X-Ray Superflares from Pre-main-sequence Stars: Flare Energetics and Frequency

Konstantin V. Getman1 and Eric D. Feigelson1,2

1 Department of Astronomy & Astrophysics, Pennsylvania State University, 525 Davey Laboratory, University Park, PA 16802, USA; kug1@psu.edu
2 Center for Exoplanetary and Habitable Worlds, Pennsylvania State University, University Park, PA 16802, USA

Abstract

Solar-type stars exhibit their highest levels of magnetic activity during their early convective pre-main-sequence (PMS) phase of evolution. The most powerful PMS flares, superflares and megaflares, have peak X-ray luminosities of $\log(L_X) = 30.5 - 34.0$ erg s$^{-1}$ and total energies $\log(E_X) = 34 - 38$ erg. Among >24,000 X-ray-selected young ($t \lesssim 5$ Myr) members of 40 nearby star-forming regions from our earlier Chandra MYStIX and SFIIsNCs surveys, we identify and analyze a well-defined sample of 1086 X-ray superflares and megaflares, the largest sample ever studied. Most are considerably more powerful than optical/X-ray superflares detected on main-sequence stars. This study presents energy estimates of these X-ray flares and the properties of their host stars. These events are produced by young stars of all masses over evolutionary stages ranging from protostars to diskless stars, with the occurrence rate positively correlated with stellar mass. Flare properties are indistinguishable for disk-bearing and diskless stars indicating star–disk magnetic fields are not involved. A slope of $\alpha \simeq 2$ in the flare energy distributions $dN/dE_X \propto E_X^{-\alpha}$ is consistent with that of optical/X-ray flaring from older stars and the Sun. Megaflares ($\log(E_X) > 36.2$ erg) from solar-mass stars have an occurrence rate of $1.7 \pm 0.2$ flares/star/year and contribute at least 10%-20% to the total PMS X-ray energetics. These explosive events may have important astrophysical effects on protoplanetary disk photoevaporation, ionization of disk gas, production of spallogenic radionuclides in disk solids, and hydrodynamic escape of young planetary atmospheres. Our following paper details plasma and magnetic loop modeling of the >50 brightest X-ray megaflares.

Unified Astronomy Thesaurus concepts: Planetary atmospheres (1244); Pre-main sequence stars (1290); Protostars (1302); Protoplanetary disks (1300); Stellar x-ray flares (1637); X-ray stars (1823)

Supporting material: figure set, machine-readable tables.

1. Introduction

1.1. Pre-main-sequence Superflares and Megaflares

X-ray imaging studies of nearby star-forming regions, such as the Taurus clouds and Orion Nebula, typically show that highly variable X-ray emission is a ubiquitous characteristic of pre-main-sequence (PMS) stars (Feigelson & Decampli 1981; Montmerle et al. 1983; Getman et al. 2005; Güdel et al. 2007). The emission arising from magnetic reconnection events similar to, but much more powerful and frequent than, flares on the contemporary Sun. Reviews relating to PMS X-ray emission are provided by Feigelson & Montmerle (1999), Güdel (2004), Feigelson et al. (2007), Gregory et al. (2010), Stelzer (2017), Feigelson (2018), Scuortino et al. (2019), and Argiroffi (2019).

Though PMS X-ray flares were surprising at first, the existence of strong magnetic dynamos in the interiors of fully convective, rapidly rotating stars, followed by eruption of field lines and violent magnetic reconnection above the stellar surface, is reasonable. The X-ray emission seems to be independent of the presence or absence of protoplanetary disks despite astrophysical calculations that star–disk magnetic field lines may be involved in X-ray emitting flares (Hayashi et al. 1996; Shu et al. 1997; Aarons et al. 2010; Lopez-Santiago et al. 2016; Colombo et al. 2019). A factor of 2 reduced X-ray activity level in accreting versus nonaccreting systems (Flaccomio et al. 2003; Preibisch et al. 2005; Telleschi et al. 2007) could have several possible causes: cooling of active regions by accreting material, attenuation of X-rays by accreting columns and/or inner disks, coronal stripping by disks, and distortion of magnetic topologies by disks/accretion (Flaccomio et al. 2003; Jardine et al. 2006; Gregory et al. 2007; Getman et al. 2008b; Flaccomio et al. 2012). Accretion shocks contribute a small fraction to the total X-ray emission from T Tauri stars in the form of soft X-ray excess emission (Telleschi et al. 2007b).

Following previous researchers (Favata et al. 2005; Getman et al. 2008a; McCleary & Wolk 2011), we focus our attention here on the most luminous PMS X-ray flares with peak X-ray luminosities exceeding $L_{Xpk} = 10^{30.5}$ erg s$^{-1}$ and/or total (time-integrated) energies exceeding $E_X = 10^{34}$ erg. In contrast, no solar flare has been observed with total X-ray energy exceeding $\sim 10^{30}$ erg, four orders of magnitude below our threshold (Schrijver et al. 2012). We call events with $10^{34} < E_X < 10^{36.2}$ erg “superflares” and events with $E_X > 10^{36.2}$ erg “megaflares,” recognizing that the superflare designation is also used for less powerful optical flares seen with the Kepler satellite in older stars. In both solar and PMS flares, X-rays constitute only a minor fraction of the total radiated energy (Flaccomio et al. 2018), and the radiated energy may be dominated by the energy in ejected magnetic fields and energetic particles.

In the present effort, we examine the largest sample of X-ray PMS super- and megaflares ever collected to seek ensemble characteristics and relationships with other properties. The sample is drawn from observations of >24,000 PMS stars detected in 40 MYStIX (Massive Young Star-forming complex study in Infrared and X-rays; Feigelson et al. 2013) and SFIIsNCs (Star Formation in Nearby Clouds; Getman et al. 2017)
star-forming regions with NASA’s Chandra X-ray Observatory (Section 2).

1.2. Astrophysical Implications

Our studies are aimed at partially addressing some of the important questions concerning flare-related physics and phenomena:

1. Previous studies using much smaller samples of X-ray flares from PMS stars (Wolk et al. 2005; Stelzer et al. 2007; Caramazza et al. 2007; Albacete Colombo et al. 2007) report similar power-law slopes \( \alpha \sim 2 \) of flare energy distributions \( dN/dE_{\text{flare}} \propto E_{\text{flare}}^{-\alpha} \), consistent with those of older stars and the Sun. Such findings may have implications for understanding the relative importance of powerful flares and nano- or microflares for heating the solar and stellar coronas (Vilangot Nhalil et al. 2020, and references therein).

2. By taking advantage of the much increased superflare sample combined with a homogeneous set of derived flare host properties (such as stellar mass), we evaluate the flare occurrence rate as a function of flare energy and stellar mass, as well as the contribution of powerful flares to the total X-ray fluence of PMS stars. Such unique estimates will provide better understanding of the effects of PMS X-ray emission on their molecular enironments including their natal molecular cloud, the infalling envelope of protostars, the protoplanetary disk around T Tauri stars, and the protoplanets revealed after the disk has dissipated.

3. Super- and megaflares are important for both their high-fluency ionizing radiation and the production of hard X-rays that can penetrate deep into molecular environments (Glassgold et al. 2000). Flare X-rays potentially can produce layers of ionization in otherwise neutral material, induce nonequilibrium ion-molecular chemistry, and sputter grain surfaces. Even low levels of ionization can couple molecular material to magnetic fields resulting (in some circumstances) in turbulent motions and (in other circumstances) in bulk outflows. There is some empirical evidence that PMS flares heat disks and diminish accretion due to the photoevaporation of disks (Drake et al. 2009; Flaccomio et al. 2018; Flaischlen et al. 2021). They may be accompanied by energetic particles that could produce spallogenic radionuclides and by coronal mass ejection shocks that could melt ices or solids.

Flare radiation plays an important role in photoevaporative flows and dispersal of protoplanetary disks (reviewed by Williams & Cieza 2011; Alexander et al. 2014; Ercolano & Pascucci 2017) and young planetary atmospheres (Lammer et al. 2003; Ribas et al. 2005; Güdel 2007; Gronoff et al. 2020). There is particular concern that the effects of violent magnetic activity in young stars—extreme ultraviolet emission and coronal mass ejections as well as X-ray emission—can erode atmospheres of planets that otherwise might be habitable (Lammer et al. 2007; Gronoff et al. 2020; Atri & Carberry Mogan 2021). The effects of energetic particles and coronal mass ejections that may be associated with superflares are still uncertain (Drake et al. 2016; Atri 2020).

4. The geometry of PMS flare plasma seems remarkable. Models of X-ray evolution of PMS superflares are usually consistent with enormous loop structures often larger than the star itself (Favata et al. 2005; Getman et al. 2008b; Reale et al. 2018). Magnetospheric calculations for classical T Tauri stars indicate that closed magnetic loops on this scale can coexist with open magnetic field lines accreting from a protoplanetary disk (Johnstone et al. 2014, and references therein). However, it is difficult to exclude other magnetic geometries such as more complicated sequentially triggered arcades (Getman et al. 2011) or eclipsed loop geometries (Johnstone et al. 2012).

5. The astrophysics of flare plasma needs to be investigated. Are the heating and cooling processes similar to those of solar flares even when the emission measure is vastly greater? To address some of these issues, in our following paper (Getman et al. 2021) we perform detailed modeling of the brightest MYStIX/SFiNCs X-ray superflares from disky and diskless PMS stars of various masses, involving the evaluation of plasma temperature and emission measure temporal profiles, flare cooling timescales, coronal loop length and thickness, and further comparison of these properties with those of numerous X-ray flares from young stars in Orion Nebula (Getman et al. 2008a) and flares from older stars (Güdel 2004) and the Sun (Aschwanden et al. 2008).

6. The explosive and thermal processes within the magnetic structures producing super- and megflare plasma needs to be investigated. Are the heating and cooling processes similar to those of solar flares even when the emission measure is vastly greater? Some flares have temporal profiles that differ from the classic “fast rise exponential decay” behavior common among solar and stellar flares (Getman et al. 2008a). Do these reflect different loop morphologies or different heating or cooling mechanisms?

1.3. Outline of the Paper

After describing the Chandra data sets (Section 2), our procedures for selecting, classifying, and characterizing X-ray superflares are described in Section 3 and the Appendices. Section 4 presents the distributions of flare luminosity and energy. Section 5 provides dependencies of super- and megflare properties on stellar mass and the presence/absence of circumstellar disk. Super- and megflare occurrence rates as functions of flare energy and stellar mass, as well as their contribution to the total X-ray fluence of PMS stars, are evaluated in Section 6. A comparison of X-ray and optical superflares is given in Section 7. Effects of super- and megflares on the environment of young stars are discussed in Section 8. Concluding remarks are presented in Section 9.

2. MYStIX and SFiNCs Data Sets

Over the past two decades, the Chandra X-ray Observatory has devoted several months to observations of star-forming regions within \( d \approx 3 \text{kpc} \) of the Sun. Chandra’s ACIS imager (Garmire et al. 2003) subtends \( 17' \times 17' \), and mosaics of multiple pointings are common. Large projects include a nearly continuous \( \sim 0.9 \text{Ms} \) exposure of the Orion Nebula Cluster, Chandra Orion Ultra-deep Project (COUP; Getman et al. 2005),
Table 1

MYStIX and SFiNCs Regions

| Region            | R.A.  | Decl. | Dis.  | N_{X,obs} | N_{tot} | Region            | R.A.  | Decl. | Dis.  | N_{X,obs} | N_{tot} |
|-------------------|-------|-------|-------|-----------|---------|-------------------|-------|-------|-------|-----------|---------|
|                   | (deg) | (deg) | (pc)  |           |         |                   | (deg) | (deg) | (pc)  |           |         |
| Be 59             | 0.6   | 67.4  | 1100  | 464       | 1703    | NGC 2264          | 100.3 | 9.6   | 738   | 898       | 1837    |
| BRC 1             | 0.0   | 68.5  | 1100  | 42        | 142     | NGC 2362          | 109.7 | -25.0 | 1332  | 467       | 512     |
| Carina Neb.       | 161.2 | -59.7 | 2620  | 6751      | 37899   | NGC 3576          | 168.0 | -61.2 | 2800  | 1131      | 12235   |
| Cep A             | 344.1 | 62.0  | 868   | 194       | 534     | NGC 6334          | 260.1 | -35.9 | 1770  | 1385      | 12661   |
| Cep B             | 343.9 | 62.6  | 868   | 1032      | 1773    | NGC 6357          | 261.4 | -34.3 | 1770  | 1952      | 12581   |
| Cep C             | 346.5 | 62.5  | 868   | 95        | 247     | NGC 7160          | 328.5 | 62.6  | 961   | 134       | 157     |
| DR 21             | 309.8 | -42.3 | 230   | 594       | 3907    | OMC 2–3           | 83.9  | -5.1  | 390   | 287       | 402     |
| Eagle Neb.        | 274.7 | -13.8 | 1740  | 2065      | 7762    | ONC Flank N       | 83.8  | -4.8  | 392   | 198       | 230     |
| Flame Neb.        | 85.4  | -1.9  | 414   | 422       | 596     | ONC Flank S       | 83.8  | -5.7  | 395   | 223       | 261     |
| GGD 12–15         | 92.7  | -6.2  | 830   | 141       | 300     | Orion Neb.        | 83.8  | -5.4  | 405   | 141       | 1700    |
| IC 348            | 56.1  | 32.1  | 324   | 307       | 284     | RCW 120           | 258.1 | -38.5 | 1680  | 262       | 1570    |
| IC 5146           | 328.4 | 47.3  | 783   | 161       | 272     | RCW 36            | 134.9 | -43.8 | 930   | 337       | 1228    |
| IRAS 20050        | 301.8 | -13.1 | 300   | 213       | 308     | RCW 38            | 134.8 | -47.5 | 1700  | 813       | 5750    |
| LDN 1251B         | 339.7 | 75.0  | 300   | 39        | 39      | Rosette Neb.      | 98.1  | 4.9   | 1560  | 206       | 7762    |
| Lagoon Neb.       | 271.0 | -24.4 | 1336  | 1828      | 5058    | Serpens Main      | 277.5 | 1.2   | 440   | 92        | 149     |
| Lkhd 101          | 67.5  | 35.3  | 364   | 197       | 311     | Serpens South     | 275.7 | -2.0  | 460   | 78        | 154     |
| M17               | 275.1 | -16.2 | 1680  | 2296      | 13412   | Sh 2–106          | 306.9 | 37.4  | 1400  | 164       | 764     |
| Mon R2            | 91.9  | -6.4  | 948   | 410       | 882     | Trifid Neb.       | 270.6 | -23.0 | 1264  | 418       | 1279    |
| NGC 1333          | 52.3  | 31.3  | 296   | 116       | 111     | W3                | 36.5  | 62.1  | 2040  | 1571      | 10274   |
| NGC 1893          | 80.7  | 33.4  | 3790  | 1110      | 5977    | W4                | 38.2  | 61.5  | 2091  | 411       | 1614    |
| NGC 2068          | 86.7  | 0.1   | 414   | 231       | 335     | W40               | 277.9 | -2.1  | 500   | 195       | 425     |

Note. Column 1: star-forming region. Columns 2–3: region’s approximate position for epoch J2000.0. Column 4: distance from the Sun. For several regions (DR 21, Flame Nebula, GGD 12–15, IRAC 20050, LDN 1251B, NGC 2068, NGC 3576, Sh 2–106, W3, and W40), we assume the original MYStIX/SFiNCs distances from Feigelson et al. (2013) and Getman et al. (2017). For other regions Gaia-based distances are adopted: RCW 36 from Fissel et al. (2019); OMC 2–3, ONC Flanking fields, and Orion Nebula from Getman et al. (2019b); Serpens Main and Serpens South from Herzeg et al. (2019); RCW 38 from Getman et al. (2019a); W4 from Cantat-Gaudin et al. (2018); and the remaining regions from Kuhn et al. (2019). Column 5: number of observed X-ray-emitting young stars from Broos et al. (2013). Column 6: total stellar population down to 0.1 M_. inferred from the X-ray luminosity function (see text).

Our experience from a wide range of MYStIX and SFiNCs studies is that the reliability of these samples is very high, although they are far from complete catalogs of the full initial mass functions of the star-forming regions. They include both Gaia-derived distances, numbers of observed X-ray young stellar objects, and estimated total intrinsic stellar populations down to around 0.1 M_., following the X-ray luminosity function (XLF) procedures of Kuhn et al. (2015b).

4 For each of the MYStIX regions, Kuhn et al. (2013a) scale the COUP XLF and Maschberger (2013) IMF to the mass-complete bright parts of the observed XLF and IMF, respectively. The XLF procedures originally applied to the MYStIX regions by Kuhn et al. (2015b) are adjusted here to use only two rather than three (as in Kuhn et al.) X-ray energy bands in order to accommodate the lower-counting SFiNCs statistics. All MYStIX+ SFiNCs XLFs are also recalibrated here to the new Gaia-derived distances.
PMS stars with and without mid-infrared excess, which we call “disk-bearing” (Class I and II) and “diskless” (Class III) PMS stars in this paper. Our previous studies have concentrated mainly on issues characterizing particular star-forming regions and relate to the formation and early evolution of young star clusters. The present study on X-ray superflares concentrates on the photon arrival times that played a small role in our previous studies.

3. Methods

3.1. Identification and Classification of X-Ray Superflares and Megaflares

Our flare selection procedure starts with two quantities tabulated for each Chandra source by the ACIS Extract Chandra and XPHOT software packages used to generate the MYStIX and SFiNCs catalogs (Broos et al. 2010; Getman et al. 2010). One is a probability measure of variability within a Chandra exposure, called ObsID. ObsID durations can range from a few hours to around 2 days early in the mission. ACIS Extract reports the most significant (among one or more ObsIDs) probability for source variability of the X-ray photon arrival times within an ObsID using a one-sample Kolmogorov–Smirnov test against a null hypothesis of constant flux. This measure is called $P_{KS}$. XPHOT reports the intrinsic X-ray luminosity of each source averaged over ObsIDs using the local exposure, the distance to the star-forming region (Feigelson et al. 2013; Getman et al. 2017), and assuming a typical PMS X-ray spectrum. This measure is called $L_{rin}$ to represent the background-subtracted X-ray luminosity in the Chandra “total” 0.5–8 keV band corrected for soft X-ray absorption derived from the observed median energy of extracted photons. Here we call this quantity $L_X$ or $L_{X,XPHOT}$.

We then filter the MYStIX and SFiNCs catalogs for the 40 star-forming regions with two criteria: $P_{KS} < 0.01$ to locate PMS stars with possible X-ray variability at least in one of the Chandra ObsIDs and log($L_X$) > 30.5 erg s$^{-1}$ to locate PMS stars with high time-averaged X-ray luminosities. These selection criteria should capture most luminous flares; rare cases where a brief powerful flare is present in an otherwise faint or undetected source may be missed. This stage flagged 3142 ObsIDs from 1713 stars as luminous and variable X-ray PMS stars.

The second stage seeks to identify the start and stop times of statistically significant variations in the photon arrival times. In the parlance of time-series analysis, this problem might be called “changepoint analysis for an inhomogeneous Poisson process” or a “Poisson regression model with multiple changepoints.” This problem is well adapted to likelihood-based fitting if a parametric model for the flare is chosen. While we could adopt an astrophysically motivated model like “fast rise exponential decay,” we choose instead a more general model of a sequence of stepwise constant flux values. This is the approach in the Bayesian Blocks procedure that is widely used to identify and characterize flaring behavior in X-ray sources (Scargle 1998; Scargle et al. 2013). Our statistical fitting procedure is adopted from methods previously used to address problems in the social sciences (Chib 1998; Frühwirth-Schnatter & Wagner 2006; Park 2010; Brandt & Sandler 2009). It is similar to Bayesian Blocks but gives some additional flexibility. Our changepoint model and its application to the MYStIX and SFiNCs X-ray lightcurves are described in Appendix B.

Flare vetting is performed in the third stage of the analyses. Graphical output from the changepoint model method, similar to Figure 11 in Appendix B, combined with observed photon arrival diagrams (Section 4.1) for the 3142 ObsIDs, is visually examined and classified into five types as illustrated in Figure 1:

**Constant:** Many of the ObsIDs are consistent with a single flux level; that is, the best-fit model has no changepoints. An example is shown in Figure 1(a). This commonly occurs when several ObsIDs are present for a MYStIX/SFiNCs star with statistically significant $P_{KS}$ in one exposure but not all ObsIDs. Sometimes a spurious changepoint is found in the first few bins due to an unphysical assumption that the lightcurve starts with zero counts. Of the 3142 ObsIDs examined, 1222 constant events are identified.

**Variable:** Often the Bayesian segmentation results in two or three flux levels without the appearance of a distinct flare (Figure 1(b)). The levels could be different by factors < 1.5 to factors up to ~ 10. Many cases are probably flares with too few counts to be clearly delineated. But other cases consist of two different, but constant, levels of X-ray emission; these may represent the emergence or disappearance of active regions due to stellar rotation. We place 834 variable lightcurves in this category.

**Flare (F):** This class requires at least three segments: a low level before the event, one or more higher levels showing rise and decay, and a low level after the event (Figure 1(c)). In most cases, a factor of $\geq$ 3 difference between the minimum and maximum levels is present. These cases are most useful for science analysis. Peak X-ray luminosities, durations, and total energies can be directly measured, and flare occurrence rates readily calculated. When more than $\geq$ 1000 counts are present in the flare, our companion study gives a detailed spectro-temporal modeling of plasma heating and cooling assuming a single-loop geometry (Getman et al. 2021). A total of 648 flare events are identified.

**Rise (R):** These are events where sudden onset is seen but the full development of the flare is truncated by the end of the ObsID exposure (Figure 1(d)). The true flare peak X-ray luminosity may or may not be the observed maximum luminosity. In most cases, a factor of $\geq$ 3 difference between the minimum and maximum levels is present. These cases can be used for flare occurrence rates but have lower limits to their peak luminosities, total energies, and durations. We find 289 rise events among the 3142 ObsIDs examined.

**Decay (D):** There are events similar to the Rise category, but with the beginning of the flare occurring before the beginning of the exposure (Figure 1(e)). The maximum luminosity is seen when the observation starts and it typically shows a slow decline in brightness to a constant low level. These cases are treated similarly to the Rise class with lower limits to flare properties. We find 149 decay events.

In a small fraction of lightcurves, the flux rises and falls more than once, and our five-component Bayesian segmentation (Appendix B) is too simple to model the behavior. These
cases are examined individually, and flares are manually extracted and placed into the appropriate F, R, or D category. The flare sample studied here thus consists of 1086 events: 648 in category F, 289 in R, and 149 in D.

3.2. Properties of Host Stars

Because the vast majority of the MYStIX/SFiNCs young stellar objects lack spectroscopic measurements, crude estimates of their source extinctions (in the visual band, $A_V$), ages, effective temperatures ($T_{\text{eff}}$), bolometric luminosities ($L_{\text{bol}}$), radii ($R$), and masses ($M$) are obtained in Appendix C using Chandra X-ray and 2MASS, UKIDSS near-infrared (NIR) photometric data described in Getman et al. (2014), Kuhn et al. (2015b), and Richert et al. (2018) complemented by fitting optical-IR spectral energy distributions (SEDs) with the VOSA (Bayo et al. 2008) using other additional numerous available optical and IR photometric catalogs. The presence or absence of circumstellar disks is acquired using the Spitzer-IRAC mid-infrared (MIR) photometry provided in Kuhn et al. (2013b), Povich et al. (2013), and Getman et al. (2017). Related methods are described in Appendix C. Table 2 lists the properties of all 1027 unique F+R+D-flare host stars.

3.3. Flare Properties

Three quantities are calculated for each flare (F), rise (R), and decay (D) event: flare duration, $t_{\text{dur}}$; peak luminosity in the 0.5–8 keV band corrected for soft X-ray absorption, $L_{\text{log} X,pk}$; and total energy, $E_X$. The time between the changepoints of elevated emission (as an output from the changepoint model) is used as an initial value for flare duration. Guided by our flare duration choices for the COUP flares (Getman et al. 2008a), upon visual inspection of the observed MYStIX/SFiNCs

![Figure 1. Examples of the five classes of segmented lightcurves from the vetting stage: constant (a), variable (b), flare (c), rise (d), and decay (e). Derived segments with averaged counts per bin are colored and labeled. Panel legends list flare names, total numbers of X-ray counts, and arrival time differences between the first and last X-ray counts. The flare name is composed of the flare host star name and the relative number of the X-ray Chandra observation, during which the flare is detected. See Appendix B for a detailed description of such lightcurves.](image-url)
photons and intrinsic X-ray luminosity averaged across all available Chandra observations using the XPHOT procedure. Columns 7–13: source’s X-ray median energy and intrinsic X-ray luminosity averaged across all available Chandra observations using the XPHOT procedure. Columns 7–13: source’s visual extinction, age, effective temperature, bolometric luminosity, mass, radius, and SED slope as derived through procedures described in Section 4.2. Stellar properties are listed in Section 4.2. Stellar properties are listed in Section 4.2.

Note. (1) This table is available in its entirety (1027 F+R+D flare host stars) in machine-readable form. A portion is shown here for guidance regarding its form and content. Column 1: star-forming region. Column 2: source name. Columns 3–4: source’s position for epoch J2000.0 in degrees. Columns 5–6: source’s X-ray median energy and intrinsic X-ray luminosity averaged across all available Chandra observations using the XPHOT procedure. Columns 7–13: source’s visual extinction, age, effective temperature, bolometric luminosity, mass, radius, and SED slope as derived through procedures described in Section 4.2. Stellar properties are listed in Section 4.2. Stellar properties are listed in Section 4.2.

4. Treatment of Rise and Decay Flares

A systematic bias is present for the R and D classes where the observed durations and flare energies are lower limits to the true values. It is not known whether the observed peak luminosities are correct or similarly biased; we consider them to be lower limits. In statistics such data are denoted as right-censored data points and the methods from “survival analysis” are used to treat the bias (Feigelson & Nelson 1985). For univariate samples that include both measured (F class) and lower limits (R and D classes), the nonparametric Kaplan–Meier (KM) estimator provides a corrected distribution function. (Kaplan & Meier 1958). The KM estimator redistributes the luminosity and energy lower limits to higher values in a maximum likelihood procedure. KM calculations are made with the survfit function from the R package survival (Therneau & Grambsch 2000; Therneau 2020).

4. Super- and Megaflare Properties

4.1. Flare Table and Atlas

Identifiers and properties of the 1086 F-, R-, and D-flare events are provided in Table 3. F flares have 19 < N_{phot} < 4, 202 cts, 0.9 < ME < 5.7 keV, 3 < E_{diss} < 154 ks, 30.7 < log L_{X,pk} < 33.8 erg s^{-1}, and 34.3 < log E_{X} < 37.6 erg. These events are thus “superflares” and “megaflares” using the criteria log L_{X,pk} > 30.5 erg s^{-1} and log E_{X} > 34 erg suggested in Section 4.2. There are 636 and 450 “superflares” and “megaflares” in our sample, respectively (Section 4.2).

We provide an atlas in which both tabulated and graphical information on each of the 1086 F-, R-, and D-flare events are

| Reg. | Src. | R.A. | Decl. | ME | log(L_{X}) | Av | t | log(T_{eff}) | log(L_{bol}) | M | R | ΩIRAC |
|------|------|------|-------|----|------------|----|---|--------------|-------------|---|---|-------|
| Be 59 | 000053.45+672615.0 | 0.222725 | 6.4737501 | 2.4 | 3.10 | 7.0 | 1.7 | 3.63 | -0.03 | 0.8 | 1.8 | -2.2 |
| Be 59 | 000054.01+672119.8 | 0.225079 | 6.4735504 | 2.4 | 3.10 | 5.5 | 1.0 | 3.56 | -0.27 | 0.5 | 1.8 | -2.0 |
| Be 59 | 000102.52+672841.0 | 0.265034 | 6.4748076 | 1.9 | 3.10 | ... | ... | ... | ... | ... | ... | -2.2 |
| Be 59 | 000138.66+672800.6 | 0.411086 | 6.4746689 | 2.0 | 3.10 | 3.3 | 0.7 | 3.52 | -0.41 | 0.2 | 1.9 | -0.5 |
| Be 59 | 000144.25+672457.3 | 0.434385 | 6.4741918 | 2.5 | 3.21 | 6.2 | 1.7 | 3.75 | 1.19 | 2.7 | 4.1 | -2.2 |
collected onto a single page. A sample atlas page is shown in Figure 2. This is an X-ray megaflare from a low-mass PMS star embedded in the M17 North Bar cloud (Broos et al. 2007). The page features two plots with the host star and flare properties extracted from Tables 2 and 3. The first plot gives X-ray photon arrival times and energies with individual photons.
marked as blue points. The red curve shows a likelihood-based local quadratic regression fit with 84% confidence intervals (black dashed lines) generated using the locfit.robust function from the CRAN locfit package (Loader 2020). This procedure and its mathematical foundations are described by Loader (1999). The flare changepoints derived in Appendix B are indicated by the green dashed lines. Time ranges from the start to the end of the Chandra ObsID exposure.

The plot title gives the star identifier from the MYStIX or SFI NICs catalog and its star formation region. The annotation gives 11 scalar quantities: distance to the star formation region; relative number of the current ObsID for this X-ray source; number of photons in the ObsID with the flare; median energy of the ObsID; infrared slope from which disk presence is inferred; and the estimated age, visual absorption, stellar effective temperature, bolometric luminosity, mass, and radius derived using optical-IR photometry data as described in Appendix C.

The second plot gives an adaptively smoothed absorption-corrected X-ray luminosity lightcurve (red) with 1σ confidence intervals for a binned histogram (blue) using analytical approximations for a Poisson distribution (Gehrels 1986). The binned histogram is composed of independent count bins, each accumulating similar numbers of X-ray counts (N_{hist}) and centered at the mean arrival time between the first and the last counts in the bin. The plot legends include flare type (Section 3.1); ObsID exposure time; number of counts per adaptive kernel (red curve); number of counts in a histogram bin (blue); flare start and stop times (green lines), which are the adjusted changepoints as described in Section 3.3; log L_{X, pk}; and log E_{X}. On the X-ray luminosity time-series plots, not all jiggles in the red curve are statistically significant, and blue circles with errors are not carefully placed with respect to possible interesting structures such as flare peaks.

Two idiosyncrasies of this particular star can be noted. First, the Spitzer-IRAC counterpart is missing so no α_{IRAC} value is given because it lies near the cloud–H_{2} interface which suffers from bright MIR nebula emission. Second, the source remains undetected in X-rays within the first half of the Chandra exposure prior to the flare.

### 4.2. Flare X-Ray Peak Luminosity and Energy Distributions

PMS flaring has been most intensively studied in the Orion Nebula Cluster for two reasons: the photon flux for flares is high due to its close distance of only 0.4 kpc, and the COUP project provides a unique almost-continuous Chandra observation over 13 days (Getman et al. 2005). Studying a sample of ~1 M_{☉} stars, Wolk et al. (2005) reported the flare energy distribution slope of dN/dE_{X} ∝ E_{X}^{-1.7} using a linear regression technique, which was revised to dN/dE_{X} ∝ E_{X}^{-1.9±0.2} by Stelzer et al. (2007) using a maximum likelihood procedure. Caramazza et al. (2007) found an X-ray flare count (C_{X}) distribution of dN/dC_{X} ∝ C_{X}^{-2.2} for 151 flares from low-mass 0.1–0.3 M_{☉} COUP stars. For the high-energy tail of a larger sample of 954 COUP flares regardless of host stellar mass, Albacete Colombo et al. (2007) find the slope of E_{X}^{-2.1}.

In addition to the Orion Nebula Cluster, an important survey of the Taurus molecular cloud was made with the XMM–Newton satellite (XEST; Güdel et al. 2007). The energy distribution of 33 X-ray flares was found to be dN/dE_{X} ∝ E_{X}^{-2.4±0.5} (Stelzer et al. 2007).

Figures 3(a)–(b) show the KM estimators of peak luminosities and total energies for the full sample of 1086 MYStIX/SFI NICs flares (red curve), together with the empirical cumulative distribution functions (c.d.f.’s) for the fully observed subsample of 648 F-class flares (blue curve). The R and D flares raise the median values from log L_{X, pk} = 31.9 to 32.2 erg s^{-1} and from log E_{X} = 35.8 to 36.1 erg. Quite reasonably, the R and D flares that are truncated by the limited duration of Chandra exposures tend to be more luminous and more energetic than those captured in their entirety as F flares. This is consistent with the finding of Albacete Colombo et al. (2007) that the flares identified within the shorter 100 ks COUP Chandra blocks appear systematically less energetic than those identified within the entire COUP observation. The black lines in Figures 3(a)–(b) and black points in Figures 3(c)–(d) represent the full sample of 1086 F+R+D flares, for which the energies and peak X-ray luminosities of the “R, D” flares are multiplied by 5 to match the KM estimator (red). The unbinned c.d.f.’s for this corrected by × 5 F+R+D sample (i.e., the c.d.f. black lines in Figures 3(a)–(b)) are fitted with the Pareto function to obtain the Pareto slope β as detailed below.

Both flares from individual regions and from a collection of regions (like MYStIX/SFI NICs) may be subject to spatially varying absorption across a region (Section 4.3). Unlike the flare luminosity/energy distributions for individual star-forming regions, the MYStIX/SFI NICs distributions may be subject to the additional effect of different distances (thus source/flare sensitivities) toward different MYStIX/SFI NICs regions. This is not a problem for the complete megafall sample (Section 4.4).

Figures 3(c)–(d) show differential distributions of the upper-panel c.d.f.’s grouped in 0.2 dex bins. Histogram error bars are approximations to 95% confidence intervals of a Poissonian distribution (Gehrels 1986). The lower bins clearly represent incomplete sampling as our superbhare selection procedure (Section 3.1) requires that the time-averaged luminosity exceeds log L_{X} = 30.5 erg s^{-1}. Many flares with 31 < log L_{X, pk} < 32 erg s^{-1} will be diluted by long periods of nonflaring emission so the time-averaged luminosity falls below our selection limit. Notice that these histogram representations of the data are shown here to visually emphasize approximate data completeness limits. Based on the peak values in these histograms, our sample appears complete above the peak of the differential distributions at log L_{X, pk} > 32.0 erg s^{-1} and log E_{X} ≈ 36.0 erg. More accurate completeness limits are derived below based on the Pareto fits to the unbinned data (i.e., those shown as c.d.f.’s in Figures 3(a)–(b)).

It is well-known that solar and stellar flares exhibit power-law (Pareto) function distributions of various properties. Choosing progressively higher X-ray luminosity and energy cutoffs following Stelzer et al. (2007), the unbinned c.d.f.’s with N = 1086 or fewer data points are fitted by the maximum likelihood estimation to the Pareto distribution function

\[
\text{c.d.f. } = 1 - (x_{\text{min}} / x)^{\beta} \quad \text{for } x \geq x_{\text{min}}, \quad \text{where} \quad (1)
\]

\[
\beta = \frac{N}{\sum_{i=1}^{N} \ln (x / x_{\text{min}})}. \quad (2)
\]

The energy (or X-ray luminosity) distributions are expressed through the Pareto slope β as log(dN/d log(E_{X})) ∝ log(E_{X})^{-\beta} or dN/dE_{X} ∝ E_{X}^{-\beta-1}.
For the entire F+R+D flare sample, the Anderson–Darling goodness-of-fit test shows statistically unacceptable fits with $p < 0.01$ at the energy cutoffs $\log(E_X) < 36.1$ erg, consistent with the shape of the differential distribution (Figure 3(d)). The fits above that energy value show acceptable and statistically indistinguishable ($p > 0.05$) solutions with the power-law slope varying between $\beta = [0.91–1.00]$ for the energy cutoffs $\log(E_X) = [36.1–36.5]$ erg, where the data samples are the richest, $N > [200–500]$ data points. Conservatively, we choose the completeness limit as $\log(E_X) = 36.2$ erg. This completeness limit is the reason we choose $\log(E_X) = 36.2$ erg as the threshold for the label “megaflare” in contrast to “superflare.”

With this energy cutoff value, the energy distribution has a power-law slope $\beta = 0.95$. Slope uncertainties (95% confidence intervals) obtained from 1000 bootstrap resamples for the KM estimators are ±0.07. The Pareto model with $\beta = 0.95 \pm 0.07$ above $\log(E_X) = 36.2$ erg is shown as the green curve in Figure 3(b).

At higher-energy cutoffs $\log(E_X) = [36.8–37]$ erg, the power-law slope changes to higher values of $\beta = [1.2–1.4]$, but with fewer sample data points $N = [80–140]$ and hence higher statistical uncertainties, ±0.2. The outlier points, visually represented by the binned point at $\log(E_X) = 37$ erg (Figure 3(d)), are likely the cause of this slope increase.

For the smaller “F” flare sample, the Pareto slope is $\beta = 1.27 \pm 0.16$ at $\log(E_X) = 36.2$ erg (the cyan lines in Figures 3(b) and (d)).

The inferred Pareto slope of $\beta = 0.95 \pm 0.07$ for the F+R+D sample leads to the X-ray flare energy distribution of $dN/dE_X \propto E_X^{-\beta-1} = E_X^{-1.95}$ within the energy range of $\log(E_X) = 36.2$ to 38 erg. This $-1.95$ power-law energy distribution is consistent with those for optical, EUV, and
X-ray solar/stellar flares captured at a very wide but lower range of energies, from $E_{\text{flare}} = 10^{24}$ erg for solar nanoflares to $E_{\text{flare}} = 10^{35}$ erg for superflares from solar-type stars (e.g., Notsu et al. 2019; Okamoto et al. 2021, and references therein). The MYStIX/SFiNCs flare energy distribution is also consistent with that of the aforementioned superflare samples from young stars in the Orion Nebula and Taurus star-forming regions (Wolk et al. 2005; Albacete Colombo et al. 2007; Caramazza et al. 2007; Stelzer et al. 2007). Compared to these previous studies, the MYStIX/SFiNCs data offer a factor of $>6$ increase in the sample of megaflares from young stars.

We thus find that the shape of the energy distribution of stellar flares at their highest levels are similar to solar and stellar flares over a remarkable 14 orders of magnitude in energy.

For the $L_{X, \text{pk}}$ distribution of the F+R+D-flare sample, the Pareto fits to the data become statistically acceptable at $\log(L_X) = 32.5$ erg s$^{-1}$ and onwards. The inferred slopes are $\beta = 1.11 \pm 0.09$ and $\beta = 1.26 \pm 0.16$ at the luminosity cutoffs of 32.5 and 32.8 erg s$^{-1}$, respectively. The latter solution is shown in green in Figures 3(a) and (c). This progression from a shallower to a bit steeper slope reflects the broken power-law morphology of the $L_{X, \text{pk}}$ histogram that is clearly seen in Figure 3(c). The latter is presented in the next section. The distributions for the “F” flare sample have a similar broken power-law shape but with systematically steeper slopes (cyan curves in Figures 3(a) and (c)).

### 4.3. Effects of Absorption

Young stellar objects are often subject to soft X-ray absorption by K- and L-shell transitions in metal atoms along the line of sight (Wilms et al. 2000). Most of this absorption occurs locally in the parental molecular clouds and, for the youngest stars, their local protostellar envelopes. Absorption is thus an indirect measure of stellar age with the youngest Class I and II systems more absorbed than older Class III systems. But for the more distant star formation regions, absorption by molecular clouds in intervening spiral arms also contributes to absorption. The line-of-sight material causes soft X-ray absorption is often quantified as $N_H$, the column density of equivalent hydrogen, which can be converted to a visual absorption $A_V$ by assuming a gas-to-dust ratio (e.g., Hasenberger et al. 2016; Zhu et al. 2017). In the MYStIX and SFiNCs studies, the median energy of X-ray photons ME in keV is used as a surrogate for $N_H$ following the calibration procedure described by Getman et al. (2010). Here we measure ME$_F$ the median energy of photons arriving between the start and end times of the flares.

Figure 4 presents the histograms of the peak X-ray luminosity and flare energy distributions for lightly (ME$_F$ $\lesssim$ 2 keV) and heavily (ME$_F$ $>$ 2 keV) absorbed flares, respectively. The 2 keV boundary corresponds approximately to $\log N_H$ $\approx$ 22.0 cm$^{-2}$ and $A_V$ $\approx$ 5 mag. We see that the shape of the $L_{X, \text{pk}}$ distribution is especially sensitive to the absorption effect. The lightly absorbed flare samples follow a power-law distribution reasonably closely over the range of $L_{X, \text{pk}}$ $\sim$ 32–34 erg s$^{-1}$ range, but the heavily absorbed samples show a deficit around $L_{X, \text{pk}}$ $\sim$ 32.2–32.8 erg s$^{-1}$.

The reason for this deficit is again probably related to our selection criteria for selecting $\geq 1000$ superflaring stars from $\geq 24,000$ MYStIX and SFiNCs stars. We require that the time-averaged X-ray luminosities exceed $L_X > 30.5$ erg s$^{-1}$ in the total Chandra 0.5–8 keV band (Section 3.1). The recovery of missing soft X-ray emission from heavily absorbed sources appears to be incomplete, and thus a significant number of stars with X-ray flare peak luminosities around $\log L_{X, \text{pk}} = 32$ erg s$^{-1}$ are excluded by our time-averaged $L_X > 30.5$ erg s$^{-1}$ selection criterion.

The flare energy distributions are less affected by absorption. Based on the Pareto fits to the unbinned data (c.d.f.’s are not shown here), both the low- and high-absorbed flare energy distributions are complete at around $\log(E_X) = [36.1–36.2]$ erg with statistically indistinguishable Pareto slopes for the F+R+D-flare samples of $\beta = [1.02–1.09] \pm 0.14$ and $\beta = [0.86–0.89] \pm 0.08$, respectively (green in Figures 4(b) and (d)).

### 4.4. Effects of Distance

To consider the possible sample bias due to the range of distances to the MYStIX/SFiNCs star-forming regions, the F+R+D-flare sample is divided into two flare groups, near ($D \leq 1500$ pc) and far ($D > 1500$ pc). Application of our analyses to the flare energy distributions of these two groups shows that the nearby flares have lower completeness limits of $\log(E_X) = [35.9–36.0] \pm 0.01$ and corresponding Pareto slope range $\beta = [0.87–0.93] \pm 0.10$. For more distant flares, the completeness limits are similar to those of the entire flare sample: $\log(E_X) = [36.1–36.2] \pm 0.04$ with $N = [387–333]$ and $\beta = [0.88–0.94] \pm 0.08$. The inferred flare energy Pareto slopes remain statistically indistinguishable between these two distance-stratified samples.

### 5. Super-/Megaflares and Host Star Properties

#### 5.1. Megaflares from Protostars

The reports of X-ray flaring in the Orion cloud Class 0 protostar HOPS 383 and other early-phase protostars push the onset of X-ray flaring into the earliest infall stages of star formation (Grosso et al. 2020 and references therein). For this Class 0 protostar, Chandra detected 28 X-ray counts with median energy 5.4 keV corresponding to the column density of $\log(N_H) \approx 23.8$ cm$^{-2}$ or $A_V \approx 200$ mag. Our MYStIX/SFiNCs superflare sample has eight heavily embedded protostellar candidates with flare X-ray median energies above 5.0 keV. These sources, which lack mass estimates, are listed in Table 4. See the electronic flare atlas (Section 4.1) for their photon arrival diagrams and lightcurve morphology. Three sources are located in the nearby ($d$ $\sim$ 300–400 pc) NGC 1333, L1251b, and Flame regions; one in the intermediate-distance ($d$ $\sim$ 900 pc) Cep A region; and four in the more distant ($d$ $\sim$ 1700 pc) M17 and RCW 38 regions. SED IRAC slopes, available for three out of eight sources, show fluxes ascending toward longer wavelengths confirm their protostellar nature.

For the five F (fully observed) MYStIX/SFiNCs protostellar flares in Table 4, the median X-ray flare peak luminosity of $L_{X, \text{pk}}$ $= 3 \times 10^{32}$ erg s$^{-1}$, a factor of 7 higher than the flare seen in HOPS 383. Using the log $E_X = 36.2$ erg boundary, at least five of the eight events are megaflares. This clearly demonstrates that the extremely high levels of flaring seen in Class II and III PMS stars is present in Class I, and possibly Class 0, protostars.

Follow-up far-IR/submillimeter observations of these eight X-ray sources would assist in unraveling their evolutionary stages and looking for spectroscopic evidence that the
penetrating superflare X-rays play a role in the ionization of their circumstellar envelopes and disks.

5.2. Comparing Disk-bearing and Diskless Stars

As outlined in Section 1, it has been long debated whether some X-ray flares from PMS stars arise from the magnetic reconnection in loops extending from the star to the disk rather than loops with both footprints in the stellar surface. If we can assume that the near- and mid-infrared photometric excess is an adequate indicator of the presence of gaseous inner protoplanetary disk, we can investigate this issue by comparing flare distributions in MYStIX/SFiNCs stars with and without disks. We associate diskless stars with infrared spectral energy distribution slopes $\alpha_{\text{IRAC}} \leq -1.9$ and disk-bearing stars with $\alpha_{\text{IRAC}} > -1.9$ (Richert et al. 2018). Two-thirds of the hosts of the 1086 superflare stars have sufficient infrared photometry to measure $\alpha_{\text{IRAC}}$, roughly evenly divided between the two classes with 397 flares from diskless and 348 flares from disk-bearing stars. Figure 5 compares the KM estimators for peak X-ray luminosity and flare energies for these two subsamples. The c.d.f. estimators for peak X-ray luminosity and flare energy of the two samples are indistinguishable. This is validated with the survival analysis logrank 2-sample test (Harrington & Fleming 1982) with $p$-value $> 0.5$ for both measures of flare strength.

Similarly, if only the most powerful flares with energies above the completeness limit of $\log(E_X) = 36.2$ erg are considered, the flare samples are reduced to 85 and 69 flares from diskless and disk-bearing stars, respectively. The c.d.f. estimators for peak X-ray luminosity and flare energy of the...
Table 4
Superflares from Protostellar Candidates

| Region       | Source_Obs          | R.A. (deg) | Decl. (deg) | o_HRAC | Flare | C_f (cmts) | ME_f (keV) | log(log(L_xpk)) (erg/s) | log(log(L_x)) (erg) |
|--------------|----------------------|------------|-------------|--------|-------|------------|------------|----------------------------|---------------------|
| CepA         | 225619.58+620223.4_1 | 344.081599 | 62.039843   | 0.3    | F     | 38         | 5.7        | 32.4                       | 36.6                |
| Flame        | 054143.54-015511.7_1 | 85.431458  | -1.919931   | 2.1    | R     | 21         | 5.4        | >31.3                      | >34.9               |
| L1251b       | 223846.92+751133.6_2 | 339.695508 | 75.192679   | ...    | F     | 77         | 5.4        | 32.3                       | 36.0                |
| M17          | 182016.85-160726.0_6 | 275.070224 | -16.123908  | ...    | R     | 22         | 5.1        | >32.8                      | >36.2               |
| M17          | 182021.76-161257.6_2 | 275.090706 | -16.216027  | ...    | F     | 296        | 5.2        | 33.8                       | 37.6                |
| M17          | 182022.11-161305.2_2 | 275.092133 | -16.218137  | ...    | R     | 53         | 5.2        | >32.8                      | >37.0               |
| NGC 1333     | 032858.43+312217.7_2 | 52.243473  | 31.371592   | 2.0    | F     | 67         | 5.4        | 32.0                       | 35.8                |
| RCW 38       | 085906.63-473021.9_1 | 134.777644 | -47.506100  | ...    | F     | 244        | 5.0        | 33.5                       | 37.5                |

Note. Column 1: star-forming region. Column 2: unique X-ray flare name, composed of the X-ray source name and the relative number of the X-ray Chandra observation, during which the flare is detected. Columns 3–4: the source position for epoch J2000.0 in degrees. Column 5: SED IRAC slope. Column 6: flare type: F = full, R = rise, D = decay. Columns 7–10: flare properties, including X-ray counts, median photon energy, peak X-ray luminosity, and flare energy. (This table is available in machine-readable form.)

disk-bearing and diskless stars remain indistinguishable, with logrank p-values > 0.5 (figure is not shown).

We thus find no statistical differences in flare strength distributions between disk-bearing and diskless MYStIX-SFiNCs samples. Both types of young stellar objects, with and without disks, produce X-ray superflares that follow similar distributions of flare peak luminosity and energy.

Consistent with our result but for much smaller numbers of X-ray flares, Stelzer et al. (2007) report no flare energy differences for flares detected in disk-bearing and diskless stellar members of the Taurus star-forming region. No differences in the flare occurrence rates and flare durations are seen between disky and diskless COUP stars (Section 5.2 in Flaccomio et al. 2012). Furthermore, no noticeable differences in the relations between the optical and X-ray flare energies are seen for flares detected from disk-bearing and diskless members of the NGC 2264 region (Flaccomio et al. 2018).

Figure 5. Kaplan–Meier cumulative distribution estimators of flare peak X-ray luminosity and flare energy for F, R, and D superflares for disk-bearing (red) and diskless (green) stars. Dashed curves give 95% confidence intervals.

5.3. Superflares and Stellar Mass

Our sample has 1027 young stellar objects that produce 1086 X-ray “F+R+D” super- and megafrares; 749 of these stars have available stellar mass estimates (Table 2). To understand the nature of these host stars, their properties are compared to the full sample of MYStIX-SFiNCs young stars with available masses and XPHOT X-ray luminosities. The six properties of interest include location on the Hertzsprung–Russell diagram (HRD), stellar mass and radius, source visual extinction and source X-ray median energy measuring line-of-sight absorption, and source X-ray luminosity averaged across all available Chandra observations ($L_{X,\text{XPHOT}}$). Figure 6 compares these six properties for all MYStIX-SFiNCs stars (upper panels) and super-/megafrares hosts only (lower panels). The derivation of the stellar properties is detailed in Appendix C.

Some intermediate-mass stars appear on the HRD diagram with ages younger/older than the chosen age boundaries of 0.4 Myr and 5 Myr, respectively because their $T_{\text{eff}}$ and $L_{\text{bol}}$ estimates were obtained from the VOSA rather than from the $J$ versus $J – H$ color–magnitude diagram (see details in Section C).

The young stellar objects are separated into five source strata associated with specific loci on the HRD diagram:
1. Fully convective low-mass stars (\( \lesssim 1 \ M_\odot \)) on Hayashi tracks, mainly M- and K-type stars (green symbols; 15,730 stars for the full sample and 237 stars for the super-megaflare sample)

2. Fully convective solar-mass stars (1 \( \lesssim M \lesssim 2.5 \ M_\odot \)) on Hayashi tracks, mainly K- and G-type stars (blue symbols; 7058 and 303 stars)

3. Intermediate-mass stars (2 \( \lesssim M \lesssim 5 \ M_\odot \)) on Henyey tracks likely developing radiative cores. These include G-, F-, A-, and some late B-type stars (red symbols; 2919 and 165 stars)

4. High-mass stars with \( M \gtrsim 5 \ M_\odot \), mainly B-type stars (orange symbols; 974 and 44 stars). Note that the X-ray emission probably is not produced by the massive primary, but rather by lower-mass unresolved secondaries in multiple systems as proposed for Orion Nebula Cluster B-type stars (Stelzer et al. 2005).
5. Young stellar objects without available HRD locations and mass estimates (black symbols; 13,360 and 278).

Anderson–Darling nonparametric two-sample tests between strata in each panel show significant differences for all of the full MYStIX+SFInCs samples. For the superflare hosts, low- and solar-mass strata (green and blue curves) have significantly different mass, radius, absorption, and with $p$-value $\sim$2% X-ray luminosity. Solar- and intermediate-mass strata (blue and red curves) differ in mass, radius, and X-ray luminosity but not absorption and median energy. Intermediate- and high-mass strata (red and orange curves) differ in mass, (with $p$-value $\sim$5%) radius, absorption, and (with $p$-value $\sim$1%) X-ray luminosity, but not median energy.

One finding from Figure 6 confirms a well-established result. The X-ray luminosity pattern for super-/megaf'are hosts is qualitatively similar to that of all MYStIX+SFInCs stars, but with $L_{X, \chi PHOT}$ shifted toward higher values due to our selection of the most powerful X-ray emitters. This strong correlation between the X-ray luminosity and stellar mass has been reported for young stars in many star-forming regions (e.g., Preibisch et al. 2005; Telleschi et al. 2007a).

Another confirmatory result is that young stellar objects across a wide mass range, from 0.1 $M_\odot$ for M-type stars to $>5$–10 $M_\odot$ for B-type and even some O-type stars, produce X-ray superflares. The X-ray flaring detected from massive stars may be associated with unresolved lower-mass stellar companions (Stelzer et al. 2005). We find that 4%–5% of solar-, intermediate-, and high-mass MYStIX+SFInCs stars are in the superflare sample. Only 1.5% of the low-mass stratum produce superflares, but this is expected from the mass–$L_X$ correlation. The MYStIX/SFInCs surveys have diminished sensitivities toward $<0.5$–$1 M_\odot$ stars and detect only a handful of young brown dwarfs.

Two new results relating specifically to super-/megaf'are host stars emerge from Figure 6. First, the X-ray spectrum of superflares from all mass strata appears indistinguishable. This result emerges from the different X-ray median energy patterns between the full MYStIX+SFInCs and the superflare samples. For the full sample (as mentioned above), the lower-mass strata exhibit lower median energies than massive PMS stars (a result known from Orion Nebula Cluster studies; Preibisch et al. 2005), but the median energies are indistinguishable among the superflare stars of different masses. This suggests that super-/megaf'are physics and production mechanisms are similar across the wide range of star masses.

Second, the superflare stars without mass estimates (black symbols) have much higher X-ray median energies indicating they are heavily absorbed. Half of these stars have median energies above 3 keV (equivalent to log $N_H > 22.5$ cm$^{-2}$ or $A_V > 15$ mag) compared to only 15% of stars with identifiable mass estimates from the HRD. These young stellar objects have X-ray luminosity distributions similar to the visible $M > 1 M_\odot$ samples. Most of them are associated with stellar clusters embedded in molecular clouds (Getman et al. 2018). It is reasonable to infer that they are very young embedded objects, supporting the evidence in Section 5.1 that the production of X-ray superflares starts very early in the PMS stages of evolution.

The X-ray properties shown in Figure 6 are based on the full Chandra exposure time, much of which may be associated with “characteristic” emission, composed of numerous weaker flares, before and after the superflare. In contrast, Figure 7 examines the dependence of the super-/megaf'are properties—flare duration, peak X-ray luminosity, and total energy—on host star mass. The figure shows KM estimators of the c.d.f.’s for the F+R+D superflare subsamples stratified by stellar mass. Note that the flare durations may be underestimated for flares lasting longer than $\sim$1 day. Several results are obtained from Figure 7.

1. Super-/megaf'are duration distributions differ little from low-mass to intermediate- and high-mass stars. Nearly all lie between $\sim$20 ks and $\sim$100 ks with a median duration of around 40 ks. This median duration for the MYStIX+SFInCs superflares is similar to that of the COUP superflares (Getman et al. 2008a).

2. Super-/megaf'are $E_X$ distributions are correlated with stellar mass: the low-mass stratum is weaker than the intermediate-mass stratum with $p$-value $\sim$0.0003 from the logrank test for equality of survival. A similar effect in the $L_{X,peak}$ distribution may be present but is not statistically significant in our samples. A reasonable explanation for a $E_X$–$M$ relation is that flare energy scales with the volume and footprint area of flaring loops involved in a single event, which in turn may depend on the stellar surface area, hence on radius and mass. Another possibility is that the flare energy is powered by the strength of surface magnetic fields that may be stronger on stars with larger stellar volumes allowing more opportunity for a convective dynamo. A similar time-averaged $L_X$–$M$ relation, with similar possible explanations, is well-known in the Orion Nebula Cluster and other PMS populations (Preibisch et al. 2005).

3. The super-/megaf'are duration, peak X-ray luminosity, and total X-ray energy distributions for the high-mass stratum (orange curve) consistently correspond to the longest and most powerful flares, albeit not all effects are statistically significant. It is unclear why X-ray flares from B-type stars should be distinct from lower-mass stars if the multiple stellar companion hypothesis for B-type star X-ray emission is correct (Stelzer et al. 2005).

4. There are 450 MYStIX/SFInCs “F+R+D” megaf'ares above the completeness limit of log($E_X$) = 36.2 erg (Section 4.2). The ratios of megaf'ares to mega+superflares are 17%, 26%, 19%, 6%, and 32% for S1, S2, S3, S4, and S5 subsamples, respectively. These values are employed in the calculations of flare frequencies below.

6. Super- and Megaf'are Occurrence Rates

6.1. Measured and Extrapolated Megaf'are Occurrence Rates

The frequency of superflares in PMS populations, in units of flares per star per year, can be estimated from our analysis of the Chandra MYStIX+SFInCs PMS sample as the ratio

$$f_{\text{SupFl}} \approx N_{\text{SupFl}} / (N_{\text{PMS}} \times \text{Med}(t_{\text{obs}})),$$

where $N_{\text{SupFl}}$ is the number of superflares above a specified flare energy limit, $N_{\text{PMS}}$ is the total intrinsic pre-main-sequence population observed with Chandra, and Med($t_{\text{obs}}$) is the median of the total Chandra exposures among the observed 24,306 X-ray young stellar objects across the 40 MYStIX/SFInCs star-forming regions. However, careful estimation of these quantities is needed:
1. We treat $N_{\text{SupFl}}$ for incompleteness by considering here only the megafarles where our sample is complete (Figure 3).

2. $N_{\text{PMS}}$ is treated for incompleteness in the Chandra MYStIX and SFiNCs PMS samples using the method of Kuhn et al. (2015b), where the X-ray luminosity function (XLF) of each region is scaled to the Orion Nebula Cluster (Section 2, Table 1). $N_{\text{PMS}} \approx 98,000$ and 14,000 for the S1 and S2+S3+S4+S5 subsamples, respectively.

3. The median Chandra observation exposure time among the observed X-ray MYStIX and SFiNCs young stellar objects ($N = 24, 306$) across the 40 regions is $t_{\text{Chandra}} = 74.4 \, \text{ks}$ with the bootstrap-derived 95% confidence band of $\pm 0.1 \, \text{ks}$ (Broos et al. 2013; Kuhn et al. 2013a; Townsley et al. 2014; Getman et al. 2017).

Figure 8 presents the result from Equation (3): observed and extrapolated occurrence rates for PMS stars based on the MYStIX+SFiNCs megafarles. Teal and green indicate S2+S3+S4+S5 (more massive stars with $M > 1 \, M_\odot$) and S1 (less massive stars with $M < 1 \, M_\odot$), respectively. The dashed lines extrapolate the frequencies with $eta = \pm \text{err}_\beta$, the solid line is from Shibayama et al. 2021, and the colored bands from Okamoto et al. 2013 with the bootstrap-derived 95% confidence bands for the KM estimators. Curve colors correspond to the five source strata from Figure 6. The legends list the numbers of flares for each of the source strata.

For each of the three flare groups, the statistical errors on $N_{\text{SupFl}}$ and $Med(t_{\text{obs}})$ contribute less than 11% to the uncertainty on $f_{\text{SupFl}}$. The systematic error on $N_{\text{PMS}}$ due to the uncertainty in the methods used to derive total stellar populations, such as XLF versus initial mass function (IMF), is about 0.2 dex on log($N_{\text{PMS}}$) (Figure 4 in Kuhn et al. 2015b). This provides the largest contribution to the error on the inferred occurrence rate.
values by accounting for the contamination from subgiants in their Kepler sample of flaring G-type main-sequence stars. The optical flare energies from these studies were multiplied by $\times 1/15$ to give equivalent X-ray energies extrapolating the PMS optical-X-ray flare relation measured by Flaccomio et al. (2018) from simultaneous multiband observations of a nearby PMS population.

The resulting frequencies of the MYStIX-SFiNCs megaflares with $\log(E_X) > 36.2$ erg using Equation (3) are $0.3^{+0.7}_{-0.3}$, $1.7^{+1.0}_{-0.3}$, and $11.0^{+6.4}_{-4.4}$ flares per star per year for the low-mass (green), full IMF (magenta), and high-mass (teal) groups, respectively. The megaflares occurrence rates, at a fixed $E_X$ value, decrease with decreasing flare host mass.

The MYStIX/SFiNCs sample includes roughly 6.5 times more $\log(E_X) > 36.2$ erg flares than the COUP flare sample of Albacete Colombo et al. (2007). Despite this large flare number difference, after renormalization of the Albacete Colombo et al. (2007) frequency from the total detected COUP point sources (1616, including contaminating AGNs) to the total stellar population in the COUP field of the Orion Nebula region ($N_{\text{PMS}} = 1700$ stars; Table 1), their COUP flare frequency of $\sim$1.4 flares per star per year is consistent with that of the full IMF MYStIX/SFiNCs megaflares occurrence rate.

The PMS flare occurrence rates inferred from Figure 8 are remarkable! The typical solar-mass PMS star is producing $\sim 1$–3 megaflares yr$^{-1}$ or $\sim 10^7$ megaflares over the early $\sim$5 Myr duration of the PMS evolutionary phase. If the extrapolation of PMS megaflares to lower unobserved energies around $\log(E_X) \approx 34$ erg is valid, then PMS stars produce $\sim 10^6$ or more superflares than older main-sequence stars. Taking a typical flare duration of 50 ks (Figure 7) and assuming superflares occur randomly in an ensemble of stars, a typical PMS star is experiencing a superflare $\sim 30\%$ of the time.

6.2. Contribution of Megaflares to PMS X-Ray Fluence

A debate has raged for decades over the relative importance of powerful flares and nano- or microflares for heating the solar and stellar coronal (Vilangot Nhali et al. 2020 and references therein). While our study gives no information on the contribution of small flares, we can address the energy contribution of the megaflares to the total time-integrated X-ray emission of PMS stars.

One can assume that many weaker flares on a given star blur together into a quasi-continuous emission of X-rays that can be called the “characteristic” level (Wolk et al. 2005; Albacete Colombo et al. 2007; Caramazza et al. 2007). For the MYStIX/SFiNCs young stars (excluding superflare hosts), this characteristic emission is shown as a local regression fit (green curve) in Figure 9. In the mass range $M > 0.3 \, M_{\odot}$, this fit is similar to the $L_X - M$ relations for young stellar members of the Orion Nebula and Taurus regions reported by Preibisch et al. (2005) and Telleschi et al. 2007A

Figure 9 shows that, for masses exceeding $\sim 1 \, M_{\odot}$ stars, superflaring stars have levels of characteristic emission that are not unusually high compared to other PMS stars. But at lower masses, superflaring stars have unusually high time-integrated X-ray luminosities. In contrast, mega-flaring stars always have time-integrated X-ray luminosities far above the characteristic emission of typical PMS stars in any mass range.

We concentrate here on the contribution of these most powerful megaflares to the total X-ray fluence of PMS stars. We integrate megaflares energetics for $36.2 < \log(E_X) < 38.0$ erg based on the flare frequency from Figure 8, and we estimate the characteristic emission energetics based on the X-ray luminosities (green line) given in Figure 9. Specifically, the megaflares energetics (in erg) released per year can be calculated as

$$E_{\text{tot}} = \frac{10^{\beta_2} \cdot \beta_2}{1 - \beta_2} \cdot (E_{\text{max}}^{1-\beta_2} - E_{\text{min}}^{1-\beta_2}),$$

where $E_{\text{min}} = 10^{36.2}$ erg and $E_{\text{max}} = 10^{38}$ erg (maximum energy of detected PMS flares). The $\beta_2$ and $\beta_2$ parameters indicate the normalization and slope of the lower and upper boundaries in the uncertainty loci, shown as colored polygons in Figure 8. $\kappa_2 = (55.43, 38.45, 3695, 32.29, 36.32, 30.94)$ and $\beta_2 = (1.55, 1.07, 1.02, 0.88, 0.98, 0.82)$ for the low-mass (green), full (magenta), and intermediate-mass (teal) flare groups, respectively.

Considering the median stellar masses and characteristic X-ray luminosities of flare host stars for the three mass strata in Figures 8 and 9, the total megaflares energies released per year are $([0.8 - 3.2]) \times 10^{36}, [0.7 - 2.0] \times 10^{37}, [0.5 - 1.4] \times 10^{38}$ erg, respectively, compared to the total “characteristic” energies of $([2.7 - 10^{37}], [8.3 \times 10^{37}], [2.2 \times 10^{39})$ erg, respectively. Megaflares with energies log($E_X$) = [36.2 - 38] erg thus contribute about 3%–11%, 8%–19%, and 17%–39% to the total X-ray energetics of the $\lesssim 5$ Myr old PMS stars in the low-mass (green), full IMF (magenta), and intermediate-mass (teal) flare groups, respectively. As with the occurrence rate, the contribution of megaflares to the total PMS energetics decreases with decreasing flare host mass.

The extrapolation of the PMS flare occurrence rate (Equation 4) toward higher energies statistically allows having monster PMS flares with energies up to $\log(E_X) = 10^{42}$ erg over the first few Myr of PMS evolution. However, the PMS magnetic fields (Sokal et al. 2020) may not be sufficiently strong to provide such enormous flaring energy. For solar-type flares, up to 10% of magnetic energy can be converted to flare energy (Okamoto et al. 2021). One can imagine a PMS star with a radius $R_* = 2.5 \, R_\odot$ hosting a large active region, which covers a quarter of the stellar surface. The region is powered by a surface magnetic field of $B = 5000$ G (currently the maximum measured average PMS field Sokal et al. 2020) with magnetic energy stored in a volume reaching a depth of $0.1 \, R_\odot$. The total magnetic energy would be $\sim B^2 \cdot V / 8 \pi \sim 2 \times 10^{39}$ erg, insufficient to power monster flares. Extreme surface magnetic fields reaching up to 20 kG for cases of powerful flares from most magnetically active stars are predicted by recent theoretical calculations (Zhuleku et al. 2021). Such fields could provide total magnetic energy of $3 \times 10^{40}$ erg, perhaps allowing some monster flares. Because these ideas are only speculative and semi-quantitative, we restrict the analysis of the megaflares energetics contribution to the observed value of $E_{\text{max}} = 10^{38}$ erg. These fractional contributions to the characteristic X-ray emission could be higher if flaring extends to energies above the $\sim 10^{38}$ erg maximum observed here.

7. Comparison of X-ray and Optical-band Superflares

7.1. Main-sequence and PMS Stars

Ilbin et al. (2021) summarize recent flare surveys using Kepler, TESS, Evryscope, and other optical photometry

---

5 Note they use earlier PMS stellar evolution models giving different mass estimates, especially for the lowest-mass range of $M < 0.3 \, M_{\odot}$. 

6 Note they use earlier PMS stellar evolution models giving different mass estimates, especially for the lowest-mass range of $M < 0.3 \, M_{\odot}$. 

---

Getman & Feigelson
Figure 9. X-ray luminosity of MYStIX/SFiNCs young stars as a function of stellar mass (black points). The superflare hosts and their corresponding local regression fits are shown as points and curves in blue, and megflare hosts are shown in magenta. A local regression fit for the MYStIX/SFiNCs stars excluding the super-/megflare hosts is marked by the green curve; these stars exhibit only “characteristic” X-ray emission. The dashed lines roughly delineate the mass ranges of the four source mass strata described in Section 5.

(Lurie et al. 2015; Chang et al. 2015; Ilin et al. 2019; Lin et al. 2019; Raetz et al. 2020; Davenport et al. 2020). The power-law slopes of these flare distributions $dN/dE \propto E^{-\alpha}$ are consistent with our slope of $\alpha = 1.95 \pm 0.07$ (Section 4.2). However, flare energies from most of these surveys do not exceed $\log(E_{\text{opt},f}) < 34–35$ erg. Only the Kepler-detected flares from G-type dwarfs (Shibayama et al. 2013) reach energies up to $\log(E_{\text{opt},f}) \sim 36$ erg. Notsu et al. (2019) show that flares from relatively younger ($t < 500$ Myr) G-type stars are more likely to reach energies of up to $10^{36}$ erg.

The bulk of solar flare energy is generally associated with the optical continuum rather than the soft X-ray band (Schrijver et al. 2012). For young star flares with $\log E < 36$ erg, the flare energy in the optical continuum exceeds the energy in the soft X-ray band on average by a factor of $\sim 6$ (Flaccomio et al. 2018).

Thus, none of these optical studies targeting main-sequence stars have flares with energies comparable to our X-ray megaflares in the energy range $\log(E_{X,f}) \sim 36–38$ erg, and most of the observed main-sequence flares are weaker than our X-ray superflares. This comparison agrees with (Getman et al. 2008a; Figures 8–9), indicating that the COUP superflares are the most powerful among known solar-stellar flares (Güdel 2004; Aschwanden et al. 2008).

For PMS stars, Jackman et al. (2019) report the detection of a single white-light megaflare from the nearby, 2 Myr old, M-type star NGTS J121939.5–355557 with flare energy of $\log(E_{\text{opt},f}) \sim 36.5$ erg. They estimate a crude flare occurrence rate of 0.2–3 flares yr$^{-1}$ for flares with $\log(E_{\text{opt},f}) > 36$ erg. Assuming the factor of 6 optical-to-X-ray energy ratio from Flaccomio et al. (2018), the crude estimate of the flare frequency for the MYStIX/SFiNCs low-mass stars with X-ray energies of $\log(E_{X,f}) > 35.2$ erg (green dashed line in Figure 8) is of several flares per star per year, similar to, but somewhat above, the occurrence rate estimated by Jackman et al.

Note that it is possible that optical superflares would appear less frequent than X-ray superflares for a star with a given flaring rate due to loop geometry. The optical emission is probably emitted at loop footprints near the photosphere whereas X-ray emission is probably emitted throughout the loop that is much larger than a stellar radius (Getman et al. 2008b). As only one hemisphere of the photosphere is visible at a given time, the apparent occurrence rate of optical superflares could be $\sim 2$ times lower than the occurrence rate of X-ray superflares. More discussion of this issue can be found in Flaccomio et al. (2018) who detected several dozen bright X-ray flares with optical and/or mid-IR flare counterparts in PMS members of the NGC 2264 star-forming region. Their estimated fraction of X-ray flares with no optical counterparts varies with the choice of flare morphology, brightness, and energetics from 19% to 48%.

7.2. Superflare Rate Dependence on Stellar Age

In light of extrapolations of flare energies and occurrence rates over wide ranges, and for different flare samples at different bands with different sensitivities, the flare rates as functions of age estimated below can only be considered suggestions rather than reliable results.

The Kepler and TESS monitoring of the young (30–150 Myr) M-type star GJ 1243 (Davenport et al. 2020) gives an occurrence rate of $E_{\text{opt},f} > 10^{34}$ erg flares as roughly 10 flares (star-yr)$^{-1}$. For this flare energy range, the extrapolated optical-to-X-ray energy ratio is around 20 (Flaccomio et al. 2018), giving an X-ray energy of $\log(E_X) \sim 32.7$ erg. Extrapolation of the MYStIX/SFiNCs trend for low-mass PMS stars with ages $< 5$ Myr (green line in Figure 8) down to $\log(E_{X,f}) \sim 32.7$ erg gives a rate of $\sim 13,000$ flares per star per year. Comparing to the rate obtained for GJ 1243, the superflare frequency rate may decrease by a factor of $\sim 1300$ between the ages of $t \lesssim 5$ Myr and 30–50 Myr. However, this rough estimate is based on a single star studied in the optical band in contrast to the large sample we have in the X-ray band for PMS stars.

Based on the optical flares collected by Ilin et al. (2021), the flare rate for $\sim 135$ Myr old low-mass stars in the Pleiades cluster at $E_{\text{opt},f} > 10^{33}$ erg is $\sim 0.4$ flares (star-yr)$^{-1}$. This suggests the superflare frequency from young low-mass stars drops by a factor of $\sim 32,000$ between the ages of $t \lesssim 5$ Myr and $t \gtrsim 100$ Myr. This is consistent with the observed X-ray superflare energetics decrease between the Orion and Pleiades populations (Guarcello et al. 2019). Our current large Chandra survey of a dozen 10–100 Myr old stellar clusters employing homogeneous data sets and methods will further help to quantify flare frequencies at this stellar evolutionary period (K. V. Getman et al., in preparation).

The tentative results for MYStIX/SFiNCs low-mass PMS stars presented here suggest that super- and megaflare occurrence rates with $\log(E_{X,f}) \gtrsim 32.7$ erg decline steeply during the early phases of stellar evolution by a factor of $\sim 1000$ from 5 to 50 Myr and a factor of $\sim 30,000$ from 5 to 135 Myr. Thus, the occurrence rate trend is roughly $t^{-3}$ over this age range. Ilin et al. (2021) report that optical flare rates in open clusters continue to decline over the age range $\sim 100–3000$ Myr. A recent estimate of the superflare occurrence rate on the Sun is $\sim 0.0002$ flares yr$^{-1}$ at a level of $\log(E_{\text{opt},f}) = 34$ erg ($\sim 1000$ class) (Okamoto et al. 2021).
8. Effects of PMS Superflares on the Environments

8.1. Implications for Protoplanetary Disk Photoevaporation

There is little doubt that PMS X-rays efficiently irradiate circumstellar disks during the early PMS phase. X-ray photoevaporation models reproduce reasonably well the line profiles of the [O I] 6300 line as a tracer of warm quasi-neutral disk wind (Picogna et al. 2019). Güdel et al. (2010) “... find indications that the production of [Ne II] emission weakly scales with the X-ray luminosity,” supporting models of disk irradiation by X-rays. Flaischlen et al. (2021) report a negative correlation between the X-ray luminosity and accretion rate for similar mass-age OMC stars as a signature of the X-ray-driven disk photoevaporation. Recurring powerful X-ray flares are also proposed to sustain the observed extended [Ne III] emission in the jets of young stars, such as DG Tau (Liu et al. 2016).

Disk photoevaporation linked to high-energy radiation from the central star is now directly detected in a number of systems, well explained by hydrodynamical calculations of the photoevaporative winds driven by stellar ultraviolet and X-ray emission (Alexander et al. 2014; Picogna et al. 2019). X-rays seem most important in the final stages of disk dispersal (Owen et al. 2013).

Because the MYStIX+SFINCs cluster samples range in age from <0.5 to ~5 Myr (Getman et al. 2014), and because protoplanetary disks have longevities around 2–8 Myr (Richert et al., 2018; references therein), we can directly estimate the total number of superflares that have irradiated a typical disk over a ~5 Myr lifetime from the occurrence rates shown in Figure 8. The result is impressive: Disks around PMS stars with masses ~1 $M_\odot$ will be irradiated by ~1 billion super- and megaflares with energies 34 < log($E_X$) < 38 erg.

If the power-law distribution extends to much higher energies, which may not be true, then a disk around a solar-mass PMS star would experience ~2 x 10^5 flares with energies log($E_X$) > 38 erg. Disks around low-mass PMS stars that will become the populous dM dwarfs will be irradiated by ~1 x 10^9 super-/megaflares with energies 34 < log($E_X$) < 38 erg, and possibly by ~7 x 10^3 flares with energies log($E_X$) > 38 erg.

Photoevaporation from X-ray super-/megaflares may affect planet formation processes. In the early stages, flare X-ray ionization will briefly penetrate into the midplane—and possibly midplane—disk layers. The plasma producing X-rays in these powerful events is unusually hot compared to plasma produced in solar magnetic reconnection events with peak temperatures ranging from 20 to 100 MK and higher (Getman et al. 2008a, 2008b, 2011; Getman et al. 2021). The bremsstrahlung spectrum from protostellar flares less luminous than our superflares has been detected out to energies of ~20 keV (Viehe et al. 2019). X-ray penetration into solar-abundance gas scales approximately with the cube of photon energy and can attain column densities log($N_H$) ~ 25–26 cm$^{-2}$ (Glassgold et al. 2000), sufficient to reach the midplane in the outer regions of some disks (Ilgner & Nelson 2006). Superflare ionization thus has the potential to increase turbulence from magnetorotational instability and induce ion-molecular chemistry in disk interiors. However, the importance of the effect depends critically on recombination rates that are difficult to estimate.

In the later stages, the removal of gas from a protoplanetary disk by flare X-ray photoionization may be sufficient to trigger the streaming instability that rapidly forms pebbles and planetesimals critical to the rapid formation of protoplanets (Lambrechts & Johansen 2012; Carrera et al. 2017). For instance, the gas in disks with initial sizes of 30 au and 100 au around a 1 $M_\odot$ young star can be removed within 1 Myr and 4 Myr, respectively (Liu et al. 2019) based on a gas removal rate by the “characteristic” X-ray emission component $M_{\text{shot}} = 6 \times 10^{-9} \times (L_X/10^{30})^{0.14}$ of 1.8 x 10$^{-8}$ $M_\odot$ yr$^{-1}$ (Owen et al. 2012). The megaflares (log($E_X$/flare) > 36.2 erg) X-ray component alone (Section 6) would increase this rate by >10%–20% and consequently speed up processes of planetesimal and planet formation. Detailed astrophysical calculations are needed to more reliably estimate possible nonlinear effects of short-lived superflare irradiation on disk.

8.2. Implications for Protoplanetary Disk Spallation and Chemistry

After the discovery of X-ray flaring in PMS stars, it was proposed that energetic particles ($E \gtrsim 10$ MeV) associated with these magnetic reconnection events could have produced short-lived radionuclides in the solar nebula through nuclear spallation (Feigelson 1982). Today the evidence indicates that, while many meteoritic radionuclides arose from supernova explosions near the Sun’s natal molecular cloud, some radionuclides were formed by spallation from an “early active Sun” with the elevated flaring behavior seen in PMS stars (Chaussidon & Gounelle 2006). Excess $^{10}$Be in Ca-Al-rich inclusions (CAIs) is particularly important as sufficient quantities cannot form in supernova events (McKeegan et al. 2000). $^{10}$Be abundances vary widely among CAIs and were produced after $^{26}$Al from decayed supernovae; these and other properties (e.g., covariation with $^{56}$Ni, irradiation products in $^{26}$Al-free hibonite-rich CAIs) point to a spallogenic origin by solar energetic particles rather than a presolar source distributed throughout the disk (Fukuda et al. 2019).

Assuming “characteristic” X-ray emission around log($L_X$) ~ 30 erg s$^{-1}$, a rough estimate is that PMS proton fluence is elevated ~10$^5$ above contemporary solar levels (Feigelson et al. 2002). This would be sufficient to produce the observed abundances of spallagenic radionuclides (Rab et al. 2017). The scaling of energetic proton flux to X-ray luminosities for superflares is unknown, so quantitative estimates of the effects of superflare particles on radionuclide production cannot be made at this time.

X-ray ionization should induce ion-molecular chemistry, and an unusual case of variable HCO$^-$ emission can be attributed to X-ray flaring. Cleeves et al. (2017) report variability in H$^2$CO$^-$ line emission in the disk of PMS IM Lup. The implied rapid abundance changes of the HCO$^+$ molecular ion can be explained by X-ray flaring that ionize the $H_2$ gas on the disk surface. This produces $H^+$ followed by the proton transfer reaction with CO to produce HCO$^+$ ions. Their flare-driven disk chemistry simulations involving X-ray flares with energies ~1 x 10$^{36}$ erg results in enhanced HCO$^+$ abundances for a period up to ~20 days.

Simulations of X-ray superflare-driven chemistry in a disk around a young solar-mass star also predict changes in the gas-phase $H_2$O abundance lasting days (Waggoner & Cleeves 2019). Their choice of superflare frequency for log($E_X$) = 37.1 erg flares once “every few years” based on the COUP studies is consistent with our rate of one flare per star every 4 yr (magenta line in Figure 8).
The Astrophysical Journal, 916:32 (27pp), 2021 July 20

Getman & Feigelson

Theoretical calculations predict many other effects of stellar X-ray irradiation on the disk, accretion and outflow astrophysics. These include stimulation of the magnetorotational instability and associated turbulence (Fromang et al. 2002), ionization necessary for launching a magnetocentrifugal disk wind (Gressel et al. 2013) and a collimated jet (Shang et al. 2002), and desorption of water ice from dust grains (Dupuy et al. 2018). The magnetohydrodynamic simulations of Colombo et al. (2019) suggest that superflares may trigger the formation of accretion funnels and influence the morphology of inner disk and accreting columns.

8.3. Implications for Young Planetary Atmospheres

Evidence for early Jovian planet formation emerges from a number of recent observations (Liu & Ji 2020 and references therein) including ALMA detections of compact rings and gaps in \( \lesssim 1 \) Myr old disks (Andrews et al. 2018), optical- and IR-band radial velocity detections of hot Jupiters around the \(~2\) Myr old PMS stars CI Tau and V830 Tau (Johns-Krull et al. 2016; Donati et al. 2016), and direct \( H_{\alpha} \) imaging of accreting protoplanets within the transition disk of the \(~5\) Myr old star PDS 70 (Haffert et al. 2019).

Recent theory based on the streaming instability and pebble accretion indicates that the formation of rocky super-Earths, necessary for the gravitational trapping of disk gas to form Jovian planets, may be extremely rapid (Raymond & Morbidelli 2020). This can be followed by the migration of resonant chains to the inner edge of the disk where the nascent rocky planets can be subject to intense radiation from superflares. Most of the resonant chains become unstable when the disk dissipates but many compact planetary systems survive. These inner rocky planets may have water-rich volatile atmospheres.

Although the super-/megafare occurrence rate rapidly decreases as the PMS star enters its main-sequence phase (Section 7.2), the flares continue after the disk dissipates and can no longer protect the inner planets from flare higher-energy photon irradiation. In addition, the planetary atmospheres can be impacted by coronal mass ejections with much greater total plasma energy than the radiative energy from super-/megafares.

Poppenhaeger et al. (2021) describe a relevant calculation for the four-planet system around the 20 Myr old solar-mass PMS star V1298 Tau. The innermost planet c orbiting 0.08 au (17 \( R_{\oplus} \)) from the star may lose a hypothetical H/He envelope within \(~100\) Myr. If such a planet was orbiting a young \( (\lesssim 5\) Myr) \( 1 M_{\odot} \) star with a “characteristic” X-ray luminosity of \( L_X = 2.5 \times 10^{30} \) erg s\(^{-1}\) (Section 6), then the hydrodynamic escape assumption of \( M = (\pi R_{\text{sun}} R_{\text{pl}} F_{\text{XUV}})/((KG_{\text{pl}})) \) (Owen et al. 2012) with the atmospheric escape efficiency \( \epsilon = 0.1 \), the Roche lobe factor of \( K = 0.8 \), the “fluffy” planetary radius at XUV wavelengths of \( 2 \times \) the radius at optical wavelengths \( (R_{\odot} = 5.6 \) \( R_{\oplus} \)) and the conservatively chosen EUV flux as \( 4 \times F_{X_{\text{ray}}} \) (Sanz-Forcada et al. 2010), would result in the H/He envelope mass-loss rate of \( 0.11 M_{\oplus} \) Myr\(^{-1}\) and complete evaporation of the envelope within \( 4.5 \) Myr.

A similar calculation by Johnstone et al. (2019) suggests even faster destruction of early planetary atmospheres. Considering the effects of extreme ultraviolet irradiating planets around a \(~100\) Myr solar-mass star, they find removal of even heavy-element atmosphere on timescales of \(~0.1 \) Myr by photodissociation and/or hydrodynamic escape.

The addition of the super-/megafarre X-ray emission component would further shorten this atmosphere evaporation process. If the disk gas is removed after \(~5\) Myr, the young planets will experience roughly one billion flares with energies \( 34 < \log(E_X) < 38 \) erg, including several million megafares with energies \( 32.6 < \log(E_X) < 38 \) erg. The megafare effects may be modest (10%-20%; Section 6) if an intense burst of X-rays has the same effect as a continuous irradiation of characteristic X-ray emission. But short-lived superflares may have nonlinear effects. Even at older ages, Atri & Carberry Mogan (2020) find that the superflare emission may dominate envelope loss for \(~20\)% of late-M-type stars. New calculations are needed to evaluate whether short-lived intense bursts of X-rays have the same evaporative effects as a weaker but continuous irradiation of X-rays.

Superflares may have other effects on young planetary atmospheres. Their ozone layer may be depleted or destroyed by stellar energetic particles leading to increased penetration of ultraviolet radiation to the planetary surface (Schaefer et al. 2000; Howard et al. 2019; Tilley et al. 2019). Superflare energetic particles may stimulate nonequilibrium atmospheric chemistry such as the production of nitrous oxide and hydrogen cyanide (Airapetian et al. 2016). This conceivably might promote surface organic chemistry, leading to the formation of life. Finally, energetic superflare photons may improve, rather than destroy, the effectiveness of photosynthesis in inhabited zone planets around late-M stars (Mullan & Bais 2018).

9. Concluding Remarks

The Chandra MYStIX and SFHNCs surveys have produced a sample of \( >30,000 \) X-ray-emitting PMS stars with ages \(<5\) Myr from 42 star-forming regions within \( d<3 \) kpc in the Galactic disk. Omitting the Carina Nebula and Orion Nebula regions, here we examine the X-ray variability among the remaining \( >24,000 \) X-ray young stars. Using a reproducible likelihood-based statistical procedure (Section 3 and Appendix B), we extract over a thousand flares. An atlas of the flare lightcurves and properties is provided (Section 4.1). Peak luminosities lie in the range \( \log(L_X) = 30.5–34.0 \) erg s\(^{-1}\) with total energies \( \log(E_X) = 34–38 \) erg in the Chandra 0.5–8 keV band. The sample is “complete” above \( \log(L_X) > 32.5 \) erg s\(^{-1}\) and \( \log(E_X) > 36.2 \) erg (Section 4.2). This is the largest collection of powerful stellar flares ever assembled in the X-ray band. They are far more luminous than the optical-band flares recently studied in main-sequence stars (Section 7).

We highlight here two themes where the findings influence important astrophysical issues.

9.1. Super-/Megaflares Do Not Arise from Star–Disk Magnetic Fields

A basic result of our study is the ubiquitous nature of the X-ray superflares among PMS stars. Averaged over an IMF ensemble, each PMS star produces several superflares (with energies \( E_X > 10^{34} \) erg) per week and one to three megafares \( (E_X > 10^{36.2} \) erg) per year (Section 6). Because the former flare frequency number is inferred using extrapolation toward lower energies, beyond our completeness limit (dashed magenta line in Figure 8), it should be considered with caution.

Super-/megaflares occur in all star-forming regions, from stars of all masses and at all stages of young stellar evolution. The flares are seen in heavily absorbed Class I protostars with enormous infrared excesses from large protoplanetary disks (Section 5.1), in Class II T Tauri stars that are still accreting.
from their disks, and in Class III diskless stars. The collective superflare energetics are indistinguishable across evolutionary classes (Section 5.2). Our companion paper will show that superflare astrophysical properties (such as peak luminosities, decay timescales, and plasma temperatures, densities, and volumes) similarly are indistinguishable between disk-bearing and diskless stars (Getman et al. 2021).

Our analyses thus provide no evidence for a distinct flaring mechanism involving the circumstellar disk, such as reconnection in field lines connecting the star and disk, or at the boundary between the stellar magnetosphere and the inner disk. Such mechanisms have been speculated from the pioneering scenarios of Hayashi et al. (1996) and Shu et al. (1997) to recent 3D+time magnetodynamical calculations of Colombo et al. (2019). The only remaining links between X-ray emission and disks are indirect, such as the possibilities that X-ray loop sizes are constrained to lie within the inner disk boundary (Getman et al. 2008b), star–disk loops are responsible for a periodic variation in X-rays seen in two ONC superflares (Reale et al. 2018), and X-ray flares trigger increased accretion from the inner disk (Espaillat et al. 2019). If there are magnetic reconnection events involving star–disk magnetic field lines, they do not manifest themselves in detectable super-/megafrares and/or the occurrence rate of such star–disk events is very low.

The similarity of coronal X-ray properties of disk-bearing and diskless stars has been seen in many studies from early observations with ROSAT (Feigelson et al. 1993) to thorough studies with XMM–Newton (Güdel et al. 2007) and Chandra (Preibisch et al. 2005). Studies of the Taurus (Stelzer et al. 2007), NGC 2264 (Flaccomio et al. 2018), Orion Nebula (Getman et al. 2008b; Flaccomio et al. 2012), and MYStIX/SFInCs regions (here) provide observational evidence that young stars with and without disks produce flares with similar properties (Section 5.2). The proposed physical processes responsible for the mild suppression of time-integrated X-ray emission in accreting versus nonaccreting PMS systems, as well as the accretion shocks producing relatively weak soft X-ray excess emission (Section 1) seem to have little or no effects on the production mechanisms and characteristics of coronal flaring including superflares of extraordinary power with \( \log(L_X) = 30.5 \pm 34.0 \text{ erg s}^{-1} \) and total energies \( \log(E_X) = 34.38 \text{ erg.} \)

When combined with solar flare properties (e.g., Aschwanden et al. 2008), the observational evidence acquired in our own studies consistently points to solar-type geometries, magnetic loops with both footprints rooted in the stellar surface, analogs of giant solar X-ray arches and streamers (Getman et al. 2008a, 2008b). The flare processes span a phenomenal range: 2–3 orders of magnitude in flare duration, 4 orders of magnitude in loop length, and 13 orders of magnitude in plasma emission measure (Getman et al. 2011). Often the same distribution functions (Section 4.2) and scaling relations are seen for both weak and powerful flares.

However, rare examples of star–disk flaring may have been found. Reale et al. (2018) report detection of quasi-periodic pulsations in COUP superflares produced by two disk-bearing young stars, V772 Ori7 and OW Ori. Their hydrodynamic flare modeling suggests single flaring loops with sizes significantly exceeding corotation radii, supporting star–disk loop configurations. While individual star–disk flares may exist, the majority of PMS super-/megafrares appear to be associated with magnetic loops anchored in the stellar surface.

9.2. Super-/Megflare Contributions to Disk and Protoplanet Irradiation

Although protoplanetary disks and planets are neutral molecular phases with typical thermodynamic temperatures in the range 100–2000 K, it is widely recognized that they are impacted by external high-energy radiation with potentially enormous effect. PMS stellar X-rays constitute only \( 10^{-3} \) of their bolometric luminosity and therefore have little effect on heating of the bulk circumstellar material. But their ionization induces disk turbulence and no-equilibrium ion-molecular chemistry, traps mostly neutral material to magnetic field lines for accretion or outflows, and photoevaporates outer layers of disk or planetary atmospheres. Furthermore, undetected bursts of energetic particles and coronal mass ejections are likely to accompany super-/megafrares with additional effects on circumstellar gaseous and solid material.

However, it is difficult to quantify these effects due to the uncertain timescales of the astrophysical response to the flare event. In Section 6, we find that the time-integrated ensemble X-ray luminosities of PMS stars is elevated at least 10%–20% by megafrares. This arises from the steep \( \alpha \approx 2 \) slope in the flare energy distribution (Section 4.2). If the response of disk ionization, planetary atmosphere escape, or other effect is much slower than the superflare timescale, then the total effect of the superflares will be modest.

But if the response is rapid, then the effects can be substantial. The observation of variable HCO+ emission in a disk by Clevees et al. (2017) and the calculation of rapidly fluctuating disk “active” and “dead” zones by Ignier & Nelson (2006) suggest that ionization effects can be sufficiently rapid that the astrophysical response is strong. More astrophysical modeling of the time dependency of such effects is needed to gain confidence in any conclusion concerning the importance of super-/megafrares on disk and planetary processes.

We are grateful to the referee for spending time and providing many useful suggestions that stimulated fresh ideas and improved the paper. This project is supported by the Chandra archive grant AR9-20002X and the Chandra ACIS Team contract SV474018 (G. Garmire and L. Townsley, Principal Investigators), issued by the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS8-03060. The Chandra Guaranteed Time Observations (GTO) data used here were selected by the ACIS Instrument Principal Investigator, Gordon P. Garmire, of the Huntington Institute for X-ray Astronomy, LLC, which is under contract to the Smithsonian Astrophysical Observatory, contract SV2-82024. This research has made use of NASA’s Astrophysics Data System Bibliographic Services.

Facility: CXO.

Appendix A: Energetics of the Brightest COUP Flares

Getman et al. (2008a, 2008b) examined and calculated the properties of flaring coronal structures associated with the
sample of the brightest 216 flares from COUP young stars. This is the largest tabulated sample of PMS superflares with modeled loop geometries. In our companion paper (Getman et al. 2021), we model the brightest MYStIX/SFiNCs superflares and compare their inferred properties with those of the COUP flares from Getman et al. (2008a, 2008b).

To better understand the energetics of these COUP flares, here we use their reported peak flare X-ray luminosity ($L_{X,pk}$) and decay e-folding timescale ($\tau_d$) of flare X-ray counts (their Table 2; Getman et al. 2008a) to estimate flare energies $E_X \approx L_{X,pk} \times \tau_d$. We fit the unbinned energies with the Pareto (power-law) function using the standard maximum likelihood procedure (e.g., Newman 2005). Most COUP flares have full observation coverage thanks to the extremely long COUP exposure time; hence no need for a KM estimator.

Figure 10 shows that this 216 COUP flare sample is complete above log($E_X$) $\geq$ 10^{36} erg with the power-law shape of $dN/dE \propto E^{-2.1}$, or equivalently $dN/d\log E \propto E^{-1.1}$. Thus, the COUP superflare energy slope is consistent with the previous analyses of COUP flares by Wolk et al. (2005), Caramazza et al. (2007), Albacete Colombo et al. (2007), and Stelzer et al. (2007) and the distribution of flare energies found here for MYStIX and SFiNCs PMS stars (Section 4.2).

**Appendix B**

**Poisson Regression Model with Multiple Changepoints**

We adopt here a statistical model of PMS X-ray variability as a sequence of stepwise constant flux values in a Poisson counting stochastic process. The statistical problem of multiple changepoint detection in a time series has a substantial history dating back to Quandt (1958) with Bayesian approaches starting with Barry & Hartigan (1993). It was an application in the famous paper by Green (1995) introducing reversible jump Markov Chain Monte Carlo methods.

A unique optimal solution to the maximum likelihood or Bayesian inference problem exists but, unless the number of partitions is known in advance, it is difficult to obtain as there is a vast number of possible partitions of the time series. Standard computational procedures (such as the EM Algorithm) cannot be directly applied because the Poisson multiple-changepoint model is discontinuous and nonlinear, and models with different numbers of changepoints are not nested.

This statistical model is widely used in high-energy astrophysics under the rubric “Bayesian Blocks” proposed by Scargle (1998). Scargle’s original algorithm was approximative and often arrived at significantly suboptimal partitions. An improved computational methodology based on dynamic programming was developed by Scargle in collaboration with a team of computer scientists, Jackson et al. (2005) and Scargle et al. (2013).

The computational procedure we use here is based on the foundational work of econometrician Chib (1998) with developments by statisticians Frühwirth-Schnatter & Wagner (2006). Chib’s procedure reparameterizes the changepoint model as a finite mixture model with latent-state variables with an unknown number of hidden time delimited regimes. It is thus an example of “state space modeling”, a powerful approach for advanced time-series modeling using hierarchical models (Durbin & Koopman 2012). This formulation is described in less technical language by Park (2010); see also the lectures by Brandt (2010). Applications in the social sciences are presented by Brandt & Sandler (2009) and Park (2011).

The statistical model for the Poisson regression model with multiple changepoints is

$$p(\beta, P, s | y) \propto \prod_{i=1}^{T} p(y_i | \beta, P, s) \prod_{i=1}^{M} p(\beta_i) p(p_i) \quad (B1)$$

$$= \text{Poisson}(y_i | \beta_{i}) \times \prod_{i=2}^{T} \sum_{m=1}^{M} \text{Poisson}(y_i | \beta_{m})$$

$$\times \Pr(s_i = m | \beta, P)$$

$$\times \prod_{i=1}^{M} \text{Normal}(\beta_{0,i}, \beta_{0,i}^2) \text{Beta}(a_i, b_i). \quad (B2)$$

Here $y_i$ are the Poisson-distributed counts in time bin centered at time $t$, $m$ is the hidden state at time $t$, $s$ are the latent-state variables with values 1, 2, ...., $M$, and $\beta_{m}$ are the regression parameters (flux levels) for each of the $M$ states. The prior

---

**Figure 10.** Cumulative distribution function of flare energies ($E_X$) left and corresponding histogram visualization of the probability distribution function with 0.2 dex bins right for 216 bright flares in the COUP observation of the Orion Nebula Cluster (data from Getman et al. 2008a). The best-fit Pareto (power-law) function fit to the high end of the c.d.f. is shown in red. Histogram error bars are approximations to 95% confidence intervals of a Poissonian distribution (Gehrels 1986).
choices are a multivariate normal for the regression parameters and a Beta distribution for the transition probabilities.

The calculation proceeds in three steps: hidden-state variables are sampled using Chib’s recursive algorithm, transition probabilities to a new state are drawn from binomial distributions, and state space variables and transition probabilities are sampled using a specialized data augmentation algorithm designed for count data. Software implementation is available in the MCMCpoissonChange function within the MCMCpack CRAN package (Martin et al. 2011) of the R public domain software environment (R Core Team 2020). Calculations are performed in R and C++. Our code and graphics are based on R scripts by Brandt (2010).

For this model, the X-ray counts must be placed into evenly spaced time bins; this is the data structure required for most methods of time-series analysis. We find that the calculated changepoints are insensitive to the choice of bin width; bin sequences can be dominated by zeros and ones, or can have many counts. We chose bin widths individually for each source so that bins typically have \( \sim 2 \) counts for the faintest sources ranging to \( \sim 16 \) counts for the brightest sources. Following the default in MCMCpoissonChange, we adopt the simple Beta(1, 1), or uniform, prior distribution for the transition probabilities.

Results seem resistant to reasonable variations in the prior, but sensitivity to weak flares is reduced for narrow priors like Beta (10, 1). We further restrict model complexity to one through five changepoints, producing two to six stepwise segments. The model with the highest Bayes factor is chosen. The chosen model may not be unique; models with fewer or more changepoints may also give satisfactory fits using the Bayes factor significance thresholds recommended by Kass & Raftery (1995). The algorithm does not permit comparison to a constant (zero-changepoint) model. For weak sources, a spurious short segment is sometimes introduced around the initial bins as the algorithm forces the first point to be zero.

Figure 11 gives an example of the graphical output for the 3143 ObsIDs identified in Section 3.1. The top panel shows the binned lightcurve with 312 counts extracted over a 148.1 ks exposure that were placed into bins with width 0.95 ks. Notice that this exposure time is the arrival time difference between the first and last X-ray counts and thus can be slightly smaller than the actual Chandra observation exposure time (Table 3), which is the difference between the start and stop times of the observation. Most bins have one to three counts but a flare is seen midway through the observation. The best changepoint model has three segments with averages of 0.8, 5.6, and 1.8 counts bin\(^{-1}\). The second panel shows the relative probabilities of the three segments based on Bayes factors for each bin of the lightcurve. The third and following panels show the Bayesian posterior probability distribution functions for the segments in the best-fit model, omitting the first segment. These are essentially the differentials of the second panel distributions.

In most cases, we choose the changepoint time to be the bin where the new component has probabilities exceeding 0.5, that is, when it dominates over all other components. These changepoint times are shown with vertical dashed lines in Figure 11. However, visual examination of the lightcurves sometimes shows low-level variations at the beginning of the flare are missed with this criterion. In these cases, the start time (and more occasionally, stop time) was manually adjusted to include the full flare.

The Chib (1998) formulation of the Poisson multiple-changepoint problem, with the software implementation of Martin et al. (2011) and Brandt (2010), should give solutions very similar to the Bayesian Blocks formulation of Scargle et al. (2013). Its graphical output, illustrated in Figure 11, allows flexible scientific interpretation. We have chosen changepoints when a component’s probability crosses the 50% boundary, but another user might choose a 90% criterion for a more conservative evaluation of change. There are also situations where several components simultaneously contribute to the model; caution might be warranted in selecting changepoints in such cases.
Appendix C
Derivation of Stellar Properties

New parallax-based distances are obtained for many MYStIX and SFiNCs star-forming regions by Cantat-Gaudin et al. (2018), Kuhn et al. (2019), and other recent Gaia-based studies. Details are given in Table 1. These new distances allow us to recalculate X-ray luminosities from Chandra fluxes, which are derived from Chandra count rates and median energies using the scalings of Getman et al. (2010).

Preliminary PMS stellar masses are estimated from the empirical scaling relation between X-ray luminosity and stellar mass obtained for the nearby Taurus star-forming region with the XMM X-ray telescope (Telleschi et al. 2007a).

For solar- and lower-mass stars, visual-band absorptions, $A_V$, are estimated by dereddening near-infrared colors to the intrinsic color locus of the Taurus low-mass stars in the $J - H$ versus $H - K_s$ color–color diagram. Details of the procedure appear in Getman et al. (2014).

Ages of these low-mass MYStIX/SFiNCs young stars are estimated in two ways. For stars with $\geq 10$ X-ray source photons, the Age$_{X}$ X-ray/near-infrared chronometer developed by Getman et al. (2014) is used. In this method, the X-ray luminosities (as surrogates for stellar masses using the Telleschi et al. relation) and $J$-band magnitudes corrected for extinction (as surrogates for bolometric luminosities) are combined with PMS evolutionary models to obtain individual
stellar ages. The inferred \( \text{Age}_{JX} \) for MYStIX/SFiNCs stellar clusters are consistent with independently known patterns of star formation histories, cluster sizes, and cluster disk fractions (Getman et al. 2014; Kuhn et al. 2015a; Richert et al. 2018). In our recent work by Richert et al. (2018) such ages were calculated for the three different evolutionary models with different treatments of interior magnetic fields (Siess et al. 2000; Choi et al. 2016; Feiden 2016). In the current work, we choose to recompute ages using the popular PARSEC 1.2S evolutionary models (Bressan et al. 2012; Chen et al. 2014). The empirical changes to the relationship of the temperature and mean optical depth across a stellar atmosphere implemented in these models may mitigate the “radius inflation” problem related to effects of surface starspots and interior magnetic pressure on the stellar structure, slowing convective energy flux and global contraction, and changing the location of PMS isochrones in the HRD (Morrell & Naylor 2019).

For low-mass stars without \( \text{Age}_{JX} \) estimates (due to very weak X-ray sources or inaccurate near-infrared photometry), and all intermediate- and high-mass stars, we assign age values to be the median ages among the seven nearest (on the sky) low-mass \( \text{Age}_{JX} \) neighbors. For cases when young stellar objects are members of highly embedded MYStIX/SFiNCs subclusters without known \( \text{Age}_{JX} \) stellar members, age estimates are obtained from the \( \text{Age}_{JX} \)–(\( J-H \)) relationship (Getman et al. 2014) transformed to the PARSEC 1.2S scale. These ages are truncated at \( \geq 0.4 \) Myr (due to the paucity of \( <0.4 \) Myr sources on the age–color diagram) and \( \leq 5 \) Myr (due to the degeneracy of PMS isochrones on the \( M_J-L_X \) diagram).

On the \( J \) versus \( J-H \) diagram (which is insensitive to the presence of circumstellar disks), we deredden each MYStIX/SFiNCs star with reliable photometry toward the intrinsic color–magnitude locus corresponding to a PARSEC 1.2S isochrone of the star’s estimated age. This gives estimates of \( A_V \), stellar \( T_{\text{eff}} \), \( L_{\text{bol}} \), \( R \), and \( M \). Such estimates are obtained for \( >26,000 \) out of \( \sim 40,000 \) MYStIX+SFiNCs young stellar objects.

Due to the degeneracy of PMS isochrones at the intermediate-mass range, the inferred properties of some stars are associated with ranges of values rather than single values. To overcome this problem, a few additional steps were taken. First, all bright (\( J<15 \) mag) MYStIX+SFiNCs stars are passed through the VOSA (Bayo et al. 2008) using (in addition to our \( JHK_s \) and Spitzer-IRAC photometry) data from numerous other optical and IR photometric catalogs, such as Gaia-DR2 (Gaia Collaboration et al. 2016, 2018), Pan-STARRS (Chambers et al. 2016), SDSS (Alam et al. 2015), VPHAS-DR2 (Drew et al. 2016), APASS-DR9 (Henden et al. 2015), CMC14 (Copenhagen University Institute of Astronomy Real Instituto Y Observatorio de La Armada et al. 2006), Tycho-2 (Høg et al. 2000), VVV-DR2 (Minniti et al. 2017), DECam (DePoy et al. 2008), and others. These star’s SEDs were fit with the BT-Settl atmospheric model (Allard et al. 2012) providing independent \( T_{\text{eff}} \) and \( L_{\text{bol}} \) estimates. Second, the \( JHK_s \) and VOSA-based outcomes are compared with each other. Third, for stars in several MYStIX+SFiNCs regions with published optical-IR spectroscopy (Getman et al. 2005; Skiff 2014; Venuti et al. 2018; Yao et al. 2018), their effective temperature and bolometric luminosity are compared with the photometric outcomes of both the \( JHK_s \) and VOSA-based methods. These extra steps allow selection of most likely \( T_{\text{eff}} \) and \( L_{\text{bol}} \) solutions among the \( JHK_s \) and VOSA choices for intermediate-mass stellar candidates. Most of these solutions are based on the VOSA modeling.

Specifically, X-ray-/NIR-derived and VOSA-derived \( T_{\text{eff}} \) and \( L_{\text{bol}} \) quantities are available for 26,681 and 12,183 (with \( J<15 \) mag) MYStIX/SFiNCs stars, respectively. In the range log(\( T_{\text{eff}} \)) = [3.4–3.8]K (roughly [0.1–3] \( M_\odot \)), the median and interquartile ranges (IQRs) of the log(\( T_{\text{eff,xray,NIR}} \)) – log(\( T_{\text{eff,VOSA}} \)) differences are 0.02 and 0.08, respectively. This log(\( T_{\text{eff}} \)) range corresponds to the log(\( L_{\text{bol}} \)) range of \([1.5 \leq 2] L_\odot \). The median and IQR of the log(\( L_{\text{bol,xray,NIR}} \)) – log(\( L_{\text{bol,VOSA}} \)) differences are \(-0.04 \) and 0.32, respectively. Hence, the distributions of the \( T_{\text{eff}} \) differences typically have small biases (5%) and dispersions (10%). The distributions of the \( L_{\text{bol}} \) differences have small biases (10%) but high spreads (100%). For the regions with published optical-IR spectroscopy, in the intermediate-mass range (log(\( T_{\text{eff}} \)) = [3.8–4.0]K), where the PMS isochrones on the \( J \) versus \( J-H \) diagram are degenerate, the VOSA-derived \( T_{\text{eff}} \) and \( L_{\text{bol}} \) quantities are reasonably well consistent with those inferred from the optical-IR spectroscopy. For 130 MYStIX/SFiNCs X-ray stars, which lie in this degeneracy locus and have highly uncertain properties obtained with the X-ray–NIR method, their final, chosen stellar properties are those derived with the VOSA method. There are also 23 X-ray stars that lie in this degeneracy locus but have unique solutions inferred using the X-ray–NIR method itself. Overall, among the 26,681 MYStIX/SFiNCs stars with available \( T_{\text{eff}} \) and \( L_{\text{bol}} \) estimates, 130 and 26,551 stars have their final (used in Section 5.3) properties obtained with the VOSA and X-ray–NIR methods, respectively.

Figure 12 compares the stellar effective temperatures and bolometric luminosities emerging from this analysis with previously published \( T_{\text{eff}} \) and \( L_{\text{bol}} \) values based on optical spectroscopy for several nearby MYStIX+SFiNCs regions. Orion Nebula Cluster values are compared to Getman et al. (2005), NGC 2264 values are compared to Venuti et al. (2018), and NGC 1333, IC 348, and Orion A values are compared to Yao et al. (2018). The comparison of red and blue curves shows that the bias is typically \(<200 K\) in \( T_{\text{eff}} \) and 0.1 dex in log(\( L_{\text{bol}} \)). However, the scatter of individual stars is larger with IQRs around \( \pm 500 K \) in \( T_{\text{eff}} \) and \( \pm 0.4 \) dex in log(\( L_{\text{bol}} \)).

The bottom-right panel of Figure 12 compares the color–\( T_{\text{eff}} \) relation for four studies. The black curve is from the PARSEC 1.2S model used in this study. The comparison green curve is from Getman et al. (2005), the blue curve is from Venuti et al. (2018), and the red curve is from the popular in the recent literature transformation by Pecaut & Mamajek (2013). This diagram shows that there is little disagreement between several color–temperature transformations for solar- and intermediate-mass stars, but discrepancies are present for cooler stars. Such discrepancies may contribute to the biases seen between our and spectroscopic-based estimates for Orion and NGC 2264 stars.

We conclude that the stellar properties derived here, based on X-ray and \( JHK_s \) photometry and modern stellar interiors models that account (directly or indirectly) for the effects of magnetic fields, are reasonably accurate to obtain trends in stellar properties related to X-ray superflare occurrence (Section 5 and 5.3). Considerable spreads are present in the \( T_{\text{eff}} \) and \( L_{\text{bol}} \) values, but we recall that \( \sim 500 K \) systematic uncertainties for \( T_{\text{eff}} \) values in PMS stars can be present even in spectroscopic studies (see Figure 3(a) in Yao et al. 2018).
Figure 12. Comparison of $T_{\text{eff}}$ and $L_{\text{bol}}$ estimates for individual stars in selected star formation regions. The ordinate gives values derived in this study based on X-ray and infrared photometry, while the abscissa gives values from a published study based on optical spectroscopy (see text for details). Blue lines show equal values, and red curves show local regression fits to the median ordinate values. The bottom-right panel presents the temperature-color scales for young stars in four studies; see text for details. The legends give the numbers of involved stars.
