ABSTRACT

WARREN III, DONALD CAMERON. Monte Carlo Simulations of Efficient Shock Acceleration during the Afterglow Phase of Gamma-Ray Bursts. (Under the direction of Donald Ellison.)

Gamma-ray bursts (GRBs) signal the violent death of massive stars, and are the brightest explosions in the Universe since the Big Bang itself. Their afterglows are relics of the phenomenal amounts of energy released in the blast, and are visible from radio to X-ray wavelengths up to years after the event. The relativistic jet that is responsible for the GRB drives a strong shock into the circumburst medium that gives rise to the afterglow. The afterglows are thus intimately related to the GRB and its mechanism of origin, so studying the afterglow can offer a great deal of insight into the physics of these extraordinary objects.

Afterglows are studied using their photon emission, which cannot be understood without a model for how they generate cosmic rays (CRs)—subatomic particles at energies much higher than the local plasma temperature. The current leading mechanism for converting the bulk energy of shock fronts into energetic particles is diffusive shock acceleration (DSA), in which charged particles gain energy by randomly scattering back and forth across the shock many times. DSA is well-understood in the non-relativistic case—where the shock speed is much lower than the speed of light—and thoroughly-studied (but with greater difficulty) in the relativistic case. At both limits of speed, DSA can be extremely efficient, placing significant amounts of energy into CRs. This must, in turn, affect the structure of the shock, as the presence of the CRs upstream of the shock acts to modify the incoming plasma flow. The modified plasma flow affects the acceleration process responsible for generating CRs, creating a highly nonlinear feedback loop between the shock profile and the CRs accelerated by that shock. Any study of efficient DSA must include this process, but such studies must be numerical due to the inherent nonlinearity.

In this dissertation, I undertake a Monte Carlo simulation of the afterglow phase of a GRB. I follow the shock from early on, while it is still relativistic, through the trans-relativistic phase at later times. The trans-relativistic regime must be approached numerically, as it defies approximations usually made in both the non-relativistic and fully relativistic limits.

I will discuss the necessary extensions to the existing Monte Carlo code needed to examine trans-relativistic shocks and the CRs they produce, as well as the parameters and tools introduced to study the acceleration of protons, heavier ions, and much lighter electrons. I will also explain how the code takes the resultant CR spectra and produces observable photon emission.

The most significant part of this dissertation, though, is the multi-stage simulation of a GRB afterglow. I will show that the nature of DSA is radically different in the fully relativistic and trans-relativistic regimes of shock speed, which in turn requires different assumptions about the parameters critical to DSA (such as the strength of the turbulent magnetic field or the amount of energy
present in CR electrons).

Starting from physically plausible external conditions and instituting just two new parameters—a compressed magnetic field for producing synchrotron emission, and a transfer of energy from ions to electrons in the shock precursor—motivated by first-principles particle-in-cell simulations, I will show that the Monte Carlo model can adequately capture the light curve of a GRB afterglow during the main decay stage. The brightness of any particular afterglow will constrain the parameter space available to the Monte Carlo code, allowing for the future possibility of working backwards to determine the conditions at the shock based on observations.
Monte Carlo Simulations of Efficient Shock Acceleration during the Afterglow Phase of Gamma-Ray Bursts

by
Donald Cameron Warren III

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Physics
Raleigh, North Carolina
2015

APPROVED BY:

________________________  _________________________
Stephen Reynolds                James Kneller

________________________  _________________________
John Griggs                      Donald Ellison
Chair of Advisory Committee
The author was born. He went to college, graduated, and hared off to Japan for a break before graduate school. While in Japan he met the woman who would become his wife, who graciously agreed to join him back in North Carolina while he undertook the fool’s errand of earning a Ph.D. There was a marriage, a pregnancy, and a child, all of which the author recommends against (at least while still in graduate school) to any future graduate students reading this work. He regrets none of those decisions, but completing the dissertation you are now reading was rather significantly delayed by them.

He is currently, as of this writing, writing a biography that is required by the NCSU dissertation format.
ACKNOWLEDGEMENTS

There are many people who deserve thanks for their role in the production of this work. In no particular order, I would like to thank all of the following people.

- My wife Celine and daughter Aurora, for their (mostly silent) patience while I was generally emotionally, sometimes physically, and often mentally unavailable writing this dissertation.

- My mother, who (like all good Jewish mothers) wanted her child to grow up and become a doctor. Done.

- My father, who I’m mostly certain is just thankful that I grew up.

- My advisor, Don Ellison, for his unflinching support in the face of a constant barrage of questions while I became familiar with this material, as well as for doing more work than anyone besides myself (which point may be arguable) in producing the manuscript you are currently reading.

- The gaming crew, for keeping me sane all these years.

Any omissions are unintentional, and are due to the aforementioned mental unavailability.

This work made use of data supplied by the UK Swift Science Data Centre at the University of Leicester. Numerous other researchers allowed me to reproduce figures from their work for comparison or clarification; they are credited in the captions of those figures, but thanked here.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................ vii

LIST OF FIGURES .................................................................... viii

CHAPTER 1 INTRODUCTION TO COSMIC RAYS AND SHOCK ACCELERATION ........ 1
  1.1 Cosmic Rays ................................................................. 1
    1.1.1 Discovery ............................................................ 1
    1.1.2 Spectrum at Earth ............................................... 2
  1.2 Diffusive Shock Acceleration ........................................... 5
    1.2.1 The macroscopic method ...................................... 6
    1.2.2 The microscopic method ...................................... 9
    1.2.3 Comments and Consequences ............................... 12
  1.3 DSA at Relativistic Shocks .............................................. 14
  1.4 DSA at Nonlinear Shocks ............................................... 16
  1.5 Approaches to DSA ..................................................... 18
    1.5.1 Analytic and Semi-analytic .................................. 18
    1.5.2 Particle-in-Cell .................................................. 19
    1.5.3 Monte Carlo ....................................................... 20
  1.6 Summary ................................................................. 21

CHAPTER 2 GAMMA-RAY BURSTS AND THEIR AFTERGLOWS ..................... 23
  2.1 History of GRBs .......................................................... 23
  2.2 Afterglow theory and methods ...................................... 32
    2.2.1 Hydrodynamical scaling relations .......................... 33
    2.2.2 Photon spectrum scaling relations .......................... 34
    2.2.3 Hydrodynamics of the jet ..................................... 37
    2.2.4 Consequences of a collimated jet ......................... 39
  2.3 Observations of afterglows ............................................ 43
    2.3.1 Pre-Swift era ..................................................... 43
    2.3.2 Swift era ........................................................ 44
  2.4 DSA parameters ......................................................... 46
  2.5 Summary ................................................................. 50
  2.6 Goals of this Dissertation ............................................. 51

CHAPTER 3 THE CODE .......................................................... 52
  3.1 Assumptions ............................................................. 53
  3.2 Injection ................................................................. 56
  3.3 The Particle Distribution and Fluxes ............................... 57
  3.4 Propagation .............................................................. 61
    3.4.1 Scattering ......................................................... 62
6.2.2 Synchrotron................................................................. 156
6.2.3 Inverse Compton ......................................................... 162
6.3 Processing ................................................................. 168
6.4 Emission shells ............................................................. 170
  6.4.1 Cooling between timesteps ........................................... 172
6.5 Summary ................................................................. 173

CHAPTER 7 COMPLETE PICTURE OF A SMOOTHED SHOCK ............ 174
  7.1 The Shock Structure ..................................................... 174
  7.2 Particle Spectra ........................................................... 177
    7.2.1 Shock Frame ......................................................... 177
    7.2.2 Plasma & ISM Frame ............................................... 179
  7.3 Photon Spectra ........................................................... 182
  7.4 Summary ................................................................. 186

CHAPTER 8 SIMULATED AFTERGLOW ....................................... 188
  8.1 The Blandford–McKee Solution ........................................ 188
  8.2 Hydrodynamics and Particle Spectra .................................. 191
  8.3 Photon Spectra ........................................................... 203
  8.4 Comparison against Observations ...................................... 206
  8.5 Summary ................................................................. 210

CHAPTER 9 CONCLUDING REMARKS ....................................... 211

CHAPTER 10 EXTENSIONS AND IMPROVEMENTS .......................... 215

BIBLIOGRAPHY .................................................................... 218
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4.1</td>
<td>Shock Parameters</td>
<td>84</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Model Parameters</td>
<td>126</td>
</tr>
<tr>
<td>Table 6.1</td>
<td>Functional forms used in the Kamae method, for a primary cosmic ray proton with energy $T_p$, producing a photon of energy $x = E_\gamma/(1 \text{ GeV})$.</td>
<td>150</td>
</tr>
<tr>
<td>Table 6.2</td>
<td>Parameters used in the Kamae method, for a primary cosmic ray proton with energy $T_p$.</td>
<td>151</td>
</tr>
<tr>
<td>Table 7.1</td>
<td>Key shock parameters</td>
<td>176</td>
</tr>
<tr>
<td>Table 8.1</td>
<td>Afterglow model, physical parameters</td>
<td>191</td>
</tr>
<tr>
<td>Table 8.2</td>
<td>Afterglow model, Monte Carlo parameters</td>
<td>192</td>
</tr>
<tr>
<td>Table 8.3</td>
<td>$0.3 - 10 \text{ keV}$ photon index $\Gamma$ by time step</td>
<td>208</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Figure 1.1</td>
<td>The spectrum of cosmic rays observed at Earth, from energies of 10^8 eV to more than 10^{20} eV.</td>
<td>4</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>The light curve of the first recorded gamma-ray burst, GRB 670702.</td>
<td>24</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>A small sample of GRBs observed by the BATSE instrument.</td>
<td>25</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Distribution of recorded GRB lengths.</td>
<td>26</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>A cartoon of the hydrodynamical structure of a relativistic jet.</td>
<td>38</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>The “canonical” afterglow model that has emerged from the large-population studies of the Swift era.</td>
<td>45</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>A histogram of the 29 values of ε_e presented in Santana et al. (2014).</td>
<td>48</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>A histogram of 31 values of n_0 found in the literature.</td>
<td>49</td>
</tr>
<tr>
<td>Figure 2.8</td>
<td>A histogram of the 30 values of ε_B presented in Santana et al. (2014).</td>
<td>50</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Thermal leakage as a mechanism for injection into the acceleration process.</td>
<td>55</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>An illustration of several key facets of the shock system.</td>
<td>58</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>A depiction of pitch-angle diffusion as implemented in the code.</td>
<td>64</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Shock rest-frame values of the flow speed, u(x), the momentum flux, F_{px}, and the energy flux, F_{en}, are shown for UM shocks with no DSA.</td>
<td>80</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Fraction of proton number flux with shock-frame momentum p or greater vs. p for a UM shock with γ_0 = 1.5 (upper panel) and γ_0 = 10 (lower panel).</td>
<td>82</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Omni-directional spectra measured in the shock frame from three unmodified shocks.</td>
<td>85</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Top panels show the shock structure in terms of u(x)/u_0, middle panels show the momentum flux F_{px}, and the bottom panels show the energy flux F_{en}.</td>
<td>87</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Same format as Figure 4.4, but for γ_0 = 1.5.</td>
<td>89</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Same format as Figures 4.4 and 4.5, but for γ_0 = 10.</td>
<td>90</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Omni-directional, shock-frame spectra for a shock with u_0 = 2 \times 10^4 km s^{-1}.</td>
<td>91</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Comparison of UM shock spectra (Model C; same as in Figure 4.3) with NL shock spectra.</td>
<td>92</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>Comparison of UM shock spectrum (Model D; same as in Figure 4.3) with NL shock spectrum.</td>
<td>93</td>
</tr>
<tr>
<td>Figure 4.10</td>
<td>Nonlinear particle distributions calculated downstream from the shock in the shock rest frame for various shock speeds as indicated (Models A–E in Table 4.1).</td>
<td>94</td>
</tr>
<tr>
<td>Figure 4.11</td>
<td>Acceleration efficiency in terms of the fraction of energy flux above shock frame p for the β_0 = 0.067 (A), γ_0 = 1.5 (C), and γ_0 = 30 (E) shocks shown in Figure 4.10.</td>
<td>95</td>
</tr>
<tr>
<td>Figure 4.12</td>
<td>Shock structure for LAS Model F with γ_0 = 10 and N_g = 100.</td>
<td>97</td>
</tr>
</tbody>
</table>
Figure 4.13  
Omni-directional, shock-frame spectra for a $\gamma_0 = 10$ shock with LAS, i.e., $N_g = 100$.  

Figure 4.14  
Top panel shows the shock structure for Model H, matching the Summerlin and Baring (2012) example discussed in the text. In the middle and bottom panels, the dashed (red) curve is the UM flux, and the solid (black) curve is the NL flux.  

Figure 4.15  
Dashed (red) curve is the downstream spectrum from the UM shock shown in Figure 4.14. The solid (black) curve is the NL spectrum.  

Figure 4.16  
Nonlinear structure of a $\gamma_0 = 15$ shock with a FEB at $x = -3 r_{g0}$ which conserves mass, momentum, and energy flux (Model F).  

Figure 4.17  
Top panel shows the shock structure for Model H, matching the Summerlin and Baring (2012) example discussed in the text. In the middle and bottom panels, the dashed (red) curve is the UM flux, and the solid (black) curve is the NL flux.  

Figure 4.18  
Structure for a $\gamma_0 = 1.5$ shock with a FEB at $-10^8 r_{g0}$ (Model I).  

Figure 4.19  
Comparison of UM downstream spectrum (dashed, red curve) for our Model I with $L_{FEB} = -10^8 r_{g0}$ against the NL spectrum (solid, black curve).  

Figure 5.1  
An example of a bin in the shock frame (Bin A), and its transformed shape (Bin A').  

Figure 5.2  
A visual representation of the projection of the 2-D transformed bin from Figure 5.1 into the 1-D array storing $dN/dp$.  

Figure 5.3  
Schematic, not-to-scale representation of a relativistic shock embedded in a conical jet, propagating with a Lorentz factor of $\gamma_0$ into material at rest.  

Figure 5.4  
This detail focuses on the volume between $L_{Ups}$ and $L_{Dws}$.  

Figure 5.5  
Protons (black curves) and electrons (red curves) from UM shocks with different $f_{ion}$ as indicated.  

Figure 5.6  
Proton and electron spectra (as labeled and multiplied by $p^{2.23}$) with $f_{ion} = 0.15$.  

Figure 5.7  
The top panel shows the shock structure for the unmodified case (Model D) and the nonlinear case (Model E). The middle and bottom panels show the momentum and energy fluxes, respectively, normalized to far upstream values.  

Figure 5.8  
Downstream, local plasma frame (LPF) spectra for the unmodified shock shown in Figure 5.7 (top panel, Model D) and the nonlinear shock shown in Figure 5.7 (bottom panel, Model E).  

Figure 5.9  
Photon emission for the UM (Model D) and NL (Model E) shocks shown in Figures 5.7 and 5.8.  

Figure 5.10  
Nonlinear downstream LPF spectra for Model E ($f_{ion} = 0.1$) and Model F ($f_{ion} = 0.4$).  

Figure 5.11  
Total observed energy flux for NL Models E ($f_{ion} = 0.1$) and F ($f_{ion} = 0.4$).  

Figure 5.12  
All curves are as in Figure 5.7, for a shock with $\gamma_0 = 1.5$.  

Figure 5.13  
Downstream, local plasma frame (LPF) spectra for the unmodified $\gamma_0 = 1.5$ shock (top panel, Model G) and the nonlinear $\gamma_0 = 1.5$ shock (bottom panel, Model H) shown in Figure 5.12.
Figure 5.14  Photon emission for the UM (Model G) and NL (Model H) shocks shown in Figures 5.12 and 5.13.  

Figure 6.1  Ion and electron spectra used for the calculation of photon production in this chapter.  

Figure 6.2  Pion-decay spectra for the particle population shown in Figure 6.1.  

Figure 6.3  Synchrotron spectrum generated by the plasma-frame electron distribution in Figure 6.1.  

Figure 6.4  Plots of $d^3N/(dt d\eta d\alpha)$ for an isotropic, monoenergetic photon field and three different electron energies.  

Figure 6.5  Same as in Figure 6.4, but with $\eta$ transformed into the ISM frame.  

Figure 6.6  Inverse Compton emission from the electron distribution in Figure 6.1.  

Figure 6.7  Depiction of how the code handles emission from multiple time steps.  

Figure 7.1  Profile and fluxes of the smoothed shock, all measured in the shock frame.  

Figure 7.2  Particle distributions as functions of shock-frame total momentum, multiplied by $p^{2.23}$.  

Figure 7.3  Same as Figure 7.2, but for plasma-frame total momentum.  

Figure 7.4  A comparison of the particle spectra for all three species, for a downstream zone and an upstream zone.  

Figure 7.5  ISM-frame momentum distribution functions, multiplied by $p^{2.23}$ as in Figures 7.2 and 7.3.  

Figure 7.6  Photon spectra, calculated in the frame of production (plasma for pion decay and synchrotron, ISM for inverse Compton) at several locations within the shock structure.  

Figure 7.7  Photon spectra, transformed into ISM frame and summed into emission regions.  

Figure 7.8  Summed photon spectra due to all processes, for each emission regions and in total.  

Figure 8.1  Comparison of the shock profiles at the discrete time steps modeled during the afterglow.  

Figure 8.2  Illustration of how downstream cohorts are handled at a given time step $i$.  

Figure 8.3  Density of the downstream fluid, as a function of current time.  

Figure 8.4  Particle distributions for all three species as they cool downstream from the shock.  

Figure 8.5  The fraction of energy density in various forms, as a function of both current time and time of original shock crossing.  

Figure 8.6  The plasma-frame spectra of each particle population at the time step when it was originally shocked.  

Figure 8.7  The radio to $\gamma$-ray spectra from all eight time steps.  

Figure 8.8  The ISM-frame photon spectrum, separated into emission process, for three time steps in the afterglow.
Figure 8.9  The contribution of individual particle populations to the total spectrum at time step 6. ................................. 207
Figure 8.10  The total flux in the $0.3 - 10$ keV band due to the sample afterglow, as well as from a sample of GRBs with radio, optical and X-ray coverage. ................................. 208
CHAPTER 1

INTRODUCTION TO COSMIC RAYS AND SHOCK ACCELERATION

1.1 Cosmic Rays

1.1.1 Discovery

The determination that Earth is constantly being bombarded by energetic particles coming from beyond the atmosphere was, like many other discoveries, accidental. In the late 1700s, scientists noticed that the charged metal leaves of an electroscope slowly lost that charge over time, suggesting that there was a previously-undetected cloud of charged particles interacting with the device.

Experiments with the ionization rate in air showed decay with increasing elevation above the ground, or by performing the experiment on a body of water. By the turn of the 20th century, the consensus was that the cloud of charged particles was coming from radioactive decay occurring in the soil. To extend the data further, scientists began attaching electroscopes to balloons and floating them to ever-greater heights.

The most precise measurements at the greatest heights (previous attempts suffered from unreliable instrumentation) were taken by Victor Hess in 1912. Using a hydrogen-filled balloon, Hess's
electroscope ascended to the (at the time record) height of more than 5 km above ground level. At first, Hess recorded the well-known decline in ionization rate, up to \(\sim 1\) km above the ground. Then, something unexpected: instead of continuing to decline, the ionization rate remained constant for the next 1000 m of altitude gain, and then increased above 2 km. It seemed, then, that the atmosphere was acting to screen a second source of charged particles, and that at greater altitude that screening effect diminished. Given the limited options for sources at/above the atmosphere, Hess advanced the conclusion that the ionizing radiation was coming from beyond Earth entirely—that is, from space.

Hess's results were not immediately accepted, given the history to that point of unreliable data from other researchers. In the next two years, a pair of still more precise balloon flights to still greater altitudes (peaking at more than 9 km above the ground, or roughly the altitude trans-continental plane travel uses) was undertaken by Werner Kolhörster. Kolhörster was skeptical of Hess's numbers and conclusions, and expected his flights to disprove the cosmic ray hypothesis. Instead, the 1913 and 1914 data confirmed Hess's results, and indeed extended the screening trend Hess had observed.

One of the most vocal critics of the extra-terrestrial origin of the ionizing radiation both Hess and Kolhörster had observed was Robert Millikan.\(^1\) By 1926, experiments above glaciers, on/in lakes, and under sheets of lead convinced Millikan that the terrestrial origin explanation was insufficient. In Millikan and Cameron (1926), he used the shorter phrase "cosmic ray" (as an analogue to the well-known \(\gamma\)-rays at the top of the electromagnetic spectrum) instead of the more cumbersome "rays of cosmic origin". The name stuck.

In the decades since cosmic rays (which I will henceforth refer to as CRs) were identified, a vast amount of work has been performed. One of the early discoveries was that cosmic rays were not, in fact, the high-energy extension of the electromagnetic spectrum. Once scientists were able to measure the direction from which cosmic rays arrived at the detector, anisotropies in distribution became apparent. In particular, the asymmetry between east-traveling and west-traveling cosmic rays meant that not only were CRs not photons, they were not electrically neutral (e.g., Johnson, 1939, and references therein, as well as other articles from that issue).\(^2\)

### 1.1.2 Spectrum at Earth

As instrumentation and theory improved, the numerous measurements of cosmic ray energy could be assembled into a spectrum. The smaller frequency of events at greater energies required building

---

1. Of oil-drop experiment fame.
2. Examination of older science textbooks highlights the slow diffusion of scientific progress to the general public. The author recalls seeing textbooks as a child that still placed CRs above \(\gamma\)-rays and X-rays, some decades after the work of Johnson, Anderson, and others in establishing the nature of CRs.
ever-larger detectors, culminating (at present) in the massive detector arrays such as the Pierre Auger Observatory, spanning thousands of square kilometers of observing area.

The full spectrum of CRs extends over many orders of magnitude in energy, as seen in Figure 1.1. The spectrum above $\sim 10^{10}$ eV shows several features of interest. First, there is a single power law relating flux and energy up to the “knee” at $3 \times 10^{15}$ eV. Above the knee the spectrum steepens for three decades in energy, to the “ankle” at a few times $10^{18}$ eV. Then the spectrum hardens again out to the highest energy detections at more than $10^{20}$ eV.

The lowest-energy cosmic rays are heavily affected by solar modulation—the Sun’s solar wind and termination shock\(^3\)—so any information about the origin of cosmic rays must come from the part of the spectrum above $\sim 10^{10}$ eV. A stricter limit arises once the gyration of charged particles in a magnetic field is considered. The Larmor radius of such gyration is

$$r_L = \frac{pc}{qB}, \quad (1.1)$$

for a particle of momentum $p$ and charge $q$ in a magnetic field of strength $B$. For reasonable values of the interplanetary magnetic field strength, protons below $\sim 10^{12}$ eV are likely to be significantly deflected on their journey to Earth, washing away any hint of their direction of origin.

Even above $10^{12}$ eV, however, detected anisotropies are small. The degree of anisotropy is less than 0.1% above $10^{14}$ eV (Hillas, 1984), and rises only slowly after that. It seems to be the case that the Earth is moving through a diffusing cloud of CRs that corotates with the rest of the matter in the Milky Way (Amenomori et al., 2006). The energy density of CRs above 1 GeV is roughly 1 eV cm\(^{-3}\) (Wdowczyk and Wolfendale, 1989), and the mean lifetime of these particles within the galaxy is $\sim 10^7$ yr (Longair, 2011). Given the volume of the Milky Way, the energy budget for replenishment every ten million years is enormous. It would require, in fact, roughly 10% of the energy of galactic supernovae be converted into CRs, as first noted by Ginzburg and Syrovatskii (1964). Supernova remnants have long been considered as the source of galactic CRs up to the knee at $10^{15}$ eV based on Equation 1.1 and expected values for the magnetic field strength and size of remnants (e.g., Blandford and Eichler, 1987). The detection of features in Tycho’s SNR consistent with a PeV-CR origin further strengthens this claim (Eriksen et al., 2011). The steepening of the CR spectrum above the knee would correspond, then to a different source for these extremely energetic particles (such as the galactic termination shock; again see Blandford and Eichler, 1987).

Cosmic rays above the ankle are very likely to be of extragalactic origin, for the simple reason that they would not be confined within the Milky Way if they were produced locally. Protons with an

\(^3\)And indeed, the so-called anomalous component of CRs at Earth is likely to be electrically neutral interstellar atoms ionized and accelerated at the termination shock.
1.1. COSMIC RAYS

CHAPTER 1. INTRO TO COSMIC RAYS AND DSA

Figure 1.1 The spectrum of cosmic rays observed at Earth, from energies of $10^8$ eV to more than $10^{20}$ eV. Individual experiments used to take the data are noted, as are comparisons to the maximum energy of terrestrial particle accelerators. Three additional features—the “knee” at $3 \times 10^{15}$ eV, the “ankle” at $10^{18}$ eV, and the cutoff in the spectrum around $3 \times 10^{20}$ eV, are discussed in the text. The green dashed line marks an $E^{-\sigma}$ power law with $\sigma \simeq 2.8$. Reproduced with permission from William Hanlon.
energy above the ankle have a Larmor radius of 300 pc in a magnetic field of 3 μG, which distance corresponds to the scale height of the Milky Way’s disk. Such protons are likely to escape the Milky Way due to the attenuation of the galactic magnetic field during their gyration, so detected particles above the ankle are more likely to have arrived from outside the galaxy.

The upper limit on the spectrum, at about $3 \times 10^{20}$ eV, is due to propagation losses against the Cosmic Microwave Background (CMB). Energetic particles above a certain energy can scatter inelastically off of CMB photons, creating pions and losing significant fractions of their energy. The minimum energy for this process to occur is approximately $5 \times 10^{19}$ eV, called the GZK limit for its independent discovery by Greisen (1966) and Zatsepin and Kuz’min (1966). The horizon for particles above this energy is approximately 50 Mpc, suggesting that CRs much above the GZK limit should be accelerated relatively close to Earth (relative being the key word here).

That the flux forms power laws—and furthermore, power laws of similar spectral index—suggests that a single process might be responsible for generating the particles detected. I will discuss in the next section the current leading theory for how particles are accelerated to the energies shown in Figure 1.1.

### 1.2 Diffusive Shock Acceleration

Energetic particles can gain energy by scattering off of magnetic field turbulence, as first published by Fermi (1949). In this model, CRs scatter back and forth between randomly-moving clouds in the interstellar medium. If the clouds are receding from each other, the particle loses some energy; if the clouds are approaching, the particle gains some energy. The gains from receding/approaching pairs of clouds are (respectively) $\sim \pm v/c$ for clouds with relative velocity $v$; if these configurations were equally likely the losses would balance out the gains for no long-term change in particle energy. However, collisions between approaching clouds are slightly more frequent. Energy gains per scattering are on the order of $(v/c)^2$ once this effect is considered, leading to the moniker “second-order Fermi process” or “Fermi-II process”.

The first-order Fermi (or Fermi-I) process requires a velocity gradient, as in a shock. This makes all scatterings between approaching objects (rather than the random distribution between approaching/receding clouds of the Fermi-II process). In this process particles gain energy by crossing back and forth across the shock front, using the great relative velocity of the upstream (unshocked) and downstream (shocked) material to provide the energy gains. This method was discovered roughly simultaneously by many different groups: Krymskii (1977), Axford et al. (1977),

---

4 The more prosaic picture is that of a tennis ball bouncing between two walls.
5 Which does indeed gain energy as $v/c$ per cycle, as I will show later.
Blandford and Ostriker (1978) and Bell (1978) all explored the consequences of shock-acceleration. As the process requires diffusion of energetic particles back and forth across the shock front, it is also called diffusive shock acceleration, or DSA.

In the next two sections I will outline the two basic approaches taken towards understanding DSA. I will then summarize the major limitations of the early work, which will set the stage for the remainder of this chapter.

### 1.2.1 The macroscopic method

The macroscopic method, explored in Krymskii (1977), Axford et al. (1977), and Blandford and Ostriker (1978), solves a diffusion equation for particles encountering a discontinuity in a flow.

This method makes many assumptions about the scenario. First among them is that all particles under consideration have a velocity much greater than the bulk flow speed anywhere in the shock structure, \( v \gg u \). The second assumption made by the macroscopic method is that scattering keeps the phase space distribution *almost* isotropic\(^6\), so that the diffusion approximation is always valid. Further, the scattering occurs in the local plasma frame, is always elastic, and is against objects of infinite mass. The shock itself is an infinite plane parallel to the \( y \) and \( z \) axes and moving in the \(-\hat{x}\) direction, and quantities (like the phase space distribution function and bulk fluid velocities) vary only in the \( x \) dimension. The bulk fluid velocity as measured in the shock frame, specifically, has the following form:

\[
 u(x) = \begin{cases} 
 u_0 & x < 0 \\
 u_2 & x \geq 0 
\end{cases}
\]  

In Equation 1.2, and elsewhere in this dissertation, the subscript “0” refers to the values far upstream from the shock. Subscript “2” refers to values downstream from the shock—allowable *a posteriori* because I will show that quantities are constant downstream from the shock—and subscript “1” is reserved for values just upstream from the shock. In this particular derivation there is no difference between far upstream and just upstream of the shock, though later this constraint will be relaxed. Finally, the scenario is steady state, so that all time derivatives are zero.

The transport equation for \( f(x,p) \), the isotropic part of the total phase space distribution \( F(\vec{x},\vec{p},t) \) is given by the Boltzmann equation (e.g., Ellison and Eichler, 1984),

\[
 \frac{\partial f}{\partial t} + u \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} \left( D(x,p) \frac{\partial f}{\partial x} \right) + \frac{1}{3} \frac{\partial u}{\partial x} p \frac{\partial f}{\partial p},
\]  

where \( D(x,p) \) is the diffusion coefficient for particles. As well, \( f \) and \( p \) are defined in the local

---

\(^6\)The “almost” is important. I will return to that later in the derivation.
plasma frame, while \( x \) and \( u \) are defined in the shock frame. I allow for the moment that particle diffusion may depend on both position and momentum. Since \( \partial f / \partial t = \partial u / \partial x = 0 \), the transport equation reduces to

\[
u \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} \left( D(x, p) \frac{\partial f}{\partial x} \right). \tag{1.4}
\]

The first (trivial) solution to Equation 1.4 is that \( f \) and \( D \) do not depend on \( x \), i.e. no change across the “shock”. Clearly the other solution is the physically significant one, and may be found using standard differential equations techniques to be

\[
f(x, p) = \begin{cases} 
Q(p) + R(p) \cdot \exp \left[ \int_{0}^{x} \frac{u_0}{D(x, p)} \, dx' \right] & x < 0 \\
S(p) + T(p) \cdot \exp \left[ \int_{0}^{x} \frac{u_2}{D(x, p)} \, dx' \right] & x \geq 0
\end{cases}, \tag{1.5}
\]

where \( Q, R, S, \) and \( T \) are unknown functions of momentum that must be determined by applying known (or inferred) boundary conditions.

At the far upstream edge of the shock structure, as \( x \to -\infty \), the phase space distribution must match \( f_{\text{UpS}}(p) \)—that of the space into which the shock is expanding (as viewed from the far upstream, not the shock, frame). Thus \( f(x \to -\infty, p) \) must approach \( f_{\text{UpS}}(p) \). Far downstream, the phase space distribution must remain finite. Actually applying both conditions depends on the dependence of \( D(x, p) \) on position. If \( D \) is independent of \( x \), or if the dependence on position is sufficiently weak, then the integrals associated with \( R(p) \) and \( T(p) \) approach \( \pm \infty \) as \( x \to \pm \infty \). In this case, \( T(p) \) must be identically 0, and the term containing \( R(p) \) vanishes far upstream. Taking into account the boundary conditions, I rewrite Equation 1.5:

\[
f(x, p) = \begin{cases} 
f_{\text{UpS}}(p) + R(p) \cdot \exp \left[ \int_{0}^{x} \frac{u_0}{D(x, p)} \, dx' \right] & x < 0 \\
f_{\text{Dws}}(p) & x \geq 0
\end{cases}. \tag{1.6}
\]

We may assume we know \( f_{\text{UpS}} \), but not \( R(p) \) or \( f_{\text{Dws}} \). To determine those, we apply continuity of both \( f \) and its derivative across the shock. As well, a new assumption is needed: that the shock is strictly parallel, such that the bulk magnetic field is perfectly aligned with the motion of the shock (and so perpendicular to the shock face everywhere). There is therefore no discontinuity in magnetic field, and minimal energy gain from crossing.

Recalling that \( F(\vec{x}, \vec{p}, t) \) was only almost isotropic, expand it in terms of an isotropic part and
moments thereof:  

\[ F(x, p, \mu) \approx f(x, p) - \lambda \frac{\partial}{\partial x} f(x, p) + O\left( \frac{\partial^2}{\partial x^2} \right). \]  

(1.7)

In the above equation, \( \lambda = 3 D(x, p)/v \) is the mean free path between scatterings of a particle with velocity \( v \). Insert Equation 1.6 into Equation 1.7, and take the limits at \( x = 0^+ \) and \( x = 0^- \) to arrive at

\[
F(x, p, \mu) \approx \begin{cases} 
    f_{\text{ups}}(p) + R(p) \cdot \frac{3 u_0}{v} & x = 0^- \\
    f_{\text{dws}}(p) & x = 0^+ 
\end{cases}.
\]  

(1.8)

From here onward, in all expansions terms of second or higher order in \( u/v \) are discarded as negligible, recalling the assumption that \( u \ll v \).

Recall that \( f_{\text{ups}} \), \( f_{\text{dws}} \), and \( p \) have all been defined in the plasma frame, while \( x \) is defined in the shock frame. Since continuity must hold in the shock frame, define a new function \( F'(x, p', \mu') \) that represents the phase space distribution \( F \) transformed into the new frame. Now, expand \( F' \) in terms of \( F \):

\[
F'(x, p', \mu') \approx F(x, p, \mu) + (p' - p) \frac{\partial F}{\partial p} + (\mu' - \mu) \frac{\partial F}{\partial \mu},
\]  

(1.9)

where \( p' \approx p(1 - \mu u/v) \) is (to first order) the Lorentz-transformed velocity, and \( \mu' \approx \mu - (u/v)(1 - \mu^2) \) is the new pitch angle. Next I plug Equation 1.8 into Equation 1.9, again eliminating terms of order \( (u/v)^2 \) or higher, and again taking the limits on either side of the shock. The result is

\[
F(x, p, \mu) \approx \begin{cases} 
    f_{\text{ups}}(p) + R(p) - \mu \left[ \frac{u_0}{v} p \frac{\partial}{\partial p} \left( f_{\text{ups}}(p) + R(p) \right) + 3 \frac{u_0}{v} R(p) \right] & x = 0^- \\
    f_{\text{dws}}(p) - \mu \left( \frac{u_0}{v} p \frac{\partial}{\partial p} f_{\text{dws}}(p) \right) & x = 0^+ 
\end{cases}.
\]  

(1.10)

Continuity across the shock must apply to particles at any pitch angle, and leads to the following continuity equations,

\[
f_{\text{ups}}(p) + R(p) = f_{\text{dws}}(p)
\]  

(1.11)

\[
u_0 p \frac{\partial}{\partial p} \left( f_{\text{ups}}(p) + R(p) \right) + 3 u_0 R(p) = u_2 p \frac{\partial}{\partial p} f_{\text{dws}}(p),
\]  

(1.12)

which are normally interpreted as conservation of particle number and particle flux across the shock.

The equation that can be solved for \( f_{\text{dws}} \) is obtained by eliminating \( R(p) \) from Equations 1.11

---

7I have replaced the arguments \( \vec{x}, \vec{p}, \) and \( t \) with \( x, p, \) and \( \mu \) here; \( \mu \) is the pitch angle of the particle, i.e. the angle between the particle’s momentum and the +\( \hat{x} \) axis. This accounts for both the steady-state nature of the solution and the anisotropy in \( F \) introduced here.
and 1.12 and defining the compression ratio $r \equiv u_0/u_2$:

$$p \frac{\partial f_{\text{DwS}}}{\partial p} = \sigma (f_{\text{UpS}} - f_{\text{DwS}}),$$

(1.13)

where $\sigma \equiv 3r/(r-1)$ is defined in terms of the compression ratio. Again using standard differential equations techniques, the solution for $f_{\text{DwS}}$ is given by

$$f_{\text{DwS}}(p) = \sigma p^{-\sigma} \int_0^p p'^{-\sigma} f_{\text{UpS}}(p') dp' + C p^{-\sigma}. \tag{1.14}$$

In Equation 1.14, $C$ is a constant of integration, and according to Drury (1983) represents thermal particles injected into the acceleration process. The integral term, per Blandford and Ostriker (1978), means that any seed population of cosmic rays exits the Fermi-I process with the same $p^{-\sigma}$ power law at high $p$, regardless of the form of $f_{\text{UpS}}(p)$.

### 1.2.2 The microscopic method

The macroscopic method outlined above was the first published, but it offers no context at the level of the individual particle; all of this information was swept under the rug with the presentation of Equation 1.3. There is essentially no information provided about any given particle beyond the assumptions that $v \sim c$ and that diffusion occurs.

The microscopic method, first discussed by Bell (1978), examines the Fermi-I process at the level of the individual particle by asking two basic questions. First, what is the probability that an energetic particle with $v \gg u$—taken from an isotropic distribution—that has made a shock crossing $\text{UpS} \rightarrow \text{DwS}$ returns to make another $\text{DwS} \rightarrow \text{UpS} \rightarrow \text{DwS}$ cycle? Second, what is the average energy gain after completing one such cycle? Before continuing, note that the microscopic method uses the same shock described in Equation 1.2.

To answer the first question, consider the population of particles downstream from the shock, moving in a plasma with bulk speed $u_2$. Such particles will be able to recross the shock headed upstream if $v_x < -u_2$. Assuming that the distribution is roughly isotropic everywhere (except, perhaps, in the immediate vicinity of the shock for particles of lower energies than we are considering), $v_x$ is a random number uniformly distributed between $+v$ and $-v$, so the flux of particles returning to

---

8The exception is if $f_{\text{UpS}}(p)$ is itself a harder power law, i.e. $f_{\text{UpS}} \propto p^{-\tau}$ with $\tau < \sigma$. In this case $f_{\text{UpS}}$ is unchanged by the acceleration process and $f_{\text{DwS}} \propto f_{\text{UpS}} \propto p^{-\tau}$.

9Note that we sidestep the injection process here also.
cross DwS→UpS is proportional to
\[ \left| \int_{-v}^{-u_2} (u_2 + v_x) \, dv_x \right| = \frac{1}{2} (u_2 - v)^2. \] (1.15)

The flux of particles crossing in the opposite direction—from upstream to downstream—is those with \( v > -u_2 \), i.e.
\[ \int_{-u_2}^{v} (u_2 + v_x) \, dv_x = \frac{1}{2} (u_2 + v)^2. \] (1.16)

The probability that a particle will return to cross the shock DwS→UpS exactly once is the ratio of the two fluxes:
\[ P(1) = \frac{\text{Flux DwS→UpS}}{\text{Flux UpS→DwS}} = \left( \frac{u_2 - v}{u_2 + v} \right)^2, \] (1.17)
a result that holds anywhere in the downstream region that isotropy is assumed, even right at the shock for sufficiently energetic particles. By mathematical coincidence, Equation 1.17 is relativistically correct (Peacock, 1981), even though it was derived above in the nonrelativistic limit. Assuming that no particles escape upstream from the shock, the probability that a particle will complete \( N \) UpS→DwS→UpS cycles is simply
\[ P(N) = \prod_{i=1}^{N} \left( \frac{u_2 - v_i}{u_2 + v_i} \right)^2 = \left[ \prod_{i=1}^{N} \left( \frac{1 - (u_2/v_i)}{1 + (u_2/v_i)} \right) \right]^2, \] (1.18)
where I allow for the possibility that the particle’s velocity might change with each crossing cycle. This is the answer to the first question.

As for the second question, about energy gains per shock crossing, consider a particle whose momentum, velocity, and pitch angle are \( p, v, \) and \( \mu \) in the upstream frame. Lorentz transformed into the shock frame, \( p_{\text{sf}} = p(1 + \mu u_0/v) \). This momentum will not change until the particle’s first scattering in the downstream region, so in the reference frame of the shocked plasma \( p_{\text{DwS}} = p(1 + \mu(u_0 - u_2)/v) \).

The probability that a particle will cross the shock is proportional to its pitch angle in the shock frame—particles moving purely in the plane of the shock will never reach it, a result I will return to in Section 3.3. Averaging over all particles directed downstream upon crossing the shock (and including the correct normalization constant),
\[ \langle \Delta p_{\text{UpS→DwS}} \rangle = \int_{0}^{1} \left[ p_{\text{UpS}} - p_{\text{DwS}} \right] \cdot 2\mu \, d\mu = \int_{0}^{1} p \cdot \frac{u_0 - u_2}{v} \cdot 2\mu^2 \, d\mu = \frac{2}{3} p \cdot \frac{u_0 - u_2}{v}. \] (1.19)
In the opposite direction, switch \( u_0 \) and \( u_2 \) and consider a different subset of pitch angles:

\[
\langle \Delta p_{\text{DWS} \rightarrow \text{UpS}} \rangle = \int_{-1}^{0} \frac{p \cdot (u_2 - u_0)}{v} \cdot 2\mu^2 \, d\mu \]

\[
= \frac{2}{3} p \cdot \frac{u_0 - u_2}{v} = \langle \Delta p_{\text{UpS} \rightarrow \text{DWS}} \rangle. \tag{1.20}
\]

The total change in momentum after a full cycle is \( \langle \Delta p_{\text{Cyc}} \rangle = \frac{4}{3} p (u_0 - u_2) / v \), to first order in \( u / v \). The momentum of a particle that has completed \( N \) such cycles, starting from momentum \( p_0 \) and velocity \( v_0 \), is

\[
p_N = p_0 \cdot \prod_{i=1}^{N} \left[ 1 + \frac{4}{3} \frac{u_0 - u_2}{v_i} \right]. \tag{1.21}
\]

Equation 1.21 should in reality be a very complicated distribution that takes into account the overall change in momentum based on pitch angle history at shock crossings. However, when \( N \) is large the central limit theorem forces the distribution towards a single sharp peak located at the value of \( p_N \).

Equations 1.18 and 1.21 represent the answers to the two questions I posed at the start of this section. All that remains is to combine them to determine the phase space distribution of particles as a function of momentum. To this, I take the natural logarithm of both equations, expanding as a Taylor series where appropriate and truncating any expansions after first order in \( u / v \).

\[
\ln p_N = \ln p_0 + \frac{4}{3} (u_0 - u_2) \sum_{i=1}^{N} \frac{1}{v_i} \quad \Rightarrow \quad \ln \left( \frac{p_N}{p_0} \right) = \frac{4}{3} (u_0 - u_2) \sum_{i=1}^{N} \frac{1}{v_i} \tag{1.22}
\]

and

\[
\ln \mathcal{P}(N) = 2 \sum_{i=1}^{N} \ln \left[ \frac{1 - \frac{u_2}{v_i}}{1 + \frac{u_2}{v_i}} \right] = -4 \sum_{i=1}^{N} \frac{u_2}{v_i}. \tag{1.23}
\]

Strictly speaking, I have performed a small sleight of hand regarding the use of \( v_i \) in the two above equations. In Equation 1.22 it is defined in the upstream plasma frame, while in Equation 1.23 it is defined in the downstream frame. However, since \( v \gg u_0 (> u_2) \), the two are equal to first order in \( u / v \).

Finally, noting the presence in both equations of a summation over \( 1 / v_i \), substitute to obtain

\[
\ln \mathcal{P}(N) = -\frac{3u_2}{u_0 - u_2} \ln \left( \frac{p_N}{p_0} \right), \tag{1.24}
\]

or

\[
\mathcal{P}(N) = \left( \frac{p_N}{p_0} \right)^{-\frac{3u_2}{u_0 - u_2}}. \tag{1.25}
\]
Thus the number of particles in the population downstream of the shock at a given momentum, as a function of that momentum, is

\[ N(p) \propto \left( \frac{p}{p_0} \right)^{-\frac{3u_2}{u_0-u_2}}, \]  

(1.26)

up to constants related to normalization. Rewriting in terms of the downstream phase space distribution \( f_{\text{DWS}} \),

\[ f_{\text{DWS}}(p) \propto \frac{1}{p^2} \frac{\partial N}{\partial p} = \left( \frac{p}{p_0} \right)^{-\frac{3u_2}{u_0-u_2}-3} = \left( \frac{p}{p_0} \right)^{-\frac{3u_0}{u_0-u_2}} \]  

(1.27)

Using the definition of the compression ratio \( r \) from the macroscopic derivation, the spectral index in phase space is exactly \( 3r/(r-1) \). The microscopic derivation thus recovers the same phase space distribution as the macroscopic method, in spite of its markedly different approach to the situation.

1.2.3 Comments and Consequences

The remarkable equivalence\(^{10} \) between the microscopic and macroscopic methods suggests that the collected assumptions, consequences, and limitations of the two are equivalent also. The first-order Fermi process possesses a great deal of explanatory power, but it is not without limitations.

- The shock in both the macro- and microscopic methods must be collisionless, i.e. the mean free path between particle-particle collisions must be much larger than the shock structure. If the shock is collisional (as is the case with all terrestrial shocks) the shocked gas quickly relaxes to a thermal distribution downstream from the shock. Only extremely thin plasmas can support the infrequent collisions needed to allow a high-energy tail to form, and those plasmas are found exclusively in astrophysical settings.

- For hydrodynamical shocks (whose equations govern even the evolution of collisionless shocks as long as magnetic field strength is sufficiently low), the compression ratio \( r \rightarrow 4 \) as shock strength increases.\(^{11} \) As a consequence of this, \( \sigma \rightarrow 4 \). This spectral index is close to what is observed here at Earth in the cosmic ray distribution up to the ankle (see Section 1.1.2), and much of that difference can be explained by residence times in the Milky Way for these very energetic particles. This was an encouraging early sign that the Fermi-I process might be responsible for the bulk of the cosmic rays arriving at Earth from outside the Solar System.

\(^{10}\)And it is a true equivalence; see Jones and Ellison (1991) for slightly different derivations of both the micro- and macroscopic approaches, and whose conclusion is that the two methods yield identical results.

\(^{11}\)This holds as long as the ratio of specific heats downstream is \( \gamma = 5/3 \), appropriate for a largely non-relativistic population of particles. As will be shown later, this assumption about the mean downstream particle energy is increasingly invalid as the shock speed increases.
The results of both methods apply to parallel shocks only (those whose propagation vector is parallel to the mean external magnetic field the shock is encountering), or at least those with $u_0 / \cos(\theta_{Bn0}) < c$, where $\theta_{Bn0}$ measures the difference between the shock normal and the upstream magnetic field. Shocks that meet this condition may be transformed into a reference frame where the two vectors are parallel, and the above derivations hold. In shocks that fail, the point traced out by the intersection of a magnetic field line with the shock surface moves faster than light, and the shock is called superluminal. This latter variety of shocks can be transformed into a reference frame where the shock normal and magnetic field vectors are completely perpendicular. In that case a different method for particle acceleration is needed to explain observations of particle acceleration at such shocks.

Neither derivation deals with the issue of injection: how initially thermal particles find themselves caught up in the acceleration process. For these particles, $u \sim v$, and the distribution cannot be isotropic near the shock. Any treatment of particle acceleration that is not completely from first principles must face this issue, and the various methods for handling DSA each take a different approach.

The energy contained in a phase space distribution $f(p) \propto p^{-4}$ is logarithmically divergent if the distribution continues to infinite momentum. There must be a cutoff at some energy, but such a cutoff does not come from within the basic picture shown above. (See Section 3.5 for more information.)

Beyond very weak assumptions about the dependence on position, the analytical models do not care how particles diffuse (i.e., the exact form of $D(x, p)$), only that they do. Turbulence of the magnetic field, off of which particles can scatter, is thus implicitly assumed. The origin of this turbulence, like the maximum energy achievable by any given cosmic ray, is beyond the scope of these initial derivations.

Two final limitations drive the work done in this dissertation. First among these is the assumption made in both derivations that $v \gg u$ everywhere within the shock structure. As mentioned above with regards to injection, this test can fail when $v$ is low. It can also fail when $u$ is high, as is the case in the transrelativistic and fully relativistic regimes. As $u \rightarrow c$, the character of the Fermi-I process changes. While some things, like Equation 1.17, continue to apply, many other facets of the above two approaches cease to do so. The study of particle acceleration by relativistic shocks is discussed further in Section 1.3.

The second limitation involves Equation 1.2, the unmodified or “test-particle” regime. It was realized very early on (Axford et al., 1977, in fact) that if particle acceleration is efficient, there might
be significant cosmic ray pressure ahead of the shock in the upstream region. The cloud of cosmic rays diffusing ahead of the shock would not leave the upstream velocity profile unchanged. The backpressure from the cosmic rays would decelerate the upstream plasma slightly,\(^\text{12}\) modifying the shock profile from the sharp discontinuity suggested by Equation 1.2. This, in turn, would impact the process of particle acceleration, which feeds back into the modification of the shock structure again. The interaction between the shock profile and the particles it accelerates is highly nonlinear, even with the assumption of a steady state, and will be taken up again in Section 1.4.

### 1.3 DSA at Relativistic Shocks

As I mentioned above, one assumption made in both the macro- and microscopic derivations of DSA is that the particles are isotropic. This requires particles with \(v \gg u_0\) (and, consequently, \(\gg u_2\)), as the angular distributions of these particles are mostly unaffected by the transformation between the various reference frames in the vicinity of an astrophysical shock.

In relativistic shocks the upstream flow speed toward the shock is a significant part of the speed of light. As \(u_0 \to c\) and the Lorentz factor\(^\text{13}\) \(\gamma_0 \to \infty\), the isotropy approximation fails, since Lorentz transformation between frames introduces significant anisotropies at ever higher energies. It is not at all certain that a diffusive process like the first order Fermi process could occur at relativistic shocks as it does in the nonrelativistic limit. It is, however, very tempting to consider these sites as candidate locations for DSA. Locations where relativistic shocks are expected to occur include gamma-ray bursts (GRBs); the termination shocks of both pulsar wind nebulae (PWNe) and jets from active galactic nuclei (AGNs); blazars; or exotic objects like relativistic supernovae. All of these objects have very large energy budgets, so even a small efficiency at producing CRs could have dramatic effects.

The issue of DSA at relativistic shocks was first examined by Peacock (1981). Using an extension of the microscopic approach of Bell (1978), Peacock (1981) assumed that suprathermal particles could diffuse in the downstream region: for test particle shocks the downstream flow speed is at most \(c/3\), which is small enough compared to \(c\) that the diffusion approximation might apply. Peacock (1981) then calculated the conditional probability that a particle crossing the shock at angle \(\theta_1\) would return at angle \(\theta_2\). By numerically averaging these probabilities and folding in the energy gain at each crossing, Peacock (1981) concluded that DSA could indeed operate at relativistic shocks, with two caveats:

\(^\text{12}\)In the shock frame, that is. In the rest frame of the upstream plasma, the fluid is accelerated slightly before being swept up by the shock front.

\(^\text{13}\)Defined as \(\gamma = 1/\sqrt{1-(u/c)^2}\) for a particle or shock moving at speed \(u\).
1. The shock in question must be parallel, as otherwise particles would be advected downstream on magnetic field lines before they could be injected into the acceleration process. (Though Peacock, 1981, did allow for the possibility of magnetic turbulence causing cross-field line diffusion, but only as an aside.)

2. The minimum energies required for electrons to be injected into the Fermi process were worryingly large: if a relativistic shock were able to compress and shock-heat the helium population impinging upon it, electrons would need to have energies above 2 GeV (corresponding to a Lorentz factor $\gamma_e \gtrsim 4000$) before they would see the shock as sufficiently transparent for the acceleration derivation to apply. This would require significant energy transfer from ions to electrons.

With the possibility of DSA at relativistic shocks established, later work focused on topics like energy gain, the spectral index of the CR population, and the maximum energy of the CR spectrum. The energy gain for particles crossing a relativistic shock almost parallel to the shock normal is a factor of $\sim \gamma_0^2$, which for the relativistic shocks expected in gamma-ray bursts (GRBs; see Chapter 2) might be a factor of 10$^5$ or more. As pointed out by Vietri (1995), if protons could execute just a few such cycles, they could attain energies at and above the ankle of the CR spectrum. The derivation of Vietri (1995) assumed that particles could undergo large changes to their orientation as they scattered, which is far from certain. A counterargument was raised by Gallant and Achterberg (1999), who used an alternate prescription for particle diffusion. Considering either (1) regular deflection about an ordered magnetic field line, or (2) small but continuous deflections due to magnetic field turbulence, Gallant and Achterberg (1999) found that the second and later crossings realized energy gains of $E_f/E_i \simeq 2 - 2.5$, far below the optimistic predictions of Vietri (1995).

As with nonrelativistic shocks, cosmic rays accelerated by relativistic shocks form a power law in energy, $N(E) \propto E^{-\sigma}$. Unlike nonrelativistic shocks, which were shown above to produce $\sigma = 2$ power laws, the resultant spectra at relativistic shocks appear to approach a so-called “universal” spectral index of $\sigma \approx 2.2 - 2.3$ (Bednarz and Ostrowski, 1998; Kirk et al., 2000; Achterberg et al., 2001; Ellison and Double, 2002; Keshet and Waxman, 2005). This result holds in the pitch-angle diffusion (PAD) limit (where $\Delta \mu \ll 1$ at any given scattering event); as particles are allowed to scatter through

---

14 If helium ions treat the shock as a true discontinuity, the length scale over which the fluid quantities change must be of order a few ion gyroradii. The gyroradius of electrons is much smaller, due to their lower mass-to-charge ratio, and so the shock stretches over more gyroradii. Thus electrons must achieve momenta roughly equal to the helium ions before the shock appears as a discontinuity to them as well.

15 Though, it must be said, at the time of publication there was very little evidence in favor of any scattering prescription in the exotic regimes used by Vietri (1995).

16 PAD stands in contrast to the large-angle scattering assumptions of Vietri (1995). Unlike in large-angle scattering, particles undergoing PAD experience frequent but very small changes to their momentum orientation as they scatter.
progressively larger angles, the spectrum becomes harder (Ellison et al., 1990b; Ellison and Double, 2002; Baring, 2004). Note that Baring (2004), citing the results of Ellison et al. (1990b), discusses a seeming non-universality of the spectral index at relativistic shocks. However, the shocks studied in Ellison et al. (1990b) were not modified by the nonlinear effects that will be discussed below; based on the results to be presented in Chapter 4, it is possible that many of those shocks could not be modified to conserve fluxes.

Regarding the upper limit of the CR spectrum produced at relativistic shocks, observations of particles above the GZK limit now date back more than five decades (Linsley, 1963). There is clear evidence that at least some particles can attain those energies, and the energy budget inherent in a relativistic shock makes them prime candidates for the acceleration site. An estimate for the maximum energy a CR can achieve in a particular environment is the Hillas (1984) criterion,

$$E_{\text{max}} \sim Ze BR_s,$$  \hspace{1cm} (1.28)

and its extension to relativistic shocks (Achterberg et al., 2001),

$$E_{\text{max}} \sim Ze BR_s \gamma_s \beta_s.$$  \hspace{1cm} (1.29)

In both above equations, the maximum energy achievable by a particle of charge $Ze$ scales with the magnetic field strength $B$ and the size $R_s$ of the acceleration region. In the relativistic extension, however, an additional gain factor of $\gamma_s \beta_s$ is realized because particles typically complete only a small fraction of their gyroperiod upstream of the shock before being caught up again. This lends credence to the possibility that relativistic shocks might be responsible for the highest-energy particles observed in the Universe.

### 1.4 DSA at Nonlinear Shocks

In the derivations outlined in Sections 1.2.1, 1.2.2, and 1.3 it was assumed that the velocity profile of fluid in the neighborhood of the shock was described by Equation 1.2. That is, the shock completely decelerates and heats incoming particles in a region of effectively infinite thinness. However, in order for DSA to place energy into cosmic rays they must spend at least some of their time ahead of the shock. Furthermore, if one makes the very reasonable assumption that the diffusion coefficient $D(x,p)$ (or, alternately, the mean free path $\lambda_{\text{mfp}}$ between scatterings) is an increasing function of energy, then one arrives at the conclusion that higher energy particles can diffuse further upstream against the unshocked flow than can their lower-energy counterparts.
If the shock is efficient at depositing energy into cosmic rays,\(^\text{17}\) then there will be a non-negligible population of accelerated particles preceding the shock. Unshocked particles approaching the shock will encounter the cloud of CRs first, which will slightly decelerate and heat them before finally being shock-heated at the subshock. This alters the strength of the subshock and changes the nature of both injection into the acceleration process and the energy gains realized once acceleration has begun. It is further the case that the population of CRs increases as cold plasma approaches the subshock, leading to a position-dependent modification of the shock’s velocity profile.

The presence of CRs ahead of the shock, and their effect on the incoming plasma, is a highly nonlinear problem that was first studied (in the nonrelativistic limit) during the 1980s (see the discussion in Jones and Ellison, 1991, and references therein). In the (comparatively) simple nonrelativistic limit, nonlinear steady-state shocks must satisfy the three Rankine-Hugoniot relations for mass, momentum, and energy flux conservation:

\[
\begin{align*}
\rho_0 u_0 &= \rho(x)u(x) \\
\rho_0 u_0^2 + P_0 &= \rho(x)u^2(x) + P(x) \\
\frac{1}{2}\rho_0 u_0^3 + \frac{\Gamma_0}{\Gamma_0 - 1} P_0 &= \frac{1}{2}\rho(x)u^3(x) + \frac{\Gamma(x)}{\Gamma(x) - 1} P(x) + Q_{\text{esc}}
\end{align*}
\]

In Equation 1.32, the quantity \(\Gamma\) is the mean adiabatic index of all plasma—thermal and CR—at a particular location. The goal of studying nonlinear shocks is to find velocity profiles \(u(x)\) that self-consistently generate cosmic-ray spectra whose effect up- and downstream of the shock is to generate sufficient pressure to recover \(u(x)\). If the wrong velocity profile is chosen, either (1) too many particles will be accelerated, and the momentum and energy fluxes will be higher than their upstream values, or (2) too few particles will be accelerated, and the right hand sides of Equations 1.31 and 1.32 will be lower. The problem becomes more difficult still once one includes the possibility of escaping energy flux, \(Q_{\text{esc}}\) (per Ellison, 1985, the escaping momentum flux is a factor \(u/c\) lower than the escaping energy flux, and is ignorable) is considered. In addition, one must admit the possibility that the mixture of nonrelativistic (unaccelerated) and fully relativistic (CR) particles have an adiabatic index that deviates from the nonrelativistic limit of \(\Gamma = 5/3\), and may in fact be position-dependent within the shock profile.

Extension of nonlinear DSA to the relativistic regime is more difficult still, and has received somewhat less attention. Prior to Particle-in-Cell (PIC) simulations (see Section 1.5.2 below)—which treat all physics self-consistently and so necessarily conserve fluxes while accelerating particles...
in a varying velocity profile—few attempts had been made to study this regime of DSA. Schneider and Kirk (1987) devised a means of calculating accelerated particle spectra for an arbitrary velocity profile, including the spectral index of the power law formed by the Fermi-I process. However, no attempt was made in that work to determine whether the profiles and accelerated populations adhered to the relativistic Rankine-Hugoniot equations (themselves given in Section 3.6). Very early PIC simulations (Hoshino et al., 1992) studied nonlinear DSA at relativistic oblique shocks, but the computational power available at that point limited the dimensionality and spatial/temporal scales of the simulation. The first Monte Carlo simulations of nonlinear relativistic DSA were presented by Ellison and Double (2002), for parallel shocks and particles undergoing pitch-angle diffusion.

Given its importance to the study of gamma-ray bursts (GRBs; see Chapter 2), acceleration of CRs at relativistic shocks has received considerable attention. Much recent work has focused on the propagation of CRs within a turbulent magnetic field (Niemiec and Ostrowski, 2006; Lemoine and Revenu, 2006; Lemoine, 2013; Plotnikov et al., 2013), or on the generation of turbulence ahead of relativistic shocks by the accelerated CRs (see, e.g., the extended discussion in Lemoine and Pelletier, 2010). However, besides the handful of Monte Carlo results mentioned above and the PIC results to be discussed below, no studies of which I am aware have attempted to consider the backreaction of the accelerated particles on the shock.

### 1.5 Approaches to DSA

Given the importance of DSA to the field of cosmic rays and the energetic environments that accelerate them, it is unsurprising that many different methods have been used in its study. In this section I will describe the three main avenues along which research has been performed. These are the analytic (and semi-analytic) approach, Particle-in-Cell (PIC) simulations, and finally the Monte Carlo method that will be used for the work presented in this dissertation. I will confine myself to discussing only work done in the relativistic regime; a staggering amount of work has been done in the non-relativistic regime which, unfortunately, must be omitted for reasons of space and time.

#### 1.5.1 Analytic and Semi-analytic

The earliest work on nonrelativistic DSA (Krymskii, 1977; Axford et al., 1977; Blandford and Ostriker, 1978; Bell, 1978) was done analytically, and this trend continued in the relativistic regime. The work of Peacock (1981) was discussed above, which concluded that, with some caveats, the Fermi-I process could operate even in shocks where the diffusion approximation breaks down.

This result has been extended by numerous analytical and semi-analytical studies, all of which
1.5. Approaches to DSA

1. INTRO TO COSMIC RAYS AND DSA

had to confront the anisotropy of cosmic rays in the vicinity of the subshock. Some, including Kirk and Schneider (1987a) and Kirk et al. (2000), treated the upstream and downstream particle distributions as an eigenfunction problem, with the primary goal being to match the two functions at the boundary (i.e., subshock). Other work (Gallant and Achterberg, 1999; Achterberg et al., 2001) made simplifying assumptions about the structure of the magnetic field and explored consequences for energy gain and acceleration time. By restricting the scattering prescription to just pitch-angle diffusion, Keshet and Waxman (2005) derived the spectral index of accelerated particles at relativistic shocks, for arbitrary upstream and downstream velocities.

What all of the above approaches have in common is exceeding quickness for use in, say, hydrodynamical computations involving relativistic shocks. However, the analytic approach is fundamentally limited by the complexity of DSA. In order to deal with shocks of arbitrary speed, unknown magnetic field strength and turbulence, and potentially including the backreaction of accelerated CRs on the shock, fully numerical methods are needed.

1.5.2 Particle-in-Cell

The gold standard for numerical study of DSA is the Particle-in-Cell (PIC) approach. A (typically) large grid is seeded with both a magnetic field and a population of charged particles, and all microphysics is handled self-consistently. At every time step, particles are moved in response to the magnetic and electric fields at their position. Then, the fields are updated based on the motion of the particles. PIC simulations can examine the formation of a relativistic shock, the generation of the magnetic turbulence responsible for particle scattering, energy transfer between particle species, and wave–particle interactions at their most fundamental level.

Prior to the work of Spitkovsky (2008), PIC simulations had shown at best inconclusive (and at worst negative) evidence that relativistic shocks could support the Fermi-I process. Two-dimensional PIC simulations have a great deal of difficulty supporting the cross-field diffusion necessary for DSA to occur (Jones et al., 1998), and to that point the only 3-D simulations performed were short in duration (e.g., Nishikawa et al., 2003; Hededal et al., 2004), admitting the possibility that the instabilities that drive relativistic shock formation were not yet saturated and calling into question claims of (non-)detection of Fermi-I acceleration.

In recent years a great deal of numerical work has been done using 3-D PIC simulations (Haugbolle, 2011; Sironi and Spitkovsky, 2011; Sironi et al., 2013) and large-scale 2-D (or “2.5-D”) simulations (Sironi and Spitkovsky, 2009a,b; Keshet et al., 2009). The simulations show that, especially in weakly magnetized plasmas typical of the interstellar medium, diffusive shock acceleration undoubtedly occurs for both electrons/positrons and much heavier ions.
1.5. APPROACHES TO DSA

The fact that simulations must be at least 2-D, and that all microphysics is calculated self-consistently, points to the single greatest drawback of PIC simulations. They have, to date, been run only on relatively small grid sizes, for relatively short periods of time. The small size of the computation boxes, particularly in the transverse direction, has profound impacts on the maximum wavelength of turbulence the simulations can support, which in turn affects the maximum energy of particles they can produce. The large mass ratio of ions to electrons requires very fine grid spacing to resolve the motion of the lightest particles. In addition, PIC simulations are limited by the length of the simulation box: to run to later times requires additional grid zones for the shock to overtake, leading to additional computation time. The conflict between these three limitations was outlined in Vladimirov et al. (2008), who concluded that—despite the profound contributions made in studying the initial stages of shock formation and particle acceleration—PIC simulations are unlikely to probe particle energies anywhere near the knee in the CR spectrum for the foreseeable future.

1.5.3 Monte Carlo

Monte Carlo simulations represent the middle ground between the computational speed of semi-analytical methods and the generality of PIC simulations. Compromises are typically made compared to the total self-consistency of PIC simulations (e.g. through parametrizing away the wave–particle interactions into a small number of scattering parameters). On the other hand, Monte Carlo simulations can self-consistently handle matters such as energy gain, acceleration times, nonlinear feedback, particle escape, and anisotropy to a degree beyond all but the most intensive semi-analytical methods—which typically deal with only a subset of the provided list rather than the entirety. As well, they do not rely on a computational grid like PIC simulations do, meaning they can (within the limitations of the assumptions made by the particular Monte Carlo implementation) study acceleration of CRs to the knee, the ankle, or beyond if the shock system allows for it.

Monte Carlo methods have been used in the study of nonrelativistic shocks almost since the initial conception of DSA as a means to accelerate cosmic rays (e.g., Ellison et al., 1981; Ellison and Eichler, 1984). The first use of a Monte Carlo code to examine relativistic shocks appears to be that of Kirk and Schneider (1987b), who simulated pitch-angle diffusion of particles in mildly relativistic test-particle shocks. This work was quickly followed by Quenby and Lieu (1989), who used isotropic diffusion to study DSA at the termination shocks of AGN jets; and Ellison et al. (1990b), who extended a successful nonrelativistic code to the (test-particle, i.e. unmodified) relativistic regime and documented the changes in acceleration time and spectral index that occurred. A parametrized turbulent amplification of the magnetic field was assumed in Bednarz and Ostrowski (1996), which was later extended to studies of DSA at oblique relativistic shocks (Bednarz and
1.6. SUMMARY

Cosmic rays have been a source of fascination since their discovery in 1912 by Victor Hess. The collisionless shocks that accelerate them to astonishing energies—and in particular, that subset of these shocks that is relativistic—occur near the most energetic and violent events in the known Universe. Studying their acceleration via the Fermi-I process means exploring high-energy environments beyond the capability of terrestrial experiments. This, in turn, requires numerical analysis to grapple with the complicated questions of magnetic field turbulence generation, injection, diffusion, acceleration efficiency, and the nonlinear feedback between cosmic rays and the shocks they accelerate was treated for the first time in Ellison and Double (2002). A set of Monte Carlo studies (Lemoine and Pelletier, 2003; Lemoine and Revenu, 2006) simulated DSA at relativistic shocks by building conditional probabilities between the pitch angles $\mu_1$ and $\mu_2$ at which CRs at a given energy crossed and returned to the shock. Niemiec and Ostrowski (2004) and Niemiec and Ostrowski (2006) studied particle diffusion and acceleration in a “realistic” magnetic field structure, with an assumed Kolmogorov turbulence spectrum. Their results suggest that large-scale turbulence is extremely inefficient at accelerating particles in relativistic shocks, and—rather than recovering the universal $\sigma \approx 2.2 - 2.3$ power law discussed earlier—the spectral index of the CR distribution (if one even formed) was highly dependent on the shock speed, obliquity, and turbulence strength. They are supported by more recent simulations (e.g., Plotnikov et al., 2011, 2013) that study the diffusion of particles in the intense short-wavelength turbulence expected in the vicinity of relativistic shocks (Lemoine, 2013).

The above-mentioned works represent only a small sampling of Monte Carlo–DSA connection. Although compromises must inevitably be made in treating the microphysics of shock acceleration, Monte Carlo codes can scale from questions of non-relativistic injection to the acceleration time- and length-scales associated with the highest energy cosmic rays. It is this versatility that makes them ideal extensions of the PIC simulations just discussed, and excellent means to test the predictions made by analytic studies.

1.6 Summary

Cosmic rays have been a source of fascination since their discovery in 1912. The collisionless shocks that accelerate them to astonishing energies—and in particular, that subset of these shocks that is relativistic—occur near the most energetic and violent events in the known Universe. Studying their acceleration via the Fermi-I process means exploring high-energy environments beyond the capability of terrestrial experiments. This, in turn, requires numerical analysis to grapple with the complicated questions of magnetic field turbulence generation, injection, diffusion, acceleration efficiency, and the nonlinear feedback between cosmic rays and the shocks they accelerate was treated for the first time in Ellison and Double (2002). A set of Monte Carlo studies (Lemoine and Pelletier, 2003; Lemoine and Revenu, 2006) simulated DSA at relativistic shocks by building conditional probabilities between the pitch angles $\mu_1$ and $\mu_2$ at which CRs at a given energy crossed and returned to the shock. Niemiec and Ostrowski (2004) and Niemiec and Ostrowski (2006) studied particle diffusion and acceleration in a “realistic” magnetic field structure, with an assumed Kolmogorov turbulence spectrum. Their results suggest that large-scale turbulence is extremely inefficient at accelerating particles in relativistic shocks, and—rather than recovering the universal $\sigma \approx 2.2 - 2.3$ power law discussed earlier—the spectral index of the CR distribution (if one even formed) was highly dependent on the shock speed, obliquity, and turbulence strength. They are supported by more recent simulations (e.g., Plotnikov et al., 2011, 2013) that study the diffusion of particles in the intense short-wavelength turbulence expected in the vicinity of relativistic shocks (Lemoine, 2013).

The above-mentioned works represent only a small sampling of Monte Carlo–DSA connection. Although compromises must inevitably be made in treating the microphysics of shock acceleration, Monte Carlo codes can scale from questions of non-relativistic injection to the acceleration time- and length-scales associated with the highest energy cosmic rays. It is this versatility that makes them ideal extensions of the PIC simulations just discussed, and excellent means to test the predictions made by analytic studies.

---

$^{18}$A problem unique to oblique relativistic shocks is the possibility of “superluminal” shocks. For shocks of speed $u_0$ whose upstream mean magnetic field makes an angle $\theta_0$ with the shock normal, the intersection of the field line and the subshock travels with speed $u_0 / \cos(\theta_0)$. In relativistic shocks with $u_0 \approx c$, any deviation from the purely parallel state will cause the intersection point to move at superluminal speeds. This has profound implications for Fermi-I acceleration, since in the absence of cross-field diffusion particles are tied to a single field line. Particles in such a situation cannot possibly travel back to the shock faster than they are advected away by the downstream flow, essentially eliminating the Fermi-I process as a means for accelerating cosmic rays. Such shocks are, however, beyond the scope of the work presented in this dissertation; see Chapter 5 for an explanation of why.
that accelerate them. PIC simulations are limited in scope due to the required computation power, making Monte Carlo models the approach of choice in this dissertation.

In the next chapter, I turn to one particular source of relativistic shocks that is of great importance to the work presented in this dissertation: gamma-ray bursts.
In this chapter I will describe the current state of research in afterglows and gamma-ray bursts (GRBs) at it relates to cosmic-ray acceleration. As with the previous chapter, I will begin with a brief history of the field to the present day. (As with the previous chapter, “brief” here is relative to the substantial body of work that exists, rather than to the size of this dissertation.) I will then discuss the theory of afterglows and analytical/numerical approaches, observations that pertain to cosmic-ray acceleration, and inferred values (or ranges) of the parameters relevant to diffusive shock acceleration. Particularly in Section 2.1, I am grateful to Kumar and Zhang (2014) for providing historical context.

2.1 History of GRBs

The political atmosphere of the Cold War was not conducive to trust, so after the signing in 1963 of a ban on most nuclear tests (specifically those in space, in the atmosphere, and underwater), the United States decided to monitor Soviet compliance with the terms of the treaty. The means chosen to do so was the Vela constellation of satellites, whose X-ray and γ-ray instruments were

capable of detecting flashes of radiation. The time resolution of the instruments on the final pair of satellites—Vela 5 and Vela 6—was fine enough that a flash could also be localized in the sky based on when the satellites detected the emission.

On July 2nd, 1967, both Vela 4 satellites detected a flash of γ-rays, showing a twin-peaked structure and lasting four seconds. The light curve is shown in Figure 2.1. As this was a very different signal than what a nuclear explosion would generate, and as no solar flares or supernovae had been observed that day, the detection was set aside.\(^1\) In the years that followed, more of these bursts were observed, and the improved timing instrumentations on the satellites allowed for localizations. While GRB 670702\(^2\) was initially ignored in the first paper on this new class of astrophysical objects (Klebesadel et al., 1973, released shortly after the data from the Vela satellites was declassified), its similarity to bursts observed with the more precise Vela 5 and Vela 6 satellites prompted its inclusion in later lists of known events (Strong et al., 1974; Klebesadel et al., 1982).

Several hundred bursts were observed between 1967 and 1991, but the launch of the Compton Gamma Ray Observatory marked a sea change in the study of GRBs. The four instruments aboard CGRO (including the Burst and Transient Source Experiment, or BATSE) covered a wide energy

\(^1\)For a more thorough, if more fanciful, account of the discovery, see Schilling (2002).

\(^2\)Although supernovae are named using the only the year of discovery, e.g. SN 1987A or SN 2013cq, gamma-ray bursts are named using the date (YYMMDDD) and a letter if multiple bursts were observed on that day, e.g. GRB 670702 or GRB 080319B.
2.1. HISTORY OF GRBS

CHAPTER 2. GRBS AND AFTERGLOWS

Figure 2.2 A small sample of GRBs observed by the BATSE instrument. Note, for example, the difference in duration of Triggers 1606 and 2514; the difference in intensity of Triggers 1406 and 2151; and the variability in number and shape of the pulses. Image reproduced with permission, from https://heasarc.gsfc.nasa.gov/docs/objects/grbs/grb_profiles.html; image credit Jerry Bonnell (NASA/GSFC).

range and a very large fraction of the sky, allowing for detection of about one burst per day with exceptional time resolution. The bursts themselves showed a wide variety of features, with very little in common between any two GRBs (besides, of course, the energy range in which they were observed). A small sample of BATSE triggers is provided in Figure 2.2. As is apparent from the figure, the tired saying that "If you’ve seen one, you’ve seen them all" hardly applies to GRBs. The more accurate, albeit tautological, statement would be "If you’ve seen one GRB, you’ve seen one GRB". GRBs are highly irregular pulses—or even sets of pulses—of X-rays and γ-rays. Their duration spans many orders of magnitude, from less than 0.01 seconds to more than 100 seconds. The number and shape of the pulses varies substantially. Some bursts are continuous over their duration, while
Figure 2.3 Distribution of recorded GRB lengths. The GRBs shown here are taken from the fourth BATSE catalog. The independent variable $T_{90}$ is roughly the time during which 90% of the GRB's total photons were detected, running from 5% to 95%. The two peaks at 0.2 seconds and $\sim 50$ seconds are plain, as is the trough occurring at $\sim 2$ seconds. Image reproduced with permission, from http://gammaray.nsstc.nasa.gov/batse/grb/duration/; figure originally published in Paciesas et al. (1999).

others have delays of seconds to minutes between bouts of activity. As well, the fluxes span several orders of magnitude.

What the GRBs of the BATSE catalog have in common is a non-thermal energy spectrum, well fitted by a broken power law (Band et al., 1993). Early results (Meegan et al., 1992) showed that the bursts were distributed isotropically across the sky, but that the source population was inhomogeneous (being biased against weak events).

Prior to the launch of CGRO, GRBs seemed to fall into two classes: “short” GRBs lasting about a second or less, and “long” GRBs lasting several seconds and more. The large number of bursts observed by the BATSE instrument illustrated the two-peaked distribution in unprecedented detail. Figure 2.3 shows that the short-long demarcation occurs at about two seconds in duration, with the peaks occurring around 0.2 and 50 seconds. The apparent third peak at 5 seconds in duration may be a statistical artifact, as it does not show up in a later catalog of GRBs recorded by a different instrument (von Kienlin et al., 2014).

The disparate distribution suggests that GRBs have multiple mechanisms of origination. The energy required to power a GRB at cosmic distances$^3$ strongly suggested that compact objects

$^3$Well outside the Solar System, but not necessarily outside the Milky Way. Not to be confused with “cosmological”,

26
like neutron stars or black holes were involved, but for lack of data theorists provided more than a hundred potential models. Just a few of the proposals for GRB sources are listed below, in no particular order:

- Neutron star mergers (e.g., Eichler et al., 1989; Narayan et al., 1992)
- “Starquakes” at the surface of neutron stars (Ellison and Kazanas, 1983; Blaes et al., 1989)
- Comet-neutron star collisions (Harwit and Salpeter, 1973)
- Collapse of massive stars (e.g., Woosley, 1993; Paczyński, 1998, though the latter was published after the debate was settled)
- Accretion-induced “novae” at the surface of neutron stars (Lamb et al., 1973; Blaes et al., 1990)
- Collisions of large chunks ($\sim 10^{15}$ g) of antimatter with normal stars (Sofia and van Horn, 1974)
- Shock breakout emission from distant supernovae (Colgate, 1968)
- Abrupt phase transition (“melting”) of the core of neutron stars (Bruk and Kugel, 1976)

Many ideas that involved neutron stars were strongly disfavored by the observed isotropy of GRB locations—given the many neutron stars within our galaxy, events with such origins should occur more frequently in the plane of the Milky Way than off of it. However, even the isotropic distribution in the sky was consistent with very nearby (i.e., distances closer than 100 parsecs, the scale height of the Milky Way’s disk at the Sun’s location), dim, sources for the bursts. The question of GRB origin remained open, waiting primarily for a detection of a burst at lower energies where localization would be more precise.

The next major step forward in understanding GRBs came with the launch of the BeppoSAX observatory in April 1996. Though primarily an X-ray mission, it was capable of nearly full-sky coverage in low-energy $\gamma$-rays (100-600 keV). The various instruments on the satellite combined to give angular positioning of a GRB to within 4 arcminutes. The mission proved its value on February 28, 1997, after the detection of GRB 970228. For the first time in almost three decades of GRB detections, a fading source of emission in both X-rays (Costa et al., 1997; Frontera et al., 1998) and visible light (van Paradijs et al., 1997)—an “afterglow”—was observed. The afterglow coincided with

---

4Even among theoretical models of events with minimal observational evidence, this idea was unusual. To their credit, the authors provided several means by which the model might be disproven, eliminating it from the pool of possibilities and making the remaining candidates more likely.
2.1. HISTORY OF GRBS

CHAPTER 2. GRBS AND AFTERGLOWS

a small galaxy of redshift $z = 0.695$, leading to the conclusion that the GRB was in fact at cosmological distances: more than 8 billion light-years from Earth.

Confirmation of the cosmological nature of GRBs settled one question, but it raised another, more troubling one. If the emission from GRBs was isotropic, the amount of energy released by a single event was staggering. When the distance to GRB 990123 was determined (Kulkarni et al., 1999), the isotropic energy release, $E_{\text{iso}}$, was more than $3 \times 10^{54}$ ergs—more than would be generated by the annihilation of a Solar mass of matter and antimatter. The GRB population spans more than seven orders of magnitude in $E_{\text{iso}}$, from $10^{48}$ to $10^{55}$ ergs, and at the bright end of the spectrum individual GRBs dwarf the emission of their host galaxies.\(^5\) Even at near perfect efficiency, the energy required to power the brightest bursts strained most models of GRB origin.

The best-equipped model to deal with the very high $E_{\text{iso}}$ requirements of GRBs claimed that the energy was not, in fact, released isotropically at all. Numerous authors had predicted that GRBs were highly relativistic collimated outflows—narrow jets of matter accelerated to nearly the speed of light—before the observation of GRB 970228’s afterglow (e.g., Paczyński, 1986; Mochkovitch et al., 1993). Confirmation that GRBs are, indeed, moving relativistically came with the detection of scintillation in radio afterglows. Scintillations (much like stars twinkling in the night sky) occur only when objects are sufficiently small as to appear almost point-like (Goodman, 1997). Once the scintillations of GRB 970508’s afterglow had faded, the size of the emitting surface could be calculated using the distance to the host galaxy ($z = 0.835$; Metzger et al., 1997), and the results supported expansion with an average Lorentz factor of 4 over the two week period (Frail et al., 1997). Later, superluminal motion was directly detected in the afterglow of GRB 030329 (Taylor et al., 2004), with radio observations spatially resolving the afterglow and allowing for determination of its size without resorting to scintillation.

One prediction (made by Rhoads, 1999) of the relativistic jet model is an achromatic (across all wavelengths) “jet break” in the afterglow light curve as the jet decelerates. The physical mechanism will be explained in greater detail in Sections 2.2 and 2.3, but observation of a sudden increase in the decay rate of afterglow emission would serve as a smoking gun for the collimated jet model. As mentioned in Rhoads (1999), the first afterglows observed showed no such signatures, but Harrison et al. (1999) reported just such a steepening in the afterglow of GRB 990510, proving that at least some of the GRB population emitted anisotropically.

There are two inescapable consequences of the anisotropic jet model for GRBs. The first is that since we only observe those GRBs whose jet is directed toward us, we do not observe the majority

\(^5\)The calculated luminosity for GRB 990123 was $3.3 \times 10^{16}$ $L_\odot$, where $L_\odot$ is the luminosity of the Sun. This is almost one million times brighter than a typical galaxy, and a thousand times brighter than the brightest sustained luminosities known as of the publication of Kulkarni et al. (1999).
of bursts. GRBs are (potentially much) more common in the Universe than initially expected. The second consequence is the resolution of the energy crisis. Since the radiation of a GRB is not emitted in all directions at once, the total amount of energy released drops precipitously. Fits to observed (or inferred) jet breaks suggest that the jets’ radii cluster around $2^{-3}$, extending out to about $10^\circ$ (Panaitescu and Kumar, 2001; Liang et al., 2008; Racusin et al., 2009). Such tightly beamed jets reduce the required energy by a factor of 100-1000 from the isotropic case. The true energy of a GRB, once this correction is taken into account, falls to the more reasonable range of $10^{49}$ to $10^{52}$ ergs (Frail et al., 2001; Cenko et al., 2010).

In spite of the images shown in Figure 2.2, BATSE and BeppoSAX observations established a more-or-less “standard” picture of a GRB and its environment (e.g., Panaitescu and Kumar, 2002). The earliest part of the afterglow sees the observed flux vary with time as $F(t) \propto t^{-1}$, and with frequency as $F(\nu) \propto \nu^{-0.9 \pm 0.5}$. Some afterglows show a jet break, after which the light curve steepens to $F(t) \propto t^{-2.2}$. Long GRBs tend to be surrounded by a uniform circumburst medium (CBM) out to $\sim 0.1$ pc, with a particle number density ranging from $0.1 - 100$ cm$^{-3}$. The highly relativistic outflow of the jet should therefore drive a very strong shock into the CBM, where shock acceleration of ambient particles (see the previous chapter) is all but certain to happen. The standard model for afterglow emission, developed well in advance of afterglow observations (for example, Paczyński and Rhoads, 1993; Mészáros and Rees, 1993a,b, 1997; see Dainotti et al., 2010 for more current work in this paradigm), assumes that emission is due to synchrotron radiation from highly energetic electrons gyrating in the magnetic field around the shock.

All of the above work represents significant progress on the nature of GRBs and afterglows, but it misses answering one very important question: what is the source of GRBs? Which of the many models suggested above is correct? As I mentioned previously, GRBs are typically divided into two classes based on their duration (Kouveliotou et al., 1993), with a cutoff at roughly 2 seconds. Somewhat unsurprisingly, then, there are two primary mechanisms to generate a GRB. Long GRBs are caused by the collapse and subsequent supernova explosion of massive stars, while short GRBs are due to the merger of compact objects.

The association between GRBs and supernovae was noted in Paczyński (1986), who commented that GRBs at cosmological distances would release roughly the same energy as a supernova. The collapsar model (Woosley, 1993; MacFadyen and Woosley, 1999) provided the physical relationship between the two objects. In the model, the core of a massive star collapses into a neutron star or black hole. Instead of a spherical supernova, however, the collapse launches a jet that pierces the star and gives rise to a GRB. This strongly suggests that the progenitor of GRBs should be a stripped-envelope star—one that has shed off its outer hydrogen layers, and perhaps also its helium—both to minimize the gravitational well and the layer of matter through which the jet must break. If a single
star caused both a GRB and a supernova, the supernova would be of type Ib or Ic.

The first major breakthrough in proving the collapsar model was GRB 980425/\(\text{SN} \ 1998\text{bw}\). As of April 25, 1998, only six GRBs had been localized by BeppoSAX, and only two of them had known distances (GRB 970508 at \(z = 0.83\), Metzger et al., 1997; and GRB 971214 at \(z = 3.42\), Odewahn et al., 1998). In followup observations of GRB 980425, Pian et al. (1998) discovered two X-ray sources within the \(\gamma\)-ray error circle. The day after the burst, the brightening supernova 1998bw was detected (Galama et al., 1998), within the \(\gamma\)-ray error circle but not within the error circle of either X-ray source. A search for the radio afterglow of GRB 980425 (Kulkarni et al., 1998) revealed an object coincident with the supernova. The broad lines of the supernova spectrum and bright radio emission indicated that the ejecta was moving with a velocity \(v \gtrsim 0.1c\). The supernova was localized to a galaxy with a measured redshift of just 0.0085, more than an order of magnitude lower than that of the previous closest burst. If the GRB was associated with the supernova, it was close enough that an optical afterglow should have been observable, but none was; in addition, the extreme closeness of the supernova made GRB 980425 extremely underpowered (\(E_{\text{iso}} \sim 10^{48}\) erg) compared to typical bursts with \(E_{\text{iso}} \gtrsim 10^{51}\) erg. The coincidence of the GRB, supernova, and radio detections was convincing, but the lack of a firm X-ray/optical afterglow detection—as well as the exceptionally low energy—prevented the connection from being conclusive.

The undeniable proof of the collapsar model came in GRB 030329/\(\text{SN} \ 2003\text{dh}\). The gamma-ray burst was detected by the HETE satellite (Vanderspek et al., 2004), and the redshift of \(z = 0.169\) (Greiner et al., 2003) meant an isotropic energy release roughly 10,000 times that of GRB 980425, firmly in line with other cosmological bursts. Also unlike GRB 980425, GRB 030329 had a bright optical afterglow (Price et al., 2003). In the days that followed, the afterglow started showing deviations from a pure power law expected from synchrotron emission. When the power law was subtracted off, a clear supernova spectrum remained: that of SN 2003dh (Hjorth et al., 2003; Matheson et al., 2003). The spectra of supernovae 1998bw and 2003dh were very similar at similar ages, confirming that SN 1998bw was, in fact, related to GRB 980425 and was the first GRB-supernova observed.

GRB-SNe remain rare, in spite of the frequency at which GRBs and afterglows are observed. As of the publication of Hjorth and Bloom (2012), only \(~25\) GRB-SNe were known to a reasonable degree of confidence. Five of them (GRB 980425/\(\text{SN} \ 1998\text{bw}\), GRB 030329/\(\text{SN} \ 2003\text{dh}\), GRB 031203/\(\text{SN} \ 2003\text{lw}\), GRB 060218/\(\text{SN} \ 2006\text{aj}\), and GRB 100316D/\(\text{SN} \ 2010\text{bh}\)) had spectroscopic confirmation, where the light spectrum of the supernova was cleanly identified. The remainder had photometric confirmation—a bump in the light curve at the right frequencies and times, but not clear enough to allow full spectra to be taken. Of the five "ironclad" GRB-SNe, only GRB 030329/\(\text{SN} \ 2003\text{dh}\) had both

---

\(6\) This is hardly a given, as the process of generating a tightly beamed jet might be incompatible with also producing a spherical supernova.
an afterglow and a supernova; the other four had SNe that were observed to be roughly simultaneous with the GRB and within the various instruments’ error circles on the sky. The remaining four ironclad GRBs were also very underluminous compared to typical values, raising the possibility of a different process at work. Since Hjorth and Bloom (2012), however, we have had GRB 130427A/SN 2013cq. GRB 130427A was so intense it saturated some of the detectors on board the Fermi satellite (Ackermann et al., 2014, supplementary materials). At $z = 0.34$ (Levan et al., 2013; Flores et al., 2013), GRB 130427A was substantially more distant than any of the other spectroscopically confirmed GRB-SNe, but still close enough that a supernova could be observed if there were one. And, indeed, 2013cq was detected with the optical afterglow about two weeks after the GRB (Xu et al., 2013). The similarities between SN 2013cq and previous GRB-SNe showed that GRBs with cosmological distances and energies—and not only nearby and underluminous GRBs—could be associated with supernovae.

The success of the collapsar model for long GRBs still left short GRBs in need of explanation. That issue was resolved in 2005, with the detection of a trio of short GRBs within a three-month period. GRB 050509B was the first short GRB to have a firm localization (Hjorth et al., 2005a), to an elliptical galaxy at $z = 0.225$ (Bloom et al., 2006); though massive stars (> 8 times the mass of the Sun, i.e., massive enough to die in a core-collapse supernova) are rare even in spiral galaxies, they are rarer still in the gas-poor environments of elliptical galaxies. Not long after, GRB 050709 was identified with an irregular dwarf galaxy at redshift $z = 0.16$ (Hjorth et al., 2005b; Fox et al., 2005). In both cases, no supernova was detected to extremely strict limits on luminosity. The third in the set, GRB 050724, also occurred in an elliptical galaxy, and in particular one with a stellar population more than 1 Gyr old (Berger et al., 2005; Barthelmy et al., 2005c); the lifetime of a star massive enough to die in a core-collapse supernova is less than 100 Myr, significantly shorter than the stellar population inferred for GRB 050724. Taken together, the three GRBs all but rule out the collapsar model for short bursts and provide convincing, if not conclusive, evidence for the binary merger model (truly conclusive evidence will arrive in the form of a gravitational wave signature contemporaneous with a short GRB; see, e.g., Kochanek and Piran, 1993; Bartos et al., 2013).

The long/short collapsar/merger distinction makes for a simple picture, but like many simple pictures the division is not as clean cut as it first appears. Two classes of objects muddy the waters: SN-less long GRBs, and long GRBs at extreme redshifts. Both GRB 060505 and GRB 060614 were long GRBs that were localized to star-forming galaxies that should have young populations (Fynbo et al., 2006, and references therein). Despite the proximity of the host galaxies ($z = 0.089$ for GRB 060505, and $z = 0.033$ for GRB 060614) no supernovae were detected to extremely strict limits. On a different

---

7Specifically, more than 100 times fainter than the (at the time) faintest known type Ia and type Ic SN, and still fainter than the faintest known supernova of any type.
2.2. AFTERGLOW THEORY AND METHODS

2.2 Afterglow theory and methods

Significant effort has gone into deriving analytical forms for afterglow emission, as this allows for rapid prediction of many important parameters of the GRB jet and its environment (e.g. whether the burst occurred in a wind-like or constant-density medium, the bulk Lorentz factor $\Gamma_0$ of the shock$^8$, the opening angle $\theta_{\text{jet}}$ of the jet, the fraction of energy $\varepsilon_e$ deposited into electrons, or the fraction of energy $\varepsilon_B$ stored in the magnetic field). In this section I will cover some of the analytical models used in the study of afterglows, as well as their numerical extensions where appropriate.

In the late 1940s, multiple scientists (Sedov, 1959, though it was originally published in 1946; von Neumann (1947), reprinted in Bethe, 1947; and Taylor, 1950) arrived independently at analytical formulations for what happens after a very large amount of energy is deposited in a very small volume. Originally applied to atomic bomb explosions and later repurposed to supernova remnants, the Sedov-Taylor-von Neumann solution lets one calculate the position of a shock—as well as related quantities like pressure, density, and fluid velocity—at an arbitrary time $t$. The solution depends only on the energy and mass involved, as well as the nature of the surrounding medium. If, however, the amount of energy $E$ that is deposited approaches or exceeds the rest mass-energy $Mc^2$ of the matter being energized, the matter will become relativistic. The Sedov-Taylor-von Neumann solution does not apply in this regime, and a different approach must be used.

The solution for the motion of a relativistic shock was presented in Blandford and McKee

---

$^8$In this chapter I defer to the notation system used within the GRB community. The capital letter $\Gamma$ is used to represent the bulk Lorentz factor of the jet, while lowercase $\gamma$ refers to the motion of individual particles within the plasma.
(1976), and will be detailed further in Chapter 7. The authors rearranged the relativistic shock jump conditions (see Section 3.6) to arrive at a self-similar solution (i.e. possessing the same shape at all times, up to a time-dependent scale factor) for both the shock and the material behind it. The two primary assumptions made are that the motion of the shock is fully relativistic \((\Gamma \gg 1)\) and that the pressure tensor of the shocked fluid is isotropic.\(^9\) The Blandford-McKee solution matches very well to numerical results while \(\Gamma_0 > 5\), and is a reasonable approximation into the trans-relativistic regime, \(\Gamma_0 \approx 2\) (Kobayashi et al., 1999).

### 2.2.1 Hydrodynamical scaling relations

In the impulsive limit (energy deposition happens rapidly, rather than over an extended period of time) of the Blandford-McKee solution, and assuming (1) a constant-density ambient medium and (2) negligible losses due to either radiation or particle escape, the energy contained in the shock structure is a constant:

\[
E(t) = \frac{4\pi}{3} R(t)^3 \cdot n_0 \bar{m} c^2 \cdot \Gamma(t)^2 = \text{const},
\]

i.e. that the energy of the blast is evenly distributed among the particles of density \(n_0\) and average mass \(\bar{m}\). One factor of \(\Gamma(t)\) is due to the energization of particles as they are initially swept up by the shock, and the second results from Lorentz-transforming their energy back into the ISM frame. Dividing through by constants, one arrives at the first scaling relation of many: \(\Gamma^2 R^3 = \text{const}\), or

\[
\Gamma \propto R^{-3/2}.
\]

From special relativistic considerations, the observer time between events (such as photon arrivals) can be significantly shorter than the time between events in the shock frame: \(t_{\text{obs}} \approx R/(2\Gamma^2 c)\). Inserting Equation 2.2 yields the relation between time, speed, and/or radius of the shock:

\[
t_{\text{obs}} \propto R \Gamma^{-2} \propto R^4, \quad \text{or} \quad R \propto t_{\text{obs}}^{1/4}.
\]

One can relax the assumption of a constant-density ambient medium around the explosion, replacing it by a power law of slope \(k\): \(n(R) = n_0 (R/R_0)^{-k}\). Note that \(k = 0\) recovers the constant-density solution, and \(k = 2\) corresponds to a wind-like environment around the blast. Repeating the same considerations as above, the quantity of material swept up by the shock becomes

\[
\int n_0 (R/R_0)^{-k} \cdot 4\pi R^2 dR,
\]

and \(R^{3-k} \Gamma^2\) is the new conserved quantity. The generalized scaling relations

\(^9\)This is equivalent to assuming that any ordered magnetic field is very weak compared to the kinetic energy of the particles it contains, since a strong field constrains particle motion and leads to anisotropies in the pressure tensor.
between $t_{\text{obs}}$, $\Gamma$, and $R$ are

$$t_{\text{obs}} \propto R \Gamma^{-2} \propto R^{4-k} \propto \Gamma^{\frac{8-2k}{k-3}},$$

(2.4)

$$\Gamma \propto t_{\text{obs}}^{-1/2} R^{1/2} \propto R^{k-3} \propto t_{\text{obs}}^{\frac{k-3}{2}},$$

(2.5)

and

$$R \propto t_{\text{obs}}^{-1/2} \propto \Gamma^{\frac{2}{k-3}} \propto t_{\text{obs}}^{-\frac{1}{k-3}}.$$  

(2.6)

The relations for $k = 2$, $t_{\text{obs}} \propto R^2 \propto \Gamma^{-4}$, are used regularly (Dai and Lu, 1998; Mészáros et al., 1998; Chevalier and Li, 1999, 2000). If one is willing to carry around extremely complicated exponents—and the community of afterglow theorists certainly is—there is no need to specialize to a medium until one attempts to fit observations.

In addition to the relations between $t_{\text{obs}}$, $\Gamma$, and $R$, one can write a scaling relation for the strength of the downstream magnetic field in the comoving (plasma) frame. Assuming that the energy density associated with the magnetic field is a fixed fraction $\epsilon_B$ of the energy density of the shocked plasma, it is straightforward to derive the following relation

$$B' \propto n^{1/2} \varepsilon_B^{1/2} \Gamma \left(\Gamma - 1\right)^{1/2}.$$

(2.7)

In the ultra-relativistic limit, $\Gamma \gg 1$ and so $B' \propto \Gamma$. In the nonrelativistic limit, $\Gamma - 1 \ll 1$ and $B' \propto \Gamma^{1/2} \propto \beta$.

Scaling arguments like those presented above lie at the heart of analytical approaches to afterglows. Until such time as GRBs and afterglows are routinely detected as other than pointlike sources, however, these hydrodynamical relations offer only limited insight regarding observations.

### 2.2.2 Photon spectrum scaling relations

The same procedure as above can be used to quantify photon production at virtually any point in the afterglow, for virtually any scenario one can envision: presence/absence of a reverse shock within the external shock, relative angle between the observer and the jet axis, optical depth of the emitting region, nature of the circumburst medium. (See Gao et al., 2013 for a “complete review”, in all the gory detail that implies.)

Virtually all such scaling relations make the same two assumptions. The first is that the extended emission, after the GRB has faded, is synchrotron radiation due to electrons. The second is that these electrons form a power law in energy, such that $dN/dE \propto E^{-p}$ from some minimum electron energy up to some maximum determined by the properties of the shock or its environment. Additionally, the canonical afterglow scaling arguments rely on three key frequencies in the synchrotron spectrum:
2.2. AFTERGLOW THEORY AND METHODS

- $\nu_m$: The peak frequency of synchrotron emission for electrons at the bottom of the power law, with minimum energy $\gamma_m$. This is typically also associated with the peak of the overall spectrum, since these electrons are more numerous than those at any higher energy.

- $\nu_c$: The “cooling frequency” (and associated cooling energy $\gamma_c$) above which electrons should be losing energy by radiative processes faster than they can gain it through whatever acceleration process is occurring.

- $\nu_a$: The self-absorption frequency. Discussed further in Section 6.2.2, any source of synchrotron radiation has a maximum intensity at low frequencies; if the intensity is greater than this limit the plasma is optically thick to photons below that frequency and the observed spectral index breaks at $\nu_a$.

The ordering of $\nu_c$ and $\nu_m$ differentiates two broad regimes within the afterglow. If $\nu_c < \nu_m$, then all electrons in the spectrum are rapidly losing energy post-acceleration, and the afterglow is considered to be in the “fast cooling” regime. If $\nu_m < \nu_c$, then at least some electrons are cooling on timescales longer than the dynamical timescale of the afterglow—the “slow cooling” regime. One typically expects fast cooling only very early in the afterglow and slow cooling thereafter, with the transition happening around (Sari et al., 1998)

$$t_0 = 210\epsilon_B^2 \epsilon_e^2 E_{52} n_1 \text{ days.}$$  

In this equation, and elsewhere in this chapter, the notation $Q = Q_x 10^x$ is used. The parameters $\epsilon_B$ and $\epsilon_e$ refer to the fraction of the overall energy density that is stored in the magnetic field and in electrons, respectively. Both of these are typically much smaller than unity, so the product of their squares is very small. The transition from fast to slow cooling may happen on the order of hours into the afterglow. Note, however, the major assumption made in Equation 2.8: the quantities $\epsilon_B$ and $\epsilon_e$ are treated as constant throughout the afterglow. This assumption makes the analytical treatment of afterglows much simpler (both mathematically and in terms of the number of necessary parameters for any given model). It is also directly testable against the results of any evolutionary model of afterglow emission—specifically, one such as will be presented in Chapter 8.

Before continuing further, it must be reiterated that virtually all results that follow assume that the electron distribution forms a single power law over the entire energy range. It ignores any thermal populations of electrons, even if such a population would be relativistic and radiating at detectable

---

10 Recall that, while uppercase $\Gamma$ is used to refer to bulk fluid velocities, lowercase $\gamma$ is used when describing individual particles.
frequencies. It also neglects the possibility that the energy distribution of electrons is poorly fitted by a power law (see Chapter 5, and Sironi and Spitkovsky, 2011).

So, given an electron distribution with spectral index $p$ and a set of break frequencies $\nu_a$, $\nu_m$, and $\nu_c$, one can derive a canonical afterglow spectrum based on the behavior of electrons in each range of energies. The spectrum assumes that the entirety of emission is due to synchrotron radiation, and that the self-absorption frequency is lower than both other break frequencies. (The conditions required for $\nu_a$ to exceed either $\nu_m$ or $\nu_c$ are extreme, even in the context of GRB afterglows.)

The resultant flux at some distance from the shock, ignoring special relativistic effects like beaming, boosting, time dilation, and the surface of equal arrival time, is (Sari et al., 1998)

$$F_{\nu,\text{fast}} = F_{\nu,\text{max}} \begin{cases} \left( \frac{\nu}{\nu_a} \right)^{1/3} \left( \frac{\nu}{\nu_c} \right)^2 & \nu < \nu_a \\ \left( \frac{\nu}{\nu_a} \right)^{1/3} \left( \frac{\nu}{\nu_c} \right)^{-1/2} & \nu_a < \nu < \nu_c \\ \left( \frac{\nu}{\nu_m} \right)^{-p/2} & \nu_c < \nu < \nu_m \end{cases} \ (2.9)$$

in the fast cooling regime, and

$$F_{\nu,\text{slow}} = F_{\nu,\text{max}} \begin{cases} \left( \frac{\nu_a}{\nu_m} \right)^{1/3} \left( \frac{\nu}{\nu_a} \right)^2 & \nu < \nu_a \\ \left( \frac{\nu}{\nu_m} \right)^{1/3} \left( \frac{\nu}{\nu_m} \right)^{(p-1)/2} & \nu_a < \nu < \nu_m \\ \left( \frac{\nu}{\nu_m} \right)^{(p-1)/2} \left( \frac{\nu}{\nu_c} \right)^{-p/2} & \nu_m < \nu < \nu_c \end{cases} \ (2.10)$$

in the slow cooling regime. The frequency at which $F_{\nu,\text{max}}$ peaks is given by $F_{\nu,\text{max}} \approx (1 + z)N_{e,\text{tot}} P_{\nu,\text{max}}' \Gamma \left( 4\pi D_L^2 \right)$. This formula includes the cosmological redshift $z$ of the GRB, the total number of electrons in the particle spectrum $N_{e,\text{tot}}$, the power per unit frequency (in the comoving frame) for a single electron at the peak of the particle spectrum, the current Lorentz factor of the shock, and the luminosity distance to the GRB (the luminosity distance $D_L$ is the total distance traveled by photons on their path to Earth, incorporating the expansion of the Universe that occurred during the travel). In light of all of the previous, the scaling relation for $F_{\nu,\text{max}}$ in a constant-density ISM is (Yost et al., 2003)

$$F_{\nu,\text{max}} = 1.6 \text{mJy}(1 + z) \epsilon_{B,2}^{1/2} E_{52} n_0^{1/2} D_L^{-2} \ (2.11)$$

\[11\] Inverse Compton emission—if even included—is used only for determining the cooling frequency, and cosmic-ray nuclei contribute negligibly to the spectrum.

\[12\] The location of the peak depends on whether the system is in the fast or slow cooling regime: $\gamma_{\text{peak}} = \min(\gamma_m, \gamma_c)$. 

36
and the locations of the three spectral breaks occur at

\[
\nu_a = 4.2 \times 10^8 \text{Hz}(1+z)^{-1} \left[ \frac{(p+2)(p+1)}{3p+2} \right]^{3/5} \epsilon_{\text{e}}^{1/5} g(p) \epsilon_{B,-2}^{1/5} n_0^{3/5} \nu_0^{1/5} \epsilon_{1/5} B_{52}^{-2} t_{\text{obs,day}}^{-3/2} \ (2.12)
\]

\[
\nu_m = 3.3 \times 10^{14} \text{Hz}(1+z)^{-1/2} \epsilon_{B,-2}^{1/2} [\epsilon_{\text{e}} g(p)]^{-1/2} E_{52}^{1/2} t_{\text{obs,day}}^{-3/2} \ (2.13)
\]

\[
\nu_c = 6.3 \times 10^{15} \text{Hz}(1+z)^{-1/2} \epsilon_{B,-2}^{-3/2} E_{52}^{-1/2} n_0^{-1} t_{\text{obs,day}}^{-1/2} \ (2.14)
\]

where \( g(p) = (p-2)/(p-1) \) if \( p > 2 \) and \( g(p) = 1/\ln(\gamma_{\text{max}}/\gamma_m) \) if \( p = 2 \). Though the consequences of particle spectra harder than \( p^{-2} \) have been explored (Dai and Cheng, 2001; Bhattacharya, 2001; Resmi and Bhattacharya, 2008), they will not be relevant to the work presented here.

The strength of the post-shock magnetic field, which is responsible for the synchrotron radiation detected in the afterglow, is a very difficult parameter to constrain. Kumar (2000) and Panaitescu and Kumar (2000) considered the relationship between the magnetic field and the highest-energy part of the synchrotron spectrum—above \( \nu_m, \nu_c \). Those authors found that the flux density scales as \( F_{\nu} \propto \epsilon_B^{-p-2/4} \). As relativistic shocks are expected to produce momentum distributions with spectral indices between 2 and 2.4 (see Section 1.3), the observational signature of the magnetic field is quite weak.

### 2.2.3 Hydrodynamics of the jet

The Blandford-McKee solution discussed in Section 2.2.1 follows only the external shock as it propagates into the circumburst medium (CBM). However, as is familiar to the astrophysical hydrodynamics community, the situation is frequently far more complicated than a single shock (for an extremely limited sampling, see Chevalier, 1982; Dwarkadas and Chevalier, 1998; Blondin and Ellison, 2001; Lee et al., 2014).

As the jet of GRB ejecta expands outward, it will drive a forward shock into the CBM. The expansion rate of the forward shock is \( \Gamma_0 \) in the Blandford-McKee solution.\(^{13}\) Separating the forward-shocked CBM and the GRB ejecta will be a contact discontinuity. The contact discontinuity is Rayleigh-Taylor unstable in multiple dimensions (Warren and Blondin, 2013; Duffell and MacFadyen, 2014), but for simplicity analytical models—and most numerical ones—assume either spherical symmetry or that the contact discontinuity has the same shape as the forward shock.

Further in from the contact discontinuity, a reverse shock may propagate down the jet into the ejecta (Mészáros and Rees, 1993a; Katz, 1994; Sari and Piran, 1995), leading to a 4-shell model of the

---

\(^{13}\) The GRB community variously uses this quantity to refer either to the expansion rate of the forward shock, to the Lorentz factor of the freshly-shocked plasma, or occasionally even to the Lorentz factor of the ejecta jet; so each derivation must be approached with that uncertainty in mind.
2.2. AFTERGLOW THEORY AND METHODS

CHAPTER 2. GRBS AND AFTERGLOWS

Figure 2.4 A cartoon of the hydrodynamical structure of a relativistic jet. From left to right are regions 1 (unshocked CBM), 2 (shocked CBM), 3 (shocked ejecta) and 4 (unshocked ejecta). Each region is separated from its neighbors by one of the three fluid discontinuities: the forward shock (FS), the contact discontinuity (CD), and the reverse shock (RS). Within each region the rest-frame density (n'_i) and pressure (p'_i) are assumed to be constant. Additionally, each region has a speed Γ_i with which it is expanding relative to the at-rest CBM (i.e., Γ_1 = 1 identically).

jet-CBM profile. This model is illustrated in Figure 2.4. Radiation from reverse shock-heated ejecta was predicted and described by, e.g., Mészáros and Rees (1993a, 1997), and Sari and Piran (1999b). During the prompt emission of GRB 990123, a bright optical flash was detected (Akerlof and McKay, 1999), which was interpreted as emission from a reverse shock crossing the ejecta (Sari and Piran, 1999a; Mészáros and Rees, 1999).

Analytical models of GRB reverse shocks (e.g. Sari and Piran, 1995) typically assume pressure balance across the contact discontinuity, i.e. p'_2 = p'_3 in Figure 2.4. (But see Beloborodov and Uhm 2006 and Uhm et al. 2012 for an alternative semi-analytical model that relies on momentum and energy conservation rather than pressure balance.) In that scenario, the relationship between n'_4/n'_1 and Γ_{41} (the Lorentz factor of region 4 as seen from region 1) determines whether the reverse shock is relativistic (n'_4/n'_1 ≪ Γ_{41}^2) or nonrelativistic (n'_4/n'_1 ≫ Γ_{41}^2).

Regardless of the speed of the reverse shock relative to the ejecta, the sound crossing time depends on the width of the ejecta (which is itself presumably related to the lifetime of the central engine that caused the GRB). Assuming a constant speed for the ejecta, Sari and Piran (1995) distinguish “thick” and “thin” shell models for reverse shock emission by comparing the Lorentz
factor $\gamma_{41}$ of the unshocked ejecta, the Sedov length $l_{\text{Sed}}$ of the blast\(^{14}\), and the observer-frame width $\Delta$ of the (pre-RS) ejecta shell. Specifically, $\xi \equiv (l_{\text{Sed}}/\Delta)^{1/2}\gamma_{41}^{-8/3} < 1$ denotes a dynamically thick shell, with other values referring to thin shells. If ejecta is stratified into sub-shells of different speed, an extremely long-lived reverse shock is possible, possibly including energy injection into the contact discontinuity and forward shock (e.g. Rees and Mészáros, 1998; Genet et al., 2007).

If no reverse shock forms, then all of the preceding discussion of its properties is irrelevant. The work of Zhang and Kobayashi (2005) extended previous (purely hydrodynamical) studies of reverse shock formation into a magneto-hydrodynamical formulation. The authors found that the magnetization parameter $\sigma = B^2/(4\pi n_0'\overline{m}c^2)$ of the unshocked ejecta plays a critical role in the formation of reverse shocks. Optical flux from the reverse shock increases rapidly as a function of $\sigma$ while $\sigma \ll 1$, peaking around magnetizations of unity. When $\sigma > 1$, only relativistic reverse shocks are possible. If the ejecta is extremely highly magnetized, with $\sigma > 100$, the large Alfvén speed in the ejecta prevents the formation of a reverse shock of any speed. Although Zhang and Kobayashi (2005) calculated the expected emission from their models, they used (by their own admission) a simplistic model for DSA and the resultant cosmic-ray spectrum.

Observations of afterglows, especially in the Swift era, offer numerous examples that show no evidence of reverse shock emission (Roming et al., 2006; Molinari et al., 2007; Rykoff et al., 2009). The GRB afterglows that do show evidence for a reverse shock fall into two broad groups:

- The first group is composed of afterglows with weak forward shock emission, so that the afterglow is dominated by the contribution from the reverse shock. This would happen near the $0.1 < \sigma < 1$ peak of Zhang and Kobayashi (2005). GRB 990123 was the first example of this class, though similar GRBs have been observed since (Fox et al., 2003; Kumar and Panaitescu, 2003; Gomboc et al., 2008).

- In the second group, the optical light curve shows two distinct peaks. The first, earlier, peak is interpreted as due to the reverse shock. The second peak is caused by the synchrotron peak of the forward-shocked CBM passing through visible wavelengths. (See, e.g., Kobayashi and Zhang, 2003; Shao and Dai, 2005)

2.2.4 Consequences of a collimated jet

As discussed earlier in this chapter, GRBs are all but certain to be associated with highly relativistic collimated jets. This has two consequences for the afterglow, both of which arise due to the inherent anisotropy of the emitting surface: the edge effect, and lateral expansion. These effects combine to

\[^{14}\text{The Sedov length is another borrowed quantity from supernova theory, and is the ratio of the blast energy to the rest mass-energy density of the CBM, } l_{\text{Sed}} = [E/(n_0'\overline{m}c^2)]^{1/3}.\]
cause an achromatic (i.e. simultaneous for all frequencies) steepening of the afterglow light curve, which is referred to as a “jet break” in the literature. Additional complications arise when considering possible structure at the forward shock or the geometrical relationship between the jet and the observer.

The edge effect (Sari et al., 1999; Mészáros and Rees, 1999; Panaitescu and Mészáros, 1999; Rhoads, 1999) is due to special relativity. A moving source beams emission into a cone of half-opening angle \( \theta_h \approx 1/\Gamma \), where \( \Gamma \) is the Lorentz factor of the source’s motion relative to the observer. Since the jet has a half-opening angle \( \theta_j \), if \( \Gamma > 1/\theta_j \) observers do not see emission from the entire jet, and the emission seen is indistinguishable from that produced by an isotropic source. But \( \Gamma \) drops with increasing time, and eventually falls below \( 1/\theta_j \). At that point the edge of the jet becomes visible and observers see a drop in flux since there is no emission beyond the edge to continue the appearance of isotropy.

Only the total emission is affected by the edge effect. The previously described scaling arguments for \( \nu_a, \nu_m, \nu_c \), and \( \Gamma \) are unchanged since the edge effect is unrelated to the fluid dynamics of the jet-CBM interaction. After the edge becomes visible the emission should decay as \( \Gamma^2 \propto t_{\text{obs}}^{3/4} \) (assuming a constant-density CBM).

The second consequence of the collimated jet is late-time lateral expansion (Rhoads, 1999; Sari et al., 1999; Granot and Piran, 2012). Early in the afterglow, the jet expands with an essentially constant half-opening angle \( \theta_j \). This due to the fact that the sound crossing time of the jet is comparable to, or slightly larger than, the dynamical age. As the jet slows, the two sides of the jet become hydrodynamically connected, and the pressure forces a lateral expansion. Conveniently, this happens at approximately the same time as the edge effect occurs: once \( \Gamma \approx 1/\theta_j \). Conservation of energy over a larger solid angle means less energy available to shock heat and accelerate particles. Unlike the edge effect, lateral expansion will impact the dynamics of the blast wave. New scaling relations may be derived for this stage (given in the above references):

\[
\begin{align*}
\Gamma & \propto t_{\text{obs}}^{-1/2} \\
\nu_m & \propto t_{\text{obs}}^{-2} \\
\nu_c & \propto t_{\text{obs}}^0 \quad \text{i.e., constant} \\
F_{\nu,\text{max}} & \propto t_{\text{obs}}^{-1} 
\end{align*}
\]
and

\[ F_\nu \propto \begin{cases} 
\nu^{1/3} t_{\text{obs}}^{-1/3} & \nu < \nu_m \\
\nu^{-(p-1)/2} t_{\text{obs}}^{p} & \nu_m < \nu < \nu_c \\
\nu^{-p/2} t_{\text{obs}}^{-p} & \nu < \nu_c
\end{cases}, \tag{2.19} \]

where \( p \) is the power law index of the electrons’ energy distribution. Note, however, that multidimensional simulations suggest that lateral expansion is not so important until the trans-relativistic regime of \( \Gamma \sim 2 - 5 \) (Kumar and Granot, 2003; De Colle et al., 2012; van Eerten and MacFadyen, 2012), as a typical jet’s transverse size grows as the logarithm of time—or slower still (this is in stark opposition to the exponential expansion predicted by Rhoads 1999; the simplifying assumption of a completely homogenous jet surface is not supported by multidimensional numerical simulations).

A further wrinkle on the standard picture of the jet is the possibility of structure on the surface of the forward shock. By far the simplest model for jet surfaces is the “top hat” model, which is uniform over the entire solid angle subtended by the jet. Many alternatives have been advanced, however. The next simplest model is a two-component jet, in which an ultrarelativistic core, with \( \Gamma \sim 100 \), is surrounded by a less relativistic sheath with \( \Gamma \sim 10 \) (Ramirez-Ruiz et al., 2002; Zhang et al., 2004; Peng et al., 2005). Several GRBs have behaved in ways consistent with a two-component jet, including GRB 030329 (Berger et al., 2003). Other options include jets whose Lorentz factor varies as a power law with angle relative to the jet axis (Zhang and Mészáros, 2002b) or as a Gaussian centered on the jet axis (Kumar and Granot, 2003). It is also possible that the jet’s surface is “patchy”, with emission that varies stochastically in position and/or time (e.g., the “subjet” model proposed in Yamazaki et al., 2004). Such a jet might, for example, be due to non-uniform magnetic field turbulence (Lyutikov and Blandford, 2002).

The jets of GRBs are beamed into a small solid angle, but probability alone dictates that most GRBs should be observed at least a little bit off-axis. This will have definite observational consequences: for example the jet break associated with the edge effect will appear different if the afterglow is observed at a non-zero angle relative to the axis, as one edge will become visible sooner than the other. Off-axis effects have been studied by, e.g., Rossi et al. (2002) and Ryan et al. (2015). One implication of the models is that observers sufficiently far from the jet axis may miss the prompt (\( \gamma \)-ray) emission entirely; instead they might view the X-ray or optical emission at later times and lower Lorentz factors as an “orphan” afterglow. The faint nature—and lack of an obvious \( \gamma \)-ray trigger—makes identifying such an event difficult, and no transients have been established as promising candidates.\(^\text{16}\)

\(^{15}\)This GRB is additionally famous for being the “smoking gun” of the GRB-supernova connection, as it was definitively associated with SN2003dh once the afterglow had faded and the supernova risen.

\(^{16}\)See Cenko et al. (2013) for a discussion of the event PTF11agg, an optical flare with radio scintillation and a power-
In the final stage of the afterglow, the forward shock has slowed to non-relativistic speeds. During this “deep Newtonian” phase (Huang and Cheng, 2003) the jet transitions from the Blandford-McKee solution toward the Sedov-Taylor solution. The transition is typically considered to occur once the mass energy of swept-up CBM is equal to the initial kinetic energy of the jetted ejecta. This leads to the canonical Sedov-Taylor radius,

\[ R_{ST} \approx \left( \frac{3E}{4 \pi n_0 m c^2} \right)^{1/3} = 1.2 \times 10^{18} E_5^{1/3} n_0^{-1/3} \text{ cm} \]  

(2.20)

for a proton-only, constant-density CBM. The scaling relations for the Sedov-Taylor solution are \( v \propto t_{\text{obs}}^{-3/5} \) and \( R \propto t_{\text{obs}}^{2/5} \), leading Dai and Lu (1999) to derive the following relationships:

- \( B' \propto v \propto t_{\text{obs}}^{-3/5} \)  
- \( \gamma_m \propto \nu^2 \propto t_{\text{obs}}^{-6/5} \)  
- \( \nu_m \propto \gamma_m^2 B' \propto t_{\text{obs}}^{-3} \)  
- \( \nu_c \propto B't^{-3} \propto t_{\text{obs}}^{-1/5} \)  
- \( F_{\nu,\text{max}} \propto N P_{\text{max}}^{\nu} \propto R^3 B' \propto t_{\text{obs}}^{3/5} \)  

(2.21)  
(2.22)  
(2.23)  
(2.24)  
(2.25)

and so

\[ F_{\nu} \propto \begin{cases} \nu^{-(p-1)/2} t_{\text{obs}}^{-(21-15p)/10} & \nu_m < \nu < \nu_c \\ \nu^{-p/2} t_{\text{obs}}^{-(4-3p)/2} & \nu < \nu_c \end{cases} \]  

(2.26)

When the spectral index of the electron distribution \( p \) takes the canonical value for relativistic shocks (or slightly inefficient nonrelativistic shocks), \( p \approx 2.2 \), the temporal decay indices are \(-1.2\) and \(-1.3\). This is steeper than the dependence of \( F_{\nu} \) on time in the Blandford-McKee solution (\( t_{\text{obs}}^{-1.15} \)), but shallower than the post jet-break case (\( t_{\text{obs}}^{-2.2} \)). Thus the softening/hardening of the afterglow light curve could discriminate whether the deep Newtonian phase begins before or after the shock has transitioned to a more spherical than collimated shape (Livio and Waxman, 2000). However, for energy distributions that go as \( p^{-2} \), this behavior disappears, and may even reverse for harder spectra still.

Finally, though the counter-jet emission is initially strongly beamed away from Earth, at late times it may become visible. This very late-time effect was explored in, e.g., Granot and Loeb (2003) and Zhang and MacFadyen (2009).
2.3 Observations of afterglows

As mentioned previously, the existence of GRB afterglows was predicted in Paczyński and Rhoads (1993) and Katz (1994), and more thoroughly expanded upon in Mészáros and Rees (1997). Definitive proof of afterglow existence came with the detection of X-ray and optical emission after GRB 970228, with radio detection of GRB 970508 following soon thereafter.

Data on afterglows was patchy in the pre-

Swift

eta (launched on 20 November 2004; Gehrels et al., 2004). Afterglow observations relied on serendipitous detections, as the repointing time for telescopes was typically some number of hours. The Swift mission carries instruments for γ-ray, X-ray, ultraviolet and optical data collection—allowing for observations of afterglows across large parts of the spectrum—and has a slew time on the order of a minute. Unsurprisingly, understanding of afterglows has blossomed in the time the mission has been operational.

In the rest of this section I will describe the general properties of afterglows as understood prior to the launch of the Swift mission. I will then consider the current picture of afterglows as supported by roughly a decade of Swift data.

2.3.1 Pre-Swift era

At optical wavelengths, afterglows do indeed decay as a power law $F_\nu \propto t^{-\alpha}$, with $\alpha \sim 1$ (Wijers et al., 1997; Harrison et al., 1999), roughly in line with the predictions of the analytical external shock model described in the previous section. After about a day, the optical afterglow of GRB 990510 steepened to a $\alpha \sim -2$ power law, which was interpreted (Harrison et al., 1999) as evidence for a jet break.

In radio, afterglows rise rather than decay initially. Higher energies reach their maximum flux earlier than do lower energies, starting at around 10 days post-GRB (Frail et al., 2000). The slower rise at lower frequencies is seen as evidence of synchrotron self-absorption suppressing the emission below $\nu_a$. The decay in frequency of the radio peak (presumably occurring at the synchrotron peak $\nu_m$) is related to decay of the minimum electron energy in the emitting power law, which accompanies the deceleration of the blast wave.

At any particular time $t_{\text{obs}}$, the radio-to-X-ray spectrum is well-fitted by a broken power law (e.g., Wijers and Galama, 1999), consistent with the scaling relations of Section 2.2. However, just as no two GRBs are alike, the observed afterglows showed a good deal of variety in time scales, decay indices, spectral breaks, and fluxes. The available observations suggested that bursts and

---

17 Specifically, the X-Ray Telescope (XRT), (Burrows et al., 2005), the Ultraviolet/Optical Telescope (UVOT) (Roming et al., 2005), and the BAT (Burst Alert Telescope) (Barthelmy et al., 2005a).
Their afterglows span a wide parameter space in, e.g., the comoving magnetic field $B'$, the fraction $\varepsilon_e$ of blast energy in electrons, or the spectral index $p$ of the electrons’ distribution function.

Furthermore, the afterglows with the best observations (like those of GRBs 020114 and 030329) show features beyond the scope of the simple power law model described above. Smooth bumps of duration $\Delta t_{\text{obs}} \sim t_{\text{obs}}$ have been interpreted as the blast wave encountering density fluctuations in the CBM (Lazzati et al., 2002; Nakar and Granot, 2007). Features with a shorter duration may signify late-time injection of energy from a still-active central engine (Katz et al., 1998; Zhang and Mészáros, 2002a) or a non-uniform jet (see discussion in Section 2.2.4).

In short, while analytical treatments were quite thorough in the pre-Swift era, the afterglows of GRBs showed a great deal of complexity. The relative paucity of detections—and particularly afterglows with excellent time coverage—made drawing firmer conclusions difficult due to the sample sizes.

### 2.3.2 Swift era

Since the launch of Swift, more than 800 GRBs and afterglows have been observed. The large population and fine temporal/spectral coverage have allowed for population studies in previously unprecedented detail. A “canonical” picture has emerged, as illustrated in Figure 2.5 (Zhang et al., 2006; Nousek et al., 2006; Evans et al., 2009). This empirically guided model of the afterglow contains five or six key pieces, which I will describe using the Zhang et al. (2006) model. Note that this is chiefly a description of the afterglow behavior in X-rays, though there may be similarities at lower (or higher) wavelengths.

In rough order of appearance, the stages of a GRB afterglow are the following:

0. The prompt emission, i.e., the GRB itself

1. A very early period of very steep ($\alpha > 2$) decay that appears to be associated with the end of the prompt emission (Barthelmy et al., 2005b; O’Brien et al., 2006)

2. A shallow decay with spectral index $\alpha \sim 0.5$, or occasionally $\alpha < 0$ (note that since $F_\nu \propto t^{-\alpha}$, this corresponds to a brightening in X-rays). This part of the light curve is typically interpreted as energy injection into the blast wave, such as from central engine activity or a stratified jet (Zhang et al., 2006; Nousek et al., 2006; Panaitescu et al., 2006b). This plateau in emission was a minor surprise, as it is not predicted by the simple external shock model. If the slowing/reversal of the decay is, in fact, due to energy injection, the transition from Section 2 to Section 3 should signify the end of the injection—the point at which the energy source is cut off.
3. The $t^{-1}$ decay phase predicted from the standard external shock model of an afterglow (see Section 2.2)

4. A late decay phase with $\alpha \approx 2$ or steeper. In the analytical models presented above this corresponds to the jet break, if the break is achromatic.

5. X-ray flares are a common occurrence, seen in roughly half of all afterglows. They can share properties (such as the fast-rise, exponential-decay shape) with pulses during the prompt emission, so they are generally accepted to be caused by late activity of the same central engine responsible for the GRB (e.g., Ioka et al., 2005; Lazzati and Perna, 2007; Margutti et al., 2010).

The break in the light curve between Sections 2 & 3, as well as that between Sections 3 & 4, are expected to be achromatic for uniform jets. Unfortunately for the clean canonical curve presented
above, in many afterglows these breaks are chromatic—occurring in the optical and X-ray bands at
different times, if the break even occurs in both (Panaitescu et al., 2006a; Liang et al., 2008).\(^{18}\)

Some GRB afterglows appear to skip Section 3 entirely, going from the plateau of Section 2 to an
extremely steep decay post-break \(t^{-3}\) or even steeper; one example is GRB 070110 (Troja et al.,
2007, but note that the afterglow later recovered to a more typical shape). Such cases are uncommon
(only a handful are discussed in Liang et al., 2007), and would be Evans et al. (2009) “oddballs”. They
defy explanation under the external shock scenario. One possibility is that these afterglows are
associated with extremely long-lived central engines that dominate the standard forward shock emission
(Troja et al., 2007; Lyons et al., 2010).\(^{19}\)

Finally, I note that afterglows of short GRBs are generally fainter (see, e.g., the discussion in
Panaitescu et al., 2001), but have enough in common with long GRB afterglows (e.g., Gehrels et al.,
2008) that the external shock scenario & environment is likely to be similar for both classes (Nyse-
wander et al., 2009).

The \textit{Swift} era has generated a wealth of afterglow observations to test the simple analytical
model described in Section 2.2. These observations both confirm parts of it and force revisions
to other parts, while displaying a wide variation. Surprises included the early plateau of Section 2,
as well as the need for multiple emission regions to simultaneously explain (1) the “canonical”
broken power law in X-rays, (2) the irregular flaring behavior exhibited by many afterglows, and
(3) the chromatic nature of many observed spectral breaks. If the sheer diversity of the afterglow
population is to be explained within the framework of shock-accelerated particles, the parameter
space sampled by afterglows must be quite large. It is to this issue that I now turn.

\section*{2.4 DSA parameters}

Despite decades of dedicated work (see the previous chapter), there still remain many open questions
about shocks and their ability to accelerate particles to ultrarelativistic energies. In the \textit{Swift} era, GRB
afterglows present several dozen unique laboratories each year to measure parameters associated
with DSA. This is especially true because GRB jets are both collisionless and relativistic—the first
condition is (at present) a sizable technical challenge for terrestrial experiments, while the second
condition is not met anywhere within the neighborhood of the Sun. Assuming the correctness both

\(^{18}\)In fact, of the \(\approx 180\) afterglows studied in Liang et al. (2008), not a single spectral break met all criteria needed to be
called a “Platinum”-level break: definite breaks occurring in at least two bands at the same time, and whose frequency
spectral indices matched derived closure relations between pre- and post-break spectra.

\(^{19}\)In the (only slightly) less exotic field of supernova remnants, there do exist cases where remnants are observed to
have an extremely dim FS (e.g. SNR0540, Williams et al., 2008) or none at all (the Crab Nebula, Frail et al., 1995). It is not
unreasonable to suppose that similar variations might exist in the progenitors of GRBs.
of the forward shock model for GRB afterglows and of DSA as a means to generate the particles whose emission is detected, observations of afterglows will constrain parameters such as $E_{\text{tot}}$, $n_{\text{amb}}$, $p$, $\varepsilon_e$, $\varepsilon_B$, $\Gamma$, and $s$.\textsuperscript{20}

Given sufficient data, the values of $p$ and $s$ are simple to constrain using closure relations and an assumption about the nature of the afterglow (see the discussion in, e.g., Racusin et al., 2009; Gao et al., 2013).\textsuperscript{21} The data used in these closure relations typically have large error bars (see, e.g., the samples provided in Evans et al., 2007, 2009); when the data is also spotty, a single burst may satisfy closure relations for multiple scenarios to within a couple of standard deviations. The uncertainty in model transfers to uncertainty in the dependence on the electron spectral index, so while it may be possible to determine $p$ precisely, the accuracy is somewhat up for debate. This uncertainty is highlighted by the work of Curran et al. (2010) and Ryan et al. (2015). In the former study, Swift afterglows were determined to cluster around $p = 2.4$, with tails extending in both directions. In the latter, the interval $2 < p < 2.3$ is found to contain the bulk of Swift afterglows, with a weak tail in the range $p > 2.3$. It can safely be said that the question of the typical electron spectrum is not yet settled.

Once $p$ is known (up to the model-dependent uncertainties just mentioned), Kumar and Zhang (2014) note that the flux above $\nu_m$ is proportional to $\varepsilon_e^{p-1}$. Thus $\varepsilon_e$ is another parameter constrained to some extent by observations. Indeed, a review of the literature by Santana et al. (2014) found a well-defined peak at $\varepsilon_e \approx 0.1 - 0.2$. A histogram of the values reported in that work is presented as Figure 2.6. Very few afterglows show evidence for $\varepsilon_e < 0.1$ (though as afterglow emission is generated by synchrotron radiation of energetic electrons, this may be a selection effect), and only a few more have $\varepsilon_e$ much greater than 0.2. This provides observational evidence that the ability of relativistic shocks to deposit energy into electrons is roughly universal.

By contrast, the ambient density is extremely poorly constrained. Values reported in the literature range from 0.001 cm$^{-3}$ (GRB 051221A; Soderberg et al., 2006a) to 680 cm$^{-3}$ (GRB 050904; Frail et al., 2006). Furthermore, the particular model assumed can make orders of magnitude of difference in the fitted density (e.g., for GRB 020405: 8 cm$^{-3}$ in Chandra and Frail 2012, but $< 0.007$ cm$^{-3}$ in Chevalier et al. 2004). In Figure 2.7 I show 31 values of the ambient density for GRBs with radio, optical, and/or X-ray afterglow observations. There may be a slight bias towards lower densities for short GRBs (Soderberg et al., 2006b), but I do not distinguish here between short and long GRB

\textsuperscript{20}Respectively: the total energy contained in the jet, the ambient density, the spectral index of the electron distribution function, the fraction of $E_{\text{tot}}$ placed in relativistic electrons, the fraction of $E_{\text{tot}}$ placed in magnetic fields and their turbulence, the (time-dependent) Lorentz factor of the forward shock, and the power-law index for the density of the ambient medium.

\textsuperscript{21}Note the implicit assumption that $p$ is constant throughout the afterglow, and recall that $\varepsilon_B$ and $\varepsilon_e$ are similarly assumed to be constant.
Figure 2.6 A histogram of the 29 values of $\varepsilon_e$ presented in Santana et al. (2014). Note the break at $\varepsilon_e = 0.1$ to a finer scale around 0. Figure reproduced with permission from the authors.

...events. While the bulk of the densities reported are clustered in a span of 2-3 orders of magnitude, long tails stretch out to almost six decades in extent between the minimum and maximum values. This is a far cry from the tightly clustered values of $\varepsilon_e$ presented in Figure 2.6, highlighting the difficulty in determining this parameter.

The situation with magnetic field strengths is not much better than with that of density. In Figure 2.8 I show the 30 values of $\varepsilon_B$ provided by Santana et al. (2014). Unlike the clear peak in $\varepsilon_e$, the reported values of $\varepsilon_B$ span more than four decades. There are three “peaks” separated by minima in the histogram, but the small-number statistics makes these unlikely to be individually significant. As these values were taken from the literature, and since $\varepsilon_B$ is a difficult parameter to constrain anyway, the wide spread may represent variation in methodology rather than in reality.

The progenitors of long GRBs are all but confirmed to be collapsing massive stars. Such stars undergo extensive mass loss in the present-day Universe, suggesting that their more distant, GRB-generating counterparts might be better fitted by wind profiles than by constant-density models. This is, surprisingly, not the case. Detailed attempts to fit afterglow profiles (Panaitescu and Kumar, 2002; Schulze et al., 2011) lead to the conclusion that not only are the majority of afterglows consistent with an ISM-like CBM, a majority of afterglows are better-fitted by constant-density CBMs than by wind-like CBMs. This may, however, simply mean that the stellar termination shock is small enough...
that it is quickly overtaken by the jet.\footnote{This scenario is more plausible given the special relativistic time dilation between local and observer times. See the discussion in Section 8.2, and particularly Table 8.1.}

The typical interstellar magnetic field in the Milky Way is $1 - 10 \mu G$, and this value is typically assumed for the CBM in afterglows. This yields a very low magnetization ($\sigma \approx 10^{-9}$) for the unshocked plasma the GRB jet is encountering. The magnetic field in the vicinity of GRB progenitors may be significantly higher, though. Studies of star-forming regions in nearby spiral and starburst galaxies have suggested ambient magnetic fields of $20 - 100 \mu G$ (Beck, 2011). By measuring Zeeman splitting of the 1667 MHz OH megamaser line, Robishaw et al. (2008) reported magnetic field strengths of $500 - 18000 \mu G$ in starburst galaxies. These latter values, however, are found in highly compressed gas clouds and are unlikely to be representative of the ISM in those galaxies. In any event, the magnetization parameter has a strong dependence on field strength: $\sigma \propto B^2$. It is questionable whether DSA could occur in GRBs exploding in such environments. As is known from Particle-in-Cell (PIC) simulations (see Section 1.5.2), strongly magnetized upstream plasma prevents the formation of the turbulence needed to pre-heat and scatter cosmic rays in the Fermi-I process (Sironi et al., 2013). This would significantly depress any forward shock emission from these GRBs.

Even in GRBs with no forward shock emission, observation of a reverse shock can place additional

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{histogram.png}
\caption{A histogram of 31 values of $n_0$ found in the literature. The bulk of the values come from Chandra and Frail (2012); a few were given in Soderberg et al. (2006b); one was listed in Fox et al. (2005).}
\end{figure}
2.5 Summary

As with cosmic rays in Chapter 1, the discovery of GRBs was accidental. Also in keeping with CRs, further exploration revealed GRBs to be extremely energetic events, offering the means to test theories of physics at vastly higher energies than any terrestrial lab can support. Much progress has been made, including identifying the progenitors of both the long and short classes of GRBs, and in particular the discovery of the afterglow emission. GRB afterglows are visible from radio to
2.6 Goals of this Dissertation

In the previous two chapters I have briefly sketched the history and current state of two very different fields. The GRB community is well aware that they must incorporate energetic particles in modeling afterglows, and the CR community has long looked at GRBs and their afterglows as potential sites for Fermi-I acceleration. However, no work has examined the production of cosmic rays in the nonlinear regime, and in particular no work has considered the DSA process as the GRB jet slows from nonrelativistic to trans-relativistic. The goals of this dissertation are, therefore, several:

- Analyze the nonlinear Fermi-I process in the fully relativistic and trans-relativistic regimes. Using proton-only shocks (since protons typically possess the bulk of the upstream kinetic energy and momentum), examine the shifts in acceleration efficiency, and in the resultant CR spectrum, as shocks slow from \( \Gamma \beta \gg 1 \) through \( \Gamma \beta \sim 1 \) to \( \Gamma \beta \ll 1 \).

- Incorporate electrons and helium nuclei into the nonlinear system to accurately model the circumburst material swept up by GRB forward shocks. Compare the simulated shocks against both PIC simulations and expectations from GRB observations.

- Using a hydrodynamical model for a relativistic jet, simulate nonlinear shocks at many times, speeds, and sizes to create a sequential profile of the evolution of the forward shock of a GRB afterglow, and the CR population it generates via DSA.

- Model the photon emission from accelerated particles. At first consider only the particles accelerated by a particular shock speed, but extend the results to incorporate particles accelerated at all previous shocks and swept downstream.

The single biggest innovation presented in this work is the use of self-consistent nonlinear DSA in the relativistic and trans-relativistic regimes, which will have significant effects on the acceleration of CRs and the resultant photon emission.
The work presented in this dissertation is computational in nature. The code I used was initially developed by my advisor during his graduate studies, and has undergone numerous major additions and revisions in the intervening time. There are many published papers that describe the code to some degree (Ellison et al., 1981; Ellison and Moebius, 1987; Ellison et al., 1990b; Jones and Ellison, 1991; Ellison et al., 1996; Ellison and Double, 2002, 2004; Vladimirov et al., 2008; Ellison et al., 2013, and the numerous references therein). For the sake of completeness I will describe all relevant portions of the code here, rather than refer the reader to any of the many citations just provided. “Relevant” here refers to just the parallel-shock scenario. While the code has in the past been extended to consider oblique geometries, for all of the work presented in this dissertation the parallel orientation is sufficient (see Chapter 5).

In brief, the code performs two major tasks. First, the code takes a population of particles and allows them to scatter in a fixed velocity profile. As the particles scatter, their positions are tracked; their kinematic properties (e.g., momentum, energy) are used to determine fluxes across grid zone boundaries at numerous places in the shock profile. Next, the code applies a smoothing algorithm that adjusts the velocity profile in response to the calculated fluxes. The goal at this stage is to satisfy the Rankine–Hugoniot equations for flux conservation everywhere in the shock profile. If necessary, the code repeats the two stages in an iterative procedure, which converges on a smoothed shock
profile that conserves energy and momentum everywhere.

3.1 Assumptions

The most important assumption made in the code is that shock acceleration can be treated by a Monte Carlo process. Instead of the completely deterministic approach used by PIC simulations, the particles tracked in the code are constantly changing their directions (and therefore their positions) via a random process that mimics scattering off of magnetic turbulence generated within the system. As mentioned in Section 1.5.3, this approach handles the microphysics numerically, offering a major advantage over analytical models. However, the Monte Carlo method parameterizes away much of the small-scale plasma physics that PIC codes treat explicitly, allowing it to handle far larger ranges in space and energy while requiring vastly less computational firepower.

With the numerical approach fixed, the next most significant assumption that the code makes is that of a steady-state system. The particles the code tracks are injected and tracked individually through many time steps (as opposed to moving forward one time step for all particles at once, as in PIC simulations). After all particles are finished (escaped the shock or culled for a different reason), the distribution is calculated at specified points within the shock structure for handling the feedback loop between particles and shock (see Section 3.6 for more information about how the momentum and energy fluxes modify the shock structure). The quality of the results is dependent on the number of particles that interact with the shock, as fluctuations due to randomness are smoothed out by the law of averages. Because of this approach, the shock structure (velocity profile, and magnetic field strength/orientation/turbulence) may only change after all particles have run through the shock. The steady-state requirement is not a severe limitation in many astrophysical contexts; the Earth’s bow shock, the forward shock of young supernova remnants, and even shocks between galactic clusters are all very slowly varying in time. In other situations the steady-state assumption is more restrictive. Rapidly-varying conditions might occur in the earliest stages of shock breakout in a supernova, the collision of shocks within a jet (such as might occur near active galactic nuclei or during the prompt emission phase of a GRB), or the striped winds expected in pulsar wind nebulae.

The Monte Carlo code assumes that the subshock\(^1\) is completely transparent, and that scattering processes are identical (up to fluid speed and magnetic field strength) on both the upstream and downstream sides of the fluid discontinuity. This assumption is slightly at odds with PIC simulations of relativistic shocks, which have shown an increase in turbulence just upstream of the shock (e.g.,

\(^1\)The fluid discontinuity separating the downstream region from any precursor that may form, or from the upstream region if no precursor is present. See Figure 3.2 for an illustration.
Sironi and Spitkovsky, 2011; Sironi et al., 2013), but additional exploration is deferred to a later date. As well, previous analytical and numerical studies (e.g., Ellison and Jones, 1988) have explored the possibility of an electric potential on either side of the shock. This potential can radically alter the ability of charged particles to cross the shock. Depending on specifics such as orientation or strength, the potential might “bounce” electrons back upstream before they can cross the shock, leading to a large energy gain before the first shock crossing; it may also prevent electrons from crossing back upstream and so keep them from entering the acceleration process at all. Either of these possibilities could be implemented in the Monte Carlo code at some point, but their addition and study is reserved for a later, post-dissertation, date. For now I simply note that, despite the (potentially gross) simplification that a transparent subshock represents, the model presented in this work retains a great deal of explanatory power.

Related to the transparent subshock is another assumption made by the code, that thermal leakage in the shock frame is the primary method for particles to enter the acceleration process. Upstream from the shock, particles are traveling supersonically, and their thermal speeds are (by the definition of a shock) less—and in the case of relativistic shocks, possibly orders of magnitude less—than the flow speed. The thermal speeds may range from non-relativistic to ultra-relativistic; the only requirement is that the upstream flow must be both supersonic and super-Alfvénic. After crossing the shock, however, the bulk fluid velocity becomes subsonic, and many particles will have thermal (local plasma frame) speeds that exceed the downstream (shock frame) flow speed. The situation is illustrated in Figure 3.1. Particles whose velocity vectors fall within the range highlighted in blue have a shock frame velocity whose $x$-component exceeds that of the downstream bulk flow speed $u_{DWS}$. This allows them to move from the downstream region back into the upstream region, which marks them as cosmic rays that have entered the acceleration process. (See Section 3.3 for more information on this distinction.)

To reduce the complexity of the simulation, the code allows fluid quantities (velocity, magnetic field strength, etc.) to vary only as functions of one dimension in the parallel case. This treats the shock as a large (but not necessarily infinite) plane. In cases like supernova remnants, where the actual shock is a closed surface around some interior volume, one can still treat the shock as a flat plane on small enough spatial scales. Care must therefore be taken to ensure that the upstream and downstream regions are close enough to the shock that the curvature is unimportant. For situations like jets—either from active galactic nuclei, gamma ray bursts, or other objects—where the shock has lateral edges, an additional assumption is needed. In these scenarios we assume that particle escape out of the sides of the jets is negligible, i.e. that the shock surface is much wider than it is deep.

The code makes a strong assumption about the strength and the power spectrum of magnetic
3.1. ASSUMPTIONS

Figure 3.1 Thermal leakage as a mechanism for injection into the acceleration process. The red circle represents a hypothetical particle near the shock in the downstream region. Surrounding the particle is a sphere with radius $|\vec{v}_{sf}|\Delta t$ representing all potential orientations of the particle’s velocity vector. As long as the velocity is greater than the downstream flow speed $u_{DwS}$, it is possible for the particle to cross the shock back into the upstream region and become a cosmic ray during this time step $\Delta t$. The possible range of velocity orientations is shown with a blue background.

field turbulence. The turbulence present around the shock determines the nature of the scattering, which is obviously a large piece of the Monte Carlo approach. The code assumes the nature of the scattering, pitch-angle diffusion, which will be discussed later in the chapter. In making the assumption of pitch-angle scattering, the code takes for granted the existence of self-generated turbulence with both the proper wavenumber and sufficient strength to deflect the even the highest-energy particles’ momentum vectors by the desired amount. There is neither evidence for nor against turbulence like this in relativistic shocks, but there is obviously emission observed from the afterglow phase of GRBs attributed to accelerated particles (see the discussion in Section 2.3). PIC simulations are still running on grid sizes too small allow particles to approach the knee in the cosmic ray spectrum (see Section 1.1.2). While Monte Carlo methods have been employed in studying this regime, no study of which I am aware has considered both the nonlinear nature of the problem and the simultaneous acceleration of electrons and ions at these shocks. As such, even if this assumption seems to be large in degree, it still represents a significant step forward in the state
3.2 Injection

In the previous section I discussed the major assumptions that were made in the design and structure of the code, before the first particle is run through the shock structure. In this section I will explain the manner in which the particles are injected to interact with the shock.

The code tracks several quantities for each particle in each species’ population. Relevant to injection are (1) the total momentum in the local plasma frame \( p_{t, pf} \), (2) the component of plasma-frame momentum parallel to a chosen axis \( p_{x, pf} \), and (3) the weighting factor assigned to each particle \( w_{pt} \). At injection, the particles are drawn from a thermal distribution at the chosen temperature. Since the thermal distribution is isotropic with respect to pitch angle, the choice of orientation axis has no effect on the distribution. However, given that the particle populations will be placed into a bulk flow and mean magnetic field, the code uses those to define the orientation from which the pitch angle \( \theta \) is measured—and therefore from which \( p_{x, pf} \) is calculated.\(^2\) Orientation vectors chosen randomly on the surface of a sphere have an equal chance of falling into any pitch angle range \( \cos(\theta), \cos(\theta + d\theta) \); this is satisfied to within statistical fluctuations by the code.

Next, each particle receives a weighting factor \( w_{pt} \). This happens for two reasons. First, although the code uses the same number of particles within each species, the fraction of the upstream plasma represented by each species varies: electrons in a fully ionized, 10% helium plasma represent more than half of the upstream particle number, while helium nuclei are less than 5%. The contribution of a given particle to the collective energy/momentum/number fluxes (see Section 3.3) must therefore be weighted proportionally to its presence in the material encountering the shock:

\[
    w_{tot, i} = \frac{n_{0, i}}{\sum_j n_{0, j}}. \tag{3.1}
\]

In the equation above, the total weight to be split among the all particles of species \( i \) is the fraction of the total number density the species represents. One noteworthy consequence of this definition of particle weight is that the total weight for all particles encountering the shock sums to unity. The second reason for assigning each particle its own weighting factor is momentum splitting, which will be discussed in greater detail in Section 3.4.2.

Finally, each particle is placed in the bulk flow a large distance upstream from the shock (with

\(^2\)In an oblique shock geometry, the bulk flow and magnetic field will not align; see (e.g., Ellison et al., 1996) for how this Monte Carlo code handles that scenario.)
one exception, described in Section 3.7). Each particle then independently propagates within the shock structure, which I discuss in the next section.

### 3.3 The Particle Distribution and Fluxes

The code divides the shock structure into a non-uniform grid. The situation is illustrated in Figure 3.2. At each grid zone boundary, we place an omnidirectional detector that is stationary in the shock frame. These detectors are oriented parallel to the shock face and are infinite in transverse extent. In other words, the detectors exist on the entire $y-z$ plane as illustrated in Figure 3.2, and their normal vectors are parallel to the $x$ axis. It is these detectors that are used to track the population of particles as they propagate within the shock structure, with consequences (specifically, the particle distribution function and various fluxes) we describe in this section.

Before discussing the tracking, a point about calculation of fluxes. PIC simulations place a number of particles in each cell at the start of the simulation and update each particle simultaneously. By contrast, the Monte Carlo code tracks each particle individually, until it either reaches the next momentum cut or escapes the shock structure (see Section 3.4.2). Additionally, the separate particles represent a steady state flux of particles that are being swept up by the shock as it expands (in the frame of the interstellar medium) into the upstream material. The time used to calculate the fluxes is therefore one “iteration length”, regardless of the physical time—in whichever reference frame—that represents.

Particles will only be detected when they cross one of the grid zone boundaries. As such, particles that have a larger velocity perpendicular to the plane of the detector (i.e. parallel to shock normal) will be detected more often. More quantitatively, the probability $P(v_{sf,x})$ that a particle will be detected is proportional to $|v_{sf,x}|$. However, each detection represents just one of a large population of particles, and for each particle that is detected we must assume that there are many more that did not cross the detector. Particles whose shock frame velocities lie almost entirely in the $y-z$ plane will be detected extremely rarely, yet in an isotropic distribution they are just as numerous as particles with velocities aligned with the $x$ axis.

We reverse this skewed detection rate by “flux-weighting” the detections. Each particle has a weight $w_{pt}$, which is specific to that particle and related to the density of that species and any momentum splitting (see Section 3.4.2) undergone. When a particle crosses a grid zone boundary, the code records its corrected weight $w_{pt}/|v_{sf,x}|$. The density of particles at a particular location on
Figure 3.2: An illustration of several key facets of the shock system. At the top left, a set of coordinate axes identifies the $+x$ direction as downstream, and the $y$ and $z$ axes are perpendicular to the shock face—and as such, interchangeable. The boundaries of the grid zones are drawn through all three sections, and the subshock is identified by the gray oval at the bottom right. **Top:** the shock frame velocity profile for a smoothed shock is provided. Note that far upstream from the shock, the plasma flows toward the subshock with a constant velocity of $u_0$ (or $\beta_0$, or $\gamma_0$, depending on how relativistic the shock is) before it encounters the precursor. Downstream from the subshock the velocity is fixed. **Middle:** the plasma frame density corresponding to the above bulk plasma speed. In keeping with the Rankine–Hugoniot equations for flux conservation, it is inversely proportional to the shock frame velocity of the fluid (however, the exact dependence on velocity changes as the shock speed increases from non-relativistic to fully relativistic speeds). **Bottom:** a cartoon of the particle population at each grid zone at a snapshot in time. The particles downstream from the subshock (i.e. to the right of the gray oval) have been omitted for clarity, but they are present at roughly the same density as in the rightmost grid zone.
the grid is therefore proportional to the sum of the corrected weights of all particle detections.

\[ n(x) \propto \sum_{pt} \frac{w_{pt}}{|v_{sf,x}|} \]  

(3.2)

Note in Equation 3.2 that the density is restricted to vary only in one dimension, as mentioned in Section 3.1.

The constant of proportionality is easily determined in the nonrelativistic case: \( u_0 \), the velocity with which the shock is propagating into the ISM. To demonstrate that \( u_0 \) is the proper constant, consider a shock encountering a region of extremely cold particles. If we break the population into \( N_{pt} \) Monte Carlo particles, each particle has an equal weight of \( n_0/N_{pt} \). Since the particles are very cold, their thermal speeds are negligible compared to the shock velocity, and so \( \bar{v}_{sf} \approx v_{sf,x} \approx u_0 \) for each particle. Far upstream from the shock, after all \( N_{pt} \) particles have crossed a grid zone boundary we see that Equation 3.2 does indeed recover the expected plasma-frame density, \( n_0 \).

In the relativistic case, however, the Lorentz transformations and velocity additions described during Section 3.4 make identifying the constant of proportionality \textit{a priori} difficult. However, we know exactly what the plasma-frame density should be at any given grid point, courtesy of Equation 3.17 (which may be found in Section 3.6 below). It is a simple matter, then, to track the particle crossings using Equation 3.2 and determine the correct proportionality constant after the fact.

I next consider how to obtain the phase space distribution function of particles \( f(\vec{x}, \vec{p}) \), the number of particles in a region \( d^3x \, d^3p \) of the six-dimensional phase space. Given that definition, integrating over momentum must recover the number density\(^3\) \( n(\vec{x}) \equiv n(x) \):

\[ \int f(\vec{x}, \vec{p}) \, d^3p = \int f(x, \vec{p}) \, p^2 dp d\Omega \equiv n(x) \]  

(3.3)

Consider now the angle-averaged phase space distribution function of a finite population of particles. The integration over \( dp \) becomes a summation over bins of size \( \Delta p_i \), and once the angular dependence has been averaged out we see that

\[ \sum_{p_i} 4\pi \bar{f}(x, p_i) \, p_i^2 \Delta p_i = n(x), \]  

(3.4)

where the sum runs over all momentum bins \( p_i, \Delta p_i \) is the width of the bin containing \( p_i \), and \( \bar{f}(x, p) \)

\(^3\)Recalling that fluid quantities are constrained to vary only in the \( x \) direction.
denotes the angle-averaged value. From Equations 3.4 and 3.2 it follows that
\[ \sum_{p_i} 4\pi p_i^2 \Delta p_i \tilde{f}(x, p) \propto \sum_{pt} \frac{w_{pt}}{|v_{sf,x}|}, \tag{3.5} \]
and so
\[ \tilde{f}(x, p_i) \propto \frac{1}{4\pi p_i^2 \Delta p_i} \sum_{pt \in \Delta p_i} \frac{w_{pt}}{|v_{sf,x}|}. \tag{3.6} \]
In Equation 3.6, the summation runs over all particles whose shock-frame momenta fell within \( \Delta p_i \) at the time of detection. For isotropic particle distributions in the shock frame, \( \tilde{f}(x, p) = f(x, p) \) and Equation 3.6 relates the phase space distribution directly to the detections made during the running of the simulation. However, the bulk flow speed is so great upstream from a relativistic shock that even ultra-relativistic particles have anisotropic distributions. Downstream from the shock, where the bulk flow speed tops out at \( \frac{c}{3} \), the cosmic ray distribution is closer to isotropy, but it is still clear that Equation 3.6 is less useful in the relativistic regime than it is for non-relativistic shocks.

In Chapter 5 I will return to this discussion; at that point I will discuss the extension of this procedure to relativistic shocks.

In order to perform the nonlinear shock smoothing I will explain in Section 3.6, the Monte Carlo code must be able to track the momentum and energy flux across the detectors. To do this, the code uses moments of Equation 3.2. For the three Rankine–Hugoniot flux conservation relations (Equations 3.17–3.19), the moments are as follows:
\[ F_{num}(x) \propto \sum_{pt} v_{sf,x} \frac{w_{pt}}{|v_{sf,x}|}, \tag{3.7} \]
\[ F_{px}(x) \propto \sum_{pt} p_{sf,x} v_{sf,x} \frac{w_{pt}}{|v_{sf,x}|}, \tag{3.8} \]
and
\[ F_{en}(x) \propto \sum_{pt} E(p) v_{sf,x} \frac{w_{pt}}{|v_{sf,x}|}. \tag{3.9} \]
In Equation 3.9 \( E(p) = \sqrt{p^2 c^2 + m^2 c^4} - mc^2 \) is the kinetic energy of a particle with mass \( m \) and momentum \( p \). This means the code does not track the rest-mass energy flux across grid zone boundaries. Note also that by construction \( F_{px}(x) \) may never decrease, regardless of the direction that particles move: if the particle is moving in the \(-x\) direction, the product of its momentum and velocity retains a positive sign. On the other hand, the fluxes \( F_{num}(x) \) and \( F_{en}(x) \) are increased by
particles moving downstream and decreased by particles moving upstream.

3.4 Propagation

Once the particles are injected upstream, the next task is propagation within the shock structure. This is the main computational loop within the Monte Carlo code, and has the following structure:

1. Determine the size of the time step to take. The basic unit of time in the code is the period of gyration of a particle about its guiding center, \( T_g = \frac{(2\pi \gamma mc)}{(qB)} \). Here \( \gamma \) is the Lorentz factor of the particle in the plasma frame, \( m \) is its mass, \( c \) the speed of light, \( q \) the particle charge, and \( B \) the local magnetic field strength. This period is divided into some number of time steps \( N_g \), a parameter that returns in Section 3.4.1. The time step is thus

\[
t_{\text{step}} = \frac{T_g}{N_g}
\]

2. Calculate the distance moved, in the shock frame, during this time step. The code first calculates the distance moved in the plasma frame during the time step, \( \Delta x_{pf} = \frac{(p_{x,pf}t_{\text{step}})}{(\gamma m)} \). The code also uses the particle’s current position on the grid to obtain the relative velocity \( u_x \) (and associated Lorentz factor \( \gamma_x \) between the plasma and shock frames). Finally, the code combines all of the above in a relativistically correct manner:

\[
\Delta x_{sf} = \gamma_x \left( \Delta x_{pf} + u_x t_{\text{step}} \right).
\]

Note that this equation is correct only in the parallel orientation; if the mean magnetic field is oblique to the shock normal, the equation must be modified. Instead of tracking the particle directly in such a case, the particle’s guiding center is translated as the particle gyrates around it. As will be discussed further in Chapter 5, however, the parallel orientation suffices for the results presented here.

3. Check to see if the particle crossed any grid zone boundaries during its motion. Perform the updates described in Section 3.3 if crossings occurred.

4. Determine whether the particle scatters this time step. (In the version of the code used for this presentation, scattering is guaranteed, but previous work with this code has allowed for the possibility that the particle will not scatter during any given time step.) If the particle crossed into a different grid zone, and the bulk fluid velocity differs from that of its initial position, then the particle’s energy will change once it scatters off of the turbulence in the new grid zone.
So the particle’s plasma-frame momentum must be transformed from the old frame, into the shock frame, and then into the new plasma frame. Additionally, establish the variables used in the scattering process if needed. (Scattering is detailed further in Section 3.4.1.)

5. Finally, decide whether another run through the above loop is necessary; if so, return to the first step. There are two conditions that would cause an exit from the loop, both of which will be discussed in their own sections. The first possibility is that the particle exceeds the current splitting momentum (Section 3.4.2), and is stored until the next momentum level. The other reason the loop might be terminated is that the particle escapes the shock region. The possibilities are explored further in (Section 3.5), but the code eliminates particles that move too far upstream or downstream from the shock, exceed a maximum energy, or hit a maximum acceleration time.

The above loop continues until one of the conditions mentioned in the final point is satisfied. It is these conditions that control the length of time particles can spend crossing and re-crossing the shock, and so impact the maximum energy achievable by a given shock configuration.

3.4.1 Scattering

Scattering is the randomness that makes this code Monte Carlo, so its importance to the code is difficult to overstate. Details of the scattering process have been presented previously (Ellison et al., 1990a, 1996; Ellison and Double, 2002, 2004; Ellison et al., 2013), so here I will merely summarize key features.

Most importantly, the code assumes that particles scatter with a mean free path that is an increasing function of momentum. Specifically,

\[ \lambda_{\text{mfp}} = \eta_{\text{mfp}} r_g, \]  

(3.12)

where \( r_g = p c/(q B) \) is the particle’s gyroradius and \( \eta_{\text{mfp}} \) (assumed constant here) parametrizes the “strength” of scattering. Note that, in addition to the dependence on the particle momentum \( p \), there are additional dependences on the particle’s electric charge \( q \) and on the local magnetic field \( B \). As well, this version of the code does not model self-generated turbulence or magnetic field amplification (e.g., McKenzie and Voelk 1982; Bell 2004, 2005—but see Vladimirov et al. 2006, 2008, 2009 for a version of the code that does). Nor do we track particle orbits in a magnetized background with a known strength and power spectrum (e.g., Niemiec et al., 2006; Lemoine and Revenu, 2006). As mentioned in 3.1, the code assumes that magnetic turbulence exists everywhere in the shock.
structure, and in sufficient strength that Equation 3.12 holds independent of the particle’s position, while acknowledging the reality to be far more complex.

The strong scattering limit of $\eta_{\text{mfp}} = 1$ is called Bohm diffusion, and is commonly assumed because of its simplicity. There is, however, little direct evidence for Bohm diffusion in actual sites where DSA takes place. Uchiyama et al. (2007) examined the variability of X-ray bright knots behind the nonrelativistic shocks of RX J1713.7-3946, and concluded that the short timescales on which the knots dimmed and brightened require $\eta_{\text{mfp}} \sim 1$. Using X-ray and GeV afterglows in GRBs, Sagi and Nakar (2012) have suggested that acceleration can approach the Bohm limit when the blast wave is ultra-relativistic. Nevertheless, these estimates are indirect and highly uncertain, and some simulations of relativistic and trans-relativistic shocks (e.g., Lemoine and Revenu, 2006) imply that $\eta_{\text{mfp}} \gtrsim 10$ or higher. Furthermore, it has been known for some time that high-energy particles scattering in small-scale turbulence may have $\lambda_{\text{mfp}} \propto p^2$ rather than $\lambda_{\text{mfp}} \propto p$ (e.g., Jokipii, 1971; Plotnikov et al., 2011).

In parallel shocks, $\eta_{\text{mfp}}$ sets the lengthscales and timescales for acceleration and $\eta_{\text{mfp}} \gg 1$ may severely limit the maximum CR energy a given shock can produce. In oblique shocks, $\eta_{\text{mfp}} \gg 1$ will reduce the injection efficiency (see Ellison and Double, 2004). For the work presented in this dissertation, I use $\eta_{\text{mfp}} = 1$. Larger values of $\eta_{\text{mfp}}$ only increase the length scales and time scales of the simulation. In the steady-state picture of a shock used in this work, the only length scale present is the location of the free escape boundary (FEB). Simulations where the FEB is fixed in diffusion lengths give results independent of $\eta_{\text{mfp}}$.

The scattering prescription used in this dissertation is pitch-angle diffusion. In pitch-angle diffusion, the probability of scattering after a time step is set to 1, but the particle’s momentum vector makes a random walk on the surface of a sphere instead of being completely randomized. One example of pitch-angle diffusion is illustrated in Figure 3.3, which shows the three important parameters to the scattering process: $\delta \theta_{\text{max}}$, $\delta \theta$, and $\delta \phi$.

The scattering parameter $\delta \theta_{\text{max}}$ determines the maximum possible variation in pitch angle for a single scattering event. In Ellison et al. (1990b), the authors found that $\delta \theta_{\text{max}}$ was related to the length of the time step: $\delta \theta_{\text{max}} = \sqrt{6 t_{\text{step}}/t_c}$. The quantity $t_c$ is the collision time, i.e. the mean free path of the particle divided by the speed. Since I earlier set the mean free path to a multiple of the gyroradius, $t_c = \eta r_g / v = \eta \tau_g / 2\pi$. Under the assumption of Bohm diffusion, $\eta = 1$, and $\delta \theta_{\text{max}} = \sqrt{12\pi t_{\text{step}}/\tau_g}$. Using Equation 3.10, the relationship between $\delta \theta_{\text{max}}$ and $N_g$ is

$$\delta \theta_{\text{max}} = \sqrt{\frac{12\pi}{N_g}}. \quad (3.13)$$
It is obvious from Equations 3.10 and 3.13 that $N_g$ controls the fineness of the scattering process. As $N_g$ increases, more scatterings happen every gyroperiod, but each scattering involves less change (on average) to the particle’s pitch angle.

The other two parameters that control scattering, $\delta \theta$ and $\delta \varphi$, are both random numbers. The change in pitch angle for a given scattering is $\delta \theta$. To find it, a value for $\cos(\delta \theta)$ is chosen at random from the interval $[\cos(\delta \theta_{\text{max}}), 1]$, with $\delta \theta$ determined from that. The parameter $\delta \varphi$, as seen from Figure 3.3, “twists” the orientation of the momentum vector and offers the second degree of freedom needed for a random walk across the surface of a sphere. The interval for $\delta \varphi$ is $[0, 2\pi)$, from which it is chosen with a uniform probability. Once both $\delta \theta$ and $\delta \varphi$ are obtained, the new pitch angle $\theta_{\text{new}}$ is determined using the formula

$$\cos \theta_{\text{new}} = \cos \theta_{\text{old}} \cdot \cos(\delta \theta) + \sin \theta_{\text{old}} \cdot \sin(\delta \theta) \cdot \cos(\delta \varphi).$$  \hspace{1cm} (3.14)$$

The pitch angle is, as mentioned earlier, measured from the orientation of the mean magnetic field, which for a parallel shock is also the direction of the bulk fluid flow throughout the shock structure. In the axially symmetric scenario associated with the parallel orientation, it is not necessary to keep track of the azimuthal angle $\varphi$ for particles; its only use is in modifying the change in pitch angle during the scattering calculations.

A final assumption is that all scatterings are elastic and isotropic in the frame where the scattering occurs, and that (as a result of their assumed infinite mass) the scattering centers are frozen in the fluid. The Monte Carlo code thus cannot model second-order Fermi acceleration or the transfer of
energy between particles and magnetic field via the production or damping of magnetic turbulence. It also neglects any cross-shock electric potential that may exist. The Monte Carlo model does self-consistently include so-called shock-drift acceleration although this process, whereby particles gain energy as they gyrate in a compressed magnetic field, is not important in the parallel-shock results presented here.

Again it is necessary to stress the intermediate nature of the Monte Carlo approach to DSA: while it is clear that the approximations used to model the wave–particle interactions are severe compared to the self-consistent nature of PIC simulations, or even Monte Carlo simulations that trace particle orbits in prescribed background turbulence, they are more general than what is currently possible with semi-analytic solutions. Given our parameterization of the plasma interactions, we can model shock smoothing and NL acceleration over a dynamic range well beyond that currently accessible with 3D PIC simulations.

A final note, discussed in Vladimirov (2009) and references therein, is that pitch-angle diffusion is a bit of a misnomer. Contrary to initial assumptions (e.g. Bell, 1978), the magnetic field turbulence that cosmic rays self-generate can reach or exceed the strength of the mean field. If the magnetic field becomes turbulence-dominated, it becomes difficult or even impossible to define an orientation for the mean field, and therefore a reference axis for the pitch angle. Fortunately, in the parallel shock orientation considered here the fluid velocity provides an equivalent reference axis against which particle orientations may be measured.

3.4.2 Momentum Splitting

As Monte Carlo processes are random by nature, statistical fluctuations are unavoidable in the results. At best they can be minimized by using large numbers of particles, which is what the momentum splitting process is designed to achieve. At runtime, a list of cutoff momenta is read in, which divides the simulation into a sequence of \( p_{\text{cut}} \)s in plasma-frame momentum. Another input parameter is a target number of particles to use at each \( p_{\text{cut}} \); this number is typically chosen large enough that the \( 1/\sqrt{N} \) statistical fluctuations are small compared to the number of particles.

While particles are injected far upstream with a Maxwell-Boltzmann distribution of momenta in the plasma frame, this will inevitably change as they propagate in the shock structure. At the end of each time step, the code checks to see if the particle’s momentum exceeds the cutoff for the current \( p_{\text{cut}} \). If so, the particle is set aside and the code moves on to the next particle. If not, the particle moves and scatters until it either exceeds the cutoff momentum (via repeated crossings of the subshock) or escapes through any of the ways described in Section 3.5. Once all particles have been banked or eliminated in this fashion, the code duplicates the surviving particles to ensure a
minimum number of particles interacts with the shock at every energy considered.

Momentum splitting also determines when the simulation ends. Since each particle has only one of two fates by the end of each $p_{\text{cut}}$ (exceeded the momentum cutoff and saved for the next $p_{\text{cut}}$, or escaped the shock structure), if no particles made it to the cutoff momentum then all particles in the system escaped. When this finally occurs, the code moves on to either the next particle species or the next iteration.

### 3.5 The Upper Limit on the Particle Spectrum

It was determined very early on in the study of shock acceleration that the predicted $p^{-4}$ particle spectrum could not continue to arbitrarily high energies. The total energy contained in such a spectrum is

$$\int_{p_{\text{min}}}^{\infty} f(p)E(p) \, d^3p \propto \int_{p_{\text{min}}}^{\infty} \frac{1}{p^4} \cdot p \cdot p^2 \, dp = \int_{p_{\text{min}}}^{\infty} \frac{dp}{p},$$

where $p_{\text{min}}$ is the momentum where the $p^{-4}$ power law begins, chosen to be high enough that $E(p) \approx pc$. It can be seen in the above equation that such a spectrum contains an infinite amount of energy. Faced with such a physically implausible consequence, the conclusion was that the spectrum must end (or at least steepen) at some maximum momentum. The implication is that particles must be able to escape the region of the shock somehow, or that there is another limit on how energetic particles may become (such as the presumably finite lifetime of the shock). In this section I explain the numerous methods the code uses to manage particle escape, all of which serve to limit the maximum energy of the particle spectrum and avoid an energetic catastrophe.

The simplest means of achieving a maximum particle energy is to specify the maximum energy particles may have during the simulation; if a particle ever exceeds this energy it is culled from the population of particles propagating within the shock structure. This restriction reflects a spatial limitation on the size of the acceleration region. At some point a particle’s gyroradius becomes comparable to the size of the shock (be it a supernova remnant, pulsar wind nebula, or even termination shock of the jet of an active galactic nucleus). A particle with this energy effectively decouples from the acceleration region, since it is no longer possible for it to easily cross and recross the shock—to say nothing of the notion of a uniformly-oriented magnetic field the size of the acceleration region. This notion holds true as long as particles’ mean free paths (and diffusion

\footnote{An alternative interpretation raised in Ellison and Jones (1991) is that the spectral index might asymptotically approach $\sigma = 2$ from above at high energies, merely giving the appearance of an energy crisis but without requiring an actual steepening or cutoff.}
lengths) are increasing functions of momentum, a very permissive assumption.

Another escape method is the notion of a free escape boundary. These may be implemented either upstream or downstream of the shock, and are specified locations on the grid. If a particle advances further upstream (downstream) from the shock than the location of the free escape boundary, it is assumed to have a zero probability of returning to the shock, and no longer contributes to the backpressure that modifies the shock in the nonlinear scenario. When the computational grid is widely spaced (as is that used in, e.g., Figure 4.4), the cessation of particle tracking introduces a sharp downtick in the observed energy flux. This is, however, purely numerical issue; and in any event it is less of a concern for the relativistic and trans-relativistic shocks to be considered in this dissertation, as very few cosmic rays diffuse far upstream to cause significant smoothing. These generalities are, of course, dependent on the particular parameters chosen for the shock to be simulated. It is possible to choose a set of input parameters that drastically reduces upstream particle diffusion in a non-relativistic shock, or dramatically enhances the same in a relativistic shock.

The “probability of return plane” is an alternative method for determining whether (and where) particles decouple from the shock and the acceleration process. Instead of a sharp cutoff downstream of the shock, the probability of return plane uses the random nature of the scattering process to analytically estimate the chance that a particle will return once it crosses the plane headed downstream. This result was reported first by Bell (1978), but has been recovered many times (Peacock, 1981; Drury, 1983; Ellison et al., 1996). It is correct in both the parallel and oblique geometries, and for shocks of all speeds from non-relativistic to ultra-relativistic. The only assumption made is that the particle distribution must be isotropic in the local plasma frame. If this condition is satisfied, then the probability that a given particle with a fixed speed will return to some arbitrary downstream location after crossing it headed downstream is

$$P_{\text{ret}} = \left( \frac{v - u_{2x}}{v + u_{2x}} \right)^2,$$

where \(v\) is the (total) plasma frame velocity of the particle in question, and \(u_{2x}\) is the \(x\)-component of the downstream fluid flow in the shock frame. This check cannot be implemented close to the shock due to the assumption of isotropy; shocks at all speeds show highly anisotropic particle distributions near the shock that only isotropize with sufficient distance (and scatterings) downstream. As implemented in the code, this check (if requested by the input) is performed after the particle has scattered approximately 7 mean free paths downstream from the shock. If the check fails, the particle is culled from the shock and considered to have escaped. If the check succeeds, however, the particle must be returned to the probability of return plane with a velocity vector pointed upstream.
This may be done one of two ways. The first is to analytically determine the distribution of such particles and assign each returning particle a state (pitch angle, phase angle) within the distribution. This method is fleshed out in Appendix A of Ellison et al. (1996). The second method is to run time “backwards” for particles that are chosen to return. In this approach, once the particle crosses the probability of return plane, it is placed in a flow moving upstream with speed $u_2$—against the particle’s current velocity vector. The particle then scatters in this medium until it crosses the plane again, at which point it is simply re-introduced downstream from the shock at the location of the probability of return plane. Both methods of handling the probability of return plane assume an essentially infinite downstream region beyond the plane in which particles can scatter, though the assumption is somewhat more obvious for the “retro time” approach.

The final method I will mention for limiting particle energy is acceleration time. At each time step, a particle-specific running counter is incremented by the length of the time step. If this exceeds a specified value, presumably the age of the acceleration region, the particle is removed from the simulation. It will not re-cross the shock, and thus gain no further energy. This method is somewhat at odds with the assumption of a steady-state shock (as shocks are sure to evolve over their lifetime), and is not used in any of the results presented in this dissertation.\footnote{Acceleration time is not a proxy for particle escape. As mentioned above it is, at best, a way to limit particle energy. However, it is useful for determining how long it takes to accelerate particles even in the steady-state case assumed for this dissertation.}

This discussion of particle escape is somewhat peculiar to the assumption of a steady-state shock. As explored by Drury (2011), in real shocks high energy particles will diffuse ahead of the shock in an expanding volume that grows faster than the shocked region. Particle escape in this environment is not rigidly defined (i.e., by the location of a free escape boundary) so much as it is a time- and energy-dependent probability. As time passes the shock (and the turbulence ahead of it) weakens to the point that it becomes essentially transparent to lower and lower particle energies. The highest-energy particles become more loosely bound to the shock system than lower-energy particles, and so diffuse ever further ahead of the shock. Individual particles reach the point of escape at different distances from the shock, at different energies, and at different times.

### 3.6 Smoothed Shocks

For a steady-state shock, several shock-frame quantities must be conserved everywhere in the shock structure: particle flux, momentum flux, and energy flux. This constraint is expressed by the Rankine–Hugoniot relations for flux conservation. The particular form of these equations depends
on the shock's speed and obliquity. For the parallel relativistic shocks examined in this work, the conservation equations read (Ellison et al., 1990b)

\[
\begin{align*}
\gamma_0^2 \beta_0 n_0 & \equiv F_{\text{num}}^0 = F_{\text{num}}(x) \equiv \gamma(x) \beta(x)n(x), \quad (3.17) \\
\gamma_0^2 \beta_0^2 \left[ e_0 + P_{xx,0} \right] + P_{xx,0} & \equiv F_{\text{px}}^0 = F_{\text{px}}(x) \equiv \gamma^2(x) \beta^2(x) \left[ e(x) + P_{xx}(x) \right] + P_{xx}(x) + Q_{\text{px}}^{\text{FEB}}, \quad (3.18) \\
\gamma_0^2 \beta_0 [e_0 + P_{xx,0}] & \equiv F_{\text{en}}^0 = F_{\text{en}}(x) \equiv \gamma^2(x) \beta(x) \left[ e(x) + P_{xx}(x) \right] + Q_{\text{en}}^{\text{FEB}}. \quad (3.19)
\end{align*}
\]

In the above equations, the subscript of “0” on the left hand side refers to the far upstream state, while quantities on the right hand side may be measured anywhere up- or downstream from the subshock itself. The quantities \( \beta \) and \( \gamma \) refer to the \( x \)-direction velocity and Lorentz factor of the bulk flow in the shock frame. Measured in the (comoving) plasma frame, however, are the local number density \( n \), internal energy \( e \), and the \( P_{xx} \) component of the pressure tensor \( P \). The pressure tensor of an isotropic fluid may be written

\[
P = \begin{pmatrix}
P_{xx} & 0 & 0 \\
0 & P_{yy} & 0 \\
0 & 0 & P_{zz}
\end{pmatrix}, \quad (3.20)
\]

with off-diagonal terms representing shear stresses that cannot exist in an unmagnetized fluid (there \emph{is} a magnetic field, of course, but its energy density is significantly lower than that of the fluid). In a parallel shock only two orientations matter—parallel and perpendicular to shock normal—so \( P_{yy} = P_{zz} \). In the work presented here we consider only isotropic pressure, where all three diagonal components of \( P \) are identical. An alternative situation might involve \( P_{xx} = P_\| \) and \( P_{yy} = P_{zz} = P_\perp \neq P_\| \). Pressure in this case is “gyrotropic” rather than isotropic; this scenario is discussed further in Double et al. (2004).

Equations 3.18 and 3.19 allow for the possibility of particle escape at an upstream free escape boundary (FEB); the fluxes lost are denoted \( Q_{\text{px}}^{\text{FEB}} \) and \( Q_{\text{en}}^{\text{FEB}} \) for momentum and energy, respectively. There is necessarily an escaping number flux of particles carrying \( Q_{\text{px}}^{\text{FEB}} \) and \( Q_{\text{en}}^{\text{FEB}} \), but the fraction of the overall population represented by these particles is vanishingly small. As such we ignore the contributions of a \( Q_{\text{num}}^{\text{FEB}} \) in Equation 3.17. However, for relativistic (and even trans-relativistic) shocks under consideration for most of this dissertation, particle escape is very unlikely and the escaping fluxes are almost exactly zero.

6The equivalent formulas for a non-relativistic, parallel shock may be found in, e.g., Ellison (1985). In a nonrelativistic oblique shock additional conservation equations exist that reflect the relevance of the magnetic field and the additional component that the bulk fluid velocity may have; see, for example, Ellison et al. (1996). In the relativistic oblique case those same relations are spectacularly complex, and are presented in Double et al. (2004).
There are six quantities relevant to parallel shocks that are initially unknown when the smoothing process begins: (1) the shock-frame fluid velocity \( u_2 \), (2) the plasma-frame density \( n_2 \), (3) the thermodynamic pressure \( P_2 \), (4) the plasma-frame energy density \( e_2 \), and (5)-(6) the escaping fluxes \( Q_{\text{FEB}}^{\text{px}} \) and \( Q_{\text{FEB}}^{\text{en}} \). With only three equations available (3.17–3.19), the system is far from closed.

One relation used to achieve closure is the equation of state of the fluid—the relation between energy density and pressure. Following Double et al. (2004), we write the fluid equation of state as

\[
e(x) = \frac{P(x)}{\Gamma(x)-1} + \rho(x)c^2,
\]

where \( P = P_{xx} \) is the isotropic scalar pressure (i.e., \( P = \text{Tr}(\mathcal{P})/3 \)), and \( \rho(x) = \overline{m}_\text{pt} \cdot n(x) \) is the mass density in the plasma frame. In using Equation 3.21, however, a new variable is introduced: the adiabatic index \( \Gamma \). This quantity is defined in the non-relativistic (\( \Gamma = 5/3 \)) and ultra-relativistic (\( \Gamma = 4/3 \)) limits of particle thermal energies. However, for intermediate, trans-relativistic particles it must be calculated directly from the particle distribution function. To avoid this difficulty, we use a simple scaling relation, which is based on the relative approach speed between any two points within the shock structure, and which behaves exactly correctly in the two limits (see Double et al., 2004):

\[
\Gamma(x) = \frac{4\gamma_{\text{rel}}(x)+1}{3\gamma_{\text{rel}}(x)},
\]

with

\[
\gamma_{\text{rel}}(x) = \frac{1}{\sqrt{1-\beta_{\text{rel}}(x)}}
\]

and

\[
\beta_{\text{rel}}(x) = \frac{\beta_0 - \beta(x)}{1 - \beta_0 \beta(x)}.
\]

In the above equations \( \beta_0 = u_0/c \) and \( \beta(x) = u(x)/c \). With this relation set, there are still two free parameters. The final equation needed relates \( Q_{\text{FEB}}^{\text{px}} \) to \( Q_{\text{FEB}}^{\text{en}} \). Since all particles escaping at an upstream FEB are fully relativistic, we may write their energy as \( E_{\text{esc}} = P_{\text{esc}}c \). In a relativistic shock any escaping particles must be almost completely aligned with the \(-x\) axis in order to diffuse against the incoming fluid. Thus, \( Q_{\text{FEB}}^{\text{en}} = c Q_{\text{FEB}}^{\text{px}} \). The system is still not closed, as we have six equations and seven unknowns. However, selecting a flow speed \( u_2 \) allows for exact determination of all other quantities on the right hand side of Equations 3.17–3.19.

In the absence of particle acceleration, i.e. if each particle crosses the shock exactly once before being swept downstream (which also implies that \( Q_{\text{FEB}}^{\text{px}} = Q_{\text{FEB}}^{\text{en}} = 0 \), Equations 3.17–3.19 and 3.21–3.22 may be balanced if the ratio between the upstream and downstream speeds is carefully chosen.

\(^7\text{Here, and elsewhere, a subscript "2" refers to downstream locations.}\)
The flux-conserving ratio is called the Rankine–Hugoniot compression ratio, $r_{RH}$.

However, the presence of accelerated particles contributes to the momentum and energy fluxes both upstream and downstream of the shock.\(^8\) This violates flux conservation, even when the Rankine–Hugoniot compression ratio is used. Furthermore, the presence of accelerated particles varies significantly with position in the shock structure; in the near upstream region many accelerated particles are present, while far upstream the accelerated particles are few in number but high in energy.

It is insufficient to simply choose a new compression ratio and continue to treat the shock in the test-particle framework. In order to preserve flux conservation everywhere within the shock structure, the shock-frame speed of the plasma must be adjusted—specifically, reduced—based on the presence of accelerated particles. This results in a smoothed velocity profile compared to the sharp test-particle case. From a physical standpoint, the presence of cosmic rays in the upstream region creates a backpressure that acts on the incoming plasma, slowing it down before it reaches the shock.\(^9\) The more cosmic rays are present ahead of the shock, the greater the smoothing effect they induce. Once the shock is smoothed, though, the previous compression ratio may no longer conserve flux between far upstream and downstream. By modifying the shock and reducing the strength of the subshock, the character of the unaccelerated downstream population is adjusted as well.

I now describe the method used to achieve flux conservation for a shock of arbitrary speed (leaving out one wrinkle that I will return to in Section 4.3.1). The first step is to select a value for the downstream flow speed $u_2$. (This allows for analytical determination of the remaining unknown fluid quantities in the downstream region.) Next a population of particles is propagated through the shock as described in previous sections. After each iteration, once all particles have either (1) escaped at the upstream FEB, or (2) been advected far downstream from the shock, we obtain from the simulation

\[
\Delta F_{px}(x) = F_{px}^0 - \left[ F_{px}(x) + Q_{px,MC}^{FEB} \right] \quad (3.25)
\]

and

\[
\Delta F_{en}(x) = F_{en}^0 - \left[ F_{en}(x) + Q_{en,MC}^{FEB} \right]. \quad (3.26)
\]

That is, we determine the difference in momentum and energy flux between the known fluxes far upstream and the Monte Carlo results. The subscript “MC” on $Q_{px,MC}^{FEB}$ and $Q_{en,MC}^{FEB}$ indicated that these quantities are determined from the Monte Carlo simulation rather than being solved

\(^8\)Accelerated particles technically add to the number flux as well, but typically are a very small fraction of the total number of particles in the system. With that in mind, their contribution to Equation 3.17 is negligible and ignored.

\(^9\)The fluid is slowed down only in the shock frame, of course. In the far upstream frame the cosmic rays compress and mildly pre-accelerate the plasma just before it is overtaken by the shock.
for separately. A consistent solution, i.e. a flux-conserving velocity profile, is one that—to within statistical limits—sees $\Delta F_{px}(x) \simeq \Delta F_{en}(x) \simeq 0$, $Q_{px,MC}^{FEB} \simeq Q_{px}^{FEB}$, and $Q_{en,MC}^{FEB} \simeq Q_{en}^{FEB}$.

If (as will likely be the case) the shock profile does not meet these conditions to a satisfactory degree, we obtain a new velocity profile $u_{\text{new}}(x)$ in the following manner. First, we calculate the pressure from the right hand side of Equation 3.18 using $F_{px}(x)$ as measured during the Monte Carlo simulation:

$$P_{MC}(x) = \frac{F_{px}(x) - Q_{px}^{FEB}}{\gamma(x)\beta(x)} - \gamma_0\beta_0 c^2 \frac{\Gamma(x)}{\Gamma(x) - 1} + \frac{1}{\gamma(x)\beta(x)},$$

where I have made the substitution $\gamma(x)\beta(x) n(x) = \gamma_0\beta_0 n_0$ (i.e., Equation 3.17). This pressure is now inserted back into Equation 3.18, but with $F_{px}(x)$ replaced by $F_{px}^0$, and the resulting equation is solved for a new velocity profile $u_{\text{new}}(x)$.

The above procedure iterates on $u(x)$ and converges to a solution, but will not yield a flux-conserving solution unless the compression ratio $r_{\text{tot}} \equiv u_0 / u_2$ is consistent with Equations 3.17–3.19 and all associated equations. If the converged $u(x)$ is not sufficiently close to conservation, a new value for $r_{\text{tot}}$ is chosen and the procedure is repeated.

3.7 Speedup

In this section, I outline two methods used to speed up the running of the code. Energy gains happen when a particle encounters a velocity difference between scattering locations. However, the vast majority of the shock structure is either far upstream (where the absence of accelerated particles ensures a constant velocity) or downstream (where the flow speed is held fixed). All methods described here reduce computation time by increasing the speed with which the code handles particles in those two regions.

The most obvious method to increase speed (keeping equal parameters like shock speed and number of particles) is to reduce the number of scatterings each particle experiences during its encounter with the shock. This is accomplished by lowering the scattering fineness $N_g$ described in Section 3.4. However, this has effects far beyond speeding up the run. Reducing $N_g$ increases both the ability of particles to cross the shock and the average energy gain for doing so, and can significantly alter the particle distribution (Ellison and Double, 2004). As such, reducing $N_g$ should be viewed with caution, if not outright suspicion, as a means to increase the speed of Monte Carlo runs.

In the shock frame, upstream thermal particles move (on average) at the bulk flow speed with
3.8 Summary

The code used in this dissertation is complex, and the work of many people over several decades. As a Monte Carlo code, it makes many assumptions about the large-scale structure of the shock and the microphysics controlling the interaction between particles and the magnetic field. It uses pitch-angle diffusion as a prescription for scattering, which itself makes significant assumptions about the character of the magnetic field turbulence. Momentum splitting is used to ensure high-quality statistics even at the highest end of the particle distribution, and the code has several ways of enforcing an upper limit on the distribution. Importantly for the results to be presented later, the code correctly handles the complex nonlinear relationship between accelerated particles and the shock that produced them. This allows for the treatment of shocks that (as is physically correct in the steady-state case) conserve number, momentum, and energy flux in accordance with the Rankine–Hugoniot conditions.

This chapter represents, more or less, the state of the code as I received it. Over the next few chapters I will describe the modifications and additions I have made in performing the work presented in this dissertation.
4.1 Abstract

We present results from a Monte Carlo simulation of a parallel collisionless shock undergoing particle acceleration. Our simulation, which contains parametrized scattering and a particular thermal leakage injection model, calculates the feedback between accelerated particles ahead of the shock, which influence the shock precursor and “smooth” the shock, and thermal particle injection. We show that there is a transition between nonrelativistic shocks, where the acceleration efficiency can be extremely high and the nonlinear compression ratio can be substantially greater than the Rankine–Hugoniot value, and fully relativistic shocks, where diffusive shock acceleration is less efficient and the compression ratio remains at the Rankine–Hugoniot value. This transition occurs
4.2. INTRODUCTION

Collisonless shocks are known to accelerate particles, and diffusive shock acceleration (DSA, also known as the first-order Fermi mechanism) is widely regarded as the most likely mechanism for this acceleration (Krymskii, 1977; Axford et al., 1977; Bell, 1978; Blandford and Ostriker, 1978; Jones and Ellison, 1991; Malkov and Drury, 2001; Bykov et al., 2012; Schure et al., 2012). Most shocks of astrophysical interest are nonrelativistic, where the speed of the shock \( u_0 \) is a small fraction of the speed of light \( c \). Relativistic shocks, where the flow speed Lorentz factor \( \gamma_0 = \sqrt{1 - (u_0/c)^2}^{-1/2} \) is significantly greater than unity, are both less common and more difficult to study analytically. Because of these differences, and the fact that nonrelativistic shocks are both ubiquitous and known to be efficient accelerators (e.g., Ellison et al., 1990a; Warren et al., 2005), most work on DSA has been directed toward understanding nonrelativistic shocks.

There are some objects, however, where relativistic shocks are likely to be important in accelerating particles (see Bykov and Treumann, 2011; Bykov et al., 2012). These include gamma-ray bursts (GRBs), relativistic supernovae (SNe), pulsar wind nebulae (PWNe), extra-galactic radio jets, and blazars. Relativistic shocks are more difficult to describe analytically because the assumption of particle isotropy breaks down. As long as \( u_0 \) is much less than the individual particle speed \( v_p \), one can make the diffusion approximation that particles are roughly isotropic in all frames. As the shock becomes relativistic, \( v_p \sim u_0 \sim c \) and the diffusion approximation no longer applies (see, however, Kirk et al., 2000; Blasi and Vietri, 2005, and references therein for analytic approaches to the problem of relativistic shocks).

To date the most fruitful approaches to DSA by relativistic shocks have been particle-in-cell (PIC) and Monte Carlo simulations. PIC simulations (e.g., Giacalone and Ellison, 2000; Nishikawa et al., 2007; Spitkovsky, 2008; Sironi and Spitkovsky, 2011) can directly model not only the particle

---

\(^1\)In Chapter 2, as is customary within the GRB community, the symbol “\( \Gamma \)” was used to refer to the bulk Lorentz factor of plasma, while “\( \gamma \)” was reserved for individual cosmic rays. In this chapter and those to follow, I return to the convention of the CR community, where “\( \gamma \)” is used for bulk flow speeds.
acceleration process but the shock formation process as well. They have a great advantage over analytic and Monte Carlo techniques (e.g., Ostrowski, 1988; Ellison et al., 1990b; Ellison and Double, 2002; Lemoine and Pelletier, 2003; Baring, 2004; Niemiec and Ostrowski, 2006) in that they can determine the self-generated magnetic turbulence, and can treat ions and electrons, self-consistently. As is well known, details of the wave–particle interactions are far more important in relativistic shocks than in nonrelativistic ones, so determining the magnetic turbulence self-consistently is critical (e.g., Pelletier et al., 2009; Plotnikov et al., 2013). The biggest disadvantage of PIC simulations is that they are extremely computationally intensive and must be fully three-dimensional (3D) in order to model the full effects of cross-field diffusion (see Jones et al., 1998, for a derivation of the restricted dimensionality constraint).

The Monte Carlo technique used in this paper treats particle scattering and transport explicitly, which, in effect, solves the Boltzmann equation with collective scattering (e.g., Ellison and Eichler, 1984). It contains a specific thermal injection model whereby some fraction of shock-heated particles are able to re-cross the shock and gain additional energy. This injection process is a direct result of the scattering assumptions and requires no additional parameters.

Once the scattering assumptions are made (as previously detailed in Section 3.4.1), the shock structure, overall compression ratio \( r_{\text{tot}} \equiv \frac{u_0}{u_2} \), and full particle distribution function \( f(x, p) \) at all positions relative to the subshock are obtained self-consistently. Here and elsewhere the subscript “0” implies far upstream values and the subscript “2” implies downstream values, i.e., \( x > 0 \).

In terms of applicability, the Monte Carlo model lies between analytic techniques and PIC simulations. Analytic techniques are computationally faster, but they have difficulty treating thermal particle injection. On the other hand, our Monte Carlo calculations are much faster than PIC simulations, and they can provide important results at all shock speeds, including trans-relativistic flows. The main disadvantages of our Monte Carlo technique are that it is intrinsically steady state and plane-parallel, it is computationally intensive compared to semi-analytic techniques, and it assumes a form for the scattering operator and so does not model the magnetic turbulence generation or the shock formation process self-consistently.\(^2\)

While we have emphasized some differences between nonrelativistic and relativistic shocks, all shocks must conserve momentum and energy, and unmodified (UM) shocks undergoing efficient particle acceleration do not (e.g., Drury, 1983; Blandford and Eichler, 1987).\(^3\) If a shock, regardless of its speed, puts a non-negligible amount of energy into superthermal particles via DSA, the back-
pressure of the accelerated particles slows the upstream plasma before it reaches the discontinuous subshock (see Figure 4.5 in Section 4.4.2.1). This effect is nonlinear (NL) in nature and acts to modify the shock structure in order to conserve momentum and energy. Apart from PIC results and the preliminary work of Ellison and Double (2002), we know of no self-consistent, NL, relativistic and trans-relativistic shock solutions other than those presented here.

The mathematical and computational difficulties in treating NL effects notwithstanding, relativistic shocks are expected to be intrinsically less efficient than nonrelativistic ones, making NL effects less important.\(^4\) While this may be the case, we show that NL effects can be important in relativistic shocks depending on the particular assumptions made regarding particle diffusion. Specifically, if large-angle scattering (LAS) occurs, where the magnetic turbulence is strong enough to randomize the particle trajectory in just a few interactions, NL effects can be dramatic and cannot be ignored in any realistic astrophysical application.

In this paper we expand on previous Monte Carlo calculations (i.e., Ellison and Double, 2002, 2004; Double et al., 2004). We include particle escape at an upstream free escape boundary (FEB) and emphasize trans-relativistic shocks. Trans-relativistic shocks are likely to be important for GRB afterglows, where cosmic-ray (CR) production, perhaps to the highest observed energies (e.g., Katz et al., 2010), may occur. A subpopulation of Type Ibc SNe, not producing observed GRBs, have observed trans-relativistic speeds, and these have been proposed as sources of ultra-high-energy cosmic rays (UHECRs; e.g., Soderberg et al., 2010; Chakraborti et al., 2011). Of course, modeling astrophysical observations normally means modeling electrons since they radiate far more efficiently than do ions. While electrons can be included in our Monte Carlo simulation (e.g., Baring et al., 1999), doing so requires additional ad hoc assumptions. We do not attempt a combined treatment here, deferring the issue to a future paper.

The application of DSA to pulsar winds brings another important difficulty to the forefront: the shock obliquity, that is, the angle between the shock normal and the upstream magnetic field, \(\theta_{\text{Bn}}\). Relativistic shocks are certain to be highly oblique since the downstream magnetic field is bent away from the downstream shock normal more strongly than in nonrelativistic shocks (see, for example, Ellison and Double, 2004; Meli et al., 2008). This effect increases with the shock Lorentz factor, so the downstream angle in ultra-relativistic pulsar wind termination shocks should be almost perpendicular. While shocks of any obliquity can be treated directly with PIC simulations if they are done in 3D, oblique geometry makes semi-analytic and Monte Carlo models more difficult.

---

\(^4\)This stems mainly from the smaller compression ratio (\(r_{\text{tot}} \sim 4\) for strong nonrelativistic shocks versus \(r_{\text{tot}} \sim 3\) for ultra-relativistic shocks). In addition, there is more uncertainty concerning the ability of relativistic shocks to self-generate the magnetic turbulence that is necessary for DSA to occur (e.g., Pelletier et al., 2009). Acceleration beyond the initial Lorentz boost remains highly uncertain in ultra-relativistic shocks.
UM, oblique, relativistic shocks have been discussed extensively (e.g., Ostrowski, 1991; Ellison and Double, 2004; Lemoine and Revenu, 2006; Niemiec et al., 2006; Summerlin and Baring, 2012), but we know of no non-PIC treatments of self-consistent, relativistic, oblique shocks. Here, we confine our discussion to parallel shocks where $\theta_{\text{BN}} = 0^\circ$ and plan to treat modified oblique shocks in future work.

Our NL parallel shock simulations show several important features: (1) shock modification results in significant changes to the accelerated particle distribution for all Lorentz factors including fully relativistic shocks, even though they are intrinsically less efficient accelerators than nonrelativistic shocks; (2) relativistic shocks with coarse LAS attempt to accelerate particles efficiently, but this produces strong shock modification that dramatically softens the accelerated particle spectrum; (3) in going from fully relativistic to nonrelativistic shock speeds, as might be relevant for modeling trans-relativistic SNe and GRB afterglows, there is a relatively sharp transition from soft spectrum relativistic shocks (i.e., phase-space spectra $f(p) \propto p^{-4.23}$) to trans-relativistic shocks with $r_{\text{tot}} > 4$ and concave upward spectra typical of efficient nonrelativistic shocks; and (4) in trans-relativistic shocks, a lowering of the ratio of specific heats can result in compression ratios greater than the Rankine–Hugoniot value without particle escape.

The structure of the remainder of this chapter is as follows. In Section 4.3, we discuss relevant aspects of the Monte Carlo approach. We present results from our simulations in Section 4.4, paying special attention to trans-relativistic cases. In Section 4.5, we discuss these results in the context of astrophysical settings like GRBs, and we conclude in Section 4.6.

4.3 Model

The code was described at length in Chapter 3, so in this section only changes from that description are included.

4.3.1 Shock Modification

A self-consistent, flux-conserving shock structure is typically found with a method that iterates on two quantities: the flow speed profile $u(x)$ and the overall shock compression ratio $r_{\text{tot}} = u_0/u_2$ (see Ellison et al., 1996, for more details). An initial $r_{\text{tot}}$ and $u(x)$ are chosen, the simulation is run, and the momentum and energy fluxes are calculated at all $x$-positions. If the fluxes are not conserved, i.e., if they are not equal at all $x$-positions to the far upstream values, a new profile $u(x)$ is chosen and the simulation repeated with the same $r_{\text{tot}}$.

In some cases, particularly fully relativistic shocks, the paired iterations on $u(x)$ and $r_{\text{tot}}$ do not
result in a consistent solution and a third iteration on the subshock scale \( x_{\text{sub}} \) must be performed. In these cases, the subshock speed is defined as

\[
    u(x) = u_2 - [u_2 - u(x_{\text{sub}})] \frac{\tan^{-1} x}{\tan^{-1} x_{\text{sub}}}
\]

over the range \(-|x_{\text{sub}}| < x < 0\), where \( x_{\text{sub}} < 0 \). This scaling is discussed for specific examples in Section 4.4.

The combined iteration over \( r_{\text{tot}} \), \( u(x) \), and \( x_{\text{sub}} \) yields consistent smooth-shock solutions over the full range of shock speeds. We note that the Monte Carlo simulation uses Equations 3.17–3.19 (the Rankine-Hugoniot equations) as a consistency check, but these equations are not used directly in propagating particles or calculating fluxes in the simulation. All fluxes are calculated directly in the shock frame from the individual particle kinematics.

### 4.3.1.1 Examples with no DSA

In the absence of DSA (i.e., in a UM shock with \( Q_{\text{px}}^{\text{FEB}} = Q_{\text{en}}^{\text{FEB}} = 0 \)), Equations 3.17–3.19 yield the Rankine-Hugoniot compression ratio \( r_{\text{RH}} \), as determined by Double et al. (2004). Typically, \( r_{\text{RH}} \) is used to start our iteration process. In Figure 4.1 we show two examples with DSA disabled in the Monte Carlo simulation. The UM shock profiles are in the panels labeled \( u(x)/u_0 \), and the momentum and energy fluxes, scaled to far upstream values, are in the panels labeled \( F_{\text{px}}/U_{\text{pS}} \) and \( F_{\text{en}}/U_{\text{pS}} \), respectively. For both shock Lorentz factors we show the fluxes for three compression ratios. The solid (black) curves are for \( r_{\text{RH}} \), and the dashed (red) and dotted (blue) curves are for lower and higher \( r_{\text{tot}} \) values, respectively.

We note that the upstream energy fluxes shown in Figure 4.1 and elsewhere in this paper differ from those in Equation 3.19 in that they ignore the incoming rest-mass energy flux. While accelerated particles may carry away significant amounts of energy and momentum at a FEB, their density compared to the upstream plasma is always negligible, independent of almost all other factors. The rest-mass energy flux entering the shock is therefore approximately equal to the rest-mass energy flux leaving via downstream convection, no matter what the shock speed considered. Thus, we ignore it and calculate escaping kinetic energy flux as a fraction of incoming kinetic energy flux.

Except for the region immediately downstream from the shock, the fluxes determined by the Monte Carlo simulation in Figure 4.1 are precisely those determined by the Rankine-Hugoniot relations for \( r_{\text{RH}} \), even for the trans-relativistic case \( \gamma_0 = 1.5 \). Results for compression ratios larger and smaller than \( r_{\text{RH}} \) clearly do not conserve energy and momentum. The fact that the Monte Carlo results are consistent with \( r_{\text{RH}} \) as determined from Equations 3.17–3.24 is somewhat remarkable. As
Figure 4.1 Shock rest-frame values of the flow speed, $u(x)$, the momentum flux, $F_{px}$, and the energy flux, $F_{en}$, are shown for UM shocks with no DSA. All quantities are scaled to upstream values, and the position $x$ is in units of the gyroradius $r_g = m_p u_0 c / (e B_0)$, where $B_0 = 3 \mu$G. The solid (black) curves assume a compression ratio $r_{tot} = r_{RH}$, i.e., the Rankine–Hugoniot value as determined in Double et al. (2004). The dashed (red) and dotted (blue) curves are for lower and higher $r_{tot}$ values as indicated.
described in detail in Ellison and Reynolds (1991), once the basic scattering assumptions given in Section 3.4.1 are made, a consistent shock solution can be obtained with no further assumptions simply by running the simulation with different $r_{\text{tot}}$ values. There is no need to calculate a ratio of specific heats in the Monte Carlo simulation, and the fluxes are calculated directly by summing over particles as they move with the bulk flow. The fluxes shown in Figure 4.1 are calculated directly in the shock frame, and there is no need to transform into the local frame to calculate a pressure or energy density as required in Equations 3.17–3.24.

Of course, the reason for defining Equations 3.17–3.24 is that, with acceleration, $r_{\text{tot}}$ can be greater than $r_{\text{RH}}$. This occurs because $\Gamma^2$ can be less than the Rankine–Hugoniot value and/or because escaping particles carry away pressure, causing $r_{\text{tot}} > r_{\text{RH}}$ (e.g., Ellison and Eichler, 1984; Jones and Ellison, 1991; Berezhko and Ellison, 1999). Both of these NL processes are included self-consistently in the Monte Carlo simulation.

### 4.3.2 Particle Injection

While background superthermal particles will be injected and accelerated if they interact with a collisionless shock, we discuss here only thermal leakage injection where all particles that get accelerated start as far upstream thermal particles.

In our Monte Carlo technique, thermal leakage injection is modeled in its simplest form. Unshocked thermal particles, injected far upstream, convect and diffuse with a mean free path in the local frame given by Equation 3.12. No restriction on $v_p/u$ is imposed. At every pitch-angle scattering event a Lorentz transformation is performed such that the particle makes an isotropic and elastic scattering in the frame of the scattering interaction. After crossing into the downstream region and interacting with the downstream plasma, some fraction of the particles will have a speed $v_p > u_z$. These particles have a certain probability to diffuse back upstream and be further accelerated. In relativistic shocks, this probability for returning back across the shock depends on the scattering parameters $N_g$ and $\lambda_{\text{mfp}}$ and is determined stochastically for each particle in the Monte Carlo simulation.

A further assumption is that the subshock is transparent, i.e., the Monte Carlo model ignores any possible cross-shock potential or effects from enhanced magnetic turbulence, which would influence the subshock crossing probability or produce an energy change in the subshock layer. We emphasize that we do not present our injection model as a “solution” to the injection problem. Rather,

---

5For an isotropic distribution of particles, it is well known that the probability of return depends only on the downstream flow speed and an individual particle's speed (see, for example, Bell, 1978; Ellison et al., 1996). Near the shock, however, the distribution is highly anisotropic, and so the scattering parameters characterize the ability of particles to cross the shock and enter the acceleration process.
it is a clearly defined process that requires no additional assumptions beyond the scattering model basic to the Monte Carlo simulation. Within statistical uncertainties, it produces self-consistent shock solutions that can be compared with observations (e.g., Ellison et al., 1990a; Baring et al., 1997) or PIC simulation results (e.g., Ellison et al., 1993). Of course, the downstream return probability and the injection fraction will change substantially if an oblique magnetic field and compressed turbulence are taken into account (e.g., Niemiec et al., 2006; Lemoine and Revenu, 2006). In principle, the smooth-shock techniques we demonstrate here can be applied to this far more complicated situation.

In Figure 4.2 we show how the injection fraction varies with $N_g$ for two UM shocks, one with...
\( \gamma_0 = 1.5 \) (top panel) and one with \( \gamma_0 = 10 \) (bottom panel). The choice of \( N_g = 15 \) implies \( \delta \theta_{\text{max}} \approx \pi/2 \) (i.e., Equation 3.13) and is near the extreme LAS limit. Larger values of \( N_g \) imply finer scattering; the fine-scattering limit is called here pitch-angle scattering, or PAS. The fraction of protons that have crossed the shock three times (up to down, down to up, and up to down again) is clearly indicated by the sharp change in the spectral shape at \( p/(m_p c) \sim \gamma_0 \).

For the trans-relativistic shock (\( \gamma_0 = 1.5 \)) there is little effect from \( N_g \), as expected, since \( N_g \) is known to be unimportant for nonrelativistic, parallel shocks. For \( \gamma_0 = 10 \), \( N_g \) does influence the injection fraction and the acceleration. For the \( \gamma_0 = 10 \) cases in the bottom panel of Figure 4.2, the smallest value of \( N_g \) results in fewer particles injected after the first shock crossing. However, the flatness of the \( N_g = 15 \) and 34 curves above \( p \sim 10 m_p c \) shows that more particles continue to be accelerated for low \( N_g \) compared to higher \( N_g \). With a small \( N_g \), particles move a larger distance between scattering events and will first interact with the downstream plasma further from the discontinuous shock (or subshock as the case may be) than with fine PAS. This reduces the probability that these particles will return upstream. Countering this effect, however, is the fact that when \( \gamma_0 \) is large, LAS produces larger energy gains, on average, than fine PAS for subsequent shock crossings. This can produce a harder spectrum, as seen in Figure 4.2 (see Ellison and Double, 2004, for a detailed discussion of this effect). While the above effect is present in the \( \gamma_0 = 1.5 \) plots, it is much less noticeable.

What is not shown in the UM examples in Figure 4.2 is that, as the shock becomes modified by the backreaction from accelerated particles, the injection fraction at the subshock automatically adjusts according to the change in the subshock compression and the pre-heating that occurs in the shock precursor.

### 4.4 Results

We now show detailed results for shock speeds spanning the range from nonrelativistic to ultra-relativistic, i.e., (A) \( \beta_0 = 0.0667 \) (i.e., \( u_0 = 2 \times 10^4 \text{ km s}^{-1} \)); (B) \( \beta_0 = 0.2 \); (C) \( \gamma_0 = 1.5 \); (D) \( \gamma_0 = 10 \); and (E) \( \gamma_0 = 30 \). Other parameters for these models, as well as for Models F–I, are given in Table 4.1.

#### 4.4.1 Unmodified Shocks

In Figure 4.3 we show spectra from UM shocks for nonrelativistic (\( u_0 = 2 \times 10^4 \text{ km s}^{-1} \); Model A), trans-relativistic (\( \gamma_0 = 1.5 \); Model C), and fully relativistic (\( \gamma_0 = 10 \); Model D) shock speeds. For the examples in Figure 4.3 we have chosen \( N_g \) large enough to obtain convergent solutions,\(^6\) and, for

---

\(^6\)By convergent we mean that no changes other than statistical fluctuations occur if \( N_g \) is increased further.
### 4.4. RESULTS

#### CHAPTER 4. TRANSRELATIVISTIC SMOOTHED SHOCKS

Table 4.1 Shock Parameters

<table>
<thead>
<tr>
<th>Model</th>
<th>([u_0])</th>
<th>([\beta_0])</th>
<th>(\gamma_0)</th>
<th>(N_g)</th>
<th>(M_S^b)</th>
<th>(r_{RH}^c)</th>
<th>(r_{NL}^d)</th>
<th>(\xi_{DSA}^e)</th>
<th>(q_{En}^{\text{FEB}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>([2 \times 10^4])</td>
<td>20</td>
<td>17</td>
<td>3.96</td>
<td>12 ± 1</td>
<td>0.93</td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>{0.2}</td>
<td>60</td>
<td>16</td>
<td>3.93</td>
<td>11 ± 1</td>
<td>0.9</td>
<td>0.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1.5</td>
<td>500</td>
<td>1900</td>
<td>3.53</td>
<td>4.0 ± 0.4</td>
<td>0.55</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>10</td>
<td>2000</td>
<td>(2 \times 10^3)</td>
<td>3.02</td>
<td>3.02</td>
<td>0.3</td>
<td>&lt;0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>30</td>
<td>(10^4)</td>
<td>(2 \times 10^3)</td>
<td>3.00</td>
<td>3.00</td>
<td>0.35</td>
<td>&lt;0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>10</td>
<td>100</td>
<td>(2 \times 10^3)</td>
<td>3.02</td>
<td>3.02</td>
<td>0.5</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>15</td>
<td>2000</td>
<td>8</td>
<td>3.00</td>
<td>3.00</td>
<td>0.3</td>
<td>&lt;1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>10.05</td>
<td>4</td>
<td>1800</td>
<td>3.02</td>
<td>3.02</td>
<td>0.8</td>
<td>&lt;0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1.5</td>
<td>500</td>
<td>1900</td>
<td>3.53</td>
<td>4.0 ± 0.4</td>
<td>0.6</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(a\) For Model I, \(B_0 = 300\mu G\). All other models have \(B_0 = 3\mu G\). All models have \(n_0 = 1\ cm^{-3}\) and \(\eta_{\text{mfp}} = 1\).

\(b\) We use the nonrelativistic definition for the sonic Mach number, \(M_S^2 = \rho_0 u_0^2 / (\Gamma_0 P_0)\). The far upstream ratio of specific heats \(\Gamma_0 = 5/3\) in all cases.

\(c\) This is the Rankine–Hugoniot value as determined by Equations 3.17–3.19 with no DSA.

\(d\) This is the self-consistent total compression ratio with DSA.

\(e\) For the NL shock, \(\xi_{DSA}\) is the fraction of shock kinetic energy flux put into superthermal particles including those that leave at the upstream FEB. The number of significant figures suggests the level of accuracy.

\(f\) For the NL shock, \(q_{En}^{\text{FEB}} = Q_{en}^{\text{FEB}} / F_{en}^{0}\) is the fraction of shock kinetic energy flux that leaves the shock at the upstream FEB.

Table 4.1 (continued) Shock Parameters

| Model | \([u_0]\) | \([\beta_0]\) | \(\gamma_0\) | \(x_{\text{sub}}\) | \(|L_{\text{FEB}}|\) | \(L_{\text{FEB}}\) (pc) |
|-------|----------|----------|----------|--------|--------|--------|
| A     | \([2 \times 10^4]\) | 0\(^g\) | \(3 \times 10^4\) | \(6 \times 10^{-4}\) |        |        |
| B     | \{0.2\}   | 0        | \(1 \times 10^4\) | \(7 \times 10^{-4}\) |        |        |
| C     | 1.5       | 0        | \(1 \times 10^4\) | \(2.5 \times 10^{-3}\) |        |        |
| D     | 10        | 3        | 1000     | \(3.4 \times 10^{-4}\) |        |        |
| E     | 30        | 9        | 100      | \(3.4 \times 10^{-5}\) |        |        |
| F     | 10        | 1.75     | 1000     | \(3.4 \times 10^{-4}\) |        |        |
| G     | 15        | 4        | 3        | \(1 \times 10^{-6}\) |        |        |
| H     | 10.05     | 3.2      | \(1 \times 10^5\) | \(3.4 \times 10^{-2}\) |        |        |
| I     | 1.5       | 0        | \(1 \times 10^8\) | 0.25 |        |        |

\(g\) A value of 0 means the self-consistent shock profile was found without using Equation 4.1.

84
4.4. RESULTS

CHAPTER 4. TRANSRELATIVISTIC SMOOTHED SHOCKS

Figure 4.3 Omni-directional spectra measured in the shock frame from three unmodified shocks: \( u_0 = 2 \times 10^4 \) km s\(^{-1}\) (A), \( \gamma_0 = 1.5 \) (C), and \( \gamma_0 = 10 \) (D). All spectra are scaled to an injected flux of 1 proton cm\(^{-2}\) s\(^{-1}\), and, for all models, the solid (black) spectra were calculated downstream from the shock, the dashed (red) spectra were calculated at \( x = -100 r_{g_0} \), and the dotted (blue) spectra were calculated at the FEB. The positions of the FEB for the three shocks are \(-3 \times 10^4 r_{g_0}\) (A), \(-10^4 r_{g_0}\) (C), and \(-10^3 r_{g_0}\) (D). The thermal portions of the upstream spectra are not shown for clarity. The dashed line labeled \( \sigma_{KW} \) is the spectral index determined from (Keshet and Waxman, 2005). Note that the vertical axis is \( p^4 f(p) \).

In each case, we plot spectra calculated at different positions relative to the subshock at \( x = 0 \). The solid (black) curves are calculated downstream from the subshock, the dashed (red) curves are calculated upstream from the subshock at \( x = -100 r_{g_0} \), and the dotted (blue) curves were calculated at the upstream FEB. The gyroradius \( r_{g_0} \equiv m_p u_0 c / (e B_0) \) varies with shock speed. These spectra, and all other \( f(p) \) plots, are shock-frame, omni-directional, phase-space distributions.\(^7\)

Now consider the spectra calculated downstream from the shock (black curves in Figure 4.3). For the nonrelativistic shock, the particle distribution function between the thermal peak and the turnover produced by the FEB is well matched by a power law with spectral index \( \sigma \approx 4 \) (i.e.,

\(^7\)A shock-frame, omni-directional spectrum is one that would be measured by a 4\(\pi\)-steradian detector stationary in the shock frame.

85
4.4. RESULTS

CHAPTER 4. TRANSRELATIVISTIC SMOOTHED SHOCKS

\( f(p) d^3p \propto p^{-\sigma} d^3p \), as expected. Also as expected, the fully relativistic case with \( \gamma_0 = 10 \) obtains a power law in the mid-energy range with \( \sigma \approx 4.23 \), the well-known ultra-relativistic result (e.g., Bednarz and Ostrowski, 1998; Kirk et al., 2000). For the trans-relativistic case (\( \gamma_0 = 1.5, r_{\text{tot}} = 3.53, \) Model C), we compare the Monte Carlo result against the expression from Keshet and Waxman (2005), i.e., \( \sigma_{\text{KW}} \approx 4.08 \) (see their Equation (23)), and find reasonable agreement for the power-law index in the range \( p \gtrsim 300 m_p c \) and below the turnover from the FEB.

The dashed and dot-dashed (red and blue) spectra calculated in the shock precursor are missing the lowest energy particles since these are not able to diffuse back upstream to the observation position.\(^8\) The particles that escape at the FEB (dotted, blue curves) are streaming freely away from the shock, which causes the omni-directional distributions to be depressed relative to the more isotropic fluxes well within the precursor. We discuss this point with the self-consistent shock solutions shown next.

4.4.2 Nonlinear Relativistic Shocks

If astrophysical shocks are efficient particle accelerators, the NL backreaction of the accelerated particles on the shock structure must be taken into account (e.g., Drury, 1983; Blandford and Eichler, 1987). To our knowledge, apart from PIC simulations, the only previous attempt to treat NL, relativistic shocks was the preliminary work of Ellison and Double (2002).

4.4.2.1 Fine-scattering solutions

We now show self-consistent, fine-scattering shock solutions. Again, a “fine”-scattering solution has \( N_g \) large enough so no changes, other than statistical, result if \( N_g \) is increased.

Figure 4.4 shows the shock structure for \( u_0 = 2 \times 10^4 \) km s\(^{-1} \) (Model A). The dashed (red) curves show the UM solution with \( r_{\text{RH}} \approx 3.96 \), and it is clear from the lower panels that the momentum and energy fluxes are not conserved by large factors. Of course, the amount of acceleration depends on the injection model used, so these results are particular to the thermal leakage model contained in the Monte Carlo simulation. The solid (black) curves show the self-consistent result obtained by iteration where the shock has been smoothed and the compression ratio has been increased to \( r_{\text{NL}} = 12 \pm 1 \). Now, once the escaping energy flux at the FEB is accounted for, the momentum and energy fluxes are conserved to within a few percent of the far upstream values. The heavy dotted (blue) horizontal line in the energy flux panels is the fractional energy flux (i.e., \( 1 - Q_{en}^{\text{FEB}} / I_{en}^0 \)), calculated from Equations 3.17–3.19 once the modification of the downstream ratio of specific heats

---

\(^8\)The thermal injected particles that have not yet crossed the subshock are present at upstream positions but are not shown in Figure 4.3 for clarity.
Figure 4.4 Top panels show the shock structure in terms of $u(x)/u_0$, (solid-black and dashed-red curves) vs. position $x$ relative to the subshock at $x = 0$. Note the split log–linear scale. Length is measured in units of $r_{g0} = m_p u_0 c / (e B_0)$. The middle panels show the momentum flux $F_{px}$, and the bottom panels show the energy flux $F_{en}$. Values in all panels are scaled to far upstream values. In all panels, the dashed (red) curves are from a UM shock with the same input values as the modified NL shock results shown with solid (black) curves. For this nonrelativistic shock, the self-consistent compression ratio is $r_{NL} = 12 \pm 1$ compared to the Rankine–Hugoniot value $r_{RH} \approx 3.96$. The thick horizontal dotted (blue) line in the bottom panels at $\sim 0.5$ shows the energy flux expected from our consistency conditions derived from Equations 3.17–3.19. The FEB is at $x = -3 \times 10^4 r_{g0}$.
has been accounted for. Except for a small deviation near the FEB at \( x = -3 \times 10^4 r_{g0} \), the Monte Carlo result is consistent with Equations 3.17–3.19. The shock structure shown in the top panels, with \( r_{NL} = 12 \pm 1 \), is a consistent solution where approximately 50% of the energy flux escapes at the upstream FEB. There is essentially no escaping momentum flux, as expected for a nonrelativistic shock (see Ellison, 1985).

Solutions with \( r_{NL} \gg r_{RH} \), such as shown in Figure 4.4, have been known for some time for nonrelativistic shocks (e.g., Jones and Ellison, 1991; Berezhko and Ellison, 1999), and our \( u_0 = 2 \times 10^4 \) km s\(^{-1}\) result is consistent with these previous results.

In Figure 4.5 we show the UM and NL shock structures for our example (C) with \( \gamma_0 = 1.5 \). The NL effects are present but less dramatic than for \( u_0 = 2 \times 10^4 \) km s\(^{-1}\). For \( \gamma_0 = 1.5 \), the self-consistent compression ratio is \( r_{NL} = 4.0 \pm 0.4 \) compared to \( r_{RH} \approx 3.53 \). There is an insignificant amount of escaping flux through the FEB at \( x = -10^4 r_{g0} \), and the momentum and energy fluxes across the shock are equal to the far upstream values to within a few percent.

For this case, the increase in \( r_{tot} \) comes about because the downstream ratio of specific heats, \( \Gamma_2 \), decreases from the Rankine–Hugoniot value when particles are accelerated. This decrease makes the shocked plasma more compressible and results in \( r_{NL} > r_{RH} \). The downstream \( \Gamma_2 \) decreases from \( \Gamma_2 \approx 1.51 \), the value obtained from Equation 3.22, to \( \Gamma_{2,MC} \approx 1.47 \) in the NL case, enough of a difference to explain the increased compression ratio. For the parameters we are using, \( \gamma_0 = 1.5 \) is a transitional speed below which strong NL effects become dominant in the fine-scattering limit.

In Figure 4.6 we show the shock structures for our fully relativistic example (D) with \( \gamma_0 = 10 \). For this case, \( r_{NL} \approx r_{RH} \approx 3.02 \). There are no significant escaping energy or momentum fluxes at the FEB and, since the shocked particles are fully relativistic without additional acceleration, \( \Gamma_2 \approx 4/3 \). Thus, in this fine-scattering example, there are no mechanisms to increase \( r_{tot} \) above \( r_{RH} \). Nevertheless, the energy and momentum fluxes in the UM shock are nearly a factor of two above the far upstream values. A noticeable smoothing of the shock is necessary to conserve momentum and energy, as shown by the \( \gamma(x) \beta(x)/(\gamma_0 \beta_0) \) curve (dot-dashed, blue) in the top panels of Figure 4.6.

The smoothing parameter \( x_{sub} = -3r_{g0} \) is shown as a vertical dotted line in the top right panel of Figure 4.6. The flow profile to the right of this line is determined by Equation 4.1. This is matched to the profile to the left by our algorithm based on Equation 3.27. The procedure is iterated until consistent momentum and energy fluxes are obtained, as shown by the solid (black) curves in the lower panels.

In Figure 4.7 we compare the UM and NL phase-space distribution functions for our \( u_0 = 2 \times 10^4 \) km s\(^{-1}\) models. There is a dramatic difference, with the superthermal NL spectra being less intense and showing the characteristic concave upward shape. The shift in the thermal peak to lower energy, which accompanies efficient DSA, is also clear in the self-consistent result. This shift is a robust
Figure 4.5 Same format as Figure 4.4, but for $\gamma_0 = 1.5$ (Model C). In all panels the dashed (red) curves show the UM case and the solid (black) curves show the self-consistent case. For this example, the expected escaping momentum and energy fluxes predicted from our consistency conditions are small, as indicated by the horizontal dotted (blue) line in the bottom panels.
Figure 4.6 Same format as Figures 4.4 and 4.5, but for $\gamma_0 = 10$ (Model D). In all panels the dashed (red) curves show the UM case and the solid (black) curves show the self-consistent case. To emphasize the shock smoothing, we show $\gamma(x)\beta(x)/(\gamma_0\beta_0)$ in the top panels with a dot-dashed (blue) curve. The vertical dotted line in the top right panel shows the position of $x_{\text{sub}}$. 
prediction of DSA. When acceleration occurs, the subshock must weaken, and injection must decrease to conserve energy and momentum.

We note that our nonrelativistic example has a sonic Mach number of \( M_S \sim 17 \), which is unrealistically low for such a high-speed shock in the typical interstellar medium. We have chosen parameters to yield a low \( M_S \) for computational convenience. As shown in Berezhko and Ellison (1999), the self-consistent compression ratio is a fairly strong function of \( M_S \) (i.e., \( r_{\text{tot}} \sim 1.3 M_S^{3/4} \)) when only adiabatic heating occurs in the precursor, as we assume here. Such high \( r_{\text{tot}} \) values make finding a self-consistent solution difficult and are almost certainly unrealistic in any case, since Alfvén wave damping is likely to be a source of heating in the precursor. The sonic Mach number is less important for relativistic shocks unless the upstream temperature is high enough for the unshocked particles to have a Lorentz factor \( \gamma_0 \). We do not consider such a situation here.

The distribution of escaping particles for the NL nonrelativistic shock is shown as a dotted (blue)
curve in Figure 4.7. As shown in Figure 4.4, this distribution contains $\sim 50\%$ of the total energy flux, but this fraction cannot be directly obtained from the omni-directional distributions we plot. (In Figure 4.11 below we give a direct result for the acceleration efficiency.) What Figure 4.7 does clearly show is that the shape of the escaping distribution is dramatically different from the downstream distribution.

In Figure 4.8 we show the UM and NL distribution functions for our trans-relativistic example with $\gamma_0 = 1.5$. The differences between the UM and NL spectra are less than for $u_0 = 2 \times 10^4 \text{ km s}^{-1}$, but they are still significant. The normalization of the superthermal NL distribution is approximately an order of magnitude less than the UM case, and there is some slight concave curvature from the shock smoothing. While the thermal peak still shifts to a lower energy in the NL case, this shift is much less than it was for the NL nonrelativistic shock shown in Figure 4.7.

The escaping flux is too small to influence the shock dynamics. This is evident both from
4.4. RESULTS

CHAPTER 4. TRANSRELATIVISTIC SMOOTHED SHOCKS

Figure 4.9 Comparison of UM shock spectrum (Model D; same as in Figure 4.3) with NL shock spectrum. The shock structures for these $\gamma_0 = 10$ shocks are shown in Figure 4.6. The solid (black curves) are calculated downstream from the subshock in the shock rest frame and the dashed (red) curves are calculated in the shock precursor at $x = -100r_{g0}$.

Figure 4.5 and from the distribution of escaping particles, shown in Figure 4.8 as a dotted (blue) curve. The normalization of the escaping particle distribution relative to the downstream distribution is more than an order of magnitude less than it is in the nonrelativistic case (see Figure 4.7).

The distribution functions for our $\gamma_0 = 10$ example (Model D) are shown in Figure 4.9. While the shapes of the UM and NL spectra are quite similar, and while they both obtain $f(p) \sim p^{-4.23}$ power laws above the thermal peak, the normalization drops by a factor of $\sim 5$ when energy and momentum conservation are taken into account. There is also a slight shift in the thermal peak to lower energy, as required when DSA occurs.

In Figure 4.10 we compare the self-consistent distribution functions for our five examples spanning the range from $\beta_0 = 0.067$ to $\gamma_0 = 30$. The transitional character of the spectrum at $\gamma_0 = 1.5$ is clear. At shock speeds below $\gamma_0 \sim 1.5$, DSA can be extremely efficient and strong NL effects from particle escape and $r_{NL} > r_{RH}$ result in concave upward spectra that are harder than $p^{-4}$ above a few
Figure 4.10 Nonlinear particle distributions calculated downstream from the shock in the shock rest frame for various shock speeds as indicated (Models A–E in Table 4.1). The spectrum for the $\gamma_0 = 1.5$ shock (dashed black curve) shows the transitional nature of nonlinear DSA. For speeds faster than $\gamma_0 \sim 1.5$, NL effects are relatively minor. For speeds below $\gamma_0 \sim 1.5$, strong NL effects occur and concave spectra are produced. The strong increase in normalization from $\beta_0 = 0.067$ to $\gamma_0 = 30$ mainly results from the fact that thermal particles gain more energy in their first shock crossing as the shock speed increases. The maximum momentum is determined by the position of the FEB in the various shocks.
4.4. RESULTS

CHAPTER 4. TRANSRELATIVISTIC SMOOTHED SHOCKS

Figure 4.11 Acceleration efficiency in terms of the fraction of energy flux above shock frame $p$ for the $\beta_0 = 0.067$ (A), $\gamma_0 = 1.5$ (C), and $\gamma_0 = 30$ (E) shocks shown in Figure 4.10. These fractions include the energy flux that escapes at the upstream FEB. The escaping fraction is $\sim 0.5$ for $\beta_0 = 0.067$ but $\lesssim 0.02$ for the two other cases.

GeV. These slower shocks will be strongly dependent on the input conditions, i.e., Mach number, the injection model, the position of the FEB (which determines $P_{\text{max}}$), assumptions concerning magnetic turbulence production and damping, and a finite speed for the magnetic scattering centers (e.g., Berezhko and Ellison, 1999; Vladimirov et al., 2008). Above $\gamma_0 \sim 1.5$, the modified shocks have compression ratios $\sim r_{\text{RH}}$ and produce spectra that are nearly $\sigma = 4.23$ power laws. These faster shocks will be strongly dependent on the details of particle scattering (our parameter $N_g$ discussed below) and on effects from oblique geometry and compressed downstream magnetic fields, which we do not model here.

4.4.2.2 Acceleration efficiency

We define the acceleration efficiency, $\varepsilon_{\text{DSA}}(\gg p)$, as the fraction of total energy flux placed in particles with shock-frame momentum $p$ or greater. This is shown in Figure 4.11 for three of the spectra shown
in Figure 4.10. The other two lie between these examples. These fractions include the escaping energy flux. If the fraction of energy above the downstream thermal peak is the measure of efficiency, all of these shocks are extremely efficient. The nonrelativistic shock ($\beta_0 = 0.067$, black curve) places over 90% of the incoming energy flux into superthermal particles. Approximately 50% of the total flux is in escaping particles. While the more relativistic shocks are less efficient, the energy flux in particles with Lorentz factors $\gtrsim \gamma_0$ is still large. $E_{\text{DSA}} \sim 0.55$ for $\gamma_0 = 1.5$ and $\sim 0.35$ for $\gamma_0 = 30$. For these cases, the escaping energy flux is insignificant. Additional factors, such as magnetic turbulence generation and wave damping, are not included here and will reduce the acceleration efficiency (e.g., Berezhko and Ellison, 1999; Caprioli et al., 2009).

### 4.4.2.3 Large-angle-scattering solutions

The fineness of scattering, as determined by our parameter $N_g$, strongly influences DSA in relativistic shocks. This effect is discussed in Ellison and Double (2004), where results for UM shocks with LAS (i.e., $N_g$ much less than the value needed for convergence) are shown. LAS produces hard spectra and implies efficient DSA, so NL effects are certain to be important. Figure 4.12 (Model F) shows the dramatic effect shock smoothing has on a $\gamma_0 = 10$ shock with $N_g = 100$ (versus $N_g = 2000$ used to obtain the convergent $\sigma \approx 4.23$ solution shown in Figure 4.6). The only difference in input parameters between this case and the shock shown in Figure 4.6 is $N_g$.

With LAS the energy flux is out of conservation by a factor of $\sim 10$ with no shock smoothing, and the UM and NL particle distribution functions are vastly different, as shown in the downstream spectra plotted in Figure 4.13. In the UM case (dashed, red curve), discrete peaks are present in $f(p)$ corresponding to a given number of shock crossings. Similar spectra were seen in Ellison et al. (1990b) and recently in Summerlin and Baring (2012). The lowest momentum peak at $p/(m_p c) \approx \gamma_0$ is from particles that have crossed the shock once. In addition, clear peaks for three, five, and seven crossings are present, each at a momentum approximately a Lorentz factor greater than the previous peak.

The consistent NL solution is shown as the solid (black) curve in Figure 4.13. The peaks from discrete shock crossings seen in the UM shock are gone. The superthermal part of the spectrum is much softer, the lowest momentum peak is at a slightly lower momentum than in the UM case, and the rest of $p^4 f(p)$ shows the concave upward curvature that is seen in efficient DSA in nonrelativistic shocks. This concave curvature is present in Figure 4.9, but it is less evident because the acceleration is less efficient with $N_g = 2000$. In Figure 4.13, the NL spectrum obtains a power-law index just below the turnover caused by the FEB, which is slightly harder than $p^{-4}$. The NL result conserves energy and momentum, and this sets the normalization of $f(p)$, which is almost a factor of 100 below the
Figure 4.12 Shock structure for LAS Model F with $\gamma_0 = 10$ and $N_g = 100$. All input parameters for this model are the same as Model D (Figure 4.6) except $N_g$. The figure format is the same as in Figure 4.6.
4.5 Discussion

From kinematics, ultra-relativistic shocks are certain to energize incoming cold particles to $p \sim \gamma_0 m c$ in the shock frame (see the “thermal peak” in any of our spectral plots). DSA beyond this, where particles are required to make repeated shock crossings, is much less certain.

4.5.1 Comparison with Previous Monte Carlo Results

Monte Carlo techniques have been used to study relativistic shocks for some time (e.g., Ostrowski, 1988, 1993; Ellison et al., 1990b; Bednarz and Ostrowski, 1998; Baring, 1999). Recently, using es-
sentially the same scattering assumptions employed here, Summerlin and Baring (2012) made a comprehensive study of the plasma parameters important for test-particle DSA, including the shock obliquity. Here we compare our NL results (Model H) with the \( \theta_{\text{scatt}} = \pi \), \( \gamma_0 \beta_0 = 10 \), \( \theta_{\text{max}} = \pi \) example shown in Figure 12 of Summerlin and Baring (2012) and in Figure 2 of Stecker et al. (2007). This LAS example, where \( \delta \theta_{\text{max}} = \pi \) implies \( N_g \approx 4 \) by Equation 3.13, is similar to, but more extreme than, our Model F shown in Figures 4.12 and 4.13.

A qualitative comparison with the test-particle Summerlin and Baring (2012) result is sufficient to show the important differences that occur when a self-consistent shock structure is considered. In Figure 4.14 we show the shock structure and momentum and energy fluxes with (solid black curves) and without (dashed red curves) considering the backreaction of accelerated particles. The extreme LAS assumption produces very efficient acceleration in the UM shock, and the momentum and energy fluxes are overproduced by nearly a factor of 100. Shock smoothing must occur unless the injection is reduced by other factors to the point where essentially only the downstream thermal distribution is present. The dot-dashed (blue) curve shows the extent of smoothing required to give a self-consistent solution.

The UM particle distribution shown in Figure 4.15 as a dashed (red) curve is similar to that shown in Figure 12 of Summerlin and Baring (2012). It shows the characteristic step-like structure of LAS in UM shocks. As shown by the solid black curve in Figure 4.15, this step-like structure disappears in the self-consistent shock. It resembles the steeper \( \delta \theta_{\text{max}} \lesssim 20^\circ \) examples shown in Figure 12 of Summerlin and Baring (2012). Even if turbulence is present that can result in LAS in relativistic shocks, basic conservation considerations preclude the formation of flat, step-like spectra.

In contrast to the simple scattering assumptions made here and in Summerlin and Baring (2012), the Monte Carlo studies of Niemiec et al. (2006), and references therein, start with a background magnetic field structure including magnetic field perturbations with various wave-power spectra in ultra-relativistic, oblique shocks. Instead of assuming pitch-angle scattering, they calculate particle trajectories directly as particles move through the background field. These test-particle results are designed to see how the background field influences the spectral shape of the accelerated population; they do not include shock smoothing or any feedback between particles and waves.

The main conclusion Niemiec et al. (2006) reach is that, for a wide variation of magnetic field structures, spectra are softer than the “universal” \( f(p) \propto p^{-4.2} \) power law and show cutoffs at energies below the maximum resonance energy determined by the background turbulence. They call
Figure 4.14 Top panel shows the shock structure for Model H, matching the Summerlin and Baring (2012) example discussed in the text. The dashed (red) curve is the UM flow speed \( u(x) \), the solid (black) curve is the self-consistent flow speed, and the dot-dashed (blue) curve is \( \gamma(x) \beta(x) / (\gamma_0 \beta_0) \) for the NL shock. In the middle and bottom panels, the dashed (red) curve is the UM flux, and the solid (black) curve is the NL flux. The momentum and energy fluxes are conserved to within a few percent in the self-consistent shock.
into question the ability of DSA to produce the radiating electrons seen in superluminal relativistic shock sources such as extra-galactic radio jets and GRB afterglows, or for ultra-relativistic DSA to be the mechanism to produce UHECRs (see Lemoine and Revenu, 2006, who reach somewhat different conclusions, for similar Monte Carlo work on perpendicular shocks).

Superluminal shocks require strong cross-field diffusion for DSA to occur. In an oblique shock, the deHoffmann–Teller (HT) frame is the frame where the $\mathbf{u} \times \mathbf{B}$ electric field is zero. This frame moves along the shock front at a speed $v_{\text{HT}} = \gamma_0 u_0 \tan \theta_{\text{Bn}}$, and shocks with $v_{\text{HT}} > c$ are superluminal. A downstream particle tied to a magnetic field line cannot re-cross the shock and be diffusively accelerated when $v_{\text{HT}} > c$. Ultra-relativistic shocks will be superluminal unless $\theta_{\text{Bn}} < 1/\gamma_0$. For $\gamma_0 = 100$, $\theta_{\text{HT}} \equiv \tan^{-1}[1/(\gamma_0 \beta_0)] \approx 0.6^\circ$, essentially precluding DSA in plasmas where the turbulence is too weak to produce strong cross-field diffusion.\(^{10}\) For $\gamma_0 = 1.5$, $\theta_{\text{HT}} \approx 42^\circ$ and far more phase

\(^{10}\)See the work of Sironi et al. (2013) on superluminal low-magnetization shocks and their ability to self-generate...
4.5. DISCUSSION

4.5.1 Discussion of Chapter 4: Transrelativistic Smoothed Shocks

space is open to DSA even without cross-field diffusion. As we have emphasized above, all of the factors that limit the efficiency of DSA in ultra-relativistic shocks, including the wave structures investigated, for example, by Niemiec et al. (2006), become less important as the shock slows to trans-relativistic speeds.

4.5.2 Comparison with PIC Results

The most compelling reason to believe that ultra-relativistic shocks do, in fact, accelerate particles beyond the initial kinematic boost comes from PIC simulations (e.g., Hoshino et al., 1992; Kato, 2007; Sironi and Spitkovsky, 2009a, 2011; Keshet et al., 2009). These simulations clearly show that acceleration occurs under some circumstances with characteristics similar to those expected from DSA. Keeping in mind the fundamental differences in how the wave–particle interactions and acceleration are treated in a PIC simulation versus a Monte Carlo simulation, we qualitatively compare our results to the $\gamma_0 = 15, \theta = 15^\circ$ example shown in Sironi and Spitkovsky (2011). Here, $\theta$ is the angle between the far upstream magnetic field and the shock normal in the simulation (i.e., wall) frame of the PIC simulation. This obliquity is subluminal for $\gamma_0 = 15$, making a comparison to our parallel shock results meaningful.

In Figure 4.16 we show Monte Carlo results for the structure of an NL shock with parameters chosen to produce spectra that can be compared to those in Figure 2 or Figure 7 in Sironi and Spitkovsky (2011). In order to obtain spectra with the low cutoff energy obtained in Sironi and Spitkovsky (2011), a short shock precursor is required. We set the FEB at $x = -3r_{g0}$, and the dot-dashed $\gamma\beta/(\gamma_0\beta_0)$ curve shows that escaping particles influence the incoming flow upstream of $x = -3r_{g0}$. The vertical dotted line shows the position of $x_{\text{sub}} = -4r_{g0}$ (i.e., Equation 4.1). For comparison, the FEB is at $x = -1000r_{g0}$ for our $\gamma_0 = 10$ Model D (Figures 4.6 and 4.9).

In the top panel of Figure 4.17 we show Monte Carlo spectra measured downstream from the subshock (dashed, blue curve) and at two upstream positions: $x = -2r_{g0}$ for the dotted (red) curve and $x = -2.5r_{g0}$ for the solid (black) curve. Considering the different nature of the simulations, we have not attempted to precisely match parameters for this illustration. Nevertheless, the spectra in Figure 4.17 are in general qualitative agreement with the proton spectra shown in Figures 2(g), (h), and (i) of Sironi and Spitkovsky (2011).\textsuperscript{11}

Common features between the PIC and Monte Carlo spectra are the following: (1) there is a broad thermal peak for the downstream spectrum with much sharper thermal peaks for the turbulence.

\textsuperscript{11}Note that our spectra are calculated in the shock frame and the Sironi and Spitkovsky (2011) spectra are calculated in the “wall” or downstream frame. This does not make a large difference since the downstream flow Lorentz factor is $\sim 1$ in the shock frame.
4.5. DISCUSSION

CHAPTER 4. TRANSRELATIVISTIC SMOOTHED SHOCKS

Figure 4.16 Nonlinear structure of a $\gamma_0 = 15$ shock with a FEB at $x = -3r_g$ which conserves mass, momentum, and energy flux (Model F). The vertical dotted line shows the position of $x_{\text{sub}} = -4r_g$.

upstream spectra; (2) the downstream thermal peak is at $\gamma \approx 10$ while the upstream peaks occur at $\gamma \approx 15$ (if measured in the plasma frame, these upstream peaks would occur at $\sim \gamma^2$ here and in the PIC simulation); (3) a relatively smooth transition can be seen in the downstream spectrum, from thermal to superthermal energies, as thermal particles are injected into the acceleration mechanism; (4) a cutoff occurs in the superthermal spectrum starting below $\gamma \sim 10^3$; and (5) the acceleration efficiency (shown in the bottom panel of Figure 4.17 for the Monte Carlo shock) is approximately 30%.

Important differences follow: (1) the downstream Monte Carlo spectrum is softer than the PIC one and doesn't flatten to $dN/d\gamma \propto \gamma^{-2.1}$ before the cutoff; and (2) the spectral cutoff in the Monte Carlo result is broader than the PIC cutoff and occurs for a different reason (the PIC spectrum cuts off because of a finite run time, while the Monte Carlo spectrum cuts off because we have imposed a FEB at $x = -3r_g$ to produce a cutoff at an energy similar to that obtained in the PIC result).

Shock modification by the backpressure of accelerated particles is also clearly present in both cases, as seen in our Figure 4.16 and the top panel of Figure 4 in Sironi and Spitkovsky (2011). Of course, the scaling of the shock structure, which depends on the magnetic field, is quite different. The self-consistent field in the PIC result is highly turbulent, time dependent, and varies strongly with location. It also includes magnetic field amplification, presumably by Bell’s instability (e.g.,

12The density ratio in Figure 4 of Sironi and Spitkovsky (2011) is $\sim 4$ because the densities in the plot are measured in the wall (i.e., downstream) rest frame. When transformed to the shock frame, the density ratio is $\sim 3$ as expected.
4.5. DISCUSSION

CHAPTER 4. TRANSRELATIVISTIC SMOOTHED SHOCKS

Figure 4.17 Top panel shows the omni-directional, shock-frame spectra for a NL $\gamma_0 = 15$ shock with a FEB at $x = -3r_{g0}$ (Model G). The solid (black) curve is measured at $x = -2.5r_{g0}$, the dotted (red) curve is measured at $x = -2r_{g0}$, and the dashed (blue) curve is measured downstream from the subshock at $x = 5r_{g0}$. The bottom panel shows the acceleration efficiency, and the horizontal dashed line marks the efficiency found by Sironi and Spitkovsky (2011) using a PIC simulation for a shock with similar parameters.
Bell, 2004, 2005). In the Monte Carlo simulation Bohm diffusion is assumed and the field is uniform and set to $B_0 = 3\mu\text{G}$. Nevertheless, shock modification must occur to satisfy energy and momentum conservation, and the fact that the PIC and Monte Carlo shocks have similar acceleration efficiencies means the gross features of the shock structure will be similar.

Shock reformation is another notable aspect of the PIC results that is not modeled by our steady-state Monte Carlo simulation. In the time-dependent PIC simulations, bunches of downstream ions can return back upstream, causing the subshock to disrupt and reform in a quasi-periodic fashion. While this process might be important for determining local particle reflection and other injection effects, the relatively narrow temporal and spatial averages applied by Sironi and Spitkovsky (2011) in their Figures 4 and 7, for example, seem to smooth the density profiles and energy spectra enough that long-term effects from reformation are no longer evident. This suggests that our steady-state Monte Carlo results can be meaningfully compared to averages taken during the evolving PIC simulations.

### 4.5.3 GRB Afterglows

The radiation observed from GRB afterglows is generally modeled as synchrotron emission from relativistic electrons, where the electrons are accelerated by the ultra-relativistic but decelerating fireball shock. A further assumption that is generally made is that the electrons obtain a power-law distribution with some minimum Lorentz factor that emerges from DSA (e.g., Mészáros, 2002; Piran, 2004; Leventis et al., 2013). Equally constraining, the shape of the electron distribution is often assumed to remain constant during the evolution of the blast wave as it slows in the circumstellar medium (CSM). As the PIC and Monte Carlo results we have discussed show, ultra-relativistic shock theory has difficulty supporting these widely used assumptions.

The ejecta moving relativistically from the central engine in a GRB will drive a forward shock (FS) into the CSM, and a reverse shock (RS) will propagate through the ejecta material (e.g., Piran, 2004, and references therein). Initially, the FS will be ultra-relativistic, but the RS may initially be nonrelativistic.

In the case of matter-dominated ejecta with a density $\Delta$ times the density of the CSM, the Lorentz factor of the RS, $\gamma_{RS}$, in the rest frame of the ejecta, can be estimated as

$$\gamma_{RS} \approx \left\{ \begin{array}{ll}
1 + \frac{4}{7} \gamma_{ej}^2 \Delta^{-1} & \gamma_{ej}^2 \ll \Delta \\
\frac{1}{\sqrt{2}} \gamma_{ej}^{1/2} \Delta^{-1/4} & \gamma_{ej}^2 \gg \Delta,
\end{array} \right. (4.2)$$

where $\gamma_{ej}$ is the Lorentz factor of the ejecta (see Piran, 2004). In the early expansion stage, when the ejecta density is large and $\gamma_{ej}^2 \ll \Delta$, the RS is nonrelativistic. It will become mildly relativistic when
4.5. DISCUSSION

While the energy in the RS is much less than in the ultra-relativistic FS, we have shown (e.g., Figure 4.10) that DSA is more efficient in trans-relativistic shocks and the spectrum produced is harder than for ultra-relativistic shocks.

In the case of GRBs driven by ejecta that are initially magnetically dominated (e.g., Lyutikov and Blandford, 2003; Zhang and Yan, 2011), a trans-relativistic RS may form during the afterglow stage if efficient magnetic dissipation occurs during the prompt emission regime (e.g., Zhang and Yan, 2011; Bykov and Treumann, 2011; Bykov et al., 2012). The main reason for considering a trans-relativistic RS, rather than the ultra-relativistic FS, which carries far more energy, as the source of the afterglow is the conceptual difficulties superluminal ultra-relativistic shocks have in diffusively injecting and accelerating particles to energies needed to model GRB afterglows (e.g., Pelletier et al., 2009). Whatever difficulties ultra-relativistic shocks have should be less for trans-relativistic shocks. Furthermore, our parallel-shock approximation is more applicable to subluminal trans-relativistic shocks.

4.5.4 Trans-relativistic Supernovae and UHECRs

Recently, mildly relativistic Type Ibc SNe have been discovered (e.g., Soderberg et al., 2010; Paragi et al., 2010; Chakraborti et al., 2011) with inferred blast-wave speeds such that \( \gamma_0 \beta_0 \sim 1 \). These speeds lie between normal supernova remnants (SNRs) with \( 10^{-3} \lesssim \gamma_0 \beta_0 \lesssim 0.1 \) and GRB jets or fireballs with \( \gamma_0 \beta_0 \gtrsim 5 \) and suggest that there may be a continuum of shock speeds from the explosions of massive stars over the full range from nonrelativistic to ultra-relativistic (see, for example, Lazzati et al., 2012).

Our trans-relativistic results are important for the interpretation of emission from shock-accelerated particles in this particular class of mildly relativistic Type Ibc SNe. We showed in Figure 4.8 that phase-space spectra \( f(p) \propto p^{-4} \) can be expected in a transitional zone around \( \gamma_0 \sim 1.5 \) (\( \gamma_0 \beta_0 \sim 1 \)). Perhaps more importantly, however, all relativistic blast-wave shocks will slow as they sweep up material, and if DSA is working as modeled in Figure 4.10, relativistic shocks will surely pass through a transitional phase.

Budnik et al. (2008) have suggested that galactic trans-relativistic SNe may produce CRs with energies in the range \( \sim 10^{15} \) to \( \sim 10^{18} \) eV. An even more dramatic assertion by Chakraborti et al. (2011) is that the sub-population of Type Ibc SNe such as SN 2009bb may be able to produce UHECRs to the Greisen–Zatsepin–Kuzmin (GZK) limit.

In our steady-state simulation, for a given set of shock parameters including \( \eta_{\text{mfp}} \) and \( N_e \), the maximum energy a shock can produce, \( E_{\text{max}} \), is proportional to the distance to the FEB, \( L_{\text{FEB}} \). Of course, in our plane-shock model, the FEB mimics the effects of a finite shock where \( L_{\text{FEB}} \)
Figure 4.18 Structure for a $\gamma_0 = 1.5$ shock with a FEB at $-10^8 r_{g0}$ (Model I). The figure format is the same as Figure 4.5. The self-consistent compression ratio is $r_{NL} = 4.0 \pm 0.4$, compared to the Rankine–Hugoniot value $r_{RH} \approx 3.53$. 

\[ r_{NL} = 4.0 \pm 0.4 \]

\[ r_{RH} \approx 3.53 \]
Figure 4.19 Comparison of UM downstream spectrum (dashed, red curve) for our Model I with $L_{\text{FEB}} = -10^8 r_{g0}$ against the NL spectrum (solid, black curve).

corresponds to the scale set by the shock radius or system size, an approximation that has been used to good effect in models of young nonrelativistic SNRs (e.g., Lee et al., 2013). In Figures 4.18 and 4.19 (Model I), we show results for a $\gamma_0 = 1.5$ shock where $L_{\text{FEB}} = -10^8 r_{g0} \sim 0.25$ pc and $B_0 = 300 \mu$G. All other parameters are the same as those used for Model C shown in Figures 4.5 and 4.8. $L_{\text{FEB}}$ and $B_0$ have been chosen to obtain $E_{\text{max}} \sim 10^{18}$ eV before the turnover from CRs escaping at the FEB. The large $B_0$ assumes that this $\gamma_0 = 1.5$ shock can produce strong magnetic field amplification or that it exists in a region with large ambient fields. While the evidence supporting strong magnetic field amplification (MFA) in nonrelativistic shocks is convincing (e.g., Cassam-Chenaï et al., 2007), little MFA is expected in ultra-relativistic shocks.\footnote{See, however, the discussion in Santana et al. (2014) about the need for magnetic field amplification in GRB afterglows to explain the observed X-ray and optical flux at early times.} However, since MFA is driven by CR production, we expect the spectral transition seen in Figure 4.10 to be accompanied by MFA as the efficiency of DSA increases and the shock speed slows.
4.6 Conclusions

As indicated by the dot-dashed (blue) curve in the top panels of Figure 4.18, the shock precursor is noticeably smoothed over the entire region from $L_{\text{FEB}}$ to the subshock. In Figure 4.19, the concave shape of the NL spectrum (solid, black curve) is obvious, but, to a fair approximation, the spectrum is harder than the UM one and $f(p) \propto p^{-4}$ over a wide momentum range. Of course, we have made the assumption that Bohm diffusion, with $B_0 \sim 300\mu G$, occurs over the scale indicated in Figure 4.18 (e.g., Milosavljević and Nakar, 2006). If this in fact happens, trans-relativistic shocks might be able to produce CRs to $\sim 10^{18}$ eV.

Another possible source of UHECRs involving DSA in transrelativistic shocks are jets in active galactic nuclei (AGNs; see, for example, Kotera and Olinto, 2011; Aharonian et al., 2012, for reviews). Shock speeds of $\beta_0 \sim 0.3$ have been inferred from multi-waveband observations in a number of radio hot spots (e.g., Meisenheimer et al., 1989), and Romero et al. (1996) suggest that the trans-relativistic jet in Centaurus A (NGC5128), the closest AGN, may accelerate CRs to $E_{\text{max}} > 10^{21}$ eV. For this $E_{\text{max}}$, a shock size of $\sim 2$ kpc and a field of $\sim 10 \mu G$ were assumed. Lemoine and Waxman (2009) established that a jet magnetic luminosity of $L_B \gtrsim 10^{45} Z^{-2}$ erg s$^{-1}$ is needed in order to accelerate particles of charge number $Z$ up to 100 EeV. Since $L_B$ for Centaurus A is estimated to be $\gtrsim 10^{43}$ erg s$^{-1}$, the production of CRs to $\sim 10^{21}$ eV may be possible.

4.6 Conclusions

If collisionless shocks accelerate particles efficiently, the shock structure must be calculated self-consistently with the acceleration and particle escape to conserve energy and momentum. We have presented a Monte Carlo model of NL DSA that, once the scattering properties are defined, determines the thermal injection and shock structure in plane-parallel, steady-state shocks of arbitrary Lorentz factor. Our injection model, parametrized diffusion, and plane-parallel, steady-state approximations are restrictive when compared to PIC simulations and some Monte Carlo calculations that trace particle orbits. However, apart from PIC simulations and the preliminary work of Ellison and Double (2002), they are the first to model DSA in NL relativistic shocks. NL effects may be critical for understanding GRB afterglows, blazars (e.g., Stecker et al., 2007), radio jets (e.g., Wykes et al., 2013), and DSA in a newly discovered class of mildly relativistic Type Ibc SNe (e.g., Budnik et al., 2008; Soderberg et al., 2010). Next is a list of our major findings.

1. In terms of the backreaction of CRs on the shock structure, NL DSA works identically in nonrelativistic and relativistic shocks, since the effects we consider depend only on energy and momentum conservation. Quantitative differences result from the fact that relativistic shocks tend to have a lower intrinsic compression ratio than nonrelativistic shocks and because
4.6. CONCLUSIONS

Particle distributions are highly anisotropic in relativistic shocks. Besides making the analysis of NL DSA more difficult, this anisotropy makes relativistic shocks much more sensitive to assumptions made regarding particle diffusion, and, unlike in nonrelativistic shocks, these assumptions can strongly influence the acceleration process. We do not specify the production mechanism of self-generated magnetic turbulence here but rather assume that Bohm diffusion applies. There are fundamental problems in how, or even if, this turbulence is created which depend on shock speed and obliquity (see, for example, Plotnikov et al., 2013, and references therein).

2. There is a smooth transition between the concave upward spectral shape long established for nonrelativistic shocks undergoing efficient DSA and the soft power-law spectrum predicted for fully relativistic shocks (i.e., Figure 4.10). For the parameters used here, this transition occurs around $\gamma_0 = 1.5$, a speed recently associated with a subclass of Type Ibc SNe (e.g., Soderberg et al., 2010). An increase in MFA can also be expected as shocks slow and cross this transition zone.

3. Even though NL effects from particle acceleration are weaker in fully relativistic shocks than in nonrelativistic ones, they nevertheless have a substantial effect on the accelerated particle spectrum. Our NL results for $\gamma_0 = 10$, in the fine-scattering limit, show the normalization of the energetic particle distribution dropping by a factor of about five compared to the UM shock (see Figure 4.9). While these specific results depend quantitatively on the details of our model, basic energy and momentum conservation demands similar backreaction effects if DSA is efficient.

4. If LAS is assumed in relativistic shocks, particles can gain a much larger amount of energy in a given shock crossing than for fine pitch-angle scattering. While this suggests that these shocks can be extremely efficient accelerators with hard spectra, NL shock smoothing reduces the resultant efficiency and softens the spectrum considerably (see Figures 4.13 and 4.15).

5. Not forgetting the fundamental differences between our Monte Carlo technique and PIC simulations, we have compared our results with those of Sironi and Spitkovsky (2011) and shown that important similarities exist in the spectral shape, average shock structure, and overall acceleration efficiency determined by each technique for a quasi-parallel, subluminal shock (see Figures 4.16 and 4.17). Our results also suggest that short-timescale shock reformation effects can be adequately modeled with a steady-state simulation.

6. The major reason for employing the Monte Carlo technique, when far more physically complete PIC simulations exist, is indicated by Figures 4.18 and 4.19. It has been suggested that
trans-relativistic shocks may be able to produce CRs to $10^{18}$ eV and above. In order to model such acceleration, large spatial scales and timescales are required, which remain well beyond PIC limits. To obtain $E_{\text{max}} \sim 10^{18}$ eV, we needed $L_{\text{FEB}} = -10^8 r_{g0}$ for our $\gamma_0 = 1.5$ shock, compared to $L_{\text{FEB}} = -3 r_{g0}$ used for our $\gamma_0 = 15$ comparison to the Sironi and Spitkovsky (2011) results (Model G). After adjusting for the different magnetic fields and shock speeds used in our length unit $r_{g0}$, this is a factor of $\sim 2.5 \times 10^5$ difference in real units, a factor easily accommodated by the Monte Carlo technique.

The most important approximations made by the Monte Carlo technique are as follows:

1. The particle scattering is parameterized using $\eta_{\text{mfp}}$ (Equation 3.12) and $N_g$ (Equation 3.13). For this study, we only show results for Bohm diffusion, i.e., $\eta_{\text{mfp}} = 1$, although our plane-parallel results are independent of $\eta_{\text{mfp}}$ if the one length scale in the problem, the FEB, is scaled in particle gyroradii. Our scattering parameterization has no spatial dependence and does not include magnetic field amplification, a process almost certain to be important in strong, nonrelativistic shocks (e.g., Bell, 2004; Vladimirov et al., 2008; Ellison et al., 2012), and one that has been invoked to explain UHECR production in trans-relativistic shocks (e.g., Chakrabarti et al., 2011). We note that the high acceleration efficiencies we find, and the related concave spectral shapes, may be less extreme in oblique shocks where values of $\eta_{\text{mfp}} > 1$. Difficulties in producing strong turbulence in ultra-relativistic, superluminal shocks may also result in larger effective $\eta_{\text{mfp}}$ values, reducing NL effects from DSA and resulting in lower cutoff energies in any given shock environment. In the critical trans-relativistic, subluminal regime, however, effects from shock obliquity are less restrictive and strong turbulence is more likely to be produced, making our Bohm diffusion approximation more plausible.

2. The damping of magnetic turbulence is not included. The strong NL effects we see will be lessened if the shock precursor is heated as energy is transferred from the turbulence to the background plasma.

3. We only model plane, parallel shocks, while relativistic shocks are expected to be oblique. Based on PIC simulation results (e.g., Sironi and Spitkovsky, 2011), the consequences of this restriction should be minor as long as the shock remains subluminal. However, important relativistic shock applications, such as the termination shock in a PWN, are likely to be strongly superluminal. While oblique, relativistic, test-particle shocks have been studied with Monte Carlo techniques (e.g., Ellison and Double, 2004; Summerlin and Baring, 2012), NL effects have not yet been modeled.
4.6. CONCLUSIONS

4. We have only modeled protons, while applications involving radiation signatures of relativistic shocks require electrons to be accelerated self-consistently with protons. Electrons can be included in the Monte Carlo model (e.g., Baring et al., 1999), but, unlike self-consistent PIC results, this requires additional assumptions and free parameters. Electrons will be considered in future work.

5. Our Monte Carlo model is steady-state, while most astrophysical shocks are in evolving systems. However, as suggested by Figure 4.10, an evolving system can be approximated with several steady-state shocks, each differing in parameters such as radius, speed, ambient density, and magnetic field strength.

DSA is the most intensively investigated acceleration mechanism in astrophysics for good reason. Collisionless shocks are common, some nonrelativistic shocks are known from \textit{in situ} spacecraft observations to be efficient accelerators with self-generated magnetic turbulence, and detailed predictions from nonrelativistic DSA theory provide excellent fits to some nonthermal sources (see Lee et al., 2012a, for a review). From its introduction in 1976–1978, the theory of DSA developed rapidly, benefiting from the combined input from direct spacecraft observations of heliospheric shocks, analytic studies, and computer simulations. The extension of DSA theory to ultra-relativistic shocks is a natural progression, but definitive results have been slowed by the inherent difficulty associated with magnetic turbulence generation and particle transport in relativistic plasmas, and by the fact that no relativistic shocks are directly observable by spacecraft.

What is certain is that there must be a smooth transition between efficient, nonrelativistic shocks, where observational evidence suggests that DSA produces strong MFA in the shock precursor, and ultra-relativistic shocks, which may be far less efficient accelerators and where self-generated turbulence may be weak or absent. The Monte Carlo model we present here provides important information on the critical trans-relativistic regime, albeit for a simple parameterized scattering model. More physically motivated diffusion has been included in nonrelativistic versions of the code (e.g., Vladimirov et al., 2008, 2009), and, in principle, similar generalizations can be made to the relativistic code.
The previous chapter involved the acceleration of protons—and only protons—at shock fronts. The Monte Carlo approach outlined in Chapter 3 is just as applicable to other species of particle, however. Here I discuss extending the simulations to include the next two most common cosmic ray species other than protons: electrons and $\alpha$ particles (a.k.a. helium nuclei). The results presented here have been submitted to Monthly Notices of the Royal Astronomical Society for publication prior to the finalization of this dissertation. Note that, while the Monte Carlo code could handle additional particle species—in particular neutrons and positrons—we do not examine those possibilities in this work.

5.1 Abstract

We have modeled the simultaneous diffusive shock acceleration (DSA) of protons, electrons, and helium nuclei by relativistic shocks. By parametrizing the particle diffusion, our steady-state Monte Carlo simulation allows us to follow particles from particle injection at nonrelativistic thermal energies to above PeV energies, including the nonlinear smoothing of the shock structure due to
5.2 Introduction

Efficient diffusive shock acceleration (DSA) is often suggested as a likely mechanism for converting the bulk kinetic energy of relativistic plasma flows into individual particle energy (e.g., Bykov and Treumann, 2011; Bykov et al., 2012). However, many aspects of particle acceleration in relativistic shocks remain uncertain because of the inherent complexity of the process. The particle distributions are highly anisotropic and the magnetic turbulence, essential for acceleration to occur, must be self-generated and is extremely difficult to characterize (Lemoine and Pelletier, 2003; Niemiec and Ostrowski, 2006; Lemoine and Pelletier, 2010; Reville and Bell, 2014). These roadblocks can be overcome with particle-in-cell (PIC) simulations and intensive work has been done in this area (e.g., Nishikawa et al., 2007; Sironi and Spitkovsky, 2011). However, current PIC simulations are computationally costly and have a limited dynamic range. The transrelativistic regime, which may be important for GRB afterglows (e.g., Mészáros, 2006; Ackermann et al., 2013) and some types of supernovae (e.g., Chakraborti et al., 2011), is largely unexplored either analytically or using PIC simulations.

In this paper we model the nonlinear acceleration of electrons and ions (protons and He$^{2+}$) at relativistic collisionless shocks using a Monte Carlo simulation of DSA (e.g., Ellison and Double, 2002; Ellison, Warren, and Bykov, 2013). The steady-state Monte Carlo simulation parameterizes magnetic turbulence generation and particle diffusion, important approximations but ones that allow a large dynamic range; the simultaneous acceleration of ions and electrons (along with the radiation they produce); and a self-consistent determination of the shock structure. We believe this is the first attempt, apart from PIC simulations, to include electrons self-consistently with ions in a nonlinear, relativistic, DSA calculation.

While it is well known that nonlinear DSA should preferentially inject and accelerate high mass-to-charge particles compared to protons in nonrelativistic shocks (e.g., Ellison et al., 1981; Eichler,
1984; Jones and Ellison, 1991), to our knowledge this process has not been investigated in relativistic or transrelativistic shocks until now. The nonrelativistic “$A/Z$” effect ($A$ is the mass in units of the proton mass $m_p$ and $Z$ is the charge in units of the electron charge $e$) has been shown to be consistent with observations of diffuse ions accelerated at the Earth bow shock (Ellison et al., 1990a) and has been used to model the shock acceleration of interstellar gas and dust, matching important aspects of the galactic cosmic-ray abundances observed at Earth (e.g., Ellison et al., 1997; Meyer et al., 1997; Meyer and Ellison, 1999; Rauch et al., 2010; Binns et al., 2013).

The $A/Z$ enhancement we model is a purely kinematic effect. It depends on the following assumptions: (i) the acceleration process is efficient enough so the backpressure from accelerated particles noticeably modifies the shock precursor; (ii) all particles have a scattering mean-free-path, $\lambda_{\text{mfp}}$, of the form

$$\lambda_{\text{mfp}} = \eta_{\text{mfp}} f_{\text{mfp}},$$

where $f_{\text{mfp}}$ is an increasing function of the local frame momentum $p$, as is generally assumed; and (iii) the parameter $\eta_{\text{mfp}}$ is similar for all particle species.

The simplest assumption for particle diffusion is that $f_{\text{mfp}}$ equals the gyroradius, i.e., $f_{\text{mfp}} = r_g = \frac{pc}{(ZeB)}$, where $p$ is the local frame momentum, $c$ is the speed of light, $Z$ is the particle charge in units of the electron charge $e$, and $B$ is the background magnetic field strength in Gauss used to scale $r_g$. Then

$$\lambda_{\text{mfp}} = \eta_{\text{mfp}} r_g,$$

and the precursor diffusion length for species $i$ is

$$L_{D}^{i} \propto \eta_{\text{mfp}} \gamma_i (A/Z) v_i^2.$$  

Here $\gamma_i (v_i)$ is the Lorentz factor (velocity) for species $i$. If $v_i \sim c$, then $\gamma \propto p/A$,

$$L_{D}^{i} \propto \eta_{\text{mfp}} (A/Z) (p/A),$$

and the precursor diffusion length for different $A/Z$ ions, at the same momentum per nucleon, scales as $A/Z$.

Therefore, if $\eta_{\text{mfp}}$ is similar for all species, at the same $p/A$, high $A/Z$ particles will diffuse further into the upstream region, “feel” a larger effective compression ratio in the modified shock structure than low $A/Z$ particles, and gain a larger momentum boost in the next shock crossing. Since $L_{D}^{i}$ increases with $p/A$, the modified shock produces a distinctive concave spectral shape (e.g., Ellison and Eichler, 1984) with high $A/Z$ particles having a harder spectrum at any $p/A$ than low $A/Z$.
particles (see Figure 5.15 below).

While this $A/Z$ effect can enhance the acceleration of heavy ions compared to protons, for electrons with $A/Z \approx 5.45 \times 10^{-4}$ it acts strongly in the opposite fashion. If only this kinematic effect is considered with Equation 5.2, electrons will be dramatically less efficiently injected and accelerated than protons in nonlinear DSA.

Electron injection was considered in a nonrelativistic Monte Carlo code similar to the one we use here in Baring et al. (1999). In that paper, in order to overcome the dramatic $A/Z$ effect and allow electrons to be injected and accelerated with efficiencies large enough to be consistent with synchrotron and IC radiation observed in young supernova remnants (SNRs), the electron $\lambda_{mfp}$ was set equal to a constant below some momentum, i.e., changing $\eta_{mfp}$ selectively for electrons. This modification gave low energy electrons a larger mean free path than Equation 5.2 would produce and allowed them to diffuse far enough upstream to overcome the shock smoothing effects. It was argued in Baring et al. (1999) that this simple modification was reasonably consistent with an electron injection model developed by Levinson (1992).

Here, we adopt a different approach. We keep Equation 5.2, but transfer some fraction of the ram kinetic energy from ions to electrons as the particles first cross the viscous subshock. We note that a “sharing” of energy between ions and electrons is clearly seen in recent PIC relativistic shock simulations (i.e., Sironi and Spitkovsky, 2011; Sironi et al., 2013), and Plotnikov, Pelletier, and Lemoine (2013) give an analytical treatment of electromagnetic instabilities transferring energy to electrons in the precursor of relativistic shocks.

In the PIC simulations, electrons are heated in the precursor by interacting with turbulence generated mainly by backstreaming protons, and obtain near equipartition with the protons before crossing the subshock. We mimic this effect by transferring a set fraction, $f_{ion}$, of ion energy to electrons as particles first cross the subshock. While this simple energy transfer model is clearly an approximation, we feel it affords a straightforward way of using plasma physics information obtained from computationally intensive PIC simulations in a calculation that can model particle acceleration and photon emission consistent with the production of high-energy cosmic rays in relativistic shocks.

We recognize, of course, that collisionless shock formation, and particle injection and acceleration, are determined by more than kinematics alone. The self generation of magnetic turbulence is critical to the process (e.g., Bednarz and Ostrowski, 1998; Lemoine and Pelletier, 2003), particularly for relativistic shocks, and, as just mentioned, energy can be transferred between electrons and protons by wave–particle interactions that are essentially independent of the kinematics. Nevertheless, if DSA is efficient, basic momentum and energy conservation demands that kinematics be taken into account, and the shock precursor must be modified by the backpressure of accelerated
5.2. INTRODUCTION

CHAPTER 5. OTHER PARTICLE SPECIES

particles.

In contrast to nonrelativistic shocks, where DSA can be directly tested against spacecraft observations (e.g., Ellison et al., 1990a; Baring et al., 1997), DSA in relativistic shocks is far less certain. Relativistic shocks cannot be directly observed with spacecraft, as can nonrelativistic heliospheric shocks, and the highly anisotropic particle distributions intrinsic to relativistic shocks make the self-generation of magnetic turbulence far more difficult to describe analytically. Despite this difficulty, intensive work continues (e.g., Lemoine and Pelletier, 2010, 2011; Plotnikov et al., 2013). Furthermore, and again in contrast to nonrelativistic shocks, the predictions for particle spectra and, therefore, photon signatures from relativistic shocks are highly uncertain. Sites harboring relativistic shocks, such as GRBs, can often be successfully modeled with alternative acceleration mechanisms (e.g., magnetic reconnection in the case of GRBs; McKinney and Uzdensky, 2012; Sironi et al., 2015), weakening the link between DSA theory and observation.

Despite this uncertainty, there is compelling evidence, primarily from PIC simulations (e.g., Hoshino et al., 1992; Kato, 2007; Sironi and Spitkovsky, 2009a; Keshet et al., 2009; Nishikawa et al., 2011; Sironi and Spitkovsky, 2011; Stockem et al., 2012), that relativistic shocks do accelerate electrons and protons beyond the initial kinematic boost from a single shock crossing. These simulations also highlight the role of the magnetization parameter $\sigma_B$ in DSA, where

$$\sigma_B = \frac{B_0^2}{4\pi n_0 m_p c^2},$$

(5.5)

$B_0$ is the upstream field, and $n_0$ is the number density (see Bykov and Treumann, 2011, for an alternative definition of $\sigma_B$).

Sironi et al. (2013) (see also Haugbolle, 2011) show that perpendicular electron–ion shocks with Lorentz factors $\gamma_0 \lesssim 150$ inject and accelerate electrons and ions efficiently when $\sigma_B \lesssim 3 \times 10^{-5}$. In these weakly magnetized plasmas, the self-generated turbulent field dominates the uniform $B$ and the shock obliquity ceases to be important, an assumption often made in DSA studies and one we make here with our plane-parallel shock assumption (c.f. Figure 5.4). Quantifying this lack of dependence on the obliquity is particularly important since ultrarelativistic shocks are essentially always highly oblique—the $B$-field component along the shock face can be highly compressed—and strongly magnetized oblique shocks are less able to inject and accelerate particles. We note that for typical interstellar medium (ISM) conditions ($B_0 \sim 3 \mu G$, $n_0 \sim 0.03 \text{ cm}^{-3}$) $\sigma_B \sim 10^{-9}$. This value is orders of magnitude below the threshold reported by Sironi et al. (2013), suggesting that ultrarelativistic shocks propagating in the normal ISM may be able to inject and accelerate electrons and ions far more efficiently than previously believed, regardless of the shock obliquity. We note that for nonrelativistic shocks the plasma $\beta \equiv n_0 k_B T_0/(B_0^2/(8\pi))$ is the relevant parameter rather
5.2. INTRODUCTION

CHAPTER 5. OTHER PARTICLE SPECIES

than Equation 5.5 and oblique geometry may be important for typical ISM parameters. Here $k_B$ is Boltzmann’s constant and $T_0$ is the ambient unshocked temperature.

Given that weakly magnetized relativistic shocks can inject particles, the maximum energy these particles obtain in a given shock remains uncertain. Arguments presented by Sironi et al. (2013) suggest that the acceleration reaches a maximum, $E_{\text{max}}$, where

$$\frac{E_{\text{max}}}{\gamma_0 m_p c^2} \sim \sigma_B^{-1/4}. \quad (5.6)$$

This is a substantial energy (for $\sigma_B = 3 \times 10^{-5}$, $E_{\text{max}} \sim 0.2$ TeV), and current simulations are still too small in terms of the particle gyroradius to reach it. In the 2D simulations of Sironi et al. (2013), the maximum transverse box size is $y_{\text{max}} \sim 4000$ cells (i.e., $\sim 100 c/\omega_{\text{pi}}$) and the maximum longitudinal size is $\sim 1.8 \times 10^5$ cells (i.e., $\sim 4600 c/\omega_{\text{pi}}$). The maximum dimensions for 3D runs are considerably smaller. Here, $\omega_{\text{pi}} = [4\pi e^2 n_e/(\gamma_0 m_p)]^{1/2}$ is the proton plasma frequency and $n_e$ is the electron number density. While the longitudinal box size is continually expanded to contain all of the reflected particles, periodic boundary conditions are imposed on the transverse dimensions. If we take $n_0 = 1 \text{ cm}^{-3}$ and $B_0 = 3000 \mu$G to obtain $\sigma_B \simeq 3 \times 10^{-5}$, we see that

$$\frac{r_{g,\text{max}}}{y_{\text{max}}} \sim 20, \quad (5.7)$$

where $r_{g,\text{max}} = E_{\text{max}}/(e B_0)$ is the gyroradius of a proton with energy $E_{\text{max}}$ as determined by Equation 5.6. The gyroradius is more than an order of magnitude larger than the lateral box size, and it is natural to question if the acceleration is saturating because the periodic boundary conditions prevent turbulence generation on large scales. While it is certain that particles that diffuse laterally will be more easily overtaken by the shock regardless of how turbulent the field is, we believe it is possible that the true $E_{\text{max}}$ may be considerably larger than determined by Sironi et al. (2013). If the lateral box size is critical and larger $E_{\text{max}}$’s are possible, the increase in PIC computational requirements will be substantial. The runtime and memory requirements, as well as the three-dimensional box size, must be increased to match the energy dynamic range. Modeling nonlinear effects from DSA to energies far above those currently obtainable with PIC simulations is the main rationale for our kinematic approach.

Next we describe the generalization of the Monte Carlo code used in Ellison et al. (2013) to include the injection and acceleration of electrons simultaneously with ions. The nonlinear shock structure is calculated including thermal leakage injection and the backreaction from all species. Full particle spectra are determined at various positions relative to the subshock, along with the distributions of
5.3. MODEL

For given values of the ambient density, magnetic field, and background photon field, we calculate the synchrotron emission using Equation (6.7a) in Rybicki and Lightman (1979), the inverse Compton (IC) emission using Equation (9) in Jones (1968), and the pion-decay emission using parameterizations given by Kamae et al. (2006, 2007) and Kelner et al. (2006). For the pion-decay from He$^{2+}$, we use the scaling relation given in Orth and Buffington (1976) and Baring et al. (1999). Once the emission is determined in the local frame it is transformed to the observer (i.e., ISM) frame. The issue of photon production is treated in significantly greater detail in Chapter 6.

We note that radiation losses for electrons are only considered during the acceleration process—once accelerated, the electrons radiate without accounting for further losses. This so-called “thin target” approximation, where the radiation length is assumed to be larger than the region between the upstream and downstream FEBs, is adequate for the steady-state examples given here and can be relaxed in models of evolving GRBs (see Chapter 8).

5.3 Model

The Monte Carlo model used in this work is described in great detail in Chapters 3 and 4. All facets of that discussion apply equally to electrons and heavy ions here. Over the rest of this section I will describe the changes made to the existing code to allow for the results presented here and in the rest of the dissertation. The lone exception is photon production, an important enough topic that it is given a chapter all its own (Chapter 6, which follows this one).

5.3.1 The Energy Distribution Function

In this section I will outline a new (to the Monte Carlo code) approach that was used extensively to obtain the results presented in this chapter and this dissertation. While not strictly related to new particle species, it does allow for relativistically-correct determination of the particle population, in any desired frame, without resorting (as do Equations 3.3–3.6) to any assumption about the isotropy of the distribution.

This new method is the energy distribution function $dn(p)/dp$, related to the phase space distribution discussed in Section 3.3 by

$$
\int \frac{dn(x, p)}{dp} dp = \int f(x, p) \cdot 4\pi p^2 dp,
$$

(5.8)

1A FEB is a position beyond which particles are assumed to decouple from the shock. Any actual shock will be finite in extent and at some point high-energy particles will obtain diffusion lengths comparable to the shock size and stream freely away. Our FEBs model a finite shock size within our steady-state, plane-parallel approximation.
or
\[
\frac{dn(x,p)}{dp} = 4\pi p^2 f(x,p).
\] (5.9)

As I will describe below, the energy distribution function is found first by tracking the phase space distribution function \(f(x,p)\) over the course of an iteration, and then integrating over pitch angle to obtain the particle density as a function of position and total momentum. As well, the phase space distribution may be converted into a different reference frame if desired. The energy distribution function may then be integrated over the volume of a grid zone to obtain the total number of particles in that zone, \(dN/dp\).

I note again that the phase space used in the code reduced from the traditional six-dimensional space, with three dimensions in position and three in momentum. Since the shock profile and all other quantities vary only according to their location on the \(x\)-axis, the three dimensions of position are reduced to one. In addition, the momenta are expressed with just two values instead of three. Under typical circumstances, a particle’s momentum could be specified by \(\langle p_{\text{tot}}, \cos(\theta), \phi \rangle\): the length of the momentum vector \(\vec{p}\), a measure of how aligned with the shock normal \(\vec{p}\) is, and a final number listing its orientation in the \(y-z\) plane. In the parallel-shock scenario, however, the system is symmetric with respect to the \(y-z\) plane, and two numbers are sufficient to specify a particle’s momentum.

Taking advantage of symmetry, then, the phase space distribution used in the code has just three dimensions: \(x\), \(p_{\text{tot}}\), and \(\cos(\theta)\).

**Setup and Tracking**

At initialization the code sets two arrays that identify the boundaries of the phase space distribution’s cells: one for the angular dimension and one for the total momentum. The bulk of the angular dimension is linearly spaced in \(\cos(\theta)\). To maximize fidelity in the upstream direction, the final bin is instead several decades of logarithmically-spaced bins in \(\theta\). This fidelity is necessary since the large shock Lorentz factors strongly beam the particle momentum vectors towards the \(-x\) direction. By contrast, the momentum array is simply logarithmically spaced from (almost) zero to the maximum momentum used in the simulation.

As discussed in Section 3.3, all fluxes are tracked in the shock frame, and the phase space distribution is no different. In the same part of the code where the energy and momentum flux arrays are updated, the phase space distribution is updated every time step for every particle with the following steps:

1. Using the current particle’s momentum vector \(\langle p_{\text{tot},\text{sf}}, \cos(\theta) \rangle\) (where \(\cos(\theta) = p_{x,\text{sf}}/p_{\text{tot},\text{sf}}\) is
defined in the shock frame), the proper bin \((i_{pt}, j_{th})\) in the phase space distribution is identified.

2. The flux-corrected particle weight is calculated (again, see Section 3.3).

3. Each grid zone boundary crossed by the particle during this time step has its \((i_{pt}, j_{th})\) bin incremented by the flux-corrected weight.

Once all particles have propagated through the shock structure and escaped (or been culled) through one of the channels described in Section 3.5, the phase space distribution contains a complete accounting of the density of particles at each grid zone. It can then be converted into \(dN/dp\)—the number of particles per momentum bin—as described below.

**Converting from Phase Space to Energy Distribution**

Converting the phase space distribution \(N(\vec{p})\) (for a given grid zone) into a particle count per momentum bin is trivial for the shock frame. One simply adds the number of particles in each row \(i_{pt}\) of the phase space distribution, eliminating any information about the angular distribution of the particles. With the total number of particles so tabulated, divide by the bin width in linear momentum space to get \(dN/dp\).  

To find \(dN/dp\) for the plasma or ISM frames is significantly more challenging. The challenge is illustrated in Figure 5.1. In orange, the shock-frame bin of the phase space distribution is a quadrilateral whose edges are parallel to the axes. Each corner, representing a momentum vector \(\langle p_{tot}, \cos(\theta) \rangle\) can then be transformed into a different frame using the equations of special relativity. The transformed bin (shown in blue in Figure 5.1) is still a quadrilateral, but the corners are now very unlikely to lie on grid points, and the edges are likely skew to the axes. The total number of particles in this bin, as a Lorentz invariant, has not changed. However, the area of the bin will almost surely differ pre- and post-transformation, making the two densities different.\(^2\)

In order to calculate \(dN/dp\) for the transformed frame, each skew bin must be divided among the momentum bins it overlaps. The situation—and the approach taken in the code—are diagrammed in Figure 5.2. The blue shape (the transformed bin from Figure 5.1) falls into three different momentum bins, with unequal areas overlapping with each of the three. The key feature of the solution is simplification of the quadrilateral by taking the two lower edges and flattening them out. All heights above the bottom edge are kept identical: the two yellow lines have the same height in both the blue (original) and the green (modified) quadrilaterals. Although this approach introduces a kink in one upper edge (and eliminates the kink between the lower edges), it reduces the computational

\(^2\)Fortunately, within a single grid zone the ratio of the two densities is constant, so the conversion introduces an error of at most a constant. This will be corrected by the rescaling later.
5.3. MODEL

CHAPTER 5. OTHER PARTICLE SPECIES

\[ \cos \theta = \frac{p_x}{p_{\text{tot}}} \]

**Figure 5.1** An example of a bin in the shock frame (Bin A), and its transformed shape (Bin A'). The axes and grid points are representative of the phase space distribution for a particular grid zone; there would be one such grid for each zone in the shock structure.

The complexity of finding the height of a region within the shape. This method also ensures that the total area of the shapes is the same, introducing no error in counting particles. The heights of the two yellow lines are known in advance, and the heights of the leftmost and rightmost corners are zero. Everything else is an interpolation between two known values, making areas straightforward to obtain.

Once the transformation procedure is complete, the relative population in each momentum bin is fixed. However, the overall scaling is not necessarily correct. To determine the normalization of the \( \frac{dN}{dp} \) curve, the code uses the shock frame flux, which per conservation must be \( \gamma_0 u_0 n_0 \) at all points in the shock structure. The shock frame crossing time is the length \( \Delta L_{sf} \) in the shock frame (see Figure 3.2) divided by the bulk fluid speed \( u(x) \). Once the surface area of the shock is known, the total area under the \( \frac{dN}{dp} \) curve for a given cell may be calculated:

\[
N(x) = \left( \gamma(x) u(x) n(x) \right) \cdot A \cdot \Delta t_{\text{cross}} \\
= \left( \gamma_0 u_0 n_0 \right) \cdot A \cdot \frac{\Delta L_{sf}}{u(x)}
\]  

This quantity is Lorentz invariant and is used to scale the energy distribution function in all
5.3. MODEL

CHAPTER 5. OTHER PARTICLE SPECIES

Figure 5.2 A visual representation of the projection of the 2-D transformed bin from Figure 5.1 into the 1-D array storing \( \frac{dN}{dp} \). The red arrows show the transformation of four key aspects of the quadrilateral as the base is flattened out for easier calculating. The black dashed lines show the locations of the bin boundaries in \( \frac{dN}{dp} \). They divide Bin \( A'' \) into three areas, which will be assigned to corresponding bins.

reference frames of interest.

5.3.2 Radiative Cooling

In the collisionless environment of a typical astrophysical shock, nuclei like \( \text{H}^+ \) or \( \text{He}^{2+} \) lose energy primarily in infrequent but large chunks due to pion production; since these events are exceedingly rare from the perspective of individual particles,\(^3\) we ignore them while these species are propagating within the shock structure. For lighter particles such as electrons, losses due to both synchrotron and inverse Compton emission occur regularly (constantly, in the case of synchrotron) during propagation. While the photon emission is calculated only at the end of an iteration, using the \( \frac{dN}{dp} \) due to all particles of all species (see Chapter 6 to follow), we include radiative losses for electrons and other light particles encountering the shock.

Synchrotron losses in a magnetic field with energy density \( U_B = B^2/8\pi \) may be written (Rybicki and Lightman, 1979)

\[
\left( \frac{dE_e}{dt} \right)_\text{syn} = \frac{4}{3} \sigma_T c \beta^2 \gamma^2 U_B. \tag{5.12}
\]

In the above equation, \( \sigma_T = 8\pi r_0^2/3 \) is the Thomson cross section, and both \( \beta \) and \( \gamma \) refer to the

\(^3\)For typical densities any given cosmic-ray proton experiences roughly one such loss event every \( 10^{14} \) seconds.
5.3. MODEL

CHAPTER 5. OTHER PARTICLE SPECIES

motion of the electron in the plasma frame. Conveniently, losses due to inverse Compton scattering have the same dependence on energy (or momentum) that synchrotron losses do. Expressing the energy density of the CMB as an equivalent magnetic field,

\[ U_{B,\text{eff}} = \frac{B_{\text{eff}}^2}{8\pi} = \frac{4\sigma}{c} T^4, \tag{5.13} \]

the corresponding emitted power via inverse Compton scattering is

\[ \left( \frac{dE_e}{dt} \right)_{IC} = \frac{4}{3} \sigma_T c \beta^2 \gamma^2 U_{B,\text{eff}}. \tag{5.14} \]

Equation 5.14 is exact below the Klein-Nishina regime, which covers the vast majority of electrons in the simulations. Since both synchrotron and inverse Compton losses depend on the same power of momentum, they are treated simultaneously while electrons are propagating within the shock structure.

5.3.3 Energy Transfer

The Monte Carlo model used in this work assumes that the scattering described in Section 3.4.1 applies to protons, helium nuclei, and electrons. This is sufficient to generate the well-known $A/Z$ enhancement in injection rate and acceleration efficiency (e.g., Ellison et al., 1981; Eichler, 1984; Ellison et al., 1990a, 1997). However, PIC simulations of relativistic shocks (Sironi and Spitkovsky, 2011; Sironi et al., 2013) show significant energy transfer from ions to electrons via wave-particle interactions, a process we do not attempt to model self-consistently. Instead, we introduce an additional parameter $f_{\text{trans}}$ to our model, representing the fraction of far upstream ion kinetic energy that is transferred to electrons.

The ion kinetic ram energy, as measured in the shock frame, is defined as $(\gamma_0 - 1)m_i c^2$, where $m_i$ is the mass of the ion species under consideration, and $\gamma_0$ is the Lorentz factor describing the shock’s expansion into the ISM. When an ion crosses the subshock from upstream to downstream the first time in the simulation, an amount of energy $f_{\text{trans}}(\gamma_0 - 1)m_i c^2$ is subtracted and stored in a running total. After all ion species have been propagated through the shock, the total energy accumulated is $f_{\text{trans}} \sum_i n_i (\gamma_0 - 1)m_i c^2$, summing over the ion species injected far upstream from the shock. This energy is divided equally among the population of electrons, being added to their energy when they cross the subshock for the first time.
5.4 Results

We approximate the geometry of a relativistic afterglow shock moving in a jet as shown in Figure 5.3. A detail of this situation is shown in Figure 5.4, where we show the additional approximations for this preliminary work that the shock is locally plane, the distances to the upstream or downstream FEBs ($L_{\text{UpS}}$ and $L_{\text{DwS}}$) measured from the subshock are small enough so the jet cone is approximately a cylinder in the region surrounding the shock, and the diameter of the “cylinder” is large enough so particle escape out the sides of the cone is negligible. The cone material outside the upstream FEB is assumed to be stationary, i.e., it is in the local ISM or explosion frame. Outside of the region between $L_{\text{UpS}}$ and $L_{\text{DwS}}$ we assume all accelerated particles have decoupled from the plasma, and we ignore any emission they might produce. For a more realistic GRB afterglow model following the evolution of a jet shock, see Chapter 8.

With these approximations we calculate the shock structure and particle spectra for a given set of parameters, as listed in Table 5.1. First we consider unmodified (UM) shocks where the backreaction of the accelerated particles on the shock structure is ignored.
5.4. RESULTS

CHAPTER 5. OTHER PARTICLE SPECIES

Figure 5.4 This detail focuses on the volume between \( L_{\text{UpS}} \) and \( L_{\text{DwS}} \). The upper panel shows the shock-frame plasma velocity profile. The lower panel shows the ISM-frame velocity at selected points in the shock structure, varying from 0 for \( x < L_{\text{UpS}} \) to \( u_0 \) just upstream from the subshock at \( x = 0 \). Note that we assume the shock is locally plane and that \( L_{\text{UpS}} \) and \( L_{\text{DwS}} \) are small compared to \( R_{\text{shk,ef}} \) as illustrated in Figure 5.3.

Table 5.1 Model Parameters

| Model | Type  | \( \gamma_0 \) | \( f_{\text{ion}} \) | \( B_0 \) [\( \mu \)G] | \(|L_{\text{UpS}}|\) \( r_g0 \) | \(|L_{\text{DwS}}|\) \( r_g0 \) | \( r_{\text{RH}} \) | \( r_{\text{tot}} \) | \( N_g \) | \( \varepsilon_H \)  | \( \varepsilon_{\text{He}} \) | \( \varepsilon_{\text{el}} \) |
|-------|-------|----------------|---------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| A     | UM    | 10             | 0                   | 100               | \( 10^4 \)        | ...               | 3.02              | 3.02              | 2000              | ...               | ...               | ...               |
| B     | UM    | 10             | 0.15                | 100               | \( 10^4 \)        | ...               | 3.02              | 3.02              | 2000              | ...               | ...               | ...               |
| C     | UM    | 10             | 0.15                | 3                 | \( 10^3 \)        | \( 10^3 \)        | 3.02              | 3.02              | 2000              | ...               | ...               | ...               |
| D     | UM    | 10             | 0.1                 | 100               | 300               | \( 10^3 \)        | 3.02              | 3.02              | 2000              | 0.60              | 0.30              | 0.10              |
| E     | NL    | 10             | 0.1                 | 100               | 300               | \( 10^3 \)        | 3.02              | 3.02              | 2000              | 0.36              | 0.20              | 0.44              |
| F     | NL    | 10             | 0.4                 | 100               | 300               | \( 10^3 \)        | 3.02              | 3.02              | 2000              | 0.66              | 0.34              | ...               |
| G     | UM    | 1.5            | 0.1                 | 100               | 300               | \( 10^3 \)        | 3.53              | 3.53              | 200               | ...               | ...               | ...               |
| H     | NL    | 1.5            | 0.1                 | 100               | 300               | \( 10^3 \)        | 3.53              | 3.9±0.4           | 200               | 0.52              | 0.27              | 0.08              |
| I     | UM    | 30             | ...                | 100               | 300               | \( 10^4 \)        | 3.00              | 3.00              | 10^4              | ...               | ...               | ...               |
| J     | NL    | 30             | ...                | 100               | 300               | \( 10^4 \)        | 3.00              | 3.00              | 10^4              | 0.66              | 0.34              | ...               |

a Models A, B, and C have \( n_p = n_e = 1 \) cm\(^{-3} \) with no helium. Models D–H have \( n_p = 1 \) cm\(^{-3} \), \( n_{\text{He}} = 0.1 \) cm\(^{-3} \), and \( n_e = 1.2 \) cm\(^{-3} \). Models I and J have \( n_p = 1 \) cm\(^{-3} \) and \( n_{\text{He}} = 0.1 \) cm\(^{-3} \) without electrons. The far upstream temperature is \( 10^6 \) K in all cases and all models use \( \eta_{\text{mfp}} = 1 \).

b In the nonlinear (NL) models the shock structure is determined self-consistently. The unmodified (UM) models have a discontinuous shock structure with no shock smoothing.

c For self-consistent NL models, this is the fraction of total energy placed in all particles with energies above 100 MeV as measured in the shock frame.
5.4. RESULTS

5.4. OTHER PARTICLE SPECIES

Figure 5.5 Protons (black curves) and electrons (red curves) from UM shocks with different \( f_{\text{ion}} \) as indicated. These spectra, multiplied by \( p^{2.23} \), are calculated downstream from the shock, in the shock frame, and have arbitrary overall normalization although the relative normalization between electrons and protons is absolute. An upstream FEB of \( L_{\text{UpS}} = -10^4 r_{g0} \) was used with no downstream FEB (i.e., a probability-of-return calculation was used to simulate an infinite downstream region). In Table 5.1 the \( f_{\text{ion}} = 0 \) case is Model A and the \( f_{\text{ion}} = 0.15 \) case is Model B.

5.4.1 Unmodified Examples

In Figure 5.5 we show number spectra, \( dN/\,dp \), calculated just downstream from an unmodified shock with \( \gamma_0 = 10 \). The solid and dotted black curves are protons and the red curves are electrons. The dotted curves were calculated with \( f_{\text{ion}} = 0 \), while the solid curves were calculated with \( f_{\text{ion}} = 0.15 \), i.e., 15% of the proton ram kinetic energy is transferred to electrons as particles first cross \( x = 0 \) headed downstream.\(^4\) Since all particles injected far upstream are nonrelativistic, the transformation from upstream to downstream frames strongly favors more massive particles in the first shock crossing. This results in the strong depression of electrons relative to \( \text{H}^+ \). Once all particles become relativistic they are treated equally and obtain similar power laws, i.e., the UM shocks in Figure 5.5 show the canonical \( dN/\,dp \propto p^{-2.23} \) power law above the “thermal” peak and below the high

\(^4\)We refer to the sharp drop in \( u(x) \) that occurs at \( x \approx 0 \) as the subshock (see Figures 5.7 and 5.12). For an unmodified shock, there is no distinction between the shock and subshock.
5.4. RESULTS

CHAPTER 5. OTHER PARTICLE SPECIES

Figure 5.6 Proton and electron spectra (as labeled and multiplied by $p^{2.23}$) with $f_{\text{ion}} = 0.15$. These spectra are measured in the local plasma frame and are normalized to the total number of particles in a given region, as indicated in Figure 5.4. The shock acceleration is limited by an upstream FEB at $L_{\text{UpS}} = -1000 r_{g0}$, and by a downstream FEB at $L_{\text{DwS}} = +1000 r_{g0}$.

momentum cutoff. For the protons, the cutoff at $\sim 10^7 m_p c$ is produced by an upstream FEB at $L_{\text{UpS}} = -1 \times 10^4 r_{g0}$, where $r_{g0} = \eta_{\text{mfp}} m_p u_0 c/(e B_0)$. The electrons cut off at a lower momentum ($\sim 2 \times 10^6 m_p c$) due to radiation losses. Without radiation losses, the electrons would obtain the same $p_{\text{max}}$ as protons, since $p_{\text{max}}$ scales as $Z$. The normalization of the electron spectra shows the dramatic effect of $f_{\text{ion}}$: the $e/p$ ratio is increased by nearly three orders of magnitude with $f_{\text{ion}} = 0.15$. The proton normalization is only slightly influenced by the change in $f_{\text{ion}}$.

To save computation time, the models in Figure 5.5 used a probability-of-return calculation instead of a downstream FEB. This mimics an infinite downstream region and allows rapid acceleration to high energies (see Ellison et al., 1996, for a discussion of the probability of return calculation). In Figure 5.6 we show spectra from an UM shock measured in the local plasma frame with both upstream and downstream FEBs. While both FEBs are present, typically one (the shorter, measured in diffusion lengths) will determine the maximum momentum $p_{\text{max}}$. For Figure 5.6, both $L_{\text{UpS}}$ and $L_{\text{DwS}}$ are $1 \times 10^3 r_{g0}$ from the subshock, but $L_{\text{DwS}}$ determines $p_{\text{max}}$ since it is much more difficult for
5.4. RESULTS

The FEBs also determine the total number of particles accelerated (as indicated in Figure 5.4) and the dN/dp spectra in Figure 5.6 are normalized to the total particle number in a region surrounding an observation position x (the position is indicated in Figure 5.6). While the particular number of accelerated particles in these examples is arbitrary, it is important to note that the Monte Carlo code determines the absolute number of accelerated particles, and subsequent radiation, for any given set of environmental and shock parameters, including $f_{\text{ion}}$. As long as $\gamma_0 \gg 1$, the large majority of accelerated particles will be in the downstream region since it is difficult for particles to stream upstream into the shock precursor. This is reflected in the higher normalizations of the black curves, measured at $x = 0+$ (i.e., just downstream), compared to the red or blue curves, measured in the shock precursor at $x = -1 \, r_{g0}$ and $x = -120 \, r_{g0}$, respectively.

Besides the downstream FEB and normalization, the spectra in Figure 5.6 differ from those in Figure 5.5 in that they are shown in the local plasma frame and they include spectra calculated upstream from the subshock (red and blue curves), as well as downstream (black curves). The effect of the Lorentz transformation from the shock frame to the local plasma frame is clearly indicated by the high-momentum upstream spectra (red and blue curves) which extend to higher momentum than the downstream spectra. Relative to the subshock, the upstream plasma frame moves with $\gamma_0 = 10$, while the downstream plasma frame moves with $\gamma_2 = [1 - (u_2/c)^2]^{-1/2} \approx 1.06$ (here, $u_2 \approx c/3$ is the downstream bulk plasma speed as measured in the shock frame).

5.4.2 Nonlinear Example: $\gamma_0 = 10$

In Figure 5.7 we show the structure of a shock where the backpressure from accelerated protons, He$^{2+}$, and electrons is taken into account. For this example, $f_{\text{ion}} = 0.1$, and the electron pressure contributes to the determination of the self-consistent shock structure. Here $L_{Ups} = -300 \, r_{g0}$ and $L_{Dws} = +1000 \, r_{g0}$. In the top panels, the solid (black) curve is the bulk flow speed, $u(x)/u_0$, and the dot-dashed (blue) curve is $\gamma(x)\beta(x)/(\gamma_0\beta_0)$, where $\beta = u/c$, and $\gamma = [1 - \beta^2]^{-1/2}$. The dashed (red) curve is $u(x)/u_0$ for the UM case. The lower panels show the momentum and energy fluxes for the NL case (solid black curves), as well as for the UM case (dashed red curves). All curves are normalized to far upstream values. Without shock smoothing, the downstream momentum and energy fluxes are nearly a factor of two out of conservation.

We note that in nonrelativistic and transrelativistic shocks, due to particle escape and a decrease in the downstream ratio of specific heats from $5/3 \rightarrow 4/3$ as nonrelativistic particles are turned...
5.4. RESULTS

CHAPTER 5. OTHER PARTICLE SPECIES

Figure 5.7 The top panel shows the shock structure for the unmodified case (dashed red curve, Model D) and the nonlinear case (solid black and dot-dashed blue curves, Model E). The solid (black) and dashed (red) curves in the top panel are the flow speed $u(x)/u_0$, while the dot-dashed (blue) curve is $\gamma(x)\beta(x)/(\gamma_0\beta_0)$, where $\gamma(x)\beta(x)$ scales as 1/density for the nonlinear shock. The middle and bottom panels show the momentum and energy fluxes, respectively, normalized to far upstream values. Note that three particle species—protons, He$^{2+}$, and electrons—are included in determining the self-consistent shock structure.
into relativistic ones, the shock smoothing required for momentum and energy conservation is accompanied by an increase in the overall shock compression ratio above the Rankine–Hugoniot value, i.e. $r_{\text{tot}} > r_{\text{RH}}$ (e.g., Berezhko and Ellison, 1999). For fully relativistic shocks, however, $r_{\text{tot}} \simeq r_{\text{RH}}$ since the downstream ratio of specific heats remains $\simeq 4/3$, and it is extremely difficult for particles to escape upstream from the subshock (for a discussion of the Rankine–Hugoniot compression ratio $r_{\text{RH}}$, and see Figure 8.1 for an illustration of the difficulty of escape from relativistic shocks. The overall compression ratio is defined as $r_{\text{tot}} = u_0/u_2$, and the subshock compression ratio is $r_{\text{sub}} = u_1/u_2$. Here, $u_2$ is the bulk plasma speed downstream from the subshock and $u_1$ is the plasma speed just upstream of the viscous subshock, both measured in the shock rest frame.\(^6\)

The top panel in Figure 5.8 shows downstream proton, He\(^{2+}\), and electron spectra for the UM shock (dashed curves in Figure 5.7). The bottom panel shows these spectra from the nonlinear shock. All parameters are the same for these two cases—the only difference is that momentum and energy are conserved in the NL case.

The effects of the smooth shock structure are clearly evident. In the NL case, the downstream spectra are noticeably curved and less intense, as is necessary to conserve energy and momentum. Significantly, the electrons are more modified than the protons or He\(^{2+}\), and the $e/p$ ratio in the quasi–power law portion of the spectra drops by more than an order of magnitude. Note that with $f_{\text{ion}} = 0$, this ratio would have dropped by several more orders. This difference is a direct result of our scattering assumption, Equation 5.2. The electrons, with their small $A/Z$, feel the effects of the smooth shock more acutely than do the heavier ions and are less efficiently injected and accelerated until they reach $p \gtrsim 10 m_p c$. Since the efficiency for accelerating electrons cannot be orders of magnitude less than it is for protons in astrophysical sources, energy must be transferred from heavy particles to electrons with a reasonable efficiency if DSA is to be important (e.g., Sironi and Spitkovsky, 2011; Sironi et al., 2013).

The heavier He\(^{2+}\), with $A/Z = 2$, is accelerated more efficiently than protons; the $\text{He}^{2+}/p$ ratio above $\sim 100 m_p c$ goes from $\text{He}^{2+}/p < 1$ in the UM shock to $\text{He}^{2+}/p > 1$ in the NL shock (see Figure 5.15 below). This is particularly significant at the high momentum cutoff. The $A/Z$ effect is discussed in more detail in Section 5.4.6. For Model E in Figure 5.8, the fraction of total ram kinetic energy placed in particles of 100 MeV or greater is $\varepsilon_\text{H} = 0.60$, $\varepsilon_{\text{He}} = 0.30$, and $\varepsilon_{\text{el}} = 0.10$, for protons, He\(^{2+}\), and electrons respectively (see Table 5.1, where it is noted that these fractions are measured in the shock frame.)

The effects of shock smoothing also show up in the broadening and shift to lower momentum

---

\(^6\)As seen in Figure 5.7, the definition of $u_1$ is imprecise because the Monte Carlo solution allows for a smooth decrease in the precursor speed into the downstream region.
Figure 5.8 Downstream, local plasma frame (LPF) spectra for the unmodified shock shown in Figure 5.7 (top panel, Model D) and the nonlinear shock shown in Figure 5.7 (bottom panel, Model E). Note the pronounced suprathermal tail on the electron distribution.
of the “thermal” peaks, as seen in the bottom panel of Figure 5.8. Of particular interest is the pronounced suprathermal tail the electrons obtain in the NL case. If synchrotron self-absorption (SSA) is unimportant, this can produce a notable effect in the synchrotron emission, as we discuss next.

5.4.3 Photon Emission, $\gamma_0 = 10$

In Figure 5.9 we show the photon emission produced by the shocks described in Figures 5.7 and 5.8. As in Figure 5.8, the top panel is for the unmodified shock and the bottom panel is for the nonlinear shock—all other parameters are the same. The solid curves labeled synchrotron, IC, and pion-decay show isotropic emission calculated in the local plasma frame (LPF) summed over the regions between the upstream and downstream FEBs (i.e., between $x = -300 r_{g0}$ and $x = +1000 r_{g0}$) as indicated in Figure 5.4. These are the fluxes that would be observed at a distance $d_{\text{Mpc}}$ Mpc if no Lorentz transformations were required. (Of course, the particle distributions will not be isotropic in the ISM frame, and Lorentz transformations are required.) The dashed (black) curves show the total emission from synchrotron, IC, and pion-decay transformed to the ISM frame and seen by an observer directly in front of the shock at $D_{\text{obs}} = -d_{\text{Mpc}}$ Mpc.

The shock simulation is done in the shock rest frame. To obtain the local plasma frame synchrotron and pion-decay spectra we first transform the particle spectra to the LPF taking into account the anisotropies introduced by the relativistic flow. (This procedure is discussed in Section 5.3.1.) We then calculate the photon emission in the LPF assuming it is produced isotropically. This is reasonable for the synchrotron emission since we implicitly assume the background magnetic field is highly turbulent, and the synchrotron photons should be produced isotropically as the electrons spiral in the turbulent field. It is also a good approximation for the pion-decay emission since the protons and He$^{2+}$ ions interact with the local plasma and produce pions that can become isotropic in the LPF before emitting a $\gamma$-ray.

From this isotropic emission, we obtain the flux in the ISM frame directly in front of the shock by employing the standard Doppler shift and Lorentz transformations (e.g., Kumar and Zhang, 2014). For an observer directly in front of the shock at $D_{\text{obs}} = -d_{\text{Mpc}}$ Mpc, these give a boost to the energy flux $\propto D^4$, where $D = \gamma(x)[1 + \beta(x)]$ is the Doppler factor at position $x$ relative to the subshock. One factor of $D$ comes from the Doppler shift, one from time dilation, and two from relativistic beaming. Here $\gamma(x)$ and $\beta(x)$ are measured relative to the ISM frame and will be different for each region between the upstream and downstream FEBs as shown in Fig. 5.4. Far upstream, $D = 1$, while downstream for $\gamma_0 = 10$ and $r_{\text{tot}} \approx 3.02$, $D \approx 14.2$.

Since the Cosmic Microwave Background (CMB) photons are nearly isotropic in the ISM frame,
5.4. RESULTS

CHAPTER 5. OTHER PARTICLE SPECIES

Figure 5.9 Photon emission for the UM (Model D) and NL (Model E) shocks shown in Figures 5.7 and 5.8. The dashed (black) curves are the total emission transformed to the ISM frame for an observer at $D_{\text{obs}} = -d_{\text{Mpc}} \, \text{Mpc}$, directly in front of the shock. With the exception of the dotted (blue) IC curves, the lower curves in each panel show emission calculated in the local plasma frame and summed over the shock from $L_{\text{ups}} = -300 \, r_\Omega$ to $L_{\text{downs}} = +1000 \, r_\Omega$. The dotted (blue) IC curves show the total IC emission transformed to the ISM frame. The dashed (red) pion-decay curves are the LPF emission from He$^{2+}$. 

134
we calculate the IC emission in a different fashion.\footnote{We only consider CMB photons here for simplicity—the techniques we present can be generalized to include IC emission from other photon fields, including synchrotron photons, if the jet parameters warrant it.} We first transform the electron distribution into the ISM frame, keeping the two-dimensional anisotropy inherent in our plane-parallel shock simulation. We then calculate the emission directed toward the observer at $D_{\text{obs}} = -d_{\text{Mpc}}\text{ Mpc}$ using Equation (9) in Jones (1968). The assumption here is that the relativistic electrons produce strongly beamed emission, so only those electrons directed toward the observer contribute to the observed flux. In Figure 5.9, the dotted (blue) curves are the fully transformed IC emission calculated in the ISM frame and directed toward the observer in the $-x$-direction. Note that the solid IC curves, obtained assuming the electrons are isotropic in the LPF and Compton scatter off of isotropic CMB photons, are only present to give an indication of the relativistic boost the IC photons receive. The dashed (black) curves contain the full observed flux from emission produced over the modified shock structure and transformed to the ISM frame for an observer looking directly down the jet axis (i.e., toward $+x$ in Figure 5.4). The dotted (blue) IC curves are part of this sum.

The effects from shock smoothing on the particle distributions (Figure 5.8) produce corresponding changes in the photon emission. Since electrons are suppressed more than ions in the NL shock, the synchrotron and IC emission drops more than does pion-decay between the UM and NL cases. In contrast, the pion-decay emission from He$^{2+}$ (dashed red pion-decay curve) is increased relative to that from protons due to the $A/Z$ effect. The “thermal” peaks near $1\,m_p\,c$ for electrons, and near $10\,m_p\,c$ for protons and He$^{2+}$, show up as clear peaks in the synchrotron, IC, and pion-decay emission at $E \approx 10^{-11}, 10^{-2}, \text{and } 10^3\text{ MeV}$, respectively. This is particularly significant for the synchrotron emission near $10^{-9}\text{ MeV}$. The NL curvature in the particle spectra shows clearly in the individual components and remains strongly evident in the summed flux (bottom panel of Figure 5.9).

5.4.4 Nonlinear, $\gamma_0 = 10$, Variation of $f_{\text{ion}}$

While we have used $f_{\text{ion}} = 0.1$ in our NL Model E, the PIC simulations of Sironi et al. (2013) show examples where $\sim 40\%$ of the energy in accelerated particles ends up in electrons (see their Figure 11). In Figure 5.10 we compare particle spectra, and in Figure 5.11 we compare the total observed energy flux, for $f_{\text{ion}} = 0.1$ (Model E) and $f_{\text{ion}} = 0.4$ (Model F). The flux between $\sim 10^{-9}\text{ MeV}$ and $\sim 1\text{ GeV}$ is $\sim 100$ times greater for the $f_{\text{ion}} = 0.4$ case, with a much smaller decrease in the GeV–TeV emission. The $f_{\text{ion}} = 0.4$ example shows the curved spectral shape that results from the NL shock smoothing but it is less pronounced between $10^{-9}$ and 1 MeV than for $f_{\text{ion}} = 0.1$. In the GeV–TeV range, the curvature is slightly greater than for $f_{\text{ion}} = 0.4$.

With $f_{\text{ion}} = 0.4$, the energy distribution above $100\text{ MeV}$ in protons, He$^{2+}$, and electrons is $\varepsilon_H = 0.36$,
Figure 5.10 Nonlinear downstream LPF spectra for Model E ($f_{\text{ion}} = 0.1$) and Model F ($f_{\text{ion}} = 0.4$). In both panels the solid (black) curves are protons, the dashed (red) curves are $\text{He}^{2+}$, and the dotted (blue) curves are electrons.
5.4. RESULTS

Figure 5.11 Total observed energy flux for NL Models E ($f_{\text{ion}} = 0.1$) and F ($f_{\text{ion}} = 0.4$).

\[ \varepsilon_{\text{He}} = 0.20, \quad \varepsilon_{\text{el}} = 0.44, \]
respectively; nearly 50% of the ram kinetic energy goes into electrons of energy 100 MeV or greater, as measured in the shock frame.

5.4.5 Trans-relativistic, $\gamma_0 = 1.5$

As described in Ellison et al. (2013), the Monte Carlo simulation smoothly treats nonrelativistic to ultrarelativistic shocks. In Figure 5.12 we show the profile of a transrelativistic $\gamma_0 = 1.5$ shock for comparison with Figure 5.7. Apart from $\gamma_0$, all input parameters are the same for the $\gamma_0 = 1.5$ and $\gamma_0 = 10$ cases (though note that the physical distances of the two sets of FEBs differ since $r_{g0}$ increases with $u_0$). Since it is easier for accelerated particles to diffuse upstream against the inflowing plasma with $\gamma_0 = 1.5$ than against $\gamma_0 = 10$, the NL shock precursor is much more extended than it is for $\gamma_0 = 10$. The bulk flow speed, $u(x)$, is noticeably modified out to $x = L_{\text{UpS}} = -300 r_{g0}$ (note the split log-linear $x$-axis in Figure 5.12).

As in the $\gamma_0 = 10$ case, the momentum and energy fluxes are not conserved in the UM shock but are within a few percent of the far upstream values once the shock structure is modified by the CR backpressure. For $\gamma_0 = 1.5$, the overall compression ratio must also be increased to conserve momentum and energy, and we find $r_{\text{tot}} = 3.9 \pm 0.4$, where the uncertainty comes from statistics and errors inherent in the Monte Carlo smoothing algorithm. This result is similar to that given
5.4. RESULTS

CHAPTER 5. OTHER PARTICLE SPECIES

Figure 5.12 All curves are as in Figure 5.7, for a shock with $\gamma_0 = 1.5$. The dashed (red) curves are the UM (Model G) case while the solid (black) and dot-dashed (blue) curves are the NL (Model H) profiles. The upstream and downstream FEBs are $L_{\text{UpS}} = -300 r_{g0}$ and $L_{\text{DwS}} = +1000 r_{g0}$, as in Figure 5.7, but these distances differ in absolute units since $u_0$ varies between $\gamma_0 = 10$ and $\gamma_0 = 1.5$. 

138
5.4. RESULTS

CHAPTER 5. OTHER PARTICLE SPECIES

in Ellison et al. (2013), except that here we have included He$^{2+}$ and electrons in determining the self-consistent shock structure, and we use a downstream FEB as well as an upstream one.

In Figure 5.13 we show the particle spectra for $\gamma_0 = 1.5$ in the same format as Figure 5.8, except that we have added the $\gamma_0 = 10$ proton spectra for comparison. The $\gamma_0 = 1.5$ spectra are harder than those for $\gamma_0 = 10$, mainly because $n_{\text{tot}}$ is larger. The maximum momentum is noticeably lower for $\gamma_0 = 1.5$, as is the “thermal” peak is also lower since, for downstream spectra, it occurs at $\sim \gamma_0 m_p c$ for protons in the UM shock. The “thermal” peak is at a noticeably lower momentum in the NL case. Another important difference is the normalization. The $\gamma_0 = 10$ proton distributions are about a factor of 100 above the $\gamma_0 = 1.5$ spectra in the quasi–power law region. This comes about for two primary reasons. The $\gamma_0 = 10$ shock has a considerably higher downstream density, since $n_2 = \gamma_0 \beta_0 n_0 / (\gamma_2 \beta_2)$; as well, with $r_g$ being larger, the volume containing this higher density is also increased. Another cause is that the spectra are multiplied by $p^{2.23}$ rather than by $p$. The area under a $dN/d\log(p)$ curve would be the total number of particles. But $dN/d\log(p) = p \cdot dN/dp$, so the extra 1.23 powers of $p$ merely enhance the perceived importance of particles at higher momentum. Since the $\gamma_0 = 10$ shock contains particles at higher momenta, it also gets plotted higher on the $p^{2.23} dN/dp$ plot. The efficiencies for producing 100 MeV or greater energy particles are $\varepsilon_{\text{H}} = 0.52$, $\varepsilon_{\text{He}} = 0.27$, and $\varepsilon_{\text{el}} = 0.08$ for Model H; i.e., the $\gamma_0 = 1.5$ shock puts $\sim 8\%$ of the shock energy into energetic electrons.

In Figure 5.14 we show the photon emission for $\gamma_0 = 1.5$ in the same format as in Figure 5.9, with the addition of the total ISM frame emission for the $\gamma_0 = 10$ cases (dotted, black curves) added to the $\gamma_0 = 1.5$ cases (dashed, black curves) for comparison. The total emission for an observer at $D_{\text{obs}} = -d_{\text{Mpc}}$ Mpc is dramatically different for the two Lorentz factors. To highlight this in the lower panel, we have plotted the total ISM-frame emission for $\gamma_0 = 1.5$, multiplied by $10^5$ (dot-dashed, purple curve). The broadband spectral shapes are very different between $\gamma_0 = 10$ and 1.5, and at $\sim 1$ keV the normalization differs by $\sim 10^5$.

Note that the NL $\gamma_0 = 1.5$ spectrum is harder than $\gamma_0 = 10$ in the GeV–TeV range but cuts off at a lower energy. The lower cutoff energy shows up dramatically in the synchrotron emission. For $\gamma_0 = 1.5$ the synchrotron peak is around 1 keV, typical of SNR observations, while for $\gamma_0 = 10$, the peak is around 1 MeV. At radio energies, the synchrotron spectra are very different because of the emission produced by the downstream thermal electrons. Of course, we have not considered SSA here; this process will produce a low-energy cutoff in the synchrotron emission which may mask the emission from the thermal electrons. We have also not considered $\gamma$-ray absorption between the GRB and Earth.
Figure 5.13 Except for the dotted (black) curves labeled “\( \gamma_0 = 10 \) p’s” all spectra are downstream, local plasma frame (LPF) spectra for the unmodified \( \gamma_0 = 1.5 \) shock (top panel, Model G) and the nonlinear \( \gamma_0 = 1.5 \) shock (bottom panel, Model H) shown in Figure 5.12. The curves labeled “\( \gamma_0 = 10 \) p’s” are identical to those in Figure 5.8.
Figure 5.14 Photon emission for the UM (Model G) and NL (Model H) shocks shown in Figures 5.12 and 5.13. Except for the dotted (black) curves labeled “Total, ISM $\gamma_0 = 10$”, all curves are for the $\gamma_0 = 1.5$ shock and are in the same format as in Figure 5.9. The dashed (black) curves are the total emission transformed to the ISM frame for an observer at $x = -d_{\text{Mpc}}$ Mpc, directly in front of the shock. The dot-dashed (purple) curve in the lower panel is the total ISM frame $\gamma_0 = 1.5$ emission, multiplied by $10^5$. 
5.4.6 $A/Z$ Enhancement of Heavy Ions

The Monte Carlo code assumes that all scatterings are elastic in the local plasma frame. This implies that an insignificant fraction of the particle energy is transferred to magnetic turbulence in the wave generation process. With elastic scattering, $\gamma_i v_i \propto p_i / A$ remains constant in a scattering event, where $\gamma_i$ and $v_i$ are local frame values for particle species $i$. The energy gain particles receive on crossing the shock, which is determined by a Lorentz transformation between the two frames, also scales as $p_i / A$. In our plane-parallel approximation, the final factor that determines the acceleration is the probability that particles make a set number of shock crossings; this also depends only on $\gamma_i v_i$. Thus an UM shock, with the Monte Carlo assumptions, will treat all particles identically in momentum per nucleon, including the thermal leakage injection.

An exception to this occurs if the acceleration is limited by a boundary at a fixed distance, as we assume here, or a maximum acceleration time (as will be discussed in Chapter 8). Since we assume Equation 5.2, diffusion length and acceleration time both scale as $(A/Z)(p_i / A)$ for $v_i \sim c$. Thus, apart from the normalization set by input parameters and the maximum momentum cutoff, all species should have identical spectra when plotted against $p / A$. With a fixed FEB, the spectra of high $A/Z$ particles will cut off at a lower $p / A$ than will the spectra of low $A/Z$ particles.

For models I, D, and G in Figure 5.15, we show downstream, shock frame, proton and He$^{2+}$ spectra for UM shocks plotted in $p / A$ units. Except for statistical variations, a factor of 10 normalization since $n_{\text{He}} = 0.1 n_p$, and the high momentum cutoff, the proton and He$^{2+}$ spectra are identical. Electrons are not plotted but would show the same effect. In the corresponding NL models (J, E, and H), a clear enhancement of He$^{2+}$, produced solely because the shock structure is smoothed by the backpressure of accelerated particles, is seen. For Model J with $\gamma_0 = 30$, we calculate $\varepsilon_{\text{He}} = 0.66$ and $\varepsilon_{\text{He}} = 0.34$.

The factor of two enhancement in the He$^{2+}/p$ ratio seen in the NL models in Figure 5.15 should be large enough to see clearly in PIC simulations. Since this enhancement is a prediction that stems directly from important assumptions of efficient DSA and thermal leakage injection, adding helium to PIC simulations can test these assumptions. If the acceleration is efficient, and the $A/Z$ effect is not seen, it implies that one or more of the following may be happening. (i) The accelerated protons and He$^{2+}$ may be sharing significant energy with each other rather than interacting mainly elastically with the background turbulence. If this is the case, it will influence all aspects of DSA. (ii) The particle mean free path may not be a monotonically increasing function of momentum, or it may differ substantially for protons and He$^{2+}$. (iii) Different $A/Z$ particles may interact differently with the viscous subshock layer. A basic assumption for thermal leakage injection is that the subshock is essentially transparent, i.e., phenomena such as cross-shock potentials or large-scale turbulence do
Figure 5.15 Models I, D, and G show proton and He$^{2+}$ spectra for UM shocks, while models J, E, and H are the corresponding spectra for the NL shocks. In all cases, the He$^{2+}$ spectra are multiplied by 10 to adjust for the ambient number density. When plotted in $p/A$ units the UM spectra are identical except for statistics and the maximum momentum cutoff. The NL shocks show a clear $A/Z$ enhancement in the He$^{2+}/p$ ratio.
not strongly influence the injection process. If these phenomena are important, it is likely they will influence different $A/Z$ particles differently, modifying the $A/Z$ enhancement seen in Figure 5.15.

5.5 Conclusions

As complicated as DSA in relativistic shocks may be, one aspect is profoundly simple: if the acceleration is efficient and a sizable fraction of the bulk plasma flow energy is put into individual accelerated particles, as is often assumed in applications (e.g., Kulkarni et al., 1999; Piran et al., 2001; Piran, 2004; Mészáros, 2006), the accelerated particles must self-consistently modify the shock structure to conserve momentum and energy, regardless of the plasma physics details. We have investigated how the kinematics of shock modification influences the relative acceleration of electrons, protons, and heavy elements (i.e., He$^{2+}$) using a Monte Carlo simulation with a dynamic range large enough to model acceleration from injection at nonrelativistic thermal energies to ultrarelativistic CR energies. Figure 5.6 shows a 12 decade range in plasma-frame momentum and greater than a 20 decade range in $dN/dp$. A corresponding range in photon emission is also obtained (e.g., Figure 5.14).

The underlying wave–particle plasma interactions, which are parameterized in the Monte Carlo code, will influence details of the shock modification and the resultant radiation; they will determine if acceleration is, in fact, efficient and set the maximum energy particles obtain. However, our results show general aspects that are largely independent of the poorly known plasma physics details if the acceleration is efficient. Considering only the kinematics, electrons will be accelerated much less efficiently than ions if the shock structure is modified by the heavy particles. This result assumes that the heavy particles and electrons diffuse in a similar fashion, as indicated in Equation 5.1. If this is the case, the $A/Z$ enhancement effect we describe increases the injection and acceleration efficiency of high $A/Z$ particles compared to low $A/Z$ ones. This dramatically decreases the abundance of accelerated electrons compared to heavier ions (e.g., Figure 5.8). The kinematics suggest that relativistic shocks will not be able to place a sizable fraction of the shock kinetic energy flux into leptons if protons are present.

Of course, beyond kinematics, the magnetic turbulence produced by wave–particle interactions plays a critical role and recent results (e.g., Sironi and Spitkovsky, 2011) show that some fraction of the proton energy can be transferred to electrons in the shock precursor via magnetic turbulence. This PIC result is particularly important for astrophysical applications where the radiating electrons presumably contain a sizable fraction of the available energy budget. In relativistic shocks, heavy elements must transfer a sizable fraction of their energy to electrons for DSA to be relevant to observed photon fluxes.

We have modeled this energy transfer by including a parameter, $f_{\text{ion}}$, that sets the fraction of
ion energy transferred to electrons as the particles first cross the subshock. While the effect of $f_{\text{ion}}$ is large (e.g., Figure 5.5), kinematics still play a role: light and heavy particles will be treated differently in relativistic flows. This is seen clearly in Figures 5.8 and 5.13 where, for a given $f_{\text{ion}}$, the $e/p$ ratio drops substantially between the UM case (where no $A/Z$ effect occurs) and the NL case. Figures 5.8 and 5.13, where $f_{\text{ion}} = 0.1$ for both the $\gamma_0 = 10$ and 1.5 shocks, also show that a larger fraction of ion energy must be transferred to electrons for high–Lorentz factor shocks to produce a significant $e/p$ ratio.

We are not aware of any analytic work on turbulence in the transrelativistic regime, where our previous results (see Figure 10 in Ellison et al., 2013) suggest there may be a significant change in the character of the NL effects. Again we emphasize that while the transfer of energy between particle species must occur through wave–particle interactions not modeled with the Monte Carlo code, the energy transfer that is required to make DSA relevant stems directly from the kinematics which the Monte Carlo technique does model.

We make a clear prediction that is directly testable with future PIC simulations. If DSA is efficient enough so the shock structure is modified by the backpressure of accelerated particles, heavy element ions will show a clear enhancement over protons (i.e., Figure 5.15). We know of no non-kinematic effects (e.g., cross-shock potentials, energy transfer via wave–particle interactions, or other electrostatic processes) that can produce such an enhancement.

The combined processes of energy transfer from heavy particles to electrons, and the kinematics of shock smoothing, produce strong signatures on the radiation emitted by these particles. In Figure 5.14 we show results for $\gamma_0 = 10$ and $\gamma_0 = 1.5$. Since there are a number of important parameters that influence the emission (such as shock Lorentz factor, ambient density, magnetic field, and size of the emitting region), it is non-trivial to characterize the emission. Nevertheless some general properties stem mainly from the kinematics and should hold regardless of the plasma physics details.

Particle spectra should harden as the shock speed decreases from fully relativistic to non-relativistic speeds, mainly because the shock compression ratio increases and, for low enough $\gamma_0$, $r_{\text{tot}} > r_{\text{RH}}$ (see Figure 10 in Ellison et al., 2013). However, even though the compression ratio (defined as $r_{\text{tot}} = u_0/u_2$) is lower for ultrarelativistic shocks, the downstream local plasma number density $n_2 = \gamma_0\beta_0 n_0/(\gamma_2\beta_2)$ can be large, enhancing the pion-decay emission more than either IC or synchrotron.

The magnetic field is a critical parameter for synchrotron emission and, with the exception of Model C, we have assumed $B_0 = 100 \mu\text{G}$ for the background field. In our plane-parallel approximation, the background field remains constant throughout the shock. Values of 100’s of $\mu\text{G}$ can be expected for a shock moving through an ambient field of a few $\mu\text{G}$ when compression and amplification
are considered. Compression will increase the field by a factor $\sim \gamma_0$. Nonlinear amplification, as believed to occur in strong, nonrelativistic shocks in young SNRs (see, for example, Bell, 2004; Vladimirov et al., 2009; Bykov et al., 2014, and references therein), may increase the strength further. For simplicity, we have not attempted to include compression or magnetic field amplification of the magnetic field here. Field compression (but not additional amplification) is included in Chapter 8.

One important aspect of the changing afterglow emission as the shock slows from ultrarelativistic to nonrelativistic speeds is the position of the synchrotron peak. As seen in Figure 5.14, the peak shifts from $\sim$MeV to $\sim$keV energies as the shock slows from $\gamma_0 = 10$ to $\gamma_0 = 1.5$. A detailed evolutionary model of GRB afterglows using Monte Carlo techniques for NL DSA is presented in Chapter 8. In this afterglow model, the Monte Carlo simulation is combined with an analytic or numerical description of the jet-shock evolution. The shock-accelerated particles and resultant radiation are calculated at various times as the shock moves through the jet, and the total emission observed at Earth is determined.
Given the likely absence of \textit{in situ} observations of GRB afterglows for the foreseeable future, the primary observational tool we have for afterglow science is the photon production by energetic particles in the vicinity of the jet. In this section I will describe the manner in which photon emission is calculated for the particle spectra.

\section{The Example Shock}

In the next section I will explain the three emission processes used in this dissertation. I will also give sample spectra to emphasize certain features of interest. The spectra are taken from the downstream region of the shock I will briefly describe in this section.

For the sample spectra, I have used an unmodified relativistic shock with $\gamma_0 = 10$, chosen so that the shocked thermal population (i.e. the particles not injected into the acceleration process) will still be highly relativistic in the plasma frame. The scattering parameter $N_s$ was chosen to be very coarse, at 240, so as to recover a $p^{-2}$ power law in the CR spectrum. This is clearly visible in Figure 6.1, which shows all particle spectra needed for calculating photon emission: proton/helium/electron in the plasma frame, and electron in the explosion (or ISM) frame.
6.2 Emission processes

This work includes three emission processes. The first I will discuss is pion production from hadronic $(p - p, p - \alpha, \alpha - \alpha)$ interactions; the produced pions then decay into $\gamma$-rays (there are other decay products, such as neutrinos and electrons/positrons; the work here considers only photons). The second mechanism for photon production used is synchrotron radiation, generated by relativistic electrons gyrating in the turbulent magnetic field; synchrotron radiation ranges from radio wavelengths all the way to $\gamma$-rays. Finally, I will discuss the inverse Compton emission process, in which ambient photons scatter off of energetic electrons and gain significant amounts of energy in so doing. For the purposes of this dissertation, IC emission is limited to the $\gamma$-ray range; at X-ray and lower energies, any IC emission would be orders of magnitude weaker than the synchrotron emission from the corresponding electron population.

6.2.1 Pion decay

When energetic nucleons scatter off of other nucleons (so phrased to include the possibility of helium nuclei), unstable pions may be produced, which then decay into $\gamma$-ray photons. The photon...
production section of the code uses two methods for calculating photon production via this process: that of Kamae et al. (2006) at lower energies, and that of Kelner et al. (2006) at higher energies.

Both procedures make two assumptions: (1) that the target protons are essentially at rest in the plasma frame, and (2) that the CRs responsible for the pion production are isotropic in that frame. Both of these assumptions are strained in the context of relativistic DSA. As shown in Figure 6.1, the thermal population of protons has a kinetic energy of \( \sim 7 \) GeV; the helium population has an even higher energy. Both of these populations are isotropic in the downstream plasma frame, but it is certainly not the case that the target population is at rest. A further issue is that the CR population is not isotropic upstream from the shock, or downstream near the shock. This might have significant effects on the angular distribution of pion-decay photons. However, both considerations are deferred to future work.

**Kamae et al. method**

At low energies, pion production is handled using the parametrization found in Kamae et al. (2006). That procedure recognizes four ways for a nucleon-nucleon interaction to create a \( \pi \) particle that then decays into photons. Two of these are the diffractive and nondiffractive scattering processes. The remaining two are resonances at high energy: \( \Delta(1232) \) particles (with a mass of 1.232 GeV) that are created during a scattering event decay into a nucleon and a pion; while a second, higher-energy resonance at 1.6 GeV decays into a nucleon and two pions.

The Kamae method proceeds by calculating the differential cross section, which is simply the linear sum of the differential cross sections for each of the four processes.

\[
\frac{\Delta \sigma_{\text{tot}}(E_{\text{sec}})}{\Delta \log_{10}(E_{\text{sec}})} = \frac{\Delta \sigma_{\text{ND}}(E_{\text{sec}})}{\Delta \log_{10}(E_{\text{sec}})} + \frac{\Delta \sigma_{\text{diff}}(E_{\text{sec}})}{\Delta \log_{10}(E_{\text{sec}})} + \frac{\Delta \sigma_{1232}(E_{\text{sec}})}{\Delta \log_{10}(E_{\text{sec}})} + \frac{\Delta \sigma_{1600}(E_{\text{sec}})}{\Delta \log_{10}(E_{\text{sec}})}.
\]  

(6.1)

Note that this equation is the differential cross section in energy of the secondary particle created, rather than of the primary cosmic ray. All four cross sections may be expressed as the product of (1) a representation of the cross section, and (2) a parameter to enforce energy and momentum conservation near the kinematic limits of the process, i.e.

\[
\frac{\Delta \sigma_i(E_{\text{sec}})}{\Delta \log_{10}(E_{\text{sec}})} = F_i(x) \cdot F_{i,kl}(x).
\]  

(6.2)

Equation 6.2 is written in terms of \( x = \log_{10}(E_{\text{sec}}) \), where \( E_{\text{sec}} \) is the energy of the secondary particle measured in GeV. The Kamae method is capable of treating seven different varieties of secondary particles—photons, electrons, positrons, electron (anti)neutrinos, and muon (anti)neutrinos—but
Table 6.1 Functional forms used in the Kamae method, for a primary cosmic ray proton with energy $T_p$, producing a photon of energy $x = E/\text{(1 GeV)}$.

<table>
<thead>
<tr>
<th>Process</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nondiffractive process</td>
<td>$F_{\text{ND}}(x) = a_0 \cdot \exp \left{ -a_1 \left[ (x - a_3) + a_2(x - a_3)^2 \right] \right} + a_4 \cdot \exp \left{ -a_5 \left[ (x - a_6) + a_6(x - a_6)^2 + a_7(x - a_6)^3 \right] \right}$</td>
</tr>
<tr>
<td></td>
<td>$F_{\text{ND},kl}(x) = \frac{1}{\exp[W_{\text{ND},lo}(L_{\text{min}} - x)] + 1} \times \frac{1}{\exp[W_{\text{ND},hi}(x - L_{\text{max}})] + 1}$</td>
</tr>
<tr>
<td>Diffractive process</td>
<td>$F_{\text{diff}}(x) = b_0 \cdot \exp \left{ -b_1 \left[ \frac{x - b_2}{1 + b_3(x - b_2)} \right]^2 \right} + b_4 \cdot \exp \left{ -b_5 \left[ \frac{x - b_6}{1 + b_7(x - b_6)} \right]^2 \right}$</td>
</tr>
<tr>
<td></td>
<td>$F_{\text{diff},kl}(x) = \frac{1}{\exp[W_{\text{diff}}(x - L_{\text{max}})] + 1}$</td>
</tr>
<tr>
<td>Δ(1232) resonance</td>
<td>$F_{1232}(x) = c_0 \cdot \exp \left{ -c_1 \left[ \frac{x - c_2}{1 + c_3(x - c_2) + c_4(x - c_2)^2} \right]^2 \right}$</td>
</tr>
<tr>
<td></td>
<td>$F_{1232,kl}(x) = F_{\text{diff},kl}(x)$</td>
</tr>
<tr>
<td>1600 GeV resonance</td>
<td>$F_{1600}(x) = d_0 \cdot \exp \left{ -d_1 \left[ \frac{x - d_2}{1 + d_3(x - d_2) + d_4(x - d_2)^2} \right]^2 \right}$</td>
</tr>
<tr>
<td></td>
<td>$F_{1600,kl}(x) = F_{\text{diff},kl}(x)$</td>
</tr>
</tbody>
</table>

\[ a \quad W_{\text{ND},lo} = 15, \quad L_{\text{min}} = -2.6, \quad W_{\text{ND},hi} = 44, \quad \text{and} \quad L_{\text{max}} = 0.96 \log_{10}[T_p/(1 \text{ GeV})]. \]

\[ b \quad W_{\text{diff}} = 75, \quad \text{and} \quad L_{\text{max}} = \log_{10}[T_p/(1 \text{ GeV})]. \]

The formulae used in calculating photon production with the Kamae method are gathered in Tables 6.1 and 6.2. The various $F(x)$ functions depend almost solely on the energy of the produced photon. The energy of the cosmic ray, measured in GeV, enters into the $F(x)$ formulae only when calculating the kinematic limits of the various processes. On the other hand, the parameters used in the functions depend heavily on the energy of the cosmic ray, measured in TeV here.

The differential cross sections are, finally, converted into photon counts through typical means: multiplying by the density of target protons, by the number of primary cosmic rays at a particular energy, and by the velocity of those primaries. Summation over all possible primary energies
## Table 6.2 Parameters used in the Kamae method, for a primary cosmic ray proton with energy $T_p$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Formula in terms of $y = \log_{10}[T_p/(1 \text{ TeV})]$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nondiffractive process</strong></td>
<td></td>
</tr>
<tr>
<td>$a_0$</td>
<td>$-0.51187(y + 3.3) + 7.6179(y + 3.3)^2 - 2.1332(y + 3.3)^3 + 0.22184(y + 3.3)^4$</td>
</tr>
<tr>
<td>$a_1$</td>
<td>$-1.2592 \times 10^{-7} + (1.4439 \times 10^{-5}) \cdot \exp[-0.2936(y + 3.4)] + (5.9363 \times 10^{-5})/(y + 4.1485) + (2.264 \times 10^{-6})y - (3.3723 \times 10^{-7})y^2$</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$-174.83 + 152.78 \cdot \log_{10}[1.5682(y + 3.4)] - 808.74/(y + 4.6157)$</td>
</tr>
<tr>
<td>$a_3$</td>
<td>$0.81177 + 0.56385y + 0.0040031y^2 - 0.0057658y^3 + 0.00012057y^4$</td>
</tr>
<tr>
<td>$a_4$</td>
<td>$0.68631(y + 3.32) + 10.145(y + 3.32)^2 - 4.6176(y + 3.32)^3 + 0.86824(y + 3.32)^4 - 0.053741(y + 3.32)^5$</td>
</tr>
<tr>
<td>$a_5$</td>
<td>$9.0466 \times 10^{-7} + (1.4539 \times 10^{-6}) \cdot \log_{10}[0.015204(y + 3.4)] - (4.1228 \times 10^{-7})y + (2.2036 \times 10^{-7})y^2 + (1.3253 \times 10^{-4})/(y + 4.7171)^2$</td>
</tr>
<tr>
<td>$a_6$</td>
<td>$-339.45 + 618.73 \cdot \log_{10}[0.31595(y + 3.9)] + 250.2/(y + 4.4395)^2$</td>
</tr>
<tr>
<td>$a_7$</td>
<td>$-35.105 + 36.167y - 9.3575y^2 + 0.33717y^3$</td>
</tr>
<tr>
<td>$a_8$</td>
<td>$0.17554 + 0.373y - 0.014938y^2 + 0.003231y^3 + 0.025579y^4$</td>
</tr>
</tbody>
</table>
| $r(y)^a$ | \[
\begin{align*}
3.05 \cdot \exp\left[-107 \left\{ (y + 3.25)/[1.0 + 8.08(y + 3.25)] \right\}^2 \right] & \quad T_p \leq 1.95 \text{ GeV} \\
1.01 & \quad T_p > 1.95 \text{ GeV}
\end{align*}
\] |

<table>
<thead>
<tr>
<th><strong>Diffractive process</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_0$</td>
<td>$60.142 \cdot \tanh[-0.37555(y + 2.2)] - 5.9564(y + 0.59913)^2 + 0.0060162(y + 9.4773)^4$</td>
</tr>
<tr>
<td>$b_1$</td>
<td>$35.322 + 3.8026 \cdot \tanh[-2.5979(y + 1.9)] - 0.0002187(y + 369.13)^2$</td>
</tr>
<tr>
<td>$b_2$</td>
<td>$-15.732 - 0.082064 \cdot \tanh[-1.9621(y + 2.1)] + 0.00023355(y + 252.43)^2$</td>
</tr>
<tr>
<td>$b_3$</td>
<td>$-0.086827 + 0.37646 \cdot \exp[-0.50353 \left{ (y + 1.0444)/[1.0 + 0.27437(y + 1.0444)] \right}^2] + 0.94131 \cdot \exp[-24.347(y + 2.45) - 0.19717(y + 2.45)^2]^2$</td>
</tr>
<tr>
<td>$b_4$</td>
<td>$2.5982 + 0.39131(y + 2.95)^2 - 0.0049693(y + 2.95)^4$</td>
</tr>
<tr>
<td>$b_5$</td>
<td>$0.11198 - 0.64582y + 0.16114y^2 + 2.2853 \cdot \exp[-0.0032432 \left{ (y - 0.83562)/[1.0 + 0.33933(y - 0.83562)] \right}^2] + 1.7843 + 0.91914y + 0.050118y^2 + 0.038096y^3 - 0.027334y^4 - 0.003556y^5 + 0.0025742y^6$</td>
</tr>
<tr>
<td>$b_6$</td>
<td>$-0.19870 - 0.071003y + 0.019328y^2 - 0.28321 \cdot \exp[-6.0516(y + 1.8441)^2]$</td>
</tr>
</tbody>
</table>

---

*a For the nondiffractive process only, a renormalization is applied once $F_{ND}(x)$ is computed.

$b_0, b_1, b_2, b_3 = 0$ if $T_p < 5.52 \text{ GeV}$.

151
Table 6.2 (continued) Parameters used in the Kamae method, as functions of primary cosmic ray proton with energy $T_p$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Formula in terms of $y = \log_{10}[T_p/(1 \text{ TeV})]$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta(1232)$ resonance</td>
</tr>
<tr>
<td>$c_0$</td>
<td>$2.4316 \cdot \exp\left(-69.484 \left{ (y + 3.1301) / \left[ 1.0 + 1.24921(y + 3.1301) \right] \right}^2 \right) - 6.3003$</td>
</tr>
<tr>
<td></td>
<td>$- 9.5349 / y + 0.38121 y^2$</td>
</tr>
<tr>
<td>$c_1$</td>
<td>$56.872 + 40.627 y + 7.7528 y^2$</td>
</tr>
<tr>
<td>$c_2$</td>
<td>$-5.4918 - 6.7872 \cdot \tanh\left[ 4.7128(y + 2.1) \right] + 0.68048 y$</td>
</tr>
<tr>
<td>$c_3$</td>
<td>$-0.36414 + 0.039777 y$</td>
</tr>
<tr>
<td>$c_4$</td>
<td>$-0.72807 - 0.48828 y - 0.092876 y^2$</td>
</tr>
<tr>
<td></td>
<td>$1600 \text{ GeV resonance}$</td>
</tr>
<tr>
<td>$d_0$</td>
<td>$3.2433 \cdot \exp\left(-57.133 \left{ (y + 2.9507) / \left[ 1.0 + 1.2912(y + 2.9507) \right] \right}^2 \right) - 1.064$</td>
</tr>
<tr>
<td></td>
<td>$- 0.43925 y$</td>
</tr>
<tr>
<td>$d_1$</td>
<td>$16.901 + 5.9539 y - 2.1257 y^2 - 0.92057 y^3$</td>
</tr>
<tr>
<td>$d_2$</td>
<td>$-6.6638 - 7.501 \cdot \tanh\left[ 30.322(y + 2.1) \right] + 0.54662 y$</td>
</tr>
<tr>
<td>$d_3$</td>
<td>$-1.50648 - 0.87211 y - 0.17097 y^2$</td>
</tr>
<tr>
<td>$d_4$</td>
<td>$0.42795 + 0.55316 y + 0.20707 y^2 + 0.027552 y^3$</td>
</tr>
</tbody>
</table>

completes the calculation.

$$\frac{d^2N(E_\gamma)}{d\log_{10}(E_\gamma)dt} = n_{\text{amb}} \sum_{T_p} N_{\text{cr}}(T_p) \cdot v(T_p) \cdot \frac{\Delta\sigma_{\text{tot}}(E_\gamma)}{\Delta\log_{10}(E_\gamma)}$$

(6.3)

The parametrizations of the Kamae method are good to within 10-20% of simulated values near the kinematic threshold for pion production, i.e. $T_p < 2 \text{ GeV}$. Fidelity improves with increasing cosmic ray energy, dropping to within a few percent, until the cosmic ray energies reach about 500 TeV. At this point the parametrizations begin to lose accuracy again. To capture the pion production by protons at or above this energy, a different method is used.
Kelner et al. method

The method given in Kelner et al. (2006) is substantially simpler than the Kamae method outlined above, and applies at higher proton energies. The procedure calculates the number of photons per second produced at a given energy by integrating over the proton distribution:

\[
\frac{d^2 N_\gamma(E_\gamma)}{dE_\gamma dt} = c \cdot n_{amb} \int_{E_p}^{\infty} \sigma(E_p) \cdot \frac{dN_{cr}(p)}{dp} \cdot F_\gamma(x, E_p) \cdot \frac{dE_p}{E_p},
\]

where \(n_{amb}\) is the ambient density of target protons, \(\sigma(E_p)\) is the cross section for \(p-p\) interactions, \(dN/dp\) counts the number of cosmic ray protons at a given momentum, \(x = E_\gamma/E_p\) is the ratio of the photon energy to the proton energy (note the difference in usage between here and the Kamae method), \(c\) is the speed of light, and \(F_\gamma(x, E_p)\) relates the energy/number of produced photons to the energy of the primary cosmic ray.

The Kelner method uses a very simple formula to compute \(\sigma(E_p)\), in stark contrast to the extremely involved calculations used in the Kamae method.

\[
\sigma(E_p) = 34.3 + 1.88L + 0.25L^2,
\]

where \(L = \ln[E_p/(1\ TeV)]\) measures the energy of the primary proton. This cross-section is a decreasing function of proton energy, and results in a concave emission spectrum for a canonical \(p^{-4}\) phase space distribution of protons. As I will show below, however, it catastrophically overestimates the cross-section at lower energies compared to the more exacting Kamae method.

The photon production function \(F_\gamma(x, E_p)\) has the following “simple” form:

\[
F_\gamma(x, E_p) = B_\gamma \frac{\ln(x)}{x} \left( \frac{1 - x^{\beta_\gamma}}{1 + k_\gamma x^{\beta_\gamma}(1 - x^{\beta_\gamma})} \right)^4 \left[ \frac{1}{\ln(x)} - \frac{4\beta_\gamma x^{\beta_\gamma}}{1 - x^{\beta_\gamma}} - \frac{4k_\gamma \beta_\gamma x^{\beta_\gamma}(1 - 2x^{\beta_\gamma})}{1 + k_\gamma x^{\beta_\gamma}(1 - x^{\beta_\gamma})} \right].
\]

In this formulation, \(x\) is defined as above, and the three parameters \(B_\gamma, \beta_\gamma,\) and \(k_\gamma\) depend only on \(L\) (again, defined as above) as follows:

\[
B_\gamma = 1.30 + 0.14L + 0.011L^2,
\]

\[
\beta_\gamma = \frac{1}{1.79 + 0.11L + 0.008L^2},
\]

This assumes that all of the primary protons are moving relativistically. Near the kinematic threshold this assumption fails dramatically, but since the Kelner method is only used for particles with energies \(E_p \gg 1\ GeV\) the approximation is valid.
6.2. EMISSION PROCESSES

CHAPTER 6. PHOTON PRODUCTION

\[ k_\gamma = \frac{1}{0.801 + 0.049L + 0.014L^2}. \] (6.9)

For cosmic ray protons in the energy range of \(0.1 - 10^5\) TeV and secondary photon energies of \(E_\gamma / E_p \geq 10^{-3}\), the Kelner method is accurate to within 10% or less. The error is greatest at low energies (which are better covered by the Kamae method) and at high energies where there is a paucity of experimental data. However, protons with energies greater than \(10^5\) TeV are exceedingly rare even in the very energetic environments considered in this dissertation, so the Kelner method is assumed to extend to arbitrary proton energies as needed.

Further wrinkles

In the preceding two sections I have described two different approaches for calculating photon emission via pion decay. The Kamae method is best for proton energies in the range \(1.2\) GeV < \(T_p < 500\) TeV, while the Kelner method works for proton energies between 0.1 TeV and \(10^5\) TeV.

In order to use both methods where they are most accurate, I use a weighted average of the two. First, both methods are used to calculate photon spectra for all protons in the distribution. Then the two are joined using a transition region of \(1 - 10^4\) GeV in photon (not proton) energy. The joining is performed as below:

\[ N_\gamma = \begin{cases} 
N_{\text{Kam}} & E_\gamma < 1 \text{ GeV} \\
E_\gamma^2 10^2 N_{\text{Kel}} & 1 \text{ GeV} < E_\gamma < 10^4 \text{ GeV} \\
N_{\text{Kel}} & 10^4 \text{ GeV} < E_\gamma 
\end{cases} \] (6.10)

where \(E_{100} = E_\gamma / (100 \text{ GeV})\) is the ratio between the photon’s energy and the transition energy of 100 GeV.

The final issue related to pion production is the use of \(\text{He}^{2+}\) as both a part of the ambient material encountering the shock and as a species of accelerated particle. High-energy helium nuclei should scatter off of the ambient medium just as high-energy protons do, and in the ultra-relativistic limit a \(\text{He}^{2+}\) nucleus may be thought of as four nucleons with \(1/4\) the energy of the nucleus. The cross section of helium is not four times the \(p - p\) cross section, however, because the nuclear forces within \(\text{He}^{2+}\) tightly bind the nucleons and reduce the effective scattering area. To incorporate these effects, I use the formula suggested in Appendix A of Orth and Buffington (1976),

\[ \sigma_{\text{heavy}} = \left( A^{3/8}_{\text{cr}} + A^{3/8}_T - 1 \right)^2 \sigma_{p - p}, \] (6.11)

where \(A_{\text{cr}}\) and \(A_T\) are—respectively—the atomic masses of the primary cosmic ray and the target.
Orth and Buffington (1976) do warn that the error associated with Equation 6.11 when both \( A_{cr} \) and \( A_T \) are greater than 1 is unknown. As the error is likely to rise with increasing \( A \), I restrict myself only to \( \text{H}^+ \) and \( \text{He}^{2+} \) nuclei in this dissertation. Extension to heavier nuclei is possible at a later date once experimental data for these nuclei are considered.

**Output**

In Figure 6.2, I show the pion-decay spectra calculated using the above methods, for the particle distributions given in Figure 6.1. Several features of note are visible in the figure. First, the spectrum due to the power law in CR energy is almost flat; there is a slight concave upwards shape due to how the Kelner et al. (2006) method calculates the cross section at high energies. While the two methods agree quite well in the averaging region of \( 1 < E_\gamma < 10^4 \) GeV—and particularly around 100 GeV—they disagree rather dramatically at low and high energies, as expected. Finally, there is a minimally visible bump at the lowest energies due to primary CRs pulled from the thermal peaks in Figure 6.1; if the injection rate for this shock were lower, and the normalization of the power law
region of the CR spectrum correspondingly depressed, the relative strength of the thermal peak would be enhanced.

### 6.2.2 Synchrotron

Synchrotron photons, generated as ultrarelativistic electrons gyrate in a magnetic field, are expected to comprise the bulk of emission from GRB afterglows (see the discussion in Chapter 2). Unlike the pion decay process described previously, the synchrotron process is extremely straightforward—if tedious—to describe and solve analytically (e.g., Pacholczyk, 1970; Rybicki and Lightman, 1979). The power per unit frequency $d\nu$ radiated by a single electron is

$$ P(\nu) = \sqrt{3} \frac{q^3 B}{m_e c^2} F \left( \frac{\nu}{\nu_c} \right), \quad (6.12) $$

where $\nu_c = 3\gamma_e^2 q B / (4\pi m_e c)$ is the critical frequency for an electron whose plasma-frame Lorentz factor is $\gamma_e$, and which is gyrating in a magnetic field of strength $B$. The frequency dependence is contained in the function $F$, defined by

$$ F(x) = x \int_x^\infty K_{5/3}(z) \, dz \quad ; \quad x = \frac{\nu}{\nu_c}, \quad (6.13) $$

which in turn contains a modified Bessel function of the second kind. Note that I omit in Equation 6.12 a factor of $\sin(\alpha)$ that is present in typical analytical treatments of synchrotron radiation. In ignoring the pitch angle $\alpha$ I am assuming that the magnetic field is always sufficiently tangled as to appear isotropic to relativistic electrons, and that it is always perpendicular to the motion of these electrons; the error incurred is a factor of at most a few, which is minor compared against other considerations. Note also that the critical frequency $\nu_c$ is not the frequency at which the spectrum given in Equation 6.12 peaks; that occurs at $\nu_m \approx 0.29\nu_c$.

### The magnetic field

The magnetic field in which the electrons are gyrating affects both the critical frequency of emitted photons and the rate at which those photons are produced. The turbulence in this field is affected by fluid compression even when (as is the case with parallel shocks) the mean field is not. To calculate the strength of the field at any given point in the shock structure, the following equation is used:

$$ B(x) = B_0 \sqrt{\frac{1}{3} + \frac{2}{3} x^2(x)} \quad (6.14) $$
(see the discussion preceding Equation (8) of Völk et al., 2002). The parameter \( \xi(x) \) is the compression of the fluid above the ambient density, i.e. \( \xi(x) = n(x)/n_0 = \gamma(x)\beta(x)/(\gamma_0\beta_0) \) by the Rankine-Hugoniot equations. This assumes that the turbulent field is tied to the fluid and evolves with it. The fractions 1/3 and 2/3 arise because only turbulence in the plane of the shock (i.e. two of the three dimensions) is compressed; turbulence along the shock normal is unaffected by passage through the shock structure.

A further factor in the strength of the magnetic field is amplification beyond even that predicted by Equation 6.14. Such amplification is commonly assumed or required in simulations of non-relativistic shocks (e.g., Vladimirov et al., 2006, 2008). And, as discussed in Santana et al. (2014), amplification is expected to occur at the shock of GRB afterglows in order to explain the intensity of emission at the start of the main decay phase. However, in the interest of simplicity no additional amplification is used in the results to be presented.

**Radiative cooling**

Equation 6.12 governs how energetic electrons lose energy via the synchrotron process. For highly energetic particles, this cooling could (conceivably) be ignored for short periods of time, but most particles interact with the shock structure for a long enough period that radiative losses must be considered.

While the code is running, at every (plasma-frame) time step \( \Delta t_{pf} \) the losses due to Equation 6.12 are computed. When that equation is integrated over losses at all frequencies, the total change in particle momentum becomes (c.f. Rybicki and Lightman, 1979, eq. 6.7b)

\[
\Delta p_{t,pf} = -\frac{1}{6\pi} \sigma_T \frac{P_{e,pf}^2}{m_e^2 c} B^2(x) \cdot \Delta t_{pf}.
\]

In the above equation, \( \sigma_T = 8\pi r_0/3 \) is the Thomson cross section in terms of the classical electron radius \( r_0 = e^2/(m_e c^2) \), and \( B(x) \) is given by Equation 6.14.

**Self-absorption and cooling breaks**

As discussed in Section 2.2.2, analytical treatments of GRB afterglow emission expect three breaks in the synchrotron spectrum. In rough order of increasing frequency, they are due to (1) the self-absorption of synchrotron photons by the electrons emitting them, (2) the electron energy at which the distribution begins (or peaks), and (3) the electron energy at which radiative losses between successive crossings of the shock exceed the energy gained from those shock crossings.

The second and third breaks arise naturally out of the Monte Carlo computation. The population
of electrons that crosses the shock naturally assumes a distribution with a sharp thermal peak and a high-energy tail; the break in the synchrotron spectrum associated with the thermal peak is typically quite clear. The presence of the cooling break depends on whether the maximum energy is limited by the size of the shock structure or by cooling losses. If energetic electrons are advected downstream—and beyond the scope of the simulation—before they can lose significant amounts of energy, then no cooling break will be observed. On the other hand, if the downstream region is sufficiently large, then a break in the electron spectrum at high energies will be visible.

The lowest spectral break, due to synchrotron self-absorption (SSA), must be included explicitly during the calculation of photon production. There are two key components of the SSA calculation: the cutoff frequency $\nu_a$ and the behavior of the synchrotron spectrum below $\nu_a$.

The physical basis of SSA is that no source of incoherent photons can be brighter at a particular frequency than a black body at an equivalent thermodynamic temperature would be (this is simply a restatement of the fact that black bodies are theoretically perfect radiators). If the intensity exceeds this limit, individual photons are likely to scatter within the source, exchanging energy and serving to quasi-thermalize the resultant photon spectrum. What the observer sees at a particular frequency, then, is only emission from a layer of material at the outer edge of the source; the situation is analogous to seeing thermal emission from the surface of the Sun. The way forward is clear: determine, based on the size of the region and the population of electrons, the frequency at which the optical depth of synchrotron emission becomes unity. For lower energies, the region will be optically thick and so the emitted synchrotron spectrum will be processed before reaching the observer; at higher energies the region is optically thin and so one observes the true synchrotron spectrum.

Skipping a somewhat lengthy and tedious calculation based on the Einstein absorption coefficients (see Rybicki and Lightman, 1979, for the full derivation), the absorption coefficient $\alpha_\nu$ at a particular frequency may be expressed as

$$\alpha_\nu = -\frac{c^2}{8\pi\nu^2} \int P(\nu, E) \cdot E^2 \frac{\partial}{\partial E} \left( \frac{1}{E^2} \frac{dn(E)}{dE} \right) dE,$$

(6.16)

where $P(\nu, E)$ is the power radiated per unit frequency at frequency $\nu$ by an electron with energy $E$ (in other words, Equation 6.12) and $dn(E)/dE$ is the number density of electrons per energy bin $dE$ (equal to $dN/d\rho$ from Section 5.3.1, with a rescaling from total number into number density). The frequency of SSA photons is typically far below the critical frequency for electrons in the thermal

\footnote{Strictly speaking, another means of particle escape is possible: diffusion upstream to a free-escape boundary (see the discussion in Section 3.5). However, for relativistic shocks the degree of upstream escape is vanishingly tiny, as particles must diffuse many mean-free-paths against a flow moving a substantial fraction of the speed of light.}

\footnote{I.e., lasers and other sources of coherent light excepted.}
population, i.e. $\nu_a/\nu_c \ll 1$. This allows for replacement of the complicated function $F(x)$ from Equation 6.13 with its asymptotic limit:

$$F(\nu/\nu_c) \rightarrow \frac{4\pi}{\sqrt{3}} \frac{1}{\Gamma(1/3)} \left( \frac{\nu}{2\nu_c} \right)^{1/3},$$  

(6.17)

with $\Gamma(x)$ being the Gamma function. I now rewrite Equation 6.12 in this limit as

$$P(\nu, E) = \frac{4\pi q^3 B}{2^{1/3} \Gamma(1/3) m_e c^2} \left( \frac{\nu}{\nu_c} \right)^{1/3},$$  

(6.18)

where $\nu_c$ (defined as in Equation 6.12) carries the energy dependence. Substitution into Equation 6.16 yields a formula for $\alpha_\nu$ in terms of $\nu$ and an involved, but $\nu$-independent, integral.

$$\alpha_\nu = \frac{1}{2^{1/3} \Gamma(1/3)} \frac{q^3 B}{m_e c^2} \nu^{-5/3} \int \nu_c^{-1/3} \left( \frac{2}{p} \frac{dn(p)}{dp} - \frac{\partial^2 N(p)}{\partial p^2} \right) dp$$  

(6.19)

Note the change of integration variables between Equations 6.16 and 6.19. I mentioned above how $n(E)$ was defined, but I restate it here. Since $n(p)$ is the number density of particles per momentum bin $dp$, the partial derivative term in the integral is the slope of the $dN/dp$ curve (after the renormalization from number to density). Equation 6.19 has a closed form when the particle distribution forms a power law, which is provided by Equation 6.53 in Rybicki and Lightman (1979).

To determine $\nu_a$ from $\alpha_\nu$, one solves the equation $\Delta R_{pf} \alpha_\nu = 1$—that is, one finds the frequency for which the physical depth of the region corresponds to an optical depth of 1.\(^4\) From the form of Equation 6.19 it can be shown (taking into account the dependence of $\nu_c$ on the magnetic field strength) that $\nu_a \propto (\Delta R_{pf})^{3/5} B^{1/5}$, but the constant of proportionality cannot be determined before the precise nature of $dN/dp$ is known.

Determining the behavior of the absorbed synchrotron spectrum below $\nu_a$ is more straightforward. The exact approach varies slightly based on whether the electrons producing the self-absorbed emission are all at one energy (the minimum energy $\gamma_m$ of the electron spectrum) or if the self-absorption frequency falls above $\nu_c$ for the lowest-energy electrons. I assumed the former case in deriving the value of $\nu_a$ above, so I will continue on those grounds below.

Although the electrons responsible for SSA are not in a thermal spectrum, one can still assign an effective temperature $T_{eff}$ to the population: $3k_B T_{eff} = \gamma_m m_e c^2$. The emission spectrum from a

\(^4\)If the population of particles changes with depth, as will be the case for the sample afterglow discussed in Chapter 8, the formula changes slightly. Instead of $\Delta R_{pf} \alpha_\nu = 1$, one uses the equation $\int \alpha_\nu dr_{pf} = 1$. Conveniently, the prefactor of $\nu^{-5/3}$ is the same no matter what population is being considered in Equation 6.19.
black body at temperature \( T_{\text{bb}} \) is given by Planck’s Law,

\[
I_\nu = \frac{2\hbar \nu^3}{c^2} \frac{1}{\exp\left(\frac{\hbar \nu}{k_B T_{\text{bb}}}\right) - 1}.
\]  

(6.20)

I am interested in the maximum radiation from the electrons at “temperature” \( T_{\text{eff}} \), so substitute \( T_{\text{bb}} = T_{\text{eff}} \). The intensity of radiation at a particular frequency depends on the ratio \( \hbar \nu / (k_B T_{\text{eff}}) \). Since the self-absorbed photons are far lower in energy than the electrons that produced them, \( \hbar \nu_a / (k_B T_{\text{bb}}) \ll 1 \). This places SSA photons firmly in the Rayleigh–Jeans limit of Planck’s Law, and so

\[
I_\nu \propto \nu^2.
\]  

(6.21)

Strictly speaking, the Rayleigh–Jeans limit is \( I_\nu \propto T_{\text{bb}} \nu^2 \), which can be used to explain the behavior in the other scenario I mentioned above. If the self-absorption frequency falls above \( \nu_m \), then the radiation at any given frequency is due to electrons of differing energies. The peak frequency of synchrotron emission varies as \( \nu \propto \gamma^2 \), or \( T_{\text{eff}} \propto \nu^{1/2} \). With the substitution of \( T_{\text{eff}} \) for \( T_{\text{bb}} \) in the Rayleigh–Jeans limit, one finds that the SSA behavior of such a system is \( I_\nu \propto \nu^{5/2} \), rather than \( \nu^2 \).

One final note about the above derivation is that it ignores any angular information about the radiation, assuming instead that the magnetic field is chaotic and the synchrotron photons are isotropic. The issue is treated with substantially more generality in the works of Ginzburg and Syrovatskii (1965, 1969), which may be necessary for handling SSA in the context of GRB afterglows.

As will be shown in Chapter 8, only synchrotron photons produced near the subshock itself encounter an isotropic electron distribution on their passage out of the jet. The velocity gradient present in the post-shock plasma ensures that photons produced further downstream see increasingly anisotropic electron distributions—if transformed into the frame in which the electrons are isotropic—as they near the surface of the jet.

Output

All features discussed in this section are present in Figure 6.3, which uses the blue (plasma frame electrons) curve of Figure 6.1. The peak of the synchrotron distribution is unusually high because of the somewhat artificial parameters chosen for this shock. The coarse scattering makes it very easy for electrons to gain energy, the magnetic field is low enough that cooling losses are minimal, and the upstream FEB was chosen large enough to allow for a power law that extends to very high energies. All of these combine in a maximum plasma-frame energy of \( 10^7 - 10^8 \) GeV for electrons, which is higher than expected for more physically-motivated parameters.
At lower photon energies than the peak, the $E d\Phi/dE$ spectrum settles into the $E^{1/2}$ power law expected from a $p^{-2}$ electron energy distribution function. Small wiggles are visible at the low end, due to the reflections of the thermal peak.\(^5\) The location of the self-absorption frequency $\nu_a$ was determined using Equation 6.19 and a reasonable downstream region size. As is visible in Figure 6.3, the bump in the synchrotron spectrum due to the thermal (i.e. uninjected into DSA) population is below $E_a = h \nu_a$, and so would be obscured by self-absorption. It should be mentioned that whether the thermal peak is absorbed depends on the magnetic field. As discussed above, $\nu_a \propto B^{1/5}$, but the location of the thermal peak will vary as $\nu_{pk} \propto B$, in keeping with the definition of $\nu_c$ in Equation 6.12. The presence or absence of the thermal peak in a radio spectrum might therefore be used to gauge the strength of the downstream magnetic field.

I note that the sharpness of the SSA break is artificial, as the SSA spectrum was added by hand after the synchrotron spectrum was already calculated. A totally self-consistent calculation would be smoother, reflecting the fact that optical depth is a probability rather than a binary switch. This would be handled by radiative transfer models that are significantly more complex than what is presented in this text.

---

\(^5\)As discussed in Section 4.4.2.3, these reflections of the thermal peak are populations of particles that have all crossed the shock once, or three times (UpS$\rightarrow$DwS$\rightarrow$UpS), or five times, etc. This is an artifact of the coarse scattering limit; the smoothing into a power law happens at much lower energies when finer scattering is assumed.
6.2. EMISSION PROCESSES

CHAPTER 6. PHOTON PRODUCTION

here, as the Monte Carlo model focuses primarily on the particle—rather than photon—diffusion.

6.2.3 Inverse Compton

Compton scattering, in which a photon scatters elastically off of an electron and loses energy, is now approaching its 100th anniversary in the literature (Compton, 1923). If viewed in reverse, the process may appear as energetic electron scattering off of—and donating energy to—a photon, and is correspondingly called the inverse Compton process. Feenberg and Primakoff (1948) were the first to treat the inverse Compton process as a means by which energetic cosmic ray electrons might lose energy during their travel to Earth. Somewhat later, the resultant photons were considered (Savedoff, 1959; Felten and Morrison, 1963).

As energies increase, the scattering process takes on a quantum-mechanical aspect, and the classical Thomson cross section,

$$\sigma_T = \frac{8}{3} \pi r_0^2 = \frac{8}{3} \pi \frac{e^2}{m_e c^2}, \tag{6.22}$$

(where \(r_0 = e/(m_e c^2)\) is the classical electron radius) must be replaced by the more complicated Klein-Nishina cross section (Klein and Nishina, 1929),

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2}{2} \left( \frac{\alpha}{\alpha_{in}} \right)^2 \left[ \frac{\alpha}{\alpha_{in}} + \frac{\alpha_{in}}{\alpha} - 1 + \cos^2(\theta) \right]. \tag{6.23}$$

In the above equation, and elsewhere in this section, \(\alpha = E_\gamma/(m_e c^2)\) is the ratio of a photon's energy to the electron rest mass-energy of 511 keV. The subscript “in” denotes photons before scattering, while no subscript refers to outgoing photons after scattering. Importantly, in Equation 6.23 the two photon energies are measured in the rest frame of the electron: either a sufficiently energetic photon or a sufficiently energetic electron can push \(\alpha_{in}\) beyond the classical limit. Note also that Equation 6.23 is a differential cross section, whose value depends on the angle of scattering \(\theta\) (it is independent of azimuth).

The Inverse Compton Kernel

At the core of the inverse Compton calculations presented here is the work of Jones (1968), who analytically determined both the rate and the spectrum of photons generated by the inverse Compton process. That derivation begins with a 6-dimensional derivative in time, photon energies (incoming and outgoing) and scattering angles, which are integrated through a series of approximations and identities. The most important single assumption is that the photon distribution is isotropic; this
sets the lab frame into which the outgoing photons will be transformed. Many authors (e.g., Baring et al., 1999; Bošnjak et al., 2009; Lee et al., 2012b) have used Equation (9) from that paper:

$$\frac{d^2N}{d\gamma d\alpha} \approx \frac{2\pi r_0^2 c}{\alpha \gamma_e^2} \left[ 2q \ln q + (1+2q)(1-q) + \frac{1}{2} \left( \frac{4\alpha_{in}\gamma_e q}{1+4\alpha_{in}\gamma_e q} \right)^2 (1-q) \right], \quad (6.24)$$

which is sufficient for isotropic electron distributions. All quantities are given in the laboratory frame. The new variable $q = \alpha / [4\alpha_{in}\gamma_e^2(1-\alpha/\gamma_e)]$ contains the information about the outgoing photon's energy, and lies within the range $1/(4\gamma_e^2) \leq q \leq 1$. The “$\approx$” signifies the termination of a series expansion at terms of order $1/\gamma_e^2$ or higher, which is typically sufficient.

The inverse Compton scattering taking place in this work is that of accelerated electrons off of photons in the cosmic microwave background (CMB). As I mentioned previously, the Jones (1968) derivation relies on the assumption that the photon distribution is isotropic. This requires that electrons be transformed into the ISM frame, where the CMB is (more or less) isotropic before calculating photon production. The transformation from the plasma frame of the electrons into the ISM frame can introduce severe anisotropies in the context of GRB afterglows, as the two frames are related by a Lorentz factor on the order of several to a few hundred.

The anisotropic electron distribution will be strongly beamed in the direction of the jet's motion; this may have significant effects on the observed photons, as the inherent anisotropy of inverse Compton emission is no longer counteracted by an isotropic electron distribution. The most self-consistent calculation of inverse Compton emission must include the angular relation between the outgoing photons and the electron's momentum. This, in turn, uses Equation (6) of Jones (1968):

$$\frac{d^3N}{d\eta d\gamma d\alpha} = \frac{\pi r_0^2 c \alpha}{2\gamma_e^5 \alpha_{in}^2 (1-\alpha/\gamma_e)} \left[ \eta^2 - 2\eta + 2 + \frac{(\alpha/\gamma_e)^2}{1-\alpha/\gamma_e} \right] \times S(\eta; \eta_1, \eta_2), \quad (6.25)$$

The angular dependence is contained within $\eta = 1-\beta_e \cos(\theta') \approx 1-\cos(\theta')$, where $\theta'$ is the scattering angle with two very large caveats. First, the angle is measured in the electron rest frame, not the laboratory/ISM frame. Second, $\theta' = 0$ points antiparallel to the electron's momentum vector, rather than parallel to it. The function $S(\eta; \eta_1, \eta_2)$ is a composition of two Heaviside step functions,

$$S(\eta; \eta_1, \eta_2) = \begin{cases} 1 & \eta_1 \leq \eta \leq \eta_2 ; \eta_1 = \frac{\alpha}{2\alpha_{in}\gamma_e^2(1-\alpha/\gamma_e)} ; \eta_2 = \frac{2\alpha}{\alpha_{in}(1-\alpha/\gamma_e)} \\ 0 & \eta < \eta_1, \eta_2 < \eta \end{cases}, \quad (6.26)$$

that together limit the energy outgoing photons may have.

The results of Equations 6.25 and 6.26 are illustrated in Figure 6.4. In the figure, three different electrons (with ISM-frame Lorentz factors of $10^5$, $10^7$, and $10^9$) have scattered off a monoenergetic,
isotropic photon field of energy $\alpha_{in} = 10^{-10}$. The horizontal axis measures ISM-frame energy of the outgoing photon. The vertical axis shows the angular dependence, with a departure from Equation 6.25: in Figure 6.4, the angle $\eta = 0$ is now parallel instead of antiparallel to the electron's momentum vector. It is, however, still measured in the rest frame of the electron. Unsurprisingly, the maximum energy attained by outgoing photons occurs when the photons have a head-on collision with the electron and return whence they came; this is true of classical elastic collisions as well. The expected $\gamma_e^2$ dependence of said maximum energy is visible in the figure, though it is slightly diminished in the bottom panel by the onset of the Klein-Nishina regime. Regarding the spectra themselves, two points must be made. First, $\eta = 1$ corresponds to scattering at right angles to the electron's momentum. Relativistic aberration beams any photons with $\eta < 1$ into a cone of half-opening angle $1/\gamma_e$. It can be seen in Figure 6.4 that this cone is almost uniform, as the angular dependence of those photons is slight. The second point is that very few of the scattered photons will be in that beamed cone. The vast majority of photons are scattered at smaller angles, so the bulk of the photon distribution is near $\eta = 2$; this trend increases sharply with increasing electron energy. Since these photons make up the bulk of the scattered population, the naïve approximation that all emission occurs within a $1/\gamma_e$ cone is a poor one.

To properly account for the photons scattered at smaller angles (i.e. larger $\eta$), the photons must be transformed back into the ISM frame. Because of the small angles and large Lorentz factors, the transformation is fraught with danger when performed numerically. Rather than transforming $\cos \theta$ via the standard relativistic aberration formula, (see Equation 6.28 below), it is safer to transform $\eta$. The results of this transformation are displayed in Figure 6.5, which shows the same three photon distributions as did Figure 6.4, but now in the ISM frame. Figure 6.5 demonstrates that the behavior of each photon distribution at its maximum energy depends quite similarly on the energy of the electron. As I wrote above, the angular dependence of photons within a $1/\gamma_e$ cone is minimal. Expanding $\eta$ about $\theta = 0$, one arrives at $\log_{10} \eta = -\log_{10}(2) - 2\log_{10}(\gamma_e)$. This is indeed where the photon spectrum steepens considerably in the upper two panels of Figure 6.5. In the bottom panel the electron is energetic enough that the highest-energy photons are all beamed into an even smaller region than the finest logarithmically-spaced bin in $\eta$.

Consequences

Figure 6.5 additionally points to a curious mathematical coincidence. The extent of the photon cone—the maximum separation between the electron's and a photon's momentum vectors in the ISM frame—is identical for all three electron energies. This is a physical effect rather than a numerical one, and is caused by the interaction of Equation 6.26 and relativistic beaming. In that equation
Figure 6.4 Plots of $d^3N/(dt d\eta d\alpha)$ for an isotropic, monoenergetic photon field and three different electron energies. The angular quantity $\eta = 1 - \cos(\theta_{\text{out}})$ is given in the electron rest frame, not the ISM frame (c.f. Figure 6.5). For all three plots, $\alpha_{\text{in}} = E_{\gamma,\text{in}}/(m_e c^2) = 10^{-10}$ in the ISM frame. In a departure from the notation of Jones (1968), $\gamma = 0$ is parallel rather than antiparallel to the electron’s ISM-frame momentum. **Top panel:** photon production from a single electron with energy $E = 10^5 m_e c^2$. **Middle panel:** as above, but for an electron with energy $E = 10^7 m_e c^2$. **Bottom panel:** as above, but for an electron with energy $E = 10^9 m_e c^2$. 
Figure 6.5 Same as in Figure 6.4, but with $\eta = 1 - \cos(\theta_{\text{out}})$ transformed into the ISM frame. The vertical scale has switched from linear to logarithmic to preserve spacing at the very small angles necessary for the calculation. Note the structure present at $\log_{10} \eta = -10.3$ and $-14.3$ in the top and middle panels, and see the text for discussion.
one sees that the minimum deflection taken by a photon corresponds to \( \eta_1 \). It is also apparent from Figures 6.4 and 6.5 that the photons that are deflected the least gain the least amount of energy in the scattering. So consider photons being scattered at almost exactly \( \eta_1 \), and restrict to photons at low enough energies that \( \alpha / \gamma_e \ll 1 \):

\[
\eta_1 = \frac{\alpha}{2 \gamma_e^2 \alpha_{\text{in}} (1 - \alpha / \gamma_e)} = \frac{\alpha}{2 \gamma_e^2 \alpha_{\text{in}}} \left[ 1 + \frac{\alpha}{\gamma_e} + \left( \frac{\alpha}{\gamma_e} \right)^2 + \left( \frac{\alpha}{\gamma_e} \right)^3 + \ldots \right] \\
\approx \frac{\alpha}{2 \gamma_e \alpha_{\text{in}}}. \quad (6.27)
\]

The formula for relativistic beaming may be found in any undergraduate textbook covering relativity, and is

\[
\cos \theta = \frac{\cos(\theta') + \beta}{1 + \beta \cos(\theta')}. \quad (6.28)
\]

In the above equation, the primed and unprimed frames are related by the velocity \( \beta \), and both cosines are oriented so that \( \theta \) (or \( \theta' \)) = 0 points in the direction of \( \beta \). This latter assumption does not hold when dealing with Equation 6.25: \( \eta \) was defined pointing opposite the electron's velocity, so the primed cosines must be negated in the formula for beaming. When this change is made, a given \( \eta_1 \) in the electron rest frame is beamed to an ISM-frame \( \eta \) of

\[
\eta_{\text{ISM}} = 1 - \cos \theta = \frac{(1 - \beta_e)(1 + \cos \theta')}{1 - \beta_e \cos(\theta')} \\
\approx \frac{1/(2 \gamma_e^2)}{2 - \eta_1} \\
= \frac{2 - \eta_1}{2 \gamma_e^2 - 2 \gamma_e^2 \beta_e + 2 \gamma_e^2 \beta_e \eta_1} \\
= \frac{2 - \eta_1}{2 \gamma_e^2 (1 - \beta_e) + \beta_e \alpha_{\text{in}}} \\
\approx \frac{2 - \eta_1}{1 + \beta_e \alpha_{\text{in}}}. \quad (6.29)
\]

At both places where \( \approx \) appears, I have expanded \( \beta_e \) in terms of \( \gamma_e \) for \( \gamma_e \gg 1 \). The denominator of Equation 6.29 is almost entirely independent of the electron energy—\( \beta_e \approx 1 \) for any electron emitting via the inverse Compton process. In the numerator there is only a weak dependence on energy, which requires that \( \eta_1 \) be small. This is merely a restatement that the outgoing photon's
6.3 PROCESSING

CHAPTER 6. PHOTON PRODUCTION

energy must be much less than the maximum.\textsuperscript{6}

The preceding discussion has two very important consequences. The first is that—as was seen in Figure 6.5—an electron’s inverse Compton emission will only be observed if the electron’s pitch angle falls within the very narrow cone given by Equation 6.29. This is a minuscule fraction of the possible pitch angle range, even once electron momenta are transformed into the ISM frame. However, the number of particles swept up by the jet is so vast that this range is well populated in physical space. The second consequence is that Equation 6.24 may be used to calculate inverse Compton emission instead of the more involved Equation 6.25. At present the code lacks the granularity to handle the very small pitch angles shown in Figure 6.5, so determining the angular dependence within that range is irrelevant.\textsuperscript{7}

The narrowness of the emitted cone informs how the Monte Carlo code determines flux from inverse Compton emission. The primary consequence is that all upscattered photons depart essentially parallel to the cosmic-ray electron. It is further assumed that the electrons’ angular distribution, on average, is normal to the shock in the ISM frame (as follows from transforming an isotropic plasma-frame distribution into the ISM frame). So the total photon production due to the inverse Compton process leaves the afterglow in an extension of the jet’s cone. The area swept out by the cone at the distance to Earth is used for calculating the flux.

Output

The emission due to the inverse Compton process is given in Figure 6.6. The spectrum shows little in the way of structure. The bumps at low energies are caused by the thermal peak and its reflections. At the highest energies the Klein-Nishina effect suppresses the emission and causes the rapid decline in flux. Although the precise shape of the turnover might contain information about the shape of the electron distribution at the highest energies, the photon counts are likely to be too low to generate good statistical data.

6.3 Processing

Because the scenario of a GRB afterglow involves reference frames (the shock frame, the local plasma frame, and the explosion—or ISM—frame) that may move at significant fractions of the speed of light, the maximum gain in energy for a photon undergoing inverse Compton scattering is a factor of \(4\gamma_e^2\) over the initial energy, or \(\alpha = 4\gamma_e^2\alpha\text{in}\). The assumption that \(\alpha \ll 4\gamma_e^2\alpha\text{in}\), when substituted into Equation 6.27, has the effect that \(\eta_1 \ll 2\), and so the \(\eta_1\) term in the numerator is essentially ignorable.

\textsuperscript{6}By “granularity” above, I refer to the number of particles rather than any limit on the code’s ability to calculate pitch angles. Even simulations with momentum splitting into tens of thousands of particles fail to populate the finest angular bins required to make use of Equation 6.25.
light relative to each other, numerous effects from special relativity must be considered.

The calculations described in Sections 6.2 result in a diverse set of spectra. Each emission process has its own minimum and maximum energy over which the spectrum was calculated. Pion production might involve multiple species, and inverse Compton might have taken place off of several photon fields. The reference frames in which the spectra were calculated differ as well: both pion decay and synchrotron are calculated in the plasma frame of the generating population, while inverse Compton is determined using the ISM-frame particle distribution. Compounding the complexity, the emission subroutines are calculated at each grid zone, so for modified shocks the relative velocity between the plasma and ISM frames varies within the shock structure.

Before further analysis may be done, the photon spectra from the various processes/species must be processed into a coherent and consistent form. This takes place in several steps, outlined below.

1. First of all, if more than one species underwent pion production, or if more than one photon field was used for inverse Compton emission, the multiple spectra are summed into a single set—one summed spectrum for each grid zone, for each of the two emission processes.

2. Next, the two processes that occur in the plasma frame (pion production and synchrotron) are transformed from the plasma frame into the ISM frame. It is assumed that both processes result in isotropic photon distributions in the plasma frame, as this makes the transformations
vastly simpler. Instead of the complicated energy- and angle-dependent transformations done for Figure 6.5 it is sufficient to split the photon emission at energy $E_{pf}$ into bins of width $\Delta \cos(\theta)$ in the plasma frame, and transform each bin's energy by the standard formula:

$$\frac{E_{\text{ISM}}}{E_{pf}} = D \equiv \gamma \cdot (1 + \beta \cos(\theta)).$$  \hspace{1cm} (6.30)

The quantities $\gamma$ and $\beta$ are the Lorentz factor and velocity separating the two frames, and $D$ is discussed further in Section 5.4.3. This transformation is done separately for all grid zones where emission occurred, because the nature of the shocks allows for significant zone-to-zone velocity differences.

3. The actual flux measured at Earth is enhanced beyond the transformation of the photon energy. Relativistic beaming adds two factors of the Doppler factor $D$ by reducing the solid angle into which the photons are emitted. As well, time dilation decreases the duration between photon arrivals at Earth. In sum, three powers of the Doppler factor must be incorporated into the plasma-frame flux to complete the transformation from what is emitted in the plasma frame to what is observed at Earth. At this early stage, it is assumed that the observer sits directly on the axis of the jet, so that $D \approx 2\gamma$. While obviously most jets are observed off-axis\(^8\), on-axis emission is a necessary first step before further complications are introduced.

4. Finally, the emission from the entire shock region is summed into a single spectrum covering all emission processes from all grid zones. This spectrum is only that of the current time step; see Section 6.4 for how previously-shocked material contributes to observed emission.

One final point about the processing of photon emission must be mentioned. Redshift is not taken into account in these calculations. At this stage it is assumed that the GRB happened relatively close to Earth. In any event, the effects of redshift are subordinate to those of the relativistic transformations, with the exception of very late times and very distant GRBs.

### 6.4 Emission shells

The process described in the previous sections is sufficient to calculate the photon production from the shock structure in the steady-state case. The afterglows of GRBs are not, however, steady-state. As the shock sweeps up material and decelerates, the downstream population of shocked thermal particles and cosmic rays must change in response. Furthermore, particles that were swept up

\(^8\)From a purely probabilistic argument as much as any other reason.
Figure 6.7 Depiction of how the code handles emission from multiple time steps. For an ISM-like circum-burst medium, the upstream plasma is identical at all time steps (though for a wind-like medium, the properties would change with time). Each time step results in a new shell of shocked material, whose properties are assumed to be constant within the shell. The temporal evolution of these shells is described in the text. Photon production is calculated for the precursor region and all existing shells at each time step.

very early in the afterglow and have since advected far downstream will still contribute to photon production, even though they are too far downstream to interact with the shock (and so beyond the scope of the Monte Carlo code). How, then, to handle the (possibly very) different particle populations?

The solution borrows a technique from the crhydro code described in, e.g., Ellison et al. (2007). As illustrated in Figure 6.7, I break the afterglow into several discrete time steps. Each time step has a particular shock speed (assumed constant throughout the time step) and duration. The material swept up during a given time step is uniform within the volume dictated by the swept-up volume and the conservation relations. Behind the current downstream material are frozen-in shells of material swept up in previous time steps.

Each shell of shocked material contributes to photon production based on its state at the current time. Between time steps the most recently shocked material is modified by two processes. The first is cooling by the same emission processes described above. The second is cooling through adiabatic expansion. Between time steps $t_{i-1}$ and $t_i$, the particle distribution function will have been
modified, and it is the updated version that is then passed to the photon production subroutines at time $t_i$.

From the above description and Figure 6.7, it should be clear that photon production is a cumulative process. All particles that have encountered the shock, from the oldest to the most recent, contribute to the observed flux. The various particle distribution functions act as a history of the shock and its dynamics.

### 6.4.1 Cooling between timesteps

Between time steps where photon production is calculated the shocked plasma is still evolving. To determine the new particle distributions at the next time step, various cooling processes are taken into account. Although particles can lose energy in a variety of ways, I include only the two most important processes here: adiabatic expansion cooling and (for electrons only) synchrotron losses.

As the shocked plasma expands behind the advancing forward shock, the magnetic field turbulence off of which particles scatter is, on average, receding from the particles. This induces an adiabatic cooling of the particle distribution. To calculate the energy losses, consider a plasma of monoenergetic particles with adiabatic index $\Gamma$ at some pressure and volume $P$ and $V$. The equation of state is $P = (\Gamma - 1) u$, where $u = E/V$ is the energy density. For an adiabatic process one has $PV^\Gamma = K$ for some constant $K$, or $u = KV^{-\Gamma}(\Gamma - 1)^{-1}$. Taking the time derivative yields

\[
\frac{du}{dt} = \frac{K}{\Gamma - 1} \cdot -\Gamma \cdot V^{-\Gamma - 1} = \frac{1}{V} \frac{dE}{dt} - \frac{E}{V^2} \frac{dV}{dt}.
\]  

After some rearranging, the loss rate for particles at energy $E$ due to adiabatic cooling is

\[
\frac{dE}{dt} = -(\Gamma - 1) \frac{E}{V} \frac{dV}{dt}.
\]  

For ultrarelativistic particles $\Gamma = 4/3$, and Equation 6.32 reduces to $\dot{E}_{\text{ad}}$ as given in Reynolds (1998).

Electron energy losses due to synchrotron and inverse Compton emission behave in almost identical manners (e.g., Reynolds, 1998), and are treated simultaneously during the inter-step cooling,

\[
\frac{dE_e}{dt} = -\frac{4}{3} \cdot \sigma_T c \cdot \left( \frac{E_e}{m_e c^2} \right)^2 \cdot \left( \frac{B_{\text{tot}}^2}{8\pi} \right).
\]  

In the above equation, $B_{\text{tot}}^2 = B^2 + B_{\text{CMB}}^2$ includes contributions from both the local magnetic field and the equivalent magnetic field strength of the CMB, $B_{\text{CMB}} \approx 3.27 \mu G$. It is also assumed that all electrons responsible for synchrotron emission are fully relativistic, so $\nu \approx c$ (or $\beta \approx 1$).

The dominant loss mechanism for ions in terms of time-averaged power is pion production.
However, this method causes extremely rare, but catastrophic, losses. For values reasonable in the context of GRB afterglows, the interaction rate for any given ion is approximately $10^{-13} \text{s}^{-1}$ or rarer. This is one collision every several million years, a significantly longer period of time than I am simulating. I therefore ignore it as a cooling mechanism.

6.5 Summary

Once the Monte Carlo code described in Chapters 3–5 produces a smoothed, flux-conserving shock and the associated particle spectra, additional subroutines are called to generate the photon spectrum due to the particles. The energies covered by the three processes range from MHz radio to TeV $\gamma$-rays, allowing for comparison against observations across a wide range of wavelengths.

However, several important effects are not considered and are reserved for future work. They include (1) propagation effects between the site of the GRB and the observer at Earth, (2) pair creation in the afterglow and other high-energy photon-photon interactions, and (3) extremely detailed consideration of the angular distribution of high-energy electrons and their inverse Compton emission. Additional work will explore additional processes (such as photo-pair production) that are subordinate to the three photon production mechanisms as methods for energy loss in afterglows.
In this chapter I use the tools and methods outlined in the previous chapters to thoroughly describe a relativistic shock. This includes examining its structure, including the precursor; the spectra of particles accelerated by the shock; and the photons produced as seen by a distant observer.

7.1 The Shock Structure

In this section I describe the profile of the smoothed shock, and major departures from the unmodified case. Key parameters of the shock are summarized in Table 7.1 below. In the upper part of the table are numbers related to DSA, such as the speed of the shock and the location of the free escape boundaries. In the lower part are parameters used in normalizing the particle distributions and calculating photon production. The profile itself is displayed in Figure 7.1. In addition to the parameters mentioned in Table 7.1, the upstream environment contained 10% helium by number, and electrons in sufficient numbers for charge balancing.

In the top row of Figure 7.1 both the shock-frame speed ($\beta_x$) and $\gamma_x \beta_x$ are given, both scaled against their upstream values. Even for a nonlinear, modified shock, the upstream region shows...
Figure 7.1 Profile and fluxes of the smoothed shock, all measured in the shock frame. All quantities are given as fractions of their far upstream value. In all three rows, the solid black line is the smoothed value, while the dashed red line is the value taken from an unmodified shock with the same parameters. Top row: flow speed (solid black, dashed red lines) and specific volume (dot-dashed blue line) versus position. The eight colored and four black vertical lines represent the positions at which spectra were measured (upstream: zones 15, 18, 21, 24, 28, 36, 43, and 50; downstream: zones 78, 84, 90, and 97), and match the curves of the corresponding color in plots to follow. Middle row: momentum flux versus position. Bottom row: energy flux versus position.
Table 7.1 Key shock parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_0$</td>
<td>8.2</td>
</tr>
<tr>
<td>$n_0$</td>
<td>1.08 cm$^{-3}$</td>
</tr>
<tr>
<td>$B_0$</td>
<td>$3.27 \times 10^{-6}$ G</td>
</tr>
<tr>
<td>$T_0$</td>
<td>$10^6$ K</td>
</tr>
<tr>
<td>$r_{g0}$</td>
<td>$10^{12}$ cm</td>
</tr>
<tr>
<td>$N_g$</td>
<td>1600</td>
</tr>
<tr>
<td>UpS FEB</td>
<td>$-1000$ $r_{g0}$</td>
</tr>
<tr>
<td>DwS FEB</td>
<td>$+300$ $r_{g0}$</td>
</tr>
<tr>
<td>$x_{\text{sub}}$</td>
<td>$-1.9$ $r_{g0}$</td>
</tr>
<tr>
<td>$r_{\text{tot}}$</td>
<td>$3.02$ ($=r_{\text{RH}}$)</td>
</tr>
<tr>
<td>$R_{\text{shock}}$</td>
<td>$1.73 \times 10^{-2}$ pc</td>
</tr>
<tr>
<td>$\theta_h$</td>
<td>13.4°</td>
</tr>
<tr>
<td>Distance</td>
<td>1000 kpc</td>
</tr>
</tbody>
</table>

effects essentially no deviation from the far upstream value. This is consistent with the expectation of negligible cosmic ray backpressure in that region for relativistic shocks, which will be borne out by the spectra shown in section 7.2. Closer to the shock the modification is apparent, including the kink at $x \sim -2$ where the artificial smoothing begins. The twelve thin vertical lines in the top row (five in the left hand panel, five in the middle, and two on the right) denote locations where particle spectra are measured. The color and dash pattern correspond to the similarly-formatted curves in the plots of section 7.2.

The middle and bottom rows of Figure 7.1 show the shock frame fluxes, again scaled against the far upstream values. The black curves confirm that the solution presented is self-consistent, conserving both energy and momentum flux everywhere in the shock structure to very close approximation (almost perfect far upstream, and less than a percent downstream). Small deviations occur starting just upstream of the shock and continuing far downstream, but for multiple reasons. Just upstream from the shock, the artificial smoothing model enforces a slightly unphysical velocity gradient—too steep in places, too shallow in others—that is responsible for the dip and recovery of the fluxes. Just downstream from the shock the fluxes are higher than unity because of numerical effects related to scattering; neither the particle energies nor their angular distribution have yet been updated post-crossing, leading to a small spike in flux. The wiggles in the unmodified downstream
7.2. PARTICLE SPECTRA

CHAPTER 7. COMPLETE PICTURE OF A SMOOTHED SHOCK

fluxes are caused by the poorer statistics used while smoothing the shock. For the modified shock, the higher particle counts resulted in much smaller statistical fluctuations.

7.2 Particle Spectra

As repeatedly demonstrated in Chapters 4 and 5, modifying a shock’s velocity profile will in turn alter the manner in which particles interact with the shock. It is in fact a necessary part of the nonlinear feedback loop discussed in Section 1.4. In this section I illustrate this effect for the shock described in Section 7.1.

7.2.1 Shock Frame

The distribution functions for all three species of particles are given as functions of shock-frame total momentum in Figure 7.2. For this figure the population of thermal particles—those that crossed the shock once from upstream to downstream, then were swept downstream without further interaction with the shock—was ignored except for normalization purposes (see Section 5.3.1; the change in height of the downstream spectra is due to increasing volumes of grid zones away from the shock). Only cosmic ray particles, which crossed the shock from downstream to upstream and were injected into the acceleration process, are shown; the thermal particles were, of course, tracked, but are not included for reasons of clarity. Each species’ spectra show three distinct stages: (1) a bump around the lowest energies accelerated particles can reach, (2) a power law shape for a few decades in momentum, and (3) a break at high momentum due to losses (either downstream free escape boundary or radiative processes). Another effect visible in all three subfigures is that the spectra at no point reach the canonical value of \( f(p) \propto p^{-4.23} \). All spectra where the power law shape extends over an appreciable momentum range show a softer spectral index. The largest difference in the three sets is the minimum and maximum momenta the spectra attain. \( A/Z \) enhancement (discussed in Chapter 5) is visible between panels (a) and (b) as an increase at both ends of the distribution function; helium nuclei have more energy when they first cross the shock, and can reach higher energies before escaping the shock structure. Electrons are limited by their charge-to-mass ratio, which affects both their mean free paths and their radiative losses; this causes the slightly lower minimum momentum in panel (c). Each of the three sets of spectra shows largely identical downstream spectra, differing only due to the increasing volumes of grid zones as one moves away from the shock.

For the upstream spectra, differences are rather more apparent. The spectra taken nearest the shock (drawn in red) are very similar to the closest downstream spectra. By this point in the shock
Figure 7.2 Particle distributions as functions of shock-frame total momentum, multiplied by $p^{2.23}$. Only the accelerated particles are shown; thermal particles are not included. In all three subfigures, the colors or dash patterns correspond to the locations shown in Figure 7.1. (a) The distribution of accelerated protons. (b) Same, for accelerated helium nuclei. (c) Same, for accelerated electrons.
7.2. PARTICLE SPECTRA  CHAPTER 7. COMPLETE PICTURE OF A SMOOTHED SHOCK

profile (c.f. Figure 7.1) almost all of the compression has occurred; combined with the lack of a strong subshock, there is little that can affect accelerated particles over such short distances, even for the lowest-energy and lowest-mass cosmic rays. At the next closest position (indicated by orange lines), the electrons show a break in behavior compared to the two heavier species. It is more difficult for the lowest-energy electrons to scatter even this far upstream, and their numbers are depressed. As the positions move farther and farther away from the shock, the lower-energy particles of all three species become absent from the spectra. Only the most energetic particles in the populations can make it even as far as the position where the violet curves were taken; this is still well within (by almost an order of magnitude) the upstream free escape boundary, and no particles escape the shock in that direction. An important point to make is that, even for a relativistic shock with no “escape” (by the arbitrary definition of a free escape boundary), energetic particles can diffuse far upstream. This may, however, be an artifact of the scattering prescription I use; the assumption of pitch-angle diffusion (Section 3.4.1) does not depend on shock speed, and takes for granted the existence of sufficient magnetic field turbulence to cause scattering in particles’ pitch angles. Ultimately particle-in-cell simulations will have to settle this question, as it is beyond the scope of the Monte Carlo approach.

7.2.2 Plasma & ISM Frame

Although the Monte Carlo code tracks fluxes in the shock frame, the photon production processes require particle energies in a different frame; I provide here the Lorentz-transformed particle spectra for reference during Section 7.3, on the photon spectra associated with this shock.

The effect of converting the momenta from the shock frame to the plasma frame is most noticeable for the upstream spectra, particularly those furthest upstream (and thus well outside the deceleration caused by cosmic ray backpressure). Even in a relativistic shock, the Lorentz factor relating the downstream shock and plasma frames is around 1.06, so there is little difference in the spectra taken downstream of the shock. Just upstream the velocity difference is still small enough that the red curves are only minimally different from their shock frame counterparts. Further upstream, however, the relativistic nature of the shock becomes apparent: both the minimum and the maximum momenta of each spectrum increase due to a Lorentz boost by a factor $\sim \gamma_0$.

In Figure 7.4 I compare the plasma-frame spectra of all three particle species for two locations in the shock structure. Unlike in Figures 7.2 and 7.3, I have included the thermal particles for the upstream region. The curves highlight the vast difference in momentum and diffusion lengths between the lowest- and highest-energy particles, as well as the range of scales on which magnetic field turbulence must exist. In this presentation it is especially clear that diffusion length of the
Figure 7.3 Same as Figure 7.2, but for plasma-frame total momentum.
7.2. PARTICLE SPECTRA

CHAPTER 7. COMPLETE PICTURE OF A SMOOTHED SHOCK

Figure 7.4 A comparison of the particle spectra for all three species, for a downstream zone and an upstream zone. The downstream zone is taken from the dot-dashed curves in Figure 7.3, while the upstream curves are from the blue curves in that figure. The upstream (non-CR) thermal population is included in the bottom panel for comparison.

Particles depend only on energy: despite the orders-of-magnitude difference in mass, far upstream from the shock all three species show essentially the same energy range. The upstream helium nuclei show the same enhancement at the highest energies that they do downstream, while the protons and electrons differ only in normalization (as, again, is the case downstream).

The only particle species whose ISM-frame spectrum is important to photon production is the electrons (though the others may be significant in the context of contributions to the galactic or intergalactic cosmic ray population; see Section 1.1.2). These ISM-frame spectra are given in Figure 7.5. The differences between the ISM-frame spectra and those from the shock frame occur at the opposite end of the shock structure from those in the plasma frame. The far upstream plasma frame is the ISM frame, so those spectra are unsurprisingly identical. Downstream, however, the relative Lorentz factor between shock and ISM frames is a significant fraction of $\gamma_0$. All particles in those regions have crossed the shock at least once and are relativistic in the plasma frame. Additionally, the lowest-energy particles (still with thermal energies of the order of $\gamma_0 mc^2$) isotropize very quickly downstream, so that particles are almost equally likely to have momenta directed
7.3 Photon Spectra

In this section I will show the photon spectra produced by the particle distributions described in the previous section. I begin with the spectra in the frame of production, which are displayed in Figure 7.6. This means the plasma frame for pion decay and synchrotron mechanisms, and the ISM frame for the inverse Compton process. The spectra are colored or formatted as in Figures 7.2-7.5, and are given as energy fluxes. The plot shows fluxes at some distance from the shock, ignoring (in the cases of synchrotron and pion decay) that production of photons happens in a different frame from observation. I am therefore disregarding special relativistic effects such as beaming—which will vary throughout the shock structure due to the varying speed in the shock frame—in presenting these spectra as fluxes. I have used these units rather than, e.g. photon production per second, to make comparison easier against the plots that follow. The one commonality in all three subfigures of Figure 7.6 is that emission is dominated by the downstream region for all radiative processes, except for the very highest-energy region of the pion-decay emission. This is due to both the higher
Figure 7.6 Photon spectra, calculated in the frame of production (plasma for pion decay and synchrotron, ISM for inverse Compton) at several locations within the shock structure. The color and style of the curves refers to the same locations as in Figures 7.2-7.5. The area under the curves is the total energy flux, not the photon count.
density and number of particles and the very high energy even the thermal particles have after encountering a shock with $\gamma_0 = 8.2$; the larger upstream volumes mitigate these effects slightly, but come nowhere close to completely counteracting them. In the pion decay spectra above $10^6$ MeV, the fluxes drop very steeply with energy as downstream photons give way to upstream photons produced by the highest-energy particles in the shock structure. By contrast, the highest-energy synchrotron photons are produced downstream of the shock due to compression of the magnetic field and the turbulence within it. The inverse Compton mechanism shows two stages of turnoff. At around $10^7$ MeV the spectrum begins to fall off due to the turnover of the electron distribution (see Figure 7.5 above). There is also a sharp drop just above $10^8$ MeV caused by the Klein-Nishina limit, even for CMB photons whose energies are $\sim 10^{-11} m_e c^2$. Both the pion decay and inverse Compton spectra downstream of the shock show bumps due to the thermal population of particles, visible at $\sim 300$ MeV and $\sim 300$ keV respectively. This point bears repeating: for the shock modeled here, there is detectable structure in the inverse Compton emission. If it is not overwhelmed by the pion decay emission in the same energy range (not a given, as we will see shortly), this structure might be observed by gamma-ray telescopes during the afterglow phase of a gamma-ray burst. A final note on the relative strengths of the emission processes: even before transformation into the ISM frame, both the pion decay and synchrotron spectra are an order of magnitude—or greater—higher than the inverse Compton spectra due to CMB photons. This does, of course, depend on the parameters used for this shock. By decreasing the upstream density, by altering the magnetic field strength, or by adjusting the one-time energy transfer from ions to electrons, inverse Compton emission could be induced to dominate the emission due to pion decay.

As outlined in Chapter 6, the next step in calculating the observed photon spectra is conversion of all photons into the ISM frame, and collection into emission regions. The results of this step are provided in Figure 7.7. (Note the change in vertical axis compared to Figure 7.6.) Both pion decay and synchrotron emission are significantly enhanced by the transformation from the plasma to ISM frame. Most of the increase in flux (about four orders of magnitude between the peaks in the two figures) is due to the summation of grid zones into emission regions. The remainder is due to the Lorentz transformation between frames.

Finally, I show in Figure 7.8 the total flux observed—summed over all three processes—for each emission process, as well as summing over all the regions into a single overall spectrum. The farthest-downstream emission region contributes the overwhelming majority of the energy flux, and is thus overlaid by the curve for the totalled spectrum. As may be inferred by comparing the plots in Figure 7.7, the observed photon spectrum is dominated by the synchrotron and pion decay processes. There is a broad peak of emission in the ultraviolet/X-ray band, around 1 keV, from synchrotron. At much higher energies, pion decay produces another, narrower, peak in the GeV
Figure 7.7 Photon spectra, transformed into ISM frame and summed into emission regions. Colored curves come from upstream regions, while black curves represent downstream regions. Red is the region closest to the shock, while blue is the farthest upstream. Relativistic beaming, Doppler shift, and power dilation have been taken into account as needed. The fluxes shown are energy fluxes, so the area under the curves is the total energy flux, not the photon count.
7.4 Summary

In this chapter I have explored all facets of a single relativistic shock that are relevant to this work: the shock profile, the distributions of accelerated particles, and the observed photon spectra. I have not discussed particle escape here; for relativistic shocks escape is extremely unlikely, as I will show...
in the next chapter.\textsuperscript{1} The modification of the shock structure due to efficient particle acceleration is apparent from both the velocity profile and the energy & momentum fluxes. The structure present in the shock profile generates a rich variety of particle spectra in all three species—protons, electrons, and helium nuclei—that shows significant variation between species and with position in the shock. All three emission processes—pion decay, synchrotron, and inverse Compton—are produced at roughly the same rate in the frame of production, but the Lorentz transformation into the ISM (observer) frame substantially enhances the pion decay and synchrotron mechanisms compared against inverse Compton. As seen by a distant observer, the multiband photon spectrum would show a two-peaked structure, with the lower ($\sim$ keV) peak due to synchrotron, and the higher ($\sim$ GeV) peak generated by pion decay. However, I again stress that these results are dependent on the particular parameters chosen for the shock (Table 7.1), and that different parameters might substantially change the results presented here. This will be emphasized in the following chapter.

\textsuperscript{1}But see, e.g., Ellison and Bykov (2011) for a discussion of particle escape from non-relativistic shocks.
In this chapter, finally, I present a simulated afterglow from early in the event until well into the trans-relativistic regime. I have not made any particular attempt to fit the data to actual afterglows. Initial conditions were chosen based on typical parameters for GRBs, but there is wide variability. As well, the microphysics of the magnetic field and scattering can have significant effects on the particle distributions; I have left these unchanged from earlier chapters, in full awareness that they will need reexamination as part of any dedicated effort to match observed afterglows to simulated shocks and spectra. My focus here is on demonstrating the infrastructure developed for simulating and processing shock acceleration into observables. Future work will involve selecting a hydrodynamical model, determining relevant DSA parameters, and calculating emission using the methods outlined below. The calculated emission can then be compared against individual GRBs to test what regions of the parameter space are consistent with observations.

8.1 The Blandford–McKee Solution

At the heart of the evolution of the shock is the Blandford and McKee (1976) solution. I briefly mentioned it in Section 2.2.1, but in this section I will elaborate on the details and how it was implemented in the simulation. Where the equations presented here differ from Blandford and
8.1. THE BLANDFORD–MCKEE SOLUTION

McKee (1976), it is because I have also relied on the discussion in Kobayashi et al. (1999).

In the Blandford–McKee solution, $\gamma_0$ is the Lorentz factor of the shock itself, as measured in the ISM frame (it is also the Lorentz factor of the far upstream plasma as measured in the shock frame, which explains the choice of subscript). It is set by considering the energy deposited in swept-up material so far:

$$\gamma_0(t) = \left( \frac{17}{8\pi} \right)^{1/2} \left( \frac{l_{\text{Sed}}}{R(t)} \right)^{3/2}$$

(8.1)

where $t$ is the time measured in the engine (or ISM) frame, and $l_{\text{Sed}}$ is the Sedov length, $(E_{\text{iso}}/\rho_0 c^2)^{1/3}$ (note that $l_{\text{Sed}}$ is defined in terms of the isotropic energy release $E_{\text{iso}}$, even though it is well-known that GRBs release their energy into only a small fraction of a sphere; the true energy of a GRB is thus lower than $E_{\text{iso}}$). The shock speed depends also on the current radius $R$, defined as

$$R(t) = c t \left( 1 - \frac{1}{8\gamma_0^2} \right).$$

(8.2)

Note that $\gamma_0$ is defined in terms of $R$, which is defined in terms of $\gamma_0$. One could solve Equations 8.1 and 8.2 for whichever of $\gamma_0$ or $R$ one wanted. Or one could step forward small increments in one of the three variables $t$, $R$, or $\gamma_0$ and update the other two variables accordingly. For the work presented here, I step forward in radius, then find the time and shock speed associated with the new distance.

It is well-established that purely compressive hydrodynamics fails to reproduce the age/radii of observed supernova remnants (Warren et al., 2005; Warren and Blondin, 2013); in these cases the fluid must be more compressible than predicted by the adiabatic index, since particle escape ahead of the forward shock removes energy from the plasma (see also Blondin and Ellison, 2001). However, the remnants studied in the cited works all involve nonrelativistic shocks, in contrast to the trans-relativistic (or faster) shocks modeled in this work. As shown in Table 8.2 below, particle escape is negligible (if not identically zero) for all times incorporated in the model, meaning that the nonlinear nature of DSA (which includes particle escape) will not affect the validity of Equations 8.1 and 8.2.

The Blandford–McKee solution is a self-similar solution whose downstream parameters all vary as some power of the function $\chi(r, t)$. $\chi = 1$ corresponds to the location of the shock at all times, increasing as one moves further downstream, and is given by

$$\chi(r, t) = 1 + 8\gamma_0^2 \left( 1 - \frac{r}{R(t)} \right).$$

(8.3)

With $\chi$ defined as above, the hydrodynamical properties of the shocked plasma (pressure, ISM-frame
Lorentz factor, and ISM-frame density, respectively) are

\[ p(r, t) = \frac{2}{3} w_0 \gamma_0^2 \chi^{-17/12}, \quad (8.4) \]

\[ \gamma^2(r, t) = \frac{1}{2} \gamma_0^2 \chi^{-1}, \quad (8.5) \]

and

\[ n'(r, t) = 2n_0 \gamma_0^2 \chi^{-7/4}. \quad (8.6) \]

In Equation 8.4 \( w_0 = \rho_0 \) is the enthalpy of the upstream plasma, which is assumed to be cold. The downstream equation of state depends on the adiabatic index \( \Gamma \) of the plasma. The Blandford–McKee solution assumes that value to be \( \Gamma = 4/3 \) everywhere, appropriate for material heated by a relativistic shock. The density \( n_0 \) in Equation 8.6 is measured in the ISM frame (also the far upstream plasma frame); however, the downstream density \( n' \) is also measured in the ISM frame rather than the rest frame of the plasma. To get density in the plasma frame, one divides by \( \gamma(r, t) \) (i.e. Equation 8.5) to yield

\[ n(r, t) = \sqrt{6} n_0 \gamma_0 \chi^{-5/4}. \quad (8.7) \]

The Blandford–McKee solution is very accurate throughout the relativistic phase of the afterglow. As the shock slows the solution diverges from numerical results, and by \( \gamma_0 = 2 \) the analytical solution is no longer a good approximation (Kobayashi et al., 1999). An explanation follows: I have mentioned that \( \chi \) measures the distance downstream at a particular time. For a shock that sweeps up a large quantity of material, \( \chi \) might have to be quite large before it encompasses all of the shocked plasma. This situation will tend to happen late in the afterglow, when the shock’s Lorentz factor is small. The tension between these two manifests itself, for example, in Equation 8.5, which may prescribe a Lorentz factor smaller than unity. This is an obviously unphysical result, but merely highlights that numerical work is needed to fully understand this phase of the afterglow.

The time-dependence of the Blandford–McKee solution is somewhat at odds with the steady-state assumption made by the Monte Carlo code. This is important early in the lifetime of the shock, when the Lorentz factor is decaying rapidly, but less important later. However, only PIC simulations are capable of fully time-dependent treatment of relativistic shocks, and at present (e.g., Sironi and Spitkovsky, 2011; Sironi et al., 2013) they are extremely limited in the time scales modeled.\(^1\) It is therefore difficult to quantify the error introduced by assuming steady-state shock at each point in time. Nonetheless, the Monte Carlo code is the most self-consistent approach currently available.

\(^1\)The physical time represented by such PIC simulations is on the order of seconds to minutes, depending on the reference frame. This is far too short to model the extended behavior of the forward shock of the GRB jet.
Table 8.1 Afterglow model, physical parameters

<table>
<thead>
<tr>
<th>Time step&lt;sup&gt;a&lt;/sup&gt;</th>
<th>( t_{\text{eng}} )&lt;sup&gt;b&lt;/sup&gt; (sec)</th>
<th>( t_{\text{obs}} )&lt;sup&gt;c&lt;/sup&gt; (sec)</th>
<th>( R_{\text{FS}} ) (cm)</th>
<th>( \gamma_0 )</th>
<th>Color in figures&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( 1.0 \times 10^7 )</td>
<td>( 7.8 \times 10^2 )</td>
<td>( 3.0 \times 10^{17} )</td>
<td>40.1</td>
<td>Magenta</td>
</tr>
<tr>
<td>2</td>
<td>( 1.4 \times 10^7 )</td>
<td>( 2.6 \times 10^3 )</td>
<td>( 4.1 \times 10^{17} )</td>
<td>25.6</td>
<td>Medium Slate Blue</td>
</tr>
<tr>
<td>3</td>
<td>( 1.8 \times 10^7 )</td>
<td>( 8.6 \times 10^3 )</td>
<td>( 5.5 \times 10^{17} )</td>
<td>16.4</td>
<td>Blue</td>
</tr>
<tr>
<td>4</td>
<td>( 2.5 \times 10^7 )</td>
<td>( 2.8 \times 10^4 )</td>
<td>( 7.4 \times 10^{17} )</td>
<td>10.5</td>
<td>Cyan</td>
</tr>
<tr>
<td>5</td>
<td>( 3.4 \times 10^7 )</td>
<td>( 9.4 \times 10^4 )</td>
<td>( 1.0 \times 10^{18} )</td>
<td>6.7</td>
<td>Green</td>
</tr>
<tr>
<td>6</td>
<td>( 4.5 \times 10^7 )</td>
<td>( 3.1 \times 10^5 )</td>
<td>( 1.4 \times 10^{18} )</td>
<td>4.3</td>
<td>Yellow</td>
</tr>
<tr>
<td>7</td>
<td>( 6.2 \times 10^7 )</td>
<td>( 1.0 \times 10^6 )</td>
<td>( 1.8 \times 10^{18} )</td>
<td>2.7</td>
<td>Orange</td>
</tr>
<tr>
<td>8</td>
<td>( 8.5 \times 10^7 )</td>
<td>( 3.5 \times 10^6 )</td>
<td>( 2.5 \times 10^{18} )</td>
<td>1.7</td>
<td>Red</td>
</tr>
</tbody>
</table>

<sup>a</sup> For figures that compare data from different time steps against each other, each time step has a color that will be consistently used to represent it.

<sup>b</sup> Elapsed time in the rest frame of the central engine; also the ISM frame.

<sup>c</sup> Elapsed time since the start of the GRB for an observer on Earth.

The particular physical parameters for the afterglow presented here are as follows \( E_{\text{iso}} = 10^{53} \) erg, \( n_0 = 1 \) proton per cm\(^3\), with 10% helium and electrons sufficient for charge balance, and a mean upstream magnetic field of 3 \( \mu \)G. These parameters are fixed through the entire simulation, indicating a constant-density circumburst medium (CBM) rather than a wind-like one (see the discussion in Section 2.4). The initial radius at which the Blandford–McKee solution starts to follow the shock depends on the maximal Lorentz factor, which was roughly 330. As discussed in Chapter 5, relativistic shocks expanding into low-magnetization plasma are extremely agnostic about the orientation of the (weak) mean magnetic field. All of the shocks presented in this chapter are sweeping up CBM material with the aforementioned average density and mean magnetic field. This is results in a magnetization parameter \( \sigma \sim 10^{-9} \), far below the threshold of \( 10^{-5} \) described in Sironi et al. (2013). So while ultra-relativistic shocks are certain to be highly oblique to the mean CBM field for all but a small fraction of their surface area (if even that), the dominance of self-generated turbulence over the bulk magnetic field allows me to treat all of these shocks in the parallel orientation without loss of generality.

8.2 Hydrodynamics and Particle Spectra

The particular physical parameters for the afterglow presented here are as follows \( E_{\text{iso}} = 10^{53} \) erg, \( n_0 = 1 \) proton per cm\(^3\), with 10% helium and electrons sufficient for charge balance, and a mean upstream magnetic field of 3 \( \mu \)G. These parameters are fixed through the entire simulation, indicating a constant-density circumburst medium (CBM) rather than a wind-like one (see the discussion in Section 2.4). The initial radius at which the Blandford–McKee solution starts to follow the shock depends on the maximal Lorentz factor, which was roughly 330. As discussed in Chapter 5, relativistic shocks expanding into low-magnetization plasma are extremely agnostic about the orientation of the (weak) mean magnetic field. All of the shocks presented in this chapter are sweeping up CBM material with the aforementioned average density and mean magnetic field. This is results in a magnetization parameter \( \sigma \sim 10^{-9} \), far below the threshold of \( 10^{-5} \) described in Sironi et al. (2013). So while ultra-relativistic shocks are certain to be highly oblique to the mean CBM field for all but a small fraction of their surface area (if even that), the dominance of self-generated turbulence over the bulk magnetic field allows me to treat all of these shocks in the parallel orientation without loss of generality.
Table 8.2 Afterglow model, Monte Carlo parameters

<table>
<thead>
<tr>
<th>Time step</th>
<th>$t_{\text{acc}, \text{max}}$</th>
<th>$\gamma_0$</th>
<th>$n_H$</th>
<th>$n_{\text{NL}}$</th>
<th>$x_{\text{sub}}$</th>
<th>$L_{\text{FEB}}$</th>
<th>$f_{\text{trans}}$</th>
<th>$q_{\text{En}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.0 \times 10^7$</td>
<td>40.1</td>
<td>7000</td>
<td>3.002</td>
<td>6.0</td>
<td>$3 \times 10^3$</td>
<td>3.1 $\times 10^{15}$</td>
<td>0.20</td>
</tr>
<tr>
<td>2</td>
<td>$1.4 \times 10^7$</td>
<td>25.6</td>
<td>4000</td>
<td>3.003</td>
<td>5.0</td>
<td>$4 \times 10^3$</td>
<td>4.2 $\times 10^{15}$</td>
<td>0.20</td>
</tr>
<tr>
<td>3</td>
<td>$1.8 \times 10^7$</td>
<td>16.4</td>
<td>3000</td>
<td>3.006</td>
<td>4.0</td>
<td>$6 \times 10^3$</td>
<td>6.2 $\times 10^{15}$</td>
<td>0.20</td>
</tr>
<tr>
<td>4</td>
<td>$2.5 \times 10^7$</td>
<td>10.5</td>
<td>2000</td>
<td>3.014</td>
<td>3.0</td>
<td>$1 \times 10^4$</td>
<td>1.0 $\times 10^{16}$</td>
<td>0.20</td>
</tr>
<tr>
<td>5</td>
<td>$3.4 \times 10^7$</td>
<td>6.7</td>
<td>1300</td>
<td>3.033</td>
<td>2.0</td>
<td>$2 \times 10^4$</td>
<td>2.0 $\times 10^{16}$</td>
<td>0.20</td>
</tr>
<tr>
<td>6</td>
<td>$4.5 \times 10^7$</td>
<td>4.3</td>
<td>800</td>
<td>3.08</td>
<td>1.0</td>
<td>$3 \times 10^4$</td>
<td>3.0 $\times 10^{16}$</td>
<td>0.20</td>
</tr>
<tr>
<td>7</td>
<td>$6.2 \times 10^7$</td>
<td>2.7</td>
<td>500</td>
<td>3.18</td>
<td>0.5</td>
<td>$6 \times 10^4$</td>
<td>5.8 $\times 10^{16}$</td>
<td>0.20</td>
</tr>
<tr>
<td>8</td>
<td>$8.5 \times 10^7$</td>
<td>1.7</td>
<td>300</td>
<td>3.41</td>
<td>4.0</td>
<td>$1.2 \times 10^5$</td>
<td>1.0 $\times 10^{17}$</td>
<td>0.07</td>
</tr>
</tbody>
</table>
at any particular $t_{\text{eng}}$ in the afterglow is

$$t_{\text{obs}} = t_{\text{eng}} - \frac{R(t_{\text{eng}})}{c},$$

(8.8)

which means that my simulated afterglow starts approximately 780 seconds after the light from the GRB first reached Earth. The final time step corresponds to an observer time of $3.5 \times 10^6$ seconds post-GRB. This is longer than most bursts are observable in the X-ray band, but there are many GRBs with X-ray detections this late (see the data presented in Evans et al., 2009; GRB 060729’s X-ray afterglow was detectable at an age of more than $5 \times 10^7$ seconds, per Grupe et al., 2010).

Eight time steps are used in simulating the afterglow. The relevant parameters (and some results) are given in Tables 8.1 and 8.2. Though I follow the shock for only a factor of $\sim 8$ in time and distance, the range of shock Lorentz factors extends from the ultra-relativistic to the trans-relativistic. The simulated afterglow is followed until $\gamma_0 = 1.7$, which occurs at $t_{\text{eng}} = 8.5 \times 10^7$ s or $t_{\text{obs}} = 3.5 \times 10^6$ s. In the trans-relativistic regime, which lies between the fully relativistic phase and the non-relativistic Sedov-Taylor phase, the Monte Carlo code excels (as was shown in Chapter 4). The Blandford–McKee solution is the source of the error in this regime, as pointed out above. This is sufficient for the proof-of-concept being presented in this dissertation, but numerical solutions will be necessary moving forward.

The downstream region ends with a probability of return plane (PRP) rather than a sharp free escape boundary (FEB). I discussed the implications of the two possibilities in Section 3.5. In Chapter 4 PRPs were used exclusively for the downstream region, while in Chapters 5 and 7 I showed results using a FEB. The motivation for returning to PRPs here is that the downstream region of each time step must include all particles that were swept up and shocked between the previous time and the current one. Since Equation 8.7 describes how the comoving density varies downstream from the shock, it is possible to calculate the depth of the shock-frame volume necessary to contain the (known) number of shocked particles. This volume is several hundreds of code units downstream from the shock, even at the earliest time step when density is highest. In light of this, I elect to treat the downstream region as effectively infinite in extent. The PRP is, however, placed far enough downstream that the bulk of cosmic rays are able to isotropize prior to reaching it, ensuring that Equation 3.16 holds.

I show in Figure 8.1 the shock profiles at each of the eight time steps listed in Tables 8.1 and 8.2. The colors here, and anywhere else time steps are compared, consistently refer to the same time step. The top two rows of Figure 8.1 show the flow speed in the shock frame, as both $\beta$ and $\gamma\beta$, normalized against $\beta_0$ or $\gamma_0\beta_0$. The profiles clearly show the trend of decreasing artificial smoothing, since the edge of the enforced precursor approaches the shock as time passes and $x_{\text{sub}}$ drops. Only
Figure 8.1 Comparison of the shock profiles at the discrete time steps modeled during the afterglow. In all panels magenta curves represent the earliest time ($\gamma_0 = 40.1$), while the red curves show the latest time ($\gamma_0 = 1.7$). All four rows show their respective quantities normalized to the far upstream values, and measured in the shock frame. Top row: the bulk fluid velocity; note that for all but the latest time step, $r_{\text{tot}} = u_0/u_2 \approx 3$, while at the last time step $r_{\text{tot}} = 4$. Second row: $\gamma \beta$ is plotted (inversely proportional to plasma-frame density). Third row: momentum flux. Bottom row: energy flux. On the left hand side of the second row, the vertical lines show the location of the upstream free escape boundary at each time step. In the third row, the lines give the maximum upstream distance at which particles were detected for that time step. Note that only the final two times see any particles escape the shock; see text for additional explanation.
time step 8 has an overall compression ratio that deviates from the Rankine-Hugoniot value. Since this measures only the velocity ratio $\beta_0/\beta_2$ (and not the density ratio $\gamma_0\beta_0/\gamma_2\beta_2$), one would expect to see this change clearly in the collection of velocity profiles, but perhaps less clearly in the $\gamma\beta$ profiles. This is indeed the case. The red $\beta_x$ curve of Figure 8.1 is markedly lower than any of the other seven downstream from the subshock, but the $\gamma\beta_x$ values show a much smoother trend.

The lower two rows of Figure 8.1 show that, to within acceptable limits, these nonlinear shocks conserve momentum and energy flux everywhere within the shock structure. The small spike at $x = 0$ is a numerical artifact of energy transfer between ions and electrons, and vanishes rapidly once the electrons have a chance to isotropize downstream from the subshock.

Comparing the two sets of vertical colored lines in Figure 8.1, and considering the final column in Table 8.2, shows that particle escape simply does not occur in the relativistic regime. Only at the final two time steps are particles able to diffuse all the way to the upstream FEB, and even then only in very small numbers: despite representing the most energetic particles in the system, escaping cosmic rays carry away just $3 \times 10^{-5}$ percent of the incoming energy flux in time step 7, and $5 \times 10^{-2}$ percent in time step 8. This demonstrates that acceleration time is the limiting factor on particle energy for the shocks modeled here. As the acceleration time increases, so does the upper limit on particle energy; given unlimited time to participate in the Fermi-I process, particles will gain energy until the FEB (and particle escape) enforces the maximum.

Over the rest of this section, I will refer to “cohorts” of particles. These are the populations of upstream particles that were swept up by the shock at a particular time step. In the absence of diffusion between emission shells (see Section 6.4, and Figure 8.2 below, for illustrations of these shells), each cohort of particles will experience the same conditions throughout the afterglow, from the initial shock speed at crossing to the cooling behavior and decaying magnetic field strength as the plasma convects downstream from the shock.

In order to calculate the photon production (Section 8.3), it is necessary to know both how many particles are present in a given emission shell (Figure 6.7) and what their plasma-frame density is. The results of the process are illustrated in Figure 8.2, and are explained here. Determining the total number of particles is simple. We know the value of $R_{FS}$ in the ISM frame at every time step, as given in Table 8.1. Assuming, as we do, that the jet maintains a conical structure with constant opening angle, it is straightforward to determine the ISM-frame volume swept out during time any given time step. Multiplication by the upstream ISM-frame density yields the total number of particles in each emission shell.

To determine the average density of each cohort, I numerically integrate the total number of particles downstream from the shock. This is achieved by starting at $R_{FS}$ in the ISM frame and counting inward (i.e. towards the central engine), dividing the downstream region into very small
8.2. HYDRODYNAMICS AND PARTICLE SPECTRA  

CHAPTER 8. SIMULATED AFTERGLOW

Figure 8.2 Illustration of how downstream cohorts are handled at a given time step $i$. Four different cohorts are shown, each with the specified number of particles ($N_j, N_{j-1}$, etc.). Along the horizontal axis are the current location of the forward shock $R_{FS}$ (given in Table 8.1), as well as the dividing radii—in the ISM frame—between different cohorts of particles at the current time step. The blue curve represents the values of Equation 8.7 downstream from the shock. Along the vertical axis are the average plasma-frame density of each cohort. Note that the density of the most recent cohort is lower than predicted by the Rankine–Hugoniot equations, and by the Blandford–McKee solution.

Each cohort $j$ is handled by calculating its density $n_j$ and radius $R_j$, and then using Equation 8.6 (again in the ISM frame) combined with the ISM-frame volume of that slice until the running total of particles reaches $N_j$, the total number of particles swept up during time step $i$. As this is happening, the ISM-frame volume of each slice is converted into the plasma frame (by way of Equation 8.5), and added to its own running total. Dividing $N_j$ by the tabulated volume for this shell gives the average plasma frame density $n_j$, as illustrated in Figure 8.2. The next most recent cohort of particles, which will have $N_{j-1}$ particles in it, resumes counting from the dividing radius $R_j$ between the two shells. Counting proceeds in exactly the same fashion, but with a different eventual particle population and different range of radii.

In Figure 8.2, the quantity $N_j$ is constant in time—the number of particles swept up by the shock during each time step is fixed once determined. However, the average densities ($n_j, n_{j-1}$, etc.) depend on Equation 8.7, which does depend on time. So, too, do the dividing radii ($R_j, R_{j-1}$, etc.) between the emission shells. As such, they would be more properly referred to as $n_{i,j}$ and $R_{i,j}$, with $i > j$: the density or radius of cohort $j$, at time $i$. For clarity reasons, this second dimension was
The calculated densities of the actual cohorts modeled, and their time dependence, are shown in Figure 8.3. Note that the curves in Figure 8.3 all start below the Blandford–McKee $n_2$ for shocks at that Lorentz factor, as expected given the discussion surrounding Figure 8.2. Note also that the locations of the stars correspond to the Rankine–Hugoniot values with no particle escape (see Section 3.6). This assumption is acceptable over the range of shock speeds considered here\(^2\); as discussed above escaping particles are negligible for these shocks. Should the model be extended to lower trans-relativistic speeds, this assumption will need to be revisited.

In Figure 8.4 I show how the very first cohort of particles—those shocked at time step 1—cool between successive time steps. The top two panels of the figure show protons and helium nuclei, which are only affected by adiabatic losses. The electrons shown in the bottom panel experience both synchrotron and adiabatic cooling. Examination of the high-energy tail of the electron distribution shows that, by the third time step, synchrotron losses are completely subordinate to adiabatic

\(^2\)At the final time step, this assumption does not hold, since nonlinear smoothing requires an increase of the total compression ratio. This is more evidence that the Monte Carlo code must be partnered with a numerical hydrodynamical solution for accurate treatment of the afterglow.
Figure 8.4 Particle distributions for all three species as they cool downstream from the shock. As in Figures 8.1 and 8.3, the colors refer to the current time step. However, only the particles initially shocked during time step 1 are shown. The number of particles in each distribution, which is conserved between time steps, is equal to the area under a $p \, dN/dp$ curve on this logarithmic axis. The reason for the apparent difference in normalization is that $p^2 \, dN/dp$ is plotted, and earlier time steps have higher momenta. In the bottom panel, it is apparent that synchrotron losses become negligible—compared against adiabatic losses—by the third time step.
cooling. It is only in the first two time steps that electrons cool much more rapidly than ions at equal momenta, and even that applies only to the most extreme upper end of the electron distribution. This particular result depends on the strength of the magnetic field. As will be demonstrated later, the magnetic field strength used in these simulations must be increased (perhaps dramatically) in order to properly fit observations. This would, in turn, result in stronger cooling of electrons at lower energies and at later times than is visible in Figure 8.4.

As discussed in Chapter 1, PIC simulations of relativistic low-magnetization shocks expect to see \( \gtrsim 40\% \) of the upstream bulk energy transferred from ions to electrons during the passage through the turbulent precursor to the shock.\(^3\) This value should be taken with some caution, though, as it may be dependent on the specific assumptions (e.g. electron/ion mass ratio, or shock Lorentz factor). In Chapter 2 I showed that the estimates for the energy fraction \( \varepsilon_e \) in GRB afterglow electrons cluster around \( 10 - 20\% \). For the simulations presented here I have aimed for a much more conservative \( 20\% \) of bulk ram energy in electrons, as can be seen from Table 8.2. However, the energy transfer process used in the Monte Carlo code handles only kinetic energy. If rest-mass energy is included in the calculation of the equipartition fraction, electrons’ share of the energy will increase slightly, but the ions’ will increase a significant amount.

The equipartition fractions of ions (protons and helium combined), electrons, and the magnetic field are plotted in Figure 8.5. As expected, using total energy (rather than simply kinetic) substantially enhances the contributions of the two ion species. If just kinetic energy were considered all eight electron curves would start at \( \varepsilon_e \approx 0.20 - 0.25 \), in keeping with the energy transfer parameter \( f_{\text{trans}} \) (the variation is due to the presence of cosmic ray electrons, which—as will be seen in Figure 8.6—interact very differently with shocks of differing speeds). Once rest mass is included, \( \varepsilon_e \) drops to as low as \( 7\% \) for time step 8. In the top two panels of Figure 8.5, there is an increase in \( \varepsilon_i \), and corresponding decrease in \( \varepsilon_e \), of cohorts at late times due to the manner in which energy transfer is handled. As discussed in Section 5.3.3, the amount of energy transfer is scaled to the shock-frame kinetic energy of the cold upstream plasma. As the shock speed decreases, so does the kinetic energy available for donating. The rest-mass is unchanged, but makes up a larger overall fraction of the total energy in the system, which further increases the faction of energy carried by ions.

Through the first six time steps the value of \( \varepsilon_e,\text{cr} \) drops with each new time step.\(^4\) This is caused by a smaller amount of kinetic energy available for donation from ions, and the general difficulty

---

\(^3\)In Figure 11 of Sironi et al. (2013) the sum of electron and ion equipartition fractions does not add to 1. This means a significant fraction of the downstream energy density is in the magnetic field; examination of Figure 10 of that paper shows small pockets of plasma where \( \varepsilon_B \) appears to approach unity.

\(^4\)Though the ratio \( \varepsilon_{e,\text{cr}}/\varepsilon_{e,\text{tot}} \) does increase slightly as a function of time.
Figure 8.5 The fraction of energy density in various forms, as a function of both current time and time of original shock crossing. Each colored curve tracks the cohort from a particular time step; colors are consistent with their use in Figure 8.1. The numbers shown here include rest-mass energy density, which enhances the contribution of ions and diminishes that of electrons. **Top panel:** energy density in the form of ions, either protons or helium nuclei. **Middle panel:** energy density in the form of electrons. The top set of curves includes both thermal and CR electrons; the bottom set tracks just cosmic-ray electrons. The total and CR-only points for time step 8 lie at \( \varepsilon_e \approx 0.07 \) and \( \varepsilon_e \approx 0.06 \), respectively; the lower point is slightly obscured by of the magenta curve. **Bottom panel:** the logarithm of the energy density in the magnetic field.
electrons have entering the acceleration process. In the final two time steps the situation reverses itself dramatically. Once the shocks become trans-relativistic, CR electrons begin to make up a larger fraction of the all-particle energy density, not just as a fraction of electron energy. This is visible in Figure 8.5, as the orange curve of time step 7 lies above the yellow curve of time step 6; the single red point of time step 8 is higher still, and denotes that the majority of electron energy when $\gamma_0 = 1.7$ is in the form of cosmic rays.

As expected, the magnetic field energy density just downstream of the shock is essentially constant in time. According to Equation 6.14, $B_2 \approx \gamma_0 B_0$, so $U_{B,2} \propto \gamma_0^2 B_0^2$. But the energy density of particles also varies as $\gamma_0^2$ (one power from shock heating, and the second from compression), meaning the contribution of the magnetic field is roughly constant—and very small. However, the energy density of the magnetic field increases with time as each population is carried farther downstream from the shock. This is due to the manner in which magnetic field strength decays. I assume that the turbulence is frozen into the field, and so the magnetic field energy density ($\propto B^2$) drops as $V^{-1}$. This is slower than the $E/V^{-1}$ rate for adiabatic cooling that particles experience (and slightly slower still for electrons that undergo synchrotron losses as well). It is therefore unsurprising that the fraction of energy density in the magnetic field increases with time. It is, though, an artifact of the assumptions made for handling turbulence. The strength of the turbulent magnetic field might be significantly enhanced above the compressed value (Santana et al., 2014), and analytical studies suggest that this enhanced microturbulence should decay much faster than the straightforward $V^{-1}$ model I have used here (Lemoine, 2013).

The points in Figure 8.5 hint that the acceleration process works differently at different times.5 The curves of Figure 8.6 prove this, showing the spectra, in the plasma frame, of each species of particle immediately after crossing the shock at each time step. The transition between time steps—and the associated increase in acceleration time—is plain, as each time step's distribution extends to higher energies and features a slightly lower thermal peak. For the ions, the transition from fully relativistic to trans-relativistic shocks is marked by a hardening of the spectrum, from $p^{-2.2}$ or so to $p^{-2}$. Once $\gamma_0 = 1.7$, the turnover at the highest CR energies is also limited by both the upstream FEB and the limited acceleration time, explaining the sharper corner for the final time step.

In Figure 8.5, I pointed out that the nature of electron acceleration changes in the trans-relativistic regime, and Figure 8.6 bears this out. Where the normalization at a given energy dropped at every time between steps 1 (magenta curve) and 6 (yellow), it rises—in addition to the increased maximum energy—at time step 7 (orange). At time step 7 I am still transferring 20% of the ions’ kinetic energy to electrons. In time step 8 (red), only 7% of the energy is transferred, but the changes in the electron

---

5 Merely a restatement of the main result of Chapter 4, that shock speed influences accelerated spectra.
8.2. HYDRODYNAMICS AND PARTICLE SPECTRA

CHAPTER 8. SIMULATED AFTERGLOW

Figure 8.6 The plasma-frame spectra of each particle population at the time step when it was originally shocked. Coloring is the same as in Figure 8.1, where time step 1 is in magenta and time step 8 is in red. From top to bottom, the panels refer to protons, helium nuclei, and electrons.
8.3. PHOTON SPECTRA

The photon spectrum have accelerated. The CR portion of the spectrum now contains the bulk of the energy, and hardens above $p^{-2}$ at high energies. As well, the height of the thermal peak relative to the CR portion is noticeably lower at time step 8 than it is at any earlier time. All of these changes will have substantial effects on the photon spectrum, as I will show in the next section.

The maximum energies of the distributions shown in Figure 8.6 are inconsistent with the results presented in Sironi et al. (2013). By considering the saturation distance of the self-generated turbulence (which is related to the diffusion length of the lowest-energy CR ions in the simulation), those authors arrived at a formula describing the maximum energy attainable by electrons in an electron-ion shock:

$$\gamma_{\text{sat,e}} \approx \gamma_0 \cdot \frac{m_i}{m_e} \cdot \frac{2}{\sigma^{1/4}}.$$ (8.9)

For proton/electron shocks with an upstream magnetization $\sigma \approx 10^{-9}$ (characteristic of typical ISM, as discussed in Chapter 5), this results in maximum electron energies of $E_{e,\text{max}} \gtrsim 300\gamma_0$ TeV, far in excess of the energies seen in any of the fully relativistic time steps. This is due to acceleration time, rather than the FEB, acting to limit CR energies. Only in the trans-relativistic regime do particles approach or exceed this putative upper bound. However, the assumption that only the lowest-energy cosmic rays cause the turbulence needs to be reconsidered in light of the harder, $p^{-2}$ spectra seen at the latest times. In that scenario it is not obvious which particles are most responsible for the saturation distance of the turbulence, and the limit of Equation 8.9 does not necessarily hold.

8.3 Photon Spectra

In this section I take the particle distributions I have just presented and show the observed photon flux from them. These spectra are presented with a handful of caveats. First, these spectra are for an observer located directly on the axis of the GRB jet. Second, I hold the opening angle of the jet fixed in time; thus, while lateral expansion does occur, I do not include the accelerated expansion due to pressure that happens post jet-break (see Section 2.2.4). Geometric effects such as these must eventually be taken into account, but the goal here is to build a simple model on which to later expand. Finally, all of the photon spectra in this section are presented without absorption effects. The absorption process is well-understood, and observers can reliably recover information like unabsorbed flux and photon index from the absorbed observations.

In Figure 8.7 I show the total, radio–to–$\gamma$-ray, spectra from all shocked particles at every time step.\footnote{Technically CRs in the precursor would contribute some emission as well, but there are vastly fewer of these particles than there are shocked particles downstream, so their contribution to these plots would be negligible.} Photon production was described in Chapter 6, and all processes are represented in this figure.
with the exception of synchrotron self-absorption. As mentioned in Chapter 6, the calculation of SSA requires an isotropic synchrotron photon and electron distributions. As the photons propagate upstream through the jet towards the observer, they encounter shocked material with an increasing Lorentz factor; necessarily, one distribution must cease to be isotropic, and so the simple description of SSA is incomplete. Preliminary results that include just the outermost cohort suggest that the self-absorption frequency increases with time, and that it lies above the thermal peak visible in Figure 8.7 at all time steps. I note, however, that the ordering of $\nu_\alpha$ and $\nu_{\text{pk}}$ depends on the magnetic field strength. As discussed in Chapter 6, $\nu_\alpha \propto B^{1/5}$, while $\nu_{\text{pk}} \propto B$; as the magnetic field rises, the peak frequency rises much more rapidly than does the self-absorption cutoff. In situations with a much stronger magnetic field one would expect the thermal peak to be detectable.

Tracing the spectrum from time step to time step highlights how different the trans-relativistic regime is from the fully-relativistic regime. The electron efficiencies and spectra are largely similar across the first seven time steps, but in time step 8 the slower shock speed allows for a dramatic shift in how electrons are accelerated to high energies. The harder electron spectrum manifests itself in the noticeably sharper synchrotron and inverse Compton peaks. (The jaggedness of the inverse Compton peak is numerical, rather than physical, in nature.)

I have split the spectrum by process for three time steps in Figure 8.8. For the environmental
Figure 8.8 The ISM-frame photon spectrum, separated into emission process, for three time steps in the afterglow. Synchrotron emission is traced in blue, inverse Compton in green, and pion decay in red; the solid red curve is due to protons, and the dashed red curve is due to helium nuclei. The heavy dashed black line is the total of all three processes. **Top panel:** spectrum at time step 1. **Middle panel:** spectrum at time step 5. **Bottom panel:** spectrum at time step 8.
parameters used in this model afterglow, emission due to electrons dominates. Pion decay photons outnumber inverse Compton photons only in a small energy range, that due to the thermal population of non-CR nuclei, and only early in the afterglow. By late times the decreased downstream densities have all but eliminated pion decay as a source of photons. At optical wavelengths emission should be an unbroken power law for all but the first handful of minutes of the shock: the synchrotron thermal peak is already at sub-millimeter wavelengths by time step 1, at $t_{\text{obs}} = 780$ s. In the radio—assuming it is not washed away by synchrotron self-absorption—the passage of the synchrotron thermal peak would be quite distinct. The height of the peak carries information about the relative number of particles in the thermal (uninjected) vs. CR parts of the electron spectrum. With sufficient granularity, then, the Monte Carlo model could use radio observations (e.g., those in Frail et al., 2000) to infer the fraction of electrons injected into DSA at relativistic shocks.

From Figure 8.8 it is apparent that the peak of the synchrotron spectrum does not drop with time. Although the magnetic field strength decays with decreasing shock speed, the increased maximum electron energy is more than enough to offset the change, and so the synchrotron peak rises with time.

The shift in the synchrotron peak is indeed due to the most recently-shocked population of particles, as illustrated in Figure 8.9. Curiously, emission at both the low and the high extremes of each emission process is due to most recently shocked particles, particularly for synchrotron and inverse Compton. At high energies this is due to the higher energy particles allowed by greater acceleration time, while at low energies the thermal peak of the most recently shocked material dominates the cooled thermal peaks of earlier populations.

### 8.4 Comparison against Observations

In this section I will compare the calculated photon spectra against observations. I start by considering total energy flux in the *Swift*-XRT range, $0.3 - 10$ keV. In Figure 8.10 the thin gray lines in the top half of the image show the light curves of a small sample of *Swift*-observed afterglows.\(^7\) These particular afterglows were selected because they had coverage at radio, optical, and X-ray wavelengths. However, it is apparent from the X-ray flux alone that my simulated afterglow very much fails to reproduce a typical afterglow. In addition to being two orders of magnitude fainter than even the dimmest afterglow of the sample, the afterglow also falls off more rapidly than the $t^{-1.2}$ power law characteristic of the *Swift* sample. Additionally, there is a distinct concavity to the light curve that is not seen in most of the sample population.

\(^7\)GRB afterglow data were pulled from the *Swift*-XRT GRB light curve repository, [http://www.swift.ac.uk/xrt_curves/](http://www.swift.ac.uk/xrt_curves/). See Evans et al. (2007) and Evans et al. (2009) for additional information.
Figure 8.9 The contribution of individual particle populations to the total spectrum at time step 6. In all panels the dashed black curves show the total emission over all six shocked populations, while the colored lines trace emission from individual populations. The color scheme is slightly different from that used in other plots. Red still represents the particles shocked at latest times, and magenta the earliest, but only six curves are shown here rather than eight. **Top panel:** synchrotron emission. **Middle panel:** pion decay from protons only. **Bottom panel:** inverse Compton emission. Note the varying vertical scales on all three panels.
8.4. COMPARISON AGAINST OBSERVATIONS

CHAPTER 8. SIMULATED AFTERGLOW

Figure 8.10 The total flux in the $0.3 - 10$ keV band due to the sample afterglow, as well as from a sample of GRBs with radio, optical and X-ray coverage. The colors correspond to time steps as given in Table 8.1. At the top right, the line traces a $t^{-1.2}$ power law, which is broadly consistent with the afterglow population, but is a poorer fit to the simulated afterglow.

Table 8.3 $0.3 - 10$ keV photon index $\Gamma$ by time step; see text for explanation of final row

<table>
<thead>
<tr>
<th>Time step</th>
<th>$\Gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.75</td>
</tr>
<tr>
<td>2</td>
<td>1.76</td>
</tr>
<tr>
<td>3</td>
<td>1.78</td>
</tr>
<tr>
<td>4</td>
<td>1.79</td>
</tr>
<tr>
<td>5</td>
<td>1.84</td>
</tr>
<tr>
<td>6</td>
<td>1.86</td>
</tr>
<tr>
<td>7</td>
<td>1.70</td>
</tr>
<tr>
<td>8</td>
<td>1.51</td>
</tr>
</tbody>
</table>

$Swift$ $1.90 \pm 0.31$

$Swift - 1$ $1.86 \pm 0.17$
Next I consider the X-ray spectra at individual time steps, which requires calculating the photon index, i.e., the spectral index $\Gamma$ of the photon distribution $dN/dE$. The spectra of Figure 8.7 are plotted given as flux per unit logarithmic energy bin. Converting to photon index, then, requires dividing by $E^2$, or subtracting two from the slope of the spectrum on the log-log plot. In Table 8.3 I have determined what the observed photon index would be for each time step. The behavior of the simulated spectra in the $0.3 - 10$ keV range is a slow softening over the first six time steps, followed by a rapid hardening at the final two. This is due at earlier time steps to the approximately balanced contributions of cohorts to the synchrotron flux, as seen for example in Figure 8.9. In the final two time steps the electron acceleration efficiency increases dramatically, leading to dominance of the latest cohort over the previous ones. Since the two latest electron spectra are also harder than those from previous time steps, the inferred X-ray spectrum hardens appreciably.

In the final rows of Table 8.3 I show the average and standard deviation of reported photon indices of the Swift afterglows used to create Figure 8.10. Many bursts have multiple observations, usually consisting of one early ($t_{\text{obs}} \lesssim 500$ s) and one or more late ($t_{\text{obs}} \gtrsim 5000$ s). I selected only observations that overlapped significantly with my own time steps. The mean of the sample of 29 photon indices is $\Gamma = 1.90$, which is softer than any of my simulated spectra. This result is highly affected by a single measurement: the reported photon index at late times for GRB 060218 is 3.28. If this single point is excluded, the average of the sample drops to $\Gamma = 1.86 \pm 0.17$, which is both (slightly) closer to and more tightly clustered around my simulated spectra.

Not shown in Table 8.3 or Figure 8.10 is that where multiple late-time ($t_{\text{obs}} \gtrsim 10^4$ s) photon indices were reported in the Swift repository, there seems to be a slight trend towards lower, i.e. harder, photon indices that is in keeping with rows 6-8 of Table 8.3.

That the simulations produce photon indices in line with Swift observations suggests that the fundamental process of acceleration is handled correctly. Obviously the normalization of the fluxes is orders of magnitude low, but this is easily correctable by modifying several assumptions I made for the parameters of this shock. First of all, I note that both my magnetic field and CR-electron energy fractions (see Figure 8.5) are much lower than expected from observations or PIC simulations. Both synchrotron emission and $\varepsilon_B$ vary as $B^2$, so an increase of the downstream magnetic field strength by a factor of hundreds would have a pronounced effect on the generated emission in the X-ray range. The same effect would occur, if on a smaller scale, with additional energy transfer from ions to electrons.

Beyond the quantity of energy in the magnetic field and electrons, several additional corrections may be necessary to more accurately reproduce observations. The simplest is adjusting the number of electrons generating synchrotron emission, i.e. adjusting either the density or the elemental...
8.5. SUMMARY

CHAPTER 8. SIMULATED AFTERGLOW

composition of the plasma encountering the shock.\textsuperscript{8} One might also question the scattering prescription used in the Monte Carlo runs. I assumed that particles scatter using pitch-angle-diffusion (the high-$N_g$ limit), and that mean free paths are dictated by Bohm diffusion (i.e. $\eta_{\text{mfp}} \approx 1$ in Equation 3.12). There is debate in the literature over whether this limit is achievable. Simulations of cosmic-ray transport in a Kolmogorov turbulence spectrum (Lemoine and Pelletier, 2003; Lemoine and Revenu, 2006) suggest that $\eta_{\text{mfp}} \gtrsim 10$ for relativistic shocks. On the other hand, considerable turbulence seems to exist at small scales (e.g., the “microturbulence” discussed in Plotnikov et al., 2013 and Lemoine, 2013). In such cases the diffusion of particles may well be Bohm-like, but using the amplified turbulent field rather than the mean field to calculate the Larmor radius (Hussein and Shalchi, 2014). It is possible, then, that the Monte Carlo simulations are overestimating the diffusion lengths of cosmic-ray electrons. This would affect the acceleration time (and so the maximum energy observed in Figure 8.6), which in turn would impact the observed radiation. Further study incorporating these effects is plainly warranted.

8.5 Summary

In this chapter I have presented the results of a series of Monte Carlo simulations that attempt to reproduce the afterglow of a GRB. Using the Blandford–McKee analytical solution for the hydrodynamics of the blast wave, I split the afterglow into eight discrete time steps covering early in the afterglow until well into the trans-relativistic stage. Each of these times was modeled using a nonlinear smoothed shock encountering a typical ISM plasma: density $n_0 = 1 \text{ cm}^{-3}$, helium fraction 10%, and mean magnetic field $B_0 = 3 \mu \text{G}$. By incorporating key results, such as energy transfer, from PIC simulations I have shown that the Fermi-I process is not identical over the lifetime of the afterglow. Some time evolution of key DSA parameters (e.g. $\varepsilon_e$, $\varepsilon_B$) may be necessary to accurately reproduce a typical light curve. However, at any individual time step, diffusive shock acceleration seems to do a very good job of reproducing the X-ray spectra observed in afterglows. The model appears sound, and the next step would be exploring the wide parameter space to find ranges that are capable of matching all facets of GRB observations.

\textsuperscript{8}The Blandford–McKee solution is quite reliable at high Lorentz factors, so the volume swept up by the shock is unlikely to vary significantly—given a constant-density, ISM-like circumburst medium.
Over the first two chapters of this dissertation, I explained the history and current state of affairs of two fields that are quite closely related: gamma-ray bursts (GRBs) and cosmic ray (CR) production via the diffusive shock acceleration (DSA) process. I made a case that studying the former requires awareness of the latter, and further that the study of CR generation must be handled numerically to account for the inherently nonlinear feedback loop between shocks and the particles they accelerate. This is true in the relativistic regime where analytic solutions exist for the unmodified (i.e., ignoring the feedback loop) case, but especially true in the trans-relativistic regime where no solutions exist under any assumptions. However, modeling the trans-relativistic regime is necessary to understanding the late afterglow phase of GRBs, since the external shocks must slow from their initial relativistic speed, through the trans-relativistic regime, and finally dissipate as non-relativistic shocks in the circumburst medium.

Over the following chapters, I described the Monte Carlo code that I used to study DSA in the afterglow phase. Although I inherited a code with a rich history of addressing problems related to DSA at non-relativistic speeds, a great deal of work was required to study nonlinear DSA in the fully relativistic and trans-relativistic domains. The results presented in Chapter 4 show that the Monte Carlo code can handle the nonlinear feedback loop over virtually all speeds relevant in the afterglow phase. Additionally, one can expect the character of the accelerated CR spectrum to vary in a smooth,
predictable manner as the shock evolves from less-efficient, relativistic speeds to ultra-efficient, non-relativistic speeds (see Figure 4.10).

Chapter 4 was concerned with proton-only shocks, however, requiring additional extensions to cover particles that are also typical of the ISM: heavier ions and electrons. In Chapter 5 I reviewed work done to expand the Monte Carlo code to include these additional species. The $A/Z$ enhancement effect seen in non-relativistic shocks, in which particles with higher mass-to-charge ratios are preferentially injected and accelerated, was observed to occur in fully relativistic and trans-relativistic shocks as well (Figures 5.8 and 5.13). Motivated by particle-in-cell simulations (and some heroic analytical work), I added a parameter to the Monte Carlo code that handles energy transfer from ions to electrons in the precursor of a shock. This energy transfer is necessary for electrons to be accelerated to any significant degree, as $A/Z$ enhancement works strongly against these light particles. With even a small amount of energy transfer, however, CR electrons can become nearly as numerous as ions in relativistic shocks. And in trans-relativistic shocks where the $A/Z$ enhancement is less severe, a small amount of energy transfer may allow electrons to become more numerous than protons.

Afterglows are observed through the photons generated by CR electrons and ions, and in Chapter 6 I outlined the three main emission processes considered in this dissertation: synchrotron, inverse Compton, and pion decay. All three mechanisms are well-documented in the literature, and this chapter is more a review than a presentation of new work.

In Chapter 7 I fully explored a single relativistic shock, verifying that all facets of shock smoothing, DSA, and photon production function as intended.

The primary results of the dissertation are presented in Chapter 8, which contains a simulated afterglow from the fully relativistic stage until well into the trans-relativistic regime. Using the Blandford–McKee solution for a relativistic blast wave expanding into a constant-density ambient medium, I described the afterglow as a series of discrete time steps, each of which swept up a different volume of the circumburst medium (CBM) and accelerated it to a time-limited CR spectrum. Downstream from the shock each previous cohort of particles expanded and cooled. I then calculated the photon spectrum from all shocked particles at each time step. Key results include the following:

1. The nature and efficiency of electron DSA changes dramatically in the trans-relativistic regime (Figure 8.6); this effect applies to all particle species, not just protons as considered in Chapter 4. In relativistic shocks, electron acceleration is highly suppressed unless energy is transferred from ions. In the trans-relativistic regime electron acceleration is so efficient that (if the energy transfer parameter $f_{\text{trans}}$ is not adjusted) electrons may carry the bulk of the energy density downstream from the shock.
2. For the parameters used in this simulation, DSA in GRBs is highly unlikely to be the source of cosmic rays above the knee in the spectrum. The age of the shock strongly impacts the peak energy of the particles accelerated during the afterglow. The highest particle energies observed in the simulation are a few PeV, and those occur later in the afterglow (again, Figure 8.6).

3. Though synchrotron self-absorption (SSA) is not explicitly treated in this work, the Monte Carlo code predicts a very clear peak at radio wavelengths that is due to the thermal—i.e., not cosmic-ray—population of electrons downstream from the shock (Figure 8.7). This peak should decay with time, and its height is directly related to the fraction of the electron spectrum cosmic rays comprise. If SSA were included, it is possible that this thermal peak would be absorbed; however, the stronger the downstream magnetic field the more likely it is that the peak will be observable.

4. Without the amplification of the ambient magnetic field by electromagnetic instabilities, the model described in previous chapters does a poor job of matching the luminosity of observed afterglows (Figure 8.10). The luminosity depends on the energy densities of both accelerated electrons and the magnetic field, both of which were lower than expectations based on particle-in-cell simulations. With larger, but more plausible, values, the simulation should match, or even exceed, the luminosity of a typical GRB afterglow.

5. The Monte Carlo code does a very good job of matching the spectrum—if not the normalization—of observations in the X-ray band, between 0.3 keV and 10 keV (Table 8.3). The simulated photon indices during most of the simulation fall within a standard deviation of the mean a Swift-observed sample of GRB afterglows. The model also predicts a hardening of the X-ray spectrum at late times as the shock slows to trans-relativistic speeds, which seems to be supported by multi-epoch Swift-XRT observations.

6. The simulation predicts non-negligible photon production at MeV energies or higher, well into the afterglow. For the parameters used in this simulation, the $\gamma$-ray emission is almost entirely due to inverse Compton scattering by ultra-relativistic electrons (Figure 8.8).

7. The temporal behavior of the simulated afterglow is a good match to observations during the relativistic phase of the simulation (again, Figure 8.10). At the latest times (where the Blandford–McKee solution is highly suspect anyway) the model is quite sensitive to the amount of energy present in cosmic-ray electrons; by adjusting the amount of energy transferred to electrons the model could explain both a hardening or a softening of the light curve.
The simulated afterglow will depend on the parameters fed into the Monte Carlo code. Given the apparent fidelity of the simulations to observations, it seems likely that the Monte Carlo approach outlined in this chapter can be used to infer properties of any given afterglow.

In summary, I believe I have shown that the Monte Carlo approach to studying DSA in relativistic and trans-relativistic shocks possesses a great deal of explanatory power, in spite of parametrizing away important aspects of the microphysics. The model presented here has the ability to self-consistently treat the acceleration of particles by fast shocks, the nonlinear interaction between cosmic rays and the shock structure, and the resultant photon production due to the shocked particle population. This capacity makes it an invaluable tool for understanding GRB afterglows.

I must note that this code, and the model it contains, is far from complete. Of course, the Monte Carlo approach requires compromise when describing wave–particle interactions or the self-generation of magnetic turbulence; while both of these are handled explicitly by particle-in-cell simulations, the tradeoff is computational speed and dynamic range. Even within the Monte Carlo paradigm, numerous refinements can still be made to more accurately treat topics like anisotropy, off-axis emission, or magnetic field amplification by turbulence. In the next, final, chapter, I will discuss extensions to the work presented here.
In this chapter I will briefly discuss potential avenues for further research on the topic of this dissertation. The point I wish to make is that the work presented in previous chapters is far from a dead-end, settled problem. There are many questions that can still be answered, and refinements to the model that can still be made.

- As discussed in Chapter 8, the Blandford-McKee solution is increasingly inadequate late into the trans-relativistic regime. Numerical hydrodynamical simulations of GRB jets are readily available, in 1-D and multi-D. Using numerical simulations to guide the choice of shock parameters would allow me to continue the afterglow simulation much deeper into the trans-relativistic phase, potentially as far as the transition to the non-relativistic Sedov-Taylor stage.

- With access to multi-dimensional hydrodynamical simulations, it becomes necessary to consider the effects of angular distance from the jet axis, as well as the jet break and lateral expansion discussed in Section 2.2.4. This will require calculating the photon production at all angles relative to the shock normal, rather than considering only photons that transform into a small cone around the jet axis. Indeed, the 1-D shock profiles assumed throughout this dissertation may prove to be inadequate, as shock speed and other DSA parameters may vary...
by angular distance from the axis.

- The smoothing method used in this dissertation is ill-behaved at high Lorentz factors, which is the principal reason that the simulation presented in Chapter 8 started as late as it did. A more detailed examination of the smoothing process—paying particular attention to the thermodynamics of populations of particles that span trans-relativistic thermal energies—might allow the afterglow simulation to begin much earlier.

- Figure 5.1 contains a significant numerical problem. The transformed bin of the phase-space diagram does not align with the axes into which it must be binned. In this dissertation it was assumed that binning the center was sufficient, but this is a clear approximation. Dealing with the transformation correctly requires identification of overlap between arbitrary polygons. This problem is well-known in computational geometry, and numerous algorithms exist for handling it. Handling the Lorentz transforms properly is thus one of the easier extensions to the dissertation, if time-consuming. Note that the same problem exists when the plasma-frame photon distributions of synchrotron and pion decay must be transformed into the ISM frame.

- Both the Kamae et al. (2006) and Kelner et al. (2006) parametrizations of pion decay emission assume that the target nucleons are essentially at rest in the plasma frame. This is a satisfactory assumption in the context of supernova remnants or molecular clouds, where temperatures are low enough that the bulk of the particles are nonrelativistic. In GRB afterglows, even the “thermal” particles are fully relativistic once they cross the shock, rendering this assumption quite invalid. More difficult still is the situation in the trans-relativistic stage of the afterglow. Once the thermal particles are only mildly relativistic, whether any given interaction exceeds the pion production threshold depends on the angle between the particles’ velocity vectors. To my knowledge, this issue has not been considered in the literature at all.

- This dissertation focused on DSA occurring at an external shock propagating into the ISM—the boundary between regions 1 and 2 in Figure 2.4. However, a reverse shock might propagate down the jet, shock-heating the ejecta and potentially energizing some of those particles. The existence of a reverse shock is highly uncertain, but if it exists the Monte Carlo code would be able to model DSA in this extreme environment. Here, again, hydrodynamical (preferably magnetohydrodynamical) simulations would be helpful in identifying conditions relevant to shock formation and DSA.

- The Monte Carlo code is computationally inexpensive enough to permit large parameter searches. With wide enough coverage of the parameter space, building a library of afterglows.
becomes possible. I may, then, be able to work backwards from a particular GRB afterglow and infer the environment of the progenitor, or details of the microphysics at the forward/reverse shock of the jet.


O. Blaes, R. Blandford, P. Madau, and S. Koonin. Slowly accreting neutron stars and the origin of


A. Bykov, N. Gehrels, H. Krawczynski, M. Lemoine, G. Pelletier, and M. Pohl. Particle Acceleration


D. C. Ellison and D. Eichler. Monte Carlo shock-like solutions to the Boltzmann equation with


S. R. Kelner, F. A. Aharonian, and V. V. Bugayov. Energy spectra of gamma rays, electrons, and


B. Reville and A. R. Bell. On the maximum energy of shock-accelerated cosmic rays at ultra-relativistic


Y. Uchiyama, F. A. Aharonian, T. Tanaka, T. Takahashi, and Y. Maeda. Extremely fast accel-


