arXiv:2212.00765v2 [astro-ph.HE] 1 Apr 2023

Multi-collision internal shock lepto-hadronic models for energetic GRBs

Annika Rudolph,[1,2] Maria Petropoulou,[3,4] Željka Bošnjak,[5] and Walter Winter[2]

[1]*Niels Bohr International Academy and DARK, Niels Bohr Institute, University of Copenhagen*
Blegdamsvej 17, 2100, Copenhagen, Denmark
[2]*Deutsches Elektronen-Synchrotron DESY*
Platanenallee 6, 15738 Zeuthen, Germany
[3]*Department of Physics, National and Kapodistrian University of Athens*
University Campus Zografos, GR 15783, Athens, Greece
[4]*Institute of Accelerating Systems & Applications*
University Campus Zografos, GR 15783, Athens, Greece
[5]*Faculty of Electrical Engineering and Computing , University of Zagreb*
Unska ul. 3 , 10000 Zagreb, Croatia

ABSTRACT

For a sub-population of energetic Gamma-Ray Bursts (GRBs), a moderate baryonic loading may suffice to power Ultra-High-Energy Cosmic Rays (UHECRs). Motivated by this, we study the radiative signatures of cosmic-ray protons in the prompt phase of energetic GRBs. Our framework is the internal shock model with multi-collision descriptions of the relativistic ejecta (with different emission regions along the jet), plus time-dependent calculations of photon and neutrino spectra. Our GRB prototypes are motivated by *Fermi*-LAT detected GRBs (including GRB 221009A) for which further, owing to the large energy flux, neutrino non-observation of single events may pose a strong limit on the baryonic loading. We study the feedback of protons on electromagnetic spectra in synchrotron- and inverse Compton-dominated scenarios to identify the multi-wavelength signatures, to constrain the maximally allowed baryonic loading, and to point out the differences between hadronic and inverse Compton signatures. We find that hadronic signatures appear as correlated flux increases in the optical-UV to soft X-ray and GeV to TeV gamma-ray ranges in the synchrotron scenarios, whereas they are difficult to identify in inverse Compton-dominated scenarios. We demonstrate that baryonic loadings around 10, which satisfy the UHECR energetic requirements, do not distort the predicted photon spectra in the *Fermi*-GBM range and are consistent with constraints from neutrino data if the collision radii are large enough (i.e., the time variability is not too short). It therefore seems plausible that under the condition of large dissipation radii a population of energetic GRBs can be the origin of the UHECRs.

Keywords: gamma-ray burst: general – cosmic rays – neutrinos – radiation mechanisms: non-thermal

1. INTRODUCTION

Gamma-ray bursts, as the most luminous extragalactic energy sources in the Universe, were proposed early on as promising candidates for the acceleration of ultra-high energy cosmic rays (UHECRs; Milgrom & Usov 1995; Waxman 1995; Vietri 1995; Waxman & Bahcall 1997; Bhattacharjee & Sigl 2000; Dermer & Atoyan 2006). The exact mechanism for accelerating UHE-CRs in the GRB outflow to rigidities up to $1.6 - 3.5 \cdot 10^{18}$ V (Heinze et al. 2019) has not been pinned down yet, with magnetic reconnection (e.g. Giannios 2010) and shock acceleration (e.g. Murase & Nagataki 2006; Dermer & Razzaque 2010) posing the two leading scenarios. The presence of very energetic protons and nuclei in the GRB outflow is accompanied by the production of High-Energy (HE) neutrinos, following the inter-

Corresponding author: Annika Rudolph
annika.lena.rudolph@nbi.ku.dk

actions of cosmic rays with the prompt emission photon field (Waxman & Bahcall 1997; Dermer & Atoyan 2006; Murase & Nagataki 2006; Hummer et al. 2012; Asano & Mészáros 2013; Zhang & Kumar 2013; Petropoulou 2014; Biehl et al. 2018; Pitik et al. 2021).

The leading models for the GRB prompt emission involve a dissipative fireball expanding relativistically (Lorentz factor $\Gamma \geq 100$), where the energy dissipation occurs in internal shocks when portions of the jet are moving outwards with varying Lorentz factors (Rees & Meszaros 1994; Daigne & Mochkovitch 1998; Kobayashi et al. 1997), or via magnetic reconnection in Poynting-flux dominated outflows (Meszaros & Rees 1997; Spruit et al. 2001; Drenkhahn 2002; Drenkhahn & Spruit 2002; Lyutikov & Blandford 2003; Kumar & Crumley 2015; Giannios & Uzdensky 2019). In both cases a fraction of the thermal electrons in the outflow is accelerated to relativistic velocities and their subsequent cooling typically produces the synchrotron spectrum observed as a prompt gamma-ray burst emission at sub-MeV energies. Energy dissipation in the sub-photospheric region resulting in electron heating can also lead to the significant contribution from photospheric photons to the observed non-thermal spectrum (e.g. Giannios 2012; Rees & Meszaros 2005; Pe'er et al. 2006; Toma et al. 2011).

If accelerated protons carry a much larger amount of the dissipated energy than electrons, then the photon spectrum below and above the MeV energy range may show additional components related to photohadronic interactions and/or photodisintegration of nuclei (see e.g. Asano et al. 2009b; Murase & Beacom 2010). The possibility that the sub-MeV prompt emission spectrum is explained by radiation from secondary leptons produced in photohadronic interactions in proton-dominated outflows has also been explored (e.g. Murase et al. 2012; Petropoulou et al. 2014).

While the Gamma-Ray Burst Monitor (GBM) on board the *Fermi* satellite remains the most prolific instrument for the detection of GRBs in the sub-MeV to MeV energy range, the Large Area Telescope offers complementary spectral information at energies above 100 MeV (Ajello et al. 2019)[1]. For instance, LAT observations of the prompt phase of several individual GRBs have revealed the presence of a hard power-law component (with photon index less than 2) extending from the lower GBM energies to the GeV band (e.g. Ackermann et al. 2010; Zhang et al. 2011; Guiriec et al.

2015; Tang et al. 2021). Among the proposed scenarios offered as an explanation for this additional spectral component, those involving UHECRs (e.g. Asano et al. 2009a) are attractive for multi-messenger connections. The possibility that UHECRs are accelerated in GRBs was also studied through the direct cross-correlations between the LAT HE ($\gtrsim 1$ GeV) events and UHECRs detected by the Telescope Array (TA) and the Pierre Auger Observatory (PAO) (Alvarez et al. 2016). However, no statistically significant correlation was found in comparisons to date. The detectable neutrino fluence predicted from the simplest theoretical models was neither confirmed in the analysis of the IceCube data for individual bright gamma-ray bursts (Gao et al. 2013), nor identified in the model-independent stacking analysis searches performed on large samples of GRBs with different spectra (Abbasi et al. 2012; Aartsen et al. 2016, 2017; Abbasi et al. 2022). The recent detections of very-high energy (> 100 GeV) photons from GRBs (Acciari et al. 2019; Abdalla et al. 2019; Yong Huang et al. 2022) have prompted the searches for the associated multi-messenger counterparts, but no neutrino detection from such GRBs has been confirmed so far.

The null results from multi-messenger searches of GRBs so far do not necessarily reject the hypothesis that UHECRs are accelerated in GRB outflows. In fact, several theoretical predictions about UHECRs, HE gamma-rays and neutrinos during the GRB prompt phase rely on one-zone models, where the emission from the whole burst is assumed to be produced in one region of the outflow, see e.g. Waxman & Bahcall (1997); Murase & Nagataki (2006); Hummer et al. (2012); Asano & Meszaros (2012); He et al. (2012); Baerwald et al. (2015); Biehl et al. (2018); Pitik et al. (2021). Since both the neutrino and UHECR production are sensitive to the locally produced photon spectrum, it is necessary to account for the temporal evolution of the relativistic outflow and evolution of the physical conditions in the emission regions. In a series of more recent works (Bustamante et al. 2015, 2017; Rudolph et al. 2020; Heinze et al. 2020), neutrino and cosmic-ray emission from multiple emission regions were considered. It was demonstrated that different messengers originate from different shock collision radii and that the neutrino fluence is dominated by the collisions close to the photosphere, leading to lower predicted neutrino fluences than the one-zone models. However, the local photon spectrum in these works was motivated from observations and not self-consistently computed. While Globus et al. (2015) considered a multi-zone model for UHECR nuclei with target photons self-consistently produced by leptonic

---

[1] Note that *Fermi* LAT probes primarily most energetic events observed by the Gamma-Ray Burst Monitor (GBM), though LAT-detected GRBs cover a large range of peak flux and fluence values when compared to the GBM population.

processes, they did not account for the feedback from hadronic processes on the electromagnetic spectrum.

In this paper we present a fully self-consistent lepto-hadronic radiative model for multiple internal shocks occurring at different locations above the GRB photosphere. Our model accounts for the different physical conditions of each collision, while taking into account the feedback of high-energy protons on the locally produced photon spectrum. We focus on energetic *Fermi*-LAT detected GRBs (i.e. $E_{\gamma,\rm iso} \gtrsim 10^{54}\,\rm erg$). These GRBs can provide the necessary energy output per GRB in order to power the UHECRs, while only a moderate baryonic loading (defined as the energy injected into non-thermal protons versus electrons) is required and even energy equipartition might be possible. The latter property is especially attractive because it can mitigate the energetic problems that models (for typical GRBs) with high baryonic loadings face when it comes to the afterglow emission. Given the moderate dissipation efficiency of internal shocks (typically less than 10%, see e.g. Panaitescu et al. (1999); Beloborodov (2000); Bosnjak et al. (2009)), a high baryonic loading implies a high outflow kinetic luminosity, a large fraction of which will end up powering the afterglow. This may be in tension with afterglow observations (see e.g. Beniamini et al. 2016), although the implications of the afterglow measurements for the prompt emission efficiency depend on the afterglow model. Nonetheless, if scenarios with high baryonic loadings for "ordinary" GRBs are disfavored, this raises the question of why energetic GRBs should be more efficient UHECR accelerators than "ordinary" GRBs. An answer to that question either points towards a different population of bursts or to "friendlier" conditions for UHECR acceleration in energetic bursts, see Sec. 7 for a deeper discussion.

The paper is structured as follows. In Section 2 we present the implementation of the internal shock scenario for the GRB prompt emission and the parameters of the model describing the physical conditions in the shocked regions. For the numerical modeling of the cooling of the injected distributions of particles we used the time-dependent code AM³ (Gao et al. 2017). We introduce two model GRBs (prototypes), a single-pulse burst and a multi-peaked event (inspired by GRB 170214A) in Section 3. We explore first the leptonic model using as test-bed the single-pulse GRB, and present our results in Section 4. In Section 5 we examine the lepto-hadronic models for both prototypes. We investigate different baryonic loadings and the impact of the variability timescales for single-peaked and multi-peaked synthetic GRBs. A multi-collision GRB model offers the possibility to investigate the time-dependence for the observation of different messengers (photons, neutrinos, UHECRs). In Section 6 we discuss possible ways to discriminate between different parameter regions using current multi-messenger observations. We continue the discussion on other aspects of our models, such as UHECR composition and future directions in Section 7. We summarize our work and present our conclusions in Section 8.

2. A MULTI-COLLISION INTERNAL SHOCK RADIATION MODEL

In the internal shock model the GRB prompt emission is generated in interactions between fast and slow parts of the relativistic ejecta (Rees & Meszaros 1994; Kobayashi et al. 1997; Daigne & Mochkovitch 1998). We follow the formalism presented in Daigne & Mochkovitch (1998) and discretize the outflow as a series of plasma layers (called 'shells') that propagate at different velocities. As fast shells catch up with slower ones they collide and energy is dissipated. The overall fireball emission is obtained by adding the contributions of all single collisions.

The relativistic outflow is discretised as shells of source-frame width $\Delta = c\mathrm{d}\tau$, where $\mathrm{d}\tau$ is calculated from the number of initial shells $N_{\rm shells}^{\rm ini}$ and the engine active time $t_{\rm eng}$ as $\mathrm{d}\tau = t_{\rm eng}/N_{\rm shells}^{\rm ini}$. Each shell is further characterised by its mass M and Lorentz factor Γ and has a volume $V = 4\pi R^2\Delta$ at a distance R from the central emitter [2]. We emphasise that the number of initial shells (and the resulting number of collisions) in our model are a discretisation choice and the results are independent of it as long as there are enough shells to adequately resolve the fireball evolution.

2.1. *Two-shell collision*

We first recapitulate the formulas describing the collision between two shells. A collision (labelled with a subscript 'Coll') between a fast (subscript 'f') and a slow ('s') shell at a radius $R_{\rm Coll}$ and time $T_{\rm Coll}$ (in the source frame) creates a new merged (subscript 'm') shell that continues in the fireball. From energy and momentum conservation the merged shell mass and Lorentz factor can be calculated as

$$M_{\rm m} = M_{\rm f} + M_{\rm s}\,, \tag{1}$$

$$\Gamma_{\rm m} = \sqrt{\frac{M_{\rm f}\Gamma_{\rm f} + M_{\rm s}\Gamma_{\rm s}}{M_{\rm f}/\Gamma_{\rm f} + M_{\rm s}/\Gamma_{\rm s}}}\,. \tag{2}$$

---

[2] We use three different frames of reference in this paper: The source (or engine) frame, the plasma comoving frame and the observers frame. Quantities in those frames will be denoted as X, X' and $X_{\rm obs}$ respectively.

The collision time as measured in the observer's frame is given by

$$T_{\rm Coll,obs} = (1+z)(T_{\rm Coll} - \frac{R_{\rm Coll}}{c}) \qquad (3)$$

where z is the redshift of the burst. Note that this also corresponds to the earliest time at which photons of a collision may be observed.

We specify next the plasma conditions in the shocked plasma of the merged shell. We assume the Lorentz factor of the emission region is the same as that of the shocked plasma region, which is given by

$$\Gamma_{\rm em} = \sqrt{\Gamma_{\rm s}\Gamma_{\rm f}}. \qquad (4)$$

This formula is obtained by assuming that most of the energy is dissipated as the less massive shell has swept up a mass comparable to its own. The mass density of the plasma is calculated as

$$\rho' = \frac{M_{\rm m}}{V'_{\rm Coll}} = \frac{M_{\rm m}}{4\pi R_{\rm Coll}^2 \Gamma_{\rm em} c\, d\tau}, \qquad (5)$$

where we used the comoving width of the shell $\Delta'_{\rm Coll} = \Gamma_{\rm em} c\, d\tau$. Energy conservation gives the dissipated energy as

$$E_{\rm diss} = (M_{\rm f}\Gamma_{\rm f} + M_{\rm s}\Gamma_{\rm s} - M_{\rm m}\Gamma_{\rm m})c^2. \qquad (6)$$

From the comoving dissipated energy $E'_{\rm diss} = E_{\rm diss}/\Gamma_{\rm em}$ we define the comoving energy density as

$$u'_{\rm diss} \equiv \frac{E'_{\rm diss}}{V'_{\rm Coll}}. \qquad (7)$$

We further define the dissipated energy per unit mass $\epsilon'_{\rm diss}$ as

$$\epsilon'_{\rm diss} = \frac{E'_{\rm diss}}{M_{\rm m}}. \qquad (8)$$

The characteristic timescale of the system (also called 'dynamical timescale') is identified as the shell expansion time

$$t'_{\rm dyn} = \frac{R_{\rm Coll}}{\Gamma_{\rm em} c}. \qquad (9)$$

We also specify the fraction of energy that is transferred to the different particle species, quantified by the microphysics parameters $\epsilon_{\rm i}$. Under the assumption that the observed prompt emission is dominated by emission of non-thermal electrons, it is convenient to relate all quantities to the fraction of energy transferred to non-thermal electrons ϵ_e. We thus define $f_{\rm p/e} = \epsilon_{\rm p}/\epsilon_e$ (where $\epsilon_{\rm p}$ is the fraction of energy transferred to non-thermal protons), $f_{\rm TH/e} = \epsilon_{\rm TH}/\epsilon_e$ (where $\epsilon_{\rm TH}$ is the fraction of energy transferred to thermal particles) and

$f_{\rm B/e} = \epsilon_{\rm B}/\epsilon_e$ (where $\epsilon_{\rm B}$ is the fraction of energy transferred to the magnetic field). Note that by this definition, $f_{\rm TH/e}$ accounts for both electrons and protons and that the parameter, although set to 0 in the following, may in reality non-negligible. The comoving magnetic field strength can then be expressed in terms of the comoving non-thermal electron energy density $u'_{\rm ele,NT} = \epsilon_e u'_{\rm diss}$ as follows:

$$B' = \sqrt{8\pi f_{\rm B/e} u'_{\rm ele,NT}}. \qquad (10)$$

2.2. Fireball energy normalisation

The initial kinetic energy of the fireball can be written as

$$
\begin{aligned}
E_{\rm kin,ini} &= \varepsilon^{-1} E_{\rm diss} = \varepsilon^{-1}\left(E_{\rm e,NT}^{\rm tot} + E_{\rm p,NT}^{\rm tot} + E_{\rm B} + E_{\rm TH}^{\rm tot}\right)\\
&= \varepsilon^{-1} E_{\rm e,NT}^{\rm tot}(1 + f_{\rm p/e} + f_{\rm B/e} + f_{\rm TH/e}), \qquad (11)
\end{aligned}
$$

where ε is the fireball (dissipation) efficiency defined as

$$\varepsilon \equiv 1 - \frac{\sum_{\rm i} M_{\rm i,fin}\Gamma_{\rm i,fin}}{\sum_{\rm i} M_{\rm i,ini}\Gamma_{\rm i,ini}}. \qquad (12)$$

In the following we will assume that the outflow is launched with a constant wind luminosity, which implies that all initial shells carry the same initial energy $E_{\rm i,ini} = d\tau/t_{\rm eng} \cdot E_{\rm kin,ini}$.

Assuming that the observed sub-MeV prompt spectrum is predominately produced by leptonic processes, it is convenient to normalise the initial fireball kinetic energy to the total energy transferred to non-thermal electrons $E_{\rm e,NT}^{\rm tot}$ that is needed to produce a given isotropic-equivalent energy in gamma rays $E_{\gamma,\rm iso}$ (the isotropic equivalent energy emitted in gamma-rays in the energy range of 1 - 100 keV). With this normalization we obtain ϵ_e for each set of $(f_{\rm p/e}, f_{\rm B/e}, f_{\rm TH/e})$ as

$$\epsilon_e = (1 + f_{\rm p/e} + f_{\rm B/e} + f_{\rm TH/e})^{-1}. \qquad (13)$$

2.3. The deceleration radius

The deceleration radius $R_{\rm dec}$ which marks the end of the prompt emission phase of the fireball is reached when the initial fireball kinetic energy is equal to the energy of the heated downstream plasma (given by $E_{\rm SW} = \Gamma^2 M_{\rm SW}(R)c^2$, where $M_{\rm SW}(R)$ is the swept-up mass at a radius R). We adjust Eq. 1 from Rees & Meszaros (1992) for a wind-like medium and find

$$R_{\rm dec} = \frac{1}{4\pi\rho_0} \sum_i \frac{M_{\rm i,ini}}{\Gamma_{\rm i,ini}}, \qquad (14)$$

where i sums over all initial shells contained in the fireball. We calculate ρ_0 as $\rho_0 = \frac{\dot{M}}{4\pi v_w}$ with typical values of the mass ejection rate $\dot{M}$ and wind velocity v_w that are

given by $\dot{M} = 10^{-5} M_\odot / \mathrm{yr}$ and $v_w = 10^3$ km/s. Note that a deviation from those values results in a different deceleration radius.

The initial masses can be re-expressed through the initial kinetic energy of the fireball:

$$\sum_i \frac{M_{\mathrm{i,ini}}}{\Gamma_{\mathrm{i,ini}}} = \frac{\mathrm{d}\tau}{t_{\mathrm{eng}} c^2} E_{\mathrm{kin,ini}} \sum_i \frac{1}{\Gamma_{\mathrm{i,ini}}^2} , \qquad (15)$$

where i again sums over the initial shells of the fireball.

For the calculating of the deceleration radius we invoke two sets of parameters: *(a)* The *equipartition* case with $f_{\mathrm{p/e}} = f_{\mathrm{B/e}} = f_{\mathrm{TH/e}} = 1$ (where thus non-thermal electrons, protons, the magnetic field and thermal particles receive equal parts of energy) and *(b)* the *hadronic* case with $f_{\mathrm{p/e}} = 30$, $f_{\mathrm{B/e}} = 1$ and $f_{\mathrm{TH/e}} = 0$ (where thus the majority of the energy is transferred to non-thermal protons). The corresponding radii will be labelled $R_{\mathrm{dec,equi}}$ and $R_{\mathrm{dec,had}}$.

2.4. Non-thermal particle distributions and radiative calculations

From the internal shock model that specifies the conditions in the shocked plasma we can obtain the parameters for the non-thermal particle distributions that are an input for the radiation modeling. Note that we do not account for the effect of photons emitted from earlier collisions.

2.4.1. Injected particle distributions

We first describe how we calculate the injected distributions of primary particles, considering a species i with mass m_{i}, Lorentz factor γ_{i}' and charge number Z_{i}. As we consider no heavy nuclei, $Z_{\mathrm{i}} = 1$. We assume that particles are accelerated at a very thin region close to the shock front before being injected into the (homogeneous) radiation zone of the shocked plasma. In the radiation zone of a single collision we do not account for any spatial dependence of quantities in the shocked plasma, such as magnetic field decay away from the shock (e.g. Lemoine 2013).

The injection rate of particle species i into the radiation zone is given by

$$\frac{\mathrm{d}n_{\mathrm{i}}'}{\mathrm{d}t' \mathrm{d}\gamma_{\mathrm{i}}'} = \dot{n}_{0,\mathrm{i}} \gamma_{\mathrm{i}}'^{-p_{\mathrm{i}}} e^{-\gamma_{\mathrm{i}}'/\gamma_{\mathrm{i,max}}'} \qquad (16)$$

for Lorentz factors above a minimum value of $\gamma_{\mathrm{i,min}}'$. The maximum Lorentz factor, $\gamma_{\mathrm{i,max}}'$, at any given time is determined by the balance between the radiative loss timescale and the acceleration timescale. The loss timescales for electrons are due to synchrotron, inverse Compton and adiabatic cooling, while for protons they

are due to photo-pion, photo-pair, synchrotron and adiabatic cooling. In general, the acceleration time can be written as

$$t_{\mathrm{acc}}' = \eta_{\mathrm{acc}} \frac{E_{\mathrm{i}}'}{cB'e} , \qquad (17)$$

where η_{acc} specifies the acceleration efficiency. Considering the fastest acceleration time we set $\eta_{\mathrm{acc}} = 1$ throughout this paper. The power-law index for electrons p_{e} is considered to be a free parameter determined by prompt observations. For protons, however, the index cannot be usually inferred by electromagnetic observations alone. So, our benchmark choice is $p_{\mathrm{p}} = 2$, but we discuss the impact of other indices in Section 6.

If ϵ_{i} defines the fraction of energy transferred to the species i (that can be calculated from the $f_{\mathrm{i/e}}$ values and ϵ_e through Eq. 13), the number fraction of accelerated particles ζ_{i} and the minimum Lorentz factor $\gamma_{\mathrm{i,min}}'$ are related through

$$\gamma_{\mathrm{i,min}}' = \frac{p_{\mathrm{i}} - 2}{p_{\mathrm{i}} - 1} \frac{\epsilon_{\mathrm{i}}}{\zeta_{\mathrm{i}}} \frac{m_{\mathrm{p}}}{m_{\mathrm{i}}} \frac{u_{\mathrm{diss}}'}{\rho' c^2} . \qquad (18)$$

This relationship is obtained by setting the total particle density n_{i}' (integrated over γ_{i}') equal to $\zeta_{\mathrm{i}} \rho'/m_{\mathrm{p}}$, assuming that the shocked plasma is composed of ionized hydrogen. For both electrons and protons we assume $\zeta_{\mathrm{i}} \propto \epsilon_{\mathrm{diss}}'$ which results in a constant $\gamma_{\mathrm{i,min}}'$ throughout the fireball evolution. As suggested in Asano et al. (2009b), the minimum Lorentz factor in mildly relativistic shocks should be of order unity and following their approach we set $\gamma_{\mathrm{p,min}}' = 10$. On the other hand, $\gamma_{\mathrm{e,min}}'$ will be adjusted to fit the peak of the spectrum in the observer's frame for each scenario. In Appendix B we further explore the additional scenario of $\zeta_e = \mathrm{const.}$ where $\gamma_{\mathrm{e,min}}'$ evolves proportional to $\epsilon_{\mathrm{diss}}'$ throughout the fireball evolution (see also Bošnjak & Daigne 2014).

The normalisation $\dot{n}_{0,\mathrm{i}}$ of the injected particle distribution is given by

$$\epsilon_{\mathrm{i}} u_{\mathrm{diss}}' = m_{\mathrm{i}} c^2 \int_{\hat{t}'=0}^{\hat{t}'=t_{\mathrm{inj}}'} \mathrm{d}\hat{t}' \dot{n}_{0,\mathrm{i}} \int_{\gamma_{\mathrm{i,min}}'}^{\infty} \mathrm{d}\gamma_{\mathrm{i}}' \gamma_{\mathrm{i}}'^{-p_{\mathrm{i}}+1} . \quad (19)$$

Here, t_{inj}' is the timescale over which particles are injected in the radiation zone. Our fiducial assumption is continuous injection of accelerated particles over the dynamical timescale, i.e. $t_{\mathrm{inj}}' = t_{\mathrm{dyn}}'$. In Appendix B we further explore an acceleration timescale much shorter than t_{dyn}', labelled as $\delta t_{\mathrm{inj}}' \to 0$. In the numerical calculations this is approximated by an injection over a timescale much shorter than t_{dyn}'; here, we choose $t_{\mathrm{inj}}' = 0.01 t_{\mathrm{dyn}}'$.

2.4.2. Numerical treatment

The numerical radiation modeling is performed with the time-dependent code AM³ (Gao et al. 2017) which follows the coupled evolution of photons, electrons, positrons, muons, pions, protons, neutrons and neutrinos. The software includes all relevant non-thermal processes such as synchrotron and synchrotron self-absorption, inverse Compton scatterings, photo-pair and photo-pion production, $\gamma\gamma$-annihilation, adiabatic cooling and escape. Secondary particles produced in these interactions are added to the overall particle distributions and undergo the same processes as primaries.

The numerical treatment is described in detail in Gao et al. (2017). Here, we briefly list the modifications to the original version of the code:

- Adiabatic cooling of charged particles is implemented with a cooling rate $d\gamma_i'/dt' = -\gamma_i'/t_{dyn}'$.

- Muons and pions are treated as separate species that may be subject to synchrotron and adiabatic cooling prior to their decay and emit synchrotron radiation.

- Quantum synchrotron radiation is implemented following Brainerd (1987).

- The treatment of photo-pair production has been updated and now follows Kelner & Aharonian (2008).

For each collision we then compute the temporal and spectral evolution of the comoving particle energy density of each particle species, $u_i'(E', t')$. We point out that feedback between different collisions is not accounted for.

2.4.3. Calculation of emitted spectra

While charged particles are assumed to be confined by the magnetic fields, neutral particles escape at a rate $t_{esc}'^{-1} = 1/t_{dyn}'$. The differential spectrum of particles that have escaped until a time t' can be calculated from the time-dependent comoving density of the relevant particle species, $u_i'(E', t') = E_i' dN_i/dE_i' dV'$, as

$$u_{esc,i}'(E', t') = \frac{1}{t_{esc}'} \int_{\hat{t}'=0}^{\hat{t}'=t'} u_i'(E', \hat{t}') d\hat{t}' . \quad (20)$$

To compute the differential emitted spectrum, $u_{em}'(E')$, we follow the system's evolution until $t' = t_{inj}' + t_{dyn}'$, at which point all primary particles have either escaped or cooled. The particles remaining in the system at this point are then added to the spectra of escaped particles.

2.4.4. Conversion into observed quantities

For the calculation of time-dependent quantities (light curves and time-resolved spectra), we take into account the curvature of the emitting surface following Granot et al. (1999). We however introduce a simplified approach where the emission from the thin shell is assumed to be released at a single time T_{Coll} and from a single shell surface of radius R_{Coll} (thus the integral over R in Granot et al. (1999) is a δ-function for each collision). For calculations that invoke only a single radiation zone, this over-simplification would severely impact the observed profile and not reproduce the Fast-Rise-Exponential-Decay (FRED) (see e.g. Granot et al. 1999; Salafia et al. 2016, for the case of an infinitesimally thin shell). When computing the emission from a large number of collisions though, as in this work (see next section), the emission profile is dominated by the fireball evolution and our simplified approach will not impact the predictions significantly.

For a given collision (with radius R_{Coll}, emission time T_{Coll}, Lorentz factor Γ_{em} and comoving width Δ_{Coll}') we calculate the energy flux, $F(T_{obs})$, integrated over the full energy range (in erg s^{-1} cm^{-2}) as

$$F(T_{obs}) = \frac{1}{2D^2} \int dE' \frac{R_{Coll}c}{(1+z)\Lambda^2} u_{emi}'(E')\Delta_{Coll}' , \quad (21)$$

where D is the comoving distance and $\Lambda = \Gamma_{em}\left[1 - \frac{c\beta_{em}}{R_{Coll}}\left(T_{Coll} - \frac{T_{obs}}{1+z}\right)\right]$.

For spectra integrated over the full duration of the burst we apply a simplified procedure. The (differential in energy) observed fluence $\mathcal{F}_E$ of a single collision is simply given by

$$E\mathcal{F}_E = u_{em}'(E')V_{Coll}'\frac{\Gamma_{em}E'}{4\pi(1+z)D^2} . \quad (22)$$

The full fireball emission is obtained by adding the contributions of all single collisions.

The effect of absorption due to interactions with the Extragalactic Background Light (EBL) are discussed in Section 6.2 and will be omitted in all other sections. We calculate it with the open-source GAMMAPY-package (Deil et al. 2018; Nigro et al. 2019), selecting the model of Dominguez et al. (2011).

3. INTRODUCING TWO PROTOTYPES

For the purposes of our study it is useful to introduce two model GRBs that differ mainly in their temporal properties. The first prototype is characterized by a smooth single-pulse (SP) light curve, while the second one has a multi-peaked (MP) light curve with short-timescale variability. Given that we will pay special attention to possible HE emission signatures, we loosely

base the parameters of our prototypes on the properties of GRBs detected by *Fermi*-LAT. As pointed out in Ajello et al. (2019), these populate the upper range of the $E_{\gamma,\mathrm{iso}}$−distribution. Consequently, we choose events with a high energy budget. The $E_{\gamma,\mathrm{iso}}$- Γ correlation (see e.g. Liang et al. 2010; Ghirlanda et al. 2018) and the requirement of being optically thin to $\gamma\gamma$ pair production, further imply comparatively high Lorentz factors of the outflow.

Leptonic scenarios will be illustrated with solely the simple, single-peaked burst. On the other hand, lepto-hadronic models (with different baryonic loadings $f_{\mathrm{p/e}}$) will be discussed for both prototypes.

3.1. *Model parameters*

Before going into detail on the two prototypes let us collect the model parameters and assumptions for better overview. Each of our prototypes will have a targeted gamma-ray isotropic energy, observed duration and observed peak energy, while it is placed at an assumed redshift (see Table 1). Targeted here refers to the order of magnitude value that we will aim to reproduce. To this end we impose *internal shock model parameters* (characterising jet parametrisation through shells) and *microphysical parameters* (characterising the conditions in the shocked plasma), collected in Table 2. The internal shock model parameters are complemented by the initial Lorentz factor distributions, as depicted in Figure 1.

Internal shock model parameters. We choose the engine active time t_{eng} that reproduces the targeted duration. Then, after selecting an initial Lorentz factor distribution of the shells (that is chosen to match a certain light curve profile), the initial energies of the shells are set such that the targeted $E_{\gamma,\mathrm{iso}}$ is approximately matched. For the latter we follow Eq. 11, assuming a set of microphysics parameters f_i.

Microphysical parameters. In the spirit of a parameter study we will explore different combinations of f_i. Namely, we will study a scenario with a strong and a weak magnetic field in the outflow. This is achieved by imposing $f_{\mathrm{B/e}} = 1$ and $f_{\mathrm{B/e}} = 10^{-3}$, which correspond respectively to a synchrotron (SYN)-dominated scenario for electron cooling, and an inverse Compton (IC)-dominated scenario. Both regimes will be explored for leptonic ($f_{\mathrm{p/e}} = 0$) and lepto-hadronic models ($f_{\mathrm{p/e}} > 0$). The benchmark baryonic loading used in the latter is $f_{\mathrm{p/e}} = 30$, but we will also explore other values in the range $0.3 - 100$.

To reproduce the sub-MeV peak energy that we use as benchmark we adjust the fraction of accelerated electrons ζ_e for each parameter set. More specifically, we
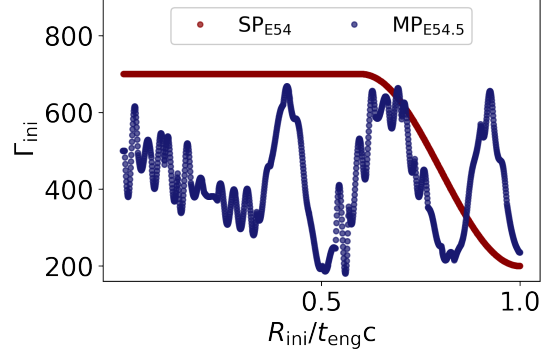
Figure 1. Initial Lorentz factor distribution for two GRB prototypes as a function of initial radius R_{ini} (normalised to $t_{\mathrm{eng}}c$).

set $E_{\mathrm{peak,obs}}$ as the synchrotron peak energy of electrons at the minimum Lorentz factor as defined in Eq. 18 and solve for $\zeta_{e,0}$. This gives $\zeta_{e,0}^{\mathrm{k}}$ for the k−th collision as

$$\zeta_{e,0}^{\mathrm{k}} = 6.2 \cdot 10^{-4} \left[\frac{(\epsilon_B u_{\mathrm{diss}}'^k)^{1/2} \Gamma_{\mathrm{em}}^{\mathrm{k}}}{(1+z) E_{\mathrm{peak,obs}}} \right]^{1/2} \frac{p_e - 2}{p_e - 1} \epsilon_e . \quad (23)$$

The reported value of $\zeta_{e,0}$ is then the average over all collisions, weighted by the dissipated energy of the collision. For completeness we will further list the minimum Lorentz factor that corresponds to the average $\zeta_{e,0}$. The initial power-law index of electrons will be set to $p_e = 2.5$ (to reproduce typical high-energy slopes of GRBs) and $p_e = 3.0$ (to reproduce a steeper high-energy slope that was observed by *Fermi*-LAT for GRB 170414A, which inspired our second prototype). For simplicity a minimum Lorentz factor of 10 and $p_p = 2.0$ will be assumed in all cases for protons.

3.2. *Single-pulse burst (SP$_{\mathrm{E54}}$)*

Our first prototype is a single-pulse burst without short-timescale variability in the light curve, referred to as SP$_{\mathrm{E54}}$ (the subscript refers to the isotropic-equivalent energy of the GRB). This type of simple example has been studied in the past (see for example Bosnjak et al. 2009; Globus et al. 2015). This light curve is representative of a single pulse in the GRB time evolution and the simple temporal and spatial structure are useful to study the effects of different emission zones and their contribution to the overall emitted spectrum.

The burst is generated from a smooth initial Lorentz factor distribution that is defined as $\Gamma(R_{\mathrm{ini}}) = \Gamma_{\mathrm{max}}$, for $R_{\mathrm{ini}}/c < 0.6\, t_{\mathrm{eng}}$ and $\Gamma(R_{\mathrm{ini}}) = (\Gamma_{\mathrm{max}} + \Gamma_{\mathrm{min}})/2 - (\Gamma_{\mathrm{max}} - \Gamma_{\mathrm{min}}) \cos\left(\pi \frac{R_{\mathrm{ini}}}{0.4 c t_{\mathrm{eng}}}\right)/2$, for $0.6\, t_{\mathrm{eng}} \leq R_{\mathrm{ini}}/c < t_{\mathrm{eng}}$ (Daigne & Mochkovitch 1998); here R_{ini} is the initial shell radius and t_{eng} is the overall engine activity time. We choose $\Gamma_{\mathrm{min}} = 200$ and $\Gamma_{\mathrm{max}} = 700$ as bench-

Table 1. Parameter values used for two GRB prototypes inspired by observations of energetic bursts detected by *Fermi*-LAT.

Parameter	Symbol	SP$_{E54}$	MP$_{E54.5}$	
			$\delta t_{\rm var} = 1.13$ s	$\delta t_{\rm var} = 0.11$ s
Targeted isotropic γ-ray energy (source frame)	$E_{\gamma,\rm iso}$	10^{54} erg	$2.9 \cdot 10^{54}$ erg	$2.9 \cdot 10^{54}$ erg
Targeted peak energy (observer frame)	$E_{\rm peak,obs}$	400 keV	566 keV	566 keV
Targeted duration (observer frame)	$T_{\rm dur,obs}$	15 s	106 s	10.6 s
Assumed redshift	z	2	2	2

Table 2. Fireball parameters and fiducial microphysics parameters used for the modeling of two energetic GRB prototypes.

| Parameter	Symbol	SP$_{E54}$	MP$_{E54.5}$ ($\delta t_{\rm var} = 1.13$ s)	MP$_{E54.5}$ ($\delta t_{\rm var} = 0.11$ s)
Number of initial shells	$N_{\rm shells}^{\rm ini}$	1000	1297	1297
Engine active time	$t_{\rm eng}$	5 s	34 s	3.4 s
Number of collisions	$N_{\rm coll}$	999	1139	1139
Total energy in non-thermal electrons	$E_{\rm e,NT}^{\rm tot}$	$1.3 \cdot 10^{54}$ erg	$3.5 \cdot 10^{54}$ erg	$3.5 \cdot 10^{54}$ erg
Average collision radius	$\langle R_{\rm Coll} \rangle$	$1.9 \cdot 10^{16}$ cm	$2.4 \cdot 10^{16}$ cm	$2.4 \cdot 10^{15}$ cm
Overall dissipation efficiency	ε	7.8%	2.98 %	2.98 %
Power-law index of non-thermal electrons	$p_{\rm e}$	2.5	3.0	3.0
Power-law index of non-thermal protons	$p_{\rm p}$	2.0	2.0	2.0
Minimum Lorentz factor of non-thermal protons	$\gamma_{\rm p,min}'$	10	10	10
Relative fraction of energy transferred to thermal particles	$f_{\rm TH/e} = \epsilon_{\rm TH}/\epsilon_{\rm e}$	0	0	0
			'SYN-dominated'	
Relative fraction of energy transferred to magnetic field	$f_{\rm B/e} = \epsilon_{\rm B}/\epsilon_{\rm e}$	1	1	1
Relative fraction of energy transferred to protons	$f_{\rm p/e} = \epsilon_{\rm p}/\epsilon_{\rm e}$	$\{\mathbf{0}, 10, \mathbf{30}, 100\}$	$\{\mathbf{0}, 3, 10, \mathbf{30}\}$	$\{\mathbf{0}, 0.3, 1, \mathbf{3}\}$
Normalisation for number fraction of accelerated electrons	$\zeta_{0,\rm e}$ [10^{-4}]	18.7	21.6	119.7
Minimum Lorentz factor of non-thermal electrons	$\gamma_{\rm e,min}$ [10^4]	1.2	1.5	0.2
			'IC-dominated'	
Relative fraction of energy transferred to magnetic field	$f_{\rm B/e} = \epsilon_{\rm B}/\epsilon_{\rm e}$	10^{-3}	10^{-3}	-
Relative fraction of energy transferred to protons	$f_{\rm p/e} = \epsilon_{\rm p}/\epsilon_{\rm e}$	$\{\mathbf{0}, 10, \mathbf{30}, 100\}$	$\{\mathbf{0}, 3, 10, \mathbf{30}\}$	-
Normalisation for number fraction of accelerated electrons	$\zeta_{0,\rm e}$ [10^{-4}]	3.3	3.8	-
Minimum Lorentz factor of non-thermal electrons	$\gamma_{\rm e,min}'$	6.5	8.6	-

Notes. – For $f_{\rm p/e}$ we list all parameter values explored in this work and mark in bold those used as benchmark values for the leptonic and lepto-hadronic models. The variability timescale in the source frame is given by $\delta t_{\rm var} = t_{\rm eng}/N_{\rm osc}$, where $N_{\rm osc} = 30$ is the number of short timescale oscillations in the initial Lorentz factor distribution, the average collision radius is obtained by weighing the distribution of $R_{\rm Coll}$ with their respective dissipated energy $E_{\rm diss}$. The number fraction of accelerated electrons in a collision in a collision can be calculated as $\zeta_e = \min\left(1, \zeta_{0,\rm e} \frac{\epsilon_{\rm diss}'}{100\,{\rm MeV/proton}}\right)$.

mark values, but we also examine different values later in Appendix B. The burst characteristics and input fireball parameters are summarised in Tables 1 and 2, respectively.

The emission from the full burst is the superposition of the emission from all individual collisions. As these occur at different locations in the outflow, the physical conditions of the emitting plasma (e.g. magnetic fields, energy densities, and so on) will be different. This is exemplified in Figure 2 (left panel) where we plot the total non-thermal electron energy, the magnetic field strength, the bulk Lorentz factor of the emission region,

and the observed time of a collision as a function of collision radius (each point represents a collision).

Because of the different physical conditions involved in each collision, it is difficult to examine in detail the role of the various physical processes at work by studying the emission of the full burst. For this reason, we will use a single *representative* collision. This is defined as the collision with the maximum dissipated energy $E_{\rm diss}$ and is indicated with a star ($\star$) symbol in Figure 2. We consider it representative in the sense that it will be the one dominating the observed emission, and its spectrum will be somewhat similar to the overall burst spectrum (this will become clearer in the next section). By com-
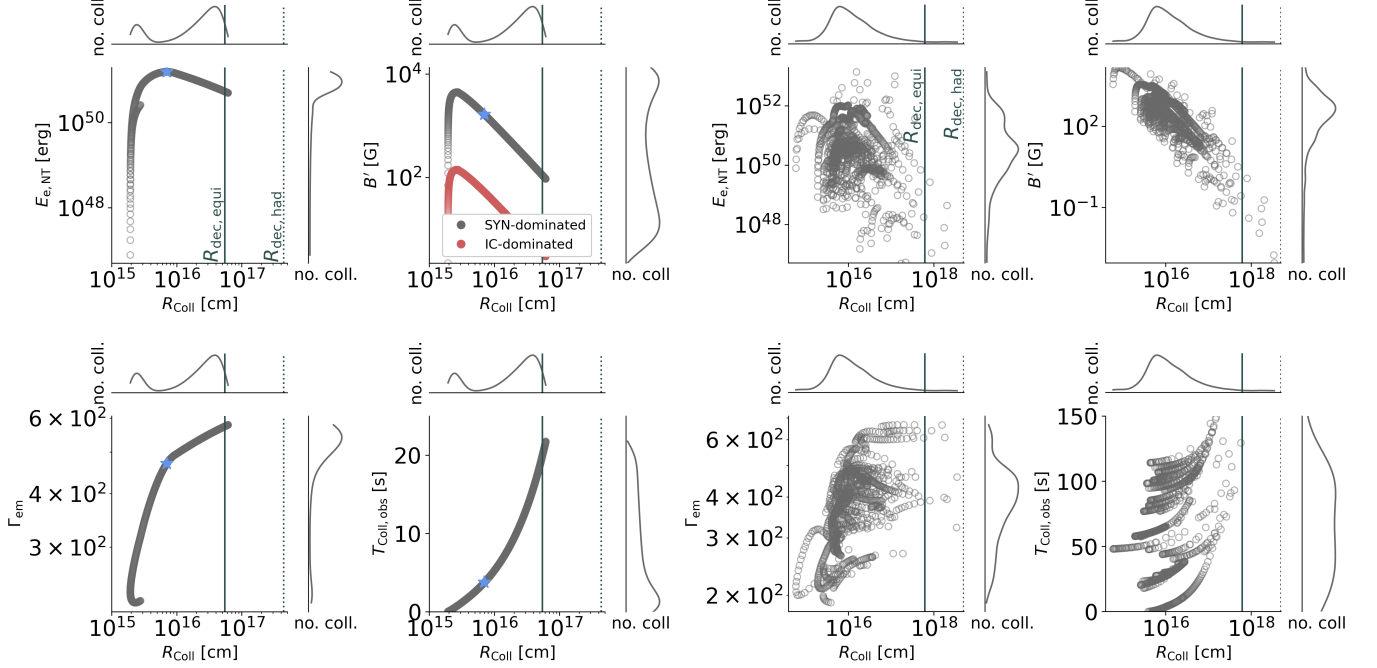
Figure 2. Evolution of outflow parameters as a function of collision radius for the *(left)* SP_{E54} and *(right)* $MP_{E54.5}$ prototypes: the energy carried by primary non-thermal electrons ($E_{e,NT}$), the Lorentz factor of the emitting shocked plasma (Γ_{em}), the comoving magnetic field (B') for the SYN- and IC-dominated scenarios, and the collision time in the observer's frame ($T_{Coll,obs}$). Vertical lines indicate the deceleration radius for two sets of microphysical parameters: $f_{p/e} = f_{B/e} = f_{TH/e} = 1$ (an equipartition (*equi*) scenario) and $f_{p/e} = 30$, $f_{B/e} = 1$ and $f_{TH/e} = 0$ (a hadronic (*had*) scenario); see Section 2.3 for details. For the SP_{E54} scenario on the left we mark the representative collision introduced in Section 3.2 with a $\star$ symbol and show magnetic fields for both the IC-and the SYN-dominated scenario. For the $MP_{E54.5}$ scenario on the right we only display the magnetic fields for the SYN-dominated scenario. For both prototypes, side panels show one-dimensional density distributions of the collisions.

paring the modeling results for the representative collision to those obtained from the full fireball evolution, we will discuss the limitations of simplified one-zone models where a single radiation zone is used for describing the full burst.

3.3. *Multi-pulse burst with short-time variability* *($MP_{E54.5}$)*

Most GRB light curves have complex structures consisting of many pulses and exhibiting variability on short timescales. These properties cannot be captured by the single-pulse model described in the previous section. We therefore introduce a second prototype burst with a multi-pulse light curve and short-timescale variability, referred to as $MP_{E54.5}$.

The prototype is inspired by GRB 170214A, a very energetic burst that was observed by both *Fermi*-GBM and *Fermi*-LAT. The onset of the LAT detection happened during the prompt phase, although delayed approximately 60 s with respect to the GBM trigger. Given that a similar short-time variability was observed in the GBM and LAT energy bands, an internal (prompt-phase) origin for the HE emission was proposed in Tang et al. (2017). In terms of energetics and (pre-

liminary) intrinsic peak energy, the burst is also similar to the recently observed GRB 221009A (S. Lesage et al. 2022; D. Frederiks et al. 2022; A. Ursi et al. 2022).

To produce an analogous burst (i.e. with similar E_{iso}, T_{90}, E_{peak}, high-energy photon index β as summarised in Table 1) we impose the fireball parameters indicated in Table 2. For the conversion to observed quantities we adopt the same redshift as for SP_{E54}, i.e. $z = 2$. The observed light curve of the burst is composed of a few long-timescale pulses with superimposed variability on short timescales. This type of pattern needs to be represented in the initial Lorentz factor distribution (which in the internal shock model is directly reflected on the light curve structure (see e.g. Bustamante et al. 2017)). Our chosen initial Lorentz factor distribution of the outflow is displayed in Figure 1.

In this case the distribution of plasma parameters as a function of collision radius deviates from the simple behaviour displayed in Figure 2. This is exemplified in Figure 2 (right), where we plot the same quantities as in Figure 2, with side panels showing the projected one-dimensional density distributions. We observe that the bulk of collisions occurs at $R_{Coll} \sim 10^{16}$ cm (similar to SP_{E54}). This is further reflected in the av-

erage dissipation radius $\langle R_{\mathrm{Coll}} \rangle$, obtained by weighing the distribution of R_{Coll} with the dissipated energy per collision. For MP$_{\mathrm{E54.5}}$ this weighted average is computed as $\langle R_{\mathrm{Coll}} \rangle = 2.4 \cdot 10^{16}$ cm (compared to $\langle R_{\mathrm{Coll}} \rangle = 1.9 \cdot 10^{16}$ cm for SP$_{\mathrm{E54}}$). This large typical collision radius is mainly driven by the adopted long duration and the relatively large Lorentz factors. Although the E_{iso} - Γ correlation may point to high dissipation radii for energetic bursts, typical quoted collision radii of GRBs in the internal shock model are $R_{\mathrm{Coll}} \sim 10^{11}$ cm to $R_{\mathrm{Coll}} \sim 10^{15}$ cm, which are significantly smaller than the values for MP$_{\mathrm{E54.5}}$ and SP$_{\mathrm{E54}}$.

In a one-zone internal shock model where a single collision is considered representative for the complete burst, the source-frame variability timescale δt_{var} can be used to estimate the collision radius through $R_{\mathrm{Coll}} = 2\Gamma_{\mathrm{em}}^2 c \delta t_{\mathrm{var}}$. Similarly, in multi-collision models the variability timescale can be used as a proxy of the collision radius; here we obtain $R_{\mathrm{Coll}} \simeq 1.7 \cdot 10^{16}$ cm for MP$_{\mathrm{E54.5}}$ using $\delta t_{\mathrm{var}} = 1.13$ s, in rough consistency with the $\langle R_{\mathrm{Coll}} \rangle = 2.4 \cdot 10^{16}$ cm obtained earlier (see also Table 2).

For the purpose of studying an event with smaller dissipation radius we introduce a modified version of MP$_{\mathrm{E54.5}}$ that has the same Lorentz factor distribution which is however ejected over a smaller engine active time t_{eng}. This yields a shorter variability timescale δt_{var}. Here $\delta t_{\mathrm{var}} \equiv t_{\mathrm{eng}}/N_{\mathrm{osc}}$, where $N_{\mathrm{osc}} = 30$ is the number of short-timescale oscillations introduced in the initial Lorentz factor distribution (see Figure 1). The corresponding parameters are also listed in Table 1. The distribution of collisions is similar to Figure 2 except for being shifted to smaller radii by a factor 10. The typical collision radius in this case is $\langle R_{\mathrm{Coll}} \rangle \sim 2.4 \cdot 10^{15}$ cm.

4. LEPTONIC MODELING OF SP$_{\mathrm{E54}}$

We begin our investigation with a leptonic radiation model for SP$_{\mathrm{E54}}$, as this is the simplest scenario and is widely used under the name "synchrotron self-Compton model". The structure is further representative for a single pulse in a light curve composed of several pulses. We show results for the SYN- and the IC-dominated scenario introduced previously (see also Table 2). To obtain a better understanding of the role of physical processes in shaping the overall spectrum, we commence with the presentation of results for the representative collision. We then illustrate how the full spectrum is built up from the contributions of single collisions. Finally we examine the predicted emission from the full burst on the relative contributions of different radiative processes (e.g. synchrotron and inverse Compton radiation from different

particle populations) and the time at which they become observable.

4.1. Results for the representative collision

For a better understanding of the physical processes at play we first present the modeling results for the representative collision (that corresponds to the point where the dissipated energy is maximal) for a SYN-dominated and a IC-dominated scenario. The collision parameters can be inferred from Figure 2 where the representative collision is marked with a star. The emitted (decomposed) spectrum and the loss rates for leptons and photons are shown in Figure 3.

The comoving emitted photon spectra (left plots of Figure 3) are decomposed into different components based on the radiating particle species (primary electrons or secondary lepton pairs from $\gamma\gamma$-annihilation) and on the radiation process (synchrotron or inverse Compton). The inverse Compton component of a particle distribution is defined as the emission produced by those particles scattering off the complete photon distribution. In the SYN-dominated case (of strong magnetic fields) the primary synchrotron component dominates the overall broadband spectrum with contributions of secondaries being small appearing at low energies and at the high end of the spectrum. On the other hand, in the IC-dominated case (of weaker magnetic fields) inverse Compton radiation of both primary and secondary leptons significantly reshapes the broadband spectrum by increasing the radiative output at high photon energies (i.e. for $E_\gamma' > m_e c^2$). Because electrons cool more efficiently via inverse Compton scatterings, the peak photon energy density of the primary synchrotron component decreases.

The plots in the middle panels of Figure 3 show the loss rates of electrons. The energy range of the injected primary distribution is overplotted for reference (grey band). Loss rates of time-dependent processes are shown at $t' \in \{0.25, 0.5, 0.75\} t_{\mathrm{dyn}}'$ with lighter (darker) colors referring to early (late) times. Naturally, the weaker magnetic field for the IC-dominated case demands a larger minimum electron Lorentz factor to reproduce the same synchrotron peak energy (recall that the synchrotron peak energy $E_{\mathrm{syn}}' \propto \gamma_{\mathrm{e,min}}'^2 B'$). Moreover, because the synchrotron cooling rate scales as B'^{-2}, electrons can be accelerated to higher Lorentz factors for weaker magnetic fields (see Eq. 17). Consequently, the primary electron distribution is shifted to higher energies. While the inverse Compton cooling rate is subdominant at all times and energy ranges for the SYN-dominated case, it overcomes the synchrotron cooling rate for low- to medium-energy electrons for the IC-
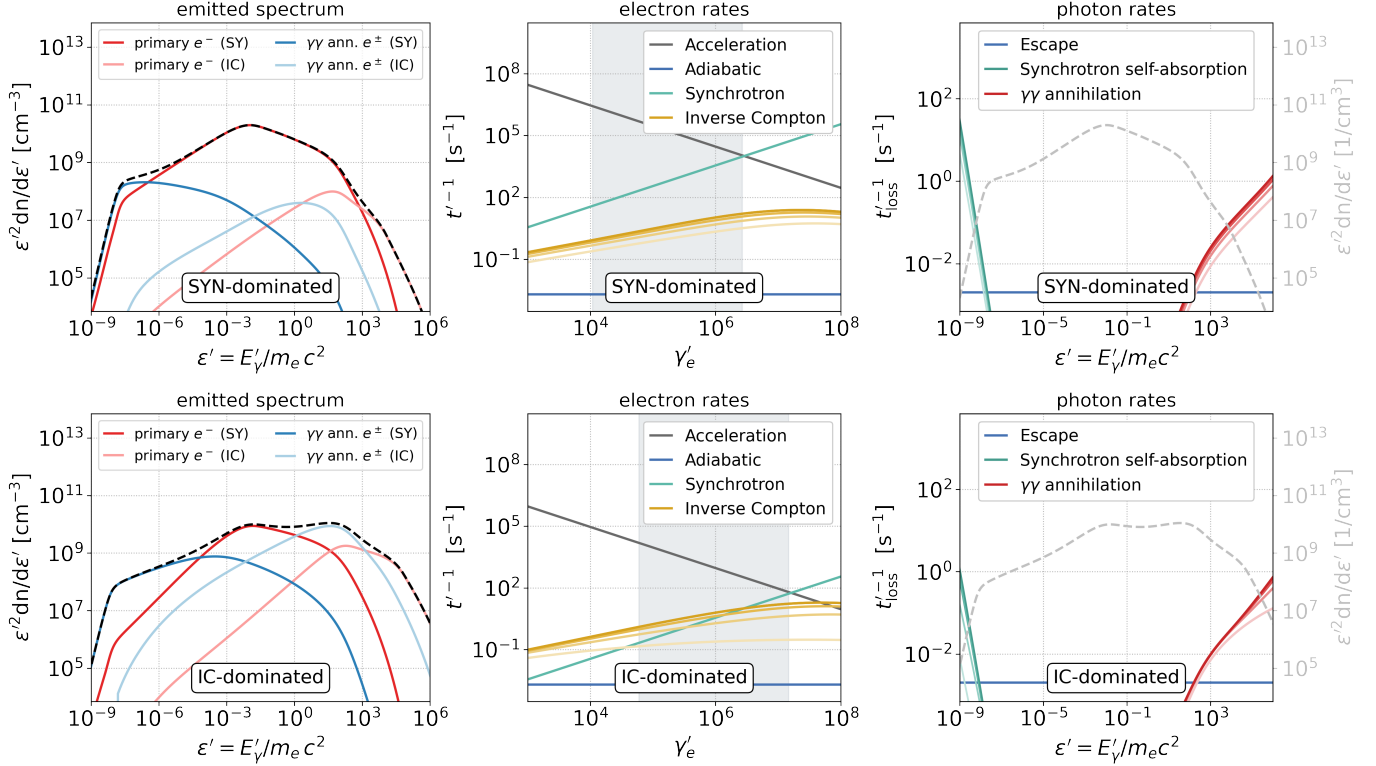
Figure 3. Examining the physical processes at play in the leptonic scenario for the representative collision of the SP$_{E54}$ prototype. We show the results for the SYN-dominated scenario (top row) and the IC-dominated scenario (bottom row). *Left:* decomposed emitted spectrum with synchrotron (SY) and inverse Compton (IC) contributions of primary electrons and secondary leptons produced in $\gamma\gamma$-annihilation, shown with different colors (see inset legends). *Middle:* loss/acceleration rates of electrons as a function of electron Lorentz factor with the energy range of injected primaries shown as a grey shaded band. *Right:* loss rates of photons with the emitted spectrum overplotted with a dashed grey line and units shown on the right y-axis. For time-dependent quantities we show the results for $t' \in \{0.25, 0.5, 0.75, 1.0\}t'_{\rm dyn}$ where light (dark) colors correspond to early (late) times. All quantities are in the comoving frame of the shocked plasma.

dominated case. Indeed, the electrons around $\gamma'_{\rm e,min}$ cool predominantly through inverse Compton scatterings for $t' \gtrsim 0.25t'_{\rm dyn}$. As these tend to occur in the Klein-Nishina regime, the dependence of the cooling rate on the electron Lorentz factor differs from the synchrotron one. As pointed out in earlier works (e.g. Daigne et al. 2011), this can lead to a synchrotron spectral shape that differs from the classical synchrotron fast- or slow-cooling predictions: the resulting spectral index is steeper, while energy extraction from electrons is still efficient.

Finally, on the right-hand side panels we show the photon loss rates as a function of photon energy. For reference, we also indicate the emitted photon spectrum (with units shown on the right y-axis). The photon fields are shaped by synchrotron self-absorption (SSA) at the lowest energies and $\gamma\gamma$-annihilation at the highest energies. The SSA dimensionless energy here is at $\sim 10^{-8}$, which in the observer's frame (assuming $\Gamma \simeq 500$) corresponds to energies of roughly 1 eV that are observed in the optical band. On the other hand,

$\gamma\gamma$-annihilation dominates over escape at a dimensionless energy of $\sim 10^{2}$ that is approximately 10 GeV in the observer's frame. Both loss rates evolve little with time. The position of the spectral breaks can be inferred from the intersection of the respective loss rates with the escape rate. We point to Appendix A for the time-evolution of photon spectra in the collision, which further illustrates e.g. the suppression of high-energy photons due to $\gamma\gamma$-annihilation.

4.2. The build-up of the full burst spectrum

From the spectrum of the representative collision we now move to the full burst spectrum. This is obtained by summing the contributions of all single collisions. This is illustrated in Figure 4, where we show the time-integrated, full spectrum (rescaled to approximately match the fluence of single collision spectra) of the IC-dominated scenario of SP$_{E54}$. Thin curves show the contribution of every 10th collision and the colour-coding indicates the observed time of the collision: light green spectra correspond to early collisions, dark blue indicate collisions to later times. The spectrum of the
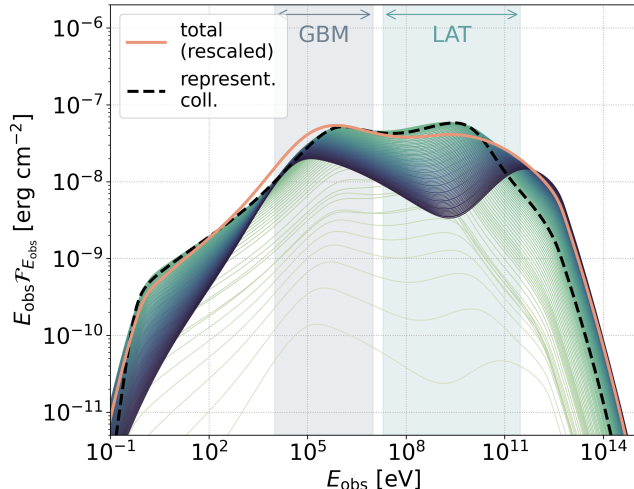
Figure 4. Decomposition of the full burst photon spectrum of SP_{E54} into the contributions of single collisions (IC-dominated). The thick orange line corresponds to the overall time-integrated spectrum that is obtained by summing the contributions of all single collisions (rescaled by the ratio of the energy dissipated in the representative collision and the total dissipated energy $E^*_{\rm diss}/E^{\rm tot}_{\rm diss}$ for easier comparison). Thin colored lines show the spectra of every 10th collision, with light green (dark blue) curves corresponding to early (late) collisions. The spectrum of the representative collision is indicated with a dashed black line. Shaded bands indicate the *Fermi*-GBM and LAT bands. All quantities are in the observer's frame assuming a redshift of $z = 2$. EBL attenuation is not included here.

representative collision discussed in the last section is indicated with a dashed line.

Although the sub-MeV synchrotron peak energy is by construction comparable in all single-collision spectra (recall that we chose $\zeta_e = $ const., which results in an almost constant $\gamma'_{e,\min}$ throughout the fireball evolution), the shape and fluence varies for the different collisions. For example, the HE (> 1 GeV) emission is largely powered by late collisions. This may be explained by differences in $\gamma\gamma$-absorption opacity in the region of the emitting plasma: close to the source, high densities increase the optical depth to $\gamma\gamma$-absorption, effectively hindering HE photons from escaping. At larger distances from the central engine, lower densities enable the escape of photons of higher energies. On the other hand, the low-energy spectrum in the eV range is mostly shaped by early collisions. The reason for this can be understood from the same reasoning: the high $\gamma\gamma$-absorption efficiency in early collisions results in a large number of secondary lepton pairs that contribute through synchrotron radiation at low energies (see also Figure 3).

Comparing the rescaled overall spectrum to the spectrum of the representative collision we find that the sub-MeV synchrotron peak of the total spectrum is broadened by the contributions of many collisions. The structures at higher energies are equally washed out, and the overall spectrum extends to higher energies than the one of the representative collision due to contributions from collisions occurring at even larger distances. These results highlight the importance of moving from single-collision to multi-collision models for an accurate description of the broadband photon spectrum.

4.3. *Full-burst decomposed spectra and light curves*

We finally proceed to evaluate the full burst results for the SYN- and IC-dominated scenarios introduced before. In Figure 5 we display the full burst spectra and light curves, decomposed in a similar manner as the representative collision before.

The contributions of the single emission processes to the full burst spectra are qualitatively similar to those of the representative collision for both scenarios (compare to Figure 3); we point out that the superposition of inverse Compton emission from many collisions leads to a relatively flat (i.e. $\propto E^{-2}_{\rm obs}$) HE spectrum for the IC-dominated scenario. As for the representative collision, the synchrotron peak is decreased in fluence. This underlines that by examining the representative collision we can indeed gain some understanding for the processes that shape the full spectrum.

The lower panels of Figure 5 show the photon index as a function of observed photon energy. In the SYN-dominated case the photon index below the synchrotron peak mostly equals the fast-cooling synchrotron prediction of $-3/2$ and is softened by contributions of secondary lepton pairs only around the synchrotron self-absorption break at $\sim$eV energies. In the IC-dominated case the low-energy photon index (in the range $10 - 10^4$ eV) is determined by the synchrotron emission of both primary and secondary electrons. The primary electron synchrotron spectrum has a photon index of ≈ -1, because electrons radiating at these energies cool mainly via inverse Compton scatterings in the Klein-Nishina regime (Nakar et al. 2009; Daigne et al. 2011; Duran et al. 2012). Still, secondary electrons cool mostly via inverse Compton scatterings in the Thomson regime, thus resulting in a photon index ≈ -1.5. As a result, the combined synchrotron spectrum has a photon index which is $\lesssim -1.5$ at low energies. Our results highlight the importance of a complete radiation treatment invoking also secondary particles and their emission, which can reshape the photon spectra.
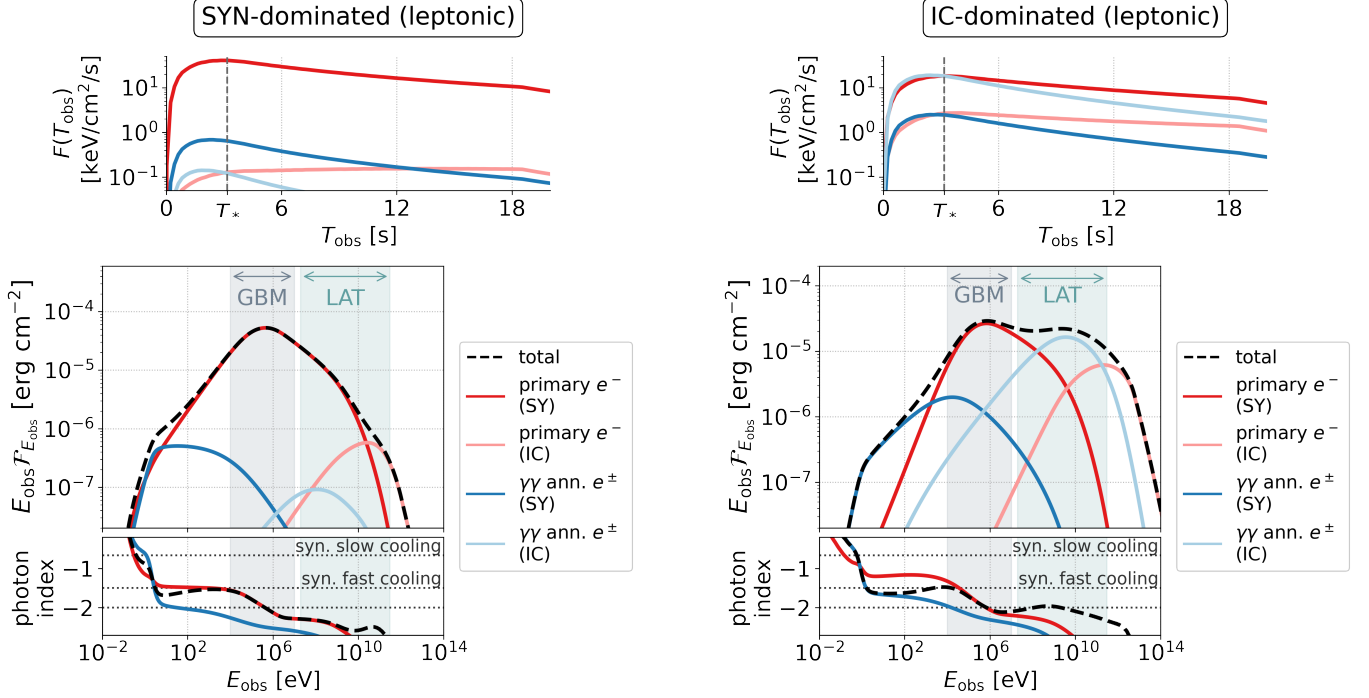
Figure 5. Full-burst decomposed light curve, spectra and photon indices for the SP$_{E54}$ leptonic model, examining *(left)* the SYN-dominated scenario and *(right)* the IC-dominated scenario. We show the synchrotron (SY) and inverse Compton (IC) of primary electrons and secondary lepton pairs from $\gamma\gamma$-annihilation. In the upper plots we show the fluence as a function of observed time for the different components. In the lower plots, shaded bands indicate the *Fermi*-GBM and LAT bands. For the photon indices in the lower panels, we only show the SY contributions of primary and secondary leptons. We further indicate the photon indices for the synchrotron slow-cooling (-2/3) and fast-cooling (-3/2) regimes as well as a photon index of -2 (that marks peaks of $E_{\rm obs}\mathcal{F}_{\rm E_{obs}}$) as dotted lines.

Besides the question *at which energy* different radiation processes contribute, one may ask *at which time* these processes dominate the overall emission. To investigate this aspect we show the time-dependent energy fluxes in the upper panels of Figure 5 for the SYN-dominated and IC-dominated scenarios. By construction, the flux of the primary synchrotron emission peaks at the observed time of the representative collision (which is when the dissipated energy is maximal). The flux created by secondary leptons (both synchrotron and inverse Compton emission) peaks at an earlier time. We recall that the flux at these early times is produced in collisions close to the source (see Figure 2), where the densities and consequently the optical depth to $\gamma\gamma$-absorption are high. The inverse Compton emission of primary electrons peaks latest. This is most evident for the SYN-dominated scenario, even though the actual contribution of this component to the overall spectrum is negligible at all times. For the IC-dominated scenario the inverse Compton emission of primary electrons peaks around the time of the representative collision T_*, while inverse Compton emission of secondary leptons peaks before T_*. As they peak at different energies, an energy-dependent shift in peak time can thus

be expected: the HE emission in the LAT range peaks *before* the keV synchrotron peak, the flux at the highest energies peaks *after* the keV synchrotron peak (see also Asano & Meszaros 2012). We point out that the early peak due to secondary leptons, was not present in Bosnjak et al. (2009) who considered scenarios of low $\gamma\gamma$-opacity.

5. LEPTO-HADRONIC MODELS

Hadronic signatures have been explored as the potential origin of the HE component that was in some cases observed by the *Fermi*-LAT (Asano & Meszaros 2012; Wang et al. 2018). In the same spirit as those studies, we investigate lepto-hadronic scenarios paying special attention to HE emission. In contrast to before, this section will include results for *both* prototypes, SP$_{E54}$ and MP$_{E54.5}$.

In the leptonic study (see previous section), no HE enhancement was found for the SYN-dominated scenario. In contrast to this, IC-dominated scenarios showed an increased HE emission, due to inverse Compton scatterings of primary and secondary leptons. Motivated by these findings, we again select the SYN- and IC-

dominated scenarios to investigate lepto-hadronic models.

As before, we will first show the results for the representative collision before moving to the emission from the full burst. While our fiducial choice of baryonic loading is $f_{\mathrm{p/e}} = 30$, we will further explore $f_{\mathrm{p/e}} \in \{0, 10, 30, 100\}$ and discuss the impact of the variability timescale for $\mathrm{MP_{E54.5}}$.
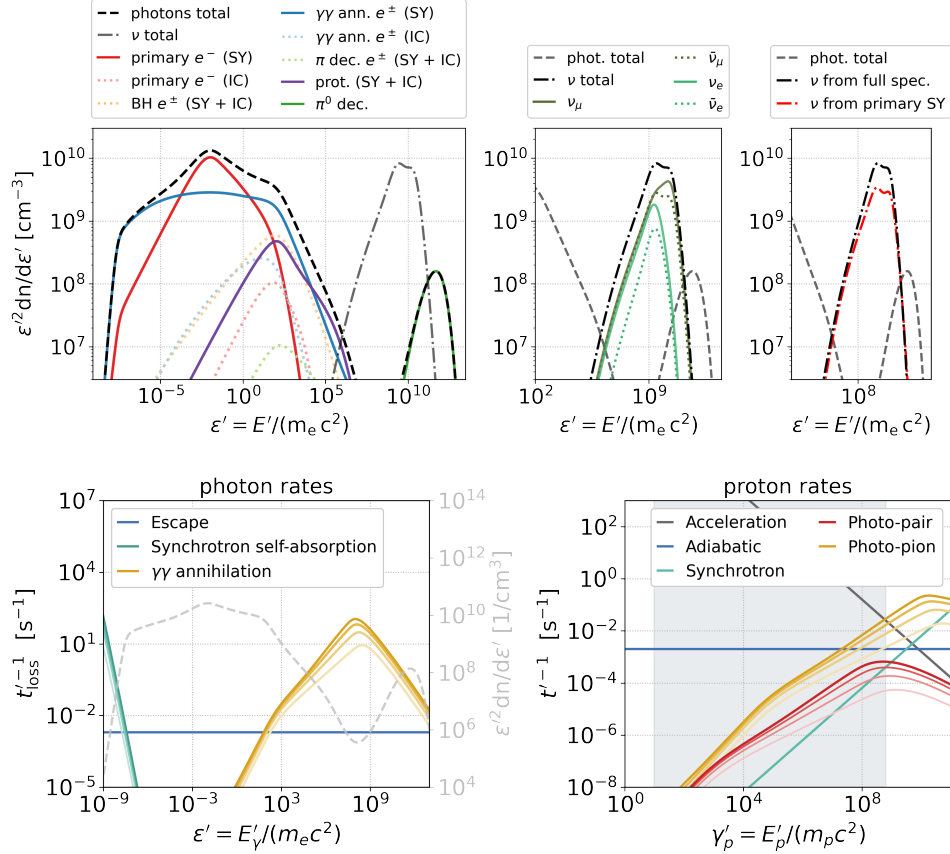
5.1. SP_{E54}: Results for the representative collision

The comoving photon and neutrino energy densities for the SYN-dominated case are shown in the upper panels of Figure 6. The left-hand side plot shows the broadband photon spectrum and its decomposition into various components: primary electrons, $e^\pm$ from photo-pair production (labelled as Bethe-Heitler, BH), secondary $e^\pm$ from $\gamma\gamma$-annihilation, $e^\pm$ from pion decays, primary protons, and photons from neutral pion decays. We indicate the synchrotron (SY) and inverse Compton (IC) emission from charged particles. Dominant contributions to the overall spectrum are shown as solid lines while sub-dominant contributions are plotted with dotted lines.

For the chosen baryonic loading, the synchrotron emission of secondary pairs from $\gamma\gamma$ annihilation follows a $n'_{\mathrm{e}}(\gamma'_e) \propto \gamma_e'^{-3}$ distribution. This creates a broad flat spectrum that dominates over all other hadronic-related contributions. Similar features were reported in earlier works on hadronic emission models for GRBs (Asano & Mészáros 2014; Petropoulou 2014; Wang et al. 2018). Contrary to the corresponding leptonic SYN-dominated scenario, the injection rate of pairs from $\gamma\gamma$ annihilation is higher in this case, because of the high luminosity of VHE photons. These photons are mainly produced from π^0 decays. The fact that the pion bump at $\varepsilon' \sim 10^{12}$ is lower in normalization than the all-flavour neutrino bump indicates the amount of attenuation; this can be also seen by the attenuation rate shown in the panel with the photon rates.

The all-flavour neutrino spectra are shown together with the photon spectra in the upper left panel, and per-flavour (ν_μ, $\bar{\nu}_\mu$, ν_e, $\bar{\nu}_e$) in the upper middle panel. The neutrino peak energy and spectral shape differ by species and depend on the parent (Baerwald et al. 2011), with ν_μ having the highest peak energies. The $\bar{\nu}_\mu$ spectrum extends to equal energies as the ν_μ spectrum, but peaks at a lower energy. The ν_e and $\bar{\nu}_e$ spectra are similar to each other and have peak energies similar to the $\bar{\nu}_\mu$ spectrum, although not extending to similarly high energies. Interestingly, the neutrino spectra peak at much lower energy than the π^0-decay photons. We attribute this difference to two effects:

- *Cooling of intermediate pions and muons.* Charged intermediate secondaries are subject to adiabatic and synchrotron cooling. For the high magnetic fields of the SYN-dominated scenario, both species cool via synchrotron radiation before decaying. This introduces a parent-dependent cooling break in the neutrino spectra, see e.g. Waxman & Bahcall (1999); Lipari et al. (2007); Baerwald et al. (2011, 2012); Tamborra & Ando (2015); Bustamante & Tamborra (2020), and reduces their energy with respect to the parent proton energy. The longer decay time of muons and their larger synchrotron cooling rate enhance this effect on the spectra of electron neutrinos (produced in muon decays). This effect and its manifestation throughout the fireball evolution are discussed in detail in Appendix C.

- *Attenuation of photons vs. neutrinos.* Neutrinos as free-streaming particles reflect the in-source distribution throughout the complete evolution. This is different for photons, which are subject to $\gamma\gamma$-annihilation. In a single collision, HE photons above $\sim$GeV can escape more easily at small t' when the radiation densities are still low (see rates in the lower panels of Figure 6). At these early t', the low radiation densities also result in a low pion-production efficiency. This in turn enables large proton maximal energies, which evolve to lower energies as the densities increase (see proton spectra and photo-pion rates in Figure 6). Thus, the escaping VHE photon spectrum is dominated by early in-source spectra with high maximal proton energies, whereas the neutrino spectra capture the complete evolution up to late t' when the maximal proton energies are lower. For further illustration of this, we also point to Appendix A and Figure 20, where we show the time evolution of single collision spectra.

It is also interesting to note that the neutrino production can be higher (and the spectra peak at higher energies) than in the simplest one zone models. First of all, cooling effects on the secondaries are not as strong as for models with smaller collision radii, which means that the cooling breaks are at comparatively high energies. Then, consider that the spectral index below the spectral peak in the synchrotron fast cooling regime (-3/2) is softer than the one typically inferred from observations (-1). This implies that the dominant energy relevant for the pion production, i.e., the energy where the photon number density peaks, can be found below the photon break at $\epsilon'_{\gamma,\mathrm{eff}} = \max\left[y_\Delta\, m_p/(20\, E'_{\nu,\mathrm{peak}}), \epsilon'_{\gamma,\mathrm{min}}\right]$ (following

Figure 6. Examining the physical processes at play in a lepto-hadronic scenario for the representative collision of SP_{E54} (SYN-dominated, $f_{p/e} = 30$). *Upper panels:* decomposed comoving spectrum of photons *(left)* and neutrinos *(middle, right)*. The red dashed-dotted in the right plot corresponds to the neutrino spectrum from the interactions of protons with the primary electron synchrotron component only. *Lower left panel:* Loss rates of photons with the emitted spectrum over-plotted with a dashed grey line and units shown on the right y-axis. For time-dependent quantities we show the results for $t' \in \{0.25, 0.5, 0.75, 1.0\} t'_{\rm dyn}$ with progressively darker colors. *Lower right panel:* Loss/acceleration rates of protons as a function of the proton Lorentz factor with the energy range of injected protons shown as a grey shaded band.

case 2 in Fiorillo et al. (2021)). Here $y_\Delta \simeq 0.2 - 0.5$ GeV is a suitable choice for the pitch angle-averaged cross section (see Fig. 4 in Hummer et al. (2010)) and $\epsilon'_{\gamma,\min}$ corresponds to the synchrotron self-absorption cutoff in our case.[3] Here we are limited by the first term, as can be seen in Figure 6, lower right panel: the photo-pion rate still increases monotonously even beyond the maximal proton energy, and the synchrotron self-absorption induced break lies significantly beyond. In order to compare to Figure 6, upper left panel, we can derive the corresponding parameter $\varepsilon'_{\gamma,\rm eff} \simeq 3 \cdot 10^{-5}$ (for $\varepsilon'_{\nu,\rm peak} \simeq 10^9$ and $y_\Delta \simeq 0.5$ GeV); this value is in the region where the blue (electromagnetic cascade) and red (primary synchrotron) curves intersect – about two orders of magnitude in energy below the peak. Thus there are two

effects which enhance the neutrino production: a) the photon number density at that point is about an order of magnitude higher than for a spectral index -1, and b) photons from the secondary cascade enhance the neutrino production by a factor of a few. We show the latter effect in Figure 6, upper right panel, where the contribution from the primary synchrotron spectrum is shown separately; here the effect is about an additional factor of 2-3 enhancement, where the additional impact of the secondary cascade depends on the hardness of the primary spectrum and the baryonic loading.

5.2. SP_{E54}: Full-burst decomposed spectra and light curves

The simulated full-burst photon and all-flavour neutrino spectra, as well as the corresponding light curves for the SYN-dominated and IC-dominated scenarios are shown in the left and right panels of Figure 7, respectively. Coloured lines indicate the various components

---

[3] Compared to Fiorillo et al. (2021), the neutrino peak neutrino energy has been used instead of the maximal proton energy, since cooling effects of the secondaries can be included in that way.

that make up the total spectrum (see inset legend for details).

Again commencing with a discussion of the spectra, we find that the spectral features of the full-burst SYN-dominated spectrum are similar to those of the representative collision discussed in the previous section (see also Figure 6). On the other hand, the neutrino peak properties relative to the photon peak are slightly different: First, the fluence of the neutrino peak relative to the photon peak fluence is ~ 45 % higher for the representative collision than for the complete burst. Also, in the representative collision the neutrino peak energies are ~ 25 % lower than for the full burst. Thus, if we scale up the neutrino spectra from the representative collision to the full burst we *overestimate* the fluence, while we *undererstimate* the peak energy. Both these effects increase the detection perspectives by instruments like IceCube, or, for non-detection, increase potential conflicts with neutrino limits.

In the IC-dominated case the broadband spectrum differs from the SYN-dominated case, and resembles that of the pure leptonic scenario (compare to Figure 5). Because of the lower $f_{B/e}$ value, pairs injected by $\gamma\gamma$-annihilation are predominantly cooling via inverse Compton scatterings. As a result, the associated inverse Compton component is much brighter than their synchrotron component, and potentially modifies the spectrum in the *Fermi*-LAT energy range. For the selected parameters, the secondary inverse Compton emission again outshines the primary inverse Compton component. The VHE peak, which is associated with the π^0 decays, has a much lower peak fluence than in the SYN-dominated case. We attribute this to two effects: Firstly, a lower maximum proton energy (which can be inferred from the lower peak energy of the pion bump) results in a reduced pion production efficiency. This lower pion production efficiency is also reflected in the lower neutrino fluxes. The lower maximum proton energy is driven by the slower acceleration in the weaker magnetic field, whereas the dominant loss processes are independent of the magnetic field (in contrast to electrons, where the weaker magnetic field for the IC-dominated scenario enables higher $\gamma_{e,max}$). Second, the opacity to $\gamma\gamma$ annihilation is higher around the VHE peak due to the lower peak energy (compare to Figure 6 lower left). This is indicated by the higher difference in energy flux of neutrinos and γ-rays when compared to the SYN-dominated case.

The temporal evolution of the observed fluxes of various components in the SYN- and IC-dominated scenarios is shown in the upper panels of Figure 7. It is useful to recall at this point that small T_{obs} cor-

respond to small collision radii R_{Coll}, small shell volumes, and high particle densities (see Figure 2). Starting with the SYN-dominated, we find that the primary electron synchrotron flux peaks at T_* as expected; the dissipated energy, a fraction of which is transferred to primary electrons, becomes maximal at this time. However, the synchrotron emission of secondary pairs from $\gamma\gamma$-annihilation and the neutrino emission reach their maximum flux at earlier times. This early emission originates closer to the central engine where radiation densities are higher. This naturally enhances the efficiency of density-dependent processes, such as $\gamma\gamma$-annihilation and photo-pion production. While the latter process is more efficient at earlier times, the photon flux from π^0 decays peak a little later, when the low-energy photon densities decrease, thus leading to a suppression of the in-source $\gamma\gamma$-annihilation rate. These results are in agreement with the findings of Bustamante et al. (2017), where neutrinos were found to originate from small radii (where the optical thickness to photohadronic interactions is high) and VHE γ-rays from large radii (where the $\gamma\gamma$ optical thickness is low).

Similar trends are found in the IC-dominated case, except for earlier peak time of the π^0 photon flux. We recall that π^0 flux depends on both the $\gamma\gamma$-annihilation and the pion production efficiency. In the SYN-dominated scenario the early flux is suppressed by $\gamma\gamma$-annihilation and thus peaks at later times. On the other hand in the IC-dominated scenario the pion production efficiency in late collisions is low, which suppresses the π^0 photon flux at late times.

5.3. *Investigating different baryonic loadings*

We continue by a systematic study of different baryonic loadings $f_{p/e}$ for both prototypes in the SYN- and IC-dominated scenarios. The spectra and photon indices are displayed in Figure 8. For SP_{E54} we explore $f_{p/e} \in \{10, 30, 100\}$, and for $MP_{E54.5}$ that has a higher isotropic energy $f_{p/e} \in \{3, 10, 30\}$. For comparison we further show the leptonic modeling results.

We observe that increasing $f_{p/e}$ leads to similar trends for both prototypes, both in the SYN- and the IC-dominated scenario. We recall that the typical emission radii are similar, however $MP_{E54.5}$ has a slightly higher E_{iso} than SP_{E54}. This implies higher energy densities which enhance the efficiency of processes such as photo-pion production and $\gamma\gamma$-annihilation. As a consequence, for the same baryonic loading the signatures of secondary particles are slightly more pronounced for $MP_{E54.5}$ than for SP_{E54}. As the differences are low we however conclude the more complex structure of $MP_{E54.5}$ thus has little effect on the predicted time-
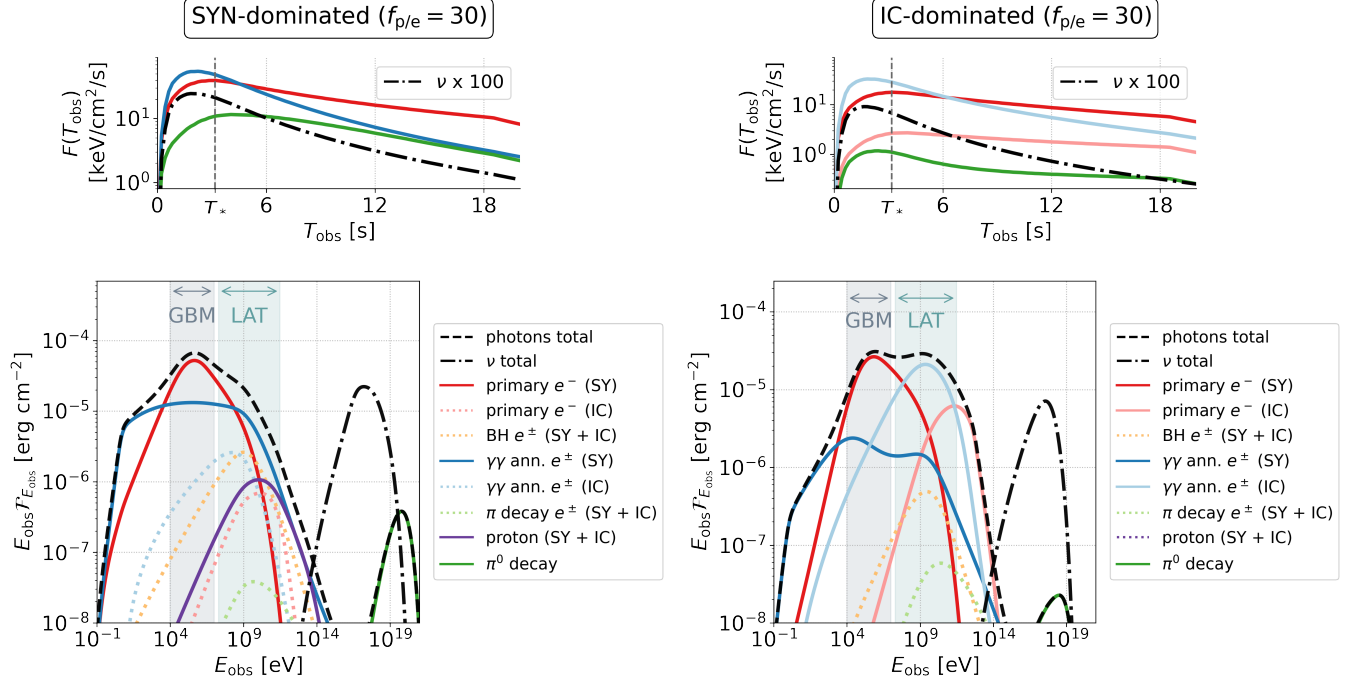
Figure 7. Full-burst decomposed light curves and spectra for the SP_{E54} lepto-hadronic model, examining *(left)* the SYN-dominated scenario and *(right)* the IC-dominated scenario with $f_{p/e} = 30$. Coloured lines show various contributions to the total spectrum which is plotted with dashed line (for details, see legends). The all-flavour neutrino fluences and fluxes are overplotted with dash-dotted black lines. For the light curves, the neutrino fluxes we scaled up by a factor 100 to match the same scale. Shaded regions indicate the energy ranges of the *Fermi*-GBM and LAT detectors. In the energy flux light curves the dashed vertical line indicates the observed time of the representative collision, marked with a star in Figure 2.

integrated spectra. Their properties are instead mostly controlled by the combination of $\langle R_{\text{Coll}} \rangle$ and E_{iso}.

Let us now discuss selected features of the photon spectra, starting with the SYN-dominated scenarios. As stated earlier, the flat additional component due to secondary synchrotron radiation introduces a wing-like broadening to the sub-MeV peak. This feature is common to both prototypes and increases in intensity for higher $f_{p/e}$. This additional component outshines the primary emission for $f_{p/e} = 100$ for SP_{E54} ($f_{p/e} = 30$ for $MP_{E54.5}$). The flat spectrum is reflected in the photon indices that are smaller than -3/2 (that is the synchrotron fast cooling prediction). For the IC-dominated case the spectra below the peak are similar for all $f_{p/e}$ explored. For the highest $f_{p/e}$ an additional peak appears in the LAT range. Overall the spectra for the IC-dominated scenarios remain mostly unaffected by the baryon loading except for the VHE π^0 peak.

To put these results in context with observations, we review the shape of the spectra in the GBM band. Measured GRB spectra are typically narrower than what would be expected from the simple synchrotron spectrum, as can be inferred from their spectral width (see Axelsson & Borgonovo (2015); Yu et al. (2015) but also the critical discussion in Burgess (2019)). Without ex-

plicitly measuring the spectral width in our models, we can by eye identify which models produce narrow or wide spectra. We compare the leptonic (dotted) to lepto-hadronic (solid) lines in Figure 8 and find that a leptonic SYN-dominated fast-cooling spectrum is the narrowest spectrum that can be achieved in our models. This is also the model with the highest photon index below the peak and lowest photon index above the peak (see lower panels).

The neutrino flux scales with the baryonic loading for all models, and all peak around 10^{18} eV. This is somewhat above the typical sensitivity of IceCube and thus radio neutrino detectors such as GRAND (Álvarez-Muñiz et al. 2020) may instead be the adequate observatories for GRB prompt phase neutrinos.

5.4. *Impact of variability timescale*

The two prototypes $MP_{E54.5}$ and SP_{E54} have rather similar typical collision radii and energy budget, hence a similar typical energy density. In Section 3 we further introduced a modified version of $MP_{E54.5}$ which has shorter duration and consequently shorter variability timescale (while using the same Lorentz factor distribution). We remind the reader that in our model, δt_{var} was defined in the *source frame* as the variability timescale of the Lorentz factor distribution and thus tracks the
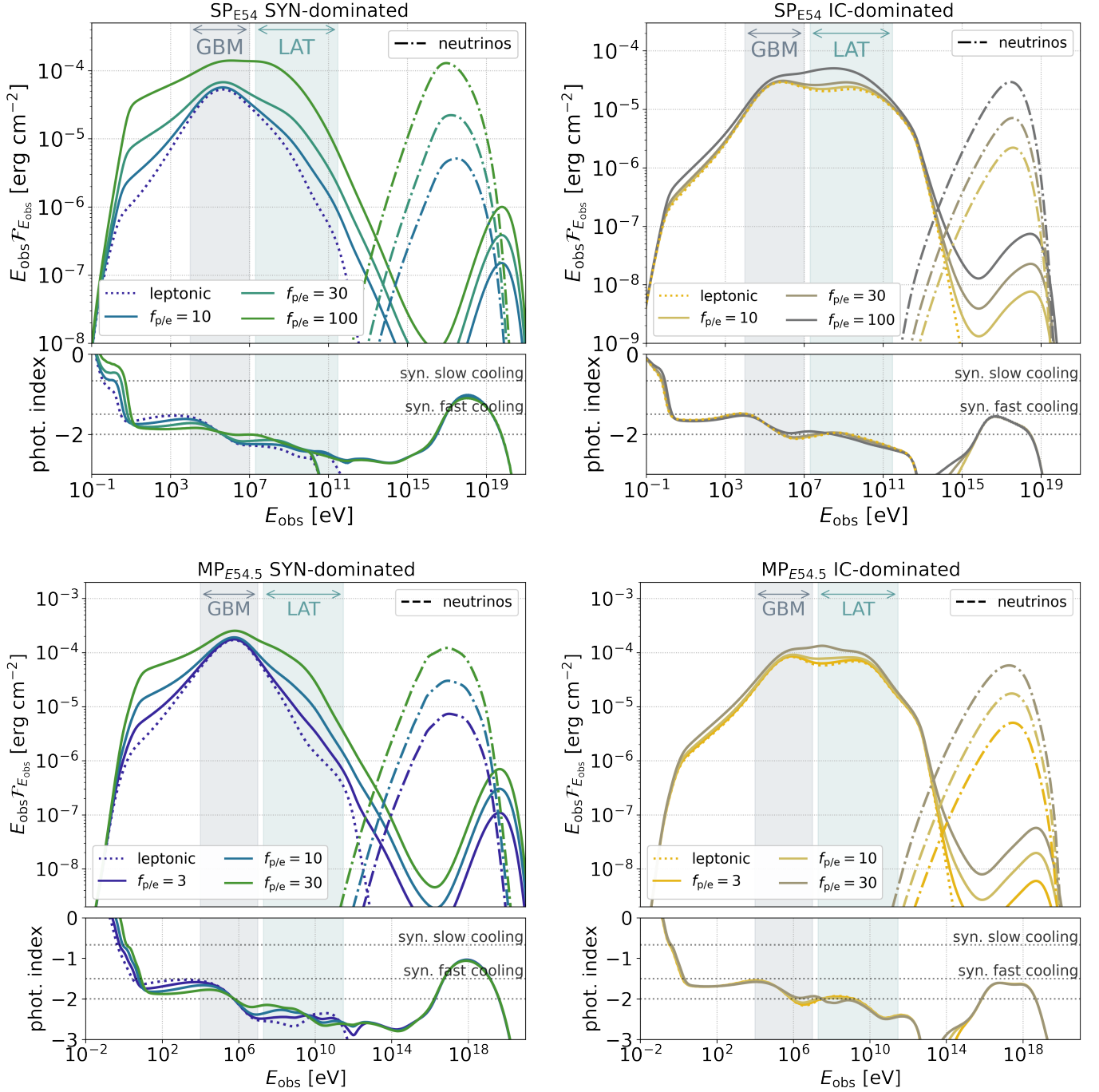
Figure 8. Lepto-hadronic spectra $E_{\rm obs}\mathcal{F}_{E_{\rm obs}}$ and photon indices for $\rm SP_{E54}$ (top panel) and $\rm SP_{E54}$ (bottom panel), examining *(left)* the SYN-dominated scenario and *(right)* the IC-dominated scenario. For all scenarios we show the leptonic case and explore $f_{\rm p/e} \in \{10, 30, 100\}$ ($f_{\rm p/e} \in \{3, 10, 30\}$) for $\rm SP_{E54}$ ($\rm MP_{E54.5}$). Dash-dotted lines mark the corresponding all-flavour neutrino fluences. For the photon indices, we indicate the synchrotron slow- and fast-cooling predictions as dashed lines and a photon index of -2 (that marks peaks of $E_{\rm obs}\mathcal{F}_{\rm E_{obs}}$) as a solid line.

variability time of the engine. Reducing $\delta t_{\rm var}$ moves the distribution of collisions inwards by a factor 10 and gives a typical collision radius of $\langle R_{\rm Coll} \rangle \sim 10^{15}$ cm.

In Figure 9 we show the results for this modified version of MP$_{\rm E54.5}$ in orange/brown colors, comparing to the results of MP$_{\rm E54.5}$ shown the last paragraph in blue colors. As before we show leptonic and lepto-hadronic models, for the latter imposing a baryonic loading of $f_{\rm p/e} = 3$.

Let us first discuss the leptonic results, i.e. comparing the blue and the orange dotted lines in Figure 9. Here, two main effects arise with decreasing $\delta t_{\rm var}$: the synchrotron self-absorption frequency increases, and the suppression in the LAT-range is enhanced. Both effects can be explained by the higher densities because of the lower collision radii that increase the efficiency of internal absorption processes. On the other hand, the spectrum in the GBM band remains unaffected. For lepto-hadronic models we additionally find that the π^0 decay peak is decreased in peak energy and intensity. This may be understood through the higher pion production efficiency (arising due to the higher densities) for short $\delta t_{\rm var}$, which reduces the maximal proton energy. This, in turn, lowers the π^0-decay peak energy. The higher pion production efficiency is further reflected in the neutrino fluxes that (for the same baryonic loading) are enhanced almost by a factor of 10 with respect to the longer $\delta t_{\rm var}$ (compare e. g. solid lines); this can be also shown analytically, see Eq. 28 (second part) later. Moreover, the neutrino spectra contain additional features/breaks which we attribute to cooling effects of primary and secondary particles.

Overall, reducing $\delta t_{\rm var}$ increases the intensity of the hadronic signatures discussed in the last sections, and a lower baryonic loading may be sustained. We point out that a similar effect can be expected by decreasing the Lorentz factor of the outflow (see also Appendix B where this is discussed for leptonic scenarios). In this sense, accurate estimates for the Lorentz factor of the outflow and the source-frame $\delta t_{\rm var}$ are essential when studying IC-dominated or lepto-hadronic models that have large contributions by density-dependent processes to the broadband emission.

6. IMPLICATIONS FOR MULTI-MESSENGER ASTROPHYSICS

Here we discuss the implications of our model for multi-messenger astrophysics. We examine first the dominant production sites of the different messengers (photons, neutrinos, UHECRs) in a GRB jet and check if these signals are time-correlated, which is relevant to neutrino alert follow-ups (Section 6.1). One of the most

uncertain parameters of lepto-hadronic models is the baryonic loading. We discuss what is the maximum allowed value by considering its impact on the broadband GRB photon spectrum (Section 6.2) and by using constraints from neutrino observations (Section 6.3). Upper bounds on the baryonic loading can also limit the predicted UHECR flux per GRB. We therefore examine the requirements for a sub-population of energetic GRBs, similar to those considered in this work, to be the source of UHECRs (Section 6.4).

6.1. *Production regions and observation times of different particle species*

Spatially resolved multi-collision models offer the possibility of investigating the production regions of different particle species (e.g. Bustamante et al. 2017; Rudolph et al. 2020; Heinze et al. 2020). In Figure 10 (lower panel) we show the energy carried by escaping neutrinos, photons (of all energies) and HE γ-rays (with energy above 1 GeV in the source frame) as a function of collision radius for the SYN- and IC-dominated scenarios assuming $f_{\rm p/e} = 30$. Results are shown for the SP$_{\rm E54}$ burst. The bolometric photon curve is similar to the radial profile of energy dissipation for this burst, which peaks around $R_{\rm Coll} \sim 10^{16}$ cm (see Figure 2). In comparison to this, the production rate of neutrinos is enhanced at the smallest radii. In other words, most of the neutrino energy is emitted at smaller radii where the photon densities are higher, even though the dissipated energy transferred into relativistic particles peaks further out. While in the lepto-hadronic model HE γ-rays are also produced via photomeson interactions, their energy is suppressed at the smallest radii due to internal $\gamma\gamma$ attenuation. The radial profiles of all messengers are similar for both SYN- and IC-dominated scenarios. Note however that the energy emitted as HE γ-rays in the IC-dominated scenario is enhanced compared to the SYN-dominated scenario, as demonstrated also in Figure 7. We point out that the range of collision radii is comparatively small (see for example the results for MP$_{\rm E54.5}$ in Figure 11), hence the differences in the location of different messengers is less pronounced compared to past publications. The upper panel of Figure 10 further shows the maximal proton energy (in the source frame) as a function of radius, as obtained by balancing the acceleration rate with the adiabatic and synchrotron loss rates. The displayed values are strictly speaking an upper bound on $E'_{\rm max,prot}$, since photo-hadronic energy losses, which for simplicity are not included here, may push the maximum energy to even lower values (see Fig. 18). The lower magnetic fields involved in the IC-dominated scenario translate to lower acceleration rates
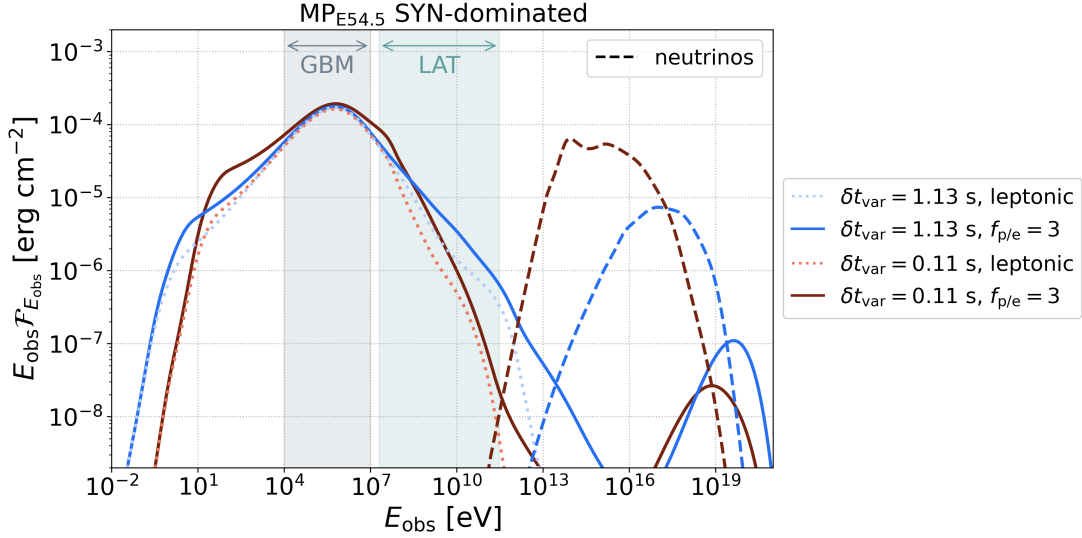
Figure 9. Time-integrated spectra for the $\mathrm{MP_{E54.5}}$ prototype (SYN-dominated scenario) illustrating the effects of the variability timescale, δt_{var} (see side-panel legend for details).

which result in lower maximal energies. For the SYN-dominated case, the high synchrotron cooling rate at the smallest radii reduces the maximal proton energy. Overall, our maximal proton energies are of the order of 10^{20}-10^{21} eV in the source frame which is sufficient to explain UHECRs, as expected for models with large collision radii and large Lorentz factors (see also Samuelsson et al. 2019).

We further investigate the production regions of neutrinos and HE γ-rays for two different variability timescales using the $\mathrm{MP_{E54.5}}$ burst (see Figure 11). As in Figure 9, we choose two engine active times that differ by a factor 10 while keeping the same Lorentz factor distribution. This setup yields variability timescales that differ by a factor of 10 in the source frame. The shorter variability timescale translates to smaller dissipation/emission radii, which can be seen by comparing the neutrino radial profiles (dotted blue and black lines in Figure 11). Besides the shift in radius, the two profiles are very similar because of the same underlying Lorentz factor distribution. As discussed before, smaller dissipation radii realised for the shorter variability timescale enhance the pion production rate and, in turn, the energy released in neutrinos. As a result, the energy emitted in neutrinos at the peak for $\delta t_{\mathrm{var}} = 0.11$ s is roughly a factor 10 higher than the maximum energy found for the longer variability burst. The impact of the variability timescale is stronger on the HE γ-ray radial profile, which is suppressed at all radii for $\delta t_{\mathrm{var}} = 0.11$ s (compare solid blue and black lines). Differences in the opacity of the emitting region to $\gamma\gamma$ pair production can also be inferred by comparing the neutrino (not attenuated)
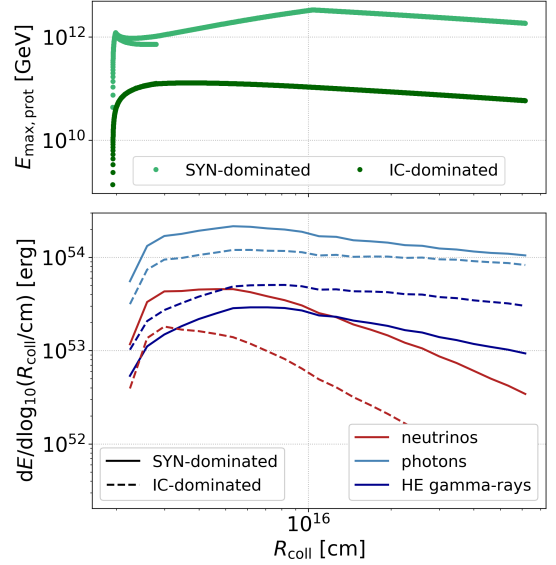
Figure 10. Spatially resolved properties of different messengers in the SYN- and IC-dominated scenarios for the $\mathrm{SP_{E54}}$ burst. *Top panel*: Upper limit on the maximum energy of protons obtained by balancing adiabatic and synchrotron losses with the acceleration time as a function of radius. The true maximal energy may be lower due to photohadronic interactions that are neglected here. *Bottom panel*: Emitted energy in different particle species as a function of radius for $f_{\mathrm{p/e}} = 30$. Here we define HE gamma-rays as those photons that have $E_\gamma \gtrsim 1$ GeV in the source frame. Recall that the deceleration radius for the hadronic case is $R_{\mathrm{dec,had}} \sim 3 \cdot 10^{17}$ cm.

and HE γ-ray curves for different variability timescales (compare solid and dotted curves).

We discuss next the time-dependence of the emission from the different messengers. UHECRs typically can-
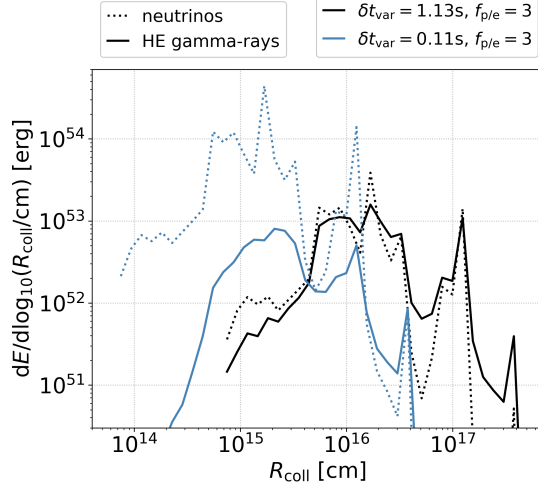
Figure 11. Effects of the variability timescale on the radial profile of emitted energy in neutrinos and HE γ-rays for the $MP_{E54.5}$ burst (SYN-dominated scenario). Different baryonic loadings are used to produce similar peak neutrino energies (see legend).
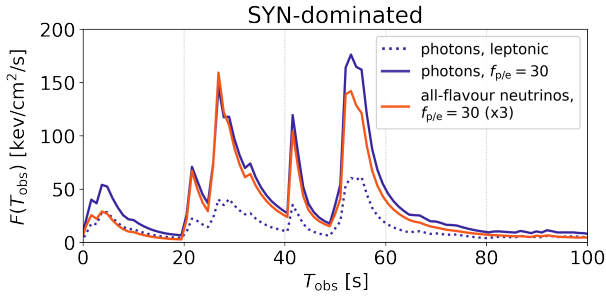
Figure 12. Energy flux for photons (integrated over the all photon energies) and neutrinos (all flavours) as a function of time for the $MP_{E54.5}$ burst. We show the SYN-dom inated scenario with $f_{p/e} = 30$ (solid lines). and the corresponding leptonic model (dotted line). The neutrino fluxes are scaled up by a factor three to match the same scale.

not be directly associated to transients because of the huge time delays caused by the deflection in intergalactic magnetic fields. We therefore focus on the other two messengers, photons and neutrinos. As an indicative example, we show in Figure 12 the energy flux emitted in neutrinos of all flavours and the bolometric photon flux as a function of time (in the observer's frame) for the $MP_{E54.5}$ burst in the SYN-dominated scenario. Both messengers have similar light curves and are strongly correlated without evidence of time lags. The photon light curve is dominated by the energy range where most of the energy is emitted – i.e. around the sub-MeV peak. The light curves, however, may look different when computed for different energy bands. For example, low-energy radiation that is produced by sec-

ondary lepton pairs will be intense at early times that correspond to small radii, (see also Figure 7), while early HE emission may be suppressed due to high opacity (as discussed in Bustamante et al. 2017). Nonetheless, the main detection channel of instruments like *Fermi*-GBM corresponds to the sub-MeV peak and we expect that the usual follow-up searches between neutrinos and electromagnetic signals in either energy band apply to this model, as differences on the order of seconds to minutes between photons and neutrinos do not have any practical impact on the follow-up strategies; the approach would always be to trigger the follow-up immediately.

6.2. *Constraints on the baryonic loading from the electromagnetic cascade*

As a next step we assess how the electromagnetic cascade may be employed to infer the maximal baryonic loading that is compatible with observations. To identify the energy bands that are potentially relevant in this context we first discuss the impact of EBL absorption. We further compare the fluences in energy bands that are expected to be observationally accessible for different models, and finally test the robustness of our results by studying the impact of the proton power-law slope p_p.

EFFECTS OF EBL ABSORPTION

In Figure 13 we commence with a redshift study, placing our prototype SP_{E54} at different distances to Earth and calculating the spectra with EBL absorption. We only show HE to VHE emission, since lower energies remain unaffected by the EBL absorption. The π^0-decay peak, which is an unambiguous hadronic signature (compare to the leptonic model, dotted lines), is expected at extremely high energies (> 10 EeV). Regardless of the degree of EBL attenuation, this feature is beyond the observing window of the Cherenkov Telescope Array (CTA) and hence not accessible. As a result, the baryonic loading of a GRB jet has to be inferred from lower energies which are shaped by secondary cascade emission. For $z \gtrsim 0.5$ (which are typical redshifts for energetic GRBs, see Poolakkil et al. (2021)), emission in the CTA range is strongly suppressed. Still, differences in the spectral shape in the LAT and CTA ranges are expected between the leptonic and lepto-hadronic models (depending on parameters like $f_{B/e}$). The prospects of differentiating between a purely leptonic and a lepto-hadronic model with combined *Fermi*-LAT and CTA analysis is worth investigating further.

MULTI-WAVELENGTH SIGNATURES: RELATIVE FLUENCE IN DIFFERENT ENERGY BANDS
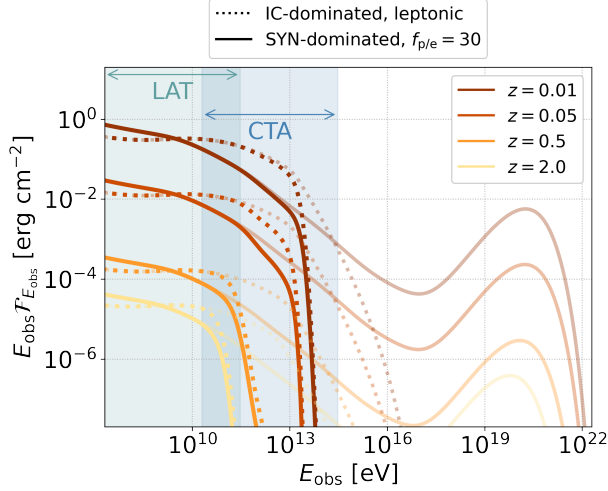
Figure 13. Examining the impact of EBL absorption on energy fluence model spectra by placing SP_{E54} at different redshifts (see inset legend). Transparent (non-transparent) curves show the results without (with) EBL absorption. Results are obtained for the EBL model of Dominguez et al. (2011).

As these are the observationally accessible energy ranges, we compare the energy fluence of our models in UV, X-rays and HE γ-rays. For each energy band we normalise with the fluence in the *Fermi*-GBM band and further give the absolute fluence that would be observed by *Fermi*-GBM. The values for the SP_{E54} burst in the SYN- and IC-dominated scenarios are listed in Table 3. The results again illustrate that the IC-dominated scenario is comparatively dim in low energies (UV and X-rays) while being bright in HE. The HE brightness has for example been analysed in Nakar & Piran (2011), who find $\mathcal{F}_{LAT}/\mathcal{F}_{GBM} \lesssim 0.1$, however larger values of up to 1 compatible with our predictions are reported for single events in Ajello et al. (2019). Further, the table illustrates that for IC-dominated case leptonic and lepto-hadronic scenarios are unlikely to be distinguishable based solely on energetics (see previous paragraph). In contrast, hadronic contributions result in a very bright UV component for the SYN-dominated scenario. The relative brightness in HE of the IC-dominated scenario cannot be recovered even with hadronic contributions. We thus conclude that a joint observation by a low-energy (optical or UV) and HE instrument would offer good diagnostics to narrow down $f_{B/e}$ (that controls the SYN- or IC-dominated regimes) and the baryonic loading $f_{p/e}$.

Let us at this point briefly comment on the efficiency in lepto-hadronic models. The table illustrates that for the scenarios with $f_{p/e} = 30$, the events would be brighter especially in the optical and HE regime. Here

it is worth remembering that this comes at the cost of a much higher initial kinetic energy of the jet (see Eq. 11), and hence the total efficiency of the fireball in converting its initial kinetic energy into radiation is reduced from $\sim 5\%$ for leptonic models to $\sim 0.5\%$ for $f_{p/e} = 30$.

RELATIVE INTENSITY OF CASCADE EMISSION

A critical question in the context of lepto-hadronic models is for which baryonic loading $f_{p/e}$ the cascade emission outshines the one of primary electrons. We investigate this aspect using the SYN-dominated scenario of SP_{E54} as an example. The ratio of fluences produced by primary electrons $\mathcal{F}_{primary}$ (sum of light and dark red lines in Figure 7) and cascade leptons produced by $\gamma\gamma$-annihilation $\mathcal{F}_{cascade}$ (sum of light and dark blue lines in Figure 7) in the *Fermi*-GBM and LAT range are shown in Figure 14 for different pairs of $(f_{p/e}, p_p)$ values. EBL absorption is neglected. The fluences are computed as

$$\mathcal{F} = \int_{E_{min}}^{E_{max}} \mathcal{F}_{E_{obs}} dE_{obs}. \tag{24}$$

where $\mathcal{F}_{E_{obs}}$ is the differential fluence, $E_{min} = 10$ keV (20 MeV), and $E_{max} = 10$ MeV (300 GeV) for the *Fermi*-GBM (LAT) instruments. The horizontal line indicates where the primary and cascade emission are equally intense ($\mathcal{F}_{cascade} = \mathcal{F}_{primary}$), but noticeable distortions of the primary synchrotron spectrum may be expected for $\mathcal{F}_{cascade} \gtrsim 0.5\,\mathcal{F}_{primary}$.

In purple we show the results for the classical proton power-law index of $p_p = 2.0$ (recall that the injected proton distribution follows $dn'_p/d\gamma'_p \propto \gamma'^{-p_p}_p$). In addition we show $p_p = 1.5$ and $p_p = 2.5$ with yellow and blue symbols, respectively. For equal energy budget of non-thermal protons and minimum proton Lorentz factor $\gamma_{p,min}$, p_p determines the energy transferred to protons at the highest energies; the lower p_p is, the more energy is transferred to high-energy protons. Due to the larger relative luminosity of high-energy protons for $p_p = 1.5$, the secondary emission outshines the primary emission already for $f_{p/e} \gtrsim 10$. With increasing p_p, larger baryonic loadings can be allowed. For instance, for soft proton power-law spectra ($p_p = 2.5$), the cascade emission does not distort the primary emission in the GBM/LAT bands even for $f_{p/e} = 300$. In other words, a lot of energy is carried by protons that do not undergo photo-hadronic interactions. The linear increase (in log-log space) is common to all choices of p_p and may therefore be interpreted as a general feature. For the usually adopted value of $p_p = 2$, the secondary emission outshines the primary one in the GBM band for baryonic loadings of $f_{p/e} \gtrsim 100$ in this SYN-dominated scenario. We again stress that this will

Table 3. Fluence in the *Fermi*-GBM range between 10 keV and 10 MeV and relative fluence for different instrumental energy ranges with respect to the *Fermi*-GBM for SP$_{E54}$.

	SYN-dominated		IC-dominated	
	leptonic	$f_{p/e} = 30$	leptonic	$f_{p/e} = 30$
Fermi-GBM $[10^{-5}\mathrm{erg/cm}^{-2}]$	9.4	12.7	5.1	5.5
ULTRASAT /*Fermi*-GBM $[10^{-4}]$	4.3	18.3	3.0	3.6
Swift-XRT /*Fermi*-GBM $[10^{-1}]$	1.2	2.1	0.85	0.87
Fermi-LAT /*Fermi*-GBM $[10^{-1}]$	2.4	4.6	12.0	14.1

Notes.− We use the energy ranges of ULTRASAT (as an example for a UV-observatory, 220 - 280 nm), *Swift*-XRT (as an example for an X-ray observatory, 0.2 - 10 keV) and *Fermi*-LAT (as an example for a γ-ray observatory, 20 MeV - 300 GeV).
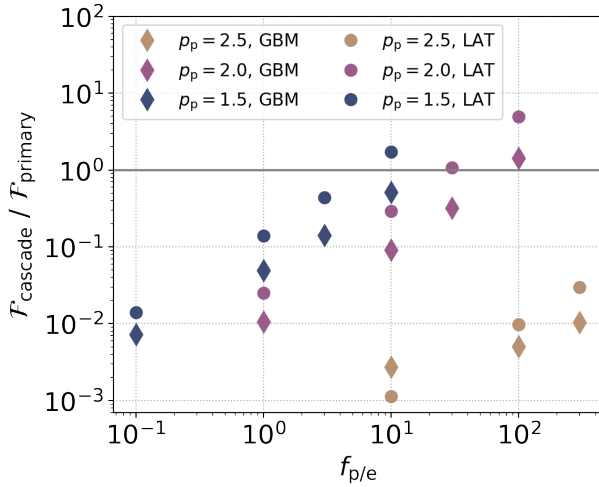
Figure 14. Ratio $\mathcal{F}_{\mathrm{cascade}}$ and $\mathcal{F}_{\mathrm{primary}}$ for SP$_{E54}$ (the integrated fluence produced by cascade/ primary leptons, see Eq. 24) in the GBM and LAT band for different values of $f_{p/e}$ and primary proton power-law slopes p_p. All calculations are performed in the SYN-dominated scenario, the solid line indicates where the fluence of the cascade equals the one of primary leptons.

not be true in an IC-dominated case where the spectra in GBM and LAT energy bands are not much affected unless $f_{p/e} \gg 100$.

6.3. *Constraints on the baryonic loading from neutrinos*

POINT SOURCE AND MULTIPLET CONSTRAINTS

No neutrinos associated with individual GRBs have been detected so far (Abbasi et al. 2012; Aartsen et al. 2017; Abbasi et al. 2022). Meanwhile, the energetic GRBs considered here are very luminous out in γ-rays, i.e., they could have been detected in most cases, see e.g. Fig. 1 in Kistler et al. (2009). Depending on the baryonic loading factor, they could also be bright neutrino

emitters (see e.g. our Figure 8). Therefore, we can place constraints on $f_{p/e}$ for an individual GRB by requiring that significantly less than one event is expected, similar to constraints derived for GRB 130427A in Gao et al. (2013), and for GRB 160625B in Fraija et al. (2017). We compute the number of expected neutrinos for each example integrating over the full energy range of neutrinos and with the "most favorable" point-source effective area, A_{eff}, between 0 and 2.29 degrees (as presented in Abbasi et al. 2021).

We find that the number of expected neutrinos per flavour for MP$_{E54.5}$ (SYN-dominated scenario) is $N_\nu \sim 3.7 \cdot 10^{-2}$ for $f_{p/e} = 10$, and for SP$_{E54}$ (SYN-dominated scenario) is $N_\nu \sim 3.4 \cdot 10^{-3}$ for $f_{p/e} = 10$, which is significantly below one in both cases. Therefore, one would not expect these GRBs to stick out in neutrino point source analyses based on neutrino data alone (for the assumed redshift of $z = 2$). For MP$_{E54.5}$ and short δt_{var} we calculate the number of expected events as $N_\nu \sim 0.14$ for $f_{p/e} = 1$; here, the lower peak energy increases the detection prospects.

In some rare cases multiplets are expected even if the predicted event rate is smaller than one. For example, taking into account the whole population of sources including redshift distribution, the maximal possible neutrino output for sources with GRB-like redshift evolution per transient has been limited to $E_\nu \lesssim 10^{53}\,\mathrm{erg}$ in Aartsen et al. (2019) (Fig. 2) for a local source rate $\dot{n}_0 \simeq 0.1\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ expected here (see below). We expect a total energy output in neutrinos of $8.4 \cdot 10^{52}$ erg for SP$_{E54}$ and $f_{p/e} = 10$ and and $5.4 \cdot 10^{53}$ erg for MP$_{E54.5}$ and $f_{p/e} = 10$, which means that the baryonic loading $f_{p/e} \lesssim 10$ is roughly compatible with these transient searches. For MP$_{E54.5}$ and $\delta t_{\mathrm{var}} = 0.11$s, we find that the energy emitted in neutrinos is $3.8 \cdot 10^{53}$ erg for $f_{p/e} = 1$, which means that for these sources even a baryonic loading $f_{p/e} = 1$ would lead to a detectable neutrino signal. This highlights the role of the collision

radius, which is related to the variability timescale, in the production rate of neutrinos from a GRB.

STACKING SEARCHES

Very strong constraints on GRBs come from stacking searches (Abbasi et al. 2012; Aartsen et al. 2017), which limit the prompt phase GRB contribution to $\lesssim 1\%$ of the diffuse astrophysical neutrino flux (Abbasi et al. 2022). Since we hypothesize that only the energetic GRBs produce neutrinos and UHECRs, the corresponding limits cannot be taken literally, but are a good indicator. The first question is how large the population of energetic GRBs is, which we address in Figure 15, left panel, for the *Fermi*-GBM catalogue presented in Poolakkil et al. (2021). From that distribution, we can estimate that about 22% and 14% of the GRBs are similar (in terms of energetics) to the SP_{E54}(green shaded) and $MP_{E54.5}$ (hashed blue) prototypes, respectively. These correspond to a local energetic GRB rate $\dot{n}_0 \simeq 0.02 - 0.1\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$, as used in earlier studies (see e.g. Table 1 in Baerwald et al. (2015)).

We compute the expected diffuse fluxes from energetic GRBs for different baryonic loadings with similar methods as those used for the stacking limits, see Figure 15, right panel. We can read off the figure that $f_{p/e} \lesssim 10$ for SP_{E54} and $f_{p/e} \lesssim 3$ for $MP_{E54.5}$ are perfectly compatible with the current stacking limits, where the difference in neutrino flux mostly comes from the different $E_{\gamma,\mathrm{iso}}$ of the prototypes. For $MP_{E54.5}$ with a shorter variability timescale, already $f_{p/e} = 0.3$ violates the stacking limit. This is also due to the lower peak energies of the neutrino spectra (that arise due to strong cooling effects of secondaries). The next-generation instrument IceCube-Gen2 (Aartsen et al. 2021) may test all scenarios in stacking searches. However, we note that the observational stacking limits are derived under the assumption that all GRBs produce neutrinos. We expect that the corresponding limits obtained selecting only energetic GRBs would be a factor of a few weaker, which does not affect our qualitative conclusions. Such a selection is, however, not trivial, as the redshift of the GRBs is not always known in the stacking samples, which means that $E_{\gamma,\mathrm{iso}}$ cannot be unambiguously derived.

6.4. *Requirements for UHECR energetics*

Following Baerwald et al. (2015), the required (isotropic-equivalent) energy ejection per GRB in UHECRs between 10^{10} GeV and 10^{12} GeV to power the UHECR flux is

$$E_{\mathrm{UHECR}} \simeq 10^{53}\,\mathrm{erg} \cdot \frac{\dot{\varepsilon}_{\mathrm{UHECR}}}{10^{44}\,\mathrm{erg}\,\mathrm{Mpc}^{-3}\,\mathrm{yr}} \cdot \frac{\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}}{\dot{n}_0}, \tag{25}$$

where $\dot{\varepsilon}_{\mathrm{UHECR}}$ is the required local energy injection density to describe data in that energy range.[4] Using the baryonic loading, one can relate this energy per GRB to $E_{\gamma,\mathrm{iso}}$ by

$$E_{\mathrm{UHECR}} = E_{\gamma,\mathrm{iso}}\, f_{p/e}\, f_{\mathrm{bol}}\, f_{\mathrm{esc}}, \tag{26}$$

where $f_{\mathrm{bol}} \simeq 0.2$ (for an E^{-2} spectrum) describes how much of the non-thermal proton energy falls into the UHECR range[5], and f_{esc} describes how many of the UHECR protons can escape; we use $f_{\mathrm{esc}} \simeq 0.5$ inspired by the values obtained in Heinze et al. (2020). Combining Eq. 25 with Eq. 26, we can estimate the required baryonic loading for our population as

$$f_{p/e} \simeq 10 \cdot \frac{10^{54}\,\mathrm{erg}}{E_{\gamma,\mathrm{iso}}} \cdot \frac{\dot{\varepsilon}_{\mathrm{UHECR}}}{10^{44}\,\mathrm{erg}\,\mathrm{Mpc}^{-3}\,\mathrm{yr}} \cdot \frac{0.1\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}}{\dot{n}_0}. \tag{27}$$

This means that the thick curves shown in Figure 15, right panel, actually correspond to the estimated baryonic loadings from the UHECR paradigm: $f_{p/e} \simeq 10$ for SP_{E54}, and $f_{p/e} \simeq 3$ for $MP_{E54.5}$. On the other hand, the required $f_{p/e}$ for powering the UHECR does not depend on δt_{var}. Hence, bursts like $MP_{E54.5}$ with short time variability cannot power the UHECRs if the stacking bounds are taken into account. This implies that the collision radius of the energetic GRBs has to be large enough to avoid efficient neutrino production. In conclusion, the population of energetic GRBs with large enough collision radii might power the UHECRs and at the same time produce a fraction of the diffuse astrophysical neutrino flux. The hypothesis is testable with the next generation of experiments.

7. DISCUSSION

PHOTON SPECTRA AND COMPATIBILITY WITH OBSERVATIONS

Let us start by comparing the simulated photon spectra to the ones of *Fermi*-LAT detected events. For GRB 170214A, which served as inspiration for our $MP_{E54.5}$ prototype, Tang et al. (2017) report that the photon spectral index in the LAT energy range is similar to or even steeper than the high-energy index β in the GBM range. This would favor equipartition be-

---

[4] Note that $\dot{n}_0$ is the apparent local GRB rate; since for UHECRs, not only GRBs emitting into our direction matter, the actual GRB rate is much higher. By using the isotropic-equivalent energy for UHECRs, that beaming effect cancels.

[5] For an $E^{-2.2}$ ($E^{-1.8}$) spectrum, one finds $f_{\mathrm{bol}} \simeq 0.03$ (0.6) for the parameters chosen in this paper. The required baryonic loading in Eq. 27 is higher (lower) by the corresponding factor seven (three). However, note that this correction critically depends on the minimal proton energy for spectra softer than E^{-2}.
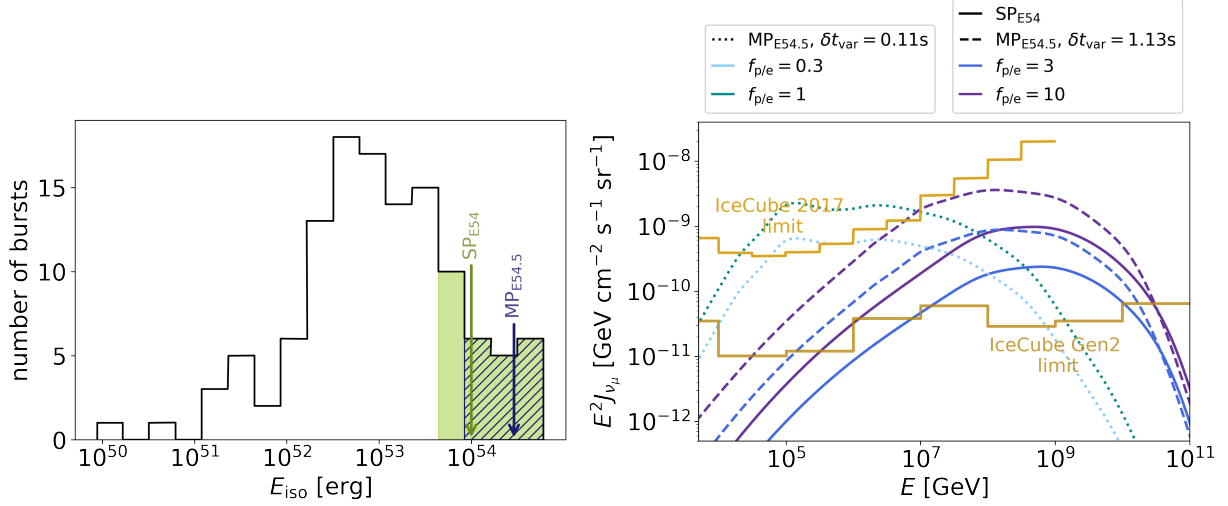
Figure 15. *Left:* Distribution of bursts as a function of source-frame $E_{\rm iso}$ for the 122 long GRBs with measured redshift of the *Fermi*-GBM catalogue presented in Poolakkil et al. (2021). Arrows indicate the $E_{\rm iso}$ of our two prototypes. The green shaded (hashed blue) areas correspond to the number of bursts that have $E_{\rm iso}$ comparable to $SP_{\rm E54}$ ($MP_{\rm E54.5}$). *Right:* All-sky quasi diffuse flux (per flavour) for the SYN-dominated scenarios for both prototype bursts. J_{ν_μ} is calculated from the neutrino fluence $\mathcal{F}_{\nu_\mu}$ from one GRB through $J_{\nu_\mu} = \mathcal{F}_{\nu_\mu}\dot{N}/(4\pi)$ assuming a redshift of $z = 2$ and $\dot{N} = 148$ yr^{-1} for $SP_{\rm E54}$ and $\dot{N} = 93$ yr^{-1} for $MP_{\rm E54.5}$. These rates were obtained by reducing the conventional number of observable GRBs $\dot{N} = 667$ yr^{-1} accounting for the $E_{\rm iso}$ distribution. For comparison we show the current IceCube limits from Aartsen et al. (2017) as well as the projected IceCube Gen2 limits (for 5000 bursts) from Aartsen et al. (2021), which assume that all GRBs contribute to the neutrino emission.

tween electrons and the magnetic field (i.e. our SYN-dominated scenario) over an IC-dominated scenario. We point out that higher values of $f_{\rm B/e}$ are in principle also possible; however in this case the magnetisation of the outflow may be considerable and as pointed out in Narayan et al. (2011), models based on magnetic energy dissipation should be considered instead of the internal shock scenario, such as Spruit et al. (2001); Giannios (2008); Zhang & Yan (2011). Further, the steep power-law index in the LAT range may be explained by a suppression of HE emission due to internal $\gamma\gamma$-absorption in case of small variability timescales/ low dissipation radii. Lepto-hadronic models have also been proposed in context of other LAT-detected bursts, e.g. in Asano et al. (2009a) for GRB 090510, and generally for a delayed HE emission in Asano & Meszaros (2012). Especially interesting is GRB 090902B, which has excess at both low and high energies (Abdo et al. 2009) – thus, a flat underlying spectrum like in our lepto-hadronic SYN-dominated scenario. Independent of the spectral shape of the sub-MeV peak supposedly incompatible with a synchrotron origin (see Ryde et al. 2010) it was pointed out that this flat component may be of hadronic origin in Wang et al. (2018). Soft power-law components that extend from low to high energies are however also present in other bursts (see Guiriec et al. 2015). Generally, a flat additional component was equally predicted by a range of optically thin lepto-hadronic models such

as Asano et al. (2009a); Murase et al. (2012), and we thus confirm their findings.

A major criticism of optically thin synchrotron models for GRBs arises from the typical slopes below the sub-MeV peak. For example, Poolakkil et al. (2021) reported value $\alpha \sim -1.1$ for fluence spectra in the recent *Fermi* GBM catalog, while Yu et al. (2016) found $\alpha \sim -0.8$ for the time resolved spectra of brightest GBM bursts. For GRB 170214A, that inspired our $MP_{\rm E54.5}$ prototype, a Band-function fit with $\alpha = -0.98$ was reported in the GBM catalogue. Such a steep slope could not be reproduced in any of our models, generally the synchrotron fast-cooling prediction of $\alpha = -1.5$ was the highest low-energy slope found. We comment on an potential additional contribution from thermal electrons: Since one of the assumptions of our model is that only a small fraction $\zeta_{\rm e}$, typically $\sim 0.1\%$, of electrons are accelerated in a non-thermal power-law distribution, the low energy slopes can be further modified by the contribution of the remaining quasi-thermal population of electrons which were not considered in this work. We point out that, in order to compare e.g. to the GBM catalogue, data should be forward-folded with the detector response (see e.g. Burgess et al. 2019). Additionally, the accuracy of reported spectra depends on the position of $E_{\rm peak}$ in the detector energy range and whether the asymptotic slopes can be reached within the fitting range. As the spectral slopes also tend to be sensitive to

the fitted function (*e.g.* Band-function or smooth broken powe-law) we further discuss the spectral width or narrowness (Axelsson & Borgonovo 2015; Burgess 2019; Yu et al. 2015) as a measure for spectra. The spectral width is related to the difference between the low- and high-energy spectral slopes (typically the photon indices α and β); simply put, if the difference between the two spectral slopes (ie., $\beta - \alpha$) is large, the spectrum is narrow. GRB spectra are generally very narrow, for example when compared to typical spectra of Active Galactic Nuclei. In our model, the narrowest spectra that we find are the leptonic SYN-dominated ones. For IC-dominated models or for models with large contributions from secondary leptons we find larger high-energy slopes that induce wider spectra.

Next, we comment on low-energy signatures in the eV range and compare with the respective conclusions of other works. The potential of constraining models and parameters by optical measurements was pointed out in Oganesyan et al. (2019), who identified an additional break in the keV-range (see also Toffano et al. 2021; Ravasio et al. 2019). The photon indices below and above the break ($-2/3$ and $-3/2$ respectively) are consistent with an interpretation of marginally fast cooling particles [6]. The possibility of a proton synchrotron interpretation was explored by Ghisellini et al. (2020); Florou et al. (2021); Bégué et al. (2022). While the observed spectra can be reproduced in both scenarios, the inferred physical conditions might be in tension with observables like the short variability timescale and/or the energetics of GRB jets. In the models we examined primary electrons were always in the fast-cooling regime, hence a cooling break in the keV range is not found. Moreover, the synchrotron emission of secondaries, whenever prominent, leads to photon indices close to -2 below a keV break energy.

Finally, we discuss the maximal baryonic loading compatible with observed photon spectra derived in the literature[7] (*e.g.* Asano et al. 2009b; Asano & Mészáros 2014; Petropoulou 2014). A summary of their findings can be found in Fig. 11 of Petropoulou (2014), where the maximum baryonic loading is plotted against the photon compactness of the emitting region, there defined as $\ell_\gamma = \sigma_{\mathrm{T}} L_\gamma / (4\pi R_{\mathrm{Coll}} \Gamma_{\mathrm{em}}^3 m_e c^3)$. This quantity depends on all parameters that affect the comoving photon density and, in turn, the photomeson interaction

rate (i.e. the observed luminosity, the collision radius, and the Lorentz factor). Overall they conclude that for low compactness ($\ell_\gamma \sim 0.1$) the baryonic loading should be below $f_{\mathrm{p/e}} \lesssim 30$, whereas in high-compactness scenarios ($\ell_\gamma \sim 100$) the baryonic loading is limited to much lower values, i.e. $f_{\mathrm{p/e}} \lesssim 10^{0.25} \simeq 1.8$. To put our work in context of these results we first compute the compactness of the representative collision of $\mathrm{SP_{E54}}$, which is $\ell_\gamma = 1.3$. Throughout the full fireball evolution we find $\max(\ell_\gamma) \simeq 7$ for $\mathrm{SP_{E54}}$ and $\max(\ell_\gamma) \simeq 41$ for $\mathrm{MP_{E54.5}}$. These high values of ℓ_γ are however achieved for collisions that do not dissipate most of the energy, and thus are not most relevant for constraining $f_{\mathrm{p/e}}$. Indeed, it is the product of both the dissipated energy and ℓ_γ that is relevant for constraining the baryonic loading in a multi-collision scenario like ours.

BARYONIC LOADING IN LIGHT OF NEUTRINO CONSTRAINTS

The absence of neutrinos from GRBs (Abbasi et al. 2012; Aartsen et al. 2016, 2017; Abbasi et al. 2022) has been a big mystery, especially if these are sources of UHECRs. In the simple one-zone model, where all internal shocks occur at the same R_{Coll}, "typical" GRB parameters, i.e. $\Gamma \simeq 300$ and $\delta t_{\mathrm{var}} \simeq 0.01\,\mathrm{s}$ (corresponding to $R_{\mathrm{Coll}} \simeq 5 \cdot 10^{13}\,\mathrm{cm}$) can be ruled out for the frequently used baryonic loading $f_{\mathrm{p/e}} \simeq 10$, see e.g. Aartsen et al. (2016).

The baryonic loading used in neutrino studies is, however, an *ad hoc* assumption only roughly motivated by the UHECR connection, whereas it can be self-consistently derived in source-propagation models describing both cosmic rays and neutrinos. Systematic parameter space studies for protons (Baerwald et al. 2015) and nuclei (Biehl et al. 2018) describing UHECR observations point towards either large R_{Coll} or low luminosities (see e.g. low-luminosity GRBs in Zhang et al. (2018); Rudolph et al. (2022a)) as possible solutions to the tension with the neutrino GRB stacking searches, where the pion production efficiency is lower. Furthermore, it has been pointed out that the assumption of a single R_{Coll} may be too strong, as different messengers are preferably produced in different regions (Bustamante et al. 2015, 2017; Rudolph et al. 2020): neutrinos may be most efficiently produced close to the photosphere (the energy density scales $\propto R^{-2}$), whereas UHECRs prefer larger R_{Coll} where the energy density (to prevent nuclear disintegration) and the magnetic field (to prevent synchrotron losses limiting the maximal energy) are lower.

GRB multi-collision models can produce lower neutrino fluxes, which allow for a combined description of

---

[6] Defined as the regime where the electron critical Lorentz factor $\gamma_{\mathrm{e,c}} \lesssim \gamma_{\mathrm{e,min}}$.

[7] Compatibility with observations basically means that the contribution of secondary particle emission does not outshine the primary photon emission.

UHECRs and neutrinos (Globus et al. 2015; Heinze et al. 2020). For example, Heinze et al. (2020) obtain a baryonic loading $f_{\mathrm{p/e}} \simeq 60-110$ to power the UHECRs with $R_{\mathrm{Coll}} \simeq 10^{15}$ to 10^{16} cm, using $E_{\gamma,\mathrm{iso}} \simeq (4-8) \cdot 10^{52}$ erg as the typical GRB isotropic energy. In the multi-collision model for energetic GRBs considered here we find that $f_{\mathrm{p/e}} \sim 3 - 10$ is consistent with the energetic requirements for powering UHECRs without violating neutrino stacking limits. Our baryonic loading is lower than the one found by Heinze et al. (2020) due to the much higher energy budget per burst that compensates for small local GRB rate for bursts with $E_{\gamma,\mathrm{iso}} \gtrsim 10^{54}$ erg. Note that spectral effects can act in the other direction and enhance the neutrino flux, as discussed in Sec. 5.

UHECR COMPOSITION

To fully describe UHECR data one also needs to take into account the information about composition. We do not explicitly describe UHECR composition in this work, as this requires a more sophisticated numerical treatment of the radiation processes in the source, including nuclear disintegration and emission processes of all by-products. This is still beyond the state-of-the-art for self-consistent hadronic radiation source models. Nevertheless, qualitative changes are not expected because our production radii are large enough that the nuclear cascade cannot develop (Biehl et al. 2018) (Figs. 7 and 8) and the energetics requirement for nuclei is similar to that of protons (Jiang et al. 2021); however, the composition should be described by a mix of injection isotopes similar to Globus et al. (2015); Heinze et al. (2020). As shown in Biehl et al. (2018), the prompt GRB neutrino fluence exhibits very little dependence on the injection composition; one reason is that the secondary cooling (which depends on B' but not on the composition) is important for the peak. The electromagnetic signatures of nuclei can also be speculated to be similar for an E^{-2} injection spectrum, as each nucleus with energy E_A can be (to first order) regarded as a superposition of A nucleons of energy E_A/A. Our qualitative conclusions are thus expected to hold roughly even if heavier nuclei are accelerated in the source. This has been demonstrated in Globus et al. (2015); Heinze et al. (2020), but still lacks the self-consistent treatment of the electromagnetic signatures of nuclei at the level used for protons in this study.

WHY SHOULD ENERGETIC GRBS BE MORE FAVOURABLE FOR UHECR PRODUCTION THAN "ORDINARY" GRBS?

In this study we argued that energetic GRBs are sufficient for powering UHECRs without violating cur-

rent neutrino constraints. An interesting question that arises then is whether these energetic GRBs are special in terms of proton/nuclei acceleration compared to the more numerous and less energetic GRBs. The pion production efficiency to photohadronic interactions can be analytically estimated from observables as (see Eq. A16 in Guetta et al. (2004))

$$f_\pi \propto \frac{L_{\gamma,\mathrm{iso}} \, \delta t_{\mathrm{var}}}{R_{\mathrm{Coll}}^2 \, \epsilon_{\gamma,\mathrm{eff}}} \propto \frac{L_{\gamma,\mathrm{iso}}}{\Gamma^4 \, \delta t_{\mathrm{var}} \, \epsilon_{\gamma,\mathrm{eff}}} \, . \qquad (28)$$

Here, the co-moving photon number density is estimated from the photon energy divided by co-moving volume, $\epsilon_{\gamma,\mathrm{eff}}$ is the photon energy where the number of photons is maximal, and the shell width is estimated from δt_{var}. This means that similar hadronic signatures on the GRB prompt emission are expected if the relationships $L_{\gamma,\mathrm{iso}} \propto R_{\mathrm{Coll}}^2$, or $L_{\gamma,\mathrm{iso}} \propto \Gamma^4$ are obeyed, neglecting the dependence on δt_{var} for the sake of simplicity; see e.g. Fig. 8 in Biehl et al. (2018) for nuclear disintegration, where this degeneracy can be clearly seen. In that sense, GRBs with very different parameters (e.g. $L_{\gamma,\mathrm{iso}}$, Γ, R, $\epsilon_{\gamma,\mathrm{eff}}$) can have similar properties in terms of the photon density, such as the low luminosity GRBs and large R_{Coll} regions mentioned earlier.

On the other hand, an empirical correlation between the isotropic energy and bulk Lorentz factor has been reported, namely $E_{\gamma,\mathrm{iso}} \propto \Gamma^{2.7}$ for a wind-like circumburst medium or $\propto \Gamma^{2.2}$ for the homogeneous case – see Eq. 20 in Ghirlanda et al. (2018). This empirical relation was obtained by using a sample of GRBs for which the peak time of the afterglow light curve (and the redshift) were measured. The bulk Lorentz factor before the deceleration can be then inferred from the peak time, assuming a density profile for the circumburst medium, and the isotropic energy of the afterglow after the prompt phase. It is clear that the $E_{\gamma,\mathrm{iso}} - \Gamma$ relationship scales differently from the above relationship and therefore breaks that degeneracy. This means that less energetic GRBs, which have lower Lorentz factors according to the $E_{\gamma,\mathrm{iso}} - \Gamma$ relationship, have larger photon densities and therefore qualitatively different properties. Here efficient photohadronic interactions (such as nuclear disintegration) might affect the acceleration of nuclei, and these GRBs may consequently have lower baryonic loadings (describing the non-thermal population).

POTENTIAL FUTURE DIRECTIONS

Possible extensions of the presented work may concern theoretical aspects (e.g. improvement of the modeling) or connection to new experimental results. First of all, a refined outflow model informed by magnetohydrodynamic simulations (as presented for example in Gottlieb

et al. 2021, 2022b,a) might provide a more realistic picture of the physical conditions in the plasma, such as locations of energy dissipation and magnetization. Alternatively, one could lift the assumption of a top-hat jet model made here, and apply our work to a structured jet, similar to the one inferred by observations of GRB 170814A (Troja et al. 2019). One might further consider the feedback between different radiation zones in the jet, which might affect the photohadronic interaction rate. In our work the properties of accelerated particle distributions were motivated by observations instead of being informed by theories of particle acceleration. For instance, the default value for the power-law slope of the accelerated protons ($p_{\mathrm{p}} = 2$) corresponds to the one usually assumed in an UHECR context. On the contrary, a steeper power law was assumed for the accelerated electrons in order to reproduce the typical photon spectrum in the GBM band. Particle acceleration processes can be studied self-consistently using Particle-In-Cell (PIC) simulations. It is noteworthy that the parameter space of mildly relativistic shocks relevant for the GRB prompt emission was explored with PIC simulations only recently (for a review see Marcowith et al. 2020). For example in the context of highly magnetised jets, where magnetic reconnection is the main energy dissipation mechanism, similar power-law indices for electrons and protons are found (e.g. Guo et al. 2016; Petropoulou et al. 2019). For non-magnetised relativistic collisionless shocks, Sironi et al. (2013) indeed find an electron power-law of $p_{\mathrm{e}} = 2.5$ differing from the proton index $p_{\mathrm{p}} = 3.0$. Weakly magnetised mildly relativistic shocks (that correspond to the conditions of the outflow assumed in our model) have been explored in Crumley et al. (2019), who report an electron index of $p_{\mathrm{e}} \sim 2.2$ but no good constraints on the ion indices are found. For magnetised mildly relativistic shocks, Ligorini et al. (2021) find a small supra-thermal population of electrons with $p_{\mathrm{e}} \sim 2$. If the acceleration process produces similar power-law indices, then a steeper proton power-law index than the one used here will affect the neutrino spectra. First and foremost, since less energy is deposited in energetic protons, the overall neutrino flux will be lower. Additionally, the neutrino spectral shape might differ, depending on the shapes of the photon and proton spectra. Finally we point out that in order to describe UHECR data, nuclei heavier than protons should be included in the radiation modeling. In the past, Globus et al. (2015); Heinze et al. (2020) have treated the nuclear cascade of elements up to iron, however without taking the feedback of nuclei on photon spectra into account. In a fully self-consistent picture, disintegration photons may produce photon cascades and the systems of nuclei and photons should not be treated separately. Radiation models can be improved further by making a better connection to observational data. For electromagnetic spectra, an adequate procedure would be to forward-fold predictions with detector responses in order to assess the feasibility of models (as shown in Burgess et al. 2019). For multi-messenger searches, separate analysis for different sub-classes of GRBs could be included in future instruments like IceCube Gen2 (Aartsen et al. 2021) and KM3NeT (Adrian-Martinez et al. 2016). Alternatively, the contribution of different GRBs could be weighted by their luminosity (see *eg.* Lucarelli et al. 2022).

8. SUMMARY & CONCLUSION

In this paper we investigated lepto-hadronic radiation models for the prompt phase of energetic GRBs with $E_{\mathrm{iso}} \gtrsim 10^{54}$ erg. Our study was motivated by the fact that a sub-class of energetic events may power the UHECR energy flux without requiring excessive baryonic loadings. Moreover, strong constraints may be placed on the baryonic loading from the non-observation of associated neutrinos from such energetic bursts. Finally, lepto-hadronic models have been proposed for the HE emission detected in some energetic GRBs by the *Fermi*-LAT.

Our aim was to understand whether the baryonic loading necessary to power the UHECR energy flux is feasible in light of multi-messenger constraints. For this we modelled the prompt phase photon and neutrino emission of two prototype GRBs (inspired by LAT-detected events) within a multi-collision internal shock model that accounts for different emission regions along the jet. In contrast to past multi-collision approaches we self-consistently accounted for the feedback between photon and proton spectra by a a time-dependent radiation model treating the coupled evolution of hadrons, leptons (including neutrinos) and photons. allowed us to examine the contribution of different radiative processes at different energies and times, including synchrotron radiation and synchrotron self-absorption, inverse Compton scattering, photo-pair and photo-pion production, and $\gamma\gamma$-annihilation. For each prototype we evaluated synchrotron (SYN)- and inverse Compton (IC)-dominated scenarios and discussed the maximal baryon loading compatible with observations (for which the secondary cascade emission would not outshine the emission of primary electrons in the GBM band). Beyond single-source spectra and light curves, we compared the diffuse neutrino flux expected for a sub-population of energetic events to current IceCube and predicted IceCube-Gen2 limits. Below we summarise our main findings:

1. *Maximum baryon loading.* This was determined by electromagnetic observations (i.e. spectral modifications in the GBM/LAT band) and neutrino observations (i.e. GRB stacking limits) and was found to be about 3–10 (depending on $E_{\gamma,\mathrm{iso}}$), assuming a E^{-2} proton spectrum for average dissipation radii around 10^{16} cm. This value is consistent with the baryonic loading needed to power the UHECRs with energetic GRBs only. Note that empirical $E_{\gamma,\mathrm{iso}} - \Gamma$ relationships points towards large dissipation radii for energetic GRBs.

2. *Neutrino signatures.* For large enough collision radii, the non-observation of single energetic events in high-energy neutrinos and the current IceCube stacking limits are consistent with the same baryonic loadings of about 3–10. The corresponding neutrino spectra peak at higher energies than previously anticipated, making energetic GRBs interesting targets for radio detection of neutrinos (with future detectors such as GRAND, Álvarez-Muñiz et al. 2020). For smaller collision radii (corresponding to shorter variability timescales), pion and muon cooling reduces the neutrino peak energies, bringing them within the range where current neutrino observatories are more sensitive. Consequently, $f_{\mathrm{p/e}}$ can be limited to smaller values. Note that the neutrino fluence can be enhanced by spectral effects, such as the spectral index in the synchrotron fast-cooling regime or electromagnetic cascade contributions – which we self-consistently include.

3. *Signatures in photon spectra.* We limit ourselves to photon energies below ∼100 GeV that will survive propagation. For SYN-dominated scenarios – even for the moderate baryonic loading of $3-10$ – the synchrotron emission from secondaries introduces an additional broadband power-law component with photon index -2. This implies correlated increase in fluence in the optical-UV, soft X-ray and HE regimes. Hence, to identify hadronic signatures in these energy ranges, prompt low-energy observations combined with HE observations by *Fermi*-LAT would be crucial. The lack of spectral breaks at low energies (UV/X-rays) and/or at high energies (LAT) would exclude low values of $\epsilon_{\mathrm{B}}/\epsilon_{\mathrm{e}}$ and constrain even more the baryonic loading. For IC-dominated scenarios and baryonic loading < 100, the broadband photon spectra obtained in leptonic and lepto-hadronic models are indistinguishable. For these cases, neutrinos are thus the only messengers that could pinpoint hadronic interactions. Generally, a IC-dominated scenario may be identified by a low relative X-Ray/UV flux, and a high HE flux.

4. *The role of different collision radii.* The typical collision radius, which is usually determined by the variability timescale and Lorentz factor of the outflow, is important for models with density-dependent processes (e.g. IC and photo-pion production). Only purely leptonic SYN-dominated models are found to be robust against variations of the collision radius. While electromagnetic spectral features at low energies (below a few keV) and high energies (above ∼ 100 MeV) are produced in collisions taking place in high-density regions of the jet, the peak (sub-MeV to MeV part) of the spectrum is mainly determined by collisions that dissipate most of the energy. Consequently, the region where most of the energy is dissipated is representative for the complete burst in that energy range; on the other hand, it is not representative for the whole GRB's neutrino emission, as it would overestimate the neutrino fluence and underestimate the peak energy.

We conclude that energetic GRBs are promising candidates for multi-messenger searchers. Baryonic loadings as high as necessary to power UHECRs probably do not modify the spectral energy distribution as much as to rule out the UHECR paradigm – in IC-dominated scenarios, these cannot even be uniquely identified. However, interesting insights can be obtained in SYN-dominated scenarios from correlated flux enhancements between the optical-UV to soft X-rays, and the GeV to TeV gamma-ray bands. Finally, the discovery of a neutrino from an energetic GRB, such as GRB 221009A, would be the smoking gun signature for our hypothesis that a class of energetic GRBs could be the sources of UHECRs.

Note. – During completion of this work GRB 221009A was detected. As this is likely to become the most prominent example for this class of energetic GRBs, we discuss the implications of our model for this burst in a companion letter (Rudolph et al. 2022b).

ACKNOWLEDGMENTS

REFERENCES

A. Ursi et al. 2022, GCN Circ. 32650

Aartsen, M. G., et al. 2016, Astrophys. J., 824, 115, doi: 10.3847/0004-637X/824/2/115

—. 2017, Astrophys. J., 843, 112, doi: 10.3847/1538-4357/aa7569

—. 2019, Phys. Rev. Lett., 122, 051102, doi: 10.1103/PhysRevLett.122.051102

—. 2021, J. Phys. G, 48, 060501, doi: 10.1088/1361-6471/abbd48

Abbasi, R., et al. 2012, Nature, 484, 351, doi: 10.1038/nature11068

—. 2021, arXiv e-prints, doi: 10.21234/CPKQ-K003

—. 2022, arXiv e-prints. https://arxiv.org/abs/2205.11410

Abdalla, H., et al. 2019, Nature, 575, 464, doi: 10.1038/s41586-019-1743-9

Abdo, A. A., et al. 2009, Astrophys. J. Lett., 706, L138, doi: 10.1088/0004-637X/706/1/L138

Acciari, V. A., et al. 2019, Nature, 575, 455, doi: 10.1038/s41586-019-1750-x

Ackermann, M., et al. 2010, Astrophys. J., 716, 1178, doi: 10.1088/0004-637X/716/2/1178

Adrian-Martinez, S., et al. 2016, J. Phys. G, 43, 084001, doi: 10.1088/0954-3899/43/8/084001

Ajello, M., et al. 2019, Astrophys. J., 878, 52, doi: 10.3847/1538-4357/ab1d4e

Alvarez, E., Cuoco, A., Mirabal, N., & Zaharijas, G. 2016, JCAP, 12, 023, doi: 10.1088/1475-7516/2016/12/023

Álvarez-Muñiz, J., Alves Batista, R., Balagopal V., A., et al. 2020, Science China Physics, Mechanics, and Astronomy, 63, 219501, doi: 10.1007/s11433-018-9385-7

Asano, K., Guiriec, S., & Meszaros, P. 2009a, Astrophys. J. Lett., 705, L191, doi: 10.1088/0004-637X/705/2/L191

Asano, K., Inoue, S., & Meszaros, P. 2009b, Astrophys. J., 699, 953, doi: 10.1088/0004-637X/699/2/953

Asano, K., & Meszaros, P. 2012, Astrophys. J., 757, 115, doi: 10.1088/0004-637X/757/2/115

Asano, K., & Mészáros, P. 2013, JCAP, 09, 008, doi: 10.1088/1475-7516/2013/09/008

Asano, K., & Mészáros, P. 2014, Astrophys. J., 785, 54, doi: 10.1088/0004-637X/785/1/54

Axelsson, M., & Borgonovo, L. 2015, Mon. Not. Roy. Astron. Soc., 447, 3150, doi: 10.1093/mnras/stu2675

Baerwald, P., Bustamante, M., & Winter, W. 2015, Astropart. Phys., 62, 66, doi: 10.1016/j.astropartphys.2014.07.007

Baerwald, P., Hummer, S., & Winter, W. 2011, Phys. Rev. D, 83, 067303, doi: 10.1103/PhysRevD.83.067303

—. 2012, Astropart. Phys., 35, 508, doi: 10.1016/j.astropartphys.2011.11.005

Bégué, D., Samuelsson, F., & Pe'er, A. 2022, Astrophys. J., 937, 101, doi: 10.3847/1538-4357/ac85b7

Beloborodov, A. M. 2000, Astrophys. J., 539, L25, doi: 10.1086/312830

Beniamini, P., Nava, L., & Piran, T. 2016, Mon. Not. Roy. Astron. Soc., 461, 51, doi: 10.1093/mnras/stw1331

Bhattacharjee, P., & Sigl, G. 2000, Phys. Rept., 327, 109,
doi: 10.1016/S0370-1573(99)00101-5

Biehl, D., Boncioli, D., Fedynitch, A., & Winter, W. 2018,
Astron. Astrophys., 611, A101,
doi: 10.1051/0004-6361/201731337

Bosnjak, Z., Daigne, F., & Dubus, G. 2009, Astron.
Astrophys., 498, 677, doi: 10.1051/0004-6361/200811375

Bošnjak, v., & Daigne, F. 2014, Astron. Astrophys., 568,
A45, doi: 10.1051/0004-6361/201322341

Brainerd, J. J. 1987, Astrophys. J., 320, 714,
doi: 10.1086/165589

Burgess, J. M. 2019, Astron. Astrophys., 629, A69,
doi: 10.1051/0004-6361/201935140

Burgess, J. M., Bégué, D., Bacelj, A., et al. 2019, Nature
Astron., 4, 174, doi: 10.1038/s41550-019-0911-z

Bustamante, M., Baerwald, P., Murase, K., & Winter, W.
2015, Nature Commun., 6, 6783,
doi: 10.1038/ncomms7783

Bustamante, M., Murase, K., Winter, W., & Heinze, J.
2017, Astrophys. J., 837, 33,
doi: 10.3847/1538-4357/837/1/33

Bustamante, M., & Tamborra, I. 2020, Phys. Rev. D, 102,
123008, doi: 10.1103/PhysRevD.102.123008

Crumley, P., Caprioli, D., Markoff, S., & Spitkovsky, A.
2019, Mon. Not. Roy. Astron. Soc., 485, 5105,
doi: 10.1093/mnras/stz232

D. Frederiks et al. 2022, GCN Circ. 32668

Daigne, F., Bosnjak, Z., & Dubus, G. 2011, Astron.
Astrophys., 526, A110,
doi: 10.1051/0004-6361/201015457

Daigne, F., & Mochkovitch, R. 1998, Mon. Not. Roy.
Astron. Soc., 296, 275,
doi: 10.1046/j.1365-8711.1998.01305.x

Deil, C., et al. 2018, PoS, ICRC2017, 766,
doi: 10.22323/1.301.0766

Dermer, C. D., & Atoyan, A. 2006, New J. Phys., 8, 122,
doi: 10.1088/1367-2630/8/7/122

Dermer, C. D., & Razzaque, S. 2010, Astrophys. J., 724,
1366, doi: 10.1088/0004-637X/724/2/1366

Dominguez, A., et al. 2011, Mon. Not. Roy. Astron. Soc.,
410, 2556, doi: 10.1111/j.1365-2966.2010.17631.x

Drenkhahn, G. 2002, Astron. Astrophys., 387, 714,
doi: 10.1051/0004-6361:20020390

Drenkhahn, G., & Spruit, H. C. 2002, A&A, 391, 1141,
doi: 10.1051/0004-6361:20020839

Duran, R., Bosnjak, Z., & Kumar, P. 2012, Mon. Not. Roy.
Astron. Soc., 424, 3192,
doi: 10.1111/j.1365-2966.2012.21533.x

Fiorillo, D. F. G., Van Vliet, A., Morisi, S., & Winter, W.
2021, JCAP, 07, 028,
doi: 10.1088/1475-7516/2021/07/028

Florou, I., Petropoulou, M., & Mastichiadis, A. 2021, Mon.
Not. Roy. Astron. Soc., 505, 1367,
doi: 10.1093/mnras/stab1285

Fraija, N., Veres, P., Zhang, B. B., et al. 2017, Astrophys.
J., 848, 15, doi: 10.3847/1538-4357/aa8a72

Gao, S., Kashiyama, K., & Mészáros, P. 2013, Astrophys. J.
Lett., 772, L4, doi: 10.1088/2041-8205/772/1/L4

Gao, S., Pohl, M., & Winter, W. 2017, Astrophys. J., 843,
109, doi: 10.3847/1538-4357/aa7754

Ghirlanda, G., Nappo, F., Ghisellini, G., et al. 2018,
Astron. Astrophys., 609, A112,
doi: 10.1051/0004-6361/201731598

Ghisellini, G., Ghirlanda, G., Oganesyan, G., et al. 2020,
Astron. Astrophys., 636, A82,
doi: 10.1051/0004-6361/201937244

Giannios, D. 2008, Astron. Astrophys., 480, 305,
doi: 10.1051/0004-6361:20079085

—. 2010, Mon. Not. Roy. Astron. Soc., 408, 46,
doi: 10.1111/j.1745-3933.2010.00925.x

—. 2012, Mon. Not. Roy. Astron. Soc., 422, 3092,
doi: 10.1111/j.1365-2966.2012.20825.x

Giannios, D., & Uzdensky, D. A. 2019, Mon. Not. Roy.
Astron. Soc., 484, 1378, doi: 10.1093/mnras/stz082

Globus, N., Allard, D., Mochkovitch, R., & Parizot, E.
2015, Mon. Not. Roy. Astron. Soc., 451, 751,
doi: 10.1093/mnras/stv893

Gottlieb, O., Bromberg, O., Levinson, A., & Nakar, E.
2021, Mon. Not. Roy. Astron. Soc., 504, 3947,
doi: 10.1093/mnras/stab1068

Gottlieb, O., Liska, M., Tchekhovskoy, A., et al. 2022a,
Astrophys. J. Lett., 933, L9,
doi: 10.3847/2041-8213/ac7530

Gottlieb, O., Moseley, S., Ramirez-Aguilar, T., et al. 2022b,
Astrophys. J. Lett., 933, L2,
doi: 10.3847/2041-8213/ac7728

Granot, J., Piran, T., & Sari, R. 1999, Astrophys. J., 513,
679, doi: 10.1086/306884

Guetta, D., Hooper, D., Alvarez-Muniz, J., Halzen, F., &
Reuveni, E. 2004, Astropart. Phys., 20, 429,
doi: 10.1016/S0927-6505(03)00211-1

Guiriec, S., et al. 2015, Astrophys. J., 807, 148,
doi: 10.1088/0004-637X/807/2/148

Guo, F., Li, X., Li, H., et al. 2016, Astrophys. J. Lett., 818,
L9, doi: 10.3847/2041-8205/818/1/L9

He, H.-N., Liu, R.-Y., Wang, X.-Y., et al. 2012, Astrophys.
J., 752, 29, doi: 10.1088/0004-637X/752/1/29

Heinze, J., Biehl, D., Fedynitch, A., et al. 2020, Mon. Not. Roy. Astron. Soc., 498, 5990, doi: 10.1093/mnras/staa2751

Heinze, J., Fedynitch, A., Boncioli, D., & Winter, W. 2019, Astrophys. J., 873, 88, doi: 10.3847/1538-4357/ab05ce

Hummer, S., Baerwald, P., & Winter, W. 2012, Phys. Rev. Lett., 108, 231101, doi: 10.1103/PhysRevLett.108.231101

Hummer, S., Ruger, M., Spanier, F., & Winter, W. 2010, Astrophys. J., 721, 630, doi: 10.1088/0004-637X/721/1/630

Jiang, Y., Zhang, B. T., & Murase, K. 2021, Phys. Rev. D, 104, 043017, doi: 10.1103/PhysRevD.104.043017

Kelner, S. R., & Aharonian, F. A. 2008, Phys. Rev. D, 78, 034013, doi: 10.1103/PhysRevD.82.099901

Kistler, M. D., Yuksel, H., Beacom, J. F., Hopkins, A. M., & Wyithe, J. S. B. 2009, Astrophys. J., 705, L104, doi: 10.1088/0004-637X/705/2/L104

Kobayashi, S., Piran, T., & Sari, R. 1997, Astrophys. J., 490, 92, doi: 10.1086/512791

Kumar, P., & Crumley, P. 2015, Mon. Not. Roy. Astron. Soc., 453, 1820, doi: 10.1093/mnras/stv1696

Lemoine, M. 2013, Mon. Not. Roy. Astron. Soc., 428, 845, doi: 10.1093/mnras/sts081

Liang, E.-W., Yi, S.-x., Zhang, J., et al. 2010, Astrophys. J., 725, 2209, doi: 10.1088/0004-637X/725/2/2209

Ligorini, A., Niemiec, J., Kobzar, O., et al. 2021, Mon. Not. Roy. Astron. Soc., 502, 5065, doi: 10.1093/mnras/stab220

Lipari, P., Lusignoli, M., & Meloni, D. 2007, Phys. Rev. D, 75, 123005, doi: 10.1103/PhysRevD.75.123005

Lucarelli, F., Oganesyan, G., Montaruli, T., et al. 2022. https://arxiv.org/abs/2208.13792

Lyutikov, M., & Blandford, R. D. 2003. https://arxiv.org/abs/astro-ph/0312347

Marcowith, A., Ferrand, G., Grech, M., et al. 2020, Liv. Rev. Comput. Astrophys., 6, 1, doi: 10.1007/s41115-020-0007-6

Meszaros, P., & Rees, M. J. 1997, Astrophys. J. Lett., 482, L29, doi: 10.1086/310692

Milgrom, M., & Usov, V. 1995, Astrophys. J. Lett., 449, L37, doi: 10.1086/309633

Murase, K., Asano, K., Terasawa, T., & Meszaros, P. 2012, Astrophys. J., 746, 164, doi: 10.1088/0004-637X/746/2/164

Murase, K., & Beacom, J. F. 2010, Phys. Rev. D, 82, 043008, doi: 10.1103/PhysRevD.82.043008

Murase, K., & Nagataki, S. 2006, Phys. Rev. D, 73, 063002, doi: 10.1103/PhysRevD.73.063002

Nakar, E., Ando, S., & Sari, R. 2009, Astrophys. J., 703, 675, doi: 10.1088/0004-637X/703/1/675

Nakar, P. B. D. G. E., & Piran, T. 2011, Mon. Not. Roy. Astron. Soc., 416, 3089, doi: 10.1111/j.1365-2966.2011.19259.x

Narayan, R., Kumar, P., & Tchekhovskoy, A. 2011, MNRAS, 416, 2193, doi: 10.1111/j.1365-2966.2011.19197.x

Nigro, C., et al. 2019, Astron. Astrophys., 625, A10, doi: 10.1051/0004-6361/201834938

Oganesyan, G., Nava, L., Ghirlanda, G., Melandri, A., & Celotti, A. 2019, Astron. Astrophys., 628, A59, doi: 10.1051/0004-6361/201935766

Panaitescu, A., Spada, M., & Meszaros, P. 1999, Astrophys. J., 522, L105, doi: 10.1086/312230

Pe'er, A., Meszaros, P., & Rees, M. J. 2006, Astrophys. J., 642, 995, doi: 10.1086/501424

Petropoulou, M. 2014, Mon. Not. Roy. Astron. Soc., 442, 3026, doi: 10.1093/mnras/stu1079

Petropoulou, M., Dimitrakoudis, S., Mastichiadis, A., & Giannios, D. 2014, Mon. Not. Roy. Astron. Soc., 444, 2186, doi: 10.1093/mnras/stu1362

Petropoulou, M., Sironi, L., Spitkovsky, A., & Giannios, D. 2019, Astrophys. J., 880, 37, doi: 10.3847/1538-4357/ab287a

Pitik, T., Tamborra, I., & Petropoulou, M. 2021, JCAP, 05, 034, doi: 10.1088/1475-7516/2021/05/034

Poolakkil, S., et al. 2021, Astrophys. J., 913, 60, doi: 10.3847/1538-4357/abf24d

Price-Whelan, A. M., et al. 2018, Astron. J., 156, 123, doi: 10.3847/1538-3881/aabc4f

Ravasio, M. E., Ghirlanda, G., Nava, L., & Ghisellini, G. 2019, Astron. Astrophys., 625, A60, doi: 10.1051/0004-6361/201834987

Rees, M. J., & Meszaros, P. 1992, Mon. Not. Roy. Astron. Soc., 258, 41

—. 1994, Astrophys. J., 430, L93, doi: 10.1086/187446

—. 2005, Astrophys. J., 628, 847, doi: 10.1086/430818

Robitaille, T. P., et al. 2013, Astron. Astrophys., 558, A33, doi: 10.1051/0004-6361/201322068

Rudolph, A., Bošnjak, v., Palladino, A., Sadeh, I., & Winter, W. 2022a, Mon. Not. Roy. Astron. Soc., 511, 5823, doi: 10.1093/mnras/stac433

Rudolph, A., Heinze, J., Fedynitch, A., & Winter, W. 2020, Astrophys. J., 893, 72, doi: 10.3847/1538-4357/ab7ea7

Rudolph, A., Petropoulou, M., Winter, W., & Bošnjak, v. 2022b. https://arxiv.org/abs/2212.00766

Ryde, F., Axelsson, M., Zhang, B. B., et al. 2010, ApJL, 709, L172, doi: 10.1088/2041-8205/709/2/L172

S. Lesage et al. 2022, GCN Circ. 32642

Salafia, O. S., Ghisellini, G., Pescalli, A., Ghirlanda, G., & Nappo, F. 2016, Mon. Not. Roy. Astron. Soc., 461, 3607, doi: 10.1093/mnras/stw1549

Samuelsson, F., Bégué, D., Ryde, F., & Pe'er, A. 2019, Astrophys. J., 876, 93, doi: 10.3847/1538-4357/ab153c

Sironi, L., Spitkovsky, A., & Arons, J. 2013, Astrophys. J., 771, 54, doi: 10.1088/0004-637X/771/1/54

Spruit, H. C., Daigne, F., & Drenkhahn, G. 2001, Astron. Astrophys., 369, 694, doi: 10.1051/0004-6361:20010131

Tamborra, I., & Ando, S. 2015, JCAP, 09, 036, doi: 10.1088/1475-7516/2015/9/036

Tang, Q.-W., Wang, K., Li, L., & Liu, R.-Y. 2021, Astrophys. J., 922, 255, doi: 10.3847/1538-4357/ac26ba

Tang, Q.-W., Wang, X.-Y., & Liu, R.-Y. 2017, Astrophys. J., 844, 56, doi: 10.3847/1538-4357/aa7a58

Toffano, M., Ghirlanda, G., Nava, L., et al. 2021, Astron. Astrophys., 652, A123, doi: 10.1051/0004-6361/202141032

Toma, K., Wu, X.-F., & Meszaros, P. 2011, Mon. Not. Roy. Astron. Soc., 415, 1663, doi: 10.1111/j.1365-2966.2011.18807.x

Troja, E., van Eerten, H., Ryan, G., et al. 2019, Mon. Not. Roy. Astron. Soc., 489, 1919, doi: 10.1093/mnras/stz2248

Vietri, M. 1995, Astrophys. J., 453, 883, doi: 10.1086/176448

Wang, K., Liu, R.-Y., Dai, Z.-G., & Asano, K. 2018, Astrophys. J., 857, 24, doi: 10.3847/1538-4357/aab667

Waxman, E. 1995, Phys. Rev. Lett., 75, 386, doi: 10.1103/PhysRevLett.75.386

Waxman, E., & Bahcall, J. N. 1997, Phys. Rev. Lett., 78, 2292, doi: 10.1103/PhysRevLett.78.2292

—. 1999, Phys. Rev. D, 59, 023002, doi: 10.1103/PhysRevD.59.023002

Yong Huang et al. 2022, GCN Circ. 32677

Yu, H.-F., van Eerten, H. J., Greiner, J., et al. 2015, Astron. Astrophys., 583, A129, doi: 10.1051/0004-6361/201527015

Yu, H.-F., et al. 2016, Astron. Astrophys., 588, A135, doi: 10.1051/0004-6361/201527509

Zhang, B., & Kumar, P. 2013, Phys. Rev. Lett., 110, 121101, doi: 10.1103/PhysRevLett.110.121101

Zhang, B., & Yan, H. 2011, Astrophys. J., 726, 90, doi: 10.1088/0004-637X/726/2/90

Zhang, B.-B., Zhang, B., Liang, E.-W., et al. 2011, Astrophys. J., 730, 141, doi: 10.1088/0004-637X/730/2/141

Zhang, B. T., Murase, K., Kimura, S. S., Horiuchi, S., & Mészáros, P. 2018, Phys. Rev., D97, 083010, doi: 10.1103/PhysRevD.97.083010
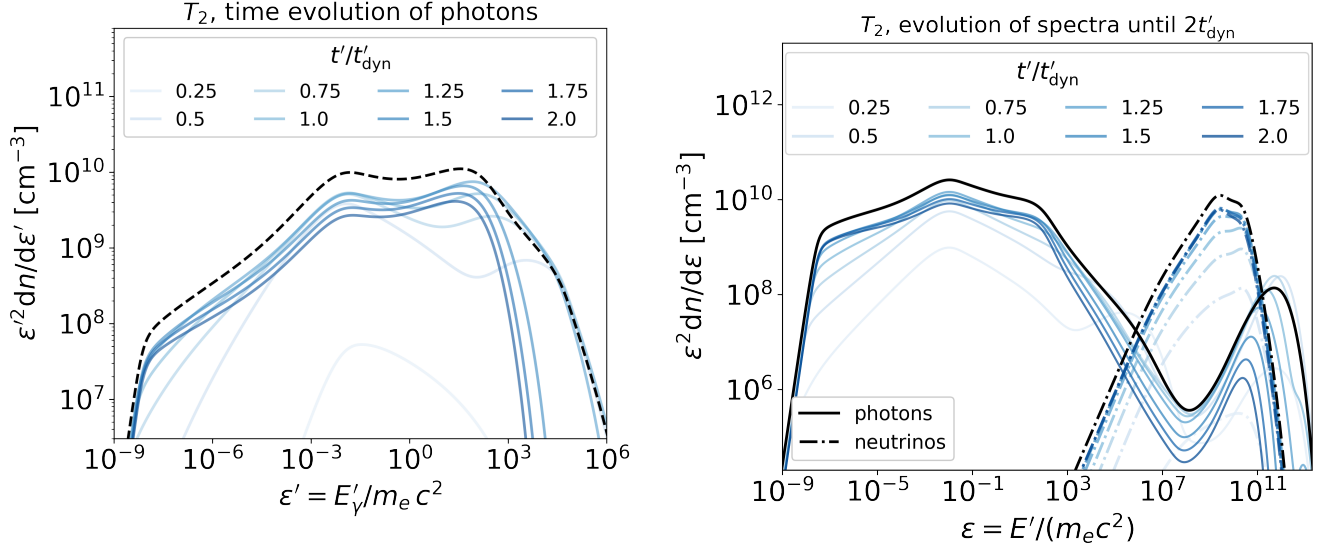
Figure 16. Temporal evolution of comoving photon spectra until $2t'_{\rm dyn}$ for our representative collision, assuming $t'_{\rm inj} = t'_{\rm dyn}$. We show the result for the *(left)* IC-dominated leptonic case and *(right)* SYN-dominated case with $f_{p/e} = 30$. Black lines indicate the emitted spectra.

APPENDIX

A. TIME EVOLUTION IN THE REPRESENTATIVE COLLISION

For each collision we evolve the system until $t' = t'_{\rm inj} + t'_{dyn}$ and calculate the emitted spectrum integrating over the escaping the photon spectrum. As this approach differs from the one of past publications (e.g. Bosnjak et al. 2009; Rudolph et al. 2022a), we illustrate its effect in Figure 16 for the representative collision of $SP_{\rm E54}$, showing the IC-dominated leptonic scenario on the left and the SYN-dominated lepto-hadronic scenario on the right.

Solid lines represent the comoving photon spectrum at different simulation times t', where early (late) times correspond to light (dark) blue colors. The emitted spectra are shown in black. We first discuss the leptonic IC-dominated case (left plot). For $t' < 0.75t'_{dyn}$ the synchrotron peak around $E' \simeq 10^{-2}m_ec^2$ and the inverse Compton peak at higher energies are clearly visible. At later times, synchrotron and inverse Compton radiation of secondary leptons enhances the photon fields at low and intermediate energies. The emitted spectrum encodes all this information. In principle, the same behaviour is observed for the photon fields of the lepto-hadronic case (right plot). Here we further notice how internal $\gamma\gamma$-annihilation suppresses VHE π^0-decay with increasing simulation time.

The neutrino spectra (that are composed of ν_μ, $\bar{\nu}_\mu$, ν_e, and $\bar{\nu}_e$) also show an evolution with time. For example, the peak of the spectrum evolves to lower energies with time. This may be understood as follows: early on, the spectrum is dominated by ν_μ from pion decays. Later, the decay products of muons dominate. However, due to the muons being subject to synchrotron cooling, those secondaries have lower peak energies.

Overall our approach equally captures features that are produced early on or late in the simulation. This may not be true e.g. for a steady-state solution.

B. IMPACT OF MODELING CHOICES

To complement the main results, we systematically explore the impact of the following parameter assumptions and modeling choices: *(a)* The injection timescale of accelerated particles, *(b)* the scaling of the number fraction of accelerated electrons throughout the fireball evolution, *(c)* the maximum and minimum (initial) Lorentz factor of the outflow and *(d)* the magnetic field strength (set through $f_{\rm B/e}$). For $f_{\rm B/e}$ the maximum (minimum) value considered corresponds to the SYN-dominated (IC-dominated) scenario presented in the main text. For simplicity we limit ourselves to leptonic models and show the simulated full-burst spectra examining these aspects in Figure 17. We will
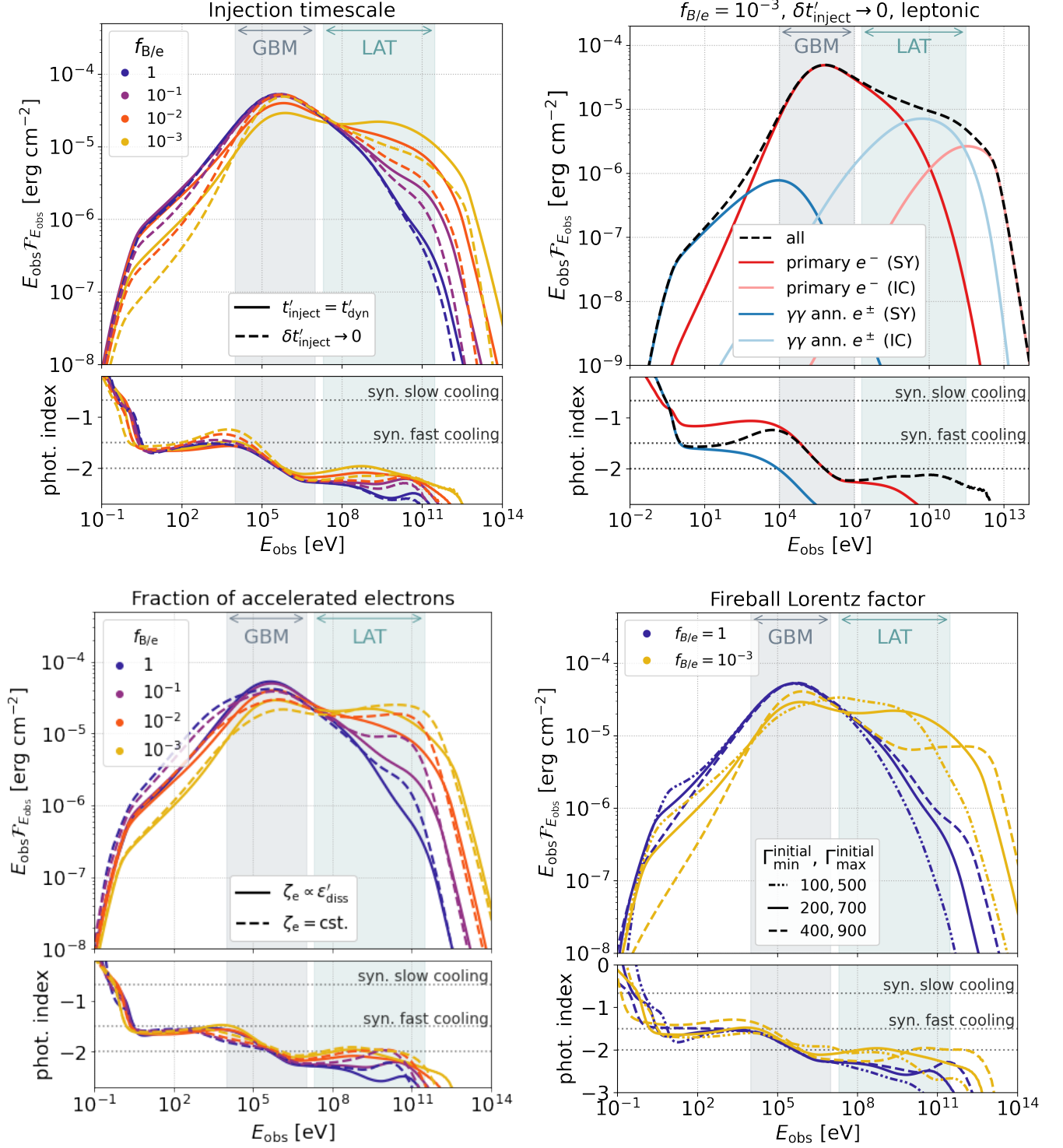
Figure 17. Examining the impact of different modeling choices in a leptonic modeling setup for the single-pulse scenario at the time-integrated spectra $E_{\rm obs}\mathcal{F}_{E_{\rm obs}}$. We show comparisons of *(top left)* different injection schemes, *(bottom left)* different choices for the fraction of accelerated electrons and *(bottom right)* different initial Lorentz factors of the outflow. In all figures, we explore different $f_{\rm B/e}$. In the *(top right)* panel we further show the decomposed spectrum for $\delta t'_{\rm inj} \to 0$ and $f_{\rm B/e} = 10^{-3}$. For the photon indices in the lower panels we indicate the synchrotron slow- and fast-cooling predictions as well as a photon index of -2 (that marks local maxima/minima of $E_{\rm obs}\mathcal{F}_{E_{\rm obs}}$) as dotted lines.

discuss the results in context to the parameter and model choices generally applied in GRB radiation models and necessary to reproduce spectra with properties similar to observed ones.

(A) INJECTION TIMESCALE OF ACCELERATED PARTICLES

In Bosnjak et al. (2009); Daigne et al. (2011); Bošnjak & Daigne (2014), the injection timescale of primary electrons (for each single-collision radiation calculation) was considered to be much smaller than the dynamical timescale. To scrutinise the implications of different choices of the injection timescale, we show the simulated spectra for $t'_{\rm inj} = t'_{\rm dyn}$ and $t'_{\rm inj} = 0.01 t'_{\rm dyn}$ (labelled as $\delta t'_{\rm inj} \to 0$) in the top left plot Figure 17.

We find that the results for the two injection timescale agree well for strong magnetic fields (high $f_{\rm B/e}$). For low $f_{\rm B/e}$, the results for $\delta t'_{\rm inj} \to 0$ differ in three main aspects: *(1)* The fluence of at the sub-MeV peak does not change with respect to the one for high $f_{\rm B/e}$, *(2)* the (V)HE component and low-energy fluences are weaker than for $t'_{\rm inj} = t'_{\rm dyn}$. and *(3)* the low-energy photon index depends systematically on $f_{\rm B/e}$ (where high $f_{\rm B/e}$ yield low photon indices and vice versa, see also Rudolph et al. (2022a)). On the other hand, the low-energy photon index is almost independent of $f_{\rm B/e}$ for $t'_{\rm inj} = t'_{\rm dyn}$. As these effects appear for low $f_{\rm B/e}$, we attribute them to inverse Compton scatterings that are more efficient in case weak magnetic fields. We relate this to the three mains points identified before: *(1)* The higher fluence at the sub-MeV peak may be explained through the dominating energy loss mechanism of primary electrons at $\gamma'_{\rm e,min}$: for $\delta t'_{\rm inj} \to 0$ energetic primaries are present only at early t'. However, at this point the photon spectrum is still building up and cannot serve as target for inverse Compton scatterings. Consequently, the efficiency of inverse Compton scatterings is low(er) and primary electrons at $\gamma'_{\rm e,min}$ predominantly cool through synchrotron emission also in low $f_{\rm B/e}$ scenarios. This, in turn, also means that inverse Compton scatterings are generally less efficient, which yields lower (V)HE photons and also less contributions from the secondary cascade. This is further illustrated in the top right plot of Figure 17, where we show the decomposed spectrum for $\delta t'_{\rm inj} \to 0$ and $f_{\rm B/e} = 10^{-3}$: Despite the fact that the parameter choices correspond to the IC-dominated scenario, primary synchrotron emission still dominates the observed spectrum. This implies that the low- and high-energy fluences which are produced by cascade emission are lower than for $t'_{\rm inj} = t'_{\rm dyn}$ (point *(2)* mentioned before). The lower contribution of the secondary cascade also explains why harder low-energy photon indices can be achieved: as detailed above, inverse Compton scatterings in Klein-Nishina regime modify the low-energy slope of the primary synchrotron such that values up to -1 can be achieved. This was also observed in Daigne et al. (2011), but could not be reproduced by our results in the main text due to the strong contribution of synchrotron emission of secondary lepton pairs. Now choosing the same setup as Daigne et al. (2011) we can however confirm their findings.

We conclude that SYN-dominated models (i.e. with high $f_{\rm B/e}$) are robust against variations of the injection timescale of energetic particles, whereas low-$f_{\rm B/e}$ scenarios are sensitive to the choice of $t'_{\rm inj}$. Meanwhile, the findings of past publications of Daigne et al. (2011) can only be reproduced for the same settings of the injection timescale.

(B) EVOLUTION OF THE NUMBER FRACTION OF ACCELERATED ELECTRONS $\zeta_{\rm E}$

GRB synchrotron models within the internal shock scenario commonly set the number fraction of accelerated electrons proportional to the dissipated energy per unit mass ($\zeta_e \propto \varepsilon'_{\rm diss}$). The intention is to keep the minimum Lorentz factor of electrons $\gamma'_{\rm e,min}$ constant throughout the fireball evolution to reproduce the correct spectral evolution. Additionally, it results in narrow(er) spectra. In the lower left panel of Figure 17, we show the results for $\zeta_e \propto \varepsilon'_{\rm diss}$ and $\zeta_e = $ cst. (with our benchmark choice of $t'_{\rm inj} = t'_{\rm dyn}$ and setting the constant ζ_e such that the same peak energy is reproduced). Independent of $f_{\rm B/e}$ the simulations with $\zeta_e = $ cst. are different in two main points. First, the synchrotron peak is broadened and reduced in energy fluence. Second, the low-energy and high-energy fluences are enhanced. The broad synchrotron peak is reflected in the photon indexes, which are relatively flat around the peak. With values that are overall below -1.5, it is questionable whether $\zeta_e = $ cst. can produce narrow spectra compatible with observations.

We conclude that $\zeta_e \propto \varepsilon'_{\rm diss}$ (which yields a constant minimum Lorentz factor of electrons $\gamma'_{\rm e,min}$ throughout the fireball evolution) is necessary to reproduce spectra compatible with observed ones within a multi-collision model. We point out that this is a somewhat fine-tuned choice and should be confirmed by simulations of particle acceleration in (mildly) relativistic shocks for a broad range of shock strengths and comoving energy densities.

(C) INITIAL JET LORENTZ FACTOR

Next, we examine the impact of the jet Lorentz factor. We choose three different initial Lorentz factor distributions, characterised by their minimal and maximal Γ:

$$(\Gamma_{\rm min}^{\rm initial}, \Gamma_{\rm max}^{\rm initial}) \in \{(100, 500), (200, 700), (400, 900)\}\,.$$

For the sake of simplicity, the lower right panel of Figure 17 only shows the results for $f_{B/e} = 1$ and $f_{B/e} = 10^{-3}$ (i.e. the SYN- and IC-dominated scenario introduced in the main text). For all cases we adjusted $\zeta_{e,0}$ such that the peak energy remains the same. We find that for the SYN-dominated scenario ($f_{B/e} = 1$), the results are robust against variations of the initial Lorentz factor distribution. However, the IC-dominated scenario of $f_{B/e} = 10^{-3}$ shows strong dependence on the Lorentz factor distribution: the lower the Lorentz factor is, the stronger the distortions of the spectrum are. A natural explanation for this behaviour is the impact of lepton pairs produced in $\gamma\gamma$-annihilation, which are produced in dense environments and distort the spectrum by synchrotron radiation at low energies and inverse Compton radiation at high energies. For low Lorentz factors, the comoving densities are higher (due to smaller collision radii and a smaller effect of boosting). Consequently, the $\gamma\gamma$-annihilation efficiency is increased and a large number of pairs reshape the spectrum.

We point out that the initial jet Lorentz factors are only one way of altering the comoving (energy) densities. Other parameters would be the overall duration of the burst (short durations yield small collision radii and hence large densities) or the isotropic equivalent energy (large $E_{\gamma,iso}$ induce large densities).

(D) MAGNETIC FIELD STRENGTH

We proceed by reviewing the impact of the magnetic field strength (set through $f_{B/e}$), which we varied for all scenarios presented above. The most extreme choices presented as blue and yellow lines in Figure 17 correspond to the SYN- and IC-dominated scenario presented in the main text. Varying $f_{B/e}$ systematically between these two values we observe a smooth transition between the two scenarios: the features of the IC-dominated scenario such as an increase in low- and high-energy fluence become more pronounced as $f_{B/e}$ is decreased. As outlined above, for $\delta t'_{inject} \to 0$ also the low-energy photon index depends systematically on $f_{B/e}$, where low (high) $f_{B/e}$ yield hard (soft) photon indices. This however cannot be confirmed in scenarios with a longer injection timescale.

C. PION AND MUON COOLING AND IMPACT ON NEUTRINO SPECTRA

In this section we study the cooling of intermediate pions and muons and its impact on neutrino spectra. We account for synchrotron and adiabatic cooling and decay. For a particle of Lorentz factor γ' and species i (with corresponding mass m_i, charge number Z_i and rest-frame lifetime $\tau_{0,i}$) the corresponding cooling/decay times in the plasma rest frame are

$$t'_{sy,i}(\gamma') = \frac{6\pi}{Z_i^4 \sigma_t} \frac{m_i c^2}{c} \left(\frac{m_i}{m_e}\right)^2 \frac{1}{B'^2 \gamma'} \tag{C1}$$

$$t'_{ad,i} = t'_{ad} \tag{C2}$$

$$t'_{dec,i}(\gamma') = \tau_{0,i} \gamma' , \tag{C3}$$

with corresponding cooling rates $t'^{-1}_{sy,i}(\gamma')$, $t'^{-1}_{ad,i}$ and $t'^{-1}_{dec,i}(\gamma')$.

C.1. *Pion and muon cooling in the representative collision and identification of cooling regimes*

We study the cooling of pions and muons first at the example of representative collision. While adiabatic cooling and decay rates are independent of the magnetic field, synchrotron cooling scales with B'^2. Similar to the results in the main text we again study the SYN- and IC-dominated scenarios.

The cooling timescales for pions and muons are shown in the upper panel of Figure 18, calculated from Eq. C1 - Eq. C3. For the SYN-dominated scenario, adiabatic cooling is sub-dominant at all energies. The longer decay time of muons shifts their critical energy $E_{c,sy/dec}$ to lower energies with respect to pions. For the IC-dominated scenario, synchrotron cooling is less efficient. For pions however either synchrotron cooling or decay domiantes while for muons adiabatic cooling dominates at intermediate energies.

The figure further displays the maximum proton energies as a function of t'/t'_{dyn} (recall that the injection time was set to $t'_{inj} = t'_{dyn}$. For the SYN-dominated scenario the proton energies are initially limited by adiabatic cooling and by photo-pion production at later t'. This implies that the maximum proton energies consecutively move to lower values with evolving t' as photo-pion production gets more efficient with the building up of the photon fields. For the IC-dominated case, the maximum proton energy is dominated by adiabatic cooling until almost t'_{dyn} and remains almost constant. We compare the maximum proton energies to the pion and muon cooling rates: For the SYN-dominated scenario, secondaries are produced at high enough energies to be able to cool prior to their decay. In contrast to this,
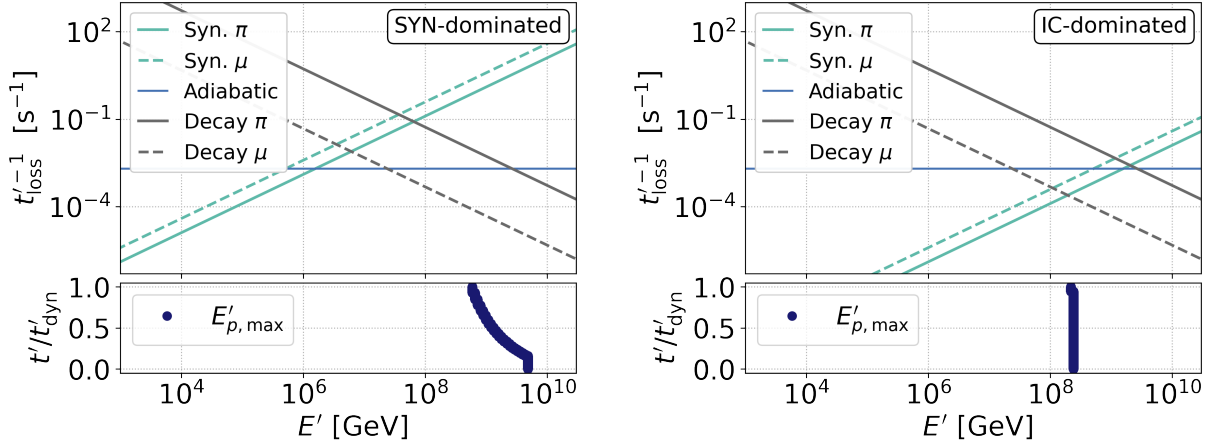
Figure 18. Pion and muon loss rates $t_{\rm cool}'^{-1}$ for the representative collision, for *(top)* the SYN-dominated scenario and *(bottom)* the IC-dominated scenario; Complemented by the corresponding evolution of (comoving) maximal proton energies obtained for the radiative calculations until $t_{\rm dyn}'$.

the pions produced in the IC-dominated scenario decay without prior cooling. The secondary muons cool adiabatically before decaying.

We label the energies at which the timescales of two different processes are equal as *critical* energies. These mark the intersections of timescale/ cooling rates in the respective plots. Three processes imply three critical energies: $E_{\rm c,ad/dec}$ (where $t_{\rm ad}' = t_{\rm dec}'$), $E_{\rm c,sy/dec}$ (where $t_{\rm sy}' = t_{\rm dec}'$) and $E_{\rm c,ad/sy}$ (where $t_{\rm ad}' = t_{\rm sy}'$).

As the decay rate decreases with particle energy, it can be assumed to be the dominating loss process at the lowest energies. Depending on the relative importance of the three processes, two scenarios are possible and can be identified in Figure 18:

(1) $E_{\rm c,ad/dec} < E_{\rm c,ad/sy} < E_{\rm c,sy/dec}$: At the highest energies synchrotron cooling dominates, at intermediate ones adiabatic cooling and at the lowest decay. It is realised for muons in the IC-dominated (right) plot of Figure 18.

(2) $E_{\rm c,ad/sy} < E_{\rm c,sy/dec} < E_{\rm c,ad/dec}$: Adiabatic cooling is subdominant for all energies. At high energies synchrotron losses dominate, at low ones decay. It is nicely illustrated for muons and pions in the SYN-dominated (left) plot of Figure 18.

C.2. *Cooling regimes throughout the fireball dynamical evolution*

From the representative collision we move to the complete fireball dynamical evolution. We define the critical energy as

$$E_{\rm c}' = \min\{E_{\rm c,sy/dec}', E_{\rm c,ad/dec}'\}. \tag{C4}$$

It may be re-phrased as energy below which particles do not cool prior to their decay. The evolution of critical energy as a function of collision radius for our educative example is shown in Figure 19. The color-coding further indicates if above the critical energy, adiabatic or synchrotron cooling dominate: Dark red/green means that adiabatic cooling dominates, for light red/green synchrotron cooling is dominant above the critical energy. Thus, light (dark) colors correspond to scenario (1) ((2)) introduced before. For comparison we further indicate the maximal proton energy calculated by balancing synchrotron and adiabatic losses with acceleration. We point out that this is rather an upper limit on $E_{\rm prot,max}'$, since photo-pair or photo-pion production may further reduce the maximal energy attainable.

For the SYN-dominated scenario, synchrotron cooling dominates for both pions and muons throughout the full fireball evolution. Further, the critical energies lie several orders of magnitude below $E_{\rm p,max}'$. In this scenario we thus expect strong impact of pion and muon cooling. For the IC-dominated case, pion critical energies are dominated partially by adiabatic, partially by synchrotron cooling. Muon critical energies are dominated by adiabatic cooling throughout the complete fireball evolution. In this scenario, the maximum proton energies lie below the pion critical energies for all collisions. Pions will thus not cool before decaying. Muon critical energies are below $E_{\rm p,max}'$ for some part of the fireball evolution, the difference in energy is however one order of magnitude at most.
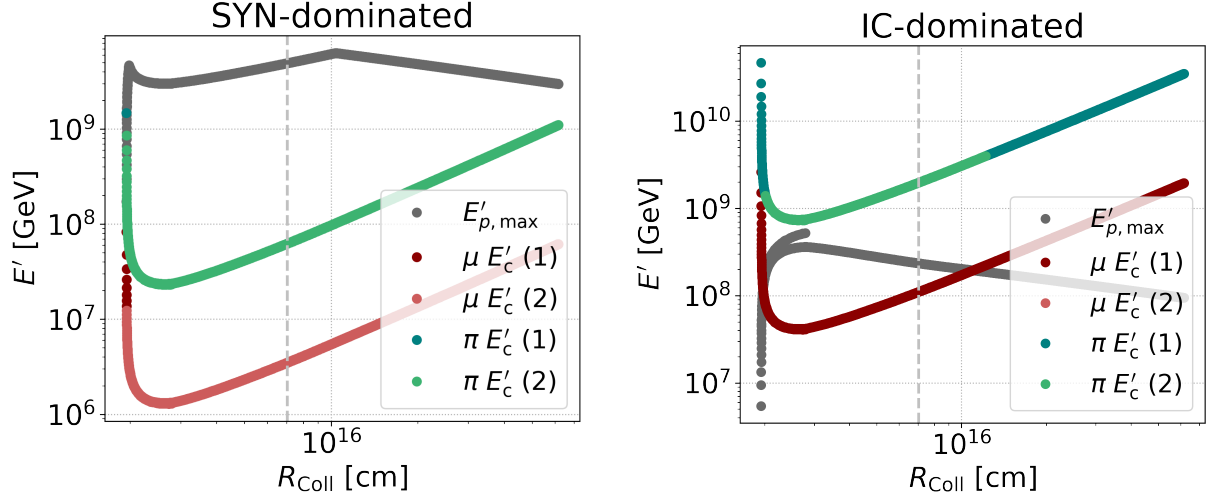
Figure 19. Critical energies E'_c (see Eq. C4) for pions and muons. The number indicates whether above the critical energy *(1)* adiabatic or *(2)* synchrotron cooling dominates. For reference we further indicate the maximal proton energies obtained by balancing acceleration with adiabatic and synchrotron losses (in contrast to Figure 18 where also photo-pion and photo-pair production were taken into account). The dashed line indicates the radius of the representative collision.
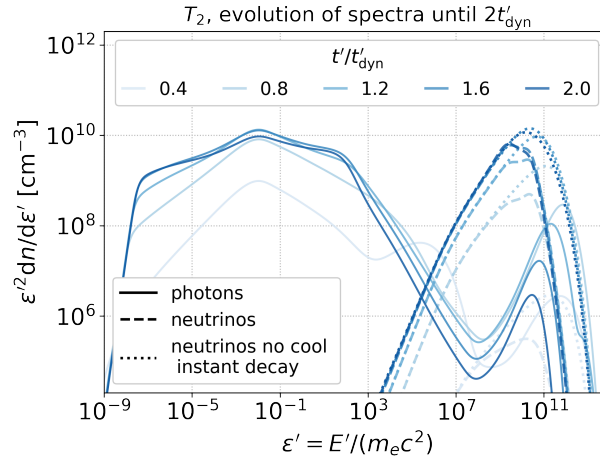
Figure 20. Evolution of comoving photon and neutrino spectra until $2t'_{\rm dyn}$ for the representative collision and the SYN-dominated scenario.

We conclude that synchrotron cooling of secondary pions and muons will reduce their energies prior to their decay for high magnetic fields (high $f_{\rm B/e}$). As this introduces a spectral break, it re-shapes the secondaries such as lepton pairs and neutrinos. For low magnetic fields (low $f_{\rm B/e}$) on the other hand, the maximal proton energies are too low to enable pion cooling before their decay. Muons may cool adiabatically before their decay. Overall, the cooling effects of pions and muons will much less impact lepton and neutrino spectra.

C.3. *Impact on neutrino spectra*

Finally, we review the impact of muon and pion cooling on neutrino spectra. We found that for the SYN-dominated scenario the neutrino spectra peak at much lower energies compared to the photons produced in π^0 decays. For the IC-dominated scenario, neutrino and π^0-decay spectra extended to approximately the same energies. In the last section we deduced that for the SYN-dominated scenario, intermediate pions and muons are subject to synchrotron cooling prior to their decay. To validate that this explains the difference between neutrino and photon spectra, we perform simulations with/without pion and muon cooling for the representative collision in the SYN-dominated scenario.

We show the evolution of comoving photon and neutrino spectra in the left plot of Figure 20. Throughout the complete evolution, the neutrino spectra peak at lower energies than the HE photons from π^0 decays. The right plot shows the evolution of neutrino spectra *(1)* with full treatment of pions/muon, *(2)* without pion and muon cooling and *(3)* assuming an infinitesimally small decay time of muons and pions (which implicitly means that pion and muon cooling are not accounted for).

For the first snapshot at $t' = 0.2t'_{\mathrm{dyn}}$, the scenarios with/ without pion and muon cooling are little different and do not extend to the highest energies. If however, pions and muons are assumed to decay instantaneously (dotted lines in the right plot), energies comparable to the photons from π^0-decays are reached. This means that at early times t' the maximal energies of neutrinos are limited by the longer decay times of high-energy pions and muons (recall the dependency of the decay time on particle Lorentz factor, Eq. C3). As the system evolves, the decay time even of high-energy particles gets smaller than t' and the dash-dotted and dotted lines in Figure 20 extend to similarly high energies. We point out that the effect of the decay timescale is not as relevant if pion and muon cooling are accounted for: By cooling down, particles move to shorter decay times.

Overall, Figure 20 confirms that the lower peak energies of neutrino spectra for a SYN-dominated scenario can be attributed to cooling effects of intermediate pions and muons.