



# Relativistic simulations of long-lived reverse shocks in stratified ejecta: the origin of flares in GRB afterglows

A. Lamberts<sup>1</sup>★ and F. Daigne<sup>2</sup>

<sup>1</sup>Theoretical Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA

<sup>2</sup>UPMC-CNRS, UMR7095, Institut d'Astrophysique de Paris, F-75014 Paris, France

Accepted 2017 November 14. Received 2017 November 14; in original form 2017 February 14

## ABSTRACT

The X-ray light curves of the early afterglow phase from gamma-ray bursts (GRBs) present a puzzling variability, including flares. The origin of these flares is still debated, and often associated with a late activity of the central engine. We discuss an alternative scenario where the central engine remains short-lived and flares are produced by the propagation of a long-lived reverse shock in a stratified ejecta. Here we focus on the hydrodynamics of the shock interactions. We perform one-dimensional ultrarelativistic hydrodynamic simulations with different initial internal structure in the GRB ejecta. We use them to extract bolometric light curves and compare with a previous study based on a simplified ballistic model. We find a good agreement between both approaches, with similar slopes and variability in the light curves, but identify several weaknesses in the ballistic model: the density is underestimated in the shocked regions, and more importantly, late shock reflections are not captured. With accurate dynamics provided by our hydrodynamic simulations, we confirm that internal shocks in the ejecta lead to the formation of dense shells. The interaction of the long-lived reverse shock with a dense shell then produces a fast and intense increase of the dissipated power. Assuming that the emission is due to the synchrotron radiation from shock-accelerated electrons, and that the external forward shock is radiatively inefficient, we find that this results in a bright flare in the X-ray light curve, with arrival times, shapes, and duration in agreement with the observed properties of X-ray flares in GRB afterglows.

**Key words:** hydrodynamics – radiation mechanisms: non-thermal – relativistic processes – shock waves – methods: numerical – gamma-ray burst: general.

## 1 INTRODUCTION

Gamma-ray bursts (GRBs) are intense flashes of gamma-rays of extragalactic origin, with an apparent rate of about once a day. Their duration shows a bimodal distribution, with long bursts (from a few seconds to a few minutes) associated with the collapse of certain massive stars (Woosley 1993; MacFadyen & Woosley 1999), and short bursts (from a few milliseconds to a few seconds) believed to originate from mergers of compact objects in a binary system (see Piran 2004; Gehrels & Mészáros 2012; Gomboc 2012; Berger 2014, for reviews). The prompt emission peaks in the keV–MeV range and is followed by a rapidly fading afterglow observed on longer time-scales from the X-ray to the radio wavelengths (see Zhang, Lü & Liang 2016, for a recent review).

The initial event leads to the formation of a new compact source, probably an accreting black hole, even if rapidly rotating magne-

tars are also discussed (Troja et al. 2007; Rowlinson et al. 2013). The highly variable prompt emission corresponds to a tremendous energy radiated in gamma-rays ( $E_{\text{iso}} \sim 10^{51}–10^{54}$  erg) and must be due to internal dissipation (to account for the fast variability) in a ultrarelativistic outflow (to avoid a strong  $\gamma\gamma$  annihilation).

The afterglow phase is due to the deceleration of the relativistic outflow by the external medium. In the standard scenario, the afterglow emission is due to the synchrotron radiation of non-thermal electrons accelerated at the ultrarelativistic (external) forward shock (FS; see Kumar & Zhang 2015, for a recent review). This naturally leads to an observed flux  $F_v \propto t_{\text{obs}}^{-\alpha}$  with  $\alpha \simeq 1$  (Mészáros & Rees 1997; Sari, Piran & Narayan 1998), in agreement with observations a few hours after the burst.

However, this simple picture was soon challenged by the first observations with the *Swift* satellite (Gehrels et al. 2004), which revealed a great diversity of light curves, especially in the early afterglow phase, and an unexpected variability in the X-ray afterglow. The latter can show an early steep decay just after the prompt emission ( $\alpha \simeq 3$ ), often followed by a plateau phase ( $\alpha \simeq 0$ )

\* E-mail: lamberts@caltech.edu

before the standard decay ( $\alpha \simeq 1$ ) (Tagliaferri et al. 2005; Nousek et al. 2006; O'Brien et al. 2006). In addition, X-ray flares are observed in  $\sim 30$  per cent of GRBs (see e.g. Burrows et al. 2005a; Chincarini et al. 2010), with some bursts presenting several flares between a few 10 s and several  $10^4$  s after the trigger. The underlying shape of the X-ray light curve remains unaffected by the flares. The flares are asymmetric, with a steep rise and a slower decay (Chincarini et al. 2007). However, contrary to the gamma-ray variability observed during the prompt phase, the X-ray flares present a common behaviour with duration increasing with time, following  $\Delta t_{\text{obs}}/t_{\text{obs}} \simeq 0.1\text{--}0.3$ .

This observed diversity and variability of the X-ray afterglow is difficult to reconcile with the standard external shock model. In particular, the emission is only mildly sensitive to the structure of the external medium (e.g. Dai & Lu 2002; Ramirez-Ruiz et al. 2005; Margutti et al. 2011; Gat, van Eerten & MacFadyen 2013). Plateaus can be reproduced with late energy injection (Rees & Mészáros 1998; Sari & Mészáros 2000; Lyutikov & Camilo Jaramillo 2017), and flares can be associated with late internal dissipation in the ejecta in the case of a long-lasting central engine (e.g. Burrows et al. 2005b; Fan & Wei 2005; Zhang et al. 2006; Margutti et al. 2011). Such possibilities put strong constraints on the energetics and lifetime of the source. Few alternatives without such constraints have been proposed (e.g. delayed magnetic dissipation proposed by Giannios 2006, or photospheric emission from slow material ejected together with the GRB relativistic ejecta; Beniamini & Kumar 2016).

On the other hand, the emission of the reverse shock (RS) propagating within the ejecta is very sensitive to the structure of the ejecta and may provide an alternative explanation. Sari & Mészáros (2000) showed that this RS emission can be long-lasting, if the ejecta end with a tail of low Lorentz factor material. Uhm & Beloborodov (2007) and Genet, Daigne & Mochkovitch (2007) proposed that such a long-lived reverse shock (LLRS) may dominate the observed emission and could easily produce plateaus. The capacity of this model to reproduce the diversity of the observed X-ray and optical light curves, including plateaus and chromatic breaks, was successfully tested by Hascoët et al. (2011), Hascoët, Daigne & Mochkovitch (2012), and Uhm et al. (2012). Several observed correlations between the prompt and plateau properties can also be accounted for (Hascoët, Daigne & Mochkovitch 2014).

Hascoët et al. (2017) suggested that this scenario could also naturally explain flares, alleviating the constraints on the lifetime and variability of the central engine. The flares are associated with the propagation of the LLRS through dense shells within the ejecta, which are expected to appear after the development of internal shocks. Hascoët et al. (2017) presented simplified simulations based on a ballistic model (Daigne & Mochkovitch 1998) and computed bolometric light curves assuming an anisotropic emission in the co-moving frame (Beloborodov et al. 2011). They obtained asymmetric flares with  $\Delta t_{\text{obs}}/t_{\text{obs}} \simeq 0.1\text{--}0.3$ . In the present paper, we validate this scenario by performing relativistic hydrodynamical simulations to compute the dynamics of the interaction of the LLRS with a structured ejecta. Although the analytic models are promising and allow the exploration of a wide range of parameters, only hydrodynamical simulations can account for the full dynamics of the shock interactions and are a necessary step for the validation of the model.

Fully modelling the propagation of an LLRS is a numerical challenge. On one hand, one needs a high enough resolution to properly model the internal shocks. On the other hand, one needs to follow the ejecta beyond the deceleration radius to follow the impact of the LLRS. As such, multidimensional simulations have specifically

focused on the internal shock phase (see e.g. Aloy et al. 2000; Duffell & MacFadyen 2015) or the deceleration phase (Meliани et al. 2007; Mimica, Giannios & Aloy 2009; van Eerten, Zhang & MacFadyen 2010; De Colle et al. 2012). Currently, parameter exploration including both phases can only be achieved with one-dimensional (1D) simulations. In this paper, we perform 1D relativistic hydrodynamic (RHD) simulations to assert the validity and limits of the ballistic model described in Hascoët et al. (2017). In Section 2 we describe our numerical methods and the set-up of our set of simulations. In Section 3 we analyse the hydrodynamic structure of the ejecta, and compute the resulting bolometric flux and compare to the ballistic model in Section 4. We discuss the observability of X-ray flares in Section 5 and conclude in Section 6.

## 2 NUMERICAL METHODS

### 2.1 Relativistic RAMSES

Our numerical method is based on the relativistic extension of the RAMSES code (Teyssier 2002), presented in Lamberts et al. (2013). The code computes the evolution of the conserved variables in the frame of the laboratory:

$$\mathbf{U} = \begin{pmatrix} D \\ M \\ E \end{pmatrix} = \begin{pmatrix} \Gamma \rho \\ \Gamma^2 \rho v \frac{h}{c^2} \\ \Gamma^2 \rho h - P \end{pmatrix}, \quad (1)$$

where  $D$  is the density,  $M$  the momentum density, and  $E$  the energy density.  $h$  is the specific enthalpy,  $\rho$  is the proper mass density,  $v$  is the fluid velocity, and  $P$  is the gas pressure. The Lorentz factor is given by  $\Gamma = (1 - v^2/c^2)^{-1/2}$ . Within the RHDs simulation, the equations are solved with all velocities normalized to the speed of light. We neglect the dynamical impact of magnetic fields. While the numerical scheme in our simulation and in Lamberts et al. (2013) are strictly identical, we use a 1D spherical grid and do not include adaptive mesh refinement in the simulations presented here.

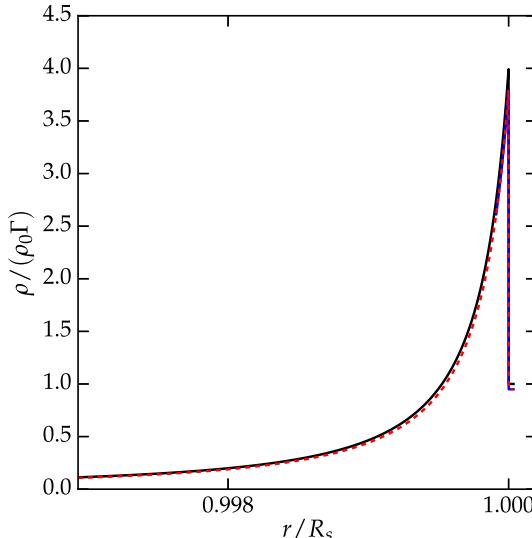
In 1D spherical coordinates, the RHD equations along the radial axis are given by

$$\begin{aligned} \frac{\partial D}{\partial t} + \frac{\partial Dv}{\partial r} &= -\frac{2}{r} Dv, \\ \frac{\partial M}{\partial t} + \frac{\partial(Mv + P)}{\partial r} &= -\frac{2}{r} Mv, \\ \frac{\partial E}{\partial t} + \frac{\partial Mc^2}{\partial r} &= -\frac{2}{r} Mc^2. \end{aligned} \quad (2)$$

The right-hand side of the equations corresponds to the so-called source terms related to the spherical coordinates. A passive scalar  $s$  is included in the simulations using  $S = s\rho\Gamma$  in the laboratory frame and  $F = \rho sv\Gamma$  to compute its flux. As explained below, this allows us to distinguish the GRB ejecta from the external medium, or different regions in the ejecta. Equation (2) is closed with the following equation of state:

$$P = (\gamma - 1)(e - \rho), \quad (3)$$

where  $e$  is the sum of the internal energy and the rest mass energy of the fluid. We assume an ultrarelativistic fluid with  $\gamma = 4/3$ . We use the Harten–Lax–van Leer–Contact (HLLC) Riemann solver and a Minmod slope limiter. Because of the spherical expansion of the ejecta, the density and pressure in the inner part of the ejecta become very small and hard to handle numerically. Therefore, we floor the density and pressure to  $10^{-10}$ .

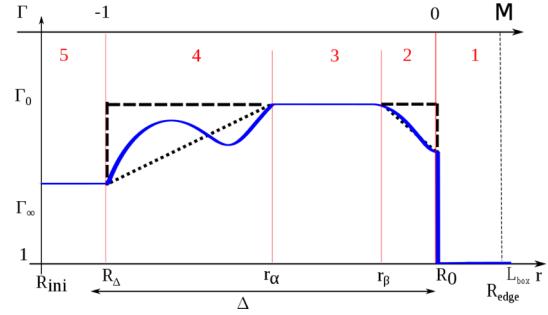


**Figure 1.** Validation of the numerical method based on the normalized density in a 1D spherical blast wave. The simulation on a standard grid (dashed red line), on a moving grid (solid blue line), and the analytic Blandford–McKee solution (solid black line) are in very good agreement, when the FS, located at  $R_s$  reaches the deceleration radius  $R_{\text{dec}}$ . The initial energy of the ejecta is  $10^{53}$  erg and it propagates in an external medium with  $n_0 = 10^3 \text{ cm}^{-3}$ . At the time of the snapshot, the Lorentz factor behind the shock is 20 and  $R_s = 4.8 \times 10^{16} \text{ cm}$ .

## 2.2 Moving grid

Modelling the interactions of the different shocks over time-scales comparable to the deceleration time-scale is the main numerical challenge in our simulations. As the width of the ejecta of the GRB  $\Delta$  is small compared to the deceleration radius  $R_{\text{dec}}$ , a static grid is not well fit. Instead, we have implemented a moving grid that follows the motion of the ejecta. The grid has a fixed physical width  $L_{\text{box}}$ , which is large enough to cover most of the ejecta at the observer time  $t_{\text{obs}}$  we are interested in. At each time step, we compute the position of the FS, defined by the outermost cell where the density is at least a thousand times higher than in the next cell. When the FS has reached  $R_{\text{edge}}$ , we shift the grid towards higher radii. The innermost  $n_{\text{shift}}$  cells are suppressed, the contents of all other cells are shifted towards the left and  $n_{\text{shift}}$  new cells are created on the right-hand side of the simulation. These cells are filled with the user-defined external medium. In our simulations we arbitrarily set  $n_{\text{shift}} = 100$  and  $R_{\text{edge}} - L_{\text{box}} = 100$ . We checked that the exact value of  $R_{\text{edge}}$  and  $n_{\text{shift}}$  does not impact the outcome of the simulation.

Fig. 1 shows the evolution of a relativistic blast wave as it propagates and expands. The simulation with the moving grid shows the same results as the simulation on the complete grid and is very close to the analytic Blandford–McKee solution (Blandford & McKee 1976). Both simulations have the same equivalent resolution and the moving grid simulation takes roughly 10 times less to complete, while modelling most of the mass of the ejecta. While this method perfectly tracks the ejecta over long distances, care has to be taken at the inner boundary. As such, in this specific simulation we set the boundary condition to match the analytic solution. For the main simulations presented here, a zero-gradient inner boundary condition is applied. By comparison with a low-resolution simulation on a static grid, we find that the zero-gradient inner boundary condition provides a proper result as long as some fraction of the



**Figure 2.** Initial distribution of the Lorentz factor in the simulations. The ejecta extends from  $R_\Delta$  to  $R_0$  and the computational box from  $R_{\text{ini}}$  to  $L_{\text{box}}$ . The distances are not to scale. In the simplest case, the ejecta is uniform (dashed line), while in more complex simulations it includes a tail and/or head region. These outer regions can present a linear variation of the Lorentz factor (dotted line) or present a sinusoidal variation (blue solid line) that will lead to internal shocks. The vertical red lines show the initial separation of each region in the simulations, with the corresponding values of the passive scalar given in red. The Lagrangian coordinates are indicated on the upper axis.  $R_{\text{edge}}$  is represented by the dashed vertical black line.

initial ejecta is present in the simulation domain. As such, we stop all our simulations when 50 per cent of the initial ejecta mass has left the domain.

## 2.3 Set-up of the simulations

We perform a set of simulations, modelling different levels of complexity in the initial ejecta by varying the initial distribution of the Lorentz factor  $\Gamma(r)$  as described below. The density is then given by

$$\rho(r) = \frac{\dot{E}}{4\pi r^2 v(r) \Gamma^2(r) c^2 \left(1 + \eta \left(\frac{\gamma}{\gamma-1} - \frac{1}{\Gamma(r)^2}\right)\right)}, \quad (4)$$

where  $\dot{E}$  is the power injected in the ejecta by the central engine and  $\eta = 10^{-3}$  the ratio between pressure  $P$  and the rest mass energy density. The initial width of the ejecta is set by the duration of the burst  $\Delta = ct_w$ . In all our simulations we have  $t_w = 100 \text{ s}$ ,  $\dot{E} = 10^{51} \text{ erg s}^{-1}$ .

The initial inner ( $R_\Delta$ ) and outer radii ( $R_0$ ) of the ejecta are set to  $R_\Delta/c = 100 \text{ s} (\simeq 3 \times 10^{12} \text{ cm})$  and  $R_0/c = 200 \text{ s} (\simeq 6 \times 10^{12} \text{ cm})$ . As such, our simulation starts at  $t = 200 \text{ s}$  after the start of the relativistic ejection. In all the following, the origin of time is the actual ejection, meaning that our first snapshot is slightly after 200 s. As such, we do not model the initial acceleration of the ejecta.

Fig. 2 provides a schematic view of different initial set-ups of the Lorentz factor. In the most general case the ejecta display variability in both the head and tail region (blue solid line) and we have

$$\begin{aligned} \Gamma_5(r) &= \Gamma_\infty, \\ \Gamma_4(r) &= \left(\Gamma_0 - (\Gamma_\infty - \Gamma_0) \frac{r - R_\Delta}{\alpha}\right) \left[1 + A \sin\left(2\pi \frac{r - R_\Delta}{\alpha}\right)\right], \\ \Gamma_3(r) &= \Gamma_0, \\ \Gamma_2(r) &= \Gamma_0 \left[\frac{1+k}{2} + \frac{1-k}{2} \cos\left(\pi \frac{r - r_\beta}{R_0 - r_\beta}\right)\right], \\ \Gamma_1(r) &= 1, \end{aligned} \quad (5)$$

with

$$r_\alpha = R_0 + (\alpha - 1)\Delta, \quad (6)$$

**Table 1.** Parameters for the simulations.

Name	$A$	$k$	$\alpha$	$\beta$	$n_0$ (cm $^{-3}$ )	Description
<i>run1</i>	0	1	0	1	$10^3$	Uniform
<i>run2</i>	0	1	0.5	1	$10^3$	Tail
<i>run3</i>	0	0.25	0.5	0.2	$10^3$	Tail+head
<i>run4</i>	0.6	1	0.5	1	$10^3$	Variable tail
<i>run5</i>	0.6	0.25	0.5	0.2	$10^3$	Variable tail+head
<i>run5b</i>	0.6	0.25	0.5	0.2	32	Variable tail+head

$$r_\beta = R_0 + (\beta - 1)\Delta. \quad (7)$$

Table 1 lists the values of  $A$ ,  $k$ ,  $\alpha$ , and  $\beta$  in our simulations. We always adopt  $\Gamma_0 = 100$  and  $\Gamma_\infty = 10$ . If the central engine switches off smoothly, one could expect  $\Gamma_\infty \simeq 1$ . However, this would give a very long deceleration time and radius that are very hard to model numerically. As we are mostly interested in the X-ray variability in the early afterglow phase, our high value is a reasonable approximation. The passive scalar tracks the evolution of different initial regions of the ejecta. The values are indicated on the top row of Fig. 2. The external density is set to  $n_0 = 1000$  cm $^{-3}$  except for *run5b*, which is identical to *run5* but has  $n_0 = 32$  cm $^{-3}$ . The high external density for most of the runs yields a small deceleration radius  $R_{\text{dec}}$ , making it possible to run our simulations. As we find no quantitative difference between the outcome of *run5* and *run5b*, we expect the main conclusions of our work to be valid for a more realistic external density.

The initial outer edge of the box is located at  $L_{\text{box}}/c = 600$  s and we use  $N = 96\,000$  grid cells. This gives a final resolution of  $\Delta r = 1.7 \times 10^8$  cm, equivalent to a temporal resolution of about 5 ms. Tests with  $N = 48\,000$  and  $192\,000$  grid cells for *run4* show that the structure of the FS is identical in all cases but the internal shocked regions are denser and narrower at higher resolution. However, the Lorentz factor is the same in the whole box for all tests. When comparing the final light curves, we find no difference for the luminosity of the FS beyond  $t_{\text{obs}} = 10$  s. We find that the total luminosity of the other shocks is higher at higher resolution, with a time-independent offset of about 13 per cent between the  $N = 48\,000$  and  $96\,000$  runs and a 5 per cent offset between the  $N = 96\,000$  and  $192\,000$  runs. In all cases, the slopes and variations in the light curves are very similar, indicating that our simulations have enough resolution to study the impact of an LLRS on the light curves. The parameters we choose are a compromise between the extreme nature of the ejecta and numerical constraints.

We stop the simulation at the end of the early afterglow phase, about  $2 \times 10^6$  s after the initial ejection. At that stage, most of the internal structure of the ejecta has been washed out by shocks and the later evolution would tend to the Blandford–McKee solution (Blandford & McKee 1976) and then a Sedov solution in the non-relativistic limit (Sedov 1969).

#### 2.4 Shock detection algorithm

To derive light curves from the hydrodynamic structure of the ejecta, we need to determine the energy dissipated in all shocked regions. Common methods to detect shocks include finding important jumps in density and/or pressure or searching for compressed regions ( $\nabla \cdot v < 0$ ; Corrales, Haiman & MacFadyen 2010). Such methods are only successful for very strong shocks, and fail to detect internal shocks in the ejecta. Therefore, we use a relativistic shock finding algorithm, inspired by Zanotti et al. (2010). The algorithm

is based on the predictions of the wave patterns of the relativistic Riemann problem (Rezzolla & Zanotti 2002; Rezzolla, Zanotti & Pons 2003). A shock occurs when the velocity gradient  $\Delta v$  between two adjacent cells is greater than a certain threshold  $\tilde{v}$ , determined by the thermodynamics of the fluid. In practice, for both directions of propagation, we compute the local minima of  $V = \tilde{v} - \Delta v$  to determine the locations of the shocks. To avoid spurious detections of very weak shocks, we smooth  $V$  over 10 computational cells. This step was not present in Zanotti et al. (2010), who instead suggested to include a small correction to  $V$  to avoid spurious detections.

The determination of the up- and downstream properties of the flow is less precise, as the shocks are smeared out due to numerical diffusion and may interact with each other and the contact discontinuity between the ejecta and the shocked external medium. The spreading of the shocks is most notable at very early times, when shocks are still forming. For a forward/backward propagating shock, we find the cell of maximal density (corrected for the spherical geometry) in the region between 150 and 100 cells before/after the shock location. The downstream density, pressure, and Lorentz factor are set to the values in that cell. For the upstream variables, we consider the mean over the cells between 100 and 150 after/before the shock. Visual inspection of the shock locations and conditions confirms the accuracy of the method for  $t > 10^5$  s. At earlier times visual determination of the up- and downstream regions is somewhat arbitrary. However, the hydrodynamic variables we recover respect the Rankine–Hugoniot jump conditions within 10 per cent after  $t = 10^4$  s. We tested alternate methods to find the downstream variables at earlier times and found no impact on the light curves beyond  $t_{\text{obs}} = 10$  s. As such, we are confident that our method yields reliable values for the up- and downstream regions and luminosity at the time-scales relevant to our study.

The detection of the contact discontinuity between the ejecta and the external medium is made possible with the passive scalar, which values are shown in Fig. 2. For the runs without head region (*run1*, *run2*, and *run4*) we set the position of the contact discontinuity  $r_{\text{CD}}$  to be the cell where the passive scalar is the closest to 2. For the runs with head region (*run3* and *run5*), we set  $r_{\text{CD}}$  to the cell where the passive scalar is closest to 1.5.

#### 2.5 Limits of 1D RHD simulations

The RHD simulations and ballistic model we present are one-dimensional, which enables us to cover a wide range of parameters and study the evolution of the ejecta from the internal shock phase until the deceleration phase. In the ultrarelativistic ejecta we are considering ( $\Gamma \geq 10$  at the end of the simulation), the lateral expansion of the jet can be safely neglected and our 1D simulations provide a good model of the global dynamics. Multidimensional simulations show the development of hydrodynamic instabilities (Aloy et al. 1999; Duffell & MacFadyen 2013) at the contact discontinuity between the outside medium and the ejecta. While the latter lead to important mixing, they do not strongly affect the structure of the internal and reverse shocks. Duffell & MacFadyen (2013) show that RS is pushed away from the FS and its emission is more delayed. We do not expect this to qualitatively impact the results presented here.

Our models also neglect the dynamical impact of the magnetic fields. While the traditional fireball model relies on thermal pressure to accelerate the ejecta, magnetic fields are also promising candidates for the acceleration and collimation of the ejecta during the early phases (see Granot et al. 2015, for a review of magnetic fields in GRB). At later stages, when the magnetization is of order

unity or above, the propagation of shocks is suppressed (Zhang & Kobayashi 2005; Giannios, Mimica & Aloy 2008; Mimica, Giannios & Aloy 2009, 2010; Mizuno et al. 2009). In such case, one has to invoke magnetic reconnection to explain the observed emission (Narayan, Kumar & Tchekhovskoy 2011; Zhang & Yan 2011). As the scenario studied in this paper relies on a complex dynamical evolution related to several generations of shocks, we clearly assume that the ejecta at large distance of the source have a low magnetization ( $\sigma \leq 0.1$ ).

In the set of simulations presented here, our assumptions on the initial structure of the ejecta are particularly important, as they lead to an internal structuration after the propagation of internal shocks, which will eventually produce flares in the light curve during the propagation of the LLRS. While we have considered highly idealized cases (see Fig. 2), negative gradients in the Lorentz factor are likely to arise naturally in the flow and a tail of slower material and can be expected from a fading central engine and/or the breakout of the ejecta through the stellar envelope in long bursts (see e.g. Duffell & MacFadyen 2015). More self-consistent modelling of the flow dynamics would require three-dimensional simulations, which come at prohibitive cost for the study presented here.

## 2.6 Ballistic model

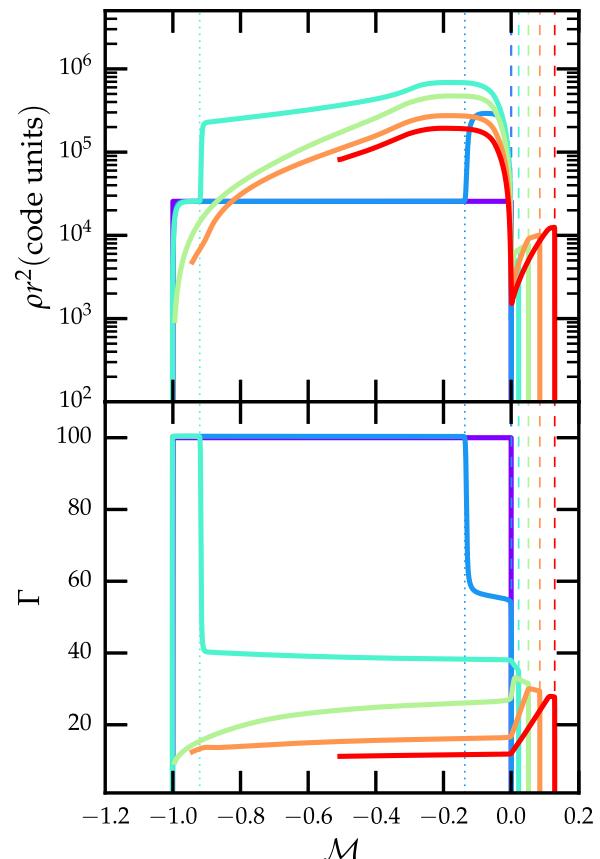
If the magnetization is small, the GRB ejecta at large distance from the source are dominated by its kinetic energy. It is then possible to model it using a ballistic approach where pressure waves are neglected. This method is described by Daigne & Mochkovitch (1998) for the internal shock phase: the outflow is discretized in a large number of shells that interact only by direct collisions. A sequence of collisions models the propagation of a shock wave. The comparison with a fully RHD simulation shows that the ballistic model recovers most features of the evolution (Daigne & Mochkovitch 2000), except for the density in the shocked regions, which is usually underestimated.

This approach has been extended by Genet et al. (2007) and Hascoët et al. (2011, 2012, 2014) to include the deceleration phase. This allows to follow the FS in the external medium and the RS in the ejecta. The limitations are the same: the radius and Lorentz factor of the different shocks are correctly estimated, as well as the dissipated power, but the density is underestimated. The advantage of the ballistic approach is its very low computational cost, which allows to explore a large domain of parameters for the model. The scenario where afterglow flares are produced by the interaction of an LLRS with dense shells in the ejecta has been explored using the ballistic approach by Hascoët et al. (2017). The promising results motivated the present study. In the following, we compare our results obtained with the fully RHD code described above with the same ejecta modelled with the ballistic approach and identify the different features that will contribute to the light curves.

## 3 HYDRODYNAMICS OF THE EJECTA

### 3.1 Uniform ejecta

Fig. 3 shows the density and Lorentz factor in *run1*, which we consider as our reference model. The initial state (purple line) corresponds to a uniform shell of width  $ct_w = 3.0 \times 10^{12}$  cm with a Lorentz factor  $\Gamma_0 = 100$ , an energy  $E_0 = 10^{53}$  erg, and a mass  $M_0 = \frac{E_0}{\Gamma_0 c^2} = 1.1 \times 10^{32}$  g. After describing the evolution in this simple case, we will progressively describe and explain the impact of various features in the set-up on the dynamical evolution of the



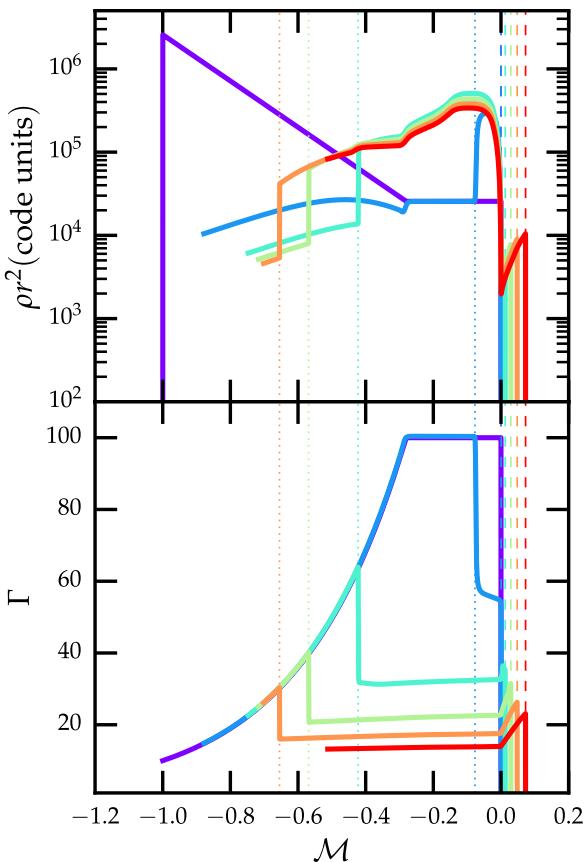
**Figure 3.** Hydrodynamics in *run1* with a uniform ejecta. Density (top) and Lorentz factor (bottom) as a function of the Lagrangian mass at  $t = 0, 2 \times 10^5, 5.2 \times 10^5, 6.8 \times 10^5, 8.0 \times 10^5$ , and  $9.17 \times 10^5$  s (from purple to red). The contact discontinuity between the ejecta and the external medium is located at  $M = 0$  by definition. The forward and reverse shocks are shown by dashed and dotted vertical lines, respectively. We have multiplied the density by  $r^2$  to remove the impact of spherical expansion and allow a better focus on shocked and expanding regions. After  $t = 6.8 \times 10^5$  s the inner part of ejecta is out of the simulation domain.

ejecta. We use a Lagrangian description to facilitate identification of different discontinuities. The  $x$ -axis shows the accumulated mass fraction normalized by the initial mass of the ejecta:

$$\mathcal{M}(R) = \frac{1}{M_0} \int_{r_{CD}}^R 4\pi r^2 \rho \Gamma dr. \quad (8)$$

As such, we use negative values for the ejecta and positive values for the shocked external medium.

With our initial conditions, the spreading radius of the ejecta is of the order of  $R_{\text{spread}} \simeq \Gamma_0^2 \Delta \simeq 3.0 \times 10^{14}$  cm and the deceleration radius  $R_{\text{dec}} \simeq (\frac{3}{4\pi} \frac{\dot{E}_{tw}}{\Gamma_0^2 n_0 m_p c^2})^{1/3} \simeq 1.2 \times 10^{16}$  cm. The condition  $R_{\text{spread}} \ll R_{\text{dec}}$  corresponds to the case where the RS is initially non-relativistic, becomes progressively relativistic, and crosses the ejecta at  $R_{\text{cross}} \simeq R_{\text{spread}}^{1/4} R_{\text{dec}}^{3/4} \simeq 4.8 \times 10^{15}$  cm  $\simeq 0.4 R_{\text{dec}}$ , before the deceleration radius (Sari & Piran 1995; Kobayashi, Piran & Sari 1999). This is indeed observed in our simulation. After  $t = 2 \times 10^5$  s, i.e.  $R \simeq 0.5 R_{\text{dec}}$  (blue line), the FS is still very close to the contact discontinuity and cannot be distinguished on the plot. On the other hand, the RS has already crossed almost 20 per cent of the mass of the ejecta. At  $t = 5.2 \times 10^5$  s, i.e.  $R \simeq 1.3 R_{\text{dec}}$ , the RS has almost finished to cross the ejecta, and the FS has just started to propagate: the shocked external medium has a mass of a few



**Figure 4.** Hydrodynamics in *run2*, which includes a slower tail region. The time steps and colours are the same as in Fig. 3.

per cent of the initial mass of the ejecta, i.e. is of the order of  $M_0/\Gamma_0$  (dark green line).

After  $t \simeq 8 \times 10^5$  s, i.e.  $R \simeq 2R_{\text{dec}}$ , the FS has accumulated a significant amount of external medium, and conversely, some of the initial ejecta has now left the simulation box on the left-hand side (orange line). At  $t = 9.17 \times 10^5$  s, 50 per cent of the initial mass of the ejecta has left the simulation domain (red line). The subsequent evolution is only driven by the FS and we recover the temporal evolution from analytic solution of Blandford & McKee (1976). However, as the shocked external medium accumulates more mass and becomes wider, it cannot be modelled properly within the size of our simulation box. After  $t \simeq 1.8 \times 10^6$  s, i.e.  $R \simeq 4.5R_{\text{dec}}$ , our simulation diverges from the Blandford & McKee (1976) and we discard its subsequent evolution.

### 3.2 Long-lived reverse shock

The luminosity of the afterglow may be altered if the RS is long-lived. Therefore, we reduce the speed of the second half of the ejecta, with a linearly decreasing Lorentz factor in *run2*. Fig. 4 shows the evolution of *run2*, at the same times as *run1*. As the energy injection in both ejecta is set to be identical, the total mass in the ejecta in *run2* is about twice as large as in *run1*. The gradient in the Lorentz factor causes the ejecta to expand and form a tail of slow material, and a fraction of the ejecta is rapidly advected out of the simulation domain. As expected for this initial state, we observe an LLRS: (i) before the RS reaches the tail of the ejecta around  $t = 4 \times 10^5$  s (transition happens between blue and green

lines), the evolution downstream of the RS is exactly the same as for *run1*; (ii) afterwards, both runs evolve differently. Because of the slower and denser tail region, the propagation of the RS is slowed down with respect to the uniform case and it leaves the simulation region at  $t = 8.7 \times 10^5$  s (compared to  $5.5 \times 10^5$  s in *run1*, transition happens between orange and red lines). Similarly, the slower tail region results in a slower FS.

### 3.3 Internal shocks

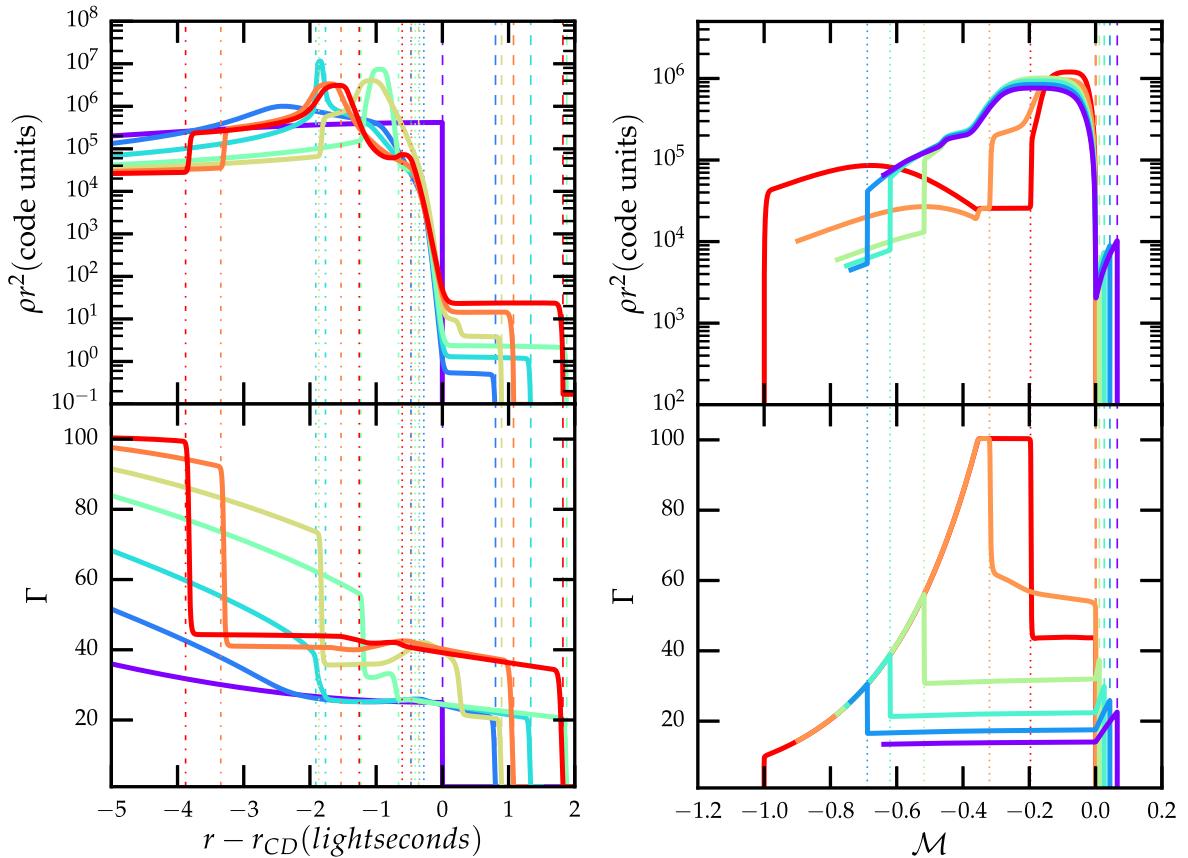
Internal shocks occur when the gradient of the Lorentz factor is strictly negative meaning faster material is behind slower material. We explore two different cases: when a head region is present ahead of the fastest part of the ejecta as in *run3*, or when a sinusoidal variation in the Lorentz factor is present in the tail region as in *run4*. In both cases, we expect internal shocks to form and structure the ejecta, but the interaction of these shocks with the RS should occur much earlier in *run3*. This is confirmed by the simulations.

Fig. 5 shows the evolution of *run3*, with the left-hand panel focusing on the early development of an internal shock. We used Eulerian coordinates as they better highlight different discontinuities in the head region at this early stage. Initially, the Lorentz factor profile is very shallow, but it quickly steepens as faster material is behind slower material. We expect the formation of internal shocks at a radius  $R_{\text{is}} \simeq 2(k\Gamma_0)^2 ct_{\text{var}}$ , where  $t_{\text{var}}$  is the variability time-scale in the initial state, of the order of  $0.5\beta t_w$ , i.e.  $R_{\text{is}} \simeq 4 \times 10^{14}$  cm. Indeed, we observe that after roughly  $10^4$  s, a denser shocked region has developed and two internal shocks appear (dot-dashed lines). However these shocks do not have time to propagate as, at  $t = 1.5 \times 10^4$  s, the RS reaches the outer edge of the dense shocked region (transition happens between green and yellow lines). This results in a reflected shock propagating forward, as can be seen at  $t = 1.8 \times 10^4$  s (yellow line). This internal shock crosses the contact discontinuity at  $t = 1.6 \times 10^4$  s (transition happens between yellow and orange lines) and eventually catches up with the FS in the external medium at  $t = 2 \times 10^4$  s, reaccelerating it. The later propagation of the RS is very similar to *run2*.

In *run4*, the negative Lorentz factor gradient is located in the tail region, where it has more time to develop into a dense shocked region before it becomes affected by the RS at  $t = 5 \times 10^5$  s. The left-hand panel of Fig. 6 shows this initial development, until  $t = 4.4 \times 10^5$  s. As the RS interacts with the outermost internal shock, a shock reflection occurs and results in a forward propagating shock (see yellow line), similarly to *run3*. The latter will propagate into the shocked external medium (see red line) and eventually catch up with the FS. The propagation of the RS lasts for almost  $10^6$  s, as it is significantly slowed down when crossing the dense region. When it eventually crosses the backward propagating internal shock, a reflected shock propagates forward again, which also eventually reaches the FS and slightly reaccelerates it (see red line). This evolution is in good agreement with the theoretical expectation described in Kumar & Piran (2000) and section 4 of Hascoët et al. (2017). As described in Hascoët et al. (2017), this interaction of the RS with the dense shell in the tail due to the formation of internal shocks will affect the luminosity and is a possible mechanism for observed flares in the afterglow.

### 3.4 Full evolution

Having identified the individual contributions of each feature, we analyse *run5*, which combines all previous features and provides



**Figure 5.** Hydrodynamics in *run3*, which includes a slower head region. Left: Eulerian view of the early evolution of the density and Lorentz factor at  $t = 0, 6 \times 10^3, 1 \times 10^4, 1.4 \times 10^4, 1.8 \times 10^4, 2.7 \times 10^4$ , and  $3.6 \times 10^4$  s (from purple to red). The dotted, dot-dashed, and dashed lines, respectively, show the position of the reverse, internal, and forward shocks. Right: Lagrangian view of the later evolution, showing the same quantities at  $t = 3.6 \times 10^4$  s, which is the latest time step on the left-hand panel (red), and  $2 \times 10^5, 5.2 \times 10^5, 6.8 \times 10^5, 8.0 \times 10^5$ , and  $9.17 \times 10^5$  (from orange to purple). The time steps on the right are the same as in Figs 3 and 4.

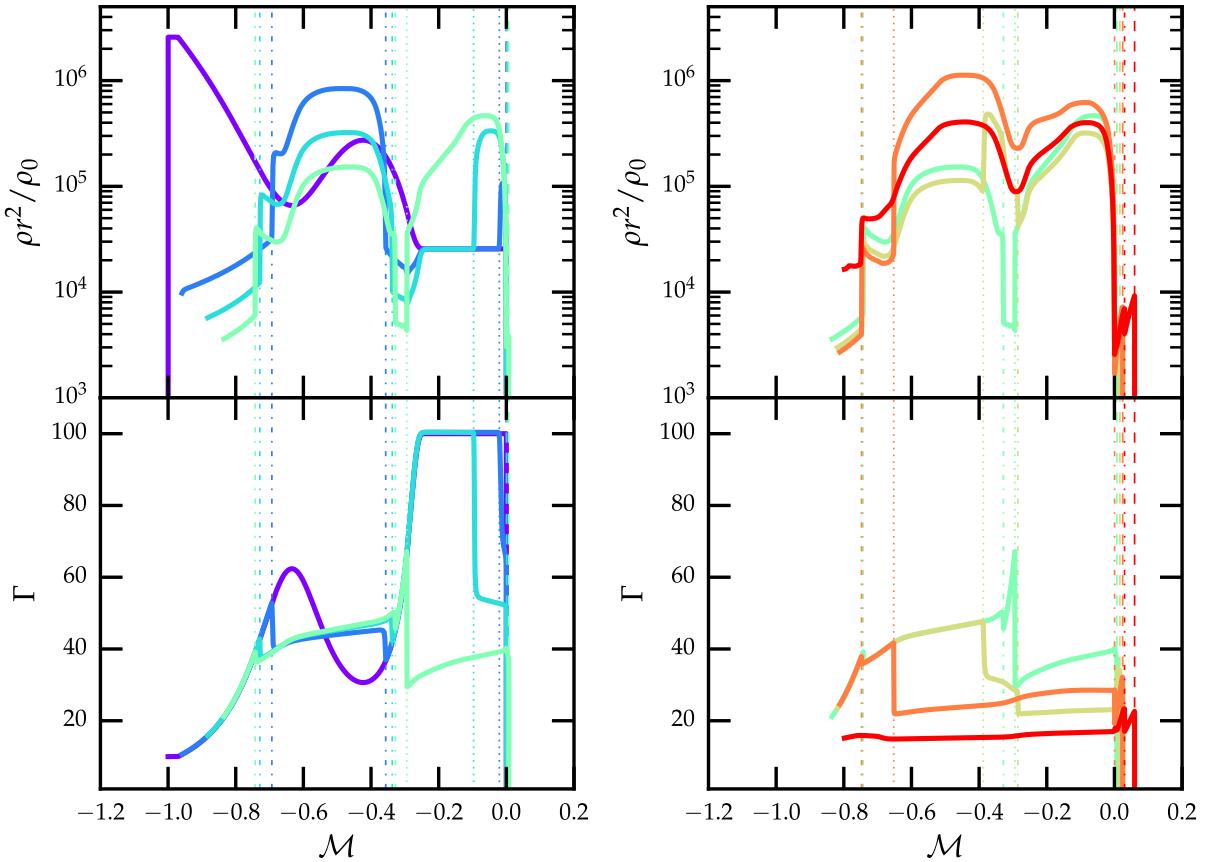
a more complete insight on the evolution of a realistic GRB afterglow. The global evolution is essentially a combination between the impact of the early internal shocks present in *run3* and the later ones from *run4*. This set-up leads to two internal shock regions, which both have a forward propagating shock and backwards propagating shock. The interaction between the latter and the RS leads to reflected shocks, which propagate back to the FS.

Fig. 7 shows the Lorentz factor downstream each of these shocks over time, with the results of the ballistic model for the same initial state shown below. At  $t = 7 \times 10^3$  s, the internal shocks in the head region become apparent (label ‘1’ in Fig. 7) and persist until the RS interacts with its forward propagating shock (label ‘2’). This effectively damps the RS, and the internal reverse shock (IRS) is now considered as the RS. A reflected shock is swiftly propagating forward until it catches up with the FS (label ‘4’) and reaccelerates it significantly. In the meantime, a shocked region has formed in the tail region as well (label ‘3’). As the RS enters the tail region it is slowed down (label ‘5’) until it is accelerated when encountering the internal shock region (label ‘6’). Again, this interaction results in a forward propagating shock that is momentarily stalled at the contact discontinuity with the shocked external medium (label ‘7’) and then catches up with the FS (label ‘10’). The RS eventually reaches the back shock of the internal shocked medium (label ‘8’) and then leaves the simulation domain (label ‘9’). In the next section we study the impact of the dynamics on the bolometric light curves of the ejecta.

Fig. 7 also highlights that in the ballistic model, the shocked regions behind the RS and the FS have the same Lorentz factor, as behind the internal forward shock (IFS) and IRS. This is due to the lack of resolution in this approach, where shells merge after each collision. Generally, the Lorentz factors in the simulations bracket the value from the ballistic model. The Lorentz factor of the FS is usually underestimated, while the RS is modelled more accurately. Conversely, the downstream density in the ballistic model (not shown here) seems well reproduced for the FS, but strongly underestimated for the RS. In the next sections we compare the resulting bolometric luminosity of both models.

#### 4 BOLOMETRIC EMISSION

In this section we compute bolometric light curves based on the different runs. The advantage is that this calculation is independent on any assumption on the microphysics in the various shocked regions. It shows where the power is dissipated and provides an indication of the shape of the observed emission at frequencies dominated by fast-cooling electrons, as it is expected for X-rays at early times. In addition, the bolometric light curves allows a direct comparison with the simulations made with the ballistic model. In Section 5, we will introduce microphysics parameters to provide a more specific discussion of the X-ray emission to explore the validity of a X-ray flare model based on an LLRS in a stratified ejecta.



**Figure 6.** Hydrodynamics in *run4*, which includes variability in the tail region. Left: early evolution showing  $t = 0, 10^5, 2.3 \times 10^5$ , and  $4.4 \times 10^5$  s (from purple to green). Right: later evolution showing  $t = 4.4 \times 10^5, 5.7 \times 10^5, 6.5 \times 10^5$ , and  $8.7 \times 10^5$  s (from green to red).

#### 4.1 Simplified emission model

We compute the bolometric light curves for our different models and determine the contributions of the FS, internal shocks, and RS. At a given  $t_{\text{obs}}$  in the observer frame, the received bolometric luminosity from one shocked region is given by (Woods & Loeb 1999)

$$L_{\text{bol}}(t_{\text{obs}}) = \int dt \frac{2\epsilon_e A(\theta) L_{\text{diss}}(t)}{\Delta t_{\text{obs}}(t)} \frac{1}{\left(1 + \frac{t_{\text{obs}} - t_{\text{obs},0}(t)}{\Delta t_{\text{obs}}(t)}\right)^3}, \quad (9)$$

where the integration over  $t$  (lab frame) is carried out during the time interval corresponding to the propagation of the considered shock. In equation (9),  $L_{\text{diss}}(t)$  is the dissipated power at the shock,  $\epsilon_e$  is the fraction of the dissipated energy injected in non-thermal electrons,

$$t_{\text{obs},0} = t - \frac{r(t)}{c} \quad (10)$$

is the observer time for the reception of the first photons emitted on-axis at time  $t$  and radius  $r(t)$ , and

$$\Delta t_{\text{obs}}(t) = \frac{r(t)}{2\Gamma^2 c} \quad (11)$$

is the observed time delay between photons emitted at time  $t$  either on-axis or at an angle  $1/\Gamma$ .

The dissipated energy at a given shock  $L_{\text{diss}}(t)$  is set by the variation of the internal energy between the downstream and upstream regions:

$$L_{\text{diss}}(t) = \dot{M}_s \Gamma_*(\epsilon_* - \epsilon) \quad (12)$$

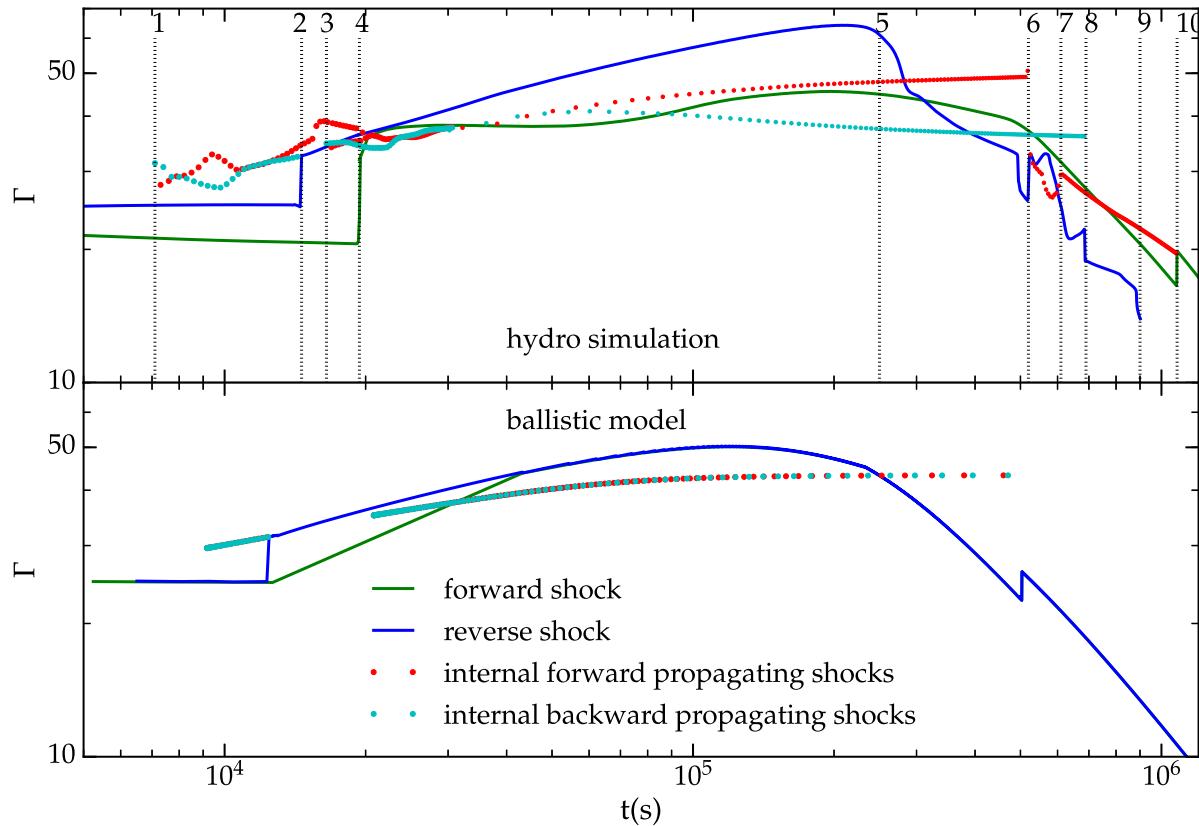
$$= 4\pi r^2 \rho_* \Gamma_*^2 (V_s - v_*)(\epsilon_* - \epsilon), \quad (13)$$

where the subscript ‘\*’ indicates quantities in a shocked region,  $\dot{M}$  is the mass flow across the shock,  $\epsilon$  the specific internal energy, and  $V_s$  is the velocity of the shock, which is determined by comparing its position between outputs.

Compared to Woods & Loeb (1999), the function  $A(\theta)$  has been introduced in equation (9) by Beloborodov et al. (2011) to account for a possible anisotropy of the synchrotron radiation in the comoving frame.  $\theta$  is the angle with respect to the radial direction in the comoving frame, such that

$$\cos \theta = \frac{\Delta t_{\text{obs}}(t) - (t_{\text{obs}} - t_{\text{obs},0}(t))}{\Delta t_{\text{obs}}(t) + (t_{\text{obs}} - t_{\text{obs},0}(t))}. \quad (14)$$

In this section, we limit our study to the bolometric light curve in the simplest case, where the emission in the comoving frame is isotropic, i.e.  $A(\theta) = 1$ . In the early X-ray afterglow and during flares, the radiating electrons are expected to be in fast-cooling regime, and therefore the bolometric light curve already gives a fair idea of the shape of the predicted emission. We introduce micro-physics parameters and discuss the effect of anisotropy in Section 5. In practice, for each output (corresponding to a given time  $t$ ), we determine the location of the different shocks, and then for each of them determine  $t_{\text{obs},0}$  and  $\Delta t_{\text{obs}}$  and the quantities in the shocked region to compute the corresponding  $L_{\text{diss}}$ . To compute the bolometric light curve due to a given shock, for each  $t_{\text{obs}}$ , we add the



**Figure 7.** Lorentz factor in *run5*, which includes variability in the tail region and a slower head region. The upper and lower panels show the Lorentz factor in the shocked material behind the forward shock (FS), reverse shock (RS), and the internal forward (IFS) and backward propagating shocks (IRS) in the hydrodynamic simulation and ballistic model, respectively. Sudden variations are indicated with numbers and vertical lines and are detailed in the text.

contributions of all  $t$  using equation (9). We eventually add up the contributions of the different types of shocks.

#### 4.2 Bolometric light curves

Fig. 8 shows the bolometric light curves for all runs and the result of the ballistic model. The left-hand column shows only the contribution of the FS propagating in the external medium. In *run1*, its bolometric luminosity initially increases until the head of the ejecta starts decelerating at  $t_{\text{obs}} \simeq 200$  s, after what it follows the expected self-similar solution with a  $t_{\text{obs}}^{-1}$  slope (Blandford & McKee 1976). When the head of the ejecta is slower (*run3* and *run5*), the initial luminosity is naturally lower and suddenly rises after  $t_{\text{obs}} = 20$  s when the FS is reaccelerated by the reflected FS (label ‘4’ in Fig. 7). The FS is more luminous than for *run1* for a brief moment and the deceleration occurs earlier. Eventually, all runs converge to the same self-similar solution, consistent with the total amount of injected energy. The bump in *run4* and *run5* around  $t_{\text{obs}} \simeq 300$  s is caused by the reacceleration of the FS due to the IFS (the latter resulting from the reflection of the RS on the internal shock region, label ‘10’ in Fig. 7). The luminosity curves are in good agreement with the ballistic model shown below. In all cases, the simulations yield a slightly later deceleration of the ejecta.

The right-hand column of Fig. 8 shows the sum of the contributions of all other shocks, which propagate within the relativistic ejecta. This includes the RS, and the internal shocks in the tail and head regions, and represents the total internal dissipation. As suggested by the complex evolution of the Lorentz factor in Fig. 7, the

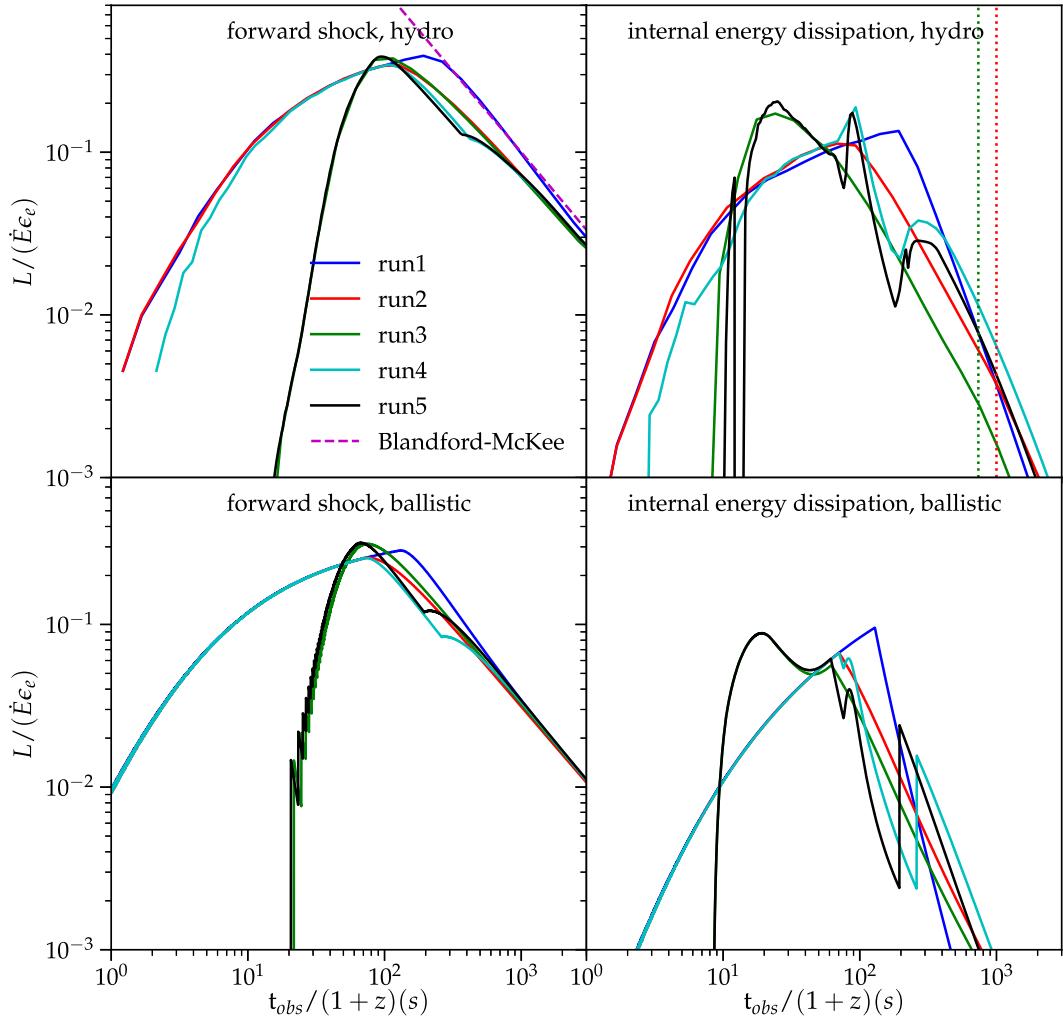
bolometric light curves are not smooth and present a much higher variability than the light curves from the FS. This important variability, which will be fully described in Section 5, confirms our initial motivation for this study.

For *run1*, the luminosity of the RS progressively increases until it reaches the inner edge of the ejecta at  $t = 5.2 \times 10^5$  s, corresponding to  $t_{\text{obs}} = 200$  s, after which only off-axis photons are received and the flux follows the expected  $t_{\text{obs}}^{-3}$  decrease. For *run2*, which includes a slower tail region and yields an LLRS, the decrease in luminosity starts earlier, when the RS reaches the tail region. As the ejecta is more extended in this case, the RSs remain active for a longer time and the luminosity ends up being larger than for *run1*. In both cases, the ballistic model very well reproduces the simulation, with a slightly lower and steeper drop-off for the luminosity.

When the head region is present, as in *run3* and *run5*, the RS is slower, and the emission very steeply rises around  $t_{\text{obs}} = 10-20$  s, when the RS encounters the shocked region (label ‘2’ in Fig. 7). The following decrease in flux is initially slow but recovers the same slope than *run2* after roughly 100 s. The small upturn seen around  $t_{\text{obs}} = 70$  s is not recovered in the ballistic model.

In *run3* and *run5*, the impact of the internal shocks in the head region is rapidly suppressed by the propagation of the RS and this configuration is unlikely to produce important variability.<sup>1</sup> On the

<sup>1</sup> However, variations of the Lorentz factor on shorter time-scales would lead to earlier internal shocks, which could propagate without being smoothed out by the RS. As our study does not focus on the prompt phase, we did not simulate such cases.



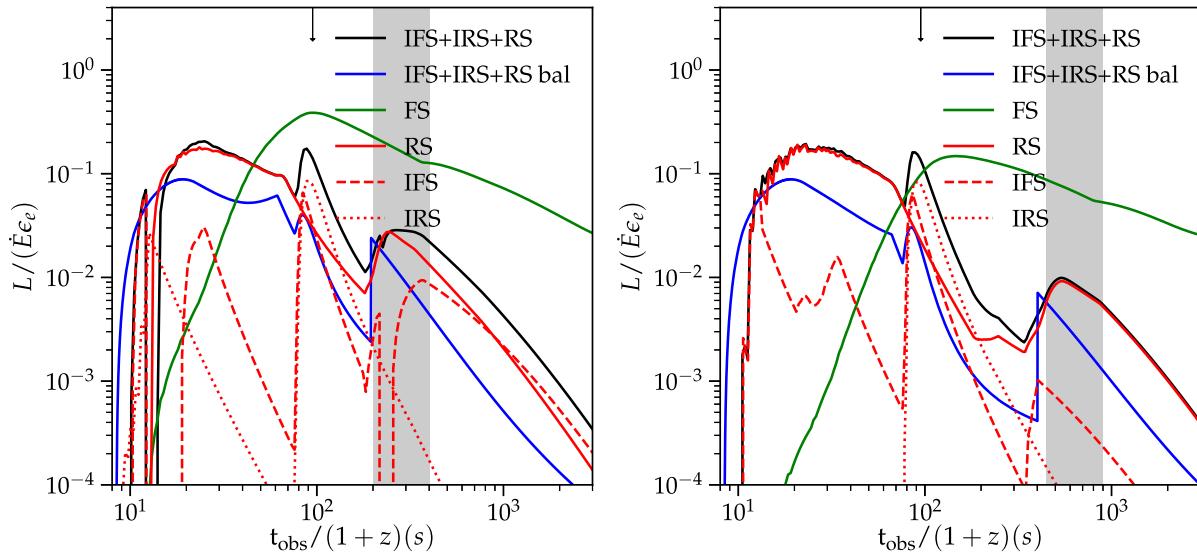
**Figure 8.** Bolometric light curves of the forward shock (left-hand column) and internal and reverse shocks (right-hand column) assuming isotropic emission in the comoving frame. The five main runs are shown (colour coded), comparing the simulations (upper row) and the ballistic model (lower row). On the top left-hand panel, the magenta dashed line shows the  $t_{\text{obs}}^{-1}$  (Blandford & McKee 1976) self-similar solution. The dotted vertical lines indicate the time beyond which some of the off-axis emission is missing because the reverse shock has left the simulation domain. As seen in the right-hand column, internal dissipation leads to flare-like features for  $t_{\text{obs}} \geq 70$  s in *run4* and *run5*.

other hand, internal shocks in the tail region produce a much stronger effect, as can be seen in *run4*. The initial evolution is identical to *run2* and the impact of the internal shock region becomes apparent around  $t_{\text{obs}} = 100$  s, just after the dimming of the RS as it enters the tail region. As the RS interacts with the internal shocked region, it gets revived and results in the reflected forward propagating shock. Both these effects account for the rebrightening at  $t_{\text{obs}} \simeq 250$  s, with the main contribution from the reflected shock. The ballistic model shows qualitatively similar results, although with lower emission at all times.

Fig. 9 shows the various contributions of the internal and reverse shocks to the bolometric light curve in the most realistic simulations: *run5* (left) and *run5b* (right), which is the same run with a lower external density. As described above, several regions can lead to flare-like features: the IFS, which produces the narrowest spikes, as suggested in Hascoët et al. (2017), and the RS. In Section 5 we will detail the observable properties of all the variability observed in our *run5* and *run5b*, and compare with the properties of the flares observed in the X-ray afterglow.

#### 4.3 Validity of the ballistic model

The first goal of this study is to establish the validity of the ballistic model used by Hascoët et al. (2017) to describe the dynamics of GRBs ejecta and more specifically flares in the early afterglow phase. The bolometric light curves in Fig. 8 show a good qualitative agreement in all cases between the ballistic model and simulation. The light curves show similar slopes, timing for flares, and relative luminosity between different configurations. However, we systematically find an increased luminosity in the hydrodynamic simulations for the reverse and internal shocks, and a slight delay for the deceleration of the FS. The increased luminosity may be due to the uncertainty in the definition of the upstream and downstream media. However, as we have detailed in Section 2.4, we are confident that our shock detection method yields the proper values soon after the beginning of the simulation. We have also tested that artificially increasing or decreasing  $L_{\text{diss}}$  and/or  $\Delta t_{\text{obs}}$  (see equation 12) at early times does not affect the luminosity at later times and cannot account for the discrepancy. As such, we expect that the difference between the ballistic model and our simulations reveals the



**Figure 9.** Contribution to the bolometric light curves of the different shocks in *run5* (left) and *run5b* (right). We separately show the energy dissipation at the forward shock (FS) and the sum of all the internal dissipation (IFS+IRS+FS), and compare with the ballistic model (blue). The flares and rebrightenings are shown by arrow and shaded regions, respectively.

weaknesses of the ballistic model. Specifically, the ballistic model underestimates the density in the shocked medium due to a poor spatial resolution intrinsic to the method, as already discussed in Daigne & Mochkovitch (2000). In addition, shock reflections at late times, which strongly impact the late flare in *run4* and *run5* at  $t_{\text{obs}} \geq 70$  s as seen in Fig. 9, are missing in the ballistic model. Their impact was predicted by Hascoët et al. (2017, see their section 4) and is confirmed here.

Despite these discrepancies, the overall shapes of the curves and their respective order is the same in the model and the hydrodynamic simulations. This validates the determination of the FS in the ballistic model and our numerical set-up with moving boundaries. The ballistic model is ideally suited to explore a wide range of parameters at negligible computational cost. However, the small quantitative differences may hinder the derivation of microphysical parameters based on direct comparison with observations.

## 5 X-RAY FLARES FROM AN LLRS IN A STRUCTURED EJECTA

Based on our 1D hydrodynamic simulations, we now explore the corresponding X-ray variability. We assume the X-ray emission results from synchrotron emission from electrons accelerated at the different shocks. Non-thermal electrons are assumed to follow a power-law distribution above a minimum Lorentz factor  $\gamma_m$ , with  $n(\gamma) \propto \gamma^{-p}$  and  $p = 2.3$ . Throughout this whole section, we consider following microphysical parameters in all internal shocked regions (RS, IFS, and IRS): (i) we assume that the accelerated electrons result from the injection of a fraction  $\epsilon_e = 0.1$  of the dissipated energy into a fraction  $\zeta = 0.01$  of the electrons, and we assume that a fraction  $\epsilon_B = 0.1$  of the energy is injected in the amplified magnetic field; (ii) we adopt the same parameters in the external FS, except that  $\epsilon_e, \epsilon_b$  are both divided by 50 to model a radiatively inefficient ultrarelativistic shock. As shown in Figs 8 and 9, the energy dissipated at the FS is higher than in the internal shocks (IFS+IRS+RS). As such, having a radiatively less efficient FS is a necessary condition for the flares to be observable in our model. As discussed in

Uhm & Beloborodov (2007), Genet et al. (2007), and Uhm et al. (2012), this can happen if the FS is initially radiatively inefficient in the ultrarelativistic regime, due to an inefficient acceleration of electrons and/or an inefficient amplification of the magnetic field in the shocked external medium.

The luminosity at observed frequency  $\nu_{\text{obs}}$  is given by a modified version of equation (9) (Woods & Loeb 1999):

$$L_{\nu_{\text{obs}}}(t_{\text{obs}}) = \int dt \frac{2\epsilon_e L_{\text{diss}}(t)}{\Delta t_{\text{obs}}(t) \nu_{p,\text{obs}}(t, t_{\text{obs}})} \mathcal{B}\left(\frac{\nu_{\text{obs}}}{\nu_{p,\text{obs}}(t, t_{\text{obs}})}\right) \times \frac{1}{\left(1 + \frac{t_{\text{obs}} - t_{\text{obs},0}(t)}{\Delta t_{\text{obs}}(t)}\right)^3}, \quad (15)$$

where  $\nu_{p,\text{obs}}$  is the peak frequency of the emission, in the observer frame,

$$\nu_{p,\text{obs}}(t, t_{\text{obs}}) = \frac{2\Gamma_*(t) \max(\nu_m(t), \nu_c(t))}{1 + \frac{t_{\text{obs}} - t_{\text{obs},0}(t)}{\Delta t_{\text{obs}}(t)}}. \quad (16)$$

The normalized spectral shape  $\mathcal{B}(x)$  in equation (15) depends on the characteristic frequencies  $\nu_m, \nu_c$  in the comoving frame. The latter are the synchrotron frequencies corresponding to the minimal electron Lorentz factor  $\gamma_m$ , and to  $\gamma_c$ , the Lorentz factor beyond which cooling is important, i.e.

$$\nu_{m,c} = \frac{3}{4\pi} B \frac{q_e}{m_e c} \gamma_{m,c}^2, \quad (17)$$

where

$$\gamma_m = \frac{\epsilon_e m_p}{\zeta m_e} \frac{p-2}{p-1} \frac{\epsilon_*}{c^2}, \quad (18)$$

$$\gamma_c = \frac{6\pi m_e c}{\sigma_T t_{\text{dyn}} B^2}, \quad (19)$$

with  $B = \sqrt{8\pi\epsilon_b\epsilon_*\rho_*}$  the magnetic field and  $t_{\text{dyn}} = \frac{r}{\Gamma_* c}$  the dynamical time-scale in the comoving frame of the shocked region.

Following Sari et al. (1998), we have for  $v_m \geq v_c$  (fast cooling):

$$\mathcal{B}(x) \propto \begin{cases} \left(\frac{v_c}{v_m}\right)^{-5/6} x^{1/3} & \text{if } x \leq \frac{v_c}{v_m}, \\ x^{-1/2} & \text{if } \frac{v_c}{v_m} \leq x \leq 1, \\ x^{-p/2} & \text{if } x \geq 1, \end{cases} \quad (20)$$

and for  $v_m \leq v_c$  (slow cooling):

$$\mathcal{B}(x) \propto \begin{cases} \left(\frac{v_m}{v_c}\right)^{-p/2+1/6} x^{1/3} & \text{if } x \leq \frac{v_m}{v_c}, \\ x^{-(p-1)/2} & \text{if } \frac{v_m}{v_c} \leq x \leq 1, \\ x^{-p/2} & \text{if } x \geq 1. \end{cases} \quad (21)$$

In both cases, the function is normalized by  $\int_0^\infty \mathcal{B}(x) dx = 1$ .

Fig. 10 shows the contribution of all shocks to the light curve at 1 keV in *run5* and *run5b*. The left-hand panel shows our standard run, while the right-hand panel shows a simulation with a lower external density. This figure is the X-ray counterpart of the bolometric emission shown in Fig. 9. We clearly see three phases. (1) *Prompt emission phase*. In both cases, the very early X-ray light curves is dominated by internal shocks for  $t_{\text{obs}} \lesssim t_w = 100$  s. (2) *Early X-ray afterglow*. For our choice of microphysics parameters, the internal energy dissipation still dominates the emission for a few thousands seconds, mainly due to the activity of the RS, but also with significant contributions from the internal shocks. At this stage, the contribution of the FS to the emission is negligible. (3) *Standard afterglow*. At later times, the external FS dominates. The transition occurs at  $t_{\text{obs}} \gtrsim 1500$  s in *run5* and appears to be rather smooth, in agreement with observations. The exact time of the transition depends of course on our arbitrary choice of microphysics parameters and could be further delayed if ultrarelativistic shocks are strongly radiatively inefficient (i.e. for even lower values of  $\epsilon_e$  and/or  $\epsilon_B$  in the FS). In addition, the transition is also delayed for a lower external density, as clearly seen in *run5b*.

We now focus on the second phase, the early X-ray afterglow, to analyse the observed variability. In both *run5* and *run5b*, the ejecta produces two flares shown with an arrow and shaded grey region in Fig. 10. The first narrow spike at  $t_{\text{obs}} \simeq 90$  s is dominated by the emission from internal shocks (IFS and IRS) before any interaction with the RS, and is therefore almost identical in both simulations, in their timing and intensity. This confirms that late internal shocks can produce flares in the very early afterglow, especially those observed during the early steep decay phase. However, the same mechanism would require a long-lasting central engine to also produce late flares (Burrows et al. 2005b; Fan & Wei 2005; Zhang et al. 2006).

The second, weaker, flare occurs at later times,  $t_{\text{obs}} \simeq 250$  s in *run5* and  $t_{\text{obs}} \simeq 550$  s in *run5b*, i.e. at observer times much longer than the duration of the relativistic ejection by the central engine,  $t_w = 100$  s. For our choice of microphysics parameters in the internal regions, the emission is dominated by the RS. More precisely in *run5* the increased emission from the RS starts at  $t_{\text{obs}} = 181$  s, when it enters the overdense internal region (label ‘6’ in Fig. 7) and drops steeply after  $t_{\text{obs}} = 335$  s, when it exits the shocked region (label ‘8’ in Fig. 7). However, as mentioned above, the interaction of the RS with the shocked region also causes the reflection and some rebrightening of the IFS towards the front of the ejecta. The emission from the IFS abruptly drops at  $t_{\text{obs}} = 250$  s, when it has reached the FS (label ‘10’ in Fig. 7). This emission always remains well below the emission of the RS, but may contribute more significantly for another choice of microphysics parameters. As this second flare is due the interactions of the RS with the internal dense shells, it happens later in *run5b*, which has a lower external density and a later development of the RS. These light curves confirm the scenario proposed by Hascoët

et al. (2017) and illustrate the capacity of this scenario to produce late flares without invoking a long-lasting central engine.

As this second flare is the novel feature of the proposed scenario, we have carefully checked our assumptions on the radiative regime. Fig. 11 shows the characteristic frequencies  $v_{m,\text{obs}}$  and  $v_{c,\text{obs}}$  from  $t_{\text{obs}} = 70$  to  $10^3$  s and for the on-axis emission of the two dominant shocks in this period (RS/IFS in *run5* and RS/IRS in *run5b*). For our choice of microphysics parameters, it shows that during the second flare, the RS that dominates the flare emission is in fast cooling with  $h\nu_{\text{obs}} = 1$  keV  $> v_{m,\text{obs}} > v_{c,\text{obs}}$ . In *run5b*, the RS emission enters in slow cooling during the decay phase, but electrons radiating in X-rays are still fast cooling as  $h\nu_{\text{obs}} = 1$  keV  $> v_{c,\text{obs}}$ . This validates our assumptions to compute the observed light curve showing an X-ray flare.<sup>2</sup> Interestingly, it appears that in *run5* the IFS, the second brightest contributor, is also in fast cooling with a peak energy in gamma-rays ( $\sim 10$  MeV), which may produce a weak flare in high-energy gamma-rays. To check if the IFS could produce a gamma-ray flare such as the one detected by *Fermi*/Large Area Telescope (LAT) in GRB 100728A (Abdo et al. 2011; Troja et al. 2015), one would require a more detailed radiative calculation including the inverse Compton scatterings.

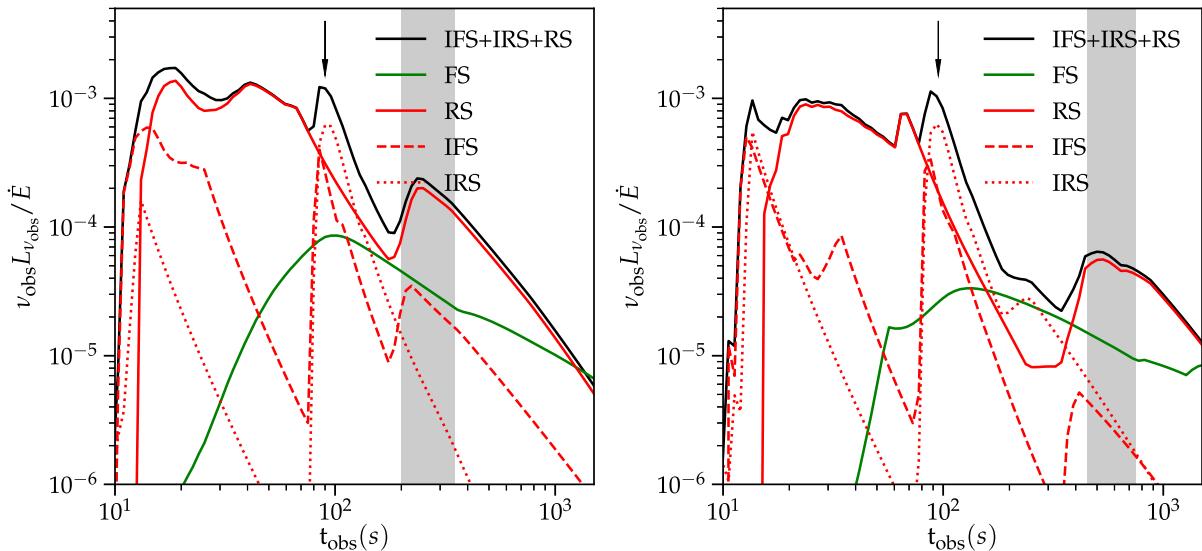
Observed X-ray flares typically have a width  $\Delta t_{\text{obs}}/t_{\text{obs}} \simeq 0.1\text{--}0.3$  (Chincarini et al. 2007), with a fast rise and a slower decay. However, in our simulations,  $\Delta t_{\text{obs}}/t_{\text{obs}} \simeq 0.7\text{--}1$ . A possible solution is an anisotropic synchrotron emission in the comoving frame, as suggested by Beloborodov et al. (2011). Indeed, equation (9) clearly shows that the flare cannot decay faster than  $t_{\text{obs}}^{-3}$  in the isotropic case, whereas a steeper slope can be obtained if  $A(\theta)$  decreases with  $\theta$ . Following Hascoët et al. (2017), we have considered the effect of a moderate anisotropy assuming  $A(\theta) \propto e^{\cos \theta / \cos \theta_0}$  with  $\theta_0 = 70^\circ$  (which leads to 90 per cent of the energy being beamed within  $\theta_0$ ). Fig. 12 compares the light curves for *run5* and *run5b* in the anisotropic and in the standard (isotropic) case. The latter shows more temporally peaked emission, with a higher peak luminosity resulting from limb darkening, in a much better agreement with observations. It remains to be understood if such a moderate anisotropy can be achieved in mildly relativistic shocks.

## 6 CONCLUSION

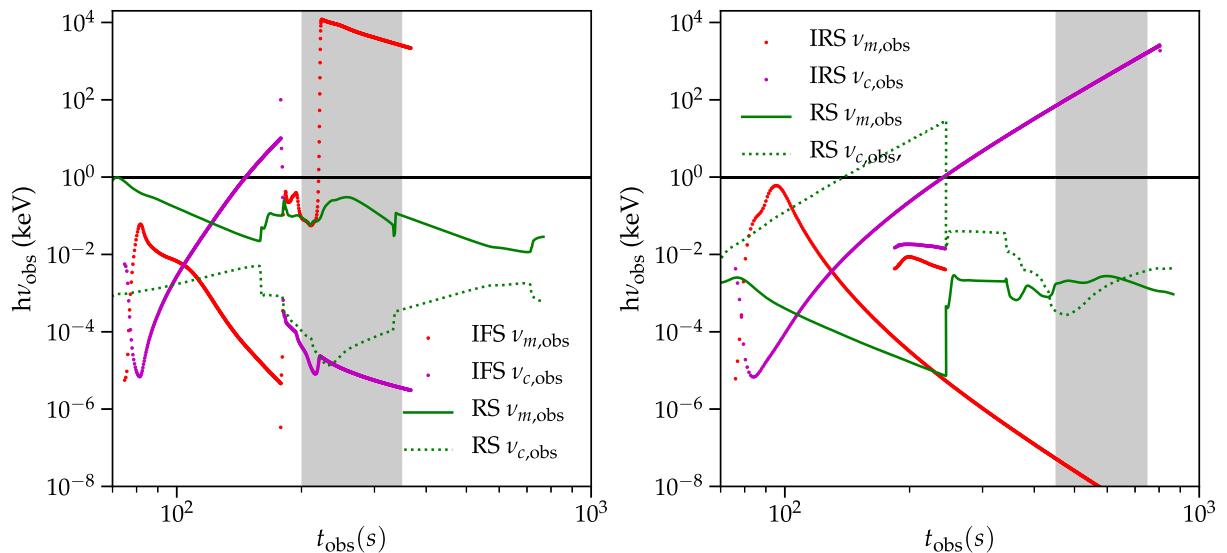
In this paper, we present a set of high-resolution one-dimensional RHD simulations of a stratified GRB ejecta interacting with a uniform external medium in order to study the variability of the early afterglow. We follow the power dissipated in each shocked region (external FS, internal shocks, and RS) and compute the resulting bolometric light curves, and X-ray light curve assuming synchrotron radiation of shock accelerated electrons in the fast-cooling regime.

The resulting bolometric light curves agree well with a ballistic model (Daigne & Mochkovitch 1998; Hascoët et al. 2017). In particular, the ballistic model is able to reproduce the slopes and relative variations of the light curves, albeit with a slightly lower total emission. However, our simulations provide a more accurate description of the dynamics: the ballistic model underestimates the

<sup>2</sup> Fig. 11 shows that, apart from the flares, X-ray photons can be produced by slow-cooling electrons, especially at late times. In this situation, the shape of the light curve should be computed more accurately as we assume here that radiating electrons are located at the shock (equation 9), whereas slow-cooling electrons are still radiating long after having been accelerated (see Beloborodov 2005; Uhm et al. 2012 for a more accurate method to compute the observed flux in this case). This tends to smooth the variability.



**Figure 10.** Separate contributions to the X-ray light curve in *run5* (left) and *run5b* with a lower external density (right). The total contributions of the internal energy dissipation are shown in black. The shaded regions and arrows indicate the flares. The light curves are computed at  $h\nu_{\text{obs}} = 1$  keV assuming a source redshift  $z = 1$ .



**Figure 11.** Characteristic synchrotron energies of the electrons at the shocks dominating the emission during the flares: RS/IFS in *run5* (left) and RS/IRS in *run5b* (right). The black line shows  $h\nu_{\text{obs}} = 1$  keV radiation for comparison. The shaded area shows the time interval of the flare.

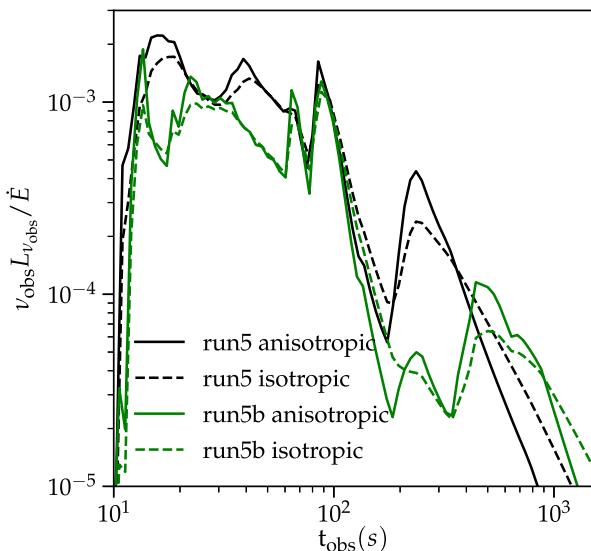
density in the shocked medium, as already pointed out by Daigne & Mochkovitch (2000), and neglects shock reflections at late times. This validates its use for preliminary studies, as its low computing time allows for the exploration of a large range of parameters, but points out its limitation that can impact the determination of micro-physics parameters when comparing the model with observations.

We have used these simulations to validate the scenario proposed by Hascoët et al. (2017) for the origin of X-ray flares in GRB afterglows. With a much more detailed description of the dynamics, our results confirm that the propagation of an LLRS in a stratified relativistic ejecta leads to the appearance of bright flares in the bolometric light curve of the internal dissipation (i.e. all shocks except for the external FS). The initial stratification of the ejecta naturally results from the internal shock phase. Internal shocks locally compress the ejecta and smoothen variations of the Lorentz factor, leading to the formation of dense shells with rather uniform

Lorentz factors. Early X-ray flares can be due to late internal shocks. However the most interesting feature is the appearance of late flares when the RS interacts with the dense shells in the ejecta.

Compared to the approach used by Hascoët et al. (2017), our simulations allow us to accurately model this interaction, which includes shock reflections that were predicted but could not be captured by the ballistic model. Adding an estimate of the synchrotron emission, we find that the variability in the bolometric light curves translates into X-ray flares with properties in agreement with observations. More specifically, the simulations show the following.

- (i) Late internal shocks can be a source of early variability in the afterglow.
- (ii) The interaction of the LLRS with a dense shell in the ejecta yields a strong flare-like rebrightening, with a fast rise and slower decay.



**Figure 12.** X-ray light curves resulting from internal energy dissipation (RS+IFS+IRS) for *run5* and *run5b*, assuming an anisotropic synchrotron emission in the comoving frame; see text. The dashed curves show the isotropic case for comparison, already shown in Fig. 10.

(iii) This bright flare is not only observed in the bolometric light curve, but also in X-rays assuming standard microphysics parameters. Indeed, the electrons radiating at 1 keV are fast cooling.

(iv) Assuming a moderate anisotropy of the synchrotron emission in the comoving frame, the width of this flare is of the order of  $\Delta t_{\text{obs}}/t_{\text{obs}} \sim 0.1\text{--}0.3$ , in agreement with the properties of observed X-ray flares.

(v) The time at which the flare is observed depends only on the initial properties of the ejecta (distribution of the Lorentz factor, kinetic energy) and on the external density. Flares can be observed at late time, without any need for a long-lasting central engine.

For this promising scenario to work, there is no constraint on the lifetime of the central engine as long as it is comparable to the GRB duration. On the other hand, several other assumptions are necessary.

(i) The GRB ejecta must be initially variable, to allow for the formation of internal shocks and the stratification of the ejecta with dense shells.

(ii) An LLRS must form. This is naturally expected if the central engine switches off smoothly, ejecting a tail of low Lorentz factor material behind at the end of the relativistic ejecta. Variability of the Lorentz factor in this slow tail is required to produce the necessary stratification for shock interaction to occur on the observed time-scale. According to our model, more complex initial variability can naturally lead to multiple flares when the RS crosses the different dense shells.

(iii) The emission of this LLRS must dominate over the emission of the external FS, at least at early times when the X-rays are emitted. This is probably the strongest assumption but it is possible if electron acceleration and/or magnetic field amplification is inefficient in the shocked external medium, behind the ultrarelativistic FS.

Additional simulations are needed to explore a larger range of parameters, both for the ejecta and the external medium (such as wind-like configurations). This would allow us to confirm that a large range of arrival times can be obtained for flares, as is found using the ballistic model. However, the set of simulations presented

in this paper already illustrates the capacity of the LLRS model to explain the observed diversity and variability of GRB early afterglows.

## ACKNOWLEDGEMENTS

The authors thank Zakaria Meliani for helpful advice on the set-up of the simulations and Robert Mochkovitch for his insight on this paper. The authors thank the anonymous referee for comments that greatly improved the discussion of the X-ray properties of the ejecta. The authors thank the French Programme National Hautes Energies (PNHE) and the Centre National d'Etudes Spatiales (CNES) for financial support. Support for AL was provided by an Alfred P. Sloan Research Fellowship, NASA ATP Grant NNX14AH35G, and NSF Collaborative Research Grant 1411920 and CAREER grant 1455342. Numerical simulations were run on supercomputer Pleiades from the NASA Supercomputing Division and supercomputer Stampede from the Texas Advanced Supercomputing Center.

## REFERENCES

- Abdo A. A. et al., 2011, ApJ, 734, L27
- Aloy M. A., Ibáñez J. M., Martí J. M., Gómez J.-L., Müller E., 1999, ApJ, 523, L125
- Aloy M. A., Müller E., Ibáñez J. M., Martí J. M., MacFadyen A., 2000, ApJ, 531, L119
- Beloborodov A. M., 2005, ApJ, 627, 346
- Beloborodov A. M., Daigne F., Mochkovitch R., Uhm Z. L., 2011, MNRAS, 410, 2422
- Beniamini P., Kumar P., 2016, MNRAS, 457, L108
- Berger E., 2014, ARA&A, 52, 43
- Blandford R. D., McKee C. F., 1976, Phys. Fluids, 19, 1130
- Burrows D. N. et al., 2005a, Space Sci. Rev., 120, 165
- Burrows D. N. et al., 2005b, Science, 309, 1833
- Chincarini G. et al., 2007, ApJ, 671, 1903
- Chincarini G. et al., 2010, MNRAS, 406, 2113
- Corrales L. R., Haiman Z., MacFadyen A., 2010, MNRAS, 404, 947
- Dai Z. G., Lu T., 2002, ApJ, 565, L87
- Daigne F., Mochkovitch R., 1998, MNRAS, 296, 275
- Daigne F., Mochkovitch R., 2000, A&A, 358, 1157
- De Colle F., Granot J., López-Cámara D., Ramirez-Ruiz E., 2012, ApJ, 746, 122
- Duffell P. C., MacFadyen A. I., 2013, ApJ, 775, 87
- Duffell P. C., MacFadyen A. I., 2015, ApJ, 806, 205
- Fan Y. Z., Wei D. M., 2005, MNRAS, 364, L42
- Gat I., van Eerten H., MacFadyen A., 2013, ApJ, 773, 2
- Gehrels N., Mészáros P., 2012, Science, 337, 932
- Gehrels N. et al., 2004, ApJ, 611, 1005
- Genet F., Daigne F., Mochkovitch R., 2007, MNRAS, 381, 732
- Giannios D., 2006, A&A, 455, L5
- Giannios D., Mimica P., Aloy M. A., 2008, A&A, 478, 747
- Gomboc A., 2012, Contemporary Phys., 53, 339
- Granot J., Piran T., Bromberg O., Racusin J. L., Daigne F., 2015, Space Sci. Rev., 191, 471
- Hascoët R., Uhm Z. L., Mochkovitch R., Daigne F., 2011, A&A, 534, A104
- Hascoët R., Daigne F., Mochkovitch R., 2012, A&A, 541, A88
- Hascoët R., Daigne F., Mochkovitch R., 2014, MNRAS, 442, 20
- Hascoët R., Beloborodov A. M., Daigne F., Mochkovitch R., 2017, MNRAS, 472, L94
- Kobayashi S., Piran T., Sari R., 1999, ApJ, 513, 669
- Kumar P., Piran T., 2000, ApJ, 532, 286
- Kumar P., Zhang B., 2015, Phys. Rep., 561, 1
- Lamberts A., Fromang S., Dubus G., Teyssier R., 2013, A&A, 560, A79
- Lyutikov M., Camilo Jaramillo J., 2017, ApJ, 835, 206
- MacFadyen A. I., Woosley S. E., 1999, ApJ, 524, 262

- Margutti R., Bernardini G., Barniol Duran R., Guidorzi C., Shen R. F., Chincarini G., 2011, MNRAS, 410, 1064
- Meliani Z., Keppens R., Casse F., Giannios D., 2007, MNRAS, 376, 1189
- Mészáros P., Rees M. J., 1997, ApJ, 476, 232
- Mimica P., Giannios D., Aloy M. A., 2009, A&A, 494, 879
- Mimica P., Giannios D., Aloy M. A., 2010, MNRAS, 407, 2501
- Mizuno Y., Zhang B., Giacomazzo B., Nishikawa K.-I., Hardee P. E., Nagataki S., Hartmann D. H., 2009, ApJ, 690, L47
- Narayan R., Kumar P., Tchekhovskoy A., 2011, MNRAS, 416, 2193
- Noosek J. A. et al., 2006, ApJ, 642, 389
- O'Brien P. T. et al., 2006, ApJ, 647, 1213
- Piran T., 2004, Rev. Modern Phys., 76, 1143
- Ramírez-Ruiz E., García-Segura G., Salmonson J. D., Pérez-Rendón B., 2005, ApJ, 631, 435
- Rees M. J., Mészáros P., 1998, ApJ, 496, L1
- Rezzolla L., Zanotti O., 2002, Phys. Rev. Lett., 89, 114501
- Rezzolla L., Zanotti O., Pons J. A., 2003, J. Fluid Mech., 479, 199
- Rowlinson A., O'Brien P. T., Metzger B. D., Tanvir N. R., Levan A. J., 2013, MNRAS, 430, 1061
- Sari R., Mészáros P., 2000, ApJ, 535, L33
- Sari R., Piran T., 1995, ApJ, 455, L143
- Sari R., Piran T., Narayan R., 1998, ApJ, 497, L17
- Sedov L., 1969, Similarity and Dimensional Methods in Mechanics. Academic Press, New York
- Tagliaferri G. et al., 2005, Nature, 436, 985
- Teyssier R., 2002, A&A, 385, 337
- Troja E. et al., 2007, ApJ, 665, 599
- Troja E., Piro L., Vasileiou V., Omodei N., Burgess J. M., Cutini S., Connaughton V., McEnery J. E., 2015, ApJ, 803, 10
- Uhm Z. L., Beloborodov A. M., 2007, ApJ, 665, L93
- Uhm Z. L., Zhang B., Hascoët R., Daigne F., Mochkovitch R., Park I. H., 2012, ApJ, 761, 147
- van Eerten H., Zhang W., MacFadyen A., 2010, ApJ, 722, 235
- Woods E., Loeb A., 1999, ApJ, 523, 187
- Woosley S. E., 1993, ApJ, 405, 273
- Zanotti O., Rezzolla L., Del Zanna L., Palenzuela C., 2010, A&A, 523, A8
- Zhang B., Kobayashi S., 2005, ApJ, 628, 315
- Zhang B., Yan H., 2011, ApJ, 726, 90
- Zhang B., Fan Y. Z., Dyks J., Kobayashi S., Mészáros P., Burrows D. N., Noosek J. A., Gehrels N., 2006, ApJ, 642, 354
- Zhang B., Lü H.-J., Liang E.-W., 2016, Space Sci. Rev., 202, 3

This paper has been typeset from a TeX/LaTeX file prepared by the author.