



Jet Dynamics in Compact Object Mergers: GW170817 Likely Had a Successful Jet

Paul C. Duffell , Eliot Quataert , Daniel Kasen, and Hannah Klion

Astronomy Department and Theoretical Astrophysics Center, University of California, Berkeley, CA, USA; duffell@berkeley.edu

Received 2018 June 27; revised 2018 August 14; accepted 2018 September 7; published 2018 October 5

Abstract

We use relativistic hydrodynamic numerical calculations to study the interaction between a jet and a homologous outflow produced dynamically during binary neutron star mergers. We quantify how the thermal energy imparted by the jet and the ability of the jet to escape the ejecta depend on the parameters of the jet engine and the ejecta. Under our assumptions, a collimated jet initiated at early times compared to the engine duration, we show that successful breakout of the forward cocoon shock necessitates a jet that successfully escapes the ejecta. This is because the ejecta is expanding, and the forward shock from a failed jet stalls before it reaches the edge of the ejecta. This conclusion can be circumvented only for very energetic wide angle jets, with parameters that are uncomfortable given short-duration GRB observations. For successful jets, we find two regimes of jet breakout from the ejecta: early breakout on timescales shorter than the engine duration, and late breakout well after the engine shuts off. A late breakout can explain the observed delay between gravitational waves and gamma rays in GW170817. We show that for the entire parameter space of jet parameters surveyed here (covering energies $\sim 10^{48}\text{--}10^{51}$ erg and opening angles $\theta_j \sim 0.07\text{--}0.4$) the thermal energy deposited by the jet is less than that produced by r -process heating on second timescales by at least an order of magnitude. Shock heating is thus energetically subdominant in setting the luminosity of thermally powered transients coincident with neutron star mergers (kilonovae).

Key words: gamma-ray burst: general – gravitational waves – hydrodynamics – ISM: jets and outflows – relativistic processes – shock waves

1. Introduction

The first detection of merging neutron stars by gravitational waves (GW170817; Abbott et al. 2017c) was accompanied by a wealth of electromagnetic counterparts. Notable was the first unambiguous detection of kilonova emission, which demonstrated that a significant amount of mass ($\gtrsim 0.04 M_\odot$) had been ejected with a velocity of $\sim 0.2c$ and kinetic energy of $\sim 10^{51}$ erg (e.g., Abbott et al. 2017a; Coulter et al. 2017; Kasliwal et al. 2017; Shappee et al. 2017; Soares-Santos et al. 2017; Valenti et al. 2017).

Approximately 1.7 s after the merger, a low-energy short-duration gamma-ray burst (GRB) was also detected (Abbott et al. 2017b). Currently there are two leading theories concerning the origin of the gamma rays: either they are produced by shock breakout from some injected energy from a central engine (Bromberg et al. 2018; Gottlieb et al. 2018), or they are produced by less relativistic material outside the core of a traditional short-GRB jet observed off-axis (a “structured jet”; Lazzati et al. 2017, 2018; Kathirgamaraju et al. 2018). Either of these models requires relativistic material to escape through the $v \sim 0.2c$ ejecta, at least some fraction of which is relatively spherically distributed (Hotokezaka et al. 2013).

The shock breakout scenario is sometimes envisioned as being associated with a failed GRB jet (e.g., Gottlieb et al. 2018). In this scenario, despite the failure of the jet to escape the surrounding ejecta, because significant energy was injected into the flow in the form of a cocoon, a shock continues to propagate forward, eventually breaking out of the ejecta, accelerating to relativistic velocities and releasing gamma rays.

The thermal energy imparted to the cocoon may also be important for the quasi-thermal kilonova emission in neutron star mergers. For GW170817 in particular, there are multiple

observed components to the thermal transient (Shappee et al. 2017): a blue component with faster velocities ($\sim 0.2c$), a red component with slower velocities ($\sim 0.1c$), and possibly an intermediate component (see Villar et al. 2017). It has been suggested that the blue component (particularly at early times) may be partially or fully powered by thermal energy in the cocoon (Kasliwal et al. 2017; Piro & Kollmeier 2018), rather than by the radioactive decay of neutron-rich elements, which is believed to power the majority of the thermal kilonova emission (Metzger et al. 2010b; Barnes & Kasen 2013).

The jet–ejecta interaction in neutron star mergers is somewhat analogous to a similar scenario in long-duration GRBs; there, the GRB jet needs to punch its way through a collapsing star (MacFadyen & Woosley 1999; Lazzati et al. 2012; Barnes et al. 2018). However, the scenario is not equivalent. The dynamical ejecta from a neutron star merger is a homologously expanding outflow, as opposed to the collapsar, which has a fixed size. This difference qualitatively affects the dynamics of the interaction. In this work, we perform analytical and numerical calculations of the jet–ejecta interaction, to develop new scalings for the conditions under which jets can break out of the ejecta, the resulting breakout time of the jet, and the thermalized jet energy as a function of time.

We find several important results. First, assuming a narrow engine whose energy is subdominant relative to the total kinetic energy in the dynamical ejecta, and assuming that the jet is initiated at early times compared with the breakout time, shock breakout of the cocoon only occurs if the jet is also successful. This is because the thermal energy from a failed jet is absorbed into the dynamical ejecta; because the ejecta is expanding, a low-energy shock never makes it to the edge of the outflow. Second, for a fixed total energy injection, the condition for success is independent of the duration of the engine; this is a consequence of scale invariance and of our assumption that the

delay between the merger and the engine activation is negligible. Therefore, success or failure of the jet only depends on the total energy injected and the solid angle within which it is injected. Third, a consequence of this is that, for very energetic jets, the breakout time is inversely proportional to the jet power. Fourth, we find two distinct regimes, “early breakout” and “late breakout.” “Late breakout” jets can break out *after* the engine turns off (a similar result appears to have been found for the long GRB case; see Lazzati et al. 2012). Fifth, the amount of energy that is thermalized can be estimated straightforwardly from the above considerations, leading to a robust prediction for the thermal energy in the cocoon at late times. For typical short-GRB parameters, this thermal energy is very low.

Analytical scalings for this problem are derived in Section 2. The numerical setup is briefly described in Section 3. Results of numerical calculations and comparison with the analytical scalings are presented in Section 4. Finally, these results are discussed in the context of short GRBs and GW170817 observations in Section 5.

2. Analytical Scalings

All scalings in this section will consider pure hydrodynamics. When the word “breakout” is used in this study, it refers to a shock rapidly accelerating at the outer layers of the ejecta. This could refer to the breakout of the head of the jet, or the breakout of a more spherical shock wave associated with the cocoon. Either way, it is defined purely in terms of hydrodynamics (the radiation itself is not considered in this study).

2.1. Success versus Failure

There are multiple components to the ejecta in binary neutron star mergers, including tidal tails (e.g., Rosswog et al. 1999), quasi-spherical ejecta produced when the neutron stars first collide and shock (e.g., Hotokezaka et al. 2013), and accretion disk winds (e.g., Metzger et al. 2008). The jet is most likely to directly interact with the quasi-spherical dynamical ejecta (the case of interaction with a wind has been studied previously by Murguia-Berthier et al. 2014, 2017). This can be modeled as a homologously expanding medium, which is qualitatively distinct from the stationary medium typically considered in the context of active galactic nuclei or long GRB jets. One notable distinction is that the initial condition of a homologous outflow has no physical length scale (so long as one purely considers hydrodynamics), whereas a long GRB progenitor has a fixed radius. The GRB engine has only one timescale, the duration of the engine (assuming that there is no significant delay between the ejection of the outflow and injection of the jet). Therefore, any physical length or timescale in the solution is proportional to the engine duration; a calculation with the same energy injected but double the engine duration would yield an identical solution, but with all physical lengths and timescales multiplied by 2.

The jet–ejecta interaction in neutron star mergers is thus less like the jet–collapsar interaction in long GRBs and more like a pulsar–ejecta interaction (Chevalier & Fransson 1992; Kasen et al. 2016), in which the only relevant dynamical timescale is the pulsar spin-down time. The main distinction between this problem and the pulsar–ejecta interaction is that the energy is now injected within a small opening angle.

Given this, the success or failure of the GRB jet cannot depend on the engine duration explicitly, but only on the total injected energy, and presumably the opening angle within which the energy is injected. We write the ability of the jet to break out of the dynamical ejecta in terms of E_{crit} , the minimum energy required for successful breakout:

$$E_j > E_{\text{crit}} \text{ (successful breakout).} \quad (1)$$

If the total energy injected is below E_{crit} , then all of the energy is adiabatically absorbed into the kinetic energy of the dynamical ejecta. It is reasonable to expect that E_{crit} must be some fraction of the kinetic energy of the dynamical outflow, because it is the only other relevant energy scale in this problem (ignoring $M_0 c^2$; see below). This assumption is in agreement with the scalings calculated by Margalit et al. (2018). It is also consistent with analytical and numerical calculations that assume a spherical injection (Chevalier & Fransson 1992; Kasen et al. 2016). Exactly what this fraction is should depend on θ_j , the opening angle of the injected jet (actually the half-opening angle as defined in this study). A likely scaling is

$$E_{\text{crit}} \propto \theta_j^2 E_{\text{ej}}, \quad (2)$$

based on the amount of energy subtended by the opening angle θ_j . Prior to performing numerical calculations, however, it is not obvious whether this scaling is correct, as the opening angle of the cocoon is not necessarily proportional to the opening angle of the injected jet. Moreover, there is a second energy scale in the problem, $M_{\text{ej}} c^2$. For highly collimated jets the speed of the head of a jet propagating through a medium is set by the momentum flux of the jet, while the speed of the cocoon spreading laterally is set by the energy content of the cocoon (Begelman & Cioffi 1989). In this case one can show that the jet propagation will depend on both E_{ej} and $M_{\text{ej}} c^2$ independently. For our problem, however, the mean ejecta velocities from neutron star mergers live in a narrow range of $\sim 0.1c$ – $0.3c$, set by the escape velocity from a neutron star. Thus, E_{ej} and $M_{\text{ej}} c^2$ are roughly proportional to each other. For simplicity, we will thus present our results in terms of E_{ej} , largely keeping the mean ejecta velocity $\sim (E_{\text{ej}}/M_{\text{ej}})^{1/2}$ fixed. Additional tests using different mean ejecta velocities of $\sim 0.1c$ – $0.4c$ show that our conclusions are all essentially unchanged within this range (Section 4).

2.2. Breakout Time

The breakout time can be estimated as in the pulsar problem (Chevalier & Soker 1989), though with geometric corrections due to the jet opening angle. We will proceed in a little bit less detail, to calculate the scalings but without calculating the dimensionless coefficients in front, which would depend on the detailed structure of the dynamical ejecta. Our numerical results will provide a calibration of these dimensionless coefficients.

The engine injects a jet with opening angle θ_j , power L , and duration T , with total injected energy $E_j = LT$. In principle, there is an additional input parameter, the delay between the ejection of the cloud and initiation of the engine. All of our numerical calculations work in the limit where this timescale is very short compared to T . In this limit, there is only one input timescale in the setup of this problem: the engine duration, T .

Under these specific assumptions, every length and timescale in the solution will be proportional to T . Specifically, one should be able to write the breakout time as

$$t_{\text{bo}} = Tf(E_j), \quad (3)$$

for some dimensionless function f of the injected energy.

One could envision a thought experiment where the duration is essentially infinite, so that the engine is only specified by a constant power. In this case, the total input energy is effectively infinite, so there is always a successful breakout. After breakout, however, one is free to shut off the engine at any time, and therefore, for a given power, the minimal engine duration T necessary for breakout is given by the relationship $LT = E_{\text{crit}}$. One can turn this formula around to estimate the breakout time: $t_{\text{bo}} \sim E_{\text{crit}}/L$ or

$$t_{\text{bo}} \sim TE_{\text{crit}}/E_j. \quad (4)$$

This breakout time is consistent with the value estimated by Margalit et al. (2018). As we shall show, this expression works for jets that break out before the engine has begun to turn off, so that the jet power can be considered a constant.

In this work, we also find a “late breakout” regime in which the jet breaks out after the engine is turned off, $t_{\text{bo}} > T$ (something similar has been seen previously in long GRB scenarios; see Lazzati et al. 2012). In this case, essentially all of E_j has been delivered to the cocoon prior to breakout, which continues to collimate the jet. The jetted outflow has a characteristic terminal velocity v_{sh} that takes time to overtake the fastest-moving ejecta. Since the breakout can occur well after the engine turns off, this regime is analogous to a collimated explosion in a homologous outflow. We will calibrate the breakout time for this regime empirically in Section 4.

2.3. Thermal Energy in the Cocoon

Some fraction of the engine power will be converted into thermal energy of the cocoon as the jet propagates through the ejecta. This, in addition to thermalized energy from radioactive decay, can potentially help power the thermal electromagnetic counterpart to neutron star mergers. The engine likely outputs a primarily kinetic (or magnetic) flow that thermalizes in a shock when colliding with the ejecta. Imagine that some fraction η of the engine power is instantly thermalized in this shock. The thermal energy as a function of time is then given by the equation

$$\frac{dE_{\text{Th}}}{dt} = \eta L(t) - E_{\text{Th}}/t, \quad (5)$$

where the second term is adiabatic conversion to kinetic energy in the homologous phase (a radiation-dominated equation of state is assumed). Note that before the jet and cocoon themselves reach homology, the actual work done on the ejecta is not the homologous expression $-E_{\text{Th}}/t$, but rather $p_c dV/dt$ (where p_c and dV/dt are the pressure and rate of change of volume of the cocoon, respectively); this in turn depends on how the jet head and cocoon propagate through the ejecta. In Equation (5) we are absorbing this physics into the factor η , which encapsulates the thermalization of the injected jet energy in the cocoon and the expansion of a cocoon-driven shock into the ejecta. We will show below that analytical considerations

result in a value of $\eta \approx 0.8$, close to the value we find in our numerical calculations.

Equation (5) can be straightforwardly solved:

$$E_{\text{Th}}(t) = \frac{\eta}{t} \int_0^t t' L(t') dt'. \quad (6)$$

If the jet breaks out, however, the forward shock detaches and the engine no longer efficiently thermalizes its energy. To first approximation, the first term in Equation (5) then vanishes, and the thermal energy drops as t^{-1} :

$$E_{\text{Th}}(t) = E_{\text{Th}}(t_{\text{bo}}) \frac{t_{\text{bo}}}{t} \quad (t > t_{\text{bo}}). \quad (7)$$

In this study, we consider an engine model with a smooth cutoff at late times:

$$L(t) = \frac{E_j/T}{(1 + t/T)^2}, \quad (8)$$

so that Equation (6) yields

$$E_{\text{Th}}(t) = \begin{cases} \eta E_j \frac{T}{t} \left[\ln(1 + t/T) - \frac{t}{T+t} \right] & t < t_{\text{bo}} \\ \eta E_j \frac{T}{t} \left[\ln(1 + t_{\text{bo}}/T) - \frac{t_{\text{bo}}}{T+t_{\text{bo}}} \right] & t > t_{\text{bo}} \end{cases}. \quad (9)$$

If the shock breaks out at sufficiently early times, then one can ignore the details of how quickly the engine shuts off. However, there is an additional complication, as the engine can still deposit thermal energy (at a diminished efficiency) onto the sides of the tunnel. In this case, the first term in Equation (5) does not strictly vanish, but is instead replaced with a diminished efficiency, η_2 . This would lead to the solution

$$E_{\text{Th}}(t) = \begin{cases} \eta E_0(t) & t < t_{\text{bo}} \\ \eta E_0(t_{\text{bo}}) \frac{t_{\text{bo}}}{t} + \eta_2 E_0(t) & t > t_{\text{bo}} \end{cases}, \quad (10)$$

where $E_0(t) = E_j \frac{T}{t} \left[\ln(1 + t/T) - \frac{t}{T+t} \right]$. The details of this formula are of course sensitive to our choice of $L(t)$, i.e., precisely how the power decays with time. More generally, a good estimate for the late-time thermal energy is

$$E_{\text{Th}}(t) \approx \begin{cases} \eta_2 E_j \frac{T}{t} & \text{early breakout} \\ \eta E_j \frac{T}{t} & \text{late breakout} \end{cases}. \quad (11)$$

In this study, we find that the transition between early and late breakout occurs around $E_j \sim 30E_{\text{crit}}$ (where E_{crit} is the transition between failed and successful jets) and that $\eta = 0.8$, $\eta_2 \approx 0.02\eta$ for our models.

Chevalier & Fransson (1992) calculate the fraction of thermalized energy for a spherically injected engine in a homologous outflow:

$$E_{\text{Th}} = \frac{5-k}{11-2k} Lt = \frac{1}{2} \eta Lt \quad (12)$$

for a pulsar wind injected into a density profile with $\rho \propto r^{-k}$ (their Equation (2.8)). This result is weakly dependent on k and leads to a thermalization efficiency of $\eta \approx 0.8$ for most of the propagation of the jet (using $k = 2.5$). Our numerical results are consistent with this value (see the analytical curves in Figure 4, which assume $\eta = 0.8$).

3. Numerical Setup

We numerically study the interaction of a jet with a homologous outflow using the equations of 2D axisymmetric relativistic hydrodynamics:

$$\partial_\mu(\rho u^\mu) = S_D \quad (13)$$

$$\partial_\mu(\rho h u^\mu u^\nu + Pg^{\mu\nu}) = S^\nu, \quad (14)$$

where ρ is proper density, $\rho h = \rho + \epsilon + P$ is enthalpy density, P is pressure, ϵ is the internal energy density, u^μ is the four-velocity, and $g^{\mu\nu}$ is the Minkowski metric (the equations are expressed in units such that $c = 1$). The source terms S_D and S^ν will be used to model the injection of mass, momentum, and energy by the central engine on small scales. The equation of state is assumed to be radiation dominated: $\epsilon = 3P$.

The equations are integrated using the moving-mesh hydrodynamics code, JET (Duffell & MacFadyen 2011, 2013). The JET code's moving mesh makes it effectively Lagrangian, which allows for accurate resolution of supersonic and relativistic flows. The inner and outer boundaries are also moved during the flow's evolution, so that the dynamic range covered at any given time is kept modest. The spatial resolution is given by $\Delta\theta \sim \Delta r/r \sim 0.01$. This value is not exact, as zones can expand and contract, but zones are refined or de-refined so as to maintain aspect ratios of order unity. Additionally, zones are more concentrated near the symmetry axis, so that resolution there is about twice as high ($\Delta\theta \sim \Delta r/r \sim 0.005$), and in practice zones near shock fronts are compressed by a factor of approximately five so that $\Delta r/r \sim 0.001$ in the most critical regions.

Initial conditions are modeled after the numerical output of a neutron star merger simulation (Hotokezaka et al. 2013; Nagakura et al. 2014). It is an outflow described by the following:

$$\rho(r, \theta, 0) = \begin{cases} f(\tilde{r}) + \rho_{\text{atm}}(r) & \tilde{r} < R_0 \\ \rho_{\text{atm}}(r) & \tilde{r} > R_0 \end{cases} \quad (15)$$

$$v_r(r, \theta, 0) = \begin{cases} r/t_0 & \tilde{r} < R_0 \\ 0 & \tilde{r} > R_0 \end{cases} \quad (16)$$

$$f(r) = \rho_c \frac{(1 - r/R_0)^{3/4}}{1 + (r/R_1)^{2.5}} \quad (17)$$

$$\rho_{\text{atm}}(r) = \rho_{\text{trans}} e^{-r/R_0} + \rho_{\text{ISM}} \quad (18)$$

$$\tilde{r}^2 = r^2(w^{4/3} \cos^2 \theta + w^{-2/3} \sin^2 \theta), \quad (19)$$

where $\rho_c = 97M_0/R_0^3$, $R_1 = 0.06R_0$, $\rho_{\text{trans}} = 10^{-8}M_0/R_0^3$, $\rho_{\text{ISM}} = 10^{-20}M_0/R_0^3$, and $R_0 = v_{\text{max}}t_0$, with $v_{\text{max}} = 0.3c$. The engine is initiated at a very early time $t_0 = 10^{-3}T$, where T is the engine duration. The ejecta model above corresponds to a mass-weighted velocity of the ejecta of $0.1c$ and kinetic energy-weighted velocity of the ejecta of $0.18c$.

In interpreting our results, the mass M_0 of the cloud is assumed to have a fiducial value of $0.07M_\odot$, which gives a kinetic energy of 10^{51} erg. The dynamics of the interaction are, however, scale invariant, so one could equally well assume, say, $M_0 = 0.01M_\odot$ and $E_{\text{ej}} = 1.4 \times 10^{50}$ erg. In cgs units, the external density $\rho_{\text{ISM}} \approx 5 \times 10^{-17}$ g cm $^{-3}$ is much larger than typical interstellar medium densities of ~ 1 proton per cm 3 , but this is irrelevant for the present study, as we are only concerned with the interaction between the jet and the ejecta, so this

density should be considered an artificial atmosphere, with low enough density not to affect the dynamics. The parameter $w = 1.3$ is the aspect ratio of the oblate cloud. The choice $w = 1.3$ has an impact on whether or not the jet can break out of the ejecta (Duffell et al. 2015), and therefore this affects any dimensionless coefficients reported in this study, but we expect that the overall scalings are independent of the details of the ejecta model.

The engine power is prescribed by Equation (8), $L(t) = (E_j/T)/(1+t/T)^2$, which is a constant power for $t \ll T$ but decays as $E_j T/t^2$ at late times. The power-law tail is nonstandard in most numerical calculations of GRB jets; typically jet power is imposed for some given duration with an instantaneous shutoff. Nature is more likely to be well modeled as a power-law decay than to have a sharp cutoff. The exponent of -2 is largely arbitrary, but it could mimic late-time power input by a central engine, e.g., magnetar spin-down or fallback accretion.

4. Results

The two dimensionless variables that are varied in this study are E_j/E_{ej} (the ratio of injected energy to the kinetic energy of the outflow) and θ_j , the opening angle of the injection (which may be different from the opening angle of the final jet). We have also performed additional runs varying v_{max} in the ejecta model (Section 3). This effectively varies the energy scale E_{ej} while keeping the mass fixed and confirms that the scaling with E_{ej} at fixed $M_{\text{ej}}c^2$ (or vice versa) is weak over the range of mean ejecta velocities of $\sim 0.1c$ – $0.4c$ expected in neutron star mergers.

Figure 1 shows the four classes of solutions that we find depending on the values of the key dimensionless numbers E_j/E_{ej} and θ_j . These are, from left to right, successful jet breakout from the homologous ejecta on a timescale $\lesssim T$, the engine duration, failed jets, late breakout on timescales $\gtrsim T$, and failed jets in which the forward shock and cocoon driven into the ejecta by the jet nonetheless break out. In what follows we delineate the regimes in which these different solutions occur and some of their key properties.

4.1. Success versus Failure

We argued in Section 2 that the ability of a jet to break out of the ejecta should depend on the jet energy relative to a critical energy E_{crit} , with

$$E_{\text{crit}} = \kappa E_{\text{ej}} \theta_j^2, \quad (20)$$

where κ is some dimensionless constant. Figure 2 shows success versus failure for a wide range of models with variable energy and opening angle. A successful jet is defined here as one that emerges from the outflow collimated and relativistic. A successful shock breakout is defined to occur if any shock at all makes it to the edge of the ejecta. For all except the cyan points in Figure 2, successful shock breakout necessitated a successful jet. We find the θ_j^2 scaling to be accurate for determining success of the jet, with the coefficient $\kappa = 0.05$ in Equation (20). The value of this coefficient should depend on the details of the initial density profile and the engine power as a function of time.

One important caveat of these numerical calculations is that they have been performed in 2D, and some of these jets fail

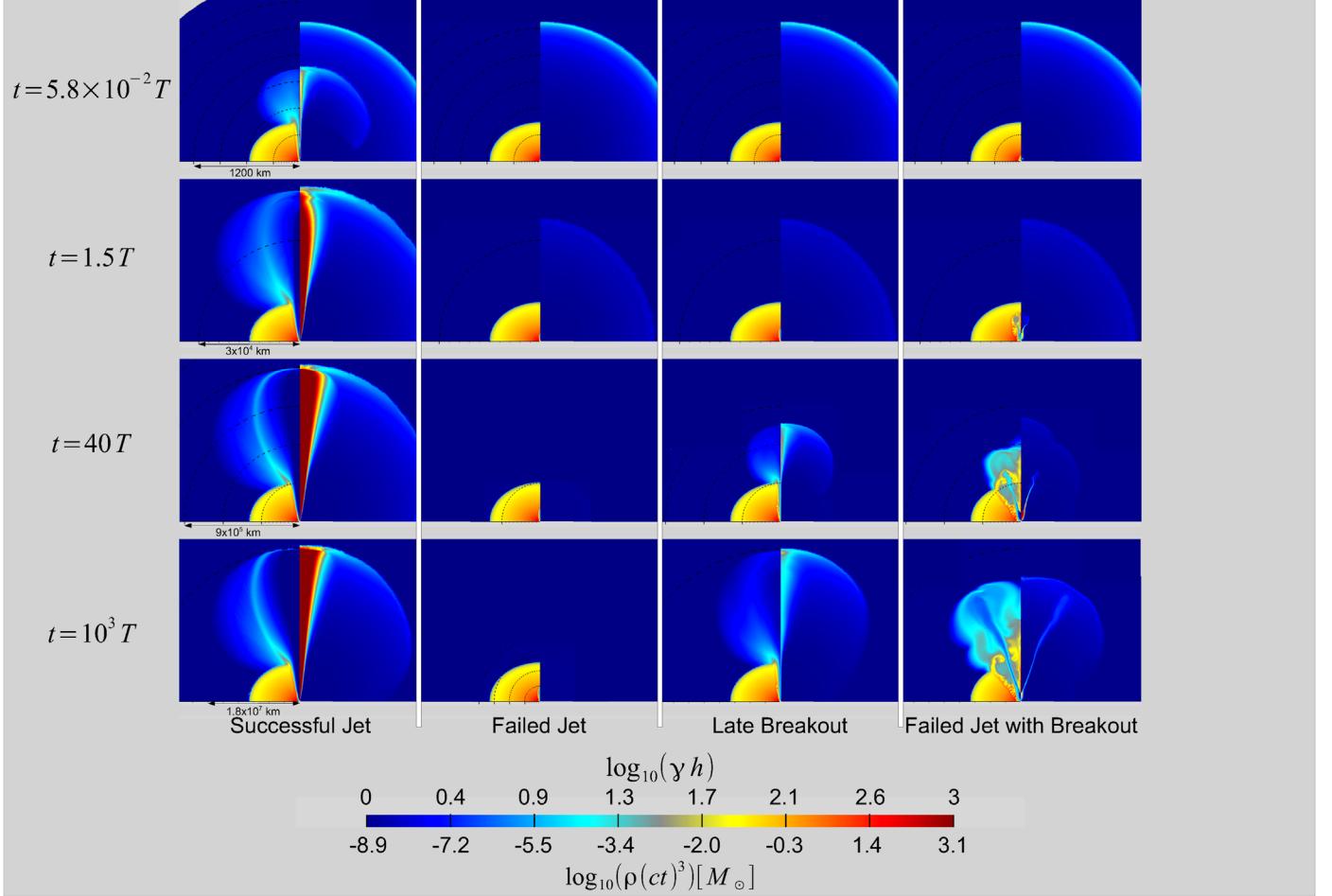


Figure 1. Various scenarios for jet propagation in a homologous ejecta. Each panel displays density on the left and γh on the right. The four columns represent different injected energies. Each row represents a different time, as indicated on the left in units of the engine duration, T . We take the ejecta energy to be 10^{51} erg to present results in physical units. The leftmost column shows an early breakout, with an opening angle of $\theta_j = 0.1$ and a large injected energy of $E_j = 10^{50}$ erg, much larger than the critical energy needed to breakout $E_{\text{crit}} \approx 5 \times 10^{47}$ erg (Equation (20)). The second column injects a much lower energy, 4×10^{47} erg (also with $\theta_j = 0.1$), below E_{crit} . This model initially drives a narrow jet, but it is too weak and no wave or shock makes it out of the ejecta. The third column shows a “late breakout,” with $\theta_j = 0.1$ and $E_j = 1.6 \times 10^{48}$ erg, just above E_{crit} . This results in a collimated and relativistic jet, with Lorentz factor ~ 30 , which breaks out an order of magnitude in time after the engine has turned off. The rightmost column has a high injected energy ($E_j = 10^{50}$ erg) and a wide injection angle ($\theta_j = 0.4$). The wide injection angle results in a “choked jet,” but the very high energy results in a successful shock/cocoon breakout, even though the choked injection effectively becomes spherical. The wide injection angle is necessary to choke such an energetic jet.

owing to “corking” (the jet is choked because a massive cork or plug is trapped ahead of the jet, which deflects the jet away from the symmetry axis; this might not occur in 3D; see Duffell & MacFadyen 2015 for a discussion of the cork and Zhang et al. 2004 for a comparison between 2D and 3D). These are labeled as such in Figure 2. Fortunately, a scaling for $E_{\text{crit}}(\theta_j)$ can be discerned despite these possibly artificially choked jets.

4.2. Breakout Time

Breakout times of successful jets are shown in Figure 3. Breakout times are consistent with the following expression:

$$t_{\text{bo}} = \begin{cases} \alpha_1 T \frac{E_{\text{crit}}}{E_j} & \text{early breakout} \\ \alpha_2 \frac{T}{\sqrt{E_j/E_{\text{crit}} - 1}} & \text{late breakout} \end{cases}, \quad (21)$$

where the “early breakout” scaling was estimated in Section 2.2, but the “late breakout” scaling is entirely empirical. For $E_j \gtrsim 30E_{\text{crit}}$, breakout times follow the “early breakout” scaling with a coefficient $\alpha_1 \simeq 3$, while for

$E_{\text{crit}} < E_j \lesssim 30E_{\text{crit}}$ there is more scatter, but breakout times reasonably track the “late breakout” scaling, with $\alpha_2 \simeq 9$.

4.3. Thermal Energy

The total thermal energy in the computational domain is shown as a function of time in Figure 4. In the top panel, this is compared with the analytical form given by Equation (11). For failed jets (green curves), the thermal energy follows the predicted formula precisely, with a thermalization efficiency $\eta = 0.8$. Successful jets with early breakout (pink curves) have a sudden transition at breakout, where the thermal energy drops and thermalization is no longer efficient because the jet vents through the channel carved in the ejecta. Jets with an early breakout exhibit a second increase in thermal energy due to the thermalization of the jet onto the sides of the tunnel. This comes with a reduced efficiency $\eta_2 \approx 0.02\eta$, though this efficiency could be higher if the jets precess or the jet opening angle varies in time owing to variability in the accretion disk driving the jet. After the engine shuts off, i.e., for $t \gtrsim T$, the

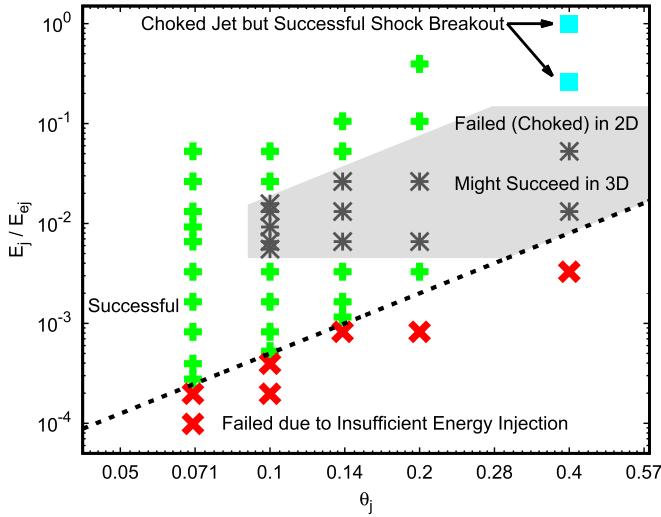


Figure 2. Diagram illustrating whether or not the jet in a given model succeeds at breaking out of the ejecta, as a function of the injection opening angle θ_j and the injected energy E_j . Green points indicate successful breakouts. Gray points are jets that were energetic but exhibited a significant cork along the symmetry axis that the jet collided with, causing failure in 2D (these jets might instead be successful in 3D; see text for details). Red crosses indicate jets that failed owing to insufficient energy injection. Cyan squares were cases that exhibited shock/cocoon breakout even with a choked jet; these require both high energy and a wide injection angle in order to choke such an energetic jet. $E_{\text{crit}}(\theta_j)$ (Equation (20)) delineating success from failure is indicated with the dashed curve.

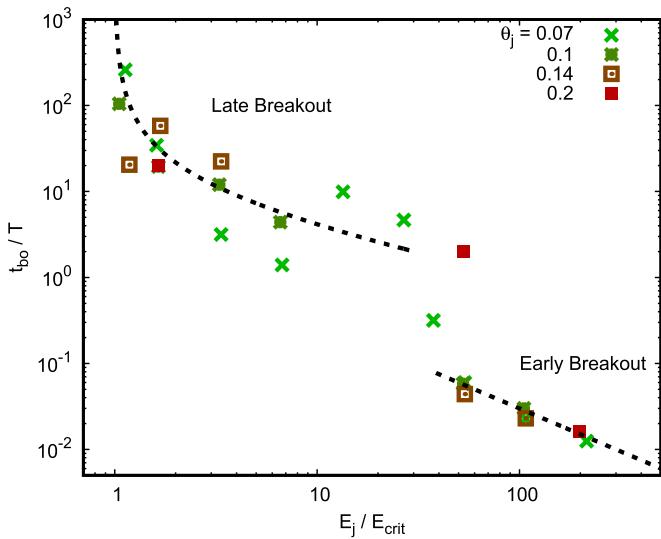


Figure 3. Jet breakout time for all successful models, compared with the prediction given by Equation (21). Jets with $E_j > 30E_{\text{crit}}$ exhibit an “early breakout,” where the engine remains on after breakout. Jets with $E_j \lesssim 30E_{\text{crit}}$ exhibit “late breakout,” breaking out well after the engine has ceased. Late breakout of the jet can produce a delay between gravitational wave and gamma-ray emission like that seen in GW170817.

thermal energy evolution transitions to a t^{-1} decline consistent with adiabatic conversion from thermal to kinetic energy.

The bottom panel of Figure 4 shows the thermal energy as a function of time in units of the ejecta kinetic energy rather than the input jet energy. This highlights that it is difficult to put a significant amount of thermal energy into the cocoon/ejecta via a jet even when the total energy of the jet approaches the initial kinetic energy of the ejecta (redder curves). This is primarily because high jet energies result in an early breakout, cutting off

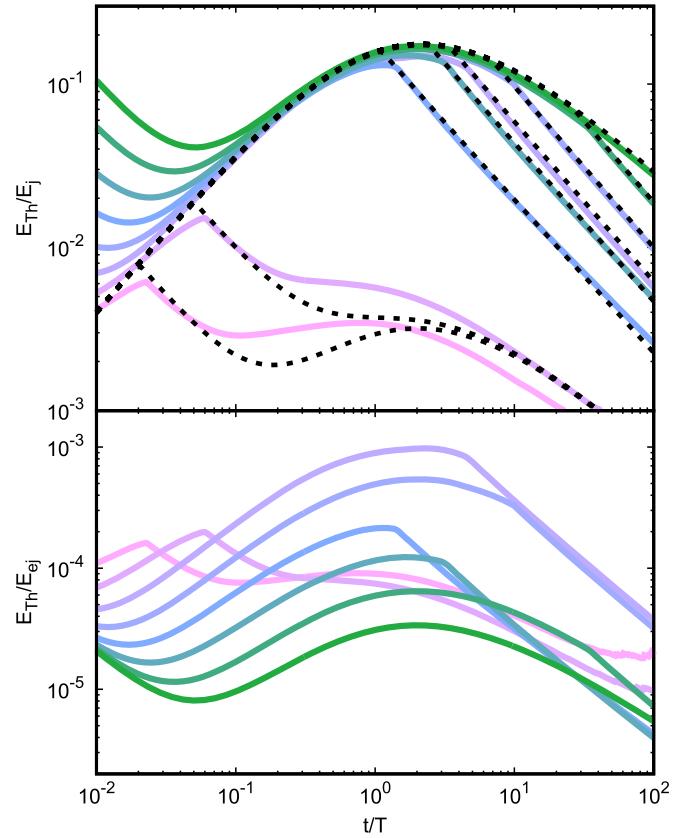


Figure 4. Total thermal energy as a function of time, for models with $\theta_j = 0.07$, and with jet energies ranging from 8×10^{47} erg (green) to 5×10^{51} erg (pink), assuming an ejecta kinetic energy of 10^{51} erg. No radioactive heating is accounted for. Failed jets precisely follow the analytic expression in Equation (10) with thermalization efficiency $\eta = 0.8$ (dashed curves). Successful jets depart from this scaling, as the forward shock breaks out of the ejecta, and the thermalization is suppressed after this point, leading to a sudden transition to a t^{-1} scaling at t_{bo} . Jets that break out early have a second, much less efficient injection of thermal energy, as a result of the engine heating the walls of the tunnel.

the thermalization and leading to most of the jet energy escaping the ejecta.

Figure 5 shows the peak thermal energy in each of our models. This peak is maximized for failed jets and late breakouts with large opening angles; for such cases, the peak thermal energy is $E_{\text{peak}} \approx 0.17E_j$, where the factor 0.17 can be determined from the peak of Equation (9). The exact value of this coefficient at the factor of few level will depend on the precise form of $L(t)$ and the ejecta structure. Figure 5 shows that only for very large opening angles and jet energies approaching $E_j \sim 10^{51}$ erg can the jet deposit an appreciable thermal energy in the ejecta. More typically, $E_{\text{peak}} \ll E_j$.

5. Discussion

The energetics of short GRBs are less constrained than those of long-duration bursts, because the opening angles of short bursts are usually more uncertain. However, current estimates from afterglow observations are that typical energies are in the range of 10^{49} – 10^{50} erg, with opening angles of $\theta_j \sim 0.05$ – 0.4 (Fong et al. 2015). An independent (but model-dependent) constraint on the opening angle is provided by the relative rate of short GRBs and neutron star mergers, which favors beaming factors of ~ 100 and thus $\theta_j \sim 0.1$.

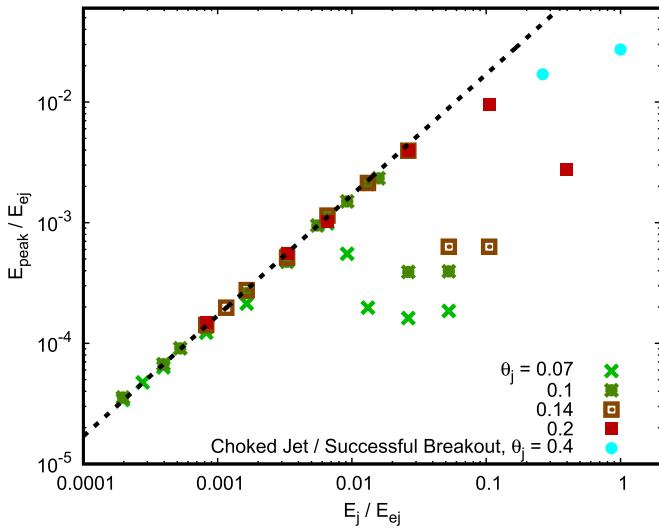


Figure 5. Peak thermal energy due to jet heating for the various models considered. Figure 4 shows that this peak thermal energy is achieved on a timescale comparable to the engine duration. Failed jets and late breakouts cluster along the curve $E_{\text{peak}} \simeq 0.17 E_j$. Early breakouts result in a reduced peak thermal energy. The most dramatic jet models had peak thermal energy of a few times $10^{-2} E_{\text{ej}} \sim 10^{49} \text{ erg}$ (taking $E_{\text{ej}} = 10^{51} \text{ erg}$). This is significantly less than the thermal energy supplied by radioactive heating of the ejecta.

Note that opening angles measured from afterglow observations utilize the “jet break” and may be different from the opening angle injected by the central engine, which is how θ_j is defined in the present study. Similarly, the duration of the burst might not be the same as the duration of the engine T , though we assume that a typical value of T should be in the typical observed duration range of 0.1–1 s. All of the calculations in this work assumed that the engine is initiated at early times, $\ll T$ and $\ll t_{\text{bo}}$. This is not guaranteed to be true, and future work should investigate the consequences of relaxing this assumption. However, this assumption is physically motivated, as an accretion disk should form on millisecond timescales. Even if a hypermassive neutron star survives for ~ 0.1 –1 s before collapsing to a black hole, it is unlikely that much mass will remain at later times to accrete onto the black hole given the short viscous times in the disk around the hypermassive neutron star (Metzger & Fernández 2014 demonstrate this explicitly in idealized axisymmetric hydrodynamic simulations). It is possible, however, that early time outflows and jets are baryon dominated because of the strong neutrino-driven wind from the neutron star (e.g., Metzger et al. 2018). In this case, a highly relativistic outflow may emerge only at later times, effectively producing a delay in the onset of the relativistic engine even though most of the accretion energy is in fact released promptly.

For jet energies and opening angles at the upper limits suggested by observations of short GRBs ($E_j \sim 10^{50} \text{ erg} \sim 0.1 E_{\text{ej}}$ and $\theta_j \sim 0.2$ –0.4), Figures 4 and 5 show that the peak thermal energy due to the jet propagating through the ejecta is $\sim 10^{-3} E_{\text{ej}} \sim 10^{-2} E_{\text{ej}} \sim 10^{48}$ – 10^{49} erg (where the latter numbers assume our fiducial $E_{\text{ej}} = 10^{51} \text{ erg}$). Moreover, this energy is reached on a timescale comparable to the engine duration $T \sim 0.1$ –1 s. For comparison, the *r*-process produces a thermal energy of $\simeq 1$ –3 MeV per nucleon on a timescale of about 1 s (Metzger et al. 2010a), with the heating then decaying as a power law at later times. For our fiducial ejecta mass of $0.07 M_{\odot}$ this corresponds to a thermal energy from the

r-process of $(1$ – $3) \times 10^{50} \text{ erg} \gtrsim 0.1 E_{\text{ej}}$ on a timescale of 1 s.¹ Figure 5 thus shows that *r*-process heating of the ejecta in neutron star mergers on second timescales dominates over jet heating in the entire parameter space of numerical models studied in this paper (usually by a large factor). In addition, the more energetic the engine, the earlier the jet breaks out, thermalizing a smaller fraction of its energy, meaning that these statements are quite robust even if one considers highly energetic engines. The *r*-process heating will thus set the overall luminosity of any associated thermally powered transient. Note that the dominance of *r*-process heating over jet heating is usually discussed in the context of day-timescale kilonova emission. Here we have shown that this is true in most cases even on *second* timescales. We do find, however, that for a small amount of mass along the pole the energy per unit mass from shock heating exceeds that produced by radioactive heating in the first few seconds. The angular dependence of this deposited energy may have observational consequences, but calculating these consequences explicitly would require full multidimensional radiation transport. The implications of this for kilonova light curves will be explored in future work.

For almost all of our parameter choices, we find that shock breakout of the cocoon only occurs if the jet itself is able to escape the ejecta. This is in contrast with some of the shock breakout scenarios presented for GRB 170817A (e.g., Gottlieb et al. 2018), which assume cocoon breakout after a choked/failed jet. The regime of shock breakout with a choked jet is only realized for large jet energies and large jet opening angles (see the right panel of Figure 1 and the cyan points in Figure 2), which are indeed the parameters used in Gottlieb et al. (2018). By contrast, for lower energies and opening angles, which appear to be somewhat more typical for short GRBs, the jet and cocoon both likely break out (Figure 2).

We find that the timescale for the jet to breakout of the ejecta scales inversely with the jet power for $E_j \gg E_{\text{crit}}$ (This is in part a consequence of our assumption of negligible delay between mass ejection and engine initiation). For energies just above E_{crit} , however, the difference between the shock velocity due to the jet and the homologous ejecta velocity is small, resulting in a late breakout on timescales that can be significantly longer than the engine duration T . In particular, reasonable engine parameters consistent with short-duration GRBs can be consistent with the 1.7 s delay between gravitational waves and gamma rays in GW170817. For example, a delay of ~ 3 – $10 T$ occurs in our late breakout regime with $E_j \sim 10 E_{\text{crit}}$ (Figure 3). Taking $\theta_j \sim 0.1$, this corresponds to $E_j \sim 10^{-2} E_{\text{ej}} \sim 10^{49} \text{ erg}$, well within reasonable short-GRB parameters. The hypothesis that the delay between gamma rays and gravitational waves in GW170817 is caused by the long breakout time can be tested by future observations. For a reasonable spread in engine parameters, it is likely that future electromagnetic (EM) counterparts to neutron star mergers will be separated into two classes: bursts with a substantial delay in gamma rays, and those without. The bursts without delay would have larger ratios of jet energy to ejecta energy. Mergers producing very little dynamical ejecta and thus fainter kilonovae would have very little delay.

¹ This conclusion holds even if E_{ej} is lower, say, $E_{\text{ej}} \sim 10^{50} \text{ erg}$ and $M_{\text{ej}} \sim 0.01 M_{\odot}$; in this case $E_j \sim 10^{50} \text{ erg} \sim E_{\text{ej}}$ leads to a maximum thermal energy of $\sim 10^{47}$ – 10^{48} erg for $\theta_j \sim 0.2$ –0.4 (Figure 5), while *r*-process heating produces an energy in the first few seconds of $\sim 3 \times 10^{49} \text{ erg}$.

The calculations presented in this paper are scale-free and can be rescaled to different assumed dynamical ejecta and jet properties. We have used a kinetic energy for the quasi-spherical dynamical ejecta of $E_{\text{ej}} = 10^{51}$ erg to anchor our discussion. This is consistent with observations of the optical-IR counterpart to GW170817 (e.g., Villar et al. 2017). Quasi-spherical dynamical ejecta with this energy can be produced in binary neutron star mergers for sufficiently small neutron star radii (e.g., Radice et al. 2016). If, however, the quasi-spherical ejecta energy is lower—which is consistent with GW170817 observations if the blue kilonova component is produced by, e.g., accretion disks or magnetar winds—a jet of a given energy and opening angle will break out more easily and will deposit less thermal energy into the cocoon as a result (Figures 3 and 5).

By using short-GRB observations to motivate engine parameters for gravitational-wave-selected events, we are assuming that at least some short GRBs are produced by binary neutron star mergers. We believe that this empirical calibration is reasonable even if short GRBs are instead due to neutron star–black hole mergers, because the mass and angular momentum content of the accretion disks are not that different. By contrast, if short GRBs are produced by an entirely different source (e.g., accretion-induced collapse of white dwarfs to neutron stars), then calibrating engine parameters of gravitational-wave-detected events to short-GRB observations may be inappropriate. It is, of course, also possible that binary neutron star mergers do produce most short GRBs, but that there is a large dispersion in short-GRB engine properties and that GW170817 just happened to be somewhat unusual relative to the bulk of the population (e.g., some short GRBs do not show any evidence for a jet break, consistent with a large opening angle of $\theta_j \sim 0.4$; see, e.g., GRB 050724A in Fong et al. 2015). Multimessenger observations in the coming decade will likely empirically settle these questions.

We thank Brian Metzger, Juna Kollmeier, and Tony Piro for stimulating conversations that helped to motivate much of this work. We thank Wen-fai Fong, Ore Gottlieb, Andrew MacFadyen, and Rosalba Perna for useful conversations. This work was supported, in part, by the Theoretical Astrophysics Center at UC Berkeley. This research was also funded by the Gordon and Betty Moore Foundation through grant GBMF5076. E.Q. was supported in part by a Simons Investigator award from the Simons Foundation. H.K. is supported by a DOE Computational Science Graduate Fellowship under grant no. DE-FG02-97ER25308. Numerical computations utilized the Savio computational cluster resource provided by the Berkeley Research Computing program at the University of California, Berkeley (supported by the UC Berkeley Chancellor, Vice Chancellor of Research, and Office of the CIO).

We would like to thank the anonymous referee for the very clear, helpful, and thorough review.

ORCID iDs

Paul C. Duffell [ID](https://orcid.org/0000-0001-7626-9629) <https://orcid.org/0000-0001-7626-9629>
Eliot Quataert [ID](https://orcid.org/0000-0001-9185-5044) <https://orcid.org/0000-0001-9185-5044>

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, *ApJL*, 848, L12
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, *ApJL*, 848, L13
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, *PhRvL*, 119, 161101
 Barnes, J., Duffell, P. C., Liu, Y., et al. 2018, *ApJ*, 860, 38
 Barnes, J., & Kasen, D. 2013, *ApJ*, 775, 18
 Begelman, M. C., & Cioffi, D. F. 1989, *ApJL*, 345, L21
 Bromberg, O., Tchekhovskoy, A., Gottlieb, O., Nakar, E., & Piran, T. 2018, *MNRAS*, 475, 2971
 Chevalier, R. A., & Fransson, C. 1992, *ApJ*, 395, 540
 Chevalier, R. A., & Soker, N. 1989, *ApJ*, 341, 867
 Coulter, D. A., Foley, R. J., Kilpatrick, C. D., et al. 2017, *Sci*, 358, 1556
 Duffell, P. C., & MacFadyen, A. I. 2011, *ApJS*, 197, 15
 Duffell, P. C., & MacFadyen, A. I. 2013, *ApJ*, 775, 87
 Duffell, P. C., & MacFadyen, A. I. 2015, *ApJ*, 806, 205
 Duffell, P. C., Quataert, E., & MacFadyen, A. I. 2015, *ApJ*, 813, 64
 Fong, W., Berger, E., Margutti, R., & Zauderer, B. A. 2015, *ApJ*, 815, 102
 Gottlieb, O., Nakar, E., Piran, T., & Hotokezaka, K. 2018, *MNRAS*, 479, 588
 Hotokezaka, K., Kiuchi, K., Kyutoku, K., et al. 2013, *PhRvD*, 87, 024001
 Kasen, D., Metzger, B. D., & Bildsten, L. 2016, *ApJ*, 821, 36
 Kasliwal, M. M., Nakar, E., Singer, L. P., et al. 2017, *Sci*, 358, 1559
 Kathirgamaraju, A., Barniol Duran, R., & Giannios, D. 2018, *MNRAS*, 473, L121
 Lazzati, D., Deich, A., Morsony, B. J., & Workman, J. C. 2017, *MNRAS*, 471, 1652
 Lazzati, D., Morsony, B. J., Blackwell, C. H., & Begelman, M. C. 2012, *ApJ*, 750, 68
 Lazzati, D., Perna, R., Morsony, B. J., et al. 2018, *PhRvL*, 120, 241103
 MacFadyen, A. I., & Woosley, S. E. 1999, *ApJ*, 524, 262
 Margalit, B., Metzger, B. D., Thompson, T. A., Nicholl, M., & Sukhbold, T. 2018, *MNRAS*, 475, 2659
 Metzger, B. D., Arcones, A., Quataert, E., & Martínez-Pinedo, G. 2010a, *MNRAS*, 402, 2771
 Metzger, B. D., & Fernández, R. 2014, *MNRAS*, 441, 3444
 Metzger, B. D., Martínez-Pinedo, G., Darbha, S., et al. 2010b, *MNRAS*, 406, 2650
 Metzger, B. D., Piro, A. L., & Quataert, E. 2008, *MNRAS*, 390, 781
 Metzger, B. D., Thompson, T. A., & Quataert, E. 2018, *ApJ*, 856, 101
 Murguia-Berthier, A., Montes, G., Ramirez-Ruiz, E., De Colle, F., & Lee, W. H. 2014, *ApJL*, 788, L8
 Murguia-Berthier, A., Ramirez-Ruiz, E., Montes, G., et al. 2017, *ApJL*, 835, L34
 Nagakura, H., Hotokezaka, K., Sekiguchi, Y., Shibata, M., & Ioka, K. 2014, *ApJL*, 784, L28
 Piro, A. L., & Kollmeier, J. A. 2018, *ApJ*, 855, 103
 Radice, D., Galeazzi, F., Lippuner, J., et al. 2016, *MNRAS*, 460, 3255
 Rosswog, S., Liebendörfer, M., Thielemann, F.-K., et al. 1999, *A&A*, 341, 499
 Shappee, B. J., Simon, J. D., Drot, M. R., et al. 2017, *Sci*, 358, 1574
 Soares-Santos, M., Holz, D. E., Annis, J., et al. 2017, *ApJL*, 848, L16
 Valenti, S., David, S. J., Yang, S., et al. 2017, *ApJL*, 848, L24
 Villar, V. A., Guillotin, J., Berger, E., et al. 2017, *ApJL*, 851, L21
 Zhang, W., Woosley, S. E., & Heger, A. 2004, *ApJ*, 608, 365