Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **498,** 1397–1405 (2020) doi:10.1093/mnras/staa2450
Advance Access publication 2020 August 17

A unified picture of Galactic and cosmological fast radio bursts

Wenbin Lu [1]★ Pawan Kumar[2] and Bing Zhang[3]

[1]*Theoretical Astrophysics, Walter Burke Institute for Theoretical Physics, Mailcode 350-17, Caltech, Pasadena, CA 91125, USA*
[2]*Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA*
[3]*Department of Physics and Astronomy, University of Nevada, Las Vegas, Las Vegas, NV 89154, USA*

Accepted 2020 August 10. Received 2020 August 7; in original form 2020 July 27

ABSTRACT

The discovery of a fast radio burst (FRB) in our Galaxy associated with a magnetar (neutron star with strong magnetic field) has provided a critical piece of information to help us finally understand these enigmatic transients. We show that the volumetric rate of Galactic-FRB like events is consistent with the faint end of the cosmological FRB rate, and hence they most likely belong to the same class of transients. The Galactic FRB had an accompanying X-ray burst, but many X-ray bursts from the same object had no radio counterpart. Their relative rates suggest that for every FRB there are roughly 10^2–10^3 X-ray bursts. The radio light curve of the Galactic FRB had two spikes, separated by 30 ms in the 400–800 MHz frequency band. This is an important clue and highly constraining of the class of models where the radio emission is produced outside the light cylinder of the magnetar. We suggest that magnetic disturbances close to the magnetar surface propagate to a distance of a few tens of neutron star radii where they damp and produce radio emission. The coincident hard X-ray spikes associated with the two FRB pulses seen in this burst and the flux ratio between the two frequency bands can be understood in this scenario. This model provides a unified picture for faint bursts like the Galactic FRB as well as the bright events seen at cosmological distances.

Key words: plasmas – waves – stars: magnetars – radio continuum: transients – fast radio bursts.

1 INTRODUCTION

On 2020 April 28, the Canadian Hydrogen Intensity Mapping Experiment (CHIME, 400–800 MHz) and the Survey for Transient Astronomical Radio Emission 2 (STARE2, 1.3–1.5 GHz) independently detected a fast radio burst (hereafter, FRB 200428), which is spatially coincident with the well-known Galactic soft gamma-ray repeater (SGR) 1935+2154 (The CHIME/FRB Collaboration 2020b; Bochenek et al. 2020a). The arrival time difference between these two frequency bands is consistent with dispersive delay due to plasma along the line of sight with dispersion measure $\mathrm{DM} = 332.8 \pm 0.1\,\mathrm{pc\,cm^{-3}}$. The burst had two $\sim$1 ms components separated by about 30 ms as measured by CHIME; the first at lower frequencies (400–550 MHz) and the second at higher frequencies (550–800 MHz). The FRB occurred in a side lobe of CHIME, so its inferred fluence of a few hundred kJy ms may suffer large uncertainty (The CHIME/FRB Collaboration 2020b). However, STARE2 provided a more accurate fluence measurement of 1.5 MJy ms (Bochenek et al. 2020a).

The SGR 1935+2154 was first detected by *Swift* with a burst of γ-rays (Stamatikos et al. 2014). Subsequent X-ray follow-up observations identified this source as a magnetar with rotational period $P = 3.24$ s and characteristic surface dipolar magnetic field $B \simeq 2.2 \times 10^{14}$ G (Israel et al. 2016). This magnetar has had multiple episodes of outbursts since the initial discovery (Lin et al. 2020a). SGR 1935+2154 is spatially associated with the supernova remnant G57.2+0.8 (Sieber & Seiradakis 1984; Gaensler 2014; Surnis et al. 2016), which is at a distance between 6.7 and 12.5 kpc from us (Kothes et al. 2018). We adopt $d \simeq 10$ kpc, but our results are unaffected by the distance uncertainty.

A hard X-ray burst was detected from SGR 1935+2154 by several instruments including INTEGRAL (Mereghetti et al. 2020), Insight-HXMT (Li et al. 2020), AGILE (Tavani et al. 2020), Konus-Wind (Ridnaia et al. 2020), and the arrival time is in agreement with that of the FRB after de-dispersion. The X-ray burst had fluence of $7 \times 10^{-7}\,\mathrm{erg\,cm^{-2}}$ in the 1–150 keV range, and the light curve in the hardest band (27–250 keV) of HXMT showed two distinct peaks separated by about 30 ms (Li et al. 2020), further confirming the association with the FRB 200428. The isotropic energy (νE_ν) ratio between the radio and the hard X-ray bands is $\sim 3 \times 10^{-5}$.

Understanding the origin of FRBs – mysterious bright millisecond-duration radio flashes first discovered about a decade ago (Lorimer et al. 2007) – has been a major scientific goal of many current or future telescopes, such as Parkes (Thornton et al. 2013; Bhandari et al. 2018), Arecibo (Spitler et al. 2016), UTMOST (Caleb et al. 2017), ASKAP (Shannon et al. 2018), CHIME (CHIME/FRB Collaboration 2019), FAST (Li, Nan & Pan 2013), and DSA (Ravi et al. 2019). Before the discovery of FRB 200428, all localized FRBs were from cosmological distances (Chatterjee et al. 2017; Bannister et al. 2019; Prochaska et al. 2019; Ravi et al. 2019; Marcote et al. 2020). Even with precise localizations of these events in their host galaxies, it is so far inconclusive what the progenitors of FRBs are and by what process the powerful radio emission is generated. Many ideas have been proposed (see Katz 2018; Petroff, Hessels & Lorimer 2019; Cordes & Chatterjee 2019, for recent reviews). They fall into two general categories: (1) emission within the magnetosphere of a neutron star (NS), and (2) emission from a relativistic outflow that interacts with the surrounding medium at large distances from the NS or black hole.

★ E-mail: wenbinlu@caltech.edu

The isotropic specific energy of FRB 200428, $E_\nu \sim 2 \times 10^{26}$ erg Hz^{-1} (for a distance of 10 kpc), is about a factor of ∼30 lower than the faintest burst detected from FRB 180916 at cosmological distances (Marcote et al. 2020), but exceeds that of the brightest known giant radio pulses from NSs by four orders of magnitude. Apart from this energetic argument, we provide further evidence based on volumetric rate (in Section 2) that FRB 200428 belongs to the faint end of the cosmological FRB population. Therefore, the detection of FRB 200428 in the Milky Way provides an extraordinary opportunity to understand the FRB phenomenon in the following three major aspects: (1) strongly magnetized NSs or magnetars can make FRBs (as advocated by many authors, e.g. Popov & Postnov 2010; Kulkarni et al. 2014; Lyubarsky 2014; Katz 2016; Beloborodov 2017; Kumar, Lu & Bhattacharya 2017; Metzger, Margalit & Sironi 2019; Wadiasingh & Timokhin 2019), (2) the associated X-ray emission (and future identifications of other counterparts) provides valuable clue for the emission mechanism, (3) the close proximity may allow us to disentangle many of the propagation effects from the intrinsic emission properties.

This paper aims to explore the implications of FRB 200428. In Section 2, we compare the rate of FRB 200428-like events with that of the cosmological FRB population. In Section 3, we compare SGR 1935+2154 with the sources of other actively repeating FRBs and discuss how they may be understood in a general framework of the magnetar progenitors from different formation channels. In Section 4, we compare the rates of magnetar X-ray bursts and FRBs, and discuss the physical link between them. Finally, we closely examine the possible emission mechanisms for FRB 200428 and for other FRBs in Section 5. A brief summary is provided in Section 6. We use the widely adopted, convenient, subscript notation of $X_n \equiv X/10^n$ in the CGS units.

2 FRB VOLUMETRIC RATE DENSITY

Based on a single detection in about 1 yr of STARE2 operation (Bochenek et al. 2020b), we roughly estimate the Galactic FRB rate to be $\sim 10\,\mathrm{yr}^{-1}$ above specific energy of $E_\nu \sim 5 \times 10^{25}$ erg Hz^{-1} (the detection threshold energy for a distance of 10 kpc). This leads to a volumetric rate of $\sim 10^8$ Gpc^{-3} yr^{-1} (see Bochenek et al. 2020a, for a more detailed calculation). This should be compared with the bright end of rate density distribution $R \sim 10^{2.6}$ Gpc^{-3} yr^{-1} above $E_\nu = 10^{32}$ erg Hz^{-1} as measured by ASKAP also at 1.4 GHz (Shannon et al. 2018; Lu & Piro 2019). We find the slope for the cumulative rate distribution to be $\beta \simeq \Delta\log R/\Delta\log E_\nu \simeq 0.8$, which is insensitive to Poisson error of a factor of a few. This agrees with the slope of the rate distribution found within the (small) ASKAP sample $0.3 \lesssim \beta \lesssim 0.9$ (Lu & Piro 2019) as well as the joint analysis of the Parkes and ASKAP samples $0.5 \lesssim \beta \lesssim 1.1$ (Luo et al. 2020). This agreement suggests that FRB 200428 contributes a significant fraction of the FRB rate density at the faint end near $E_\nu \sim 10^{26}$ erg Hz^{-1}, as illustrated by Fig. 1. Combining this with the fact that the specific energy of FRB 200428 is only a factor of ∼30 below the faintest known cosmological FRB (Marcote et al. 2020), we conclude that the magnetar nature of the progenitor and emission mechanism of FRB 200428 is likely representative of the whole FRB population.

3 NATURE OF FRB PROGENITORS

The number density of Galactic magnetars, SGR 1935+2154 being one of them, is of the order of 3×10^8 Gpc^{-3}. The progenitors of highly active repeaters like FRB 180916 (CHIME/FRB Collaboration 2019; Marcote et al. 2020) are much rarer in the Universe

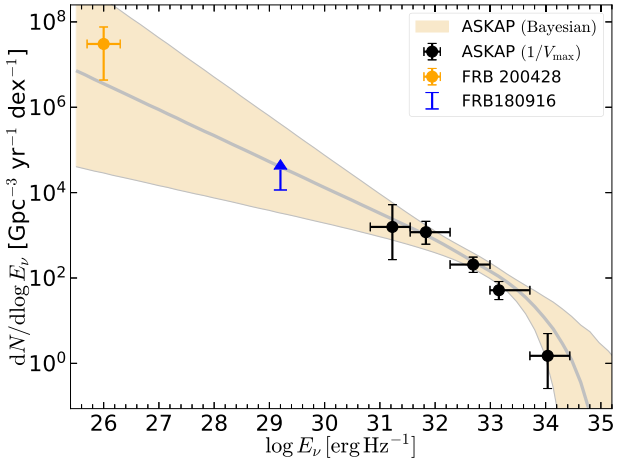

Figure 1. The volumetric rate at the faint end as inferred from FRB 200428 (orange point with 68 per cent C.L. Poisson errors, Bochenek et al. 2020a), as compared to the rate at the bright end inferred from the ASKAP sample (the shaded region). The silver lines mark the 16 (-1σ), 50 (median), and 84 per cent ($+1\sigma$) percentiles based on the Bayesian posterior shown in fig. 4 of Lu & Piro (2019) and evaluated at redshift $z = 0.3$ (where the ASKAP constraints are the strongest). We do not expect significant rate evolution between $z = 0.3$ and the local Universe at $z = 0$. The black points (with 68 per cent C.L. Poisson errors) are from an independent analysis of the ASKAP sample based on the classical $1/V_{\rm max}$ estimator (Schmidt 1968). The blue arrow shows the 90 per cent C.L. lower limit for the contribution to the total volumetric rate density by FRB 180916 (The CHIME/FRB Collaboration 2020a), although this is measured at ∼0.6 GHz rather than 1.4 GHz.

with a number density of ∼7–700 Gpc^{-3}, which is estimated by the expectation number between 0.05 and 4.7 (90 per cent C.L., Gehrels 1986) of such repeaters in about half of the sphere (visible by CHIME) within 150 Mpc. If these active repeaters are also powered by magnetars, they must belong to a type of 'active magnetars' not seen in the Milky Way. If one assumes that all active magnetars will evolve to normal magnetars over time by reducing the bursting rate, the volume density of these active magnetars' 'descendants' would be at most a factor of $\sim 10^4/30 \sim 300$ times greater than the volume density of active magnetars, where 10^4 yr is the typical age of SGR 1935+2154-like Galactic magnetars and ∼30 yr is a conservative estimate of the characteristic age of active magnetars. This gives a volume density of $\sim 2 \times 10^3$–2×10^5 Gpc^{-3} of these descendants, still 3–4 orders of magnitude smaller than the Galactic magnetar volume density. The discrepancy is even larger if the characteristic age of active magnetars is longer than 30 yr. This deficit cannot be fully reconciled by reasonable beaming correction,[1] because we would not have seen FRB 200428 from one of the ∼30 magnetars in our Galaxy if the average beaming fraction is $\ll 1/30$. We can then draw the conclusion that 'active magnetars' and SGR 1935+2154-like Galactic magnetars must be two distinct populations (as also suggested by Margalit et al. 2020b).

One possibility is that the progenitors of FRB 180916 (or FRB 121102, Spitler et al. 2016) may be produced from rare, extreme explosions such as long gamma-ray bursts (LGRBs), superluminous supernovae (SLSNe; e.g. Metzger, Berger & Margalit 2017), or NS

[1]Here, beaming correction is given by the *total* beaming factor, which is defined as the fraction of the whole 4π sky occupied by the union of the solid angles spanned by all bursts from a given source. Due to the star's rotation, the *total* beaming factor is typically substantially larger than that for individual bursts.

mergers (e.g. Margalit, Berger & Metzger 2019; Wang et al. 2020a), so that they have relatively short (e.g. millisecond) periods at births (Usov 1992; Zhang & Mészáros 2001; Metzger et al. 2011). These magnetars likely stored more toroidal magnetic energy inside the star, which provides a larger energy reservoir to power bursting activities (e.g. Thompson & Duncan 1993). In contrast, Galactic magnetars were likely born with a more moderate initial spin, as evidenced by the limited energy in their surrounding supernova remnants (e.g. Vink & Kuiper 2006). These magnetars may store less toroidal magnetic energy inside the star and are relatively less active compared with their active cousins. The possible dichotomy of FRB magnetar progenitor is consistent with the host-galaxy data of the localized FRBs (Li & Zhang 2020), whereas FRB 121102 has a host galaxy similar to that of LGRBs or SLSNe (Metzger et al. 2017; Nicholl et al. 2017; Tendulkar et al. 2017); other four hosts resemble the Milky Way Galaxy that hosts regular magnetars (Bannister et al. 2019; Prochaska et al. 2019; Ravi et al. 2019; Marcote et al. 2020).

4 LINK BETWEEN X-RAY AND RADIO EMISSION

The hard X-ray burst associated with FRB 200428 was one of the numerous X-ray bursts that SGRs generate during their active periods. The ratio of the energy release in the radio and X-ray bands is $f_{\rm r} \sim 3 \times 10^{-5}$. In the following, we discuss the implications on the physical link between emission in these two bands and possible beaming of the radio emission.

The Galactic SGR X-ray burst rate is of the order of $0.1\,{\rm yr}^{-1}$ (volumetric rate $\sim 2 \times 10^6\,{\rm Gpc}^{-3}\,{\rm yr}^{-1}$) above $E_{\rm x} = 10^{44}$ erg, and the energy dependence has a similar power-law form as that for FRBs (e.g. Ofek 2007; Kulkarni et al. 2014; Beniamini et al. 2019). For a fiducial value of the radio-to-X-ray flux ratio $f_{\rm r} = 10^{-4} f_{\rm r,-4}$ to connect the rates of X-ray bursts to FRBs (Chen, Ravi & Lu 2020), $E_{\rm x} = 10^{44}$ erg corresponds to FRB specific energy of $E_\nu \simeq 10^{31} f_{\rm r,-4}\,{\rm erg\,Hz}^{-1}$ (for 1 GHz bandwidth), above which the volumetric rate is $\sim 3 \times 10^3 f_{\rm r,-4}^{-0.8}\,{\rm Gpc}^{-3}\,{\rm yr}^{-1}$ (Lu & Piro 2019; Luo et al. 2020). We see that only a small fraction (10^{-3} to 10^{-2}) of X-ray bursts may be associated with FRBs. This also agrees with the fact that 29 of the X-ray bursts from SGR 1935+2154 had concurrent observations by FAST, but no radio signal was detected down to fluence limit of $\sim$10 mJy ms (Lin et al. 2020b).

This small fraction of association may be explained by (a combination of) the following two possible reasons. The first is that most X-ray bursts are accompanied by an FRB, but the radio emission is highly beamed, with a beaming fraction of $\Omega_{\rm frb}/4\pi \sim 10^{-3}$–$10^{-2}$. This may be realized if FRBs are only generated along magnetic field lines near the poles. The second explanation is that only a small fraction of X-ray bursts may be physically associated with FRB, but in each association the FRB beaming fraction is order unity.

In the next section, we discuss the implications of the association between FRBs and magnetar X-ray bursts on the coherent emission mechanism.

5 EMISSION MECHANISM

Models for the generation of FRB coherent radio emission can be divided into two broad classes based on the distance from the NS where they operate. The first class consists of the 'far-away' models, where relativistic ejecta from an NS (or black hole) dissipates its energy at large distances by interacting with the circumstellar medium (CSM) and the radio emission is generated by a maser process (Lyubarsky 2014; Beloborodov 2017, 2019; Waxman 2017; Metzger et al. 2019; Margalit, Metzger & Sironi 2020a). The second class are the 'close-in' models, which describe that the coherent processes occur within the magnetosphere of an NS (Pen & Connor 2015; Cordes & Wasserman 2016; Lyutikov, Burzawa & Popov 2016; Kumar et al. 2017; Zhang 2017; Lu & Kumar 2018; Yang & Zhang 2018; Kumar & Bošnjak 2020; Lyubarsky 2020; Wang, Xu & Chen 2020b). These two general classes of models have very different predictions regarding the FRB temporal and spectral properties, and multiwavelength counterparts. In the Appendix, we present a detailed analysis of the 'far-away' models and show that they face a number of difficulties explaining the available radio and X-ray data for FRB 200428.

In this section, we focus on the generation of FRB radiation in the magnetosphere of a magnetar. An additional motivation for our consideration of this model is that at least some FRBs show very rapid variability time, as short as tens of microseconds (Farah et al. 2018; Prochaska et al. 2019; Cho et al. 2020), which corresponds to the light-crossing time of a few km and suggests that the radiation might be produced close to an NS.[2]

The basic scenario we suggest is that a disturbance emanating from the NS surface spreads through the magnetosphere. The dissipation of the energy near the surface in the closed field line region produces X-ray emission. The disturbance propagating to distances much larger than the NS radius, above the magnetic poles, is converted into coherent radio waves (Fig. 2).

Let us first consider that the energy in the outburst near the surface of the NS is carried by a beam of $e^\pm$ pairs of isotropic equivalent luminosity $L_{\rm b}$ and Lorentz factor $\gamma_{\rm b}$. The $e^\pm$ number density at distance $R = 10^8 R_8$ cm from the NS in the beam comoving frame is given by $n_{\rm b} \sim 3 \times 10^{16}\,{\rm cm}^{-3}\,R_8^{-2}\gamma_{\rm b}^{-2}$ for $L_{\rm b} = 4\pi R^2 \gamma_{\rm b}^2 n_{\rm b} m_e c^3 \sim 10^{38}\,{\rm erg\,s}^{-1}$, which is the minimum particle beam luminosity so as to generate the observed radio flux of FRB 200428. This corresponds to a plasma frequency of $\nu_{\rm p} \sim 10^{13} R_8^{-1}$ Hz in the observer frame. Moreover, the cyclotron frequency is $\nu_{\rm B} \sim 3 \times 10^{15} R_8^{-3}$ Hz for surface dipolar magnetic field strength of $B_{\rm ns} = 10^{15}$ G. Most maser processes resulting from an interaction between highly relativistic beam of particles and mildly or sub-relativistic plasma produce radiation near the plasma frequency or the appropriately Doppler shifted cyclotron frequency. The estimates for these frequencies show the difficulty for the particle beam based class of maser models to produce GHz radiation with the observed luminosity of FRB 200428.

We consider another possibility, that the energy released near the polar region of the NS is carried by magnetic disturbances – Alfvén waves – which damp far away from the surface, but well inside the light cylinder, and produce radio waves (Kumar & Bošnjak 2020). Let us consider that the amplitude of the Alfvén wave at the NS surface is δB and its transverse wavenumber is $k_\perp = 2\pi/\lambda_\perp$, where $\lambda_\perp$ is the wavelength perpendicular to the NS magnetic field. Both δB and $k_\perp$ decrease with radius as $R^{-3/2}$, as the wave packet follows the curved magnetic field lines and fans out such that its transverse size increases as $R^{3/2}$. The wave becomes charge starved at a radius R, where the plasma density is below the critical density

$$n_c = \frac{|\nabla \times \delta \mathbf{B}|}{4\pi q} \approx \frac{k_\perp \delta B}{4\pi q} \simeq (1 \times 10^{12}\,{\rm cm}^{-3})\, R_7^{-3} \frac{\delta B_{10}}{\lambda_{\perp 4}}, \quad (1)$$

[2]The transverse size of the source and the distance from the compact object is larger when the radiation is produced in a relativistic outflow moving towards the observer with a Lorentz factor γ by a factor γ and γ^2, respectively (e.g. Katz 2019).

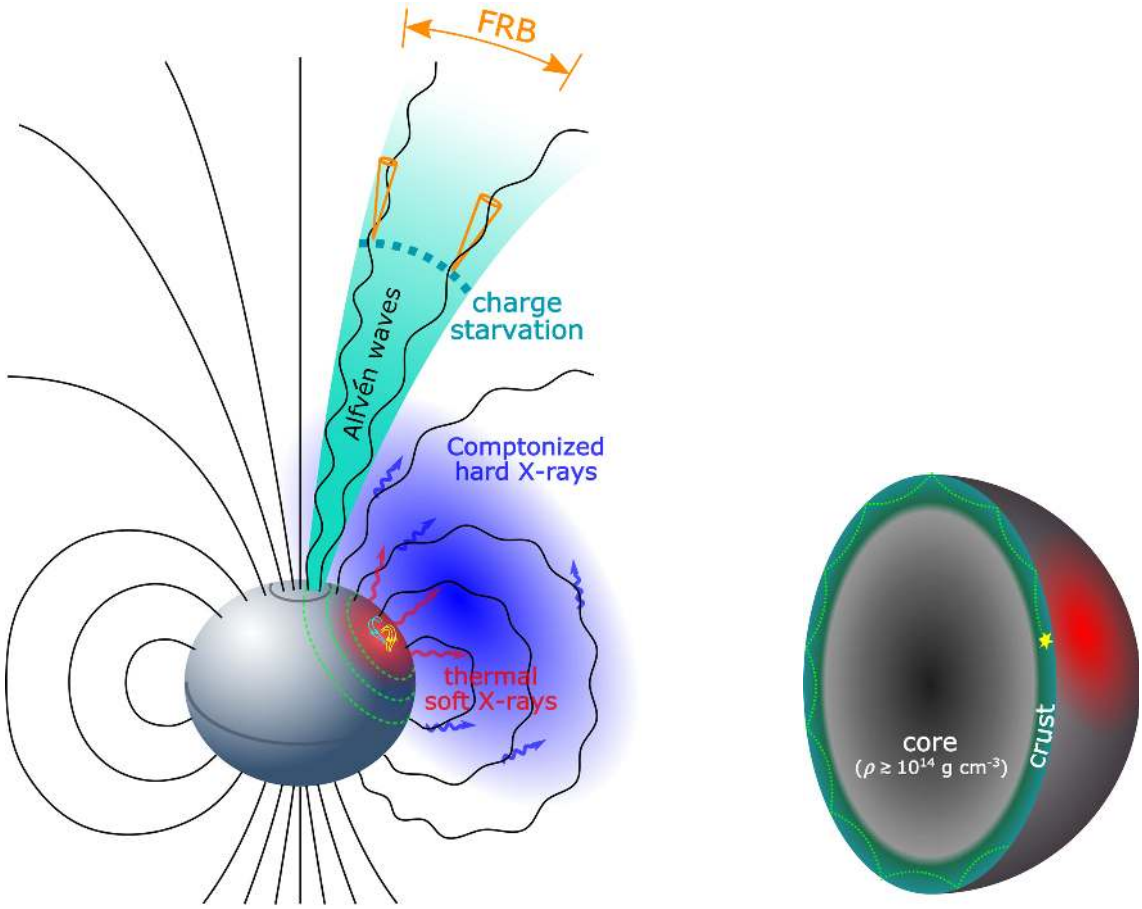

Figure 2. Sketch of the model described in this paper. *Left-hand panel*: Sudden magnetic energy dissipation heats up the NS surface and generate $e^{\pm}$ pair fireball, which is trapped by the closed field lines. X-rays are produced by the heated surface (red shaded region) and then inverse Compton scattered by $e^{\pm}$ pairs (blue shaded region) in the magnetosphere to higher energies. The disturbance spreads across the NS surface (green dashed circles) and launches Alfvén waves (shown as wiggles with exaggerated amplitude), which propagate along magnetic field lines. Near the magnetic poles, Alfvén waves can reach distances much larger than the NS radius, where the charge density is too low to sustain the plasma current associated with the wave (marked by a teal dashed curve). This is because the plasma density in the magnetosphere drops rapidly with the distance to the NS. As a result of charge starvation, a strong electric field parallel to the background magnetic field develops, and charge clumps are accelerated to high Lorentz factors and coherently produce curvature emission in the radio band (marked as orange cones). In this picture, the FRB emission is narrowly beamed into the region spanned by the orange arrows, whereas the X-rays are visible from a large fraction of the sky. The double radio pulses seen in FRB 200428 are produced by two separate eruptions, which also enhance Comptonization and give rise to the two hard X-ray peaks. *Right-hand panel*: Crustal deformations due to sudden magnetic energy release excite shear mode oscillations. The shear wave propagates along the crust, and when it reaches the NS surface, a fraction of energy is transmitted into the magnetosphere as Alfvén waves and the rest is reflected back into the crust. The FRB duration is given by shear wave propagation delay between different paths, $t_{\rm frb} \sim 1$ ms for typical wave speed $v_{\rm s} \sim 0.01c$.

where q is electron charge, $\delta B_{10} = \delta B/10^{10}$ G and $\lambda_{\perp 4} = \lambda_{\perp}/10^4$ cm are measured at the NS surface.

When the wave arrives at the charge starvation radius R, a strong electric field develops along the background magnetic field and accelerates clumps of particles that were formed due to two-stream instability associated with the Alfvén wave current density. These particle clumps move along curved field lines and produce coherent curvature radiation. The clumps that form due to two-stream instability have a broad spectrum of longitudinal sizes $\ell_{\shortparallel} \lesssim c/\nu_{\rm p}$ (c being the speed of light), and radio emission is generated by those ones with $\ell_{\parallel} \simeq \lambda_{\rm frb}/2 = 15\nu_9^{-1}$ cm. The number of particles that can radiate coherently is $N_{\rm coh} \simeq \pi n_{\rm c} \ell_{\parallel} \ell_{\perp}^2$, where the transverse size is given by $\ell_{\perp} \simeq \sqrt{R\lambda_{\rm frb}}$ such that the photon arrival time does not differ by more than half an FRB wave period. This choice of $\ell_{\perp}$ is because the other two relevant length-scales – the Alfvén transverse wavelength $\lambda_{\perp}$ and the causal length R/γ – are typically much longer than $\sqrt{R\lambda_{\rm frb}}$. The clump Lorentz factor γ is related to the characteristic frequency of curvature emission $\nu = 3\gamma^3 c/(4\pi R_{\rm B})$ and the curvature radius of magnetic field lines $R_{\rm B}$,

$$\gamma \simeq 240\,(\nu_9 R_{\rm B,8})^{1/3}. \tag{2}$$

The total luminosity is $N_{\rm coh}^2$ times the curvature luminosity $L_{\rm curv} \simeq 16\gamma^8 q^2 c/3R_{\rm B}^2$ from an individual particle, provided that the observer is located within the relativistic beaming cone (of angular size $\sim\gamma^{-1}$), so we obtain

$$L_{\rm frb} \simeq 7\times 10^{39}\,{\rm erg\,s^{-1}}\,\frac{(\delta B_{10}/\lambda_{\perp 4})^2}{R_7^{11/3}\theta_{-1.5}^{2/3}\nu_9^{4/3}}, \tag{3}$$

where we have denoted the magnetic colatitude of the field line on the NS surface as $\theta = 10^{-1.5}\theta_{-1.5}$ rad and the corresponding curvature radius for a dipolar geometry is $R_{\rm B} \simeq 0.8(R_{\rm ns}/\theta)(R/R_{\rm ns})^{1/2}$.

The luminosity is mainly set by the charge starvation radius R, and the initial amplitude δB as well as the transverse wavelength $\lambda_{\perp}$ of the Alfvén waves. Our poor understanding of the charge

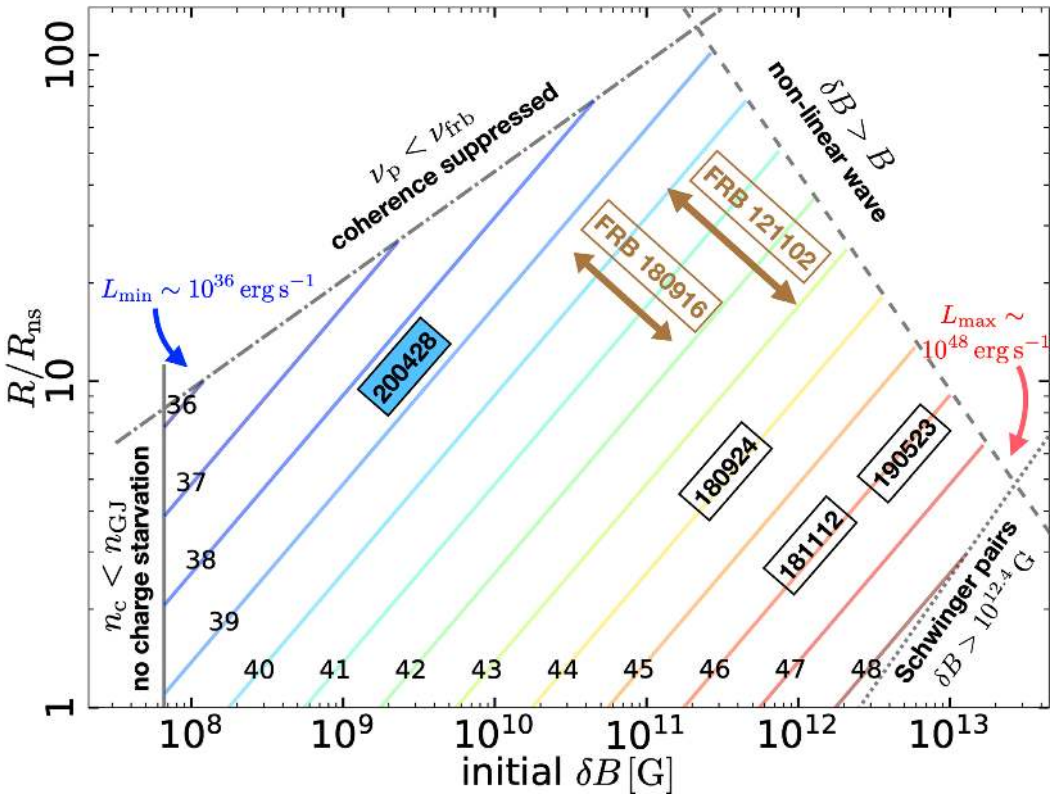

Figure 3. Physically allowed initial Alfvén wave amplitude δB, charge starvation radius R (in units of NS radius $R_{\rm ns}$), and FRB luminosity for the model described in this paper. The boundaries of the parameter space are given by the following constraints (shown by the grey lines): (1) the critical density $n_{\rm c}(R)$ must be greater than the Goldreich–Julian density $n_{\rm GJ}$ for charge starvation to be possible (solid), (2) the plasma frequency $\nu_{\rm p}$ must exceed the FRB frequency $\nu_{\rm frb}$ so as to allow charge clumps of size $\ell_{\shortparallel} \sim \lambda_{\rm frb}$ (dash–dotted), (3) the wave amplitude at the critical radius $\delta B(R)$ is less than the background magnetic field B so the wave remains linear (dashed), and (4) the wave amplitude at the critical radius must not exceed ∼5 per cent of the quantum critical field strength to avoid rapid production of Schwinger pairs (dotted). The solutions for different FRB luminosities from 10^{36} to 10^{48} erg s^{-1} lie on the coloured lines (with $\log L_{\rm frb}$ [erg s^{-1}] marked on each line). Since the charge density profile in the NS magnetosphere is poorly understood, our current model cannot provide a unique solution. We mark the localized sources with known (ranges of) luminosities in boxes, with the repeaters (FRB 121102 and 180916) in brown. The parameters used for this plot are transverse wavelength $\lambda_\perp = 10^4$ cm, surface magnetic field $B_{\rm ns} = 3 \times 10^{14}$ G, spin period $P = 3$ s, magnetic colatitude of the field line $\theta = 10^{-1.5}$ rad, and observing frequency $\nu_{\rm frb} = 1$ GHz.

density profile of the magnetosphere does not allow us to directly determine R. Generally, Alfvén waves launched near the magnetic poles where field lines extend to large distances are much more likely to become charge starved and produce coherent radiation. Here, we can use observed luminosity of FRB 200428, $L_{\rm frb} \sim 3 \times 10^{38}$ erg s^{-1} (assuming a distance of ∼10 kpc and frequency bandwidth of $\Delta\nu \sim 1.4$ GHz), to constrain $R/R_{\rm ns} \sim 20\,(\delta B_{10}/\lambda_{\perp 4})^{6/11}$.

The spectrum of the emergent radio waves depends on the size distribution of particle clumps and their Lorentz factors. The emergent power at frequency ν depends on the Fourier transform of particle number density $\tilde{n}(k)$ at wavenumber $k \sim 2\pi\nu/c$, and the distribution of particle Lorentz factor (γ) on this scale. We note that the transverse size of the coherent patch $\ell_\perp$ is typically much smaller than the causal length R/γ, which means that Doppler effect only slightly broadens the spectrum by $\Delta\nu/\nu \simeq (\gamma\ell_\perp/R)^2 \sim 0.1$, and the spectrum can have large intrinsic variations over a narrow band as radiation arrives from different clumps at different observer time.

The FRB emission is produced at a radius of $R \sim 20R_{\rm ns}$, as described above, with an uncertainty by a factor of a few. Thus, Alfvén waves should be launched within the polar angle $\theta \lesssim 0.1$ rad in order to ensure that these waves are able to propagate out to $\sim 10^2 R_{\rm ns}$ and pass through charge starvation point. Furthermore, θ cannot be much smaller than 0.02 rad because otherwise the beaming cone of field lines at $\sim 20R_{\rm ns}$ would rotate outside observer line of sight in 30 ms[3] and the second radio pulse seen from FRB 200428 would have been missed. These constraints on the magnetic colatitude motivate the choice of $\theta = 0.03$ rad as our fiducial value in equation (3). All things being similar for different X-ray bursts from SGR 1935+2154, we expect to see one FRB for $\sim 10^2$ X-ray bursts (this seems consistent with the available data for this object, Lin et al. 2020b). If the Alfvén wave packet has an azimuthal angular span of $\delta\phi \sim 1$ rad, then the solid angle of FRB emission at the charge starvation radius is $\Omega_{\rm frb} \sim \delta\phi\,\theta^2 (R/R_{\rm ns}) \sim 10^{-2}$ sr. The beaming fraction of $\Omega_{\rm frb}/4\pi \sim 10^{-3}$ is consistent with that inferred from the volumetric rate of X-ray bursts and FRBs in Section 4.

In Fig. 3, we show the solutions for different FRBs with a wide range of luminosities, along with a number of physical constraints

[3]It might be tempting to consider the possibility that the two radio pulses separated by 30 ms were in fact due to one continuous event that produced a hollow cone of radio emission, and the two pulses corresponded to the sweep of the cone across the line of sight as the NS rotated. However, two hard X-ray pulses also separated by ∼30 ms cast doubt on this possibility, since it requires that the hard X-ray emission is also beamed into the same hollow cone as the radio emission whereas the softer X-rays were presumably not beamed.

on the charge starvation radius and the initial amplitude of Alfvén disturbance. For simplicity, we fix $B_{\rm ns} = 3 \times 10^{14}$ G, $P = 3$ s, $\theta = 10^{-1.5}$ rad, and $\nu_{\rm frb} = 1$ GHz. The biggest uncertainty lies on $\lambda_\perp$, the transverse wavelength of the Alfvén waves on the NS surface, which depends on how the initial disturbance is launched. For $\lambda_\perp = 10^4$ cm, our model predicts FRB luminosities in the range 10^{36}–10^{48} erg s^{-1} and hence provides a viable explanation for faint bursts like FRB 200428 as well as bright events like FRB 190523 (Ravi et al. 2019). The maximum luminosity is due to the wave electric field, parallel to the magnetic field, at the charge starvation radius exceeding the Schwinger limit (Lu & Kumar 2019). We also predict that the FRB luminosity function must have a (so far unobserved) flattening at the lower end, although the exact minimum luminosity $L_{\rm min}$ depends on the unknown $\lambda_\perp$. This is because, for very small initial Alfvén amplitude, charge starvation occurs far away from the NS surface where the plasma frequency is below the GHz band, and in this case all charge clumps have longitudinal sizes $\ell_{\shortparallel} > \lambda_{\rm frb}$ and hence the coherent emission at GHz frequencies is strongly suppressed. When the line of sight is outside the beaming cone of angular size $\sim\gamma^{-1}$, the observed luminosity is heavily suppressed by relativistic effects and hence may be below $L_{\rm min}$, but the chance of detection is very small.

What fraction of energy in this event reached near the magnetic poles and contributed to FRB emission? Suppose initially the outburst started far away from the magnetic pole, since most free energy in the tangled magnetic fields is near the equator (Thompson, Lyutikov & Kulkarni 2002; Gourgouliatos, Wood & Hollerbach 2016). Crustal deformations during the flare excite seismic oscillations, preferentially toroidal shear modes that preserve the shape of the star (Duncan 1998; Piro 2005), and the disturbance propagates along the crust to other parts of the star. Due to the small thickness of the crust $h \sim 0.5$–1 km, the wave undergoes many reflections off the surface before reaching the polar region. The distance travelled by the wave between two consecutive reflections is $\ell \sim \sqrt{hR_{\rm ns}}$, and the minimum number of reflections between the trigger to the magnetic pole is $\pi R_{ns}/2\ell \sim 4$. The FRB duration is given by propagation delay between different paths $t_{\rm frb} \sim \ell/v_{\rm s} \sim 1$ ms for wave speed $v_{\rm s} \sim 0.01c$. Each time the waves reach the surface, high-frequency ($\gg 10^4$ Hz) Fourier components are largely transmitted into magnetospheric Alfvén modes (Blaes et al. 1989). The Alfvén waves launched at $\theta \gtrsim 0.1$ rad are trapped in closed field lines, cascade to smaller scales, and create an $e^{\pm}$-photon plasma that radiates most of the energy as X-rays. For low-frequency seismic components $\lesssim 10^4$ Hz, since the corresponding Alfvén wavelength is $\gtrsim 3R_{\rm ns}$, their transmission to the magnetosphere preferentially occurs near the poles where the magnetic field lines are sufficiently extended (Thompson & Duncan 1995). The energy per unit surface area transmitted into the magnetospheric Alfvén waves in the polar region can be estimated to be $F_{\rm A}/F \sim T(1-T)^{N_{\rm r}} h/R_{\rm ns}$, where the fluence normalization $F = E/4\pi R_{\rm ns}^2$ is from uniformly distributing the total energy over the NS surface, T is the transmission coefficient from crustal shear waves to Alfvén waves, and $N_{\rm r} \sim 5$ is the typical number of reflections. The frequency spectrum of seismic oscillations and their propagation properties are still highly uncertain. For $0.03 \lesssim T \lesssim 0.5$ (Blaes et al. 1989; Bransgrove, Beloborodov & Levin 2020), we roughly estimate $F_{\rm A}/F$ to be of the order of 10^{-3}. Following the field lines from the NS surface to the charge starvation radius, the energy per solid angle drops by another factor of $R_{\rm ns}/R \sim 0.1$. Assuming a fraction of order unity of the Alfvén luminosity is converted into coherent radio emission, we expect the FRB to X-ray luminosity ratio to be of the order of 10^{-4}, which is in rough agreement with observed fluence ratio between these two bands.

6 SUMMARY

The first Galactic FRB from a magnetar, with its associated X-ray counterpart, provides an extraordinary opportunity to understand the FRB phenomenon as a whole. We explore the implications of FRB 200428 in various aspects. We find that FRB 200428-like events likely contribute a significant fraction of the cosmological FRB rate function at the faint end near specific energy $E_\nu \sim 10^{26}$ erg Hz^{-1}. We compare SGR 1935+2154 with the sources of other active repeaters (e.g. FRB 121102) and discuss how they may be understood in a general framework of the magnetar progenitors from different formation channels. Then, we compare the rates of SGR X-ray bursts and FRBs and find that only a small fraction (of the order of 10^{-3}–10^{-2}) of X-ray bursts may be accompanied by FRBs.

We consider two broad classes of FRB emission mechanisms. First, the 'far-away' models describe that a relativistic outflow drives a shock into the surrounding medium at large distances and generates radio emission by a plasma maser process. We carried out a detailed analysis of these models and found a number of difficulties explaining the radio and X-ray data from FRB 200428. The second class are the 'close-in' models, where radio emission is generated by a coherent process within the NS magnetosphere. We propose a scenario that magnetic disturbance near the stellar surface propagates to larger radii in the form of Alfvén waves, which then damp and produce radio emission. FRB 200428 was associated with an X-ray burst, and the hard X-ray light curve had two prominent spikes that occurred at nearly the same time as the two FRB pulses. The coincidence of hard X-ray and radio peaks and their relative fluxes can be understood in this scenario. This model provides a unified picture for faint bursts like FRB 200428 as well as the bright events seen at cosmological distances.

ACKNOWLEDGEMENTS

WL thank Shri Kulkarni for his encouragement throughout this project. This work has been funded in part by an NSF grant AST-2009619. WL was supported by the David and Ellen Lee Fellowship at Caltech.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Bannister K. W. et al., 2019, Science, 365, 565
Beloborodov A. M., 2017, ApJ, 843, L26
Beloborodov A. M., 2020, ApJ, 896, 142
Beniamini P., Hotokezaka K., van der Horst A., Kouveliotou C., 2019, MNRAS, 487, 1426
Bhandari S. et al., 2018, MNRAS, 475, 1427
Blaes O., Blandford R., Goldreich P., Madau P., 1989, ApJ, 343, 839
Bochenek C. D., Ravi V., Belov K. V., Hallinan G., Kocz J., Kulkarni S. R., McKenna D. L., 2020a, preprint (arXiv:2005.10828)
Bochenek C. D., McKenna D. L., Belov K. V., Kocz J., Kulkarni S. R., Lamb J., Ravi V., Woody D., 2020b, PASP, 132, 034202
Bransgrove A., Beloborodov A. M., Levin Y., 2020, ApJ, 897, 173
Caleb M. et al., 2017, MNRAS, 468, 3746
Chatterjee S. et al., 2017, Nature, 541, 58
Chen G., Ravi V., Lu W., 2020, ApJ, 897, 146
CHIME/FRB Collaboration, 2019, ApJ, 885, L24
Cho H. et al., 2020, ApJ, 891, L38
Cordes J. M., Chatterjee S., 2019, ARA&A, 57, 417

Cordes J. M., Wasserman I., 2016, MNRAS, 457, 232
Duncan R. C., 1998, ApJ, 498, L45
Farah W. et al., 2018, MNRAS, 478, 1209
Gaensler B. M., 2014, GCN Circ., 16533, 1
Gehrels N., 1986, ApJ, 303, 336
Gourgouliatos K. N., Wood T. S., Hollerbach R., 2016, Proc. Natl. Acad. Sci., 113, 3944
Hardy L. K. et al., 2017, MNRAS, 472, 2800
Israel G. L. et al., 2016, MNRAS, 457, 3448
Katz J. I., 2016, ApJ, 826, 226
Katz J. I., 2018, Prog. Part. Nucl. Phys., 103, 1
Katz J. I., 2019, MNRAS, 487, 491
Kennel C. F., Coroniti F. V., 1984, ApJ, 283, 694
Kothes R., Sun X., Gaensler B., Reich W., 2018, ApJ, 852, 54
Kulkarni S. R., Ofek E. O., Neill J. D., Zheng Z., Juric M., 2014, ApJ, 797, 70
Kumar P., Bošnjak Ž., 2020, MNRAS, 494, 2385
Kumar P., Lu W., 2020, MNRAS, 494, 1217
Kumar P., Panaitescu A., 2000, ApJ, 541, L51
Kumar P., Zhang B., 2015, Phys. Rep., 561, 1
Kumar P., Lu W., Bhattacharya M., 2017, MNRAS, 468, 2726
Li Y., Zhang B., 2020, ApJ, 899, L6
Li C. K. et al., 2020, preprint (arXiv:2005.11071)
Li D., Nan R., Pan Z., 2013, in van Leeuwen J., ed., Proc. IAU Symp. 291, Neutron Stars and Pulsars: Challenges and Opportunities after 80 years. Kluwer, Dordrecht, p. 325
Lin L., Göğüş E., , Roberts O. J., Kouveliotou C., Kaneko Y., van der Horst A. J., Younes G., 2020a, ApJ, 893, 156
Lin L. et al., 2020b, preprint (arXiv:2005.11479)
Lorimer D. R., Bailes M., McLaughlin M. A., Narkevic D. J., Crawford F., 2007, Science, 318, 777
Lu W., Kumar P., 2018, MNRAS, 477, 2470
Lu W., Kumar P., 2019, MNRAS, 483, L93
Lu W., Piro A. L., 2019, ApJ, 883, 40
Luo R., Men Y., Lee K., Wang W., Lorimer D. R., Zhang B., 2020, MNRAS, 494, 665
Lyubarsky Y., 2008, ApJ, 682, 1443
Lyubarsky Y., 2014, MNRAS, 442, L9
Lyubarsky Y., 2020, ApJ, 897, 1
Lyutikov M., Burzawa L., Popov S. B., 2016, MNRAS, 462, 941
Marcote B. et al., 2020, Nature, 577, 190
Margalit B., Berger E., Metzger B. D., 2019, ApJ, 886, 110
Margalit B., Metzger B. D., Sironi L., 2020a, MNRAS, 494, 4627
Margalit B., Beniamini P., Sridhar N., Metzger B. D., 2020b, ApJ, 899, L27
Mereghet ti S. et al., 2020, ApJ, 898, L29
Metzger B. D., Giannios D., Thompson T. A., Bucciantini N., Quataert E., 2011, MNRAS, 413, 2031
Metzger B. D., Berger E., Margalit B., 2017, ApJ, 841, 14
Metzger B. D., Margalit B., Sironi L., 2019, MNRAS, 485, 4091
Nicholl M., Williams P. K. G., Berger E., Villar V. A., Alexander K. D., Eftekhari T., Metzger B. D., 2017, ApJ, 843, 84
Ofek E. O., 2007, ApJ, 659, 339
Pen U. L., Connor L., 2015, ApJ, 807, 179
Petroff E., Hessels J. W. T., Lorimer D. R., 2019, A&AR, 27, 4
Piro A. L., 2005, ApJ, 634, L153
Plotnikov I., Sironi L., 2019, MNRAS, 485, 3816
Popov S. B., Postnov K. A., 2010, in Harutyunian H. A., Mickaelian A. M., Terzian Y., eds, Evolution of Cosmic Objects through their Physical Activity. Gitutyun Publishing House, Armenia, p. 129
Prochaska J. X. et al., 2019, Science, 366, 231
Ravi V. et al., 2019, Nature, 572, 352
Ridnaia A. et al., 2020, preprint (arXiv:2005.11178)
Schmidt M., 1968, ApJ, 151, 393
Shannon R. M. et al., 2018, Nature, 562, 386
Sieber W., Seiradakis J. H., 1984, A&A, 130, 257
Sironi L., Keshet U., Lemoine M., 2015, Space Sci. Rev., 191, 519
Spitler L. G. et al., 2016, Nature, 531, 202
Stamatikos M., Malesani D., Page K. L., Sakamoto T., 2014, GCN Circ., 16520, 1
Surnis M. P., Joshi B. C., Maan Y., Krishnakumar M. A., Manoharan P. K., Naidu A., 2016, ApJ, 826, 184
Tavani M., Ursi A., Verrecchia F., Casentini C., Pittori C., 2020, Astron. Telegram, 13686, 1
Tendulkar S. P. et al., 2017, ApJ, 834, L7
The CHIME/FRB Collaboration, 2020a, Nature, 582, 351
The CHIME/FRB Collaboration, 2020b, preprint (arXiv:2005.10324)
Thompson C., Duncan R. C., 1993, ApJ, 408, 194
Thompson C., Duncan R. C., 1995, MNRAS, 275, 255
Thompson C., Lyutikov M., Kulkarni S. R., 2002, ApJ, 574, 332
Thornton D. et al., 2013, Science, 341, 53
Usov V. V., 1992, Nature, 357, 472
Vink J., Kuiper L., 2006, MNRAS, 370, L14
Wadiasingh Z., Timokhin A., 2019, ApJ, 879, 4
Wang F. Y., Wang Y. Y., Yang Y. P., Yu Y. W., Zuo Z. Y., Dai Z. G., 2020a, ApJ, 891, 72
Wang W. Y., Xu R., Chen X., 2020b, ApJ, 899, 109
Waxman E., 2017, ApJ, 842, 34
Yang Y. P., Zhang B., 2018, ApJ, 868, 31
Zhang B., 2017, ApJ, 836, L32
Zhang B., Mészáros P., 2001, ApJ, 552, L35

APPENDIX A: 'FAR-AWAY' MODELS – EMISSION FROM BEYOND THE LIGHT CYLINDER

In this Appendix, we study the other class of 'far-away' models, where a relativistic outflow drives a shock into the circumstellar medium (CSM) at large distances, and FRB is generated by a plasma maser process, as proposed by Lyubarsky (2014), Beloborodov (2017, 2019), Metzger et al. (2019), and further developed by Plotnikov & Sironi (2019) and Margalit et al. (2020a).

The properties of the CSM may be highly diverse as it is shaped by the pulsar wind, flares from the NS, and the supernova remnant. These complications can be avoided by considering one snapshot in the FRB light curve. The observed flux at a given time can be shown to be produced when the shock front is at some effective radius r. We take the average density of the material swept up by the shock front up to radius r to be ρ_0, bulk Lorentz factor of the unshocked, upstream medium to be Γ_0, and magnetization parameter $\sigma = 1 + B_0^2/4\pi\rho_0 c^2$; where B_0, the magnetic field strength of the upstream fluid, and ρ_0 are measured in the comoving frame of the upstream medium. The CSM is initially cold. The ejecta drives a shock into the CSM and the shock Lorentz factor in the lab or NS rest frame is Γ_s. The energy of the shocked CSM at radius r is

$$E \simeq 4\pi r^3 u_0 \Gamma_{\rm rel}^2 = 4\pi r^3 \rho_0 c^2 (\Gamma_s/\Gamma_0)^2, \tag{A1}$$

where we have used $u_0 = \sigma\rho_0 c^2$ as the average energy density of unshocked CSM up to radius r, and the relative Lorentz factor between the shocked and pre-shock plasma $\Gamma_{\rm rel} = \Gamma_s/(\Gamma_0\sigma^{1/2})$ as given by the Rankine–Hugoniot jump conditions (e.g. Kennel & Coroniti 1984). The emission frequency of the maser emission $\omega = 2\pi\nu$ is roughly given by (Plotnikov & Sironi 2019)

$$\omega \simeq 3\Gamma_s\omega_{\rm p}, \tag{A2}$$

where $\omega_{\rm p} = \sqrt{4\pi n_0 q^2/m_{\rm e}}$ is the plasma frequency, $n_0 = \rho_0/m$ is the electron number density of the upstream plasma, and m is the mean

mass per electron. The emission duration $t_{\rm frb}$ is given by[4]

$$r \simeq 2\Gamma_{\rm s}^2 t_{\rm frb} c. \tag{A3}$$

It is straightforward to solve the above three equations for the emission radius r, shock Lorentz factor $\Gamma_{\rm s}$, and pre-shock number density n_0. And we find

$$\begin{aligned} r &\simeq (8.9\times10^{11}\,{\rm cm})\left(\frac{m}{m_{\rm e}}\right)^{-\frac{1}{3}}\Gamma_0^{\frac{2}{3}}E_{40}^{\frac{1}{3}}\nu_9^{-\frac{2}{3}},\\ \Gamma_{\rm s} &\simeq 1.2\times10^{2}\left(\frac{m}{m_{\rm e}}\right)^{-\frac{1}{6}}\Gamma_0^{\frac{1}{3}}E_{40}^{\frac{1}{6}}\nu_9^{-\frac{1}{3}}t_{\rm frb,ms}^{-\frac{1}{2}},\\ n_0 &\simeq (9.0\times10^{4}\,{\rm cm^{-3}})\left(\frac{m}{m_{\rm e}}\right)^{\frac{1}{3}}\Gamma_0^{-\frac{2}{3}}E_{40}^{-\frac{1}{3}}\nu_9^{\frac{8}{3}}t_{\rm frb,ms}. \end{aligned} \tag{A4}$$

The optical depth of the upstream plasma for induced Compton scattering is given by (e.g. Lyubarsky 2008; Kumar & Lu 2020)

$$\tau_{\rm IC} \simeq \frac{3\sigma_{\rm T}E_{\rm frb}\Gamma_0^2 n_0 c}{32\pi^2 r^2 m_{\rm e}\nu^3} \simeq 23\frac{m}{m_{\rm e}}f_{\rm r,-4}\,\nu_9\,t_{\rm frb,ms}. \tag{A5}$$

To allow GHz coherent radio waves to escape, we find that the upstream composition must be dominated by electron–positron pairs[5] with $m \simeq m_{\rm e}$. In fact, the baryonic shock model is ruled out by the data since it overproduces X-ray luminosity by a factor of $\gtrsim 10^3$. This is because the baryonic shock must have much larger energy to get around the induced Compton constraints. Hereafter, we take $m = m_{\rm e}$ and then the luminosity of upstream material is

$$L_0 \simeq (2.2\times10^{34}\,{\rm erg\,s^{-1}})\,\sigma\Gamma_0^{\frac{8}{3}}E_{40}^{\frac{1}{3}}\nu_9^{\frac{4}{3}}t_{\rm frb,ms}. \tag{A6}$$

In Fig. A1, we show how the FRB frequency is related to the upstream luminosity L_0 and Lorentz factor Γ_0 according to the above relation, while fixing $E_{40} = 1$ and $t_{\rm frb,\,ms} = 1$ as motivated by FRB 200428. We see that, to generate GHz radio emission, the upstream plasma conditions must lie along a narrow valley in the otherwise very wide parameter space.

The next step is to consider that there are two radio pulses separated by about 30 ms as detected by CHIME. The first one spans from 400 MHz (lower end of the observing band) up to 550 MHz, and the second one spans from 550 to 800 MHz (upper end of the observing band). One should be cautious about the details of the spectrum because FRB 200428 is detected in the far side lobe where the spectral response may not be well understood. However, since CHIME's response is not expected to change significantly on a time-scale of 30 ms, the *difference* between the spectra of the two pulses should be physical. Each pulse has duration of about 1 ms, after correcting for scattering broadening. We also note that the associated X-ray burst also had two distinct peaks separated by 30 ms in the hardest band (27–250 keV) of HXMT (Li et al. 2020), which were temporally coincident with the two radio peaks. This suggests that the two radio pulses are generated by two separated ejectas. The first ejecta interacts with the (perhaps temporarily enhanced) magnetar wind.

[4]From pressure balance between the shocked regions, one obtains the relative Lorentz factor $\Gamma_{\rm rel} \simeq (L/L_0)^{1/4}$, where $L = E/t_{\rm ej}$ is the luminosity of the ejecta, $t_{\rm ej}$ is the launching duration, and $L_0 = 4\pi r^2 u_0 \Gamma_0^2 c$ is the luminosity of the outflowing CSM. This combined with equation (A3) then gives $t_{\rm frb} \simeq t_{\rm ej}/(2\sigma)$ (Kumar & Zhang 2015), which means the FRB duration is much shorter than the ejection duration if $\sigma \gg 1$.

[5]For a baryonic (electron–proton) composition, the radiative efficiency must be extremely low $f_{\rm r} \lesssim 10^{-7}$ in order to have $\tau_{\rm IC} \lesssim 10$. And that requires an ejecta energy of $E \gtrsim 10^{43}$ erg, which is 3 orders of magnitude higher than seen in the associated hard X-ray burst.

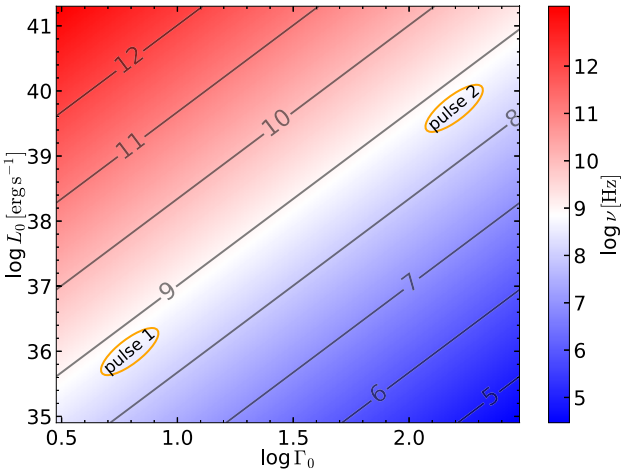

Figure A1. This graph (with black contour lines) shows the FRB frequency ν as a function of the upstream plasma luminosity L_0 and Lorentz factor Γ_0 for the shock-maser model described in this Appendix. We see that, to generate GHz radio emission, the upstream plasma parameters must lie in a very narrow range of the allowed space (along $L_0 \propto \Gamma_0^{8/3}$ line). In reality, the physical properties of the CSM (that the relativistic, magnetar flare, ejecta runs into) is expected to be highly diverse as it is shaped by many different processes, and the probability that the maser emission from the shock falls in the observing band is extremely small (the figure shows that the frequency of the emergent maser emission could be anywhere between 10^6 Hz and 10^{12} Hz for the parameters of FRB 200428 according to the shock model). We also show the locations of the two radio pulses as observed by CHIME by orange ellipses. For this plot, we fix the upstream magnetization $\sigma = 1$, ejecta energy $E = 10^{40}$ erg, and FRB duration $t_{\rm frb} = 1$ ms, as motivated by FRB 200428.

The second ejecta interacts with the slower tail of the first ejecta or the magnetar wind in between the two flare ejectas responsible for the two radio pulses.

The second ejecta will catch up with the tail of the previous ejecta or the wind following the previous ejecta, which we take to be moving with Lorentz factor $\Gamma_{\rm t}$, at $\delta t \simeq 30$ ms in the observer's frame, at the radius

$$r \simeq 2\Gamma_{\rm t}^2 \delta t\, c \simeq 1.8\times10^{9}\,{\rm cm}\,\Gamma_{\rm t}^2. \tag{A7}$$

We combine equation (A7) with the expressions in equation (A4) to obtain

$$\begin{aligned} r &\simeq (2.0\times10^{13}\,{\rm cm})\,E_{40}^{\frac{1}{2}}\nu_9^{-1},\\ \Gamma_{\rm s} &\simeq 5.7\times10^{2}\,E_{40}^{\frac{1}{4}}\nu_9^{-\frac{1}{2}}t_{\rm frb,ms}^{-\frac{1}{2}},\\ n_{\rm t} &\simeq (4.0\times10^{3}\,{\rm cm^{-3}})\,E_{40}^{-\frac{1}{2}}\nu_9^{3}t_{\rm frb,ms},\\ \Gamma_{\rm t} &\simeq 105\,E_{40}^{1/4}\nu_9^{-\frac{1}{2}}, \end{aligned} \tag{A8}$$

where $n_{\rm t}$ is number density of the upstream plasma in its comoving frame. We see that the dynamics of the second ejecta is well determined,[6] thanks to the resolved X-ray light curve by HXMT. The upstream Lorentz factor $\Gamma_{\rm t}$ is reasonable if the first ejecta has

[6]In fact, the dynamics of the first ejecta can also be determined if we assume the density profile of the upstream plasma ahead of the first shock to be $n_0 \propto r^{-2}$ (or other power–law forms). This is because, at the observer's time $t \simeq 1$ ms (during the first radio pulse of FRB 200428), the first ejecta is at its deceleration radius, which is a factor of $30^{1/2}$ less than the first expression in equation (A8). Then, one can plug the deceleration radius back into equation (A4) to solve for the unknown Γ_0 and hence other quantities as well.

most of the energy near the front end with high Lorentz factor $\gg 100$ (which is responsible for the first radio pulse) and a small fraction of energy in the tail with relatively low Lorentz factor ~ 100 (which is responsible for decelerating the second ejecta and hence generate the second radio pulse).

Can these shocks produce the non-thermal hard X-rays observed by HXMT and other instruments? The answer turns out to be no. The reason is that the characteristic synchrotron frequencies ($\nu_{\rm m}$) for an electron–positron CSM, for the parameters of the two shocks we determined above, are of the order of 10^{13} Hz and 10^{15} Hz respectively, much smaller than X-ray frequencies. Simulations suggest that shocks in a magnetized pair plasma might not produce an extended power-law particle spectrum above the average energy per particle (Sironi, Keshet & Lemoine 2015), i.e. little emission above $\nu_{\rm m}$. Even ignoring this difficulty, let us assume that the Fermi acceleration operates in the $e^{\pm}$ magnetized shock and produces power-law particle distribution with index $p \simeq 2$. The emergent synchrotron spectrum then is $F_\nu \propto \nu^{-0.5}$, which is consistent with the observed soft X-ray power law. The spectrum should extend with the same slope up to $\sim$100 MeV for the shock parameters of equations (A4) and (A8). However, Konus-Wind detected no significant emission above 250 keV from this event (Ridnaia et al. 2020), which suggests that the hard X-rays did not originate in these shocks.

In the following, we point out several problems with the shock model.

Since the two radio components have similar frequencies and durations to within factors of order unity and the upstream magnetization is modest $\sigma \lesssim 2$ (otherwise the FRB duration will be much less than a few ms), we infer that the upstream plasma for both shocks must have similar ratio of $L_0/\Gamma_0^{8/3}$ as given by equation (A6). This poses a problem for this model because the physical conditions of the upstream plasma before the two shocks are largely unrelated. The ratio $L_0/\Gamma_0^{8/3}$ could change by many orders of magnitude from one pulse to another in an FRB event – especially, considering that the shock is being driven into the tail end of the previous outburst, or outflow preceding the current flare, which contains a tiny fraction of the total energy of the outburst – and the resulting synchrotron maser emission will generally be at widely separated frequencies and produce pulses of very different durations in the observer frame. For instance, if L_0 were to be different for the two shocks by a factor of 2, then the maser frequency in the observer frame would be different by a factor of 1.7 (for the same pulse duration). A factor of 2 change in Γ_0 would lead to a factor of 4 change in the maser frequency. The same argument applies to other close burst pairs such as the ones seen in FRB 121102 (Hardy et al. 2017).

The typical variability time of the emission from a relativistic shock should be of the order of the signal arrival time in the observer's frame, $\Delta t \sim t$, because the observed flux at a given moment comes from a wide range of emitting radii of $\Delta r \sim r$ and angles with respect to the line of sight $\theta \sim 1/\Gamma_{\rm s}$. However, the de-dispersed light curves of the two pulses in FRB 200428 show extremely rapid rise with $\Delta t/t \equiv \xi \sim 0.1$. Some other FRBs also show very rapid variability time as short as tens of microseconds (Cho et al. 2020). An external shock can account for this sharp rise time, provided that the observed flux is produced in a very small emission area $A \sim \xi^2 (r/\Gamma_{\rm s})^2$, which is much smaller than the size of the causally connected region $r/\Gamma_{\rm s}$. However, in this case the blastwave energy should be larger by a factor of $\xi^{-2} \sim 10^2$ to account for the observed FRB flux. Then, the efficiency decreases to $f_{\rm r} \sim 10^{-7}$, and the energy required in the relativistic shock is $\sim 10^2$ larger than seen in the X-ray band. Furthermore, an even more serious problem is the requirement that the size of the emission patch in the two completely unrelated shocks should be nearly of the same area and similar location wrt the observer line of sight in order that the observed flux of the two pulses and their rise times are similar.

The observed spectrum of the first (or second) pulse cuts off abruptly above (or below) about 550 MHz. This also poses a problem. The spectrum for the maser-in-shock mechanism is expected to be broad with $\Delta\nu \sim \nu$ due to slightly different Doppler shift for different points on the shock surface within an angle $1/\Gamma_{\rm s}$ from the line of sight. Particularly worrisome is the cutoff of the spectrum of the second pulse below 550 MHz. Even if the maser mechanism is terminated suddenly when the shock is at some radius r, we will continue to receive radiation for at least a few times $r/(2c\Gamma_{\rm s}^2)$, which drifts down in frequency as $1/t$ and the flux declines roughly as $1/t^{-2}$ (Kumar & Panaitescu 2000); t is the observer frame time. Therefore, in the shock scenario, it is not possible to cutoff the observed emission below 550 MHz except by invoking some propagation effects, but then that makes it problematic to explain the first pulse which extends down to 400 MHz merely 30 ms earlier.

Another concern, at least for the second radio pulse, is that the predicted downward frequency drift by the shock model is not observed; the observed frequency should decrease with time as the shock decelerates. From the expression of $\Gamma_{\rm s}$ in equation (A8), the observer's time scales as $t \propto \Gamma_{\rm s}^{-2}\nu^{-1}$ since the blastwave energy is conserved. For two different frequencies $\nu^{(1)} > \nu^{(2)}$, the shock Lorentz factor must satisfy $\Gamma_{\rm s}^{(1)} > \Gamma_{\rm s}^{(2)}$, so we obtain $t^{(1)}/t^{(2)} < \nu^{(2)}/\nu^{(1)}$, meaning that the observed frequency evolution is steeper than $t \propto \nu^{-1}$. However, no significant drift is seen for the second pulse between 550 and 800 MHz, despite the fact that a factor of $\gtrsim$1.5 in arrival time difference should be measurable.

We end this Appendix by concluding that the 'far-away' shock-maser model does a good job of explaining the radio emission efficiency $f_{\rm r} \sim 10^{-5}$–10^{-4}. The radio waves can escape the upstream plasma without being significantly scattered by the induced Compton process, provided that the upstream composition is electron–positron. However, there are a number of serious problems with the model: (1) The two radio pulses are generated by two shocks driven by different ejectas separated by 30 ms, but the frequency and duration of the radio pulses require that the upstream plasma with which these ejectas collide must have almost identical values of $L_0/\Gamma_0^{8/3}$, even though they are expected to be physically unrelated and their values could have been easily different by a factor of $\gtrsim$10. (2) The rapid variability time $\Delta t \ll t$ can be explained by the model by invoking that the flux at a given time only comes from a small patch of size much smaller than the causally connected region, but that decreases the efficiency by another factor of $\sim 10^2$, i.e. the energy requirement for the relativistic ejecta exceeds X-ray emission by a factor of $\sim 10^2$ in this case. Furthermore, it will need to invoke an additional uncomfortable assumption that the size of the emitting patch is nearly the same for the two pulses produced by unrelated shocks. (3) The narrow frequency band, particularly in the second pulse (550–800 MHz), is problematic for the emission from a relativistic shock. This is because even if the shock and the maser emission is suddenly turned off at a certain radius, we would continue to see photons of frequency smaller than 550 MHz arriving to us from an angle wrt to our line of sight just slightly larger than $\Gamma_{\rm s}^{-1}$ with flux barely a factor of 2 smaller than that at 550 MHz; CHIME should have detected the emission down to 400 MHz. (4) The emission from a decelerating shock drifts downwards in frequency with time, but the expected drift is not observed by CHIME.

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