Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **490,** 4858–4876 (2019)
Advance Access publication 2019 October 03

doi:10.1093/mnras/stz2729

Instability of twisted magnetar magnetospheres

J. F. Mahlmann,[1★] T. Akgün,[2] J. A. Pons,[2] M.A. Aloy[1] and P. Cerdá-Durán[1]

[1]*Departament d'Astronomia i Astrofísica, Universitat de València, E-46100 Burjassot, Spain*
[2]*Departament de Fisica Aplicada, Universitat d'Alacant, E-03690 Alicante, Spain*

Accepted 2019 September 25. Received 2019 September 22; in original form 2019 July 31

ABSTRACT

We present 3D force-free electrodynamics simulations of magnetar magnetospheres that demonstrate the instability of certain degenerate, high energy equilibrium solutions of the Grad–Shafranov equation. This result indicates the existence of an unstable branch of twisted magnetospheric solutions and allows us to formulate an instability criterion. The rearrangement of magnetic field lines as a consequence of this instability triggers the dissipation of up to 30 per cent of the magnetospheric energy on a thin layer above the magnetar surface. During this process, we predict an increase of the mechanical stresses on to the stellar crust, which can potentially result in a global mechanical failure of a significant fraction of it. We find that the estimated energy release and the emission properties are compatible with the observed giant flare events. The newly identified instability is a candidate for recurrent energy dissipation, which could explain part of the phenomenology observed in magnetars.

Key words: magnetic fields – methods: numerical – stars: magnetars – stars: neutron – X-rays: bursts.

1 INTRODUCTION

Soft gamma-ray repeaters (SGRs) are neutron stars with recurrent X-ray activity in the form of short bursts with duration ∼0.1 s and luminosities in the range 10^{36}–10^{43} erg s^{-1}. Over the last 40 yr, three bursts have been uniquely energetic, the so-called *giant flares* (GFs) with luminosities of the order of 10^{44}–10^{47} erg s^{-1} (SGR 0525-66, SGR 1900 + 14, and SGR 1806-20; see Cline et al. 1980; Hurley et al. 1999, 2005). In the three referenced cases, a short initial peak was followed by a softer X-ray tail lasting for 50–400 s. The engine behind these extraordinary events are *magnetars*, neutron stars with the strongest known magnetic fields (10^{14}–10^{16} G; see comprehensive reviews of magnetar observations and physics, e.g. in Woods & Thompson 2006; Rea & Esposito 2011; Mereghetti, Pons & Melatos 2015; Turolla, Zane & Watts 2015; Kaspi & Beloborodov 2017).

The precise mechanism producing such energetic events is still unclear. Strong magnetic fields are a gigantic energy reservoir in magnetars, generally of the order

$$\mathcal{E}_{\rm magnetar} \sim 1.6 \times 10^{47}\,{\rm erg} \left(\frac{B}{10^{15}\,{\rm G}}\right)^2 \left(\frac{R_*}{10\,{\rm km}}\right)^3, \qquad (1)$$

where we consider a neutron star with radius R_*.

The time-scale on which the magnetar is evolving, mainly due to Hall drift and Ohmic dissipation in the crust, is of the order of 10^3–10^6 yr (Jones 1988; Goldreich & Reisenegger 1992; Pons & Geppert 2007; Pons, Miralles & Geppert 2009; Gourgouliatos, Wood & Hollerbach 2016), by itself too slow to explain this phenomenology. Two complementary models have tried to explain these observations. In the *crustquake* model (Thompson & Duncan 1996; Perna & Pons 2011) the dynamical trigger is the mechanical failure of patches of the magnetar crust due to large stresses built during its magnetothermal evolution. Numerical simulations of the Hall evolution of the crust (Viganò et al. 2013) show that it is possible to recurrently reach the maximum stress supported by the very same (Horowitz & Kadau 2009; Baiko & Chugunov 2018). At this point, the crust likely becomes plastic (Levin & Lyutikov 2012), i.e. the crust generates thermoplastic waves emerging from such a localized trigger, or in other words *yields* (Beloborodov & Levin 2014; Li, Levin & Beloborodov 2016). The waves propagate into the magnetosphere, probably resulting in rapid dissipation through a turbulent cascade triggered by reconnection on slightly displaced flux surfaces (Thompson & Duncan 1996, 2001; Li, Zrake & Beloborodov 2018). The energy released in those events suffices to explain the observed luminosities, even for GFs (Thompson & Duncan 1996; Lander et al. 2015). The burst duration (∼0.1 s) is related to the crossing time of shear waves through the whole crust (1–100 ms). A limitation is that, if stressed for long periods of time (∼ 1 yr) as it is the case due to the slow magnetothermal evolution, the crust may yield at significantly lower breaking stresses (Chugunov & Horowitz 2010). In that case, it would effectively deform as a plastic flow, and, depending on its (unknown) properties, cease to yield altogether (Lyutikov 2015; Lander & Gourgouliatos 2019). Thompson, Yang & Ortiz (2017) has argued that even in this case the crust could yield.

★ E-mail: jens.mahlmann@uv.es

© 2019 The Author(s)

Published by Oxford University Press on behalf of the Royal Astronomical Society

The *magnetospheric instability* model requires a strongly twisted magnetosphere that becomes unstable and leads to a rapid reconnection event (Lyutikov 2003). The existence of long-lived magnetospheric twists is supported by the observation of hard X-ray emission in persistent magnetars (Beloborodov 2013a; Hascoët, Beloborodov & den Hartog 2014). During the magnetothermal evolution of the crust, the displacement of the magnetic field footprints can generate large twists in the magnetosphere (Akgün et al. 2017, 2018b). Above a critical twist, the magnetosphere becomes unstable and undergoes a rapid rearrangement where energy is dissipated by reconnection (Lyutikov 2003; Gill & Heyl 2010; Elenbaas et al. 2016) in a similar fashion as in the crustquake model. The main challenge of this scenario is the ability of the crust to produce significant twists in the magnetosphere. Beloborodov (2009) estimated that currents supporting magnetospheric twist are bound to dissipate on time-scales of years, effectively leading to a progressive untwisting. Therefore, Hall evolution is required to proceed relatively fast in order to allow for significant twists. Plastic viscosity may also be a problem for similar reasons (Lander & Gourgouliatos 2019). The latter authors have also suggested that the dynamical crust fractures of the crustquake model could be substituted by sustained episodes of accelerated plastic flows which are able to generate large magnetospheric twists on times shorter than the untwisting time-scale.

Numerical simulations by Parfrey, Beloborodov & Hui (2012), Parfrey, Beloborodov & Hui (2013), and Carrasco et al. (2019) confirm the instability of the magnetosphere beyond a critical twist, accompanied by the formation of plasmoids. These results are an analogy to the context of eruption processes in the solar corona as found in numerical experiments by Roumeliotis, Sturrock & Antiochos (1994), Mikic & Linker (1994). The energy dissipated in the reconnection events is sufficient to explain the GF processes (Parfrey et al. 2012). A caveat to these simulations is that the applied twisting rate is larger than the one expected from the respective magnetothermal evolution, although it would be fine if the trigger was a rapid plastic deformation.

An alternative approach to the above is the study of stability properties of magnetospheres. A number of authors have constructed equilibrium solutions to the Grad–Shafranov equation (GSE) for neutron star magnetospheres (Fujisawa & Kisaka 2014; Glampedakis, Lander & Andersson 2014; Pili, Bucciantini & Del Zanna 2015; Akgün et al. 2016; Kojima 2017, 2018; Kojima & Okamoto 2018; Akgün et al. 2018a). Akgün et al. (2017) performed magnetothermal evolutions coupling the crustal magnetic field at the stellar surface with an exterior equilibrium solution. The results showed that large twists grow in the magnetosphere up to a critical point beyond which no stable equilibrium solutions where found. A more detailed analysis by Akgün et al. (2018a) showed that, for sufficiently large twists, the solutions of the GSE are degenerate with several possible configurations of different energies but matching boundary conditions at the surface. This suggests the possibility of an unstable branch of the solutions and, thus, a possible explanation for the occurrence of bursts and GFs. In this work we explore this possibility by performing 3D numerical simulations of the equilibrium models in Akgün et al. (2018a). We assess their stability properties and their potential as candidates for transient magnetar phenomenology.

This work is organized as follows. In Section 2 we review and discuss the physics involved in magnetars relevant to the processes that we want to study. In Section 3 we briefly review the equations of force-free electrodynamics (FFE) implemented for simulations conducted on the infrastructure of the `Einstein Toolkit` (supplemented by Appendix A1). A detailed description of the derivation of initial models according to Akgün et al. (2018a) is given in Section 4. In Section 5 we present the numerical setup of our simulations as well as the outcome of the conducted 3D force-free simulations of twisted magnetospheres (reviewing details on maintaining the force-free regime in Appendix A2). The observed rapid dissipation of electromagnetic energy through the magnetar crust is interpreted and related to observable quantities, such as luminosity estimates, shear stresses on the stellar crust, and opacity models, in Section 6. Along this paper we use Gaussian units in CGS, except for Section 3 in which we use Heaviside-Lorentz with geometrized units ($G = c = \mathrm{M}_{\odot} = 1$). For convenience we express current densities in A m^{-2} and voltages in V, instead of the corresponding CGS units.

2 PHYSICS OF MAGNETARS

The basic structure of the magnetar interior is a (likely) fluid core of $\sim$10 km radius, amounting for most of the mass of the object, surrounded by a solid crust of about 1 km size. Outside, there is a tenuous, corotating magnetosphere connected to the NS by magnetic field lines (threading the central object) that extend up to the light cylinder at distances larger than 10^5 km. We start by discussing some basic properties of the different parts of the magnetosphere relevant for the interpretations and models presented later in this work.

2.1 Currents supporting the magnetosphere

For the typical rotation periods of magnetars ($P \sim$ 1–10 s) the Goldreich–Julian particle density (Goldreich & Julian 1969) for a magnetar magnetosphere has the typical value

$$n_{\rm GJ} = 7 \times 10^{12}\,\mathrm{cm}^{-3} \left(\frac{B_{\rm pole}}{10^{15}\,\mathrm{G}}\right) \left(\frac{R_*}{r}\right)^3 \left(\frac{10\,\mathrm{s}}{P}\right), \tag{2}$$

where $B_{\rm pole}$ is the magnetic field strength at the magnetar pole, R_* the magnetar radius, and r the distance to the centre of the star. This limits the magnetospheric current density close to the surface to $J < e\,c\,n_{\rm GJ} \approx 3 \times 10^8\,\mathrm{A\,m}^{-2}$, much below the typical values needed to support currents in strongly twisted magnetospheres of magnetars, of the order of

$$J \sim \frac{Bc}{4\pi r} \sim 8.2 \times 10^{12}\,\mathrm{A\,m}^{-2} \left(\frac{B_{\rm pole}}{10^{15}\,\mathrm{G}}\right) \left(\frac{R_*}{10\,\mathrm{km}}\right)^{-1}. \tag{3}$$

In general, magnetospheric currents in magnetars cannot be supported neither by Goldreich–Julian charges nor by charges lifted from the surface. Beloborodov & Thompson (2007) proposed that the currents are supported by e^+–e^- pairs generated in the magnetosphere in an intermittent discharge process that can be sustained for voltages along magnetic field lines of about 10^8–10^9 V. This voltage can be maintained by self-induction in untwisting magnetospheres (Beloborodov 2009). This untwisting is driven by the effective resistivity of the magnetosphere; the thermal photons from the magnetar's surface scatter resonantly off the charges supporting the magnetospheric currents, taking energy away, at the same time that pairs are produced. The untwisting time-scale is $\sim$1 yr, and it may explain the spectral evolution of some magnetars (Beloborodov 2009).

2.2 Time-scales

Changes in magnetars take place during two different time-scales. On the one hand, there is a *secular time-scale* of thousands of years during which the magnetar is essentially in equilibrium. On the other hand, there is a *dynamical time-scale* associated to energetic events (burst, flares) that can produce observable variations on time-scales as fast as 0.1s. The latter are likely associated to out-of-equilibrium states.

2.2.1 Secular time-scales

The *secular time-scale* is set by the slow magnetothermal evolution of the cooling object. The interior magnetic field evolution is dominated by Hall drift and Ohmic diffusion at the crust (see e.g. Viganò, Pons & Miralles 2012; Fujisawa & Kisaka 2014, and references therein), which proceeds on typical time-scales of 10^3–10^6 yr. The long-term evolution of the magnetosphere is driven by the changes in the crustal magnetic field, which displaces the footprints of the magnetospheric magnetic field lines. Since this evolution is much slower than the dynamical time-scale of the magnetosphere (see below), it can be considered that the magnetosphere evolves through a series of equilibrium states. This evolution creates a twist in the magnetosphere supported by currents – until a critical maximum twist is reached ($\varphi_{\rm crit} \sim 1$ rad) beyond which no magnetospheric equilibrium solutions exist (Akgün et al. 2017). The stability of the magnetosphere close to this critical point is the subject of this paper.

At the same time as the crustal magnetic field evolves, other processes in the magnetosphere can also contribute to the evolution. The untwisting of the magnetosphere on time-scales of $\sim$1 yr (Beloborodov 2009, and discussion in Section 2.2.1), may be a competing action to the twisting process described above.

Although the velocity of the footprints is typically very slow, numerical simulations of the magnetothermal evolution of magnetars including the magnetosphere show that, close to the critical point, it can be as fast as $v_\varphi \sim 1\,{\rm km\,yr^{-1}}$ (see Akgün et al. 2017) in the most optimistic scenario. Therefore, close to the critical twist, the magnetosphere twists slowly ($\dot{\varphi}_{\rm max,crit} \lesssim 0.1\,{\rm rad\,yr^{-1}}$), evolving on time-scales $\gtrsim$ 10 yr. In the best case scenario, this time-scale is comparable to the untwisting time-scale ($\sim$1 yr) and, hence, parts of the magnetosphere could sustain a significant twist. This time-scale is still much longer than the dynamical time-scale of the system (see below). Therefore, in our study of the dynamical behaviour we can neglect the secular evolution of the field.

2.2.2 Dynamical time-scales

The *dynamical time-scale* is set by the traveltime of waves propagating in the different regions of the magnetar. In the magnetosphere, the mass density can be neglected in view of the dominating magnetic field energy density. Also, the velocity of Alfvén and fast magnetosonic waves is degenerated to the speed of light. Hence, within $\sim$100 km the whole magnetosphere is coupled through time-scales smaller than 1 ms, which sets the scale for the dynamical evolution of the magnetosphere. In this region it is possible to neglect the inertia of the fluid in the evolution equations of the so-called FFE, which is used in the numerical simulations of this work.

In the outermost parts of the crust, the force-free condition still holds because of low densities. At sufficiently high densities, elastic forces of the solid crust and pressure gradients break this condition.

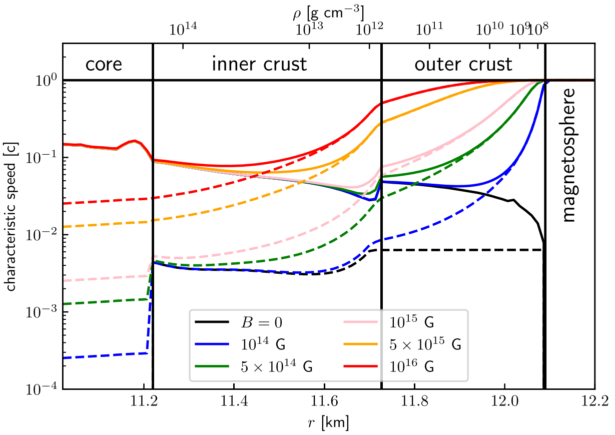

Figure 1. Fast magnetosonic (solid lines) and magnetoelastic (dashed lines) speed in the outer layers of a magnetar, for different magnetic field strengths ranging from 0 to 10^{16} G. The neutron star model corresponds to the $1.4M_\odot$ mass APR + DH model of Gabler et al. (2012). The magnetic field is considered to be constant for simplicity.

To estimate the transition density one may consider the depth at which waves propagate at a velocity significantly different to the speed of light. Two possible waves can travel in the interior of the magnetized crust, the so-called magnetosonic (ms) waves, related to sound waves, and magneto-elastic (me) waves, a combination of Alfvén and shear waves. The complete eigenvalue structure of relativistic ideal MHD equations in the presence of an elastic solid is not known. To make a simple order of magnitude estimate of the different wave speeds, we use the expression of magnetoelastic torsional waves parallel to the magnetic field derived in Gabler et al. (2012) as well as the expression for fast magnetosonic waves perpendicular to the field:[1]

$$v_{\rm me}/c = \sqrt{\frac{\mu_s + B^2}{e + B^2}} \qquad v_{\rm ms}/c = \sqrt{\frac{e c_s^2 + B^2}{e + B^2}}, \tag{4}$$

where e is the energy density and μ_s the shear modulus. Note that in the limit of low magnetic field ($B^2 \ll \mu_s,\ B^2 \ll e$) we recover the shear and sound speed, respectively. In the high magnetic field limit ($B^2 \gg \mu_s,\ B^2 \gg e$) both, $v_{\rm me}$ and $v_{\rm ms}$, coincide with the speed of light. Inside the fluid core ($\mu_s = 0$) the magnetoelastic speed becomes the Alfvén speed.

Fig. 1 shows the value of the characteristic speeds in the outer layers of a typical NS model for different magnetic fields in the magnetar range. Indeed, both fast magnetosonic waves and Alfvén waves have a degenerate speed equal to the speed of light in the magnetosphere. Inside the outer crust ($\rho < 4 \times 10^{11}\,{\rm g\,cm^{-3}}$), all characteristic speeds transition from the speed of light to a significantly lower value, in a region that can still be considered force-free. This transition depends on the magnetic field strength, happening deeper inside the star for larger values of $B_{\rm pole}$. Given these characteristic speeds, any global rearrangement of the magnetosphere can modify the entire structure of the crust (of size $\sim 2\pi R_*$) on a time-scale of $\sim$1 ms for magnetosonic waves and $\sim$10 ms for magnetoelastic waves.

---

[1] Slow magnetosonic waves are also possible but their velocity is much smaller and not relevant for this work, in fact, for the case of waves perpendicular to the magnetic field their speed is zero.

One last aspect to consider is the ability of magnetospheric waves to transmit energy into the crust. The discussion should be limited to Alfvén waves, which become magnetoelastic waves once they penetrate the crust; the energy carried by fast magnetosonic waves in the magnetosphere can be neglected due to the small density, which renders the compressibility effects of fast-magnetosonic waves unimportant.

Since the characteristic time in the magnetosphere is ∼1 ms, the typical frequency of the waves generated during its dynamics is in the kHz range. At this frequency, the crust can be considered as a thin layer because its thickness (∼1 km) is much smaller than the typical wavelength in the magnetosphere ($\lambda \sim 100$ km). In this case the energy transmission coefficient for waves perpendicular to the surface is approximately (cf. Link 2014; Li & Beloborodov 2015)

$$\mathcal{T} = \frac{4 v_{\rm me}/c}{(1 + v_{\rm me}/c)^2} \approx 0.04 \left(\frac{v_{\rm me}/c}{0.01} \right), \tag{5}$$

for typical physical conditions in the magnetar crust. Given the low transmission coefficients of magnetospheric Alfvén waves hitting the crust as well as the differences on time-scales between the crust and the magnetosphere (typically ∼10 times shorter in the later) it is reasonable to consider that most of the crust remains rigid during any dynamical rearrangement of the magnetosphere.

In our magnetar model we will consider two regions: A force-free region (*exterior*, hereafter) consisting of the magnetosphere and the force-free part of the outer crust as well as the magnetar *interior* for the remainder of the NS, which we will consider to be fixed during our simulations. The limit between both regions is a spherical surface below the NS surface, where magnetic field lines are anchored, and is located below the transition density between inner and outer crust at a density $\rho < 4 \times 10^{11}\,{\rm g\,cm^{-3}}$. For the purpose of describing the simulations we will refer to this transition point simply as *surface*.

3 FORCE-FREE ELECTRODYNAMICS

In analogy to Komissarov (2004) and Parfrey, Spitkovsky & Beloborodov (2017) we solve Maxwell's equations in the force-free limit:

$$\frac{\partial \tilde{\boldsymbol{B}}}{\partial t} = -\nabla \times \tilde{\boldsymbol{E}} \quad \text{and} \quad \frac{\partial \tilde{\boldsymbol{E}}}{\partial t} = \nabla \times \tilde{\boldsymbol{B}} - \tilde{\boldsymbol{J}}_{\rm FF}, \tag{6}$$

where $\tilde{\boldsymbol{E}}$, $\tilde{\boldsymbol{B}}$, and $\tilde{\boldsymbol{J}}_{\rm FF}$ are the electric field, the magnetic field, and the so-called force-free current, respectively. We place a tilde to distinguish quantities expressed in our Heaviside-Lorentz *geometrized* (HLG) system of units, while the same symbols without tilde express quantities in the Gaussian *non-geometrized* (GNG) system of units (see Table 1). We explicitly include the charge conservation equation

$$\frac{\partial \tilde{\rho}_{\rm e}}{\partial t} + \nabla \cdot \tilde{\boldsymbol{J}}_{\rm FF} = 0, \tag{7}$$

where $\tilde{\rho}_{\rm e}$ represents the charge density. Furthermore, we use a mixed hyperbolic/parabolic correction by the introduction of additional potentials (further discussed in Appendix A1) in order to numerically ensure the constraints $\nabla \cdot \tilde{\boldsymbol{B}} = 0$ and $\nabla \cdot \tilde{\boldsymbol{E}} = \tilde{\rho}_{\rm e}$ (Dedner et al. 2002; Palenzuela et al. 2009; Mignone & Tzeferacos 2010).

Table 1. Conversion table between code output in Heaviside-Lorentz *geometrized* units ($M_\odot = G = c = 1$) and *non-geometrized* Gaussian units. In order to transform the respective quantities from code quantities to the *non-geometrized* system, one has to multiply the *geometrized* quantity by its conversion factor expressed in CGS.

Quantity	Non-geometrized unit	Conversion factor
Mass	$\rm M$	$\rm M_\odot$
Length	$\rm L$	$\mathrm{M}_\odot G c^{-2}$
Time	$\rm T$	$\mathrm{M}_\odot G c^{-3}$
Electric charge	$\rm L^{3/2} M^{1/2} T^{-1}$	$(4\pi)^{-1/2} \mathrm{M}_\odot G^{1/2}$
Electric field	$\rm L^{-1/2} M^{1/2} T^{-1}$	$(4\pi)^{1/2} \mathrm{M}_\odot^{-1} G^{-3/2} c^4$
Magnetic field	$\rm L^{-1/2} M^{1/2} T^{-1}$	$(4\pi)^{1/2} \mathrm{M}_\odot^{-1} G^{-3/2} c^4$
Current density	$\rm L^{-1/2} M^{1/2} T^{-2}$	$(4\pi)^{-1/2} \mathrm{M}_\odot^{-2} G^{-5/2} c^7$
(EM) Energy	$\rm L^2 M\, T^{-2}$	$\mathrm{M}_\odot\, c^2$
(EM) Stress	$\rm L^{-1} M\, T^{-2}$	$\mathrm{M}_\odot^{-2} G^{-3} c^8$

In the force-free limit it is necessary to guarantee that there are either no forces acting on the system or, more generally, that the forces of the system balance each other. This is equivalent to a vanishing net Lorentz force on the charges $\tilde{\rho}_{\rm e}$ (see e.g. Camenzind 2007):

$$\tilde{\boldsymbol{E}} \cdot \tilde{\boldsymbol{J}}_{\rm FF} = 0 \tag{8}$$

$$\tilde{\rho}_{\rm e} \tilde{\boldsymbol{E}} + \tilde{\boldsymbol{J}}_{\rm FF} \times \tilde{\boldsymbol{B}} = 0. \tag{9}$$

From equation (9) one readily obtains the degeneracy condition

$$\tilde{\boldsymbol{E}} \cdot \tilde{\boldsymbol{B}} = 0. \tag{10}$$

Additionally, force-free fields are required to be magnetically dominant, the magnetic field being always stronger than the electric one, such that the following condition must hold:

$$\tilde{\boldsymbol{B}}^2 - \tilde{\boldsymbol{E}}^2 \geq 0. \tag{11}$$

Conditions (10) and (11), as well as the conservation condition $\partial_t \left(\tilde{\boldsymbol{E}} \cdot \tilde{\boldsymbol{B}} \right) = 0$ can be combined in order to obtain an explicit expression for $\tilde{\boldsymbol{J}}_{\rm FF}$ (cf. Komissarov 2011; Parfrey et al. 2017):

$$\tilde{\boldsymbol{J}}_{\rm FF} = \left[\tilde{\boldsymbol{B}} \cdot \nabla \times \tilde{\boldsymbol{B}} - \tilde{\boldsymbol{E}} \cdot \nabla \times \tilde{\boldsymbol{E}} \right] \frac{\tilde{\boldsymbol{B}}}{\tilde{\boldsymbol{B}}^2} + \tilde{\rho}_{\rm e} \frac{\tilde{\boldsymbol{E}} \times \tilde{\boldsymbol{B}}}{\tilde{\boldsymbol{B}}^2}. \tag{12}$$

Across the literature (e.g. Komissarov 2004; Alic et al. 2012; Parfrey et al. 2017) we find various modifications in the definition of $\tilde{\boldsymbol{J}}_{\rm FF}$ in order to drive the numerical solution of the system of partial differential equations (6) towards a state which fulfils equation (10) by introducing a suitable cross-field conductivity. In the numerical setup, we choose to combine the prescription of Komissarov (2004) with the force-free current given above. This strategy effectively minimizes the violations of equations (10) and (11) by exponentially damping the (numerically induced) components of the electric field parallel to $\tilde{\boldsymbol{B}}$ and suitably adjusting the electric field in magnetospheric current sheets in order to obtain $\tilde{\boldsymbol{B}}^2 - \tilde{\boldsymbol{E}}^2 \rightarrow 0$ at these locations.

Throughout the literature, the magnetic dominance condition (11) condensates to a necessary condition of FFE (e.g. Uchida 1997; McKinney 2006). For some authors (e.g. McKinney 2006) the breakdown of the magnetic dominance implies the invalidity of the numerical model. Others (e.g. Uchida 1997) claim that some physical processes (e.g. radiation losses) taking place in the regions where condition (11) is breached may restore the magnetic dominance condition. Indeed, Uchida (1997) explicitly allows for transient phases violating condition (11) – these regions are then interpreted as abandoning the freezing of magnetic flux on to the flux of matter, being necessarily accompanied by dissipation. Following Uchida (1997), the force-free regime continues to be a valid approximation as long as the dissipative effects are only a small fraction of the total energy. The violation of the perpendicularity condition (10) is an additional source of (Ohmic) dissipation (studied for example in the context of Alfvén waves in force-free electrodynamics by Li et al. 2019). In practice, this channel of dissipation occurs when $\tilde{\boldsymbol{E}} \cdot \tilde{\boldsymbol{B}} \neq 0$ such that $\tilde{\boldsymbol{J}} \cdot \tilde{\boldsymbol{E}} \neq 0$. Currently used force-free codes aim to avoid the transient into this regime by numerically cutting back all violations of condition (11) (e.g. Palenzuela et al. 2010; Paschalidis & Shapiro 2013; Carrasco & Reula 2016) or include a suitable Ohm's law (e.g. Komissarov 2004; Spitkovsky 2006; Alic et al. 2012; Parfrey et al. 2017) in order to minimize these violations during a transient phase. Fig. 5 shows the breakdown of condition (11) during the simulation and hints towards the aforementioned dissipative processes. We refer to Appendix A2 as well as, for example, Lyutikov (2003) for further details on the necessary constraint preservation and limitations of the highly magnetized regime (such as the lack of physical reconnection). We will give a thorough review of the procedures employed in our code in a subsequent technical paper.

4 TWISTED MAGNETAR MAGNETOSPHERE MODELS

4.1 Magnetospheres

Due to the long rotational period of observed magnetars pushing the location of the light cylinder to great distances, it is possible to neglect the rotation of the neutron star when building numerical models of magnetospheres in the near zone. The equilibrium structure of a non-rotating axisymmetric force-free magnetosphere is given through the well-known GSE (Lüst & Schlüter 1954; Grad & Rubin 1958; Shafranov 1966). This approach has been followed in several recent papers (e.g. Spitkovsky 2006; Beskin 2010; Viganò, Pons & Miralles 2011; Fujisawa & Kisaka 2014; Glampedakis et al. 2014; Pili et al. 2015; Akgün et al. 2016, 2018a; Kojima 2017, 2018; Kojima & Okamoto 2018). In most of these works, the toroidal field is confined within a magnetic surface near the equator, smoothly transitioning to vacuum at large distances. In stationary, non-rotating, axisymmetric magnetosphere models, the toroidal field cannot extend to the poles. Otherwise, the toroidal field would extend all the way to infinity, thus, violating the requirements of finite magnetic energy. Following the notation of Akgün et al. (2016, 2018a), we write the axisymmetric magnetic field in terms of its poloidal and toroidal components:

$$\boldsymbol{B} = \nabla P \times \nabla\varphi + T\nabla\varphi, \tag{13}$$

where φ is the azimuthal angle in spherical coordinates. Here, P and T are the poloidal and toroidal stream functions. Expressed in the orthonormal spherical basis corresponding to the coordinates (r, θ, φ), the magnetic field can be explicitly computed from the potentials P and T as

$$B^r = \frac{1}{r^2 \sin\theta}\partial_\theta P, \tag{14}$$

$$B^\theta = -\frac{1}{r\sin\theta}\partial_r P, \tag{15}$$

$$B^\varphi = \frac{T}{r\sin\theta}. \tag{16}$$

For an axially symmetric force-free field, the functions T and P may be expressed in terms of each other and appear as solutions of the force-free GSE:

$$\left[\partial_r^2 + \frac{1-\mu^2}{r^2}\partial_\mu^2\right] P + T\frac{\mathrm{d}T}{\mathrm{d}P} = 0, \tag{17}$$

where $\mu = \cos\theta$. P and T are constant on magnetic surfaces or, equivalently, along magnetic field lines. P is related to the magnetic flux passing through the area centred on the axis and delineated by the magnetic surface. Therefore, its value at the poles is zero and increases towards the equator. The function T is related to the current passing through the same area. Its functional dependence on P can be chosen freely (consistently with any continuity and convergence requirements, particularly for the currents), which is equivalent to setting boundary conditions for T at the surface of the star. Here, we invoke the same functional form for $T(P)$ as in Akgün et al. (2016, 2018a). Thus, the toroidal field is confined within some *critical* magnetic surface ($P = P_\mathrm{c}$),

$$T(P) = \begin{cases} s \times (P - P_\mathrm{c})^\sigma & : \ P \geqslant P_\mathrm{c} \\ 0 & : \ \text{else} \end{cases}, \tag{18}$$

s being a parameter determining the relative strength of the toroidal field with respect to the poloidal field. In order to avoid divergences in the currents we must demand that the power index satisfies $\sigma \geq 1$. For a pure dipolar field, the poloidal stream function in the magnetosphere is

$$P = \frac{1}{2}B_\mathrm{pole}\frac{R_*^3}{r}\sin^2\theta, \tag{19}$$

while the toroidal stream function is $T = 0$ everywhere. We will consider the simplest cases where the boundary value of P at the surface of the magnetar coincides with that of a dipolar field, and, therefore, the initial data are symmetric with respect to the equator. For different choices of the functional relation $T(P)$ given by equation (18) we solve the GSE and obtain a twisted magnetospheric initial model. We would like to note that all equations can be rescaled with B_pole, hence, the results of our numerical simulations can be normalized to the field strength of interest.

The energy stored in the magnetosphere can be computed as a volume integral

$$\mathcal{E} = \frac{1}{8\pi}\int(\boldsymbol{B}^2 + \boldsymbol{E}^2)\,\mathrm{d}V. \tag{20}$$

For later reference and in order to normalize the energetic content of our models, we provide the energy stored in the magnetosphere of a pure dipolar magnetic field ($\vec{E} = 0$, $B^r = B_\mathrm{pole}(R_*/r)^3\cos\theta$, $B^\theta = (B_\mathrm{pole}/2)(R_*/r)^3\sin\theta$, $B^\varphi = 0$):

$$\mathcal{E}_\mathrm{d} = \frac{1}{12}B_\mathrm{pole}^2 R_*^3 = 8.3\times 10^{46}\,\mathrm{erg}\left(\frac{B_\mathrm{pole}}{10^{15}\,\mathrm{G}}\right)^2\left(\frac{R_*}{10\,\mathrm{km}}\right)^3. \tag{21}$$

Once the surface value of P and the functional relation $T(P)$ are defined, one can solve the GSE iteratively (as it is a non-linear equation), while imposing vacuum boundary conditions at large distances. We use the numerical code described in Akgün et al. (2018a) to build our initial data. Using this parametrization, the boundary condition at the surface of the neutron star for the GSE (values of P and T) is fully determined by four parameters B_pole, s, P_c, and σ. However, the solution of the GSE with this fixed boundary condition is not necessarily unique. Akgün et al. (2018a) showed that for sufficiently large magnetospheric twists, there exist degeneracies, i.e. different solutions of the GSE for the same boundary conditions (the same set of four parameters). These solutions differ in their energy, twist, and the radial extent of the toroidal currents.

Table 2 shows the parameters used to construct the initial data for our numerical simulations. Each of the series A, B, and C of initial models were chosen to have identical parameters but *different* magnetospheric energies and, hence, represent degenerate

Table 2. Overview of initial data models used in our simulations. s, σ, and $P_{\rm c}$ are the parameters determining the boundary condition at the surface of the neutron star (see Section 4.1). $\mathcal{E}$ denotes the total electromagnetic energy in the magnetospheres, which is normalized to the vacuum dipole energy $\mathcal{E}_{\rm d}$ (equation 21), hence without dimension. $\tilde{J}_{\max}$ denotes the maximum current density at $t=0$ (see Section 2.1 as well as Table 1 for unit conversion). The maximum initial electromagnetic stresses on the magnetar surface (equation 24) at $t=0$ are shown in the last two columns (i.e. $\tilde{T}^{ra}_{\max} := \max_{\{|\boldsymbol{x}|=R_*\}}\{\tilde{T}^{ra}(t=0,\boldsymbol{x})\}$, with $a=\theta,\varphi$). Values of $\tilde{J}_{\max}$ and $\tilde{T}^{ra}_{\max}$ are given in HLG units for an NS with $B_{\rm pole}=10^{15}$ G and $R_*=13.7$ km.

	s	σ	$P_{\rm c}$	$\mathcal{E}/\mathcal{E}_{\rm d}$	$\tilde{J}_{\max}$	$\tilde{T}^{r\varphi}_{\max}$	$\tilde{T}^{r\theta}_{\max}$
A1	2	2	0.3294	1.1553	1.71e-6	8.97e-10	1.44e-9
A2	2	2	0.3303	1.3356	1.58e-6	8.95e-10	1.24e-9
B1	1	1	0.3717	1.1547	1.08e-6	7.68e-10	1.39e-9
B2	1	1	0.3720	1.2276	1.07e-6	7.68e-10	1.31e-9
C1	1	1	0.4400	1.0653	1.95e-6	6.68e-10	1.56e-9
C2	1	1	0.4412	1.1943	1.03e-6	6.68e-10	1.44e-9
C3	1	1	0.4396	1.2738	1.03e-6	6.71e-10	1.35e-9

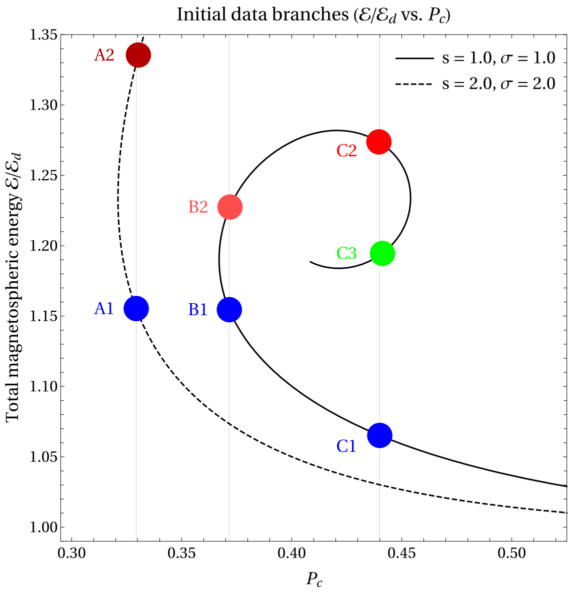

Figure 2. Magnetospheric energy normalized to the vacuum dipole energy (equation 21) of the initial equilibrium models, for different values of the parameter $P_{\rm c}$ (in units of P at the equator). The solid and dashed lines correspond to a series of models with constant s and σ. The coloured dots correspond to the initial data models used in our simulations.

magnetospheric models. We would like to point out that the value of $P_{\rm c}$ is only equal, within each series, up to the second significant digit, due to numerical reasons. Fig. 2 shows the energy of the initial models as a function of the parameter $P_{\rm c}$. Models within each spiral curve (constant s and σ) and with the same value of $P_{\rm c}$ have identical boundary conditions but different energies. In the interpretation made by Akgün et al. (2018a), the lower energy state for each series of degenerate models (i.e. A1, B1, and C1) corresponds to stable configurations, while high energy states (i.e. A2, B2, C2, and C3) may be unstable and would evolve towards the stable configuration releasing the respective energy difference. This instability is a possible scenario for the flare activity observed in magnetars.

The lowest energy solutions are the ones that are most similar to the vacuum solutions, with all field lines connected to the surface, while the higher energy solutions are more radially extended, and can contain disconnected field lines.

4.2 Magnetar interior

The initial models described above provide solutions only for the magnetosphere. For each possible magnetospheric model one can build infinite solutions to describe the neutron star interior. The magnetospheric (exterior) values of P and T determine the magnetic field $\boldsymbol{B}$ at the exterior side of the surface (equations 14 to 16). To match this solution to the interior, one has to ensure the continuity of B^r at the surface. This is valid if P is continuous and, hence, T and B^φ are continuous as well. However, B^θ does not necessarily match continuously to the neutron star interior because current sheets (thin current-carrying layers across which the magnetic field changes either direction or magnitude) in the φ direction may occur. Even if all components of $\boldsymbol{B}$ are continuous at the surface, the magnetic field structure in the interior depends completely on how currents are internally distributed.

In the astrophysical scenario we are considering, the magnetar reaches the initial state in which we start our numerical simulation after a slow magnetothermal evolution that proceeds in a long time-scale compared to the dynamical time-scales (cf. Section 2.2) of the magnetosphere (∼1 ms) or the crust (∼10 ms). On such long time-scales, any current close to the surface of the NS is expected to be dissipated by Ohmic diffusion. Therefore, we consider that initially all fields are continuous across the surface. We build our interior solution by extrapolating the exterior magnetic field towards the stellar interior across a number of grid cells as needed by the reconstruction algorithm used for the magnetospheric evolution in our simulations. Since the neutron star is basically a perfect conductor, the initial charge density and electric field in the interior (and the magnetosphere) are set to zero.

The surface values of B^r and B^φ are coincident for degenerate models (e.g. within the series C1, C2, and C3 in Fig. 2) because P and T at the surface are identical. However, since P and T may have a different radial dependence outside of the magnetar, and B^θ depends on the radial derivative of P (equation 15), it is different for every model of the same series.

5 SIMULATIONS

We have performed numerical simulations of the neutron star magnetosphere using the initial models in Table 2. For all the simulations we employ our own implementation of a General Relativistic FFE code in the framework of the `Einstein Toolkit`[2] (Löffler et al. 2012). The EINSTEIN TOOLKIT is an open-source software package utilizing the modularity of the `Cactus`[3] code (Goodale et al. 2003) which enables the user to specify the so-called thorns in order to set up customary simulations. There exist other code packages such as `GiRaFFE` (Etienne et al. 2017), which integrate the equations of force-free electrodynamics employing an evolution scheme based on the Poynting flux as a conserved quantity (cf. McKinney 2006; Paschalidis & Shapiro 2013) rather than the electric field and its current sources (as formulated in e.g. Komissarov 2004; Parfrey et al. 2017). The `Einstein Toolkit` employs units where $\mathrm{M}_\odot = G = c = 1$, which sets the respective time and length scales to be $1\,\mathrm{M}_\odot \equiv 4.93\times10^{-6}$ s $\equiv 1477.98$ m. This unit system is a variation of the so-called system of *geometrized units* (as introduced in appendix F of Wald 2010), with the additional normalization of the mass to 1 $M_\odot$ (i.e. our HLG units, as introduced in Section 3). For easy reference, we provide a set of conversion factors for relevant physical quantities in Table 1.

---

[2] http://www.einsteintoolkit.org

[3] http://www.cactuscode.org

5.1 Numerical setup

All shown simulations are conducted on a 3D box with dimensions $[4741.12\,M_\odot \times 4741.12\,M_\odot \times 4741.12\,M_\odot]$ with a grid spacing of $\Delta_{x,y,z} = 74.08\,M_\odot$ on the coarsest grid level. For the chosen magnetar model of radius $R_* = 9.26\,M_\odot$ ($\simeq 13.7$ km) this corresponds to a $[512R_* \times 512R_* \times 512R_*]$ box with a grid spacing of $\Delta_{x,\,y,\,z} = 8R_*$. For the low and high resolution tests we employ seven and eight additional levels of mesh refinement, each increasing the resolution by a factor of two and encompassing the central object, respectively. This means that the finest resolution of our models (close to the magnetar surface) are $\Delta^{\min}_{x,y,z} = 0.0625 \times R_* = 0.5787\,M_\odot$ and $\Delta^{\min}_{x,y,z} = 0.03125 \times R_* = 0.2894\,M_\odot$ for the low and high resolution models, or in other words 16 and 32 points per R_*, respectively. The initial data are evolved for a period of $t = 1185.28\,M_\odot \simeq 5.84$ ms, which is chosen to be well below the dynamical time-scale of the magnetar crust, which can be considered as a fixed boundary (see Section 2.2).

In order to ensure the conservation properties of the algorithm, it is critical to employ *refluxing* techniques correcting numerical fluxes across different levels of mesh refinement (see e.g. Collins et al. 2010). Specifically, we make use of the thorn `Refluxing`[4] in combination with a cell-centred refinement structure (cf. Shibata 2015). We highlight the fact that employing the refluxing algorithm makes the numerical code 2–4 times slower for the benefit of enforcing the conservation properties of the numerical method (especially of the charge). Refluxing also reduces the numerical instabilities, which tend to develop at mesh refinement boundaries.

In conservative schemes, numerical reconstruction algorithms (we employ an MP7 scheme; cf. Suresh & Huynh 1997) derive intercell approximations of the conservative variables by making use of their values at several adjacent grid-points (for MP7, one requires seven points). As a result of the numerical coupling between the magnetosphere and the magnetar crust introduced by the intercell reconstruction at the stellar surface, the field dynamics induce a mismatch in the current flowing through the surface and effectively trigger a (numerical) flow of charges leaving or entering the domain. In order to avoid this artefact, we replace the reconstructed values of the radial current $\tilde{J}^r_{\mathrm{FFE}}$ at interfaces between the stellar interior and exterior by the cell-centred value in the stellar interior. This procedure ensures a conservation of magnetospheric charge.

The (3D) initial data are imported from the (2D) initial models (see section 4.1) by bicubic spline interpolation. Throughout the numerical evolution, all quantities on grid-points inside of the magnetar radius are fixed to their initial values.

5.2 Instability onset and magnetospheric energy balance

We have performed simulations with initial models in the low energy branch (A1, B1, and C1) and in the high energy branch (A2, B2, C2, C3). We observe a differentiated behaviour in the evolution of the system depending on the class of initial model. For models in the low energy branch we find that the magnetosphere is stable and that the system remains essentially unchanged. The energy of the system remains constant throughout the simulation (see blue lines in Fig. 3), confirming the stability of these configurations, at least on dynamical time-scales. This is especially true in the high resolution models, which exhibit a smaller numerical dissipation. The slightly larger numerical dissipation of the low resolution models explains the small drift in time with respect to the initial energy displayed by the blue dashed lines in Fig. 3. On the other hand, models in the high energy branch become unstable on a time-scale of a few milliseconds and the magnetosphere changes its shape roughly at the same time as the energy of the magnetosphere decreases (see red and green lines in Fig. 3). This numerical experiment confirms the hypothesis of Akgün et al. (2018a) that, for degenerate initial models, only the lowest energy state is stable, and that all corresponding degenerate cases of high energy are unstable. In addition, we note that the lower energy states are closer to a purely dipolar magnetosphere, hence, the minimized circumference of the magnetic surfaces minimize the magnetospheric energy content (cf. Thompson & Duncan 1996).

For configurations in the unstable branch, the onset of the instability proceeds earlier for lower numerical resolution. This is expected because a coarser grid contains larger numerical discretization errors acting as a seed for the instability onset. However, the rapid drop in energy during the instability proceeds in a similar fashion for both numerical resolutions, indicating that the instability has a physical origin and is not a numerical artefact. In the case of the high energy initial model C2 we observe a rearrangement of the lobes of magnetic twist towards a dipolar structure (see Fig. 4) prior to a significant drop of magnetospheric energy (by approximately 30 per cent of its initial value). During the phase of full validity of the force-free condition (see equation 11) the loss of magnetospheric energy is dominated by an outgoing Poynting flux at the innermost boundary (see Fig. 5). For our boundary condition it can be interpreted as the formation of a strong current on a thin layer below the surface, where energy can be efficiently dissipated.

Following Parfrey et al. (2013) in the context of twisted magnetar fields and Li et al. (2019) in a study of energy dissipation in collisions of force-free Alfvén waves, the onset of the (topological) relaxation is likely to be linked to Ohmic heating $\boldsymbol{J} \cdot \boldsymbol{E} \neq 0$, which occurs as a result of (minor) violations of the force-free condition (10), as can be seen in the bottom panel of Fig. 5 (note the much smaller scale of that panel compared to the middle one). We give a more detailed review of the treatment of these violations in our code and throughout the literature in Appendix A2.

5.3 Surface currents and long-time evolution

Following the initial instability and subsequent rapid rearrangement of the magnetar magnetosphere (Section 5.2), thin currents form at the magnetar surface (see Figs 6 and 7). These currents are expected to appear as the initial model in the high energy state tries to relax to the lowest energy magnetospheric configuration, while keeping the interior field fixed (see the discussion in Section 4.2). There are two possible fates for these currents: (i) they could propagate inwards, inside the magnetar crust, deforming the magnetic field inside, and creating a mechanical stress in the crust, on a time-scale of several 10 ms, or (ii) they could form a thin surface current dissipating on a time-scale shorter that the time it takes to deform the crust. These two possibilities are not mutually exclusive and a combination of both is possible. In none of the cases our simulations can give a conclusive answer because (i) we are not evolving the

---

[4] Refluxing at mesh refinement interfaces by Erik Schnetter: https://svn.cct.lsu.edu/repos/numrel/LSUThorns/Refluxing/trunk

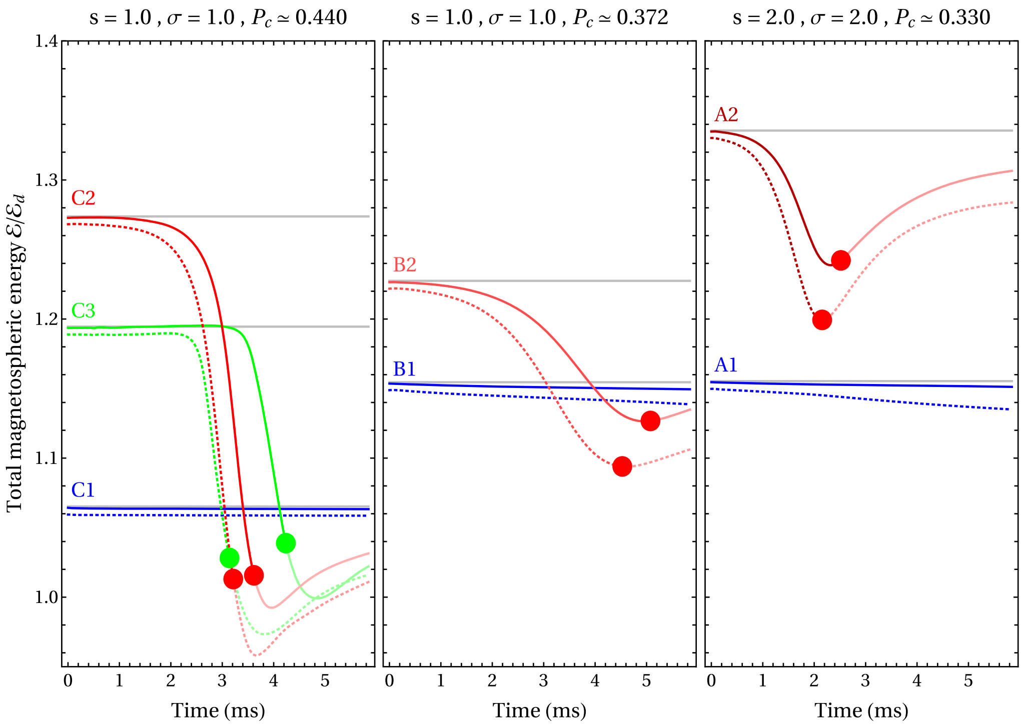

Figure 3. Time evolution of total magnetospheric energy content for the models in Table 2. The instability of the field configuration for degenerate solutions of higher energy triggers the rearrangement of magnetic field lines as well as a release of energy into the magnetosphere and on to the magnetar surface. The simulated time-scale on which the instabilities are observed falls within the dynamical time-scale of the magnetar crust. Low resolution simulations (16 points per R_*) are shown in dotted lines, high resolution simulations (32 points per R_*) in solid lines. The initial (analytical) value of total magnetospheric energy for each configuration is indicated by grey lines. The approximate time of the breakdown of the force-free condition $\tilde{\boldsymbol{E}}^2 - \tilde{\boldsymbol{B}}^2 < 0$ (see Appendix A2) is depicted by coloured dots.

magnetar interior as we are considering only time-scales smaller than the dynamical time-scale of the crust, (ii) the formation of thin surface currents is numerically challenging (would require a computationally prohibitively high resolution near the magnetar surface), and (iii) it would eventually violate the FF conditions (10) and (11), hence invalidating our current numerical approach.

The aforementioned current layers are expected to be regions of strong energy dissipation and the breakdown of the force-free conditions (see e.g. Uchida 1997; McKinney 2006; Palenzuela et al. 2010; Parfrey et al. 2013). Figs 5 and 7 link the breakdown of the force-free condition (11) and the occurrence of surface currents with the opening of dissipation channels different to the energy flow through the magnetar surface (see Appendix A2 for a short review of the force-free breakdown). We find the violation of condition (10) to be continuously occurring with peaks at the instance of rapid energy dissipation. Condition (11) starts to fail on longer time-scales at the moment of fastest transfer of magnetic energy through the surface. At this time, further dissipation mechanisms (see Fig. 5) come into play, as is expected throughout the literature (Uchida 1997; McKinney 2006; Li et al. 2019).

It should be noted that the total magnetospheric energy for the models B2, C2, and C3 drops below the energy of their respective low energy equilibrium solutions, and even below the magnetospheric energy of the vacuum dipole (equation 21). However, this energy drop is (slightly) smaller for the high resolution simulations, and shows some dependence on the chosen setup of the hyperbolic/parabolic cleaning procedures (see Appendix A1) at the magnetar surface. The sensitivity of this behaviour to the numerical details at the location of the (3D Cartesian) crust may be attributed to the numerical dissipation of the employed code.

6 DISCUSSION

6.1 Energy release during the instability

During the rearrangement of magnetic field lines in the high energy models A2, B2, C2, and C3, an amount $\Delta\mathcal{E}$ of electromagnetic energy is released into the magnetosphere and on to the magnetar crust (Poynting flux through the stellar surface, see Fig. 5). The amount of released energy in CGS units, $\mathcal{E}_r$, can be quantified directly from Table 3 by employing the conversion formula

$$\mathcal{E}_\mathrm{r} = 2.14 \times 10^{47}\,\mathrm{erg}\left(\frac{\Delta\mathcal{E}}{\mathcal{E}_\mathrm{d}}\right)\left(\frac{B_\mathrm{pole}}{10^{15}\,\mathrm{G}}\right)^2\left(\frac{R_*}{13.7\,\mathrm{km}}\right)^3 . \quad (22)$$

For the changes in energy ($\Delta\mathcal{E}/\mathcal{E}_\mathrm{d} \approx 0.1 - 0.3$) observed in our simulations with the highest energy within each series (C2, C3, B2, and A2) the released energy is in the range $\mathcal{E}_\mathrm{r} \approx 2.1 \times 10^{46} - 6.4 \times 10^{46}$ erg. This energy range is compatible with that of observed GFs (10^{45}–10^{48} erg). For instance, the energy liberated during the peak of the GF of SGR 1806-20 is $\sim 3.7 \times 10^{46}$ erg (Hurley et al. 2005), which is compatible with values $\Delta\mathcal{E}/\mathcal{E}_\mathrm{d} \simeq 0.17$. However, the other two known GF events (SGR 0525-66 and SGR 1900 + 14; see Cline et al. 1980; Hurley et al. 1999) display significantly smaller amounts of energy during their initial peaks.

The range of $\Delta\mathcal{E}/\mathcal{E}_\mathrm{d}$ in our simulations depends on the choice of initial models. The detailed analysis in Akgün et al. (2018a) shows that $\Delta\mathcal{E}/\mathcal{E}_\mathrm{d}$ could, in principle, be as large as 0.8 for models with the appropriate values of s and σ and the value of P_c to be at the maximum of the corresponding sequence (see fig. 3 in Akgün et al. 2018a). However, the astrophysical path that could lead to an unstable configuration this far away from the equilibrium branch is

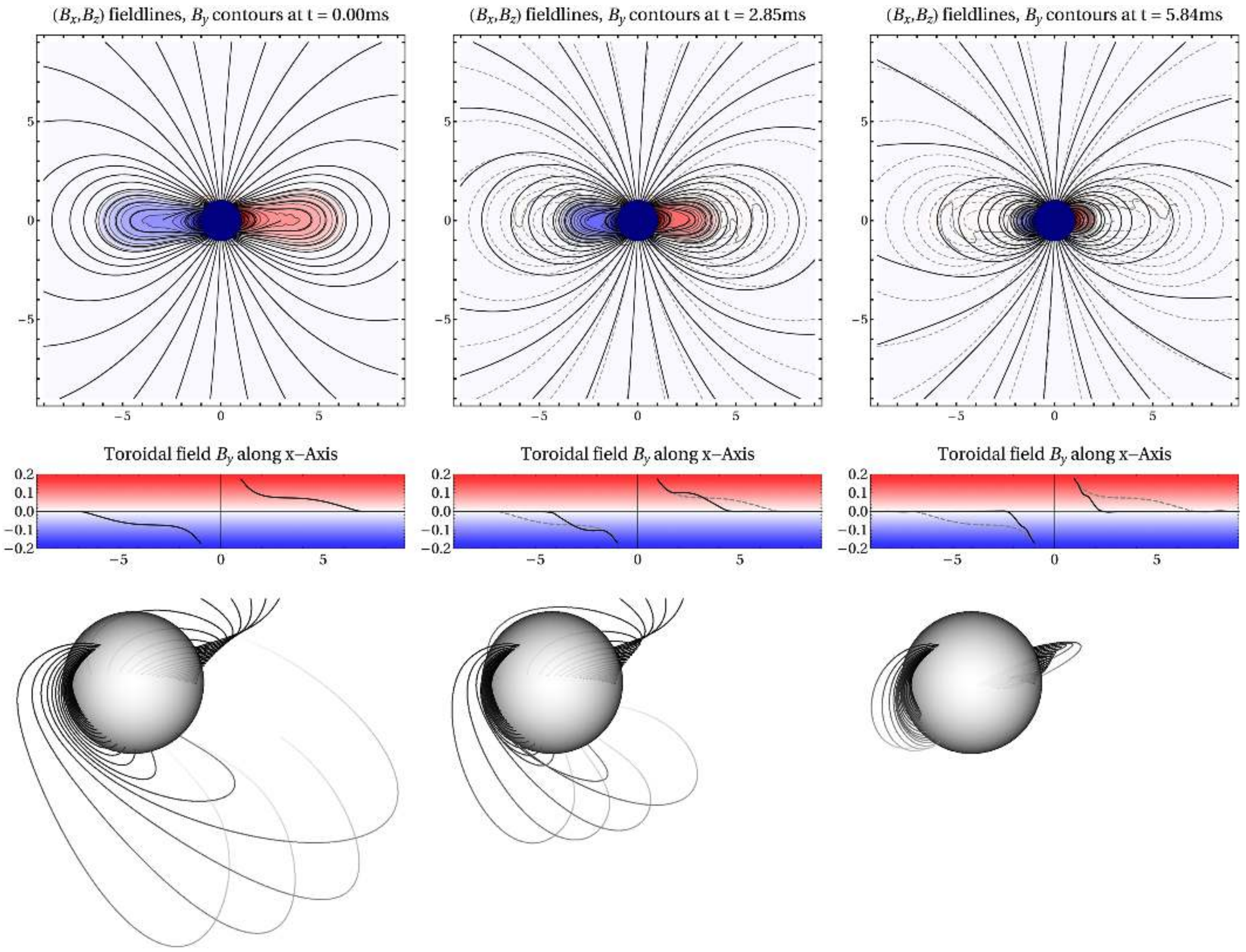

Figure 4. Field line evolution (high resolution, 32 points per R_*) of the C2 initial model (Table 2). The initially extended lobes of magnetic twist relax towards a dipolar structure and fall towards the central object. Strong energy dissipation (see Fig. 3) occurs when the magnetic twist collapses on to the magnetar crust. The final configuration is dipole-like, though it fully relaxes on a much longer dynamical time-scale. *Top:* Poloidal field lines (cross-section through the 3D data) and colour-filled contours of the toroidal magnetic field (same colour coding as below). The initial field line configuration is indicated by grey dashed lines. *Middle:* Toroidal field distribution along the *x*-axis. The initial toroidal magnetic field is denoted by grey dashed lines. *Bottom:* Evolution of selected field lines in 3D, displaying the twist relaxation. *Click for animation:* Evolution of total magnetospheric energy and selected field lines in 3D of the (high resolution) C2 initial model (Acrobat Reader only).

unclear. Speaking in terms of evolution, models close to the stability threshold for which $\Delta\mathcal{E}/\mathcal{E}_\mathrm{d}$ could be a small fraction of the energy encountered in our simulations are much more likely than models with values of e.g. $\Delta\mathcal{E}/\mathcal{E}_\mathrm{d} > 0.2$.

The time-scale on which $\mathcal{E}_\mathrm{r}$ is released ($\Delta t_\mathrm{r} \sim 1{-}5\,\mathrm{ms}$; see Table 3) is consistent with the dynamical time-scales in the magnetosphere (Section 2.2.2). If we estimate the luminosity of the energy released as

$$L_0 := \frac{\mathcal{E}_\mathrm{r}}{\Delta t_\mathrm{r}}\,, \tag{23}$$

we find that $L_0 \sim (0.7 - 4) \times 10^{49}\,\mathrm{erg\,s^{-1}}$ for the unstable models listed in Table 4. This dynamical luminosity is significantly larger than the peak luminosity of GFs (e.g. the peak luminosity of SGR 1806-20 is $\sim 2 \times 10^{47}\,\mathrm{erg\,s^{-1}}$; Hurley et al. 2005), and suggests that only a fraction of the released energy contributes to explain the thermal properties of GFs in SGRs. As an alternative, not necessarily exclusive, we consider different mechanisms to broaden the time-scale over which the energy leaks out of the system, hence reducing L_0, in the following sections.

6.2 Stresses induced in the crust

Fig. 5 suggests that a significant part of the released energy is transferred into the magnetar crust during the (fully force-free) evolution. We would like to point out that an exact modelling of magnetar crust physics will be necessary in order to simulate respective feedback mechanisms between the stellar surface and the magnetosphere. However, in this section we make some crude estimates regarding the stresses induced in the crust as a result of the magnetospheric evolution of our models.

The stresses induced in the crust by the evolving magnetosphere can be computed studying the momentum-transfer from the magnetosphere to the crust. The stress tensor in the (force-free)

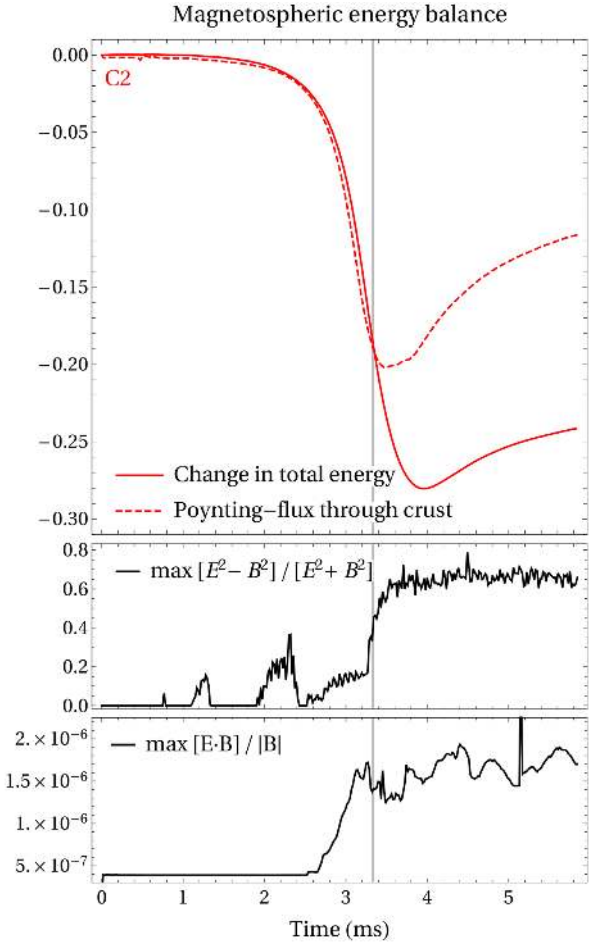

Figure 5. Energy balance during the evolution of the high resolution model C2 (Table 2). *Top:* Comparison of the change in total magnetospheric energy, normalized to the energy of a magnetosphere equipped with a pure dipolar magnetic field, $\Delta\mathcal{E}/\mathcal{E}_{\rm d}$, as well as the Poynting flux through the magnetar surface. Up to a simulation time of $t \sim 3.33$ ms the energy change is dominated by Poynting flux on to the magnetar crust. *Middle:* Maximum violation of the $\tilde{\boldsymbol{B}}^2 - \tilde{\boldsymbol{E}}^2 \geq 0$ condition throughout the numerical grid. *Bottom:* Maximum violation of the $\tilde{\boldsymbol{B}} \cdot \tilde{\boldsymbol{E}} = 0$ constraint throughout the numerical grid. At the time of the breakdown of conditions (10) and (11), the energy change is dominated by secondary (possibly numerical) effects.

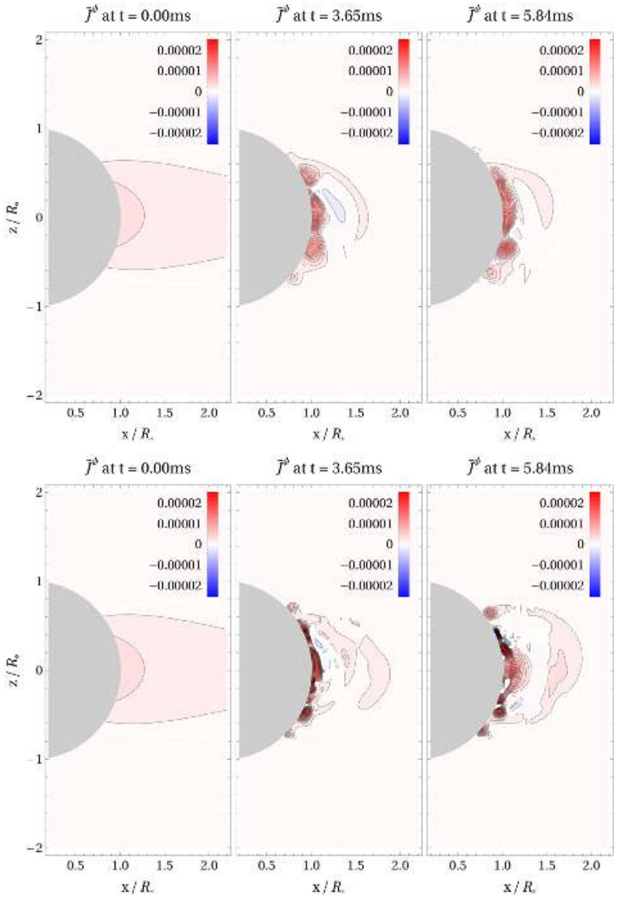

Figure 6. *xz*-cross-sections of the toroidal current in geometrized units showing the development of strong surface currents during the evolution, in addition to other currents extended on larger magnetospheric volumes. *Top:* Low resolution model C2 (16 points per R_*). *Bottom:* High resolution model C2 (32 points per R_*). The high resolution evolution shows currents located around the magnetar surface with more detailed structures, emphasizing their interpretation as surface currents. The spatial coincidence of the currents in both resolutions reinforce the argument that the observed currents are of physical origin (in spite of the – relatively small – differences among different resolutions).

magnetosphere is

$$T^{ij}_{\rm ms} = \frac{1}{4\pi}\left(\frac{1}{2}\delta^{ij}\left(E^2_{\rm ms}+B^2_{\rm ms}\right) - E^i_{\rm ms}E^j_{\rm ms} - B^i_{\rm ms}B^j_{\rm ms}\right), \qquad (24)$$

where $B^i_{\rm ms}$, and $E^i_{\rm ms}$ are the magnetic and electric fields in the magnetosphere. The stress tensor in the crust consists of the contribution of the magnetic field, the fluid, and the stress of the solid

$$T^{ij}_{\rm c} = \mathcal{P}\delta^{ij} + \frac{1}{4\pi}\left(\frac{1}{2}\delta^{ij}B^2_{\rm c} - B^i_{\rm c}B^j_{\rm c}\right) + \sigma^{ij}, \qquad (25)$$

where $\mathcal{P}$ is the pressure of the fluid, $B^i_{\rm c}$ the magnetic field inside the crust, and σ^{ij} is the stress tensor of the deformed solid. Especially, $\sigma^{ij} = 0$ for a non-deformed solid – which holds at the beginning of the presented simulations in which the crust is relaxed after the long-term magneto-thermal evolution during which plastic deformations can keep this relaxed state. Throughout the instability phase captured in our simulations, the magnetosphere induces a stress in the crust that effectively deforms it. The Lagrangian displacement of any point in the crust with respect to the relaxed state is given by the deformation vector ξ^i. For linear displacements, the stress tensor can be expressed in terms of the deformation vector (Landau & Lifshitz 2012) as follows:

$$\sigma^{ij} = K\xi^k_{;k} f^{ij} + 2\mu\left(\frac{1}{2}(\xi^{j;i}+\xi^{i;j}) - \frac{1}{3}f^{ij}\xi^k_{;k}\right), \qquad (26)$$

where semicolon indicates the covariant derivative, f_{ij} the flat 3-metric, K is the bulk modulus, and μ the shear modulus. Crust and magnetosphere can only interchange momentum through $T^{r\theta}$ and $T^{r\varphi}$. Hence, these are the only relevant components. Imposing continuity of these two components at the surface of the star ($\mathcal{P} = 0$) one finds

$$-\frac{1}{4\pi}\left(E^r_{\rm ms}E^a_{\rm ms} + B^r_{\rm ms}B^a_{\rm ms}\right) = -\frac{1}{4\pi}B^r_c B^a_c + \sigma^{ra} \qquad a=\{\theta,\varphi\}, \qquad (27)$$

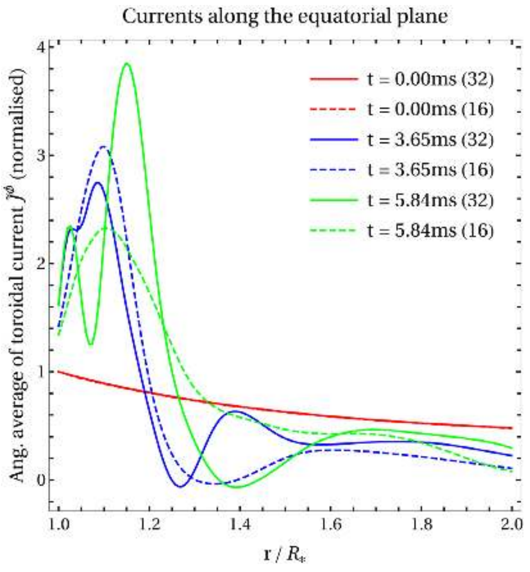

Figure 7. Azimuthal angular averages of the toroidal current (normalized to its initial value at the stellar surface) in the equatorial plane showing the development of surface currents during the evolution of the C2 initial model. We display the current evolution for both low resolution (16 points per R_*, denoted by dashed lines), and high resolution (32 points per R_*, denoted by solid lines) models. The increase of the toroidal current during the transient of energy dissipation (see Fig. 3) at the lower resolution (compare the two blue lines) may be attributed to a faster onset of the twist instability for this model.

Table 3. Selection of electromagnetic quantities monitored throughout the (high resolution, 32 points per R_*) simulation of the models of Table 2. The total change in energy $\Delta\mathcal{E}$ (displayed as a fraction of the vacuum dipole energy; equation 21) corresponds to the maximum drop of electromagnetic energy during the total runtime (see Section 6.1). The operator $\Delta_{\rm mx}$ acting on any quantity $A(t, \boldsymbol{x})$ is defined as $\Delta_{\rm mx}A := \max_{\{t,|\boldsymbol{x}|=R_*\}}\{A(t,\boldsymbol{x}) - A(0,\boldsymbol{x})\}/\max_{\{|\boldsymbol{x}|=R_*\}}A(0,\boldsymbol{x})$. Hence, $\Delta_{\rm mx}J$ is the maximum increase in current density in the magnetosphere during the relaxation relative to the initial values (see Section 2.1). In the right-hand columns, $\Delta_{\rm mx}T^{r\varphi}$ and $\Delta_{\rm mx}T^{r\theta}$ denote the maximum increase of electromagnetic stresses relative to their corresponding initial values (see Section 6.2) on the stellar surface compared to its initial value. We highlight with bold face the maximum values of each of the last four columns.

Model	$\Delta t_{\rm r}$(ms)	$\Delta\mathcal{E}/\mathcal{E}_{\rm d}$	$\Delta_{\rm mx}J$	$\Delta_{\rm mx}T^{r\varphi}$	$\Delta_{\rm mx}T^{r\theta}$
A1	5.8400	0.0033	0.0159	0.0012	0.0010
A2	1.4162	0.0963	1.6350	0.0295	0.0150
B1	5.8400	0.0042	0.0363	0.0012	0.0014
B2	3.0427	0.1002	0.9805	0.0358	0.0232
C1	5.8400	0.0009	0.0640	0.0008	0.0013
C2	2.1604	**0.2808**	**3.5400**	0.0851	0.0414
C3	1.0490	0.1962	3.1720	**0.1008**	**0.0811**

and therefore

$$\sigma^{ra} = \frac{1}{4\pi}\left(B^r_{\rm c}B^a_{\rm c} - E^r_{\rm ms}E^a_{\rm ms} - B^r_{\rm ms}B^a_{\rm ms}\right) \qquad a = \{\theta, \varphi\}. \tag{28}$$

For the equilibrium configuration at the beginning of the simulation, in which $\boldsymbol{E} = 0$ and $\boldsymbol{B}$ is continuous (no initial current sheets), the mechanical stress is zero ($\sigma^{ra} = 0$) and, hence, the stress at the surface is just $T^{ra}_{\rm c} = -B^r_{\rm ms}(t = 0)B^a_{\rm ms}(t = 0)/(4\pi)$. Therefore, we can compute the mechanical stress at any time as

$$\sigma^{ra} = T^{ra}_{\rm ms} - T^{ra}_{\rm ms}(t = 0). \tag{29}$$

As discussed in Section 2.2.2, the magnetic fields are dominant in the outermost low-density part of the crust and can be considered to be force-free (Beloborodov 2009). The point at which the magnetic field lines are anchored is not the surface of the star, but some radius, $r_{\rm c}$, below it (see also the discussion referencing Fig. 1). However, equation (28) still holds at this radius, because P is continuous, and the relevant terms cancel out. In other words, from the point of view of the numerical simulation, the inner boundary condition therein used corresponds to $r_{\rm c}$, and not the radius of the star. The force-free region of the crust corresponds to the region where shear stresses do not play a role in the dynamics, i.e. $\mu \ll B^2$. For typical magnetar magnetic fields of $B \sim 10^{15}$ G this is fulfilled for $\mu_{\rm c} \ll 10^{30}\,{\rm erg\,cm^{-3}}$, which typically and for a large variety of equations of state (Steiner & Watts 2009) corresponds to densities of $\rho \ll 10^{14}\,{\rm g\,cm^{-3}}$.

For the discussion at hand, we will consider that the anchoring is produced at some point between the inner crust outer boundary ($\rho \approx 4 \times 10^{11}\,{\rm g\,cm^{-3}}$), with $\mu_{\rm IC} \approx 1.4 \times 10^{28}\,{\rm erg\,cm^{-3}}$, and $\mu_{14} \sim 10^{30}\,{\rm erg\,cm^{-3}}$, its value close to the core-crust transition, at about $10^{14}\,{\rm g\,cm^{-3}}$. The relevant components of the stress tensor in spherical coordinates are

$$\sigma^{r\theta} = 2\mu\, s^{r\theta} = \mu\left[r\partial_r\left(\frac{\xi^\theta}{r}\right) + \frac{1}{r}\partial_\theta\xi^r\right], \tag{30}$$

$$\sigma^{r\varphi} = 2\mu\, s^{r\varphi} = \mu\left[r\partial_r\left(\frac{\xi^\varphi}{r}\right) + \frac{1}{r\sin\theta}\partial_\varphi\xi^r\right], \tag{31}$$

where s^{ij} is the strain tensor. For sufficiently large strains the crust will fail and a rapid plastic deformation will deform the crust persistently. The breaking strain of the crust has been estimated to be about 0.1 (Horowitz & Kadau 2009). Therefore, any stress larger than $\sim 0.2\mu_{\rm c}$ will likely produce a failure in the crust. The maximum mechanical stress exerted on the magnetar crust, $\sigma^{ra}_{\rm max}$, can be quantified directly from the results shown in Tables 2 and 3 by employing the conversion formula

$$\sigma^{ra}_{\rm max} = 5.55 \times 10^{28}\,{\rm erg\,cm^{-3}}\left(\frac{\Delta_{\rm mx}T^{ra}}{0.1}\right)\left(\frac{\tilde{T}^{ra}_{\rm max}}{10^{-9}}\right)\left(\frac{B_{\rm pole}}{10^{15}\,{\rm G}}\right)^2. \tag{32}$$

The maximum mechanical stress (see Fig. 8) on the magnetar crust measured throughout the shown simulations (see Tables 2 and 3) correspond to $\sigma^{ra} \approx 10^{28}\,{\rm erg\,cm^{-3}}$ for $B_{\rm pole} \approx 10^{15}$ G. Considering the quadratic leverage of the magnetic field strength, mechanical stresses of $\sigma^{ra} \approx 10^{30}\,{\rm erg\,cm^{-3}}$ are likely to be reached for $B_{\rm pole} \approx 10^{16}$ G and beyond. The largest mechanical stresses are exerted in case of the high energy models A2, B2, C2, and C3.

Our numerical simulations indicate that the instability occurs in a quasi-axisymmetric way (cf. Fig. 4), with deviations from axisymmetry of less than 1 per cent.[5] In axisymmetry, axial displacements (ξ^φ) and polar displacements (ξ^r, ξ^θ) decouple and it is possible to estimate the axial displacement from the $\sigma^{r\varphi}$ component of the stress tensor. Although the magnetospheric dynamics can, in principle, induce radial deformations, ξ^r, in reality those deformations are strongly suppressed because they involve the motion of matter parallel to the gravitational field (not included in our calculation). Therefore, in practice one can consider $\xi^r = 0$, such that

$$\sigma^{r\varphi} = \mu\, r\partial_r\left(\frac{\xi^\varphi}{r}\right), \tag{33}$$

---

[5] We quantify these deviations performing a multipolar expansion of the eletromagnetic energy and evaluating the energy stored in modes with azimuthal numbers $m > 0$.

Table 4. Energetics of our models scaled to a magnetic field strength $B_{\rm pole} = 10^{15}$ G. (i) Energy released. (ii) Estimates of dynamic luminosity L_0 (equation 23). (iii) Estimates of photospheric luminosity $L_{\rm ph}$ (equation 53). (iv) Estimates of photospheric temperature $k_{\rm B}T_{\rm ph}$ (equation 52). Rows (v) and (vi) display the estimated photospheric luminosity $\mathcal{L}_{\rm ph}$ and temperature $k_{\rm B}\mathcal{T}_{\rm ph}$ computed for the case in which $\eta < \eta_*$, assuming that the energy is released over a time-scale $\Delta t_{\rm spike} = 0.1$ s (equation 56). Finally, rows (vii) and (viii) show the initial luminosity $\mathcal{L}_0$ (equation 56) and temperature $k_{\rm B}\mathcal{T}_0$ also assuming that the energy is released over a time-scale $\Delta t_{\rm spike} = 0.1$ s. Note that the last two rows coincide with the photospheric values if $\eta > \eta_*$.

		C2	C3	B2	A2
(i)	$\mathcal{E}_{\rm r}$ (erg)	6.03×10^{46}	4.21×10^{46}	2.15×10^{46}	2.07×10^{46}
(ii)	L_0 (erg s^{-1})	2.78×10^{49}	4.00×10^{49}	7.05×10^{48}	1.46×10^{48}
(iii)	$L_{\rm ph}$ (erg s^{-1})	9.32×10^{47}	2.31×10^{47}	5.9×10^{47}	3.49×10^{47}
(iv)	$k_{\rm B}T_{\rm ph}$ (keV)	25	21	43	84
(v)	$\mathcal{L}_{\rm ph}$ (erg s^{-1})	2.60×10^{47}	2.31×10^{47}	1.84×10^{47}	1.82×10^{47}
(vi)	$k_{\rm B}\mathcal{T}_{\rm ph}$ (keV)	121	140	186	189
(vii)	$\mathcal{L}_0$ (erg s^{-1})	6.03×10^{47}	4.21×10^{47}	2.15×10^{47}	2.07×10^{47}
(viii)	$k_{\rm B}\mathcal{T}_0$ (keV)	281	257	217	215

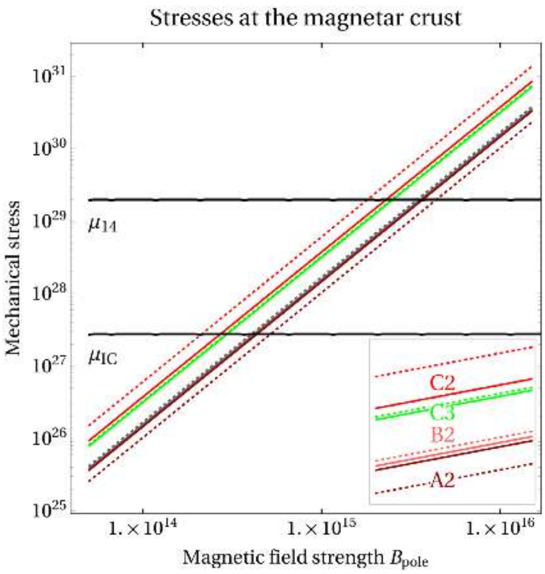

Figure 8. Mechanical stresses exerted on the magnetar crust (according to equation 32) for the maximum stresses (Tables 2 and 3) observed during the high resolution simulations of models A2, B2, C2, and C3. The stress component $\sigma^{r\varphi}$ is denoted by solid lines, the component $\sigma^{r\theta}$ by dotted lines. The colour coding corresponds to the initial models as introduced in Fig. 3. The black lines denote the approximate breaking stresses $\sim 0.2\mu_{\rm IC}$, and $\sim 0.2\mu_{14}$, at the inner crust boundary and near the core-crust transition, respectively. The high energy models reach the limit of a possible breaking of field lines for field strengths of $B_{\rm pole} \approx 10^{15}$–$10^{16}$ G.

$$\sigma^{r\theta} = \mu\, r \partial_r \left(\tfrac{\xi^\theta}{r}\right). \tag{34}$$

The transition at the anchoring point happens across a small distance, $h \equiv R_* - r_{\rm c}$, over which we can consider that $\mu = \mu_{\rm c}$ and σ^{ra} are constant. Integrating the stress tensor along this distance we obtain:

$$\xi^a_{\rm c} = r_{\rm c} \frac{\sigma^{ra}}{\mu_{\rm c}} \ln\left(\frac{r_{\rm c}}{r_{\rm c}+h}\right) \approx R_* \frac{\sigma^{ra}}{\mu_{\rm c}}, \tag{35}$$

for $h \ll r_{\rm c}$, R_*, and independent of the size of the transition layer, h. The radial force per unit volume induced by the applied stress is (Landau & Lifshitz 2012)

$$f^r = \sigma^{rk}{}_{;k} = \frac{1}{r}\left(\partial_\theta \sigma^{r\theta} + \cot\theta\, \sigma^{r\theta}\right) + \frac{1}{r\sin\theta}\partial_\varphi \sigma^{r\varphi}, \tag{36}$$

where we have considered that the only non-vanishing components are $\sigma^{r\theta}$ and $\sigma^{r\varphi}$. We can estimate the radial displacement ξ^r balancing this force with the gravitational force on the displaced mass, taken out of hydrostatic equilibrium. We can make an order of magnitude estimate using linear perturbation theory if one neglects terms including gradients of background quantities and perturbations of the gravitational potential. In that case, the force balance reads:

$$c_s^2\, \rho\, \partial_{rr}\xi^r \approx -f_r. \tag{37}$$

Integrating over the transition length h we get

$$\xi^r \approx -\frac{f_r\, h^2}{2\, c_s^2\, \rho} \approx -\frac{c^2_{\rm shear}}{c_s^2} h^2 s^{rk}{}_{;k}, \tag{38}$$

where $c^2_{\rm shear} \equiv \mu/\rho$ is the shear speed. For typical values in the crust one assumes $c^2_{\rm shear}/c_s^2 \sim 10^{-2}$. If we consider the maximum possible strain, i.e. the breaking strain, $s^{ij} \sim h s^{ij}{}_{;j} \sim 0.1$ (Horowitz & Kadau 2009), and the maximum possible value for $h \sim \Delta R \sim 1$ km, the size of the crust, one finds an upper limit for the radial displacement of $\xi^r_{\rm max} \sim 100$ cm. At the same time, the displacement components may be estimated directly from the results displayed in Fig. 8 by employing equation (35) and $\mu_{\rm c} = 0.5 \times (\mu_{14} + \mu_{\rm IC})$:

$$\xi^a_{\rm c} \approx 2.7 \times 10^4\ {\rm cm} \left(\frac{\sigma^{ra}}{10^{28}\ {\rm erg\ cm^{-3}}}\right)\left(\frac{B_{\rm pole}}{10^{15}\ {\rm G}}\right)^2\left(\frac{R_*}{13.7\ {\rm km}}\right). \tag{39}$$

Our results show that for typical magnetar field strengths ($B \gtrsim 10^{15}$ G) the instability is likely to break a large fraction of the crust down to the inner crust. For the largest magnetic fields ($B \gtrsim 10^{16}$ G) the stresses induced in the crust are sufficient to shatter the entire crust. We should mention that the three magnetars that have showed GFs are among the more magnetized known ones and all three exceed 5×10^{14} G.

6.3 Emission processes

6.3.1 Estimation of observational properties of the energy release

We have advanced that our models may release $\mathcal{E}_{\rm r} \approx 2.1 \times 10^{46} - 6.4 \times 10^{46}$ erg on time-scales of milliseconds, producing, therefore, dynamic luminosities $L_0 \sim (0.7 - 4) \times 10^{49}$ erg s^{-1} for the unstable models listed in Table 4. Following the reasoning of Thompson & Duncan (1995), confining this energy in the form of a photon-pair plasma by a closed magnetic flux loop of outer radius r requires that the field pressure at the outer boundary of the loop exceed the deposited energy density

$$\frac{B(r)^2}{8\pi} \gtrsim \frac{\mathcal{E}_\mathrm{r}}{4\pi/3(r^3 - R^3)}. \tag{40}$$

In a dominantly dipolar magnetosphere, $B(r) \sim B_\mathrm{pole}(R_*/r)^3$, the plasma can be confined if $\mathcal{E}_\mathrm{r} < \mathcal{E}_\mathrm{d}/2$ within a radius r in the range

$$\frac{\mathcal{E}_\mathrm{d}}{\mathcal{E}_\mathrm{r}}\left[1 - \sqrt{1 - 2\frac{\mathcal{E}_\mathrm{r}}{\mathcal{E}_\mathrm{d}}}\right] \lesssim \left(\frac{r}{R_*}\right)^3 \lesssim \frac{\mathcal{E}_\mathrm{d}}{\mathcal{E}_\mathrm{r}}\left[1 + \sqrt{1 - 2\frac{\mathcal{E}_\mathrm{r}}{\mathcal{E}_\mathrm{d}}}\right]. \tag{41}$$

For the range of values of $\mathcal{E}_\mathrm{r}/\mathcal{E}_\mathrm{d} \sim \Delta\mathcal{E}/\mathcal{E}_\mathrm{d}$ from our models (Table 3) we obtain that the size of the confinement region, $\Delta R \equiv r - R_*$, is limited by

$$(1.8 - 6) \times 10^{-2}\, R_* \lesssim \Delta R \lesssim (0.8 - 1.7)\, R_*. \tag{42}$$

Note that this result is independent of the magnetic field strength B_pole.

Our numerical simulations show that most of the energy is released in a thin and numerically unresolved surface current of the star, that we measure as a Poynting flux (see Fig. 5) and in a region close to the surface ($r \lesssim 1.25R_*$) with large currents (see Figs 6 and 7). Energy deposited there, essentially at the footprints of magnetic field lines, is expected to distribute efficiently along those lines aided by the flowing pair plasma. As a result, we expect that the energy will fill an extended region of the magnetosphere comparable in size to the region filled with currents (see Fig. 4). This region can be as large as $\sim 4R_*$ at the time of maximum energy dissipation. For magnetic field lines extending within the limits given by equation (42), the energy is expected to be confined. However, for lines extending beyond $(0.8{-}1.7)R_*$, the energy will not be confined and it may yield an ultrarelativistic fireball composed of pairs, photons, and a small amount of baryons lifted up from the outer crust by the large energy released there. Depending on the structure of the magnetosphere, the energy released in this form can be a significant fraction of $\mathcal{E}_\mathrm{r}$. Obviously, our methodology does not allow us to track the evolution of the released energy, but we may obtain a rough estimation of its bolometric properties. For the estimate we will consider that most of the energy is released in the fireball, which gives us upper limits.

The physics of such expanding fireball has been considered in many papers (e.g. Goodman 1986; Paczynski 1986; Shemi & Piran 1990; Meszaros, Laguna & Rees 1993; Piran, Shemi & Narayan 1993), especially addressing the generation of gamma-ray burst (GRBs), but also applied to SGRs (e.g. Nakar, Piran & Sari 2005). The sudden energy release results into a thermal burst carrying *most* of the initial energy, and according to the canonical interpretation (e.g. Hurley et al. 2005), with roughly the original temperature and a fraction of the energy in the form of relativistic pairs. The observed thermal spectrum of the flare and its temperature support this idea.

Here we follow the model of Mészáros & Rees (2000), which suffices for the basic estimates we aim at. Assuming that in a region of size $R_0 \simeq R_*$ (initially at rest), energy is released at a rate L_0, the initial temperature of the fireball in units of the electron rest mass is (Mészáros & Rees 2000, equation 2)

$$\Theta_0 = \left(\frac{k_\mathrm{B}}{m_e c^2}\right)\left(\frac{L_0}{4\pi R_0^2 c a_r}\right)^{1/4}$$

$$= 1.43 \left(\frac{L_0}{2.8 \times 10^{49}\,\mathrm{erg\,s^{-1}}}\right)^{1/4}\left(\frac{R_0}{1.37 \times 10^6\,\mathrm{cm}}\right)^{-1/2}, \tag{43}$$

where $m_e = 9.1095 \times 10^{-28}$ g is the electron mass, $a_r = 7.57 \times 10^{-15}$ g cm^{-1} s^{-2} K^{-4} is the radiation constant, and $k_\mathrm{B} \simeq 1.38 \times 10^{-16}$ erg K^{-1} is the Boltzmann constant. In the previous equation (and hereafter) we have scaled the luminosity to the dynamical luminosity estimated for model C2, but a similar exercise has been undertaken for models C3, B2, and A2, being the results listed in Table 4. The value of Θ_0 in equation (43) corresponds to a comoving temperature $k_\mathrm{B}T_0 \simeq 732$ keV. Starting from its initial radius, R_0, the fireball expands and accelerates until it converts most of its internal energy into kinetic energy at a distance R_s, commonly called the saturation radius (see equations 50 and 51 below). The Lorentz factor, Γ, of the expanding fireball is approximately given by

$$\Gamma = \begin{cases} \dfrac{r}{R_0} & \text{if } r < R_\mathrm{s}, \\ \dfrac{R_\mathrm{s}}{R_0} & \text{if } r \geq R_\mathrm{s}. \end{cases} \tag{44}$$

The amount of mass that may be unbound due to an energy release as large as suggested by our models ($\mathcal{E}_\mathrm{r}$) is uncertain, but we may estimate it to be as small as $M_\mathrm{ex} \simeq 3 \times 10^{-10}\,\mathrm{M}_\odot$. The period over which this mass is extracted we assume to be the same as that over which the energy is released, Δt_r. This implies a mass-loss rate from the magnetar surface $\dot{M} \simeq M_\mathrm{ex}/\Delta t_\mathrm{r} \simeq 2.8 \times 10^{26}\,\mathrm{g\,s^{-1}}$. The dimensionless entropy of the fireball for this baryon load is

$$\eta = \frac{L_0}{\dot{M}c^2}$$

$$\simeq 110 \left(\frac{L_0}{2.8 \times 10^{49}\,\mathrm{erg\,s^{-1}}}\right)\left(\frac{\dot{M}}{2.8 \times 10^{26}\,\mathrm{g\,s^{-1}}}\right)^{-1}. \tag{45}$$

As usual, we define the photospheric radius as the distance at which the fireball becomes optically thin, which may happen before the Lorentz factor saturates or after that, i.e. in the regime where the fireball coasts

$$R_\mathrm{ph} \simeq \frac{L_0 \sigma_\mathrm{T} Y}{4\pi m_\mathrm{p} c^3 \eta^3}, \qquad (R_\mathrm{ph} > R_\mathrm{s}) \tag{46}$$

$$R_\mathrm{ph} \simeq \left(\frac{L_0 \sigma_\mathrm{T} Y}{4\pi m_\mathrm{p} c^3 \eta}\right)^{1/3}. \qquad (R_\mathrm{ph} \leq R_\mathrm{s}) \tag{47}$$

Here, $\sigma_\mathrm{T} = 6.6525 \times 10^{-25}$ cm^2 and $m_\mathrm{p} = 1.6726 \times 10^{-24}$ g are the Thompson cross-section and the proton mass, respectively. Y represents the number of electrons per baryon. In the following, we will take $Y \simeq 1$, which is appropriate once pairs are not present in the system. Indeed, this shall be the case for radii larger than R_p (Mészáros & Rees 2000, equation 3)

$$R_\mathrm{p} = R_0 \frac{\Theta_0}{\Theta_p} \simeq 5.8 \times 10^7\,\mathrm{cm}$$

$$\times \left(\frac{L_0}{2.8 \times 10^{49}\,\mathrm{erg\,s^{-1}}}\right)^{1/4}\left(\frac{R_0}{1.37 \times 10^6\,\mathrm{cm}}\right)^{1/2}\left(\frac{\Theta_p}{0.03}\right)^{-1}, \tag{48}$$

where the comoving dimensionless temperature below which $e^\pm$ pairs drop out of equilibrium is $\Theta_p \simeq 0.03$ (equivalently, $k_\mathrm{B}T_p \simeq 17.4$ keV). Note that $R_p \ll R_\mathrm{ph}$ (see equations 46 and 47).

The critical baryon load, η_* for which the photospheric radius equals the saturation radius, i.e. $R_\mathrm{ph} = R_\mathrm{s}$, is given by (Mészáros & Rees 2000, equation 5)

$$\eta_* = \left(\frac{L_0 \sigma_\mathrm{T}}{4\pi m_p c^3 R_0}\right)^{1/4}$$

$$\simeq 393 \left(\frac{L_0}{2.8 \times 10^{49}\,\mathrm{erg\,s^{-1}}}\right)^{1/4}\left(\frac{R_0}{1.37 \times 10^6\,\mathrm{cm}}\right)^{-1/4}. \tag{49}$$

Depending on the value of the parameter η, there are two regimes. Either the photospheric radius happens beyond the saturation radius ($\eta < \eta_*$) or, otherwise, the saturation radius happens when the fireball is still expanding ($\eta > \eta_*$). In the former case, the saturation radius is

$$R_{\rm s} = \eta R_0 \simeq 1.5 \times 10^8\,{\rm cm} \left(\frac{R_0}{1.37 \times 10^6\,{\rm cm}} \right) \left(\frac{\eta}{110} \right), \qquad (50)$$

where we have used for η the value computed in equation (45) for the assumed value of $\dot{M}$. If the photosphere appears when the fireball is still accelerating, the saturation radius is attained at a distance (Mészáros & Rees 2000, equation 11)

$$R_{\rm s} = \eta_* R_0 \simeq 5.4 \times 10^8\,{\rm cm} \times \left(\frac{L_0}{2.8 \times 10^{49}\,{\rm erg\,s^{-1}}} \right)^{1/4} \left(\frac{R_0}{1.37 \times 10^6\,{\rm cm}} \right)^{3/4}, \qquad (51)$$

Interestingly, Hurley et al. (2005) model the peak of SGR 1806-20 assuming that the dimensionless entropy of the fireball is $\eta > \eta_*$ because for the observed peak luminosity (much smaller than that implied in our models, namely, $\sim 2 \times 10^{47}\,{\rm erg\,s^{-1}}$), the critical baryon load would be 3−4 times smaller than estimated in equation (49) and, hence, Hurley et al. (2005) naturally obtain $\eta \gtrsim \eta_*^{(1806-20)}$. The observational difference between the two described regimes is notable for our models as we see next in the estimation of the photospheric temperature and luminosity of the events. In the case $\eta < \eta_*$, the photospheric temperature and luminosity are, respectively,

$$k_{\rm B} T_{\rm ph} = k_{\rm B} T_0 \left(\frac{R_{\rm ph}}{R_{\rm s}} \right)^{-2/3} \simeq 25\,{\rm keV} \left(\frac{L_0}{2.8 \times 10^{49}\,{\rm erg\,s^{-1}}} \right)^{-5/12} \times \left(\frac{\eta}{110} \right)^{8/3} \left(\frac{R_0}{1.37 \times 10^6\,{\rm cm}} \right)^{-5/6}, \qquad (52)$$

and (Mészáros & Rees 2000, equation 9),

$$L_{\rm ph} = L_0 \left(\frac{R_{\rm ph}}{R_{\rm s}} \right)^{-2/3} \simeq 9.3 \times 10^{47}\,{\rm erg\,s^{-1}} \left(\frac{L_0}{2.8 \times 10^{49}\,{\rm erg\,s^{-1}}} \right)^{7/12} \times \left(\frac{\eta}{110} \right)^{8/3} \left(\frac{R_0}{1.37 \times 10^6\,{\rm cm}} \right)^{-5/6}. \qquad (53)$$

The value obtained in equation (52) must be compared with the ones obtained from observations, namely $k_{\rm B} T_{\rm peak}^{\rm obs} \simeq 175 - 250$ keV. Our result underestimates the observed temperature significantly. However, we are neglecting Comptonization effects, which may slightly raise the estimated photospheric temperature (still below the observational data). Note that smaller values of L_0, in line with the observed luminosities at peak for SGRs, would bring the observed photospheric temperature to the observed values, but, at the same time, they would significantly raise the photospheric luminosity, hence yielding events much more luminous than observed. The dependence on $\eta^{8/3}$ is the same in both equations (52) and (53), therefore, changes in the assumed baryon loading may not improve the consistency of our estimated photospheric values with the observed ones. However, if the baryon load is sufficiently small such that $\eta > \eta_*$ (as assumed in Hurley et al. 2005), the declining temperature and luminosity in the outflow are compensated by the relativistic blueshift. In this case, we would estimate the following photospheric temperature

$$k_{\rm B} T_{\rm ph} = k_{\rm B} T_0 \simeq 723\,{\rm keV} \left(\frac{L_0}{2.8 \times 10^{49}\,{\rm erg\,s^{-1}}} \right)^{1/4} \left(\frac{R_0}{1.37 \times 10^6\,{\rm cm}} \right)^{-1/2}, \qquad (54)$$

and luminosity

$$L_{\rm ph} = L_0 \simeq 2.8 \times 10^{49}\,{\rm erg\,s^{-1}}\,{\rm erg\,s^{-1}}. \qquad (55)$$

In this case, both estimations for $T_{\rm ph}$ and $L_{\rm ph}$ significantly overestimate the observed values for SGRs.

We have found in this section that independently of whether the photosphere of the expanding fireball happens in the acceleration phase or in the coasting phase, the estimated values of $T_{\rm ph}$ and $L_{\rm ph}$ are not compatible with observations. The root for the discrepancies found are the very large dynamic luminosities (L_0) of most of our models. These large values result from considering magnetospheric initial data where the twist is so large that they release a large amount of energy on time-scales of milliseconds. We note that models with larger relative toroidal fields (as induced by a power-index $\sigma = 2$, and $s = 2$) spanning a larger fraction of the magnetar surface (due to their smaller values of $P_{\rm c}$), e.g. model A2 (Table 4), show values of $T_{\rm ph}$ and $L_{\rm ph}$ broadly compatible with the most energetic GFs observed so far (see e.g. Hurley et al. 2005; Coti Zelati et al. 2018). This is in contrast to models where we have built up the magnetosphere with $s = \sigma = 1$ (namely, C2, C3, and B2), which systematically yield overluminous and too cold photospheric conditions. Thus, our results suggest that twisting magnetospheres to the largest (theoretical) levels we have considered here may not be realized in nature. Well before reaching the largest twists of models C3, C2, or B2 the dynamical instability may set in releasing smaller amounts of energy (and hence, producing smaller dynamical luminosities).

A potential handicap in our models is the duration of the observational signal that yield the fireballs modelled so far. In the canonical fireball model, the energy release leads to a *frozen pulse* whose duration approximately equals the time-scale over which the energy is deposited, $\Delta t_{\rm r}$ (e.g. Piran et al. 1993, but see Janka et al. 2006). Since $\Delta t_{\rm r} \ll \Delta t_{\rm spike}$, the quasi-thermal radiation bursts that we have estimated are too short to account for the typical time-scale of the initial spike of GFs in SGRs ($\Delta t_{\rm spike} \sim 0.1$ s). In our simulations, the energy change in the magnetosphere is driven by the Poynting flux through the star surface. However, the ability of the crust to absorb all this energy on the dynamical time-scale of the magnetosphere is limited because of the low transmission coefficient (see equation 5). So far we have considered that all this energy is temporarily stored in a thin layer above the magnetospheric surface, where intense currents may convert the stored magnetic energy into thermal energy. This is consistent with the boundary conditions imposed in our numerical simulations. Alternatively, we could have chosen boundary conditions that avoid the formation of strong thin surface currents (as e.g. in Carrasco & Reula 2016). In that case, Alfvén waves propagating towards the surface of the star get reflected and collide at some distance from the surface. This forces the formation of reconnection points at some distance from the neutron star surface. Li et al. (2019) have estimated that this process is relatively inefficient in dissipating the energy of the magnetosphere and that it may take multiple bounces in the magnetosphere to dissipate all the energy. This may allow for a slower energy deposition on time-scales $\sim \Delta t_{\rm spike}$.

Unfortunately, our numerical models do not include the relevant microphysics to fully address the conversion of magnetic into thermal energy. Thus, we can only warn the reader that the milliseconds time-scales over which we have made our (simple) estimations of the dynamical luminosity of the models at hand are only lower bounds of the true time-scales on which the released energy may leave the magnetosphere. Taking into account this caveat, the values of L_0 listed in Table 4 are upper bounds to the effective initial luminosity, $\mathcal{L}_0$,

$$\mathcal{L}_0 := \frac{\mathcal{E}_{\rm r}}{\Delta t_{\rm spike}} \simeq 10^{47}\,{\rm erg\,s^{-1}} \left(\frac{\mathcal{E}_{\rm r}}{10^{46}\,{\rm erg}}\right) \left(\frac{\Delta t_{\rm spike}}{0.1\,{\rm s}}\right)^{-1}. \qquad (56)$$

Redoing the previous estimations for the photospheric conditions, we find the values $\mathcal{L}_{\rm ph}$ and $k_{\rm B}\mathcal{T}_{\rm ph}$ listed in Table 4. In addition to these estimates of the photospheric luminosity and temperature corresponding to the values of the initial luminosity given by equation (56) when the photosphere happens beyond the saturation radius (i.e. for $\eta < \eta_*$), we also provide the estimation of the photospheric luminosity ($\mathcal{L}_0$) and temperature ($k_{\rm B}\mathcal{T}_0$) in the complementary case when the photospheric conditions are reached during the acceleration phase of the fireball (i.e. $\eta > \eta_*$). All these new values of the photospheric luminosity and temperature are perfectly compatible with observational data. Not surprisingly, we find that depending on whether we assume that photospheric conditions are met in the accelerating phase or in the coasting phase of the fireball, the values obtained for the photospheric temperature *bracket* the typical values found for the spike of SGRs.

6.3.2 *Optical depth of the magnetosphere*

The observed maximum current density throughout the magnetosphere, $J_{\rm max}$, can be quantified directly from the results shown in Tables 2 and 3 by employing the conversion formula

$$J_{\rm max} = 4.4926 \times 10^{12}\,{\rm A\,m^{-2}} \left(\frac{\tilde{J}_{\rm max}}{10^{-6}}\right) \left(\frac{B_{\rm pole}}{10^{15}\,{\rm G}}\right). \qquad (57)$$

The presented results compare well to the expected current density stated in equation (3). Close to the surface of the star, where the highest currents appear, the particle density is

$$n_e = \frac{J}{ce}\mathcal{M} \sim 10^{19}\,{\rm cm^{-3}}, \qquad (58)$$

where $\mathcal{M}$ is the multiplicity. Beloborodov (2013b) has estimated that in extended regions close to the poles the multiplicity can be as large as $\mathcal{M} \sim 100$, while close to the equator $\mathcal{M} \sim 1$.

The dominant contribution to the opacity in the magnetosphere is the resonant cyclotron scattering of thermal photons off charge particles in the vicinity of the neutron star.[6] Thompson, Lyutikov & Kulkarni (2002) have estimated that for twists of $\Delta\varphi \sim 1$ the typical optical depth in the magnetosphere is ∼1. In general, computing the optical depth for magnetar magnetospheres is a complicated problem, because one needs a self-consistent solution of the photon field and the momentum distribution of charged particles travelling along the magnetic field lines (see Beloborodov 2013b). In this work we make an estimation for radially streaming photons and a simplified momentum distribution of charged particles. We only consider 1 keV photons, which are typical for the observed surface temperature in magnetars. Inspired by Beloborodov (2013b) we use a simple waterbag momentum distribution (see Appendix B) which is characterized by two parameters, the mean specific momentum ($\bar{p}$, where $p = vW$) and $\mathcal{M}$. We integrate the optical depth (τ) radially inwards (see appendix B, equation B1 for details on the computation) and identify the photosphere as the place where $\tau = 1$.

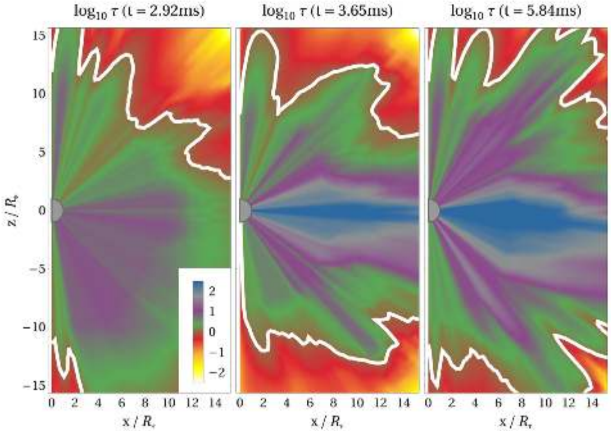

Figure 9. Snapshots of the logarithm of the optical thickness during the evolution of the high resolution version of model C2. The logarithm of the optical thickness for the $\{\mathcal{M} = 100, \gamma = 30\}$ model is displayed by the colour scale, the photosphere ($\tau = 1$) is displayed as a white solid line. See appendix B for further details.

Fig. 9 shows estimates for the optical thickness of the magnetosphere at three different times (during and after the rapid drop of magnetospheric energy) computed with parameters $\{\mathcal{M} = 100, \gamma = 30\}$. During the rearrangement of the magnetosphere, the coronal region along the equator becomes optically thick. The initial configuration is optically thin and, hence, not shown here. An important conclusion is that close to the critical point, most of the magnetosphere, if not all, is optically thin, which gives rise to a blackbody spectrum with the typical temperature of the NS surface (∼ 1 keV) plus a possible non-thermal contribution of up-scattered photons. However, during the instability, the increase of the magnetospheric currents, makes a large fraction of the magnetosphere of a few stellar radii optically thick. This region is filled up with pair plasma and will emit thermal radiation through its photosphere. Its lifetime is related to the presence of strong currents in the magnetosphere and may be an explanation for the X-ray tail ($k_{\rm B}T \sim 30\,$keV) observed after GFs and lasting for a few 100 s. We note that only a relatively small fraction of the total energy released in the magnetosphere by the instability may contribute to the tail, while most of it may contribute to the initial peak characteristic of GFs (see discussion in Section 6.3.1).

Our model to compute the magnetospheric optical thickness for resonant cyclotron scattering assumes uniform values of the multiplicity and of the electron Lorentz factor. Neither for the multiplicity (as we have argued above) nor for γ this is completely correct. Modelling locally the values of the parameters $\{\mathcal{M}, \gamma\}$ is beyond the scope of this paper. However, we may test the robustness of our results by exploring the parameter space determined by $\mathcal{M}$ and γ. In Fig. 10, we display the time evolution of the optical thickness at the equator of the magnetar for various parameter sets. As expected, the larger the value of $\mathcal{M}$, the larger the number density of leptons and, consistently, the larger the opacity (note the

---

[6] If there is a dynamical mass ejection a result of the large energy release close to the magnetar surface (Section 6.3.1), the Thompson scattering (in the expanding fireball) may be the dominant source of opacity at sufficiently large distances.

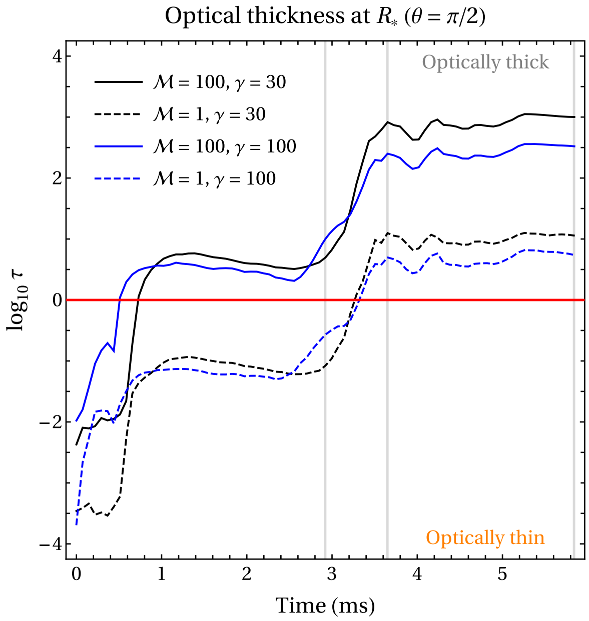

Figure 10. Evolution of the optical thickness at the equator of the magnetar ($\theta = \pi/2$, $r = R_*$) of the high resolution initial model C2 ($B_{\rm pole} = 10^{15}$ G) and various parameter sets (see legends). The modelled optical thickness depends sensitively on the chosen multiplicity $\mathcal{M}$ and Lorentz factor γ (cf. Section 6.3.2 and Appendix B). Initially, the magnetosphere is optically thin. During the onset of the instability (∼1 ms) charges are produced in the magnetosphere. Depending on the chosen model ($\mathcal{M}$, γ), the magnetosphere becomes optically thick at different times. The times used for the visualization of optical thickness in Fig. 9 are denoted in grey lines.

nearly two orders of magnitude difference between the solid lines with $\mathcal{M} = 100$ and the dashed lines with $\mathcal{M} = 1$). The effect of the variation of the Lorentz factor (electrons or positrons) is small compared to the strong impact of $\mathcal{M}$ on the opacity. Although the magnetosphere becomes eventually optically thick for all the parameter sets under investigation, models with $\mathcal{M} = 100$ develop regions with $\tau > 1$ very early ($t \lesssim 0.8$ ms), while models computed with $\mathcal{M} = 1$ become optically thick only when the instability in the magnetosphere fully develops.

Emission by resonant scattering in magnetar magnetospheres may be subject to ($\perp$ or $\shortparallel$) polarization (see e.g. Fernández & Davis 2011; Beloborodov 2013b). In the presented (approximate) modelling of optical thickness, however, we have found differences in these polarization states of < 1 per cent. We will further explore the emission properties of force-free twisted magnetospheres on suitable high-resolution numerical data in our future work.

7 CONCLUSIONS

In this work, we explore the stability properties of force-free equilibrium configurations of magnetar magnetospheres by performing numerical simulations of a selection of the models computed in Akgün et al. (2018a). For the case of degenerate magnetospheres (i.e. the same boundary conditions but different energies) we validate the hypothesis of Akgün et al. (2018a) that configurations in the high-energy branches are unstable while those in the lowest energy branch are stable. This confirms the existence of an unstable branch of twisted magnetospheres. It also allows us to formulate an instability criterion for the sequences of models computed in Akgün et al. (2018a). Our results are consistent with an interesting scenario where bursts and GFs in magnetars are triggered without involving crustal failures. The twist that is naturally produced in the magnetosphere by the Hall evolution of the crust (Akgün et al. 2017) can lead to unstable configurations that will release up to a 10 per cent of the energy stored in the magnetosphere, sufficient to explain the observations.

Akgün et al. (2017) have shown that the magnetothermal evolution of the crust leads naturally to configurations close to the instability threshold. However, the amount of energy released depends on how far away from the stable branch can the evolution drive the configuration. This is essentially a problem of comparing the evolution time-scale and the instability time-scale. For the models studied in this work the instability time-scale is of the order of milliseconds, much shorter than the magnetothermal evolution time-scales of the object (see Section 2.2.1). However, close to the critical point, the growth rate of the instability could be significantly smaller (actually, it should be zero at the critical point) which would allow us to overshoot the instability threshold. Note that, since the energy reservoir is large (∼10^{46} erg), even a very small fraction of energy release could explain many of the phenomenology of magnetars. Alternatively, there could be phenomena leading to fast dynamics in the crust such as sustained episodes of accelerated plastic flows triggered by the magnetic stresses in the crust (Lander & Gourgouliatos 2019).

For the unstable models, we observe the development of almost axisymmetric instabilities on a time-scale of a few ms rearranging the magnetic field to a configuration similar to those in the (stable) lower energy branch. The energy of the magnetosphere also decreases towards the value of the stable configuration. Differences with respect to the corresponding stable configuration can be attributed to the influence of the non-preservation of the force-free constraints (10) and (11). Using (much) larger numerical resolution (beyond the scope of our computational resources) we envision that the violation of the force-free constraints would be significantly reduced and the expected (low-energy) states would be the endpoint of the evolution after a full relaxation of the magnetosphere takes place. The energy decrease is explained, mainly, by a flow of energy towards the surface of the star, where it is dissipated efficiently. A large fraction of this energy is also dissipated in the magnetosphere at locations where the force-free conditions break. This contrasts with the work of Beloborodov (2011), Parfrey et al. (2013), and Carrasco et al. (2019) in which most of the energy is dissipated by the formation and ejection of plasmoids. The different setup used in these works (dynamically twisting versus unstable equilibrium configurations) makes a direct comparison difficult. A possible source for the qualitative discrepancy may be differences in the boundary condition at the surface of the star. While we use a boundary condition that dissipates very efficiently any strong currents formed at the surface, in their work, their use of essentially non-dissipative boundary conditions make the surface perfectly reflective. For the future it would be interesting to compare more closely the differences in the boundary condition and to develop a better physical model for dissipation at the NS surface.

The magnetic field remains nearly axisymmetric throughout the simulation indicating that the instability is mostly an $m = 0$ instability. A complete theoretical analysis of the origin of the instability and its properties is beyond the scope of this paper. However, we anticipate that such analysis has to be carried out on a global scale either by calculating the eigenmodes or by using the so-called energy principle of Bernstein et al. (1958) and is not trivial due to the presence of both poloidal and toroidal components (Akgün et al. 2013, and references therein). However, we note that, since the poloidal field structure changes somewhat less than the toroidal field, this instability could be compared to the interchange instability discussed by Tayler (1973), where displacing the toroidal field radially decreases the energy (even in the absence of a fluid).

We have made a crude estimation of the observational properties of the energy liberated in the magnetosphere as a result of the instability. The fact that large amounts of energy (in excess of 10^{46} erg) are released on milliseconds time-scales results in dynamical luminosities significantly larger than 10^{48} erg s^{-1} (reaching in some models 4×10^{49} erg s^{-1}). This should trigger the expansion of a pair-photon fireball polluted with baryons unbound from the magnetar crust. The bolometric signature of these fireballs seems incompatible with the observations of the initial spikes observed in GFs. With our simple analytic model, most of the unstable magnetospheres produce overluminous, too cool, and excessively short flashes. However, this problem can be solved if the energy can be liberated on longer time-scales, of the order of the observed GF spikes ($\Delta t_{\rm spike} \sim 0.1$ s). This could be possible in a scenario of slow energy dissipation as the one proposed by Li et al. (2019), which we plan to explore in the future.

The currents produced during the instability increase significantly the amount of pairs in the magnetosphere, a large fraction of which, of size $\sim 10 R_*$, becomes optically thick. The hot plasma magnetically confined in this region could be responsible for the extended thermal X-ray emission lasting for 50−300 s after GFs.

Our force-free numerical method cannot properly deal with the evolution of extremely thin surface currents. Therefore, the dynamical millisecond time-scales computed in our models should be taken as a lower bound for the physical time-scales. The magnetic dissipation taking place at these locations can be due to, e.g. Ohmic processes or to non-linear Alfvén wave interactions. Assuming that energy is released on $\sim \Delta t_{\rm spike}$, our estimate of the electromagnetic signature yields photospheric luminosities and temperatures compatible with observational data. Since this is a sound physical assumption, we conclude that observed GFs in SGRs are broadly compatible with the development of instabilities in twisted magnetospheres.

ACKNOWLEDGEMENTS

We thank Amir Levinson for his support in challenging our numerical code prior to the production of the presented results. We also thank Oscar Reula, Federico Carrasco, and Carlos Palenzuela for their valuable discussions on the boundary conditions. We acknowledge the support from the grants AYA2015-66899-C2-1-P and PROMETEO-II-2014-069. JM acknowledges a Ph.D. grant of the *Studienstiftung des Deutschen Volkes*. PC acknowledges the Ramon y Cajal funding (RYC-2015-19074) supporting his research. We acknowledge the partial support of the PHAROS COST Action CA16214 and GWverse COST Action CA16104. The shown numerical simulations have been conducted on *MareNostrum 4* of the *Red Española de Supercomputación* (AECT-2019-1-0004) as well as on the computational infrastructure of the University of Valencia. We thank the *EWASS 2019* and *GR22* conferences for the possibility to disseminate the results of this work.

REFERENCES

Akgün T., Reisenegger A., Mastrano A., Marchant P., 2013, MNRAS, 433, 2445
Akgün T., Miralles J. A., Pons J. A., Cerdá-Durán P., 2016, MNRAS, 462, 1894
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2017, MNRAS, 472, 3914
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2018a, MNRAS, 474, 625
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2018b, MNRAS, 481, 5331
Alic D., Moesta P., Rezzolla L., Zanotti O., Jaramillo J. L., 2012, ApJ, 754, 36
Baiko D. A., Chugunov A. I., 2018, MNRAS, 480, 5511
Baumgarte T. W., Montero P. J., Cordero-Carrión I., Müller E., 2013, Phys. Rev. D, 87, 044026
Beloborodov A. M., 2009, ApJ, 703, 1044
Beloborodov A. M., 2011, Astrophys. Space Sci. Proc., 21, 299
Beloborodov A. M., 2013a, ApJ, 762, 13.
Beloborodov A. M., 2013b, ApJ, 777, 114
Beloborodov A. M., Levin Y., 2014, ApJ, 794, L24
Beloborodov A. M., Thompson C., 2007, ApJ, 657, 967
Bernstein I. B., Frieman E. A., Kruskal M. D., Kulsrud R. M., 1958, Proc. R. Soc. A, 244, 17
Beskin V. S., 2010, MHD Flows in Compact Astrophysical Objects. Springer-Verlag, Heidelberg
Camenzind M., 2007, Compact Objects in Astrophysics: White Dwarfs, Neutron Stars, and Black Holes. Springer, Heidelberg
Carrasco F., Viganò D., Palenzuela C., Pons J. A., 2019, MNRAS, 484, L124
Carrasco F. L., Reula O. A., 2016, Phys. Rev. D, 93, 085013
Chugunov A. I., Horowitz C. J., 2010, MNRAS, 407, L54
Cline T. L. et al., 1980, ApJ, 237, L1
Collins D. C., Xu H., Norman M. L., Li H., Li S., 2010, ApJS, 186, 308
Coti Zelati F., Rea N., Pons J. A., Campana S., Esposito P., 2018, MNRAS, 474, 961
Dedner A., Kemm F., Kröner D., Munz C.-D., Schnitzer T., Wesenberg M., 2002, J. Comput. Phys., 175, 645
Elenbaas C., Watts A. L., Turolla R., Heyl J. S., 2016, MNRAS, 456, 3282
Etienne Z. B., Wan M.-B., Babiuc M. C., McWilliams S. T., Choudhary A., 2017, Class. Quantum Gravity, 34, 215001
Fernández R., Davis S. W., 2011, ApJ, 730, 131
Fujisawa K., Kisaka S., 2014, MNRAS, 445, 2777
Gabler M., Cerdá-Durán P., Stergioulas N., Font J. A., Müller E., 2012, MNRAS, 421, 2054
Gill R., Heyl J. S., 2010, MNRAS, 407, 1926
Glampedakis K., Lander S. K., Andersson N., 2014, MNRAS, 437, 2
Goldreich P., Julian W. H., 1969, ApJ, 157, 869
Goldreich P., Reisenegger A., 1992, ApJ, 395, 250
Goodale T., Allen G., Lanfermann G., Massó J., Radke T., Seidel E., Shalf J., 2003, in Vector and Parallel Processing – VECPAR'2002, 5th International Conference, Lecture Notes in Computer Science. Springer, Berlin, Available at: http://edoc.mpg.de/3341
Goodman J., 1986, ApJ, 308, L47
Gourgouliatos K. N., Wood T. S., Hollerbach R., 2016, Proc. Natl. Acad. Sci., 113, 3944
Grad H., Rubin H., 1958, in Proceedings of the Second United Nations Conference on the Peaceful Uses of Atomic Energy, Vol. 31. UN, Geneva, p. 190
Hascoët R., Beloborodov A. M., den Hartog P. R., 2014, ApJ, 786, L1
Horowitz C. J., Kadau K., 2009, Phys. Rev. Lett., 102, 191102
Hurley K. et al., 1999, Nature, 397, 41
Hurley K. et al., 2005, Nature, 434, 1098
Janka H. T., Aloy M. A., Mazzali P. A., Pian E., 2006, ApJ, 645, 1305
Jones P. B., 1988, MNRAS, 233, 875
Kaspi V. M., Beloborodov A. M., 2017, ARA&A, 55, 261
Kojima Y., 2017, MNRAS, 468, 2011
Kojima Y., 2018, MNRAS, 477, 3530
Kojima Y., Okamoto S., 2018, MNRAS, 475, 5290
Komissarov S. S., 2004, MNRAS, 350, 427
Komissarov S. S., 2011, MNRAS, 418, L94
Landau L. D., Lifshitz E. M., 2012, Theory of Elasticity. Elsevier
Lander S. K., Gourgouliatos K. N., 2019, MNRAS, 486, 4130

Lander S. K., Andersson N., Antonopoulou D., Watts A. L., 2015, MNRAS, 449, 2047
Levin Y., Lyutikov M., 2012, MNRAS, 427, 1574
Li X., Beloborodov A. M., 2015, ApJ, 815, 25
Li X., Levin Y., Beloborodov A. M., 2016, ApJ, 833, 189
Li X., Zrake J., Beloborodov A. M., 2019, ApJ, 881, 13
Link B., 2014, MNRAS, 441, 2676
Löffler F. et al., 2012, Class. Quantum Gravity, 29, 115001
Lüst R., Schlüter A., 1954, Z. Astrophys., 34, 263
Lyutikov M., 2003, MNRAS, 346, 540
Lyutikov M., 2015, MNRAS, 447, 1407
McKinney J. C., 2006, MNRAS, 367, 1797
Mereghetti S., Pons J. A., Melatos A., 2015, Space Sci. Rev., 191, 315
Mészáros P., Rees M. J., 2000, ApJ, 530, 292
Meszaros P., Laguna P., Rees M. J., 1993, ApJ, 415, 181
Mignone A., Tzeferacos P., 2010, J. Comput. Phys., 229, 2117
Mikic Z., Linker J. A., 1994, ApJ, 430, 898
Miranda-Aranguren S., Aloy M. A., Rembiasz T., 2018, MNRAS, 476, 3837
Montero P. J., Baumgarte T. W., Müller E., 2014, Phys. Rev. D, 89, 084043
Nakar E., Piran T., Sari R., 2005, ApJ, 635, 516
Paczynski B., 1986, ApJ, 308, L43
Palenzuela C., Lehner L., Reula O., Rezzolla L., 2009, MNRAS, 394, 1727
Palenzuela C., Garrett T., Lehner L., Liebling S. L., 2010, Phys. Rev. D, 82, 044045
Parfrey K., Beloborodov A. M., Hui L., 2012, ApJ, 754, L12
Parfrey K., Beloborodov A. M., Hui L., 2013, ApJ, 774, 92
Parfrey K., Spitkovsky A., Beloborodov A. M., 2017, MNRAS, 469, 3656
Paschalidis V., Shapiro S. L., 2013, Phys. Rev. D, 88, 104031
Perna R., Pons J. A., 2011, ApJ, 727, L51
Pili A. G., Bucciantini N., Del Zanna L., 2015, MNRAS, 447, 2821
Piran T., Shemi A., Narayan R., 1993, MNRAS, 263, 861
Pons J. A., Geppert U., 2007, A&A, 470, 303
Pons J. A., Miralles J. A., Geppert U., 2009, A&A, 496, 207
Rea N., Esposito P., 2011, Astrophys. Space Sci. Proc., 21, 247
Roumeliotis G., Sturrock P. A., Antiochos S. K., 1994, ApJ, 423, 847
Shafranov V., 1966, Rev. Plasma Phys., 2, 103
Shemi A., Piran T., 1990, ApJ, 365, L55
Shibata M., 2015, Numerical Relativity. 100 Years of General Relativity. World Scientific Publishing Company, Singapore
Spitkovsky A., 2006, ApJ, 648, L51
Steiner A. W., Watts A. L., 2009, Phys. Rev. Lett., 103, 181101
Suresh A., Huynh H. T., 1997, J. Comput. Phys., 136, 83
Tayler R. J., 1973, MNRAS, 161, 365
Thompson C., Duncan R. C., 1995, MNRAS, 275, 255
Thompson C., Duncan R. C., 1996, ApJ, 473, 322
Thompson C., Duncan R. C., 2001, ApJ, 561, 980
Thompson C., Lyutikov M., Kulkarni S. R., 2002, ApJ, 574, 332
Thompson C., Yang H., Ortiz N., 2017, ApJ, 841, 54
Turolla R., Zane S., Watts A. L., 2015, Rep. Prog. Phys., 78, 116901
Uchida T., 1997, Phys. Rev. E, 56, 2181–2197
Viganò D., Pons J. A., Miralles J. A., 2011, A&A, 533, A125
Viganò D., Pons J. A., Miralles J. A., 2012, Comput. Phys. Commun., 183, 2042
Viganò D., Rea N., Pons J. A., Perna R., Aguilera D. N., Miralles J. A., 2013, MNRAS, 434, 123
Wald R., 2010, General Relativity. University of Chicago Press, Chicago
Woods P. M., Thompson C., 2006, Soft Gamma Repeaters and Anomalous X-Ray Pulsars: Magnetar Candidates. Cambridge University Press, Cambridge, UK, p. 547

APPENDIX A: NUMERICAL DETAILS

A1 The augmented system

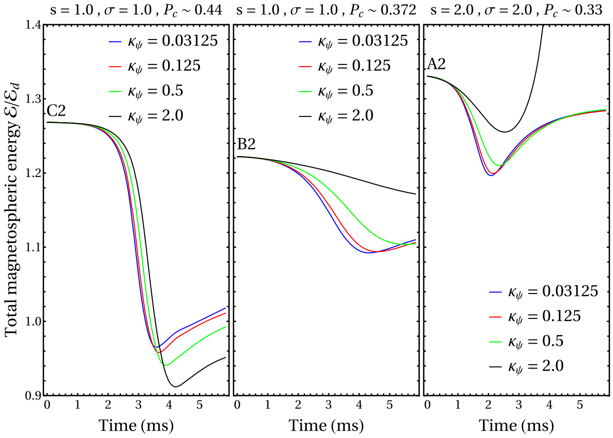

Figure A1. Energy evolution of the high energy initial data models A2, B2, and C2 using different damping constants κ_ψ (divergence cleaning) in a low resolution study (16 points per R_*). While one observes a converging evolution for the lower cleaning potentials $\kappa_\psi = 0.03125$ and $\kappa_\psi = 0.125$, the energy evolution shows a strong (non-physical) dependence on κ_ψ for larger damping constants. This effect is amplified in the high resolution (32 points per R_*).

In order to preserve the physical conditions div$\tilde{\boldsymbol{B}} = 0$ and div$\tilde{\boldsymbol{E}} = \tilde{\rho}_e$ we make use of hyperbolic/parabolic cleaning potentials (Dedner et al. 2002; Palenzuela et al. 2009; Mignone & Tzeferacos 2010). Specifically, we implement an augmented system of Maxwell's equations as follows (Palenzuela et al. 2009; Miranda-Aranguren, Aloy & Rembiasz 2018):

$$\partial_t \phi - \partial_i \tilde{E}^i = -\tilde{\rho}_e - \kappa_\phi \phi \tag{A1}$$

$$\partial_t \tilde{E}^i - \partial_j \left(\epsilon^{ijk} \tilde{B}_k + \delta^{ij} \phi\right) = -\tilde{J}^i_{\mathrm{FF}} \tag{A2}$$

$$\partial_t \psi + \partial_i \tilde{B}^i = -\kappa_\psi \psi \tag{A3}$$

$$\partial_t \tilde{B}^i + \partial_j \left(\epsilon^{ijk} \tilde{E}_k + \delta^{ij} \psi\right) = 0. \tag{A4}$$

Here, ψ (divergence cleaning) and ϕ (charge conservation) are the scalar potentials, κ_ϕ and κ_ψ the respective damping constants and δ^{ij} denotes the Kronecker delta. As for the practical implementation, we follow a Strang splitting approach (as employed e.g. in Komissarov 2004), effectively solving part of the scalar equations (A1) and (A3) analytically. Prior (before `MoL_Step`) and after (before `MoL_PostStep`) the time integration of the EINSTEIN TOOLKIT thorn `MoL` we evolve in time the equations

$$\phi\,(t) = \phi_0 \exp\left[-\kappa_\phi t\right], \tag{A5}$$

$$\psi\,(t) = \psi_0 \exp\left[-\kappa_\psi t\right], \tag{A6}$$

for a time $t = \Delta t/2$. The coefficients κ_ϕ and κ_ψ have to be chosen by optimization in accordance with the grid properties.

We find it beneficial to choose a large value for κ_ϕ, effectively dissipating charge conservation errors on very short time-scales. As for the divergence cleaning, we conducted a series of tests, optimizing κ_ψ to yield stable and converging evolution for all shown resolutions, ultimately resorting to $\kappa_\psi = 0.125$ (see Fig. A1 for a review of the optimization process).

It should be noted at this point that Mignone & Tzeferacos (2010) present a promising scheme of choosing κ_ψ according to the grid resolution that has also been used in Miranda-Aranguren et al. (2018). In the framework of mesh refinement of the Einstein Toolkit, this would result in a different damping of the cleaning potentials across the refinement levels. We have found that the optimization of the hyperbolic/parabolic cleaning becomes a very subtle issue and may experience strong numerical effects when increasing the overall resolution. This observation may, however, be an artefact of the fixed boundary of the magnetar surface – which on a Cartesian grid, resembles an accumulation of boxes rather than a perfectly aligned spherical boundary. The exploration of these effects and the transition to a fully spherical version of this force-free thorn (as introduced in Baumgarte et al. 2013; Montero, Baumgarte & Müller 2014) will be a subject of future efforts.

A2 Conservation of force-free constraints

FFE codes are valid in the limit of high electromagnetic energy compared to the rest mass and thermal energy of the respective plasma. The dynamics of force-free fields is described entirely without the plasma four velocity. However, demanding the existence of a physical, time-like velocity field $\boldsymbol{u}$ with $F_{\mu\nu}u^{\nu} = 0$, as well as the degeneracy condition $F_{\mu\nu}J^{\nu} = 0$ (see Uchida 1997, for a detailed algebraic review) one is left with the aforementioned constraints:

$$\tilde{\boldsymbol{E}} \cdot \tilde{\boldsymbol{B}} = 0 \tag{A7}$$

$$\tilde{\boldsymbol{B}}^2 - \tilde{\boldsymbol{E}}^2 \geq 0. \tag{A8}$$

Within the shown simulations we find it beneficial to employ an approach presented in Komissarov (2011) and Parfrey et al. (2017) in order to archive $\partial_t \left(\tilde{\boldsymbol{E}} \cdot \tilde{\boldsymbol{B}}\right) = 0$ throughout the evolution (by making use of the force-free current as in equation 12) without the employment of target currents (as discussed in Parfrey et al. 2017). Additionally, we include a suitable Ohm's law (Komissarov 2004, section C3) into our Strang splitting approach aiming towards an evolution minimizing the violation of conditions (A7), and (A8).

In order to build up a force-free current, Komissarov (2004) introduces a generalized Ohm's law in the context of FFE:

$$\tilde{\boldsymbol{J}} = \sigma_{\|}\tilde{\boldsymbol{E}}_{\|} + \sigma_{\perp}\tilde{\boldsymbol{E}}_{\perp} + \tilde{\boldsymbol{j}}_{d}, \tag{A9}$$

where the subscripts $\|$ and $\perp$ denote the components parallel and perpendicular to the magnetic field, $\tilde{\boldsymbol{B}}$. A to be specified model for σ introduces a suitable resistivity into the force-free system (see also Lyutikov 2003, for further comments on resistive FFE), while $\tilde{\boldsymbol{j}}_d$ is the drift current perpendicular to the electric and magnetic fields. In its general form, (A9) plays the central role in ensuring the force-free conditions (A7) and (A8). Komissarov (2004) suggests a resistivity model that depends on the time-step of the evolution Δt (throughout the presented simulations we employ CFL =0.2), where

$$\sigma_{\|} = \frac{d}{\Delta t}. \tag{A10}$$

The cross-field resistivity $\sigma_{\perp}$ is strongly linked to the violation of condition (A8),

$$\sigma_{\perp} = \begin{cases} 0 & : \boldsymbol{B}^2 \geq \boldsymbol{E}^2 \\ b\dfrac{\left(\tilde{E}_{\perp} - \tilde{E}_{\perp}^{*}\right)}{\tilde{E}_{\perp}^{*}} & : \tilde{\boldsymbol{B}}^2 < \tilde{\boldsymbol{E}}^2 \end{cases}, \tag{A11}$$

where $\tilde{E}_{\perp} = \left|\tilde{\boldsymbol{E}}_{\perp}\right|$ and $\left(\tilde{E}_{\perp}^{*}\right)^2 = \left(\tilde{\boldsymbol{B}} - \tilde{\boldsymbol{E}}_{\|}\right)^2$ and b is an scalar parameter controlling the magnitude of $\sigma_{\perp}$. Equations (A10) and (A11) have a pair of analytic solutions:

$$\tilde{\boldsymbol{E}}_{\|}(t) = \tilde{\boldsymbol{E}}_{\|}(0) \times e^{-\sigma_{\|}t} \tag{A12}$$

$$\tilde{\boldsymbol{E}}_{\perp}(t) = \left[\tilde{E}_{\perp}^{*}(0) + \frac{\tilde{E}_{\perp}^{*}(0)\left[\tilde{E}_{\perp}(0) - \tilde{E}_{\perp}^{*}(0)\right] \times e^{-b\sigma_{\|}t}}{\tilde{E}_{\perp}(0) - \left[\tilde{E}_{\perp}(0) - \tilde{E}_{\perp}^{*}(0)\right] \times e^{-b\sigma_{\|}t}}\right] \times \frac{\tilde{\boldsymbol{E}}_{\perp}(0)}{\tilde{E}_{\perp}(0)}. \tag{A13}$$

During our numerical simulations, we usually choose $d = 5.0$, and $b = 0.1$, and solve equation (A12) prior to equation (A13) in a Strang splitting scheme in direct analogy to the implementation described in Section A1. This resistivity model ensures the validity of the force-free regime throughout time, in other words, the evolution is driven towards a force-free state

$$\begin{aligned} &\tilde{\boldsymbol{E}} \cdot \tilde{\boldsymbol{B}} \rightarrow 0 \\ &\tilde{\boldsymbol{B}}^2 - \tilde{\boldsymbol{E}}^2 \rightarrow 0 \qquad : \tilde{\boldsymbol{B}}^2 < \tilde{\boldsymbol{E}}^2. \end{aligned} \tag{A14}$$

APPENDIX B: OPTICAL DEPTH TO RESONANT CYCLOTRON SCATTERING

For the presented modelling of the optical thickness of highly magnetized force-free plasmas around magnetars (see Section 6.3), we adapt the techniques describing resonant scattering as presented by Beloborodov (2013b) (from now on Be13). In the following, we will give a short review of the underlying equations. In order to derive the optical thickness τ, we integrate equation (Be13/A15),

$$\frac{\mathrm{d}\tau}{\mathrm{d}s} = 2\pi^2 r_e \frac{c}{\omega}\frac{\xi}{|\tilde{\mu}|} n_e \left[f_e\left(p_1\right) + f_e\left(p_2\right)\right]. \tag{B1}$$

Here, $r_e = e^2/m_e c^2$ denotes the photon wavelength, ω the frequency of the seed photon (we consider 1 keV photons), and $\xi = 1$ or $\xi = \tilde{\mu}^2$ depending on the photon polarization ($\perp$ or $\|$, respectively). The relativistic particles require the specification of the quantities $\mu = \cos\vartheta$ and $\tilde{\mu} = \cos\tilde{\vartheta}$, where ϑ is the angle between the photon path and the magnetic field $\boldsymbol{B}$ in the lab frame and $\tilde{\vartheta}$ in the rest frame of the electron. The dimensionless momenta $p_{1,2}$ correspond to the electron (or positron) velocities favoured by the resonant scattering model. As both polarizations yield similar results, we only consider the slightly dominant $\perp$ orientation for our model. Beloborodov (2013b) estimated that the contribution of non-resonant scattering to the optical depth is negligible and will not be considered in our calculations (see, however, footnote 6).

Following Be13, we employ the so-called waterbag model as a distribution function for electron (or positron) momenta. In analogy to a two-fluid model, the distribution function is characterized by the two parameters (dimensionless momenta) p_+ and p_-, with the overall shape

$$f_e(p) = \begin{cases} \left(p_+ - p_-\right)^{-1} & : p_- < p < p_+ \\ 0 & : \text{else} \end{cases}. \tag{B2}$$

Applying the waterbag model (B2) in equation (B1) selects the relevant electron (or positron) momenta for the scattering process. The distribution of this normalization factor throughout the magnetosphere especially depends on the flow direction of charges along $\boldsymbol{B}$. As described in Section 5.2 of Be13, we adjust their model according to a flow of electrons (or positrons) which turns back to the central object when field lines cross the equator. We apply this to all field lines crossing regions with $B < 10^{13}$ G (this holds everywhere except in the inner coronal region of strong closed magnetic field lines).

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