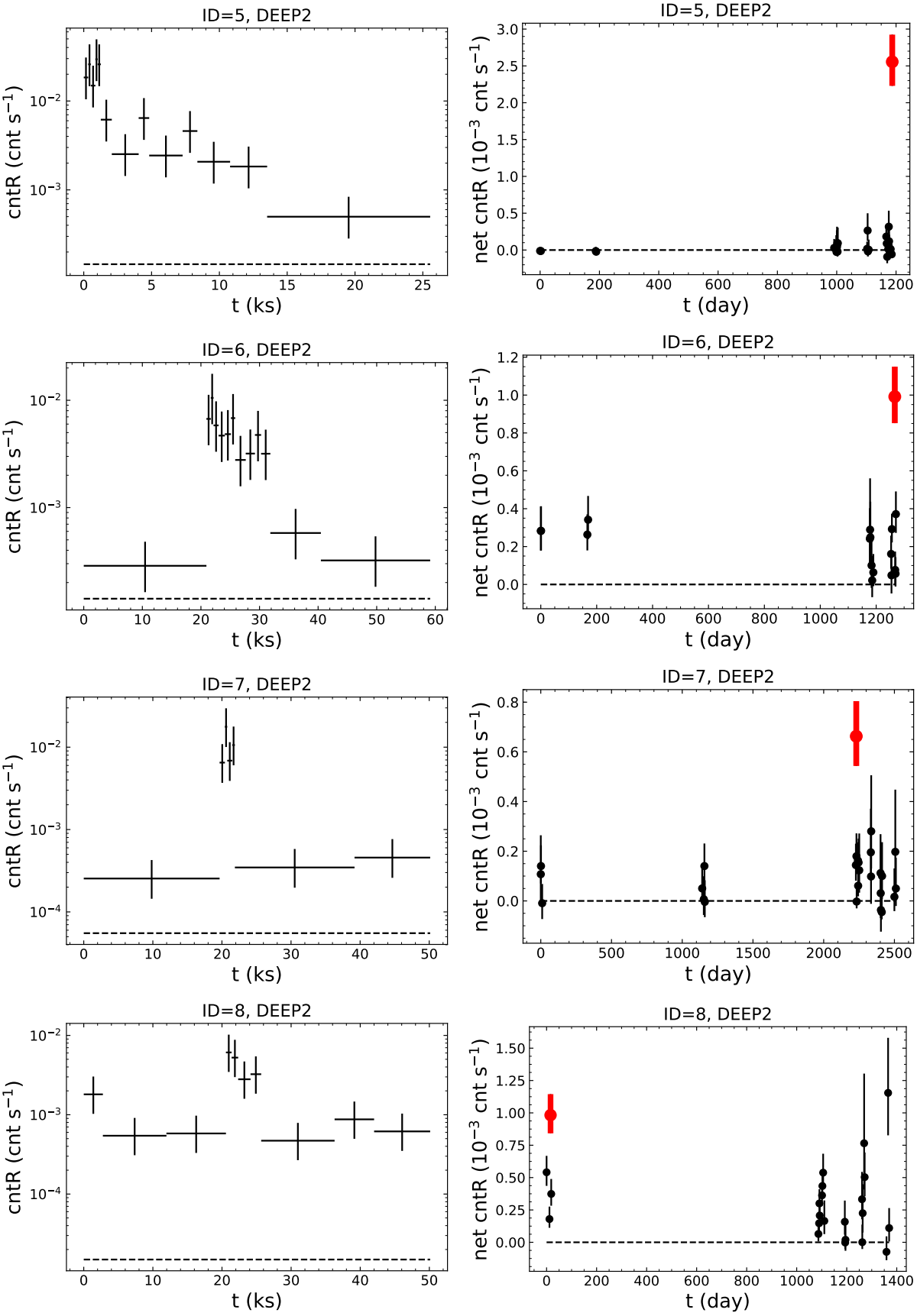


**Figure 6 (Continued).**

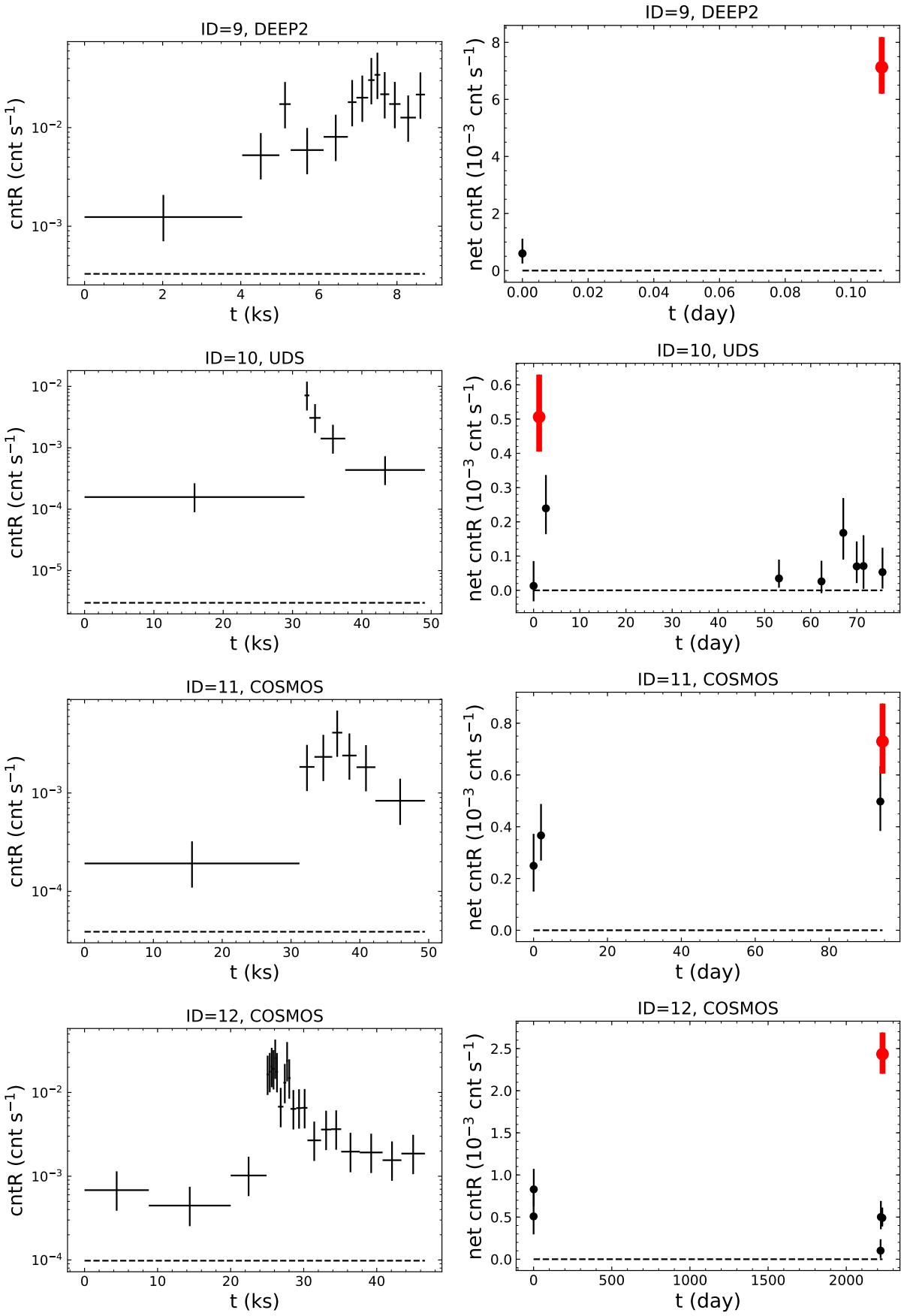


**Figure 6 (Continued).**

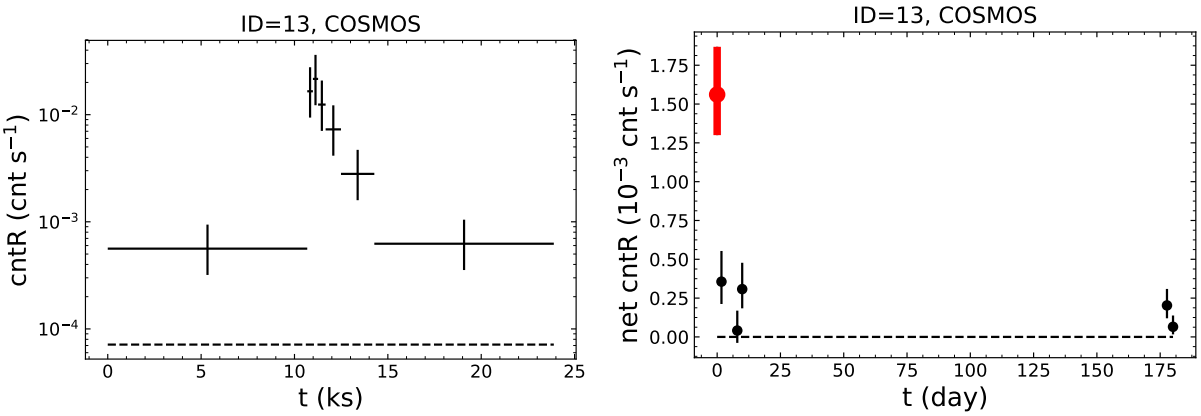


**Figure 6 (Continued).**

main scope of this paper is to study extragalactic transients similar to CDF-S XT1 and XT2, we do not further classify the stellar objects into, e.g. "transient stars" vs. "variable stars".

Since our algorithm is optimized for selecting CDF-S XT-like transients (see §2.2), the fact that only two such transients are found indicates such events are relatively rare. We further estimate the CDF-S XT-like event rate in §4. The prevalence of stars among our transient candidates is likely because stellar flares are intrinsically more common than CDF-S XT-like extragalactic transients, and it does not necessarily indicate that our algorithm is more sensitive in selecting stellar flares. There should be even more stellar flares in the survey data not identified by our algorithm, which is designed to select XT-like transients rather than stellar flares. In fact, we have tested adjusting our algorithm slightly, and the resulting stellar sample changes while the extragalactic sample remains the same. For example, if we chop the exposures to $t_{\rm exp} < 70$ ks instead of $t_{\rm exp} < 50$ ks (§2.1), CDF-S XT1 and XT2 will be still identified. However, this change will select 6 new stellar flares while missing 3 old stellar flares.

## 4 EVENT RATE AND FUTURE PROSPECTS

Our transient-search algorithm is able to find CDF-S XT-like transients with $\log F_{\rm peak} \gtrsim -12.6$ (cgs) effectively (§2.2.3). We remind that the limiting peak flux here is estimated for a typical off-axis angle of 5′ (see §2). For an off-axis angle of 0.5′ (nearly on-axis) and 8′ (the maximum value accepted by our algorithm; §2.1), the limiting flux changes slightly ($\approx 0.1$ dex; see Appendix B). However, we do not find any new extragalactic transients that are similar to CDF-S XT1 and XT2, despite searching *Chandra* observations totaling 19 Ms exposure (§3.3). Based on this search result, we estimate the event rate of CDF-S XT-like transients in §4.1. From the estimated event rate, we discuss the prospects of future missions (*Athena* and *Einstein Probe*) in detecting CDF-S XT-like transients.

### 4.1 Event-Rate Estimation

Since our simulations in §2.2.3 show that the efficiency of our transient selection in short *Chandra* exposures ($t_{\rm exp} \lesssim 8$ ks) is low, we do not include exposures shorter than 8 ks in when estimating the event rate below. These short exposures only add up to 0.022 Ms of observation time in total, which is negligible compared to the total observation time analyzed (19.3 Ms).

For a set of *Chandra* observations, the expected number of transients brighter than the flux limit ($\log F_{\rm peak} \gtrsim -12.6$, cgs) can be written as

$$\mathcal{N} = \sum_i R_i \Omega_i t_i, \tag{4}$$

where $R_i$ is the event rate; $\Omega_i$ and $t_i$ are the field of view (FOV) and exposure time, respectively; the subscript ($i$) denotes different exposures. In general, $R_i$ is a function of the sky coordinate of the telescope pointing. However, considering that our focus is extragalactic transients and the Universe is largely isotropic, we assume that $R_i$ is a constant and denote it as $R$. $\Omega_i$ depends on the instrument used. All of our analyzed survey data are from *Chandra*/ACIS-I imaging observations, and thus $\Omega_i$ is a constant and we denote it as $\Omega = \pi \times (8')^2 = 201$ arcmin$^2$. Eq. 4 can then be simplified as

$$\mathcal{N} = \sum_i R\Omega t_i = R\Omega \sum_i t_i, \tag{5}$$

i.e. $\mathcal{N}$ only depends on the total exposure time of these observations. In other words, it does not matter whether our analyzed 19 Ms of data are from a single sky zone or multiple sky zones. From Eq. 6, the event rate $R$ can be calculated as

$$R = \frac{\mathcal{N}}{\Omega \sum_i t_i} \tag{6}$$

Based on the fact that 2 events are detected in 19 Ms of data, we estimate $R \approx 59^{+77}_{-38}$ evt yr$^{-1}$deg$^{-2}$, where the uncertainties are Poisson $1\sigma$ errors, calculated with the ASTROPY.STATS package. We stress that the event rate estimated throughout this paper refers to that of a particular type of transients (i.e. similar to CDF-S XT1 and XT2 with $\log F_{\rm peak} \gtrsim -12.6$, cgs) rather than general extragalactic transients.

Given the event rate estimated above, we can estimate

**Table 3.** Counterpart Properties of Transient Candidates

| ID (1) | Source (2) | RAc (3) | DECc (4) | Offset (5) | $\mathrm{Mag}_z$ (6) | $z$ (7) | $z$ type (8) | *Gaia* (9) |
|---|---|---|---|---|---|---|---|---|
| 1 | CANDELS | 53.16157 | −27.85936 | 0.07″ | 27.9 | 2.14 | phot | n/a |
| 2 | CANDELS | 53.07659 | −27.87329 | 0.50″ | 24.5 | 0.74 | spec | n/a |
| 3 | WIRCam | 189.02037 | 62.33728 | 0.14″ | 16.8 | 0.00 | spec | star |
| 4 | CANDELS | 189.10575 | 62.23467 | 0.21″ | 16.4 | 0.00 | spec | star |
| 5 | DEEP2-1 | 215.07411 | 53.10657 | 0.26″ | 19.6 | 0.00 | spec | n/a |
| 6 | DEEP2-1 | 214.95966 | 52.74351 | 1.10″ | 14.1 | n/a | n/a | star |
| 7 | DEEP2-1 | 214.61031 | 52.54338 | 0.62″ | 17.1 | 0.00 | spec | n/a |
| 8 | DEEP2-1 | 214.66805 | 52.66666 | 0.33″ | 16.8 | n/a | n/a | star |
| 9 | DEEP2-2 | 252.12746 | 34.96339 | 0.45″ | 15.8 | 0.00 | spec | star |
| 10 | HSC | 34.48311 | −5.09118 | 0.25″ | 18.1 | 0.00 | spec | star |
| 11 | UltraVISTA | 149.75412 | 2.14183 | 0.38″ | 16.7 | 0.00 | spec | star |
| 12 | UltraVISTA | 149.82649 | 2.71803 | 0.41″ | 15.8 | 0.00 | spec | star |
| 13 | UltraVISTA | 149.99794 | 2.77960 | 0.40″ | 16.5 | 0.00 | spec | n/a |

NOTE. — (1) Transient ID in this work. (2) Source of the counterpart: CANDELS (Grogin et al. 2011; Koekemoer et al. 2011), WIRCam (Wang et al. 2010), DEEP2 (Coil et al. 2004), HSC (Aihara et al. 2018), and UltraVISTA (Laigle et al. 2016). (3) and (4) The position of the optical/NIR counterpart. (5) The distance between the X-ray position and the counterpart. (6) $z$-band AB magnitude of the counterpart. For ID1 and ID2, the $z$-band filter refers to *HST* F850LP; for other sources, the filter refers to SDSS $z$. (7) and (8) redshift and its type. "0.00" means stellar object. "n/a" means redshift unavailable. $z$ = 0.00 and $z$ type = phot mean the source's SED prefers a stellar template rather than a quasar/galaxy template. For ID7, we adopt the redshift from SDSS, since redshift information is not provided in the X-ray catalog (Goulding et al. 2012). (9) *Gaia* classification. "star" indicates the source has non-zero parallax and/or proper motion (S/N > 5) measured from *Gaia*; otherwise, "n/a" is listed.

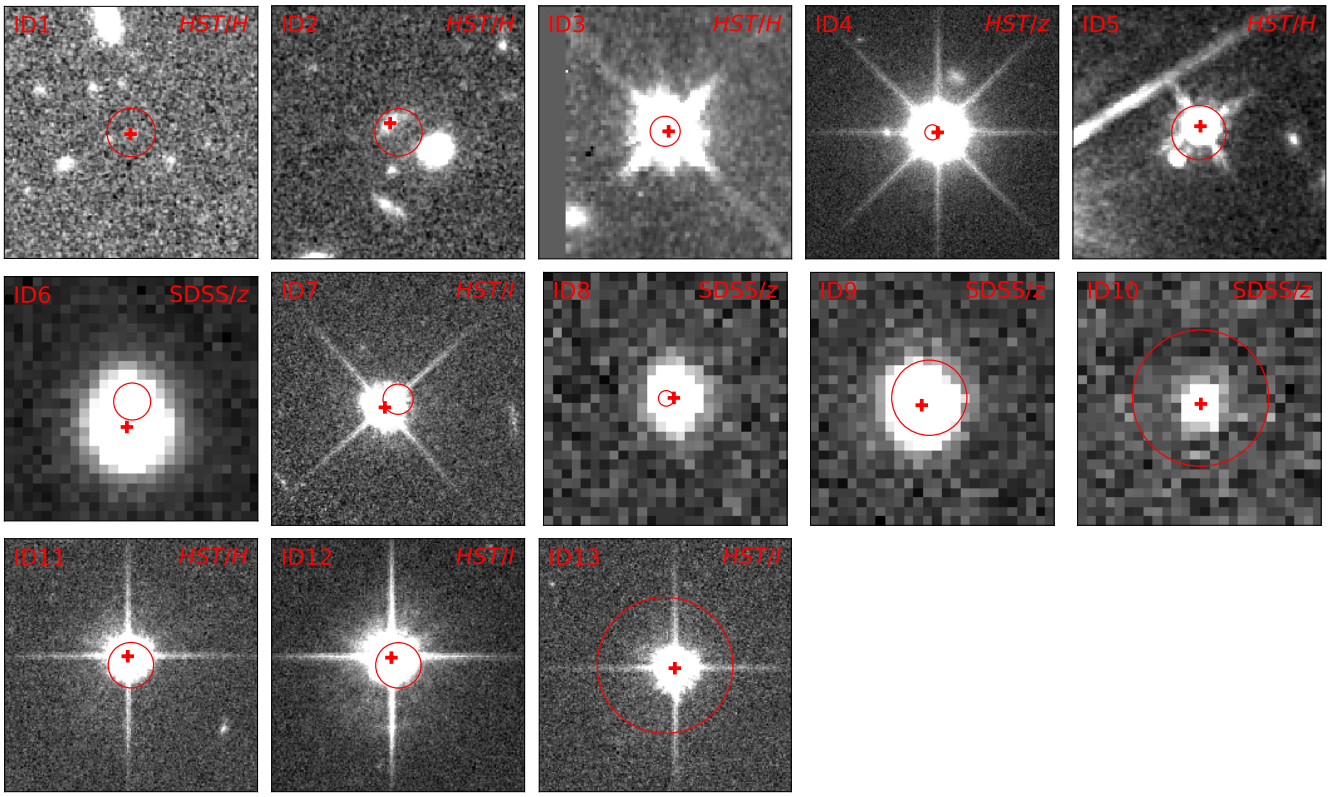


**Figure 7.** Optical/NIR image 10″ × 10″ cutouts of the transients. Each cutout is centered at the X-ray position. The central red circle denotes the X-ray positional uncertainty, and has a radius $3 \times \Delta X$, where $\Delta X$ is the $1\sigma$ X-ray positional error listed in Table 2. The red cross marks the position of the optical/NIR counterpart (Table 3). The cutouts are from the *HST* bands (as labelled) when available or the SDSS $z$ band. The *HST* cutouts are from Koekemoer et al. (2007) and the *Hubble* Legacy Archive (https://hla.stsci.edu/). The X-ray and optical/NIR positions are generally consistent with each other.

the number of CDF-S XT-like transients potentially existing in the *Chandra* archive. As of March 2019, there are 95 Ms and 94 Ms of ACIS-I and ACIS-S archival imaging observations (excluding < 8 ks exposures) at Galactic latitudes of $|b| > 20°$.[8] ACIS-I and ACIS-S consist of 4 and 6 CCD chips, respectively. For ACIS-I, all the chips are front-illuminated (FI); for ACIS-S, 4 and 2 chips are FI and back-illuminated (BI), respectively. The BI chips have a slightly higher ($\approx 10\%$) flux-to-counts conversion factor than the FI chips.[9] The former have a higher background ($\approx 2$ times) than the latter, but still at a low level (only $\approx 6$ counts for a 50 ks exposure). After considering these differences in conversion factor and background in our simulations (§2.2), we find the flux limits of our transient detection are similar for FI and BI chips ($\log F_{\rm lim} \approx -12.6$ for both). Therefore, the differences between the FI and BI chips should not affect our estimation of the transient number in *Chandra* archival observations below.

As for ACIS-I, we only account for the CCD area with off-axis angle < 8′ for ACIS-S, which covers the S2, S3, and S4 CCD chips. However, unlike the case for ACIS-I generally, ACIS-S may have some chips turned off during an observation (S3 is always on as it covers the aimpoint). When one (S3), two (S3+S4 or S2+S3), and three (S2+S3+S4) relevant chips are on, the CCD areas are $\Omega_{\rm S,1} \approx 69$ arcmin$^2$, $\Omega_{\rm S,2} \approx 96$ arcmin$^2$, and $\Omega_{\rm S,3} \approx 123$ arcmin$^2$, respectively. The total exposure times for the three cases are $T_{\rm S,1} \approx 16$ Ms, $T_{\rm S,1} \approx 25$ Ms, $T_{\rm S,1} \approx 53$ Ms, respectively. Therefore, we can estimate the total number of CDF-S XT-like transients in these archival observations as

$$\mathcal{N} = R(\Omega_{\rm I} T_{\rm I} + \Omega_{\rm S,1} T_{\rm S,1} + \Omega_{\rm S,2} T_{\rm S,2} + \Omega_{\rm S,2} T_{\rm S,3}) = 15^{+20}_{-10}, \quad (7)$$

where $\Omega_{\rm I}$ ($\Omega_{\rm S}$) and $T_{\rm I}$ ($T_{\rm S}$) are the FOV and total exposure time ($|b| > 20°$) of ACIS-I (ACIS-S) in the *Chandra* archive. We note that, at $|b| > 20°$, Galactic absorption is typically low, with column density of $N_{\rm H} \lesssim 10^{21}$ cm$^{-2}$ (e.g. Stark et al. 1992), and such absorption only reduces the observed flux by $\lesssim 10\%$ (estimated with PIMMS). Therefore, Galactic absorption is unlikely to significantly affect the estimated number of transients above.

We will perform an extensive *Chandra* archival search in a separate paper (Quirola Vásquez et al. in prep.). From our results (§3.3), the stellar objects found in archival data are likely to have bright optical/NIR counterparts ($z$-band magnitudes $\lesssim 20$), and thus their stellar nature can be largely determined with current wide-field surveys, e.g. SDSS, UKIDSS, and *Gaia*. In contrast, the counterparts of extragalactic transients will likely be faint in the optical/NIR, and follow-up observations with large ground-based telescopes will be helpful to study their properties such as redshift and host-galaxy stellar mass. These counterparts may also be studied with future deep wide-field surveys such as LSST (Ivezić et al. 2019) and *Euclid* (Laureijs et al. 2011). *XMM-Newton* has had a similar operational time as *Chandra*, and it notably has a larger effective area and FOV but also higher background than *Chandra*. Future work could also search *XMM-Newton* archival data for CDF-S XT-like transients (e.g. the EXTraS project; De Luca et al. 2016).

### 4.2 The Perspectives for Future Missions

Future X-ray missions such as *Athena* and *Einstein Probe* should be able to discover a large number of extragalactic transients similar to CDF-S XT1 and XT2. Now, we estimate the sample sizes of transients that will be potentially detected by *Athena* and *Einstein Probe*. As a first-order approximation, we assume that the event-rate density (event rate per dex of flux) is a power-law function, i.e.

$$\frac{dR}{d\log F_{\rm peak}} \propto F_{\rm peak}^{-\gamma} \quad (8)$$

Here, the power-law index ($\gamma$) is positive, because otherwise the event rate above a given $F_{\rm peak}$ would be divergent. By integrating Eq. 8 from $\log F_{\rm lim}$ (limiting peak flux of the mission)[10] to $\infty$ and applying Eq. 4, we can estimate the number of detected CDF-S XT-like transients as

$$\begin{aligned} \mathcal{N} &\propto F_{\rm lim}^{-\gamma} \Omega T \\ &\propto F_{\rm lim}^{-\gamma+1} \frac{\Omega}{F_{\rm lim}} T \\ &\propto A^{\gamma-1} (\Omega A) T \\ &\propto A^{\gamma-1} G T, \end{aligned} \quad (9)$$

where $A$ and $G$ are the effective area and grasp (defined as $\Omega \times A$) of the mission. In Eq. 9, we adopt the approximation of $F_{\rm lim} \propto A^{-1}$. If further assuming $\gamma = 1$ and $T$ is similar for different missions, we have $N \propto G$. Since both *Athena* and *Einstein Probe* have $G$ values $\sim 200$ times larger than that of *Chandra* (e.g. Nandra et al. 2013; Burrows et al. 2018; Yuan et al. 2018) which can detect $\sim 15$ transients (see above), we expect that *Athena* and *Einstein Probe* will each detect $\sim 3000$ sources if they operate for $\approx 20$ years. These samples will be sufficiently large for detailed sample studies. Note that the estimated sample sizes depend on the assumption that $\gamma = 1$. If $\gamma > 1$, *Athena* (*Einstein Probe*) will detect more (fewer) transients; if $0 < \gamma < 1$, the situation is the opposite.

Our estimation above is based on the assumption of a power-law function of event-rate density (Eq. 8) with $\gamma = 1$. A natural prediction of this power-law function is that there are more faint sources than bright sources in general. One might be concerned that this prediction contradicts our results, i.e. the 19 Ms of *Chandra* data only contains two relatively bright sources (XT1 and XT2, both having $\log F_{\rm peak} > -12.2$; see Table 2) but no fainter sources.

[8] Here, we do not consider the 7 Ms of observations performed by HRC, because the sensitivities and thereby flux limits of HRC and ACIS are different. We also do not include ACIS subarray-mode observations to avoid complexity in the calculation of FOV. Such observations only contribute 1% and 17% of the exposure time for ACIS-I and ACIS-S, respectively. Accounting for these observations is technically challenging, but would only affect our estimated transient number by a few percent at most.

[9] http://cxc.harvard.edu/proposer/POG/html/chap6.html

[10] Here, we integrate from $\log F_{\rm lim}$ rather than −12.6 (the limiting flux of *Chandra*; §2.2.3). This is because, in this Section, our goal is to estimate the number of XT-like transients detectable by future missions (i.e. sources with $F_{\rm peak}$ above $F_{\rm lim}$ of these missions). Therefore, the integration lower limit should be $F_{\rm lim}$ of the mission of interest.

We now test whether this apparent inconsistency is statistically significant or not. Assuming there are two transients above the *Chandra* flux limit ($\log F_{\rm lim} = -12.6$) detected in the 19 Ms of data, we estimate the chance for these two both to have $\log F_{\rm peak} > -12.2$. From Eq. 8 ($\gamma = 1$), the probability for one detected transient to be bright ($\log F_{\rm peak} > -12.2$) is

$$P_{\rm bright} = \frac{\int_{-12.2}^{\infty} F_{\rm peak}^{-1} d\log F_{\rm peak}}{\int_{-12.6}^{\infty} F_{\rm peak}^{-1} d\log F_{\rm peak}} = 0.40. \quad (10)$$

Then, according to the binomial distribution, the probability ($p$-value) for both sources to be bright ($\log F_{\rm peak} > -12.2$) is $0.40^2 = 0.16$, only corresponding to $1.4\sigma$ significance. Therefore, the assumption of Eq. 8 ($\gamma = 1$) does not contradict our results significantly. Actually, we find Eq. 8 is always consistent with our results at a $3\sigma$ level, as long as $0 < \gamma < 3.1$.

## 5 SUMMARY

We have performed a systematic search for CDF-S XT-like extragalactic transients in 19 Ms of *Chandra* surveys, including CDF-S, CDF-N, DEEP2, UDS, COSMOS, and E-CDF-S. Our main results are summarized below.

(i) We developed a method to select transients within a *Chandra* observation (§2). From simulations, we show that our method is efficient in discovering transients with 0.5–7 keV peak flux $\log F_{\rm peak} \gtrsim -12.6$ (erg cm$^{-2}$ s$^{-1}$).

(ii) Our selection yields 13 transient candidates (§3), including CDF-S XT1 and XT2 which have been reported in previous works (Bauer et al. 2017; Xue et al. 2019). All the candidates have optical/NIR counterparts (§3.3). Except for CDF-S XT1 and XT2, all other sources are stellar objects.

(iii) The lack of new CDF-S XT-like transients in our search indicates that such objects are rare (§4). We estimate an event rate of $59^{+77}_{-38}$ evt yr$^{-1}$deg$^{-2}$, corresponding to a total of $15^{+20}_{-10}$ events in *Chandra* archival observations at $|b| > 20°$. Future X-ray missions such as *Athena* and the *Einstein Probe* with large grasps might be able to find thousands of extragalactic transients, and sample studies will be feasible then.

## ACKNOWLEDGEMENTS

We thank the referee for helpful feedback that improved this work. We thank David Burrows, Qingling Ni, John Timlin, and Fabio Vito for helpful discussions. GY, WNB, and SFZ acknowledge support from CXC grant AR8-19016X, CXC grant AR8-19011X, and NASA ADP grant 80NSSC18K0878. FEB acknowledges support from CONICYT-Chile (Basal AFB-170002, FONDO ALMA 31160033) and the Ministry of Economy, Development, and Tourism's Millennium Science Initiative through grant IC120009, awarded to The Millennium Institute of Astrophysics, MAS. YQX acknowledges support from the 973 Program (2015CB857004), NSFC (11890693, 11421303), and the CAS Frontier Science Key Research Program (QYZDJ-SSW-SLH006). The Guaranteed Time Observations (GTO) for the CDF-N included here were selected by the ACIS Instrument Principal Investigator, Gordon P. Garmire, currently of the Huntingdon Institute for X-ray Astronomy, LLC, which is under contract to the Smithsonian Astrophysical Observatory; Contract SV2-82024. This project uses Astropy (a Python package; see Astropy Collaboration et al. 2018).

## REFERENCES

Aihara H., et al., 2018, PASJ, 70, S8
Astropy Collaboration et al., 2018, preprint, (arXiv:1801.02634)
Bauer F. E., et al., 2017, MNRAS, 467, 4841
Belloni T. M., Stella L., 2014, Space Sci. Rev., 183, 43
Brandt W. N., Alexander D. M., 2015, A&ARv, 23, 1
Burrows D. N., et al., 2018, in Space Telescopes and Instrumentation 2018: Ultraviolet to Gamma Ray. p. 106991J (arXiv:1808.02883), doi:10.1117/12.2312785
Civano F., et al., 2016, ApJ, 819, 62
Coil A. L., Newman J. A., Kaiser N., Davis M., Ma C.-P., Kocevski D. D., Koo D. C., 2004, ApJ, 617, 765
De Luca A., Salvaterra R., Tiengo A., D'Agostino D., Watson M. G., Haberl F., Wilms J., 2016, in Napolitano N. R., Longo G., Marconi M., Paolillo M., Iodice E., eds, Vol. 42, The Universe of Digital Sky Surveys. p. 291 (arXiv:1503.01497), doi:10.1007/978-3-319-19330-4_46
Evans I. N., et al., 2010, ApJS, 189, 37
Gaia Collaboration et al., 2018, A&A, 616, A1
Gallo L., 2018, in Revisiting narrow-line Seyfert 1 galaxies and their place in the Universe. 9-13 April 2018. Padova Botanical Garden. p. 34 (arXiv:1807.09838)
Glennie A., Jonker P. G., Fender R. P., Nagayama T., Pretorius M. L., 2015, MNRAS, 450, 3765
Goulding A. D., et al., 2012, ApJS, 202, 6
Grogin N. A., et al., 2011, ApJS, 197, 35
Güdel M., Nazé Y., 2009, Astronomy and Astrophysics Review, 17, 309
Haisch B., Strong K. T., Rodono M., 1991, ARA&A, 29, 275
Ivezić Ž., et al., 2019, ApJ, 873, 111
Kara E., Miller J. M., Reynolds C., Dai L., 2016, Nature, 535, 388
Kocevski D. D., et al., 2018, ApJS, 236, 48
Koekemoer A. M., et al., 2007, ApJS, 172, 196
Koekemoer A. M., et al., 2011, ApJS, 197, 36
Komossa S., 2015, Journal of High Energy Astrophysics, 7, 148
Krishnamoorthy K., Thomson J., 2004, Journal of Statistical Planning and Inference, 119, 23
Laigle C., et al., 2016, ApJS, 224, 24
Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193
Lawrence A., et al., 2007, MNRAS, 379, 1599
Luo B., et al., 2017, ApJS, 228, 2
Marchesi S., et al., 2016, ApJ, 817, 34
Markowitz A., et al., 2003a, ApJ, 593, 96
Markowitz A., Edelson R., Vaughan S., 2003b, ApJ, 598, 935
Nandra K., et al., 2013, arXiv e-prints, p. arXiv:1306.2307
Nandra K., et al., 2015, ApJS, 220, 10
Paolillo M., et al., 2017, MNRAS, 471, 4398
Park T., Kashyap V. L., Siemiginowska A., van Dyk D. A., Zezas A., Heinke C., Wargelin B. J., 2006, ApJ, 652, 610
Pooley D., Kumar P., Wheeler J. C., Grossan B., 2018, ApJ, 859, L23
Ricci C., et al., 2016, ApJ, 820, 5
Stark A. A., Gammie C. F., Wilson R. W., Bally J., Linke R. A., Heiles C., Hurwitz M., 1992, ApJS, 79, 77

Vito F., et al., 2016, MNRAS, 463, 348
Wang W.-H., Cowie L. L., Barger A. J., Keenan R. C., Ting H.-C., 2010, ApJS, 187, 251
Xue Y. Q., 2017, New Astron. Rev., 79, 59
Xue Y. Q., Luo B., Brandt W. N., Alexander D. M., Bauer F. E., Lehmer B. D., Yang G., 2016, ApJS, 224, 15
Xue Y. Q., et al., 2019, Nature, 568, 198
Yang G., et al., 2016, ApJ, 831, 145
York D. G., et al., 2000, AJ, 120, 1579
Yuan W., et al., 2018, Scientia Sinica Physica, Mechanica and Astronomica, 48, 039502
Zheng X. C., et al., 2017, ApJ, 849, 127
van der Klis M., 1989, Annual Review of Astronomy and Astrophysics, 27, 517

## APPENDIX A: EFFICIENCY OF THE SELECTION ALGORITHM FOR DIFFERENT TRANSIENT MODELS

The simulations in §2.2 are based on a fiducial transient model similar to the CDF-S XTs. The employment of this fiducial model is driven by the main aim of this paper, i.e. investigating CDF-S XT-like transients in *Chandra* surveys. However, our algorithm might also be able to identify other types of transients as a "bonus". In this Appendix, we perform Monte Carlo simulations for some other transient models as examples, although pursuing them is not the main focus of our paper.

The first additional transient model we test is a "time-reversed" version of our fiducial model (see Fig. A1 top for the light curve). The fiducial light curve has features of a fast rise and slow decline (Fig. 1), and thus the reverse has features of a slow rise and fast decline. The reversed model has the same flux-to-counts conversion factor and timescale as the fiducial model. We then apply the simulation process in §2.2.2 to the reversed model, and show $P_{\rm eff}$ as a function of $t_{\rm exp}$ in Fig. A1 (bottom). The simulation results are similar to those of the fiducial model, e.g. for $\log F_{\rm peak} \gtrsim -12.6$ (cgs, corresponding to $\approx 30$ counts), $P_{\rm eff}$ is $\approx 1$ for a wide range of $t_{\rm exp} = 8$–$50$ ks. We have also tested some other light curves with different shapes but similar timescales, and found the sensitivity of our algorithm for these models is similar to the fiducial model. These results indicate that our algorithm is also capable of detecting different types of transients with timescales similar to that of the CDF-S XTs.

Another additional transient model we test is based on the ultrafast transient discovered by Glennie et al. (2015). This transient lasts only $\approx 100$ s with $\log F_{\rm peak} = -9.9$ (cgs), and has a spectral shape of $\Gamma \approx 1.4$. The nature of the transient remains unknown, as the optical/NIR counterpart has not been found due to the lack of deep multiwavelength data (§1). The light curve can also be approximated by the general formula in Eq. 1, with $(t_1, t_2, \alpha_1, \alpha_2) \approx (10\,{\rm s}, 30\,{\rm s}, 0, -4)$. This light-curve model is displayed in Fig. A2 (top). The flux-to-counts conversion factor (Eq. 2) for this model is $3.2 \times 10^{12}$, and the $T_{90}$ is 47 s. Here, the conversion factor is much lower than that in Eq. 2. This is mainly because the ultrafast model has a timescale much shorter than the fiducial model, and to reach similar counts, the former must have a much higher peak flux than the latter. We show the simulation results in Fig. A2 (bottom). Unlike $P_{\rm eff}$ in Fig. 1, $P_{\rm eff}$ in Fig. A2 does not drop below $t_{\rm exp} \approx 8$ ks. The drop in Fig. 1 is because, when the exposure time becomes shorter than the transient timescale, the observed light curve will be similar to a normal variable source (§2.2.3). However, this is not the case in Fig. A2, since the ultrafast-transient timescale ($T_{90} = 47$ s) is even shorter than our shortest exposures (3 ks). In Fig. A2, for $\log F_{\rm peak} \lesssim -11.1$, $P_{\rm eff}$ declines toward high $t_{\rm exp}$ due to high background levels for long exposures (§2.1). For $\log F_{\rm peak} \gtrsim -11.0$ (corresponding to $\approx 30$ counts), $P_{\rm eff}$ is stable for different $t_{\rm exp}$, because the X-ray signal is dominated by the source rather than the background.

Glennie's model tested above is faster than our fiducial model. Now, we test another transient model which is "slower" than the fiducial model. We extend the plateau phase of the fiducial model (§2.2.1) by setting $t_2 = t_1 + 5$ ks (Eq. 1), while keeping the other parameters the same. The light curve of this slower model is displayed in Fig. A3 (top). The flux-to-counts conversion factor (Eq. 2) for this model is $6.0 \times 10^{14}$, and the $T_{90}$ is 16.7 ks. The simulation results are displayed in Fig. A3 (bottom). For a given $F_{\rm peak}$, $P_{\rm eff}$ rises toward high $t_{\rm exp}$ for the aforementioned reason, i.e. our algorithm may not be able to differentiate the transient from normal variable sources when $t_{\rm exp} \lesssim$ transient timescale. Since most ($\approx 90\%$; §3.2) of our exposures are longer than the timescale of the slower model, our algorithm is largely capable of detecting such transients in our data.

In summary, our algorithm can detect different types of transients with timescales similar to or below that of the CDF-S XTs, as long as $\gtrsim 30$ counts are available. For transients with longer timescales, only observations with $t_{\rm exp} \gtrsim$ transient timescale can have high detection probabilities. Since 80% of our exposures are longer than 25 ks (§3.2), we are potentially able to detect transients with timescales shorter than $\approx 25$ ks in our data.

## APPENDIX B: EFFICIENCY OF SELECTION ALGORITHM AT DIFFERENT OFF-AXIS ANGLES

The simulations in §2.2 are performed for a typical off-axis angle of $5'$. In this Appendix, we perform simulations at off-axis angles of $0.5'$ (nearly on-axis) and $8'$ (the maximum value accepted by our algorithm; §2.1). In our simulation configurations (§2.2.1), there are two parameters dependent on off-axis angle, i.e. flux-to-counts conversion factor and background noise. The conversion factors (Eq. 2) are $\approx 1.7 \times 10^{14}$ and $\approx 1.5 \times 10^{14}$ (cgs) at $0.5'$ and $8'$, respectively; the typical background count rates are $5.9 \times 10^{-6}$ cnt s$^{-1}$ to $2.5 \times 10^{-4}$ cnt s$^{-1}$.

We perform our simulations under these new configurations, and display the results in Fig. B1. Similar to the results for $5'$, $P_{\rm eff}$ drops significantly below $t_{\rm exp} \approx 8$ ks, because short exposures cannot differentiate between variable sources and transients (§2.2.3). Compared to that for $5'$, $P_{\rm eff}$ for $0.5'$ ($8'$) generally increases (decreases) for a given $F_{\rm peak}$ and $t_{\rm exp}$, as expected. As a consequence, the peak-flux limit could change if using the simulation configurations for $0.5'$ ($8'$). In §2.2.3, we choose the peak-flux limit as the minimum flux above which $P_{\rm eff}$ is $\approx 1$ for $t_{\rm exp} = 8$–$50$ ks. Applying the same criteria to Fig. B1, the peak-flux limits are $\log F_{\rm peak} \approx -12.7$ ($\approx -12.5$) for $0.5'$ ($8'$).

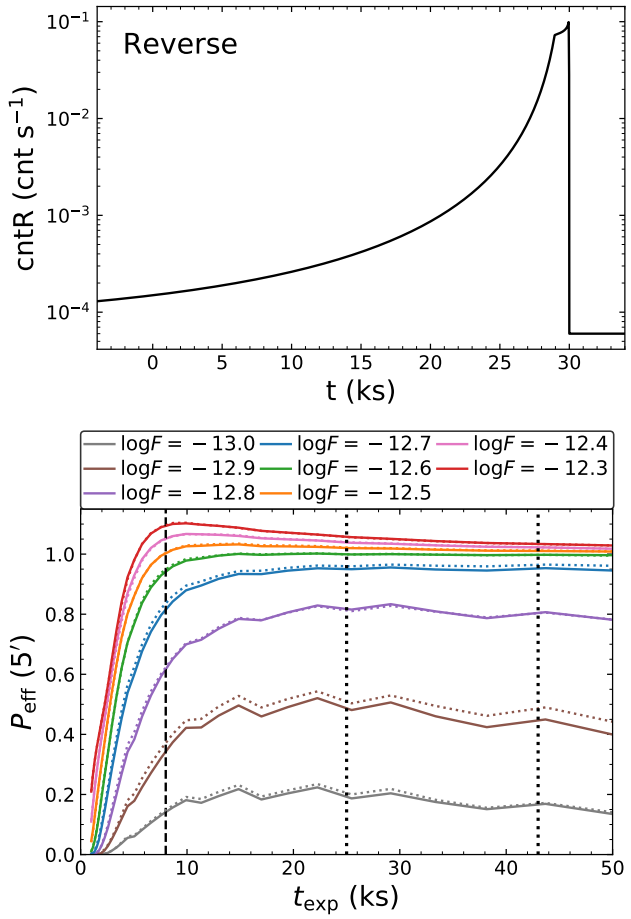


**Figure A1.** Top: Same format as Fig. 1 but for a time-reversed fiducial model. Bottom: Same format as Fig. 3 but for the time-reversed model in the top panel. For comparison, the $P_{\rm eff}$ for the fiducial model are also plotted as the dotted curves.

This paper has been typeset from a TEX/LATEX file prepared by the author.

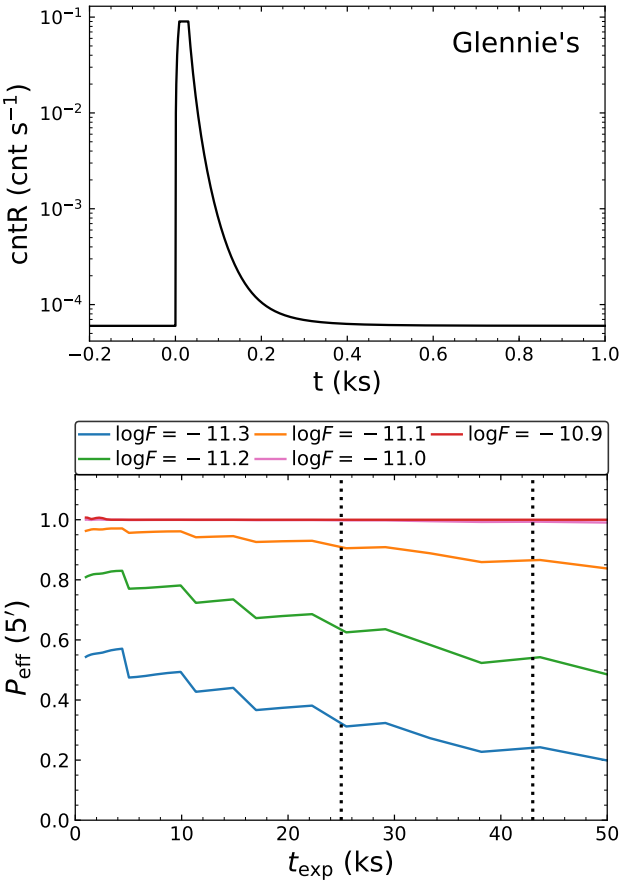


**Figure A2.** Top: Same format as Fig. 1 but for an ultrafast transient model similar to Glennie's event. Bottom: Same format as Fig. 3 but for the ultrafast model in the top panel.

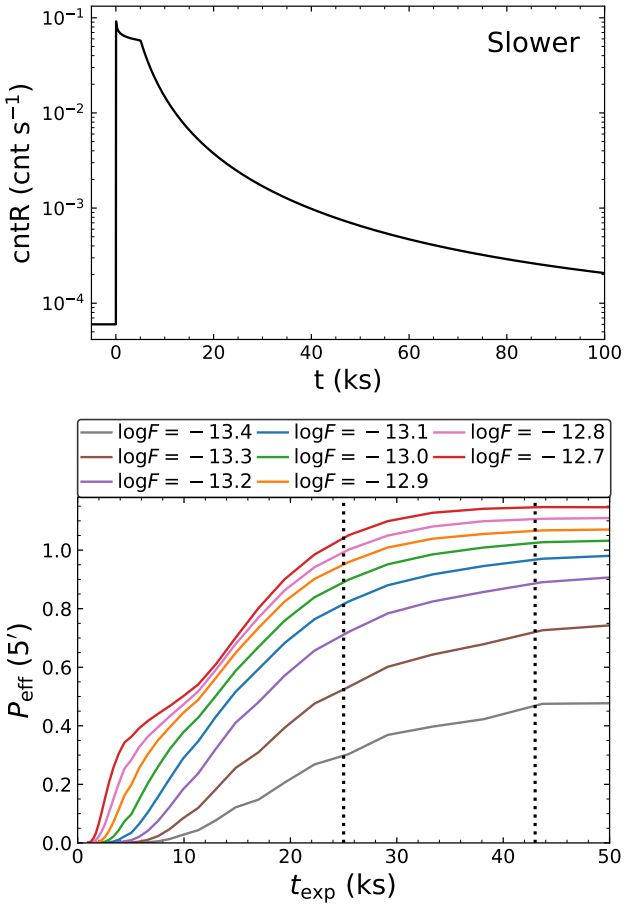


**Figure A3.** Top: Same format as Fig. 1 but for a slower transient model with $t_2 = t_1 + 5$ ks (Eq. 1). Bottom: Same format as Fig. 3 but for the slower model in the top panel.

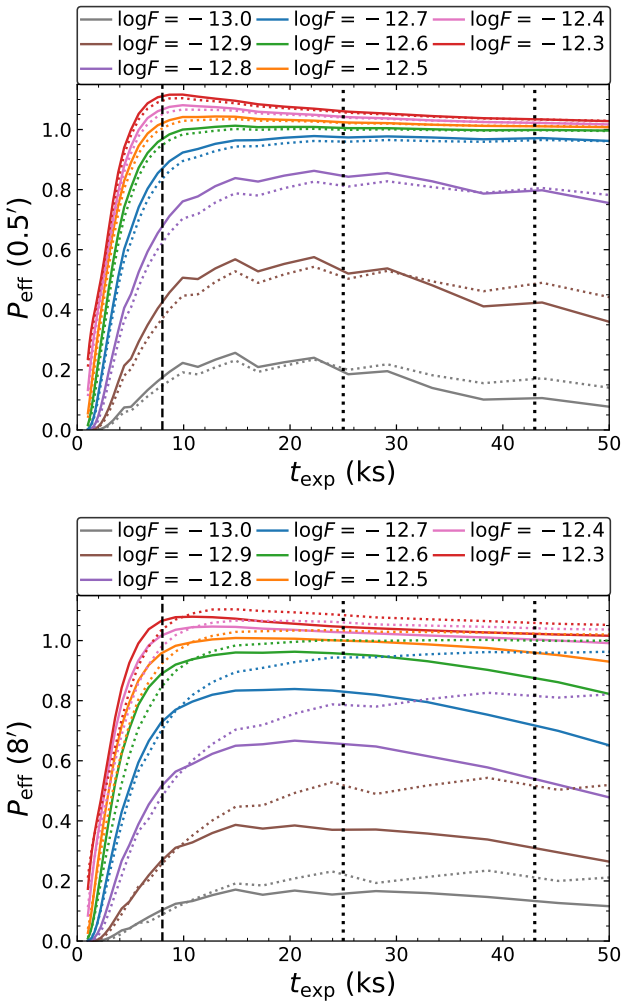


**Figure B1.** Same format as Fig. 3 but for off-axis angles of 0.5′ (top) and 8′ (bottom). For comparison, the $P_{\rm eff}$ for 5′ are also plotted as the dotted curves.