Magnetic Reconnection and the Extreme Plasmas of Blazar Jets

by

J. M. Mehlhaff

B.S., University of Washington, 2016

M.S., University of Colorado, 2019

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Physics

2021

Committee Members:

Dmitri Uzdensky, Chair

Gregory Werner

Mitchell Begelman

John Cary

Jason Dexter

Mehlhaff, J. M. (Ph.D., Physics)

Magnetic Reconnection and the Extreme Plasmas of Blazar Jets

Thesis directed by Prof. Dmitri Uzdensky

Magnetic reconnection is a plasma physical process in which a magnetic field is topologically rearranged, releasing free energy and imparting it to plasma matter. Though seemingly ubiquitous as a magnetic energy conversion channel in laboratory, space, and astrophysical settings, magnetic reconnection is by no means the same in these diverse environments. Astrophysical reconnection, in particular, is quite different from the type that occurs in the solar system. This difference stems from several unique and extreme properties of astrophysical plasmas. While plasmas on Earth and in nearby space are often non-relativistic, made of electrons and ions (e.g., protons), and non-radiative, astrophysical plasmas are quite frequently relativistic (in both the special and general senses), made of electron-positron pairs, and highly radiative – exhibiting non-trivial radiation-matter (e.g., QED) interactions.

In this dissertation, I focus on radiative relativistic reconnection regimes that prominently feature many of the above extreme effects. Here, the initial magnetic energy per particle exceeds the electron rest mass, enabling reconnection to accelerate relativistic particles. These particles go on to suffer strong radiative losses, emitting photons that may later be absorbed inside the system to produce pairs. Feedback from both direct radiative cooling and pair production contributes to a rich collective interplay between reconnection and high-energy radiative physics.

Throughout this work, I treat reconnection as an idealized physical process without explicit reference to many aspects of the global astrophysical environments where it unfolds. Instead, I obtain astrophysical relevance through the relativistic and radiative physics and through the parameter space that I explore. These mimic reconnection conditions in blazar jets – collimated relativistic outflows from active galactic nuclei traveling toward the observer. Adopting a simplified microphysical perspective enables me to model the plasma (including radiation) with high fidelity

and from first principles, especially in the large portion of this work based on numerical simulations. In addition to revealing the nature and observable features of reconnection in blazar jets, the resulting insight sometimes actually constrains the global features of the jet itself, illustrating the utility of a bottom-up, physics-first approach.

Dedication

To Harvey Mehlhaff, whose paraphrase of Psalm 118:24 has been especially helpful during life as a graduate student:

> Today is the best day of my life because it's the only day I've got. Yesterday's gone and tomorrow's not here yet.

Acknowledgements

The list of those to whom this work is indebted extends back many years before my time at the University of Colorado. To many dear family members, friends, teachers, and mentors who I cannot list by name, please forgive my omission and know that I am truly grateful for your support.

I would first like to thank my PhD research advisor, Dmitri Uzdensky. Few students are privileged with an advisor who has such a firm and compassionate commitment to their success. Of all the many things I have learned from Dmitri, I am especially grateful for his commitment to communicate important intangibles, including the value of clear and thoughtful exposition in making one's work appreciable by a large audience.

I would also like to thank my co-advisors, Greg Werner and Mitch Begelman, who provided patient, crucial guidance on, respectively, the computational and astrophysical aspects of my work.

To my fellow graduate students, especially Kai Wong, Lia Hankla, and Muni Zhou: your camaraderie, insight, and encouragement have been vital to my studies and well-being. Thank you.

I am further grateful to several good friends from The Well Church, many of whom are also graduate students, who have been a refreshing and thoughtful community to me while in Boulder.

Additionally, I am deeply indebted to another community – the staff and residents of International Crossroads – especially to Dwain Upton, Edi Idrissi, and Karston Christensen. Your friendship and prayers have made the last several years a gift in ways I never could have imagined.

Finally, to my dear parents, Mark and Kimberly, and to my brother, Joshua, whose enduring support in all aspects of my life has remained undiminished in these graduate school years – none of this would have been possible without you.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Introduction

This Introduction begins very broadly in order to orient the uninitiated reader (for example, an advanced undergraduate or beginning graduate student in physics). Plasma physics experts without strong astrophysics background may wish to begin in section 1.2; plasma astrophysics practitioners may skip to section 1.5 or 1.6.

1.1 Magnetic reconnection: a microcosm of the plasma universe

Plasmas, often called the 'fourth state of matter', are by far the most abundant type of ordinary matter (excluding dark matter and dark energy) in the visible universe (e.g., Gurnett & Bhattacharjee, 2017). As one might expect from this fact, a thorough grasp of fundamental plasma physics is essential to understand many of the most fascinating and important systems in the universe. In nature, such systems include: the interior and exterior of our Sun; the launch and propagation of the solar wind – including its interaction with the Earth's magnetosphere (one consequence of which is the aurora borealis) – and some of the most extreme astrophysical objects: neutron stars, black holes, and their accretion disks and jets. The applications of plasma physics in the lab and in high technology are equally varied and important. These include the development of: nuclear fusion (a potentially abundant source of clean energy); plasma-based particle accelerators; new experimental facilities to probe the interaction of radiation and matter in extreme uncharted regimes; plasma etching of integrated circuits; and plasma surface sterilization against biological contaminants (National Academies of Sciences, Engineering, and Medicine, 2021).

In addition to its many societally and scientifically impactful applications, plasma physics is also a highly fascinating discipline from a purely intellectual perspective. A plasma is essentially a gas that has been ionized, may therefore conduct current and, as a result, both produces and responds to electromagnetic fields. The long range electromagnetic forces that couple to a plasma give rise to complex and fascinating collective behavior, and plasma dynamics are hence much richer than those of ordinary gases. In fact, so interwoven is the interplay between the material of a plasma and the fields in which it is immersed, that it is arguably more appropriate to use the term 'plasma' to refer to the matter *and* the fields as one collective object (rather than just to the current-carrying charged-gas component).

The study of magnetic reconnection, one particular plasma process, very much reflects the spirit of plasma physics as an overarching discipline: in its pursuit one finds both a wealth of practical applications and rich fundamental insight. As a physical process, magnetic reconnection prominently features the nonlinear collective interactions characteristic of plasma physics. At a glance, reconnection occurs when two magnetized plasma regions of opposite polarity (magnetic field direction) come into close proximity. The plasma current sheet that self-consistently forms between the spatial polarity domains (satisfying Ampère's Law) is unstable. Often, it quickly evolves into a configuration with one or more *X-points* where the current-perpendicular magnetic field goes to zero. At these cites, the magnetic field lines that originally belonged to one or the other polarity domains come together and reconnect (Fig. 1.1). This relaxes the magnetic field to a lower energy configuration and transfers much of its initial free energy to the plasma matter.

Besides being an intrinsically interesting collective plasma process in its own right, reconnection turns out to have ubiquitous applications in many of the laboratory, space, and astrophysical[1] plasma systems already mentioned (e.g., Zweibel & Yamada, 2009). Either the system is configured at a large scale to produce reconnecting current sheets, or they form as the inevitable result of mi-

[1] I adopt the standard terminology that *space physics* generally refers to studies of regions of space close enough to Earth to feasibly be probed *in situ* with spacecraft. Space physics is therefore generally restricted to inside of, or at least very close to, the heliosphere, which is the Sun's domain of magnetic influence. *Astrophysics*, in contrast, refers to phenomena that are generally so far away that they can only be observed remotely.

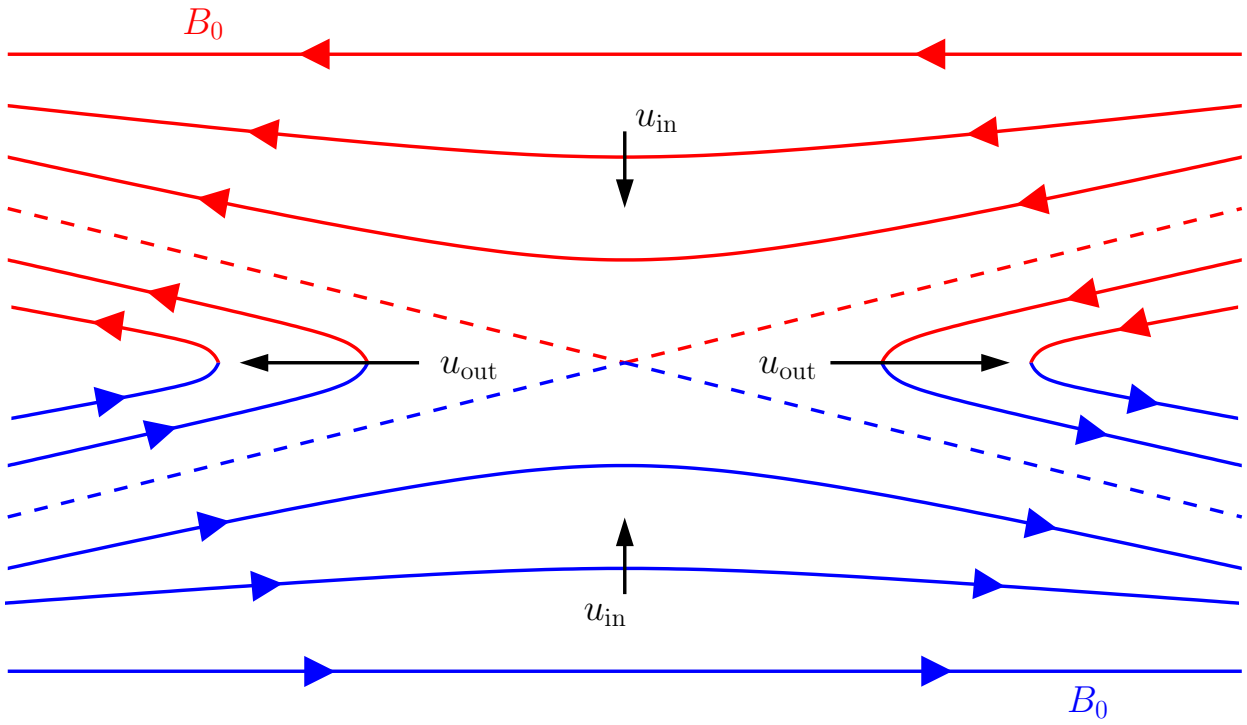

Figure 1.1: Cartoon diagram of reconnection centered around a single X-point. Plasma material and antiparallel magnetic field lines are carried toward the midplane with velocity u_{in}. The field lines are bent toward the separatrix (dashed lines), which distinguishes the regions of unreconnected flux (where, in the diagram, the field lines bear a single color) from the region of reconnected flux (where the field lines each possess two colors). At the X-point where the dashed lines cross, the in-plane magnetic field vanishes, and two field lines that were once separate come together and reconnect. Field lines leaving the X-point carry plasma material at the outflow velocity $u_{\text{out}} > u_{\text{in}}$ and have two 'halves': one that was reconnected from the upper (red) domain, and another that was reconnected from the lower (blue) domain.

crophysical turbulent dynamics. In both cases, the resultant change in the topology of the magnetic field as well as the plasma matter energization – despite happening at very localized, microscopic locations – directly affect the global structure and evolution of the system. This underscores the need for detailed, rigorous, microphysically grounded theories of reconnection to build up a confident understanding of how this process impacts the global properties of the plasma systems where it occurs.

This need is very poignantly felt in the quest to understand the astrophysical systems of interest to this work. In order to more thoroughly elucidate the relationship between reconnection plasma microphysics and global astrophysical dynamics, I now turn to some of the unique theoretical challenges faced in the study of astrophysical plasmas. Very much emblematic of the discussion thus far, these challenges are both intrinsically interesting and highly rewarding in terms of the applied astrophysical insight that stands to be gleaned.

1.2 Extreme high-energy astrophysical plasmas

(This section, especially its technical notions of 'extreme' and 'traditional', borrows heavily from Uzdensky et al. 2019.)

Many of the traditional thrusts in plasma physics – especially the development of fusion energy and research into space and solar physics – explicitly deal with plasmas that share a number of common traits. Usually, these plasmas are non-relativistic, comprise electrons and ions, and are non-radiative (in the sense that radiation does not significantly impact the plasma dynamics). By contrast, because the objects of interest to high-energy astrophysics – particularly neutron stars, black holes, their accretion modes, outflows, and mergers – imbue their environments with high material and electromagnetic energy densities, the plasmas they host are often extreme and exotic in several ways compared to those encountered in traditional laboratory and space settings. Astrophysical plasmas must very often be treated using a relativistic formalism, which frequently includes general relativity, especially where black holes are concerned. In addition, the high energy densities permeating astrophysical plasmas can render them highly radiative, with strong cooling

losses significantly modifying the plasma behavior. However, radiative cooling is by no means the only channel through which astrophysical plasmas interact radiatively. The energy scales near relativistic compact objects are often high enough to spark (i.e., are above threshold for) a number of quantum electrodynamics interactions, including pair production. Thus, the photons emitted by the radiating particles come into their own as a fundamental, dynamically important 'particle' species, significantly contributing to the rich collective plasma behavior. The reality of pair production in astrophysics also means that one often deals with plasmas containing (if not composed solely of) electron-positron pairs that *do not*, in general, satisfy particle number conservation, as do traditional electron-ion plasmas.

All of these departures from traditional plasma physical regimes, with the exception of general relativity, are of critical importance in this dissertation. However, they are not the only complications to the story. Besides their extreme properties, the plasmas encountered in this work also possess a number of more common (e.g., that may also be encountered in more traditional settings) but still nontrivial properties that need to be addressed.

Plasmas found in neutron star and black hole environments are usually highly magnetized and quite tenuous, with mean-free-paths to Coulomb collisions commonly exceeding the macroscopic system size. This deprives the plasma particles of a direct mechanism through which to thermalize, and means that it is unknown *a priori* what the distribution of particle velocities will be at any given point in space. Thus, one cannot treat the plasma as a conducting fluid (e.g., as in magnetohydrodynamics, or MHD), which entails assuming the particle velocity distribution. One instead needs a kinetic paradigm in which the density of particles is not just calculated dynamically in real space, but in full 6-dimensional phase space.

The need for a kinetic paradigm is actually twofold in high-energy plasma astrophysics. Besides that it is required for a realistic physical description, it also provides a critical connection point to two of the main diagnostics available from observations: the time variability and frequency spectrum of received light. Unlike laboratory and space plasmas, which can be highly spatially resolved when observed remotely, and even probed *in situ* at individual points in space, astrophysical sources

often appear point-like on the sky. This places enhanced burden on timing data, which take the form of lightcurves – plots of received radiation power as a function of time – and spectral data, which take the form of spectral energy densities – distributions of received radiation decomposed by photon frequency. Because the spectrum of emitted light depends directly on the distribution of radiating particles, a kinetic plasma description allows one to construct a theory with predictions that may be compared against critical spectrum-oriented diagnostics. When it comes to lightcurves, a kinetic paradigm also enhances one's ability to make frequency-resolved timing predictions.

A powerful computational tool for studying, at a kinetic level, extreme, magnetized, collisionless astrophysical plasmas is the particle-in-cell (PIC) technique (e.g., Birdsall et al., 2005). In PIC codes, Maxwell's equations are solved on a computational grid, the resulting fields are used to update the positions of a large number of particles, and then the charges/currents produced by the particles are deposited back onto the grid, enabling Maxwell's equations to be solved again at the next timestep. Hence, PIC codes self-consistently account for feedback between the plasma particles (including their velocity distribution) and the electromagnetic fields. The method is also massively parallelizable and can be straightforwardly decorated with extreme plasma physical processes like radiation-reaction and pair production. For these reasons, PIC simulations, along with analytic theory, form an integral component of this dissertation.

To summarize, astrophysical plasmas are often highly magnetized and collisionless, and they commonly possess a number of extreme properties (e.g., relativity, radiation, and pair production). A kinetic description is needed both to accurately model the plasma behavior and to calculate observable characteristics of the radiated light. The required kinetic description can be implemented numerically using the PIC technique, and PIC simulations constitute one of the main tools I use to study astrophysical plasmas in this work.

1.3 Reconnection couples microphysics to macrophysics in blazars

Having overviewed some of the basic properties of astrophysical plasmas, I now introduce blazars, the primary type of astrophysical system of interest to this work. I then bring magnetic

reconnection back into the picture, discussing some of its important roles in these objects.

A blazar is an observational manifestation of an active galactic nucleus – a supermassive black hole at the center of a galaxy that is actively accreting matter – launching counterpropagating relativistic jets, one of which travels almost parallel to the observing line-of-sight. These jets are vast and powerful, both in ways that strain the imagination. Though launched from black holes that would fit comfortably within our solar system, $\sim 1\,\mathrm{AU}$ in size, they can remain well-collimated, in the most extreme cases, out to almost $\sim 1\,\mathrm{Mpc}$, i.e., well beyond their host galaxy and even bordering on typical intergalactic distances (Blandford et al., 2019).[2] Considering, in addition, that the microphysical plasma length scales near the black hole where the jet is launched can be about 10^7 times smaller than the black hole itself (Crinquand et al., 2020), the total scale separation – from microscopic plasma physics to the size of the jet – is $\sim 10^7\,\mathrm{Mpc}/1\,\mathrm{AU} \sim 10^{18}$. To put this in perspective, blazar jets dwarf the microphysical length scales governing their operation by even more than macroscopic human-sized objects (typical size 1 m) dwarf atomic nuclei (typical size 10^{-15} m). Equally large (in a literal sense) is the breadth of the electromagnetic spectrum across which blazars are observed: all the way from radio (upwards of hundreds of MHz) to TeV gamma-rays, giving a total frequency range of $1\,\mathrm{TeV}/100\,h\,\mathrm{MHz} \sim 10^{18}$, where h is Planck's constant (Begelman et al., 1984; Madejski & Sikora, 2016; Blandford et al., 2019). Because blazar jets are relativistic and pointed almost directly toward the observer, their emission is strongly Doppler boosted. This enhances both their apparent brightness – a blazar can outshine its entire host galaxy – and the energies of the emitted photons – as a class, blazars dominate the point-like sources on the extragalactic gamma-ray sky (e.g., `http://tevcat.uchicago.edu/`; Abdollahi et al., 2020).

Some of the main mysteries surrounding blazars, most of which are still outstanding research questions, include (but are by no means limited to): how the jet is launched; where along the jet axis most of the emission is produced; how energy is transported and dissipated along the jet (which is related to the previous question); how the radiation at the highest energies is produced; and what the jet is made of.

[2] For reference, $1\,\mathrm{AU} = 1.5 \times 10^{13}\,\mathrm{cm}$, $1\,\mathrm{Mpc} = 10^6\,\mathrm{pc}$, $1\,\mathrm{pc} = 3.3\,c\,\mathrm{yr} = 3.1 \times 10^{18}\,\mathrm{cm}$ and c is the speed of light.

As for the first question – how the jet is launched – one common paradigm involves the advection, by the accretion flow, of magnetic fields onto the rotating supermassive black hole central engine.[3] Once the field lines are planted onto the black hole, the rapid rotation of spacetime twists them up into a kind of spring/slinky configuration. Then, like a compressed spring, the magnetic field lines expand, driving counterpropagating outflows along the black hole spin axis (Blandford & Znajek, 1977; Davis & Tchekhovskoy, 2020).

As for the rest of the blazar questions mentioned above, it turns out that magnetic reconnection may play a critical role in each of them. When it comes to the issue of energy transport and dissipation along the jet, reconnection comports nicely with the magnetic launch mechanism of Blandford and Znajek and the requirement for some *in situ* particle acceleration process to offset adiabatic expansion and radiative cooling of the particles that shine all along the jet's length (Blandford & Königl, 1979; Begelman et al., 1984). For example, magnetic reconnection between 'striped' domains of opposite field polarity – imprinted within the jet at its launch point – may gradually dissipate energy along the jet as it propagates (Giannios & Uzdensky, 2019).

Local jet reconnection further provides an attractive explanation for the observed emission. Reconnection occurs naturally within a moderately-to-highly magnetized jet, and the current sheets that cause it may not only form as the result of large-scale plasma motion (as in the picture of Giannios & Uzdensky, 2019) but also as the result, at small scales, of turbulence (e.g., Zhdankin et al., 2013, 2020; Comisso & Sironi, 2018, 2019; Loureiro & Boldyrev, 2020; Boldyrev & Loureiro, 2020; Nättilä & Beloborodov, 2020; Sobacchi et al., 2021). Additionally, magnetic reconnection may explain the non-thermal (e.g., power-law) spectra that are generally observed from blazars (e.g., Fossati et al., 1998; Ghisellini, 2011). This is because collisionless relativistic reconnection tends to yield non-thermal distributions of accelerated particles (a phenomenon called *non-thermal particle acceleration*, or NTPA) as found by numerous PIC simulation studies of both pair plasmas

[3] In highly conductive media, including plasmas, magnetic field lines are carried along with, and twisted up by, the bulk flow. In turn, the field lines exert pressure (they do not like to be bunched together) and tension (they prefer to be straight rather than bent – much like guitar strings) on the underlying material. Thus, in plasma physics, magnetic field lines are, in some sense, more than just a convenient tool for visualizing the underlying field; they take on genuine material properties of their own.

(Zenitani & Hoshino, 2001, 2007, 2008; Jaroschek et al., 2004; Bessho & Bhattacharjee, 2007, 2012; Lyubarsky & Liverts, 2008; Cerutti et al., 2012b, 2013, 2014a,b; Guo et al., 2014, 2015, 2019, 2020; Sironi & Spitkovsky, 2014; Sironi et al., 2016; Werner et al., 2016; Werner & Uzdensky, 2017; Petropoulou & Sironi, 2018; Schoeffler et al., 2019; Werner et al., 2019; Sironi & Beloborodov, 2020; Hakobyan et al., 2021; Ortuño-Macías & Nalewajko, 2020; Mehlhaff et al., 2020) and electron-ion plasmas (Melzani et al., 2014a,b; Dahlin et al., 2015, 2017; Guo et al., 2016; Ball et al., 2018; Werner et al., 2018). Finally, shock acceleration, a potential alternative NTPA mechanism, is likely to be suppressed in the magnetized, relativistically propagating jet (Sironi & Spitkovsky, 2009; Sironi et al., 2015).

As far as the material composition of the jet is concerned, reconnection may mix material from a geometrically thick accretion flow into the nascent jet, feeding in electron-ion plasma at its base (Ripperda et al., 2020; Wong et al., 2021). Then, farther downstream, reconnection may contribute to *in situ* pair production (Mehlhaff et al., 2021).

All throughout this (very truncated) list of ways that reconnection may influence critical global properties of blazar jets, microscopic kinetic plasma physics plays a decisive role. As discussed already, the spectrum of radiation produced by reconnection in collisionless plasmas requires a kinetic description in order to compute. But even more than that, crucial issues, such as the influence of pair production, depend sensitively on the distribution of radiating particles and on the distribution of photons that those particles produce – information that can only be self-consistently calculated in a kinetic plasma description. Likewise, reconnection-mediated material mixing between the accretion flow and the young jet, as well as how far the jet propagates before dissipating most of its initial magnetic energy, depend on how fast collisionless reconnection proceeds.

Thus, it is simply not the case that reconnection microphysics can be studied in isolation from the macroscopic dynamics of the overarching system. On the contrary, this fundamental plasma process plays a critical role in deciding key global properties. Hence, there is a poignant need for self-consistent, rigorous, first-principles theories of reconnection in extreme astrophysical regimes.

In this dissertation, I aim to address this need in ways it has not yet been confronted. To

do so, I start at the level of the fundamental plasma physics of magnetic reconnection. Then, using insight gleaned there, I extrapolate outward, pointing out the implications for the global host system, especially its observable features. The type of reconnection that occupies the main focus of this dissertation is one that is highly relevant to blazar jets: radiative relativistic magnetic reconnection. Next, I discuss the general properties of reconnection as it proceeds in this regime.

1.4 Radiative relativistic magnetic reconnection

In a scenario where magnetic reconnection powers observed blazar emission, the magnetic field is often envisioned to dominate the overall energy budget of the jet (over the energy contained in matter and radiation) until reconnection liberates this energy, delivering it to particles which may then radiate the light that is seen (e.g. Giannios et al., 2009, 2010; Nalewajko et al., 2011, 2012; Giannios, 2013; Sironi et al., 2015; Petropoulou et al., 2016; Werner et al., 2018). The plasma feeding reconnection is thus magnetically dominated, but not just in the traditional (e.g., solar, space, and laboratory) plasma physical sense, in which the plasma magnetic-to-kinetic energy density ratio is large. In addition, the magnetic energy per particle exceeds the electron *rest mass*. Thus, reconnection proceeds in the relativistic regime (Blackman & Field, 1994; Lyutikov & Uzdensky, 2003; Lyubarsky, 2005), in which it accelerates relativistic individual particles and bulk flows. However, relativistic motion is far from the only extreme feature of reconnection in blazar jets. In addition, radiation can also dramatically impact the qualitative features of reconnection. In fact, radiative processes are always at work by necessity in astrophysical reconnection: there must be at play some radiative mechanism(s) that channel(s) the kinetic energy delivered, through reconnection, to particles into observed light. The question is whether the two steps in this energy conversion process – from the magnetic field to plasma through reconnection, and from plasma to light through radiation – can be imagined as happening separately. This is certainly the case, for example, if the plasma radiative cooling time is longer than the characteristic time-scale on which reconnection occurs. The plasma particles are *first* energized by reconnection and *subsequently* radiate. In such circumstances, the observable signatures of reconnection are, in some sense, simple

to compute: one first characterizes the particle energy distribution produced by reconnection in the absence of radiative cooling. Then, as a second and independent step, one may use this distribution to compute the radiation spectrum.

If, however (as is often thought to occur at X- and gamma-ray energies, especially in blazars), the emitting particles cool on time-scales comparable to – or even much shorter than – the reconnection dynamical time, then radiative and reconnection physics are inseparable. The coupling between them may, furthermore, be facilitated by much more than just optically thin radiative cooling, where the emitting particles suffer radiative drag but the produced photons passively escape the system. In some situations, the optical depths to various processes, including Thomson scattering and pair-creation, may exceed unity, affording the emitted photons further opportunity to impact the ongoing reconnection process. Even when these optical depths are small – but especially when they are not – a self-consistent approach that models radiation and reconnection simultaneously is required to capture the modifications that the various radiative interactions may make to the reconnection-powered photon spectrum. In this dissertation, I term any regime of this kind, in which reconnection and photon processes are inextricably coupled, as a 'radiative' regime of magnetic reconnection (Uzdensky & McKinney, 2011; Uzdensky, 2016).

Although blazars provide the main astrophysical impetus to this dissertation, it is worth mentioning that the desire to understand radiative relativistic reconnection more generally, because of its broad applications in astrophysics, has spurred a large number of recent studies on this type of reconnection regime. Some of the astrophysical contexts for this work have been: pulsar winds (Pétri, 2012; Cerutti & Philippov, 2017; Cerutti et al., 2020), pulsar wind nebulae (Uzdensky et al., 2011; Cerutti et al., 2012a, 2013, 2014a,b; Yuan et al., 2016), pulsar magnetospheres (Lyubarskii, 1996; Uzdensky & Spitkovsky, 2014; Philippov et al., 2015; Cerutti et al., 2016; Philippov & Spitkovsky, 2018; Hakobyan et al., 2019), magnetar magnetospheres (Schoeffler et al., 2019), gamma-ray bursts (McKinney & Uzdensky, 2012), accreting black holes (Beloborodov, 2017; Werner et al., 2019; Sironi & Beloborodov, 2020; Mehlhaff et al., 2021), blazars (Nalewajko et al., 2011, 2012, 2018; Ortuño-Macías & Nalewajko, 2020; Sobacchi et al., 2021; Mehlhaff et al., 2020,

2021), and black hole magnetospheres (Parfrey et al., 2019; Crinquand et al., 2021). There have also been several recent radiative reconnection[4] studies that have not focused on a single object class, but have still been motivated by some combination of the above systems (e.g., Jaroschek & Hoshino, 2009; Uzdensky & McKinney, 2011; Uzdensky, 2011, 2016; Hakobyan et al., 2021; Nättilä & Beloborodov, 2020).

Let me now discuss some of the specific cooling mechanisms that may facilitate radiative reconnection, honing in on those that are the main focus of this study. Two common emission mechanisms in astrophysics are synchrotron and inverse Compton (IC) cooling. Cooling as a result of these processes can be strong in its own right, initiating radiative reconnection by causing the emitting particles to lose their energy rapidly. Additionally, these mechanisms may enable radiative reconnection by supplying photons that go on to feed back on the evolving plasma (e.g., through pair production).

In the context of blazars, both the synchrotron and IC channels, as well as secondary photon interactions, can all be important. I focus in this work on radiative reconnection mediated by IC cooling, including pair production that may be initiated by the high-energy IC-produced photons (Mehlhaff et al., 2021). This specialization constitutes a crucial and necessary simplification, narrowing this dissertation to a feasible scope. However, the results are still applicable to a broad class of blazars (namely, the flat-spectrum radio quasars, which I discuss in more detail in chapters 4 and 8) where IC cooling is expected to dominate.

In brief (I supply more technical discussions in chapters 2 and 6), inverse Compton cooling occurs when a relativistic particle (in this work, an electron or positron) traverses an ambient soft radiation field – a bath of low-energy *seed* photons. This radiation is assumed not to be beamed to within a tighter half-opening angle than the inverse of the particle's Lorentz factor (quite reasonable for highly relativistic particles). The particle suffers Compton collisions with the ambient photons as it travels, and, due to its motion, the photons tend to impart momentum in the direction opposite

[4] From here onward, I sometimes refer to radiative relativistic magnetic reconnection as just 'radiative reconnection'. Although reconnection does not, in general, have to be relativistic in order to be radiative, all instances of radiative reconnection in this work are also relativistic.

its initial velocity. Thus, the particle loses energy on average, upscattering the background photons to higher energies than they originally possessed.

At low energies (defined by the product of the scattering particle and seed photon energies normalized by the square of the electron rest mass energy), Compton scattering in the particle's rest frame reduces to Thomson scattering. Then, the particle does not suffer appreciable recoil in a single scattering, and radiative losses become essentially continuous. On the other hand, at high energies (defined in the same sense) – the Klein-Nishina regime – each photon collision causes the particle to lose an order-unity fraction of its energy and thus IC losses are essentially discrete. I treat both the Thomson and Klein-Nishina IC regimes in this work.

Additionally in this dissertation, I specialize to IC cooling where the soft seed photons are static, homogeneous, isotropic, and roughly monochromatic. They are envisioned to impinge on the reconnection layer from some external, time-independent source. Like the specialization to IC cooling in the first place, this is both a helpful simplification and is also highly relevant to a wide range of blazars. Within this general framework, it turns out that, at approximately the same energies where IC losses become discrete, pair production also kicks in between the Comptonized photons and the ambient radiation field. I therefore treat pair production, together with the self-regulating effect that the produced pairs have on reconnection, when considering high-energy Klein-Nishina Compton losses.

Before getting to a concrete roadmap for this dissertation, I now recapitulate and consolidate the main motivational and introductory material discussed so far. With this summary, I hope to provide large-scale conceptual pillars that can help anchor one's reading of the detailed technical discussions to follow – aids against missing the forest for the trees.

1.5 The story so far

At this point, I have introduced all of the key players that underpin this dissertation. These players come in essentially three main varieties.

(1) First, I have summarized a small sample of general ideas that motivate, and lay out the

general landscape for, studies of astrophysical reconnection and astrophysical plasmas more broadly. In this category, I have pointed out several ways that astrophysical plasmas are 'extreme' relative to their more traditionally studied counterparts found in the laboratory and in space. Though these extreme aspects – which include relativity, radiation, and QED (in particular, pair production) – are highly intrinsically interesting from a strictly intellectual perspective, giving rise to exotic new collective plasma physical effects, they also present a promising landscape for illuminating, from first principles, the inner workings of many astrophysical systems. Thus, though the physics itself is novel with respect to conventional plasma studies, the combination of both fundamental and applied insight is very much in the spirit of plasma physics as a broader discipline.

(2) I have introduced the main *plasma description* and *physical process* relevant to this work. As for the former, a kinetic plasma description, in which the full 6D phase-space density of the plasma particles is evolved self-consistently, is essential. The particle velocity distribution needs to be calculated from first principles both to faithfully describe collisionless and often-radiative astrophysical plasmas and to calculate realistic radiative signatures. With respect to physical processes, this dissertation focuses entirely on one plasma physics phenomenon, magnetic reconnection, but in the relativistic radiative regime where it is decorated with extra, extreme (in the sense described above) radiative processes: IC cooling and pair production. These significantly alter the reconnection dynamics and must be included in order to properly interpret astrophysical observations.

(3) I have introduced the main *astrophysical object* that motivates, and provides the primary application for, this study: blazars. In the process, I have also described how magnetic reconnection – though a fundamental plasma physical process involving scales that are microscopic relative to the size of the blazar host system – is essential to understand (in a radiative relativistic regime) in order to answer many basic questions concerning the global nature of blazars.

With these basic background ideas in tow, I turn now to concretely outline the remainder of this dissertation. Here, I summarize both the particular observationally oriented questions that I confront and the additional, more specific aspects of IC radiative reconnection I investigate in order to address those questions.

1.6 Outline

One of the main observational mysteries I seek to address in this work is rapid gamma-ray blazar flares: flares where the characteristic variability time-scale is shorter than the light crossing time of an important macroscopic system length-scale, such as the radius of the supermassive black hole central engine's event horizon or the transverse width of the jet. A number of such flares have been observed from blazars (Aharonian et al., 2007; Albert et al., 2007; Aleksić et al., 2011; Ackermann et al., 2016; Britto et al., 2016; MAGIC Collaboration et al., 2019) and also from pulsar wind nebulae (in which context the relevant light crossing time is usually taken to be that of the nebula itself; Abdo et al., 2011; Tavani et al., 2011; Buehler et al., 2012) and radio galaxies (Acciari et al., 2009; Abramowski et al., 2012; Aleksić et al., 2014).

Reconnection offers a compelling potential explanation for rapid flares from all three of these diverse object types. They are all expected to host magnetized collisionless plasmas that, as discussed earlier, are prone to dissipation through magnetic reconnection. Furthermore, and also as previously mentioned, reconnection-powered NTPA may explain the observed non-thermal flaring spectra. In addition, a reconnection scenario offers an attractive explanation for the dramatically short flaring time-scales, including the rigid constraints these imply on the underlying driving mechanism for the flares. This latter capability for reconnection to deliver rapid flaring variability relies on an aspect of the reconnection process that I have not yet introduced, but that is very important to this work. That aspect is *kinetic beaming*.

Kinetic beaming as the result of reconnection was originally discovered and studied in the context of the Crab Nebula flares. Cerutti et al. (2012b) showed that, near reconnection X-points (spatial locations where the magnetic field reconnects and goes to zero; see Fig. 1.1), particles are both accelerated and collimated. The acceleration and collimation are correlated in the sense that the particles that receive more energy are focused more tightly than those that receive less. Thus, the collimation is inherently kinetic – it can only be described within a kinetic plasma framework – hence the name *kinetic beaming*. Focused via this mechanism, a beam of high-energy

particles may sweep across the observer's line of sight, and the synchrotron or IC emission of the bunch – also emitted as a beam – may then create a short blip in the measured lightcurve: a 'lighthouse effect'. As a result, for the most energetic radiation, the lightcurve can change on time-scale as short as one-tenth the light crossing time of the reconnection layer (Cerutti et al., 2013, 2014a,b). Moreover, kinetic beaming predicts faster variability at higher energies, and is therefore observationally distinguishable from Doppler beaming (e.g., Rees, 1966), where relativistic bulk motion focuses emission achromatically.

Several lines of evidence suggest that kinetic beaming may also underlie rapid gamma-ray IC flares from blazars (unlike the synchrotron Crab flares, gamma-ray emission in blazars is typically thought to be furnished by the IC process; e.g., Madejski & Sikora, 2016). For a 2010 outburst from PKS 1222+21 (Aleksić et al., 2011; Tanaka et al., 2011), kinetic beaming relaxes the energy density required to feed the emission region to a level accessible in the blazar's jet (Nalewajko et al., 2012). On the observational side, there have been at least two TeV events with approximately symmetric rise and decay times (consistent with a sweeping beam of light) and for which the TeV variability increases with photon energy (Aharonian et al., 2007; H. E. S. S. Collaboration et al., 2010). Yet another line of evidence, from numerical simulations, suggests that kinetic beaming at reconnecting current sheets may naturally occur within the turbulent environment of a magnetized blazar jet (Zhdankin et al., 2020, 2021), and may be mediated by the strong radiative cooling to which the emitting particles are likely subjected (Mehlhaff et al., 2020).

I focus here on the last aspect: the connection between radiative cooling and kinetic beaming. As first argued by Kagan et al. (2016), strong radiative losses may play a critical role in allowing kinetic beaming to manifest observationally. Without them, particles may radiate most of their energy well after their momentum distribution – initially collimated during X-point acceleration – has been isotropized by the magnetic field external to the acceleration site. However, when strongly cooled, particles are likely to radiate their energy before dispersing, leading to sweeping beams of light that may manifest as rapid flares.

This is potentially quite important as far as reconnection models for rapid gamma-ray flares

are concerned: it means that reconnection may need to operate in a highly radiative regime in order to deliver the observed flaring time-scales. Indeed, as one expects from the logic of Kagan et al. (2016), large simulations conducted by Sironi & Spitkovsky (2014) and Sironi et al. (2016), which did not incorporate radiative losses, showed little to no anisotropy in the momentum distributions of the high-energy particles. Additionally, Yuan et al. (2016) observed beaming to cease at late times in radiatively inefficient simulations. These results contrast starkly, however, with highly radiative simulations (Cerutti et al., 2013, 2014a,b), tailored to the Crab Nebula flares, that showed prominent kinetic beaming in angular distributions of high-energy emitted photons.

Thus, kinetic beaming appears to be a generic byproduct of the reconnection acceleration mechanism (Uzdensky et al., 2011; Cerutti et al., 2012a,b, 2013, 2014a,b), but it may not always be observationally relevant (Sironi & Spitkovsky, 2014; Kagan et al., 2016; Sironi et al., 2016; Yuan et al., 2016). The particles emitting synchrotron or IC photons at the energies of interest must do so efficiently. Otherwise, the observable signatures of kinetic beaming (i.e., rapid lightcurve variability) may become washed out by particle isotropization. Moreover, the issue of what level of radiative efficiency is required for kinetic beaming to yield rapid flares is of tremendous astrophysical importance. It decides whether and under what conditions a kinetic beaming scenario may be exported to other astrophysical contexts (beyond pulsar wind nebulae), such as blazars, in which it also holds promise for explaining fast outbursts (Nalewajko et al., 2012; Zhdankin et al., 2020, 2021; Mehlhaff et al., 2020).

The initial chapters of this dissertation are therefore dedicated to exploring the dependence of observable kinetic beaming in reconnection on radiative cooling, and they mainly concern material first published by Mehlhaff et al. (2020). The research proceeds in two main thrusts. The first of these thrusts focuses on the plasma physics of radiative reconnection in the Thomson IC regime. Here, reconnection-energized particles IC-scatter externally supplied soft seed photons purely in the Thomson limit. I detail some the basic physical parameters, cast as energy scales, governing Thomson IC radiative reconnection in chapter 2. Also in that chapter, I classify some of the radiative reconnection regimes that one may expect based on the ordering of these scales. This

exercise provides foundational notions of radiative efficiency to anchor the subsequent discussion. Thereafter, in chapter 3, I describe PIC simulations of Thomson radiative reconnection run using my research group's in-house code, ZELTRON (Cerutti et al., 2013; Cerutti & Werner, 2019). By analyzing these simulations, I explicitly show that efficient cooling is required in order for kinetic beaming to imprint itself upon the observed light.

After presenting the numerical findings, I shift to the second main thrust of the part of this dissertation dedicated to exploring kinetic beaming in radiative reconnection (as first presented by Mehlhaff et al., 2020): astrophysical implications of the plasma physical results. I move, in chapter 4, to discuss how a strong radiative efficiency requirement for reconnection-based flare models, in the context of blazars, can constrain the transverse structure of the jet. Within a Thomson IC framework, the most appropriate global system configuration involves a structured jet (for example, a spine-sheath setup Ghisellini et al., 2005; Sikora et al., 2016) wherein a slower moving (e.g., a sheath) region of the jet supplies photons that seed reconnection-powered TeV IC emission in a faster moving (e.g., a spine) region. This result is quite powerful, and illustrates how studying microscopic plasma physics from first principles may inform theories of the overarching macroscopic astrophysical system.

One question that remains unanswered after the initial chapters of the dissertation, outlined above, is the influence of Klein-Nishina effects on IC radiative reconnection, including the observational consequences for blazar emission. For example, Klein-Nishina effects may change the radiative conditions necessary for kinetic beaming to deliver rapid gamma-ray flares, potentially allowing global system configurations (for example, a transversely homogeneous jet) that are ruled out within an exclusively Thomson IC framework. In fact, if the seed photons for IC gamma-ray flares come, as is more conventionally thought, from beyond the jet itself – a scenario disfavored by my Thomson radiative reconnection analysis (Mehlhaff et al., 2020) – then observations strongly suggest the seed photons are Comptonized in the Klein-Nishina regime. Thus, to either firmly rule out, or potentially restore viability to, jet-external seed photon sources in reconnection-based flare models requires self-consistently treating Klein-Nishina effects in radiative reconnection. Thus,

starting in chapter 6, I transition this dissertation to its second main component, which focuses on Klein-Nishina IC radiative reconnection. This includes findings first presented by Mehlhaff et al. (2021) and also preliminary unpublished results.

In chapter 6, I cover the basic parameters governing Klein-Nishina radiative reconnection. This discussion is in the same style as, and generalizes, that of chapter 2. In it, as in chapter 2, I phrase the fundamental reconnection parameters as energy scales, and classify the radiative regimes one may expect by permuting the scale hierarchy. The difference from chapter 2 is that, with Klein-Nishina effects present, there are more scales to deal with, increasing the dimensionality of the parameter space and uncovering many new and unexplored regimes.

Because Klein-Nishina reconnection involves fundamentally new and unstudied collective effects, which include pair-production of the Comptonized photons with their parent seed photon population, I do not jump, as I do with Thomson radiative reconnection, straight into simulations after organizing the scales in the problem and classifying different regimes. Instead, to lay the foundation for future Klein-Nishina reconnection studies – both numerical and analytic – in chapter 7, I provide a detailed analytic model of Klein-Nishina reconnection in a wide swath of the parameter space outlined in chapter 6. The most salient feature of this model is the self-regulating role played by pairs produced as the result of reconnection on the subsequent dynamics. This is a fundamentally new way in which radiative processes impact reconnection (beyond pure radiative cooling); it cannot occur in the Thomson IC regime.

Then, in chapter 8, I overview some of the most salient observable features of the analytic model of pair-regulated Klein-Nishina reconnection. I point out, in particular, connections to kinetic beaming and TeV blazar flares, but I also explore more general consequences of pair-regulated Klein-Nishina reconnection on quiescent blazar emission. This latter focus is important because very little work has been done on Klein-Nishina reconnection and, therefore, little is known about what it may look like when powering emission of any time-scale (not just rapid flares). I further list connections to observations of other objects, namely accreting black holes, in which Klein-Nishina IC effects may also modify the observed appearance of reconnection-powered emission.

The basic picture of self-regulated Klein-Nishina reconnection put forward in chapter 7 and the exploration of its observable consequences in chapter 8 together establish basic expectations for how Klein-Nishina reconnection might operate and appear through a telescope, laying the foundation for future, more detailed studies. Such studies include numerical work to test basic features of the model from chapter 7 and to refine its observationally oriented predictions. I present preliminary results from such a numerical study in chapter 10. I also present in that chapter the significant modifications I made to our in-house PIC code, ZELTRON, in order to model Klein-Nishina radiative cooling and pair production.

Finally, in chapter 11, I summarize the main results from this dissertation and offer an outlook for future study.

Chapter 2

Thomson Radiative Reconnection Basics

In this chapter, I provide (section 2.1) an overview of the main radiative process of interest to this work, inverse Compton (IC) cooling, specializing, for now, to the Thomson regime (chapter 6 generalizes the discussion in this chapter to the Klein-Nishina limit). I then discuss (section 2.2) how Thomson IC cooling can cause relativistic reconnection to enter a highly radiative regime, significantly modifying the dynamics from the non-radiative case.

2.1 Single-particle Thomson IC cooling

Consider a static, homogeneous, isotropic bath of ambient radiation with spectral energy density $u_{\rm ph}(\epsilon)$. An ultrarelativistic electron or positron with energy $E = \gamma m_e c^2 \gg m_e c^2$ traversing this radiation field preferentially IC-scatters photons that appear strongly blueshifted in its rest frame: the lab-frame photon energy ϵ transforms to the rest frame energy $\epsilon' \sim \gamma\epsilon$ (except for a negligible few photons that travel within an angle $1/\gamma \ll 1$ of the particle's velocity and so transform to much smaller energies). If $\epsilon' \ll m_e c^2$, then the encounter reduces to Thomson scattering, which is approximately elastic, yielding final (scattered) rest-frame photon energy $\epsilon'_{\rm f} \simeq \epsilon'$. Moreover, the Thomson differential cross section lacks a strong angular dependence. Thus, very few photons are emitted to within angle $1/\gamma$ opposite the scattering particle's velocity, and the vast majority of scatterings yield typical final lab-frame photon energy $\epsilon_{\rm f} \sim \gamma\epsilon'_{\rm f} \sim \gamma^2\epsilon$.

The same assumptions that place us in the Thomson regime also imply that the particle loses a very small fraction of its energy in one collision: $\epsilon_{\rm f}/\gamma m_e c^2 \sim \epsilon'/m_e c^2 \sim \gamma\epsilon/m_e c^2 \ll 1$. This allows

IC radiative losses to be treated classically – as a continuous drag force, $\boldsymbol{f}_{\rm T} = -(4/3)\sigma_{\rm T}\gamma^2 U_{\rm ph}\boldsymbol{\beta}$, where $\sigma_{\rm T} = 8\pi e^4/3m_e^2c^4$ is the Thomson cross section, $U_{\rm ph} \equiv \int {\rm d}\epsilon\, u_{\rm ph}(\epsilon)$ is the total background radiation energy density, and $c\boldsymbol{\beta}$ is the scattering particle's velocity vector (cf. Blumenthal & Gould, 1970; Rybicki & Lightman, 1979; Phinney, 1982; Pozdnyakov et al., 1983; Uzdensky, 2016; Werner et al., 2019; Sironi & Beloborodov, 2020; Mehlhaff et al., 2020). Importantly, $\boldsymbol{f}_{\rm T}$ depends only on $U_{\rm ph}$ and not on the spectral distribution $u_{\rm ph}(\epsilon)$ of IC seed photons. Thus, provided a plasma radiates purely in the Thomson IC regime and is also optically thin to Thomson scattering, the collective dynamics are insensitive to the incident spectrum (though the *Comptonized* photon spectrum is not). The Thomson IC radiated power per particle is

$$P_{\rm T}(\gamma) = |c\boldsymbol{\beta}\cdot\boldsymbol{f}_{\rm T}| = \frac{4}{3}\sigma_{\rm T}c\beta^2\gamma^2 U_{\rm ph}\,. \tag{2.1}$$

All of the features discussed here – continuous emission, $\epsilon_{\rm f} \propto \gamma^2$, and incident-spectrum independence – change significantly when I later (chapter 6) allow particles to experience general Klein-Nishina Compton losses. However, in the first few chapters of this work, I focus solely on Thomson IC cooling and its effects on reconnection. Thus, I now move from this single-particle picture of IC cooling and chart out how, purely in the Thomson regime, the IC mechanism may interact with, and alter, the collective reconnection dynamics.

2.2 Thomson IC effects on collective plasma behavior

I diagnose the impact of radiative cooling on magnetic reconnection in two stages. First, I introduce a set of parameters, cast as particle energy scales (Lorentz factors), that characterize Thomson IC scattering in an explicit reconnection context. I then examine the radiative regimes represented by the possible scale hierarchies.

2.2.1 Reconnection energy scales

The reconnecting magnetic field, B_0, can be recast in terms of a length scale: the nominal relativistic gyroradius $\rho_0 \equiv m_e c^2/eB_0$. This is a useful form for comparing B_0 with other length

scales in the problem. For example, introducing the system size, L, allows one to write down a fiducial 'Hillas criterion': the energy,

$$\gamma_{\rm Hillas} \equiv \frac{eB_0 L}{m_e c^2} = \frac{L}{\rho_0}\,, \tag{2.2}$$

of a particle with Larmor radius $\gamma_{\rm Hillas}\rho_0$ equal to the system size (Hillas, 1984).[1] Equivalently, $\gamma_{\rm Hillas}$ is the energy imparted to a particle accelerated across the system by an electric field of strength B_0.

While $\gamma_{\rm Hillas}$ gives a firm upper bound on the achievable particle energies, often a more practical scale for reconnection problems is

$$\gamma_{\rm max} \equiv 0.1\gamma_{\rm Hillas}\,, \tag{2.3}$$

corresponding to extreme acceleration (Aharonian et al., 2002; Uzdensky et al., 2011; Cerutti et al., 2012a). I use $\gamma_{\rm max}$ instead of $\gamma_{\rm Hillas}$ because the reconnection electric field, $E_{\rm rec}$, is not quite as strong as B_0, but about equal to $\beta_{\rm rec}B_0$ where $\beta_{\rm rec}$ is the dimensionless reconnection rate. For collisionless reconnection, $\beta_{\rm rec}$ can be expressed in terms of the Alfvén speed $v_{\rm A}$ as $\beta_{\rm rec} \simeq 0.1 v_{\rm A}/c$. In the relativistic limit treated in this work, $v_{\rm A} \simeq c$ and hence $\beta_{\rm rec} \simeq 0.1$.

It is sometimes helpful to equivalently define $\gamma_{\rm max}$ by equating time-scales. I therefore introduce the acceleration time for a particle being linearly accelerated by the electric field $E_{\rm rec} \simeq 0.1B_0$:

$$t_{\rm X}(\gamma) \equiv \frac{\gamma m_e c^2}{ecE_{\rm rec}} = \frac{\gamma m_e c^2}{\beta_{\rm rec} ecB_0} \simeq \frac{10\gamma\rho_0}{c} \simeq 1.6\frac{2\pi\gamma\rho_0}{c}\,, \tag{2.4}$$

equal to $10/2\pi \simeq 1.6$ cyclotron periods in the magnetic field B_0. Here, subscript 'X' denotes that the spatial regions where this type of linear acceleration is active usually surround magnetic X-points – locations where the magnetic field reconnects and near which particles become unmagnetized (e.g., Uzdensky et al., 2011; Cerutti et al., 2012a). Equivalently to (2.3), one may define $\gamma_{\rm max}$ such that a particle is X-point-accelerated over one system light crossing time:

$$t_{\rm X}(\gamma_{\rm max}) \equiv \frac{L}{c}\,. \tag{2.5}$$

[1] For simplicity, I use a single length scale L to characterize the size of the system in all 3 spatial dimensions.

Another important reconnecting system parameter is the (combined electron + positron) upstream plasma density n_0. With B_0 fixed, n_0 can be cast in terms of the (dimensionless) upstream cold magnetization

$$\sigma_{\rm c,0} \equiv \frac{B_0^2}{4\pi n_0 m_e c^2}\,, \tag{2.6}$$

Physically, $\sigma_{\rm c,0}$ represents the initial magnetic energy per particle: $\sigma_{\rm c,0} m_e c^2/2$. Since reconnection delivers an appreciable fraction of the magnetic energy to the plasma, $\sigma_{\rm c,0}$ also characterizes the average Lorentz factor $\langle\gamma\rangle$ of reconnection-energized particles (absent radiative cooling). Assuming that half of the initial magnetic energy is dissipated gives $\langle\gamma\rangle \sim \sigma_{\rm c,0}/4$ (cf. Sironi et al., 2015; Werner et al., 2016; Sironi & Beloborodov, 2020). In this way, just like $\gamma_{\rm max}$ furnishes a characteristic particle energy scale to stand in for the system size L, the cold magnetization provides an energy scale that acts as a proxy for the upstream number density n_0.

I report here, for reference, one more important dimensionless parameter, the hot magnetization (cf. Melzani et al., 2014a; Werner et al., 2018)

$$\sigma_{\rm h,0} \equiv \frac{B_0^2}{4\pi w_0}\,. \tag{2.7}$$

Here, $w_0 = p_0 + u_0$ is the relativistic plasma enthalpy density with p_0 the upstream pressure and u_0 the internal energy density. For a relativistically hot upstream plasma with temperature $\theta_0 = T_0/m_e c^2 \gg 1$ the enthalpy density is $w_0 \simeq 4\theta_0 n_0 m_e c^2$ ($p_0 = \theta_0 n_0 m_e c^2 = u_0/3$); for a cold $\theta_0 \ll 1$ plasma, $w_0 \simeq u_0 \simeq n_0 m_e c^2$ is dominated by rest-mass energy. Thus, in the relativistically hot case, $\sigma_{\rm h,0} \simeq 1/(2\beta_{\rm pl}) = \sigma_{\rm c,0}/(4\theta_0) \ll \sigma_{\rm c,0}$ where $\beta_{\rm pl} = 8\pi p_0/B_0^2$ is the plasma beta parameter, but in the opposite limit, $\sigma_{\rm h,0} \simeq \sigma_{\rm c,0}$. Physically, the hot magnetization determines whether the energy flux into the reconnection region is dominated by the magnetic field ($\sigma_{\rm h,0} \gg 1$) or by the matter ($\sigma_{\rm h,0} \ll 1$). In the former $\sigma_{\rm h,0} \gg 1$ limit, the Alfvén speed, $v_{\rm A} = c\sqrt{\sigma_{\rm h,0}/(1+\sigma_{\rm h,0})}$, approaches c and, as a result, magnetic reconnection may drive not only relativistic individual particle motion (which merely requires $\sigma_{\rm c,0} \gg 1$) but also relativistic bulk flows. Thus, $\sigma_{\rm h,0} \gg 1$ is really the defining feature of 'relativistic' reconnection. In the spirit of the present discussion, one

may regard $\sigma_{\rm h,0}$ as a proxy for the upstream temperature θ_0, especially when $\theta_0 \gg 1$. However, I do not explicitly use $\sigma_{\rm h,0}$ until I first present PIC simulations in chapter 3. For now, I simply assume the relativistic limit of reconnection, $\sigma_{\rm h,0} \gg 1$, and, within that context, scope out the possible regimes by ordering the other important energy scales (which so far include $\gamma_{\rm max}$ and $\sigma_{\rm c,0}$).

Armed with $\gamma_{\rm max}$ and $\sigma_{\rm c,0}$ (and also assuming $\sigma_{\rm h,0} \gg 1$), I can describe the size of the relativistic reconnection system. A 'large' system satisfies $\gamma_{\rm max} \gg \sigma_{\rm c,0}$. In terms of the global geometry, this implies a high aspect ratio: the system is much longer than the microscopic current sheet thickness, of order $\langle\gamma\rangle\rho_0$ – the typical Larmor radius of reconnection-energized particles (specifically, $\gamma_{\rm max} \gg \sigma_{\rm c,0} \Rightarrow L \gg 10\sigma_{\rm c,0}\rho_0$; cf. Werner et al. 2016). This renders the layer plasmoid-unstable and initiates plasmoid-dominated reconnection (e.g., Ji & Daughton, 2011).

In terms of individual particles, $\gamma_{\rm max} \gg \sigma_{\rm c,0}$ also alleviates system-size constraints on particle energization, at least up to the mean Lorentz factor $\langle\gamma\rangle \sim \sigma_{\rm c,0}/4$. In addition, when $1 \ll \theta_0 \ll \langle\gamma\rangle \sim \sigma_{\rm c,0}/4$ and, hence, $\sigma_{\rm h,0} \gg 1$, previous non-radiative 2D particle-in-cell (PIC) simulations have found that direct/fast acceleration by the reconnection electric field saturates at $\gamma_{\rm X} \sim$ (several)$\sigma_{\rm c,0}$ (e.g., $4\sigma_{\rm c,0}$; Werner et al., 2016; Kagan et al., 2018). Thus, when $\gamma_{\rm max} \gg \sigma_{\rm c,0}$, both the physics governing average-energy particle acceleration and the direct acceleration mechanism are unencumbered by the system size.

However, I remark that there are, in addition to primary X-point acceleration, other energization channels in the large-system plasmoid-dominated regime: most of them take place on slower time-scales but are not limited to Lorentz factors of order several $\sigma_{\rm c,0}$. I term these 'secondary' acceleration processes because they typically operate on plasma that has already been processed into the region of reconnected magnetic flux. One example of such a process is adiabatic heating by slowly compressing reconnected magnetic fields inside plasmoids (Petropoulou & Sironi, 2018; Hakobyan et al., 2021). A different, but related, example is a Fermi-type mechanism (Drake et al., 2006; Dahlin et al., 2014; Guo et al., 2014, 2015, 2016, 2019, 2020) in which particles bounce from end to end across contracting plasmoids. (See Uzdensky 2020 for a review of secondary acceleration mechanisms in 2D reconnection.) The existence of such acceleration channels may, without radia-

tive cooling, allow the highest particle energies to grow well beyond $\gamma_{\rm X}$ – even if the dynamics of average particles and X-point acceleration top out at much lower energies. This is one aspect where even weakly radiative reconnection differs from its non-radiative counterpart. Radiative losses can impose a system-size-independent high-energy cutoff even on acceleration processes with no intrinsic upper energy limit (e.g., Hakobyan et al., 2021; Mehlhaff et al., 2020), potentially allowing the spectrum of accelerated particles to become independent of L.

I now introduce the main energy scale associated with Thomson IC radiative cooling in the context of reconnection. As discussed above, particles emitting in the Thomson limit suffer a drag force $\boldsymbol{f}_{\rm IC}$ determined by the total background radiation energy density $U_{\rm ph}$. Thus, Thomson radiation introduces just one extra parameter, $U_{\rm ph}$, into the reconnection problem. To cast $U_{\rm ph}$ as an energy scale, I define the Lorentz factor, $\gamma_{\rm rad,T}$, at which radiative drag matches the acceleration force from the reconnection electric field, $eE_{\rm rec} \simeq 0.1eB_0$, or (equivalently) such that the X-point acceleration time, $t_{\rm X}(\gamma)$, equals the Thomson cooling time, $t_{\rm cool,T}(\gamma) \equiv \gamma m_e c^2/P_{\rm T}(\gamma)$, over which a particle radiates a significant fraction of its energy γ. Putting $f_{\rm IC}(\gamma_{\rm rad,T}) \equiv 0.1eB_0$ [or $t_{\rm X}(\gamma_{\rm rad,T}) \equiv t_{\rm cool,T}(\gamma_{\rm rad,T})$] yields (cf. Uzdensky, 2016; Nalewajko, 2016; Werner et al., 2019; Sironi & Beloborodov, 2020; Mehlhaff et al., 2020)

$$\gamma_{\rm rad,T} \equiv \sqrt{\frac{0.3eB_0}{4\sigma_{\rm T}U_{\rm ph}}}\,. \tag{2.8}$$

There is also a third way to define $\gamma_{\rm rad,T}$: it is the Lorentz factor of a particle that cools in ~ 1 cyclotron period in the magnetic field B_0 (cf. Uzdensky, 2016).

Despite the fact that it is defined only by balancing radiative cooling against X-point energization, the energy $\gamma_{\rm rad,T}$ firmly radiatively caps the achievable particle energies (Werner et al., 2019; Mehlhaff et al., 2020). This is because acceleration by the coherent reconnection electric field may be the fastest significant particle energization channel in magnetic reconnection.[2] Other channels (such as those discussed by Petropoulou & Sironi, 2018; Guo et al., 2019; Hakobyan et al.,

[2] It is unclear whether motional electric fields $\boldsymbol{E}_{\rm motion} = -\boldsymbol{v}\times\boldsymbol{B}/c$ ($v \simeq v_{\rm A} \simeq c$) – e.g., due to rapidly moving compact plasmoid cores where $B > B_0$ and, hence, $E_{\rm motion} \sim v_{\rm A}B/c > B_0$ – yield an overall faster effective acceleration than X-point regions. A similar remark applies to momentary pulses of high electric fields associated with waves launched at plasmoid mergers (Philippov et al., 2019). In both cases, the issue is not just one of field strength but also of spatio-temporal coherence.

2021; Guo et al., 2020), while not saturating at $\gamma_{\rm X}$ like X-point acceleration, are much slower and hence radiatively stall at Lorentz factors less than $\gamma_{\rm rad,T}$ (e.g., Mehlhaff et al., 2020).

To sum up, I now have three energy scales, each characterizing different physical parameters in relativistic reconnection. Two of these, $\gamma_{\rm max}$ and $\sigma_{\rm c,0}$ – representing, respectively, the system size L and the upstream particle density n_0 – are non-radiative, common to all reconnection problems. The third scale, $\gamma_{\rm rad,T}$, encodes the energy density of ambient radiation and is unique to reconnection with IC cooling. Furthermore, one of these energies, $\sigma_{\rm c,0}$, splits into two: it represents both the average energy of reconnection-energized particles $\langle\gamma\rangle \sim \sigma_{\rm c,0}/4$ and the intrinsic maximum energy deliverable by the X-point acceleration mechanism $\gamma_{\rm X} \sim 4\sigma_{\rm c,0}$ in the large-system, plasmoid-mediated regime. I later argue (chapters 3 and 8) that, under some circumstances, $\gamma_{\rm X}$ may exceed the nominal value $\gamma_{\rm X} = 4\sigma_{\rm c,0}$ found from 2D PIC simulations (Werner et al., 2016; Kagan et al., 2018), but in the case where $\gamma_{\rm X}$ truly is of order $4\sigma_{\rm c,0}$, $\gamma_{\rm X}$ and $\langle\gamma\rangle$ are only offset by about a factor of 16.

2.2.2 Regimes of Thomson radiative reconnection

Next, I enumerate the various orderings of these scales and examine the physical regimes each ordering represents. To simplify this program, I concentrate solely on the large-system regime with a relativistic amount of magnetic energy per particle: $\gamma_{\rm max} \gg \sigma_{\rm c,0} \gg 1$ (or, splitting the scale $\sigma_{\rm c,0}$ into $\gamma_{\rm X} \sim 4\sigma_{\rm c,0}$ and $\langle\gamma\rangle \sim \sigma_{\rm c,0}/4$, the regime $\gamma_{\rm max} \gg \gamma_{\rm X} \gg \langle\gamma\rangle \gg 1$). Different scale hierarchies are then realized by inserting $\gamma_{\rm rad,T}$ into various positions of the base ordering $\gamma_{\rm max} \gg \gamma_{\rm X} \gg \langle\gamma\rangle \gg 1$. The possible orderings are summarized in Table 2.1.

This procedure is made more conceptually transparent if I introduce the derived scale $\gamma_{\rm cool}$, the Lorentz factor of a particle that cools in one dynamical time L/c of the system. Writing $t_{\rm cool,T}(\gamma_{\rm cool}) \equiv L/c$ yields

$$\gamma_{\rm cool} \equiv \gamma_{\rm rad,T}^2/\gamma_{\rm max}\,. \tag{2.9}$$

Interestingly, the radiatively limited Lorentz factor $\gamma_{\rm rad,T}$ is always intermediate between $\gamma_{\rm cool}$

Table 2.1: Scale hierarchies and associated radiative regimes of relativistic reconnection subject to Thomson IC losses. A relativistic large-system ordering $\gamma_{\rm max} > \gamma_{\rm X} > \langle\gamma\rangle > 1$ is assumed throughout. Inserting $\gamma_{\rm rad,T}$ and $\gamma_{\rm cool}$ at different locations in this base ordering corresponds to different regimes. To guide the eye, the 'independent' parameters (i.e., independent of $\gamma_{\rm max}$, $\gamma_{\rm X}$, and $\langle\gamma\rangle$ but not of each other) $\gamma_{\rm rad,T}$ and $\gamma_{\rm cool}$ are typeset in red. The regimes are semantically distinguished based on which particles – either those in the bulk of the particle energy distribution near the average energy $\langle\gamma\rangle$ or those in the high-energy tail – cool on time-scales shorter than L/c (strong cooling) versus those that cool in one gyroperiod (saturated cooling). For Thomson IC radiation reaction (but not in the more general Klein-Nishina limit) saturated cooling is more efficient than, and therefore implies, strong cooling.

		Bulk particles		High-energy particles	
Scale hierarchy	Regime name	Strong cooling (Y/N)	Saturated cooling (Y/N)	Strong cooling (Y/N)	Saturated cooling (Y/N)
$\gamma_{\rm cool} > \gamma_{\rm rad,T} > \gamma_{\rm max} > \gamma_{\rm X} > \langle\gamma\rangle$	Non-radiative	N	N	N	N
$\gamma_{\rm max} > \gamma_{\rm rad,T} > \gamma_{\rm cool} > \gamma_{\rm X} > \langle\gamma\rangle$	Quasi non-radiative	N	N	Y*	N
$\gamma_{\rm max} > \gamma_{\rm rad,T} > \gamma_{\rm X} > \gamma_{\rm cool} > \langle\gamma\rangle$		N	N	Y	N
$\gamma_{\rm max} > \gamma_{\rm rad,T} > \gamma_{\rm X} > \langle\gamma\rangle > \gamma_{\rm cool}$		Y	N	Y	N
$\gamma_{\rm max} > \gamma_{\rm X} > \gamma_{\rm rad,T} > \gamma_{\rm cool} > \langle\gamma\rangle$		N	N	Y	Y
$\gamma_{\rm max} > \gamma_{\rm X} > \gamma_{\rm rad,T} > \langle\gamma\rangle > \gamma_{\rm cool}$		Y	N	Y	Y
$\gamma_{\rm max} > \gamma_{\rm X} > \langle\gamma\rangle > \gamma_{\rm rad,T} > \gamma_{\rm cool}$	Extremely radiative	Y	Y	Y	Y

*Highest-energy particles may or may not achieve strong cooling. If they do, it is not through impulsive X-point acceleration (see text).

and $\gamma_{\rm max}$, equal to the geometric mean of those two scales. Note that one may have $\gamma_{\rm cool} < 1$, in which case $\gamma_{\rm cool}$ does not correspond to a physical Lorentz factor. In that case, all particles cool to non-relativistic energies in $t_{\rm cool,T}(1) < L/c$. Related to $\gamma_{\rm cool}$ is the compactness of the system

$$\ell \equiv \frac{U_{\rm ph}\sigma_{\rm T}L}{m_e c^2} = \frac{3}{4\gamma_{\rm cool}}\,. \tag{2.10}$$

The time for a particle to cool from any initial γ to $\gamma = 1$ is $\sim L/c\ell$.

Let me begin exploring the various radiative regimes by quantifying the *non-radiative* limit $\gamma_{\rm cool} > \gamma_{\rm rad,T} > \gamma_{\rm max} \gg \sigma_{\rm c,0} \gg 1$. Hereafter, I do not list $\sigma_{\rm c,0} \gg 1$ explicitly, and I only use '>' (not '≫') symbols. One should bear in mind that all regimes become more distinct when the corresponding scales are well-separated. The regime $\gamma_{\rm cool} > \gamma_{\rm rad,T} > \gamma_{\rm max} > \sigma_{\rm c,0}$ corresponds to the limit $U_{\rm ph} \to 0$ and, hence, $\gamma_{\rm rad,T} \to \infty$. Here, no particle radiates a significant fraction of its energy within one dynamical time L/c, effectively decoupling radiation from reconnection. This is the regime mentioned in the Introduction chapter where magnetic reconnection can, in principle, be studied on its own and the radiative signatures calculated independently.

The first step up in radiative efficiency might be called the *quasi non-radiative regime* $\gamma_{\rm max} > \gamma_{\rm rad,T} > \gamma_{\rm cool} > \gamma_{\rm X} > \langle\gamma\rangle$. Here, primary X-point acceleration does not impart enough energy to particles so that they significantly radiate on one dynamical time L/c, let alone so that they achieve radiative saturation $\gamma_{\rm rad,T}$. Secondary acceleration channels, on the other hand, might be able to deliver particles to energies $\geq \gamma_{\rm cool}$ so that those particles radiate faster than the global time-scale. However, this depends on the detailed nature of each secondary acceleration process – whether any of them radiatively stall above $\gamma_{\rm cool}$ is not guaranteed.

As an example of a secondary acceleration mechanism relevant to the quasi non-radiative regime, one may consider particles slowly energized inside of adiabatically compressing magnetic islands (also 'plasmoids'), as detailed by Petropoulou & Sironi (2018) and Hakobyan et al. (2021). The energy where radiative losses shut this process down is determined by matching the plasmoid compression time to the particle cooling time-scale. As reported by Hakobyan et al. (2021), this upper-limit energy is $\gamma_{\rm sec} \sim m_e c^2 \beta_{\rm rec}/w\sigma_{\rm T}U_{\rm ph}$, where $w \simeq 0.1L$ is the size of the largest plasmoids

formed by reconnection. [I ignore that smaller plasmoids compress faster and so may yield, for the smaller number of particles they contain, higher $\gamma_{\rm sec}$ (Hakobyan et al., 2021).] In effect, $\gamma_{\rm sec} \sim \ell^{-1} = (4/3)\gamma_{\rm cool}$. Thus, this secondary energization channel only barely accelerates some particles up to $\gamma_{\rm cool}$, and most reach energies much less than this. Hence, considering only this secondary mechanism, the reconnection process can be regarded as marginally non-radiative, with only very few highest-energy particles cooling in less than one dynamical time.

Increasing the cooling efficiency once more brings us to the first of several truly *radiative* regimes where radiation is dynamically important for at least some of the particles. Aptly naming these regimes is cumbersome because different particles can experience varying degrees of radiative efficiency. The high-energy particles, for example, may be rapidly cooled and the particles at the average energy $\langle\gamma\rangle$ cooled quite slowly. Therefore, I do not classify these regimes globally, calling the entire system weakly or strongly radiative, but I refer to them based on which populations of particles cool on various time-scales.

I call particles *strongly cooled* if radiation reaction causes them to lose an appreciable fraction of energy in less than one dynamical time L/c – i.e., if their Lorentz factors exceed $\gamma_{\rm cool}$. This agrees with typical notions of strong cooling in astrophysics, which indicate that radiative cooling occurs faster than some macroscopic system time-scale. Correspondingly, I call particles with $\gamma < \gamma_{\rm cool}$ *weakly cooled* (because they are not strongly cooled) and sometimes *non-radiative* (because they do not radiate appreciably in a dynamical time). The particles with much higher energies, close to $\gamma_{\rm rad,T} > \gamma_{\rm cool}$, I say exhibit *saturated cooling*: their Lorentz factors are radiatively saturated because intense emission prevents further energization ($t_{\rm cool,T} \lesssim t_{\rm X}$). Although, in the Thomson IC limit, particles undergoing saturated cooling radiate much more efficiently than just strongly cooled particles, this is not always true once Klein-Nishina effects come into play (chapter 6). Thus, I wish to avoid associating the saturated cooling regime with a term connoting excessively efficient or fast cooling (e.g., 'very strong cooling').

Using these terms, one sees that the scale hierarchy $\gamma_{\rm max} > \gamma_{\rm rad,T} > \gamma_{\rm X} > \gamma_{\rm cool} > \langle\gamma\rangle$ indicates that average (or 'bulk') particles are weakly radiative; most of them cool slower than L/c ($\langle\gamma\rangle <$

$\gamma_{\rm cool}$). Meanwhile, at least some of the high-energy particles accelerated by the primary X-point channel are strongly cooled ($\gamma_{\rm X} > \gamma_{\rm cool}$). Even so, radiative losses are not so fast as to hinder direct X-point acceleration, with $t_{\rm X}(\gamma_{\rm X})$ faster than $t_{\rm cool,T}(\gamma_{\rm X})$ because $\gamma_{\rm X} < \gamma_{\rm rad,T}$. Thus, X-point acceleration (because it is intrinsically capped to below $\gamma_{\rm rad,T}$) – and, hence, all other (known) secondary energization channels (because they are slow) – cannot deliver particles up to the radiative saturation limit $\gamma_{\rm rad,T}$.

Permuting scales again by swapping the positions of $\gamma_{\rm cool}$ and $\langle\gamma\rangle$, one arrives at the regime $\gamma_{\rm max} > \gamma_{\rm rad,T} > \gamma_{\rm X} > \langle\gamma\rangle > \gamma_{\rm cool}$. Here, most particles radiate strongly because $\gamma_{\rm cool} < \langle\gamma\rangle < \gamma_{\rm X}$. However, like in the previous regime, virtually no particles are expected to achieve radiative saturation ($\gamma_{\rm X} < \gamma_{\rm rad,T}$), and X-point acceleration, while unaffected by radiative cooling on the short time-scale on which it occurs [$t_{\rm X}(\gamma_{\rm X}) < t_{\rm cool,T}(\gamma_{\rm X})$], does produce particles of sufficiently high energies ($> \gamma_{\rm cool}$) to be strongly cooled.

Next come the scale hierarchies where at least a few particles exhibit saturated cooling. One such domain is $\gamma_{\rm max} > \gamma_{\rm X} > \gamma_{\rm rad,T} > \gamma_{\rm cool} > \langle\gamma\rangle$. Here, some particles are promptly accelerated near X-points to the upper-limit energy $\gamma_{\rm rad,T} < \gamma_{\rm X}$, but the bulk particles, with $\langle\gamma\rangle < \gamma_{\rm cool}$, barely radiate even on global L/c time-scales. This is perhaps the most extreme example of how vastly different the cooling rates can be for different reconnection-energized particles. However, unless $\gamma_{\rm X}$ can substantially exceed its nominal $4\sigma_{\rm c,0}$ value, this regime may not be realized in astrophysical contexts. This is because, if $\gamma_{\rm X} = 4\sigma_{\rm c,0}$, then $\gamma_{\rm X}/\langle\gamma\rangle \simeq 16$, implying that $\gamma_{\rm rad,T}/\gamma_{\rm cool} < 16$, and, through equation (2.9), that $\gamma_{\rm max}/\gamma_{\rm rad,T} < 16$, whereas $\gamma_{\rm max}$, $\gamma_{\rm rad,T}$, and $\gamma_{\rm cool}$ are each usually separated by several decades in astrophysical systems (e.g., chapter 8).

Moving to the last two possible orderings, one has $\gamma_{\rm max} > \gamma_{\rm X} > \gamma_{\rm rad,T} > \langle\gamma\rangle > \gamma_{\rm cool}$, in which the high-energy particles accelerated near X-points attain radiative saturation ($\gamma_{\rm rad,T} < \gamma_{\rm X}$) and the bulk particles are strongly cooled ($\gamma_{\rm cool} < \langle\gamma\rangle$). Finally, there is an *extremely radiative* regime, $\gamma_{\rm max} > \gamma_{\rm X} > \langle\gamma\rangle > \gamma_{\rm rad,T} > \gamma_{\rm cool}$. Here, IC losses firmly cap the acceleration of nearly all particles – not even the formal mean energy, $\langle\gamma\rangle$, available per particle can be attained – and should have dramatic effects on the large-scale reconnection dynamics (Uzdensky, 2016). Table 2.1

summarizes the radiative reconnection regimes discussed in this section.

To complete this tour of the Thomson IC reconnection landscape, I review some of the previously identified physical effects that occur in these regimes. Several systematic PIC studies of Thomson radiative reconnection have been conducted in recent years, including those by Werner et al. (2019), Sironi & Beloborodov (2020), and Mehlhaff et al. (2020). There have also been radiative PIC studies of reconnection with strong synchrotron cooling (e.g., Cerutti et al., 2013, 2014a,b; Yuan et al., 2016) and QED effects (like pair-production; Schoeffler et al., 2019; Hakobyan et al., 2019). In some cases, the qualitative features of reconnection regimes mediated by different radiative processes are similar, but, in this chapter, I focus only on those effects studied within the context of Thomson IC losses.

The three radiative PIC studies conducted by Werner et al. (2019), Sironi & Beloborodov (2020), and Mehlhaff et al. (2020) explored a number of regimes outlined in this section. Numerically exploring the scenario first outlined by Beloborodov (2017), Sironi & Beloborodov (2020) focused primarily on the regime $\gamma_{\rm max} > \gamma_{\rm X} > \gamma_{\rm rad,T} > \langle\gamma\rangle > \gamma_{\rm cool}$ where the bulk particles are strongly cooled and the highest-energy particles saturate at $\gamma = \gamma_{\rm rad,T}$. They found that, here, plasmoids are generally filled with cold plasma that has already released much of its energy through Compton losses. The plasma kinetic energy inside plasmoids is then dominated by bulk motion. This motion is also subject to radiative drag and, hence, is slower than in the non-radiative case. Sironi & Beloborodov (2020) also confirmed that the highest-energy particles are accelerated near reconnection X-points and top out at Lorentz factors close to $\gamma_{\rm rad,T}$. Werner et al. (2019) studied the effect of Compton losses on large-scale reconnection dynamics and on non-thermal particle acceleration, exploring all the way from the non-radiative regime to that of fully saturated high-energy cooling and strong bulk cooling ($\gamma_{\rm max} > \gamma_{\rm X} > \gamma_{\rm rad,T} > \langle\gamma\rangle > \gamma_{\rm cool}$). They found that radiation steepens the high-energy part of the non-thermal tail of reconnection-accelerated particles but that the overall reconnection rate is virtually unaffected. Mehlhaff et al. (2020) focused on the angular distributions of high-energy particles, showing that particles approaching radiative saturation ($\gamma_{\rm rad,T} \lesssim \gamma_{\rm X}$) exhibit energy-dependent collimation in momentum space, forming nar-

row beams at the highest energies – the *kinetic beaming* effect (see the Introduction). In the next chapter, I detail the analysis of Mehlhaff et al. (2020), demonstrating from PIC simulations that strong radiative losses appear to be an essential ingredient in mediating observable signatures of kinetic beaming.

Chapter 3

A Numerical Study of Kinetic Beaming in Thomson Radiative Reconnection

In this chapter, I present particle-in-cell (PIC) simulations of reconnection with Thomson IC cooling. The main series of simulations scans across the IC radiative cooling strength ($\gamma_{\rm rad,T}$), progressing through several of the regimes outlined in the previous chapter (e.g., Table 2.1). In analyzing the simulations, I concentrate on determining whether the particles accelerated through reconnection (and, consequently, whether the photons that they emit) are kinetically beamed. The main result of this chapter is that kinetic beaming requires a high degree of radiative efficiency in order to manifest itself within the reconnection-yielded distributions of high-energy particles and photons. Together, this chapter and the next detail work first published by Mehlhaff et al. (2020).

3.1 Simulations

In this initial section, I begin by describing the PIC reconnection simulations, including how they self-consistently incorporate IC cooling. Then, I overview the time evolution of these simulations. Afterwards, I introduce the basic units of angular information – energy-resolved *angular maps* of the directional distributions of particle momenta – that form the basis for the quantitative beaming metrics developed in the following section (3.2).

3.1.1 Set-up

I present relativistic pair-plasma simulations run using the radiative electromagnetic PIC code ZELTRON (Cerutti et al., 2013). The simulation domain is a 2D box of size $L_x \times L_y$ with $L_y =$

$2L_x = 2L$ and periodic boundary conditions enforced in both dimensions. Spatial dependence is limited to the (x, y) dimensions, but vectorial quantities, including velocities and field components, are fully three-dimensional.

I initialize the simulations with standard double Harris current sheets (Kirk & Skjæraasen, 2003) of half-thickness δ carrying anti-aligned currents in the $\pm z$-directions and centered on the planes $y_1 = L_y/4$ and $y_2 = 3L_y/4$. This double-sheet configuration is chosen because it is consistent with the periodic box boundaries. Namely, the currents establish an in-plane magnetic field $B_x(y) = \pm B_0 \tanh\left[(y - y_{1,2})/\delta\right]$ that reverses twice – once at each sheet – and is therefore periodic in y. Additionally, I add a uniform initial guide field $\boldsymbol{B}_{\rm g} = (B_0/4)\hat{\boldsymbol{z}}$. This serves chiefly as a numerical device to support magnetic islands against radiative cooling-induced contraction to the point where the Debye length becomes unresolved. The value $B_{\rm g} = B_0/4$ does not substantially alter non-thermal particle acceleration (NTPA; Werner & Uzdensky, 2017; Werner et al., 2019).

The current in the Harris layers is carried by a drifting plasma component. In each layer, drifting electrons and positrons begin the simulation counterstreaming with bulk velocities $\pm c\beta_{\rm d}\hat{\boldsymbol{z}}$ and a combined drifting lab-frame number density $n_{\rm d}(y) = n_{\rm d,0}\,{\rm sech}^2\left[(y - y_{1,2})/\delta\right]$. The resultant current profile $J_z(y) = \pm c\beta_{\rm d} e n_{\rm d}(y)$ is precisely that necessary to generate the field $B_x(y)$ as dictated by Ampère's Law. In addition to the current-governing mean velocity $\beta_{\rm d}$, I initialize the counterstreaming species with relativistic comoving temperature $\theta_{\rm d} = T_{\rm d}/m_{\rm e}c^2 = 1050 \gg 1$ to support the current layers against the upstream magnetic pressure.

With the Harris equilibrium satisfied by the drifting particles, each simulation also contains an initially stationary uniform background plasma of combined density n_0 and relativistically hot temperature $\theta_0 = T_0/m_{\rm e}c^2 = 25 \gg 1$ that provides the inflow material for reconnection. As discussed in chapter 2, n_0 and θ_0 can be recast in terms of the cold and hot magnetizations, $\sigma_{\rm c,0} \equiv B_0^2/4\pi n_0 m_e c^2$ and $\sigma_{\rm h,0} \equiv B_0^2/4\pi w_0$ respectively, where the enthalpy density is $w_0 \simeq 4n_0\theta_{\rm b}m_e c^2$ because $\theta_0 \gg 1$. I set $\sigma_{\rm c,0} = 10^4$, allowing individual particles to be accelerated to relativistic energies ($\langle\gamma\rangle \sim \sigma_{\rm c,0}/4 = 2.5 \times 10^3$). I also take $\sigma_{\rm h} = 10^2$, operating in the magnetically dominant limit, which also implies $v_{\rm A} \simeq c$ and places reconnection in the relativistic regime.

To minimize system-size effects, the simulation box must considerably exceed the largest kinetic scale in the problem: the typical Larmor radius of energetic particles (see Werner et al., 2016). Because the liberated magnetic energy per upstream particle is $\langle\gamma\rangle m_e c^2 \sim \sigma_{\rm c,0} m_e c^2/4$, an average particle accelerated through the reconnection layer emerges with Larmor radius $\langle\gamma\rangle\rho_0$ where $\rho_0 = m_{\rm e}c^2/eB_0$ is a nominal Larmor radius. I conduct simulations in the large system regime $\gamma_{\rm max} \gg \langle\gamma\rangle$ (i.e., $L_x \gg 10\langle\gamma\rangle\rho_0$; cf. chapter 2 and Werner et al., 2016) and set $\gamma_{\rm max} = 32\sigma_{\rm c,0} \sim 128\langle\gamma\rangle$ (implying $L_x/\sigma_{\rm c,0}\rho_0 = 320$). To confirm the insensitivity of my results to the simulation box size, I also run a series of simulations at different L_x.

At the small-scale end, the cell size is $\Delta x = \Delta y = \sigma\rho_0/24$, which is just smaller than the Debye length: $\lambda_{\rm D} = 1.2\Delta x$. With the cell size set, I employ a corresponding time-step satisfying the Courant–Friedrichs–Lewy condition $c\Delta t = 0.7\Delta x < \Delta x/\sqrt{2}$. The initial number of simulation particles per grid cell is 80.

I incorporate radiative cooling into the simulations via inverse Compton scattering of a background (seed) radiation field that is static, homogeneous, and isotropic. The photons comprising this field are not tracked simulation entities, but, in the Thomson limit, give rise to a continuous radiative drag force $\boldsymbol{f}_{\rm T}$ that enters self-consistently into the PIC particle push (Tamburini et al., 2010). As detailed in chapter 2, a particle of 4-velocity $(c\gamma, c\gamma\boldsymbol{\beta})$ suffers the drag force $\boldsymbol{f}_{\rm T} = -(4/3)\sigma_{\rm T}U_{\rm ph}\gamma^2\boldsymbol{\beta}$ where $\sigma_{\rm T} = 8\pi e^4/3m_{\rm e}^2c^4$ is the Thomson cross section and $U_{\rm ph} \equiv \int d\epsilon\, u(\epsilon)$ is the total energy density of background photons. Because the drag force depends only on the total energy density $U_{\rm ph}$, I adopt, without loss of generality, a simple monochromatic spectral energy density $u(\epsilon) = U_{\rm ph}\delta(\epsilon - \epsilon_{\rm ph})$. The Thomson limit is satisfied if the photon energies encountered by a particle of Lorentz factor γ in its rest frame are small compared with its mass:

$$\frac{\gamma\epsilon_{\rm ph}}{m_{\rm e}c^2} \ll 1 \tag{3.1}$$

(see chapter 2). This condition allows the recoil in any single scattering event to be neglected and justifies a continuous treatment of radiative losses (Blumenthal & Gould, 1970). When equation (3.1) is not satisfied, IC cooling transitions to the discrete Klein–Nishina regime, with particles

delivering an order-unity fraction of their energies to single photons. I discuss the more general Klein-Nishina case in chapter 6.

In this chapter, I specify $U_{\rm ph}$ through the particle Lorentz factor, $\gamma_{\rm rad,T} \equiv (0.3eB_0/4\sigma_{\rm T}U_{\rm ph})^{1/2}$ [equation (2.8)], at which radiative losses match the expected acceleration force from the reconnection electric field. My main results come from a series of simulations scanning across $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = [1/2, 1, 2, 4, 6, 8, 16, 32, 64, \infty]$ at the fiducial box size $L_x/\sigma\rho_0 = 320$. Recall from chapter 2 that *higher* $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$ corresponds to *weaker* IC cooling. In particular, the case $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = \infty$ implies $U_{\rm ph} = 0$: no cooling. For each respective value of $\gamma_{\rm rad,T}$, the corresponding Lorentz factor $\gamma_{\rm cool} = \gamma_{\rm rad,T}^2/\gamma_{\rm max}$ from equation (2.9) is $\gamma_{\rm cool}/\sigma_{\rm c,0} = [1/128, 1/32, 1/8, 1/2, 9/8, 2, 8, 32, 128, \infty]$. Adopting $\langle\gamma\rangle = \sigma_{\rm c,0}/4$ and an intrinsic X-point acceleration saturation Lorentz factor $\gamma_{\rm X} = 4\sigma_{\rm c,0}$ allows one to examine what regimes, in the sense of chapter 2 (section 2.2.2), are covered by the simulations:

(1) For $\gamma_{\rm rad,T} = 1/2, 1,$ and 2: the high-energy particles reach radiative saturation and the bulk particles are strongly cooled, $\gamma_{\rm max} > \gamma_{\rm X} > \gamma_{\rm rad,T} > \langle\gamma\rangle > \gamma_{\rm cool}$. This scale hierarchy corresponds to the 6th row of Table 2.1.

(2) For $\gamma_{\rm rad,T} = 4, 6,$ and 8: the high-energy particles no longer reach radiative saturation (except, perhaps, for $\gamma_{\rm rad,T} = 4\sigma_{\rm c,0} = \gamma_{\rm X}$), but both the high-energy and the bulk particles are strongly cooled: $\gamma_{\rm max} > \gamma_{\rm rad,T} \geq \gamma_{\rm X} > \langle\gamma\rangle > \gamma_{\rm cool}$. This scale hierarchy is the 4th row of Table 2.1.

(3) For $\gamma_{\rm rad,T} = 16$: the high-energy particles are still strongly cooled, but the bulk particles are not (and no particles reach radiative saturation): $\gamma_{\rm max} > \gamma_{\rm rad,T} > \gamma_{\rm X} > \gamma_{\rm cool} > \langle\gamma\rangle$. This is the 3rd row of Table 2.1.

(4) For $\gamma_{\rm rad,T} = 32, 64,$ and ∞: these simulations are non-radiative, satisfying $\gamma_{\rm cool} \geq \gamma_{\rm rad,T} \geq \gamma_{\rm max} > \gamma_{\rm X} > \langle\gamma\rangle$. This is the 1st row of Table 2.1.

Thus, this series of simulations explores most of the potential regimes of Thomson radiative recon-

nection.

In addition to scanning across $\gamma_{\rm rad,T}$, I run a few simulations with a uniform radiative efficiency $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 4$ but differing $L_x/\sigma\rho_0 \in [80, 120, 160, 240, 320]$ as a first step towards characterizing the system-size dependence of my findings. Table 3.1 summarizes these values and those of the other parameters discussed so far.

The wide-ranging scan in $\gamma_{\rm rad,T}$ discussed above is limited on the strong-cooling end by radiative losses in the upstream region, which can cause the background plasma feeding reconnection to change in time. In order to avoid this effect, one should require that the upstream IC cooling time, $t_{\rm cool,T}$, exceed the duration, $3L_{\rm x}/c$, of the simulations. Using the average thermal Lorentz factor, $3\theta_0$, of the background plasma, $t_{\rm cool,T}/(3L_x/c)$ can be written as

$$\begin{aligned}\frac{t_{\rm cool,T}}{3L_x/c} &= \frac{3\theta_0 m_e c^2/P_{\rm T}(3\theta_0)}{3L_x/c} = \frac{1}{3}\frac{\sigma_{\rm c,0}}{\gamma_{\rm max}}\frac{\gamma_{\rm rad,T}}{3\theta_0}\frac{\gamma_{\rm rad,T}}{\sigma_{\rm c,0}} \\ &= \frac{4\sigma_{\rm h}}{9}\left(\frac{\sigma_{\rm c,0}}{\gamma_{\rm max}}\right)\left(\frac{\gamma_{\rm rad,T}}{\sigma_{\rm c,0}}\right)^2 \\ &\simeq 1.4\left(\frac{\sigma_{\rm h}}{100}\right)\left(\frac{\gamma_{\rm max}/\sigma_{\rm c,0}}{32}\right)^{-1}\left(\frac{\gamma_{\rm rad,T}}{\sigma_{\rm c,0}}\right)^2 . \end{aligned} \tag{3.2}$$

Here, I used $P_{\rm T}(\gamma) = |c\boldsymbol{\beta}\cdot\boldsymbol{f}_{\rm T}| = (4/3)c\sigma_{\rm T}U_{\rm ph}\gamma^2\beta^2$ and the relativistic limits $\beta \simeq 1$ and $1 \ll \theta_0 = \sigma_{\rm c,0}/4\sigma_{\rm h}$. Let me go one step farther, employing equation (3.2) along with $\mathrm{d}\gamma/\,\mathrm{d}t = -P_{\rm IC}(\gamma)/m_e c^2$ to estimate the amount by which the upstream plasma cools during a simulation. The temperature, $\theta_{\rm f}$, reached by the upstream plasma at $t = 3L_{\rm x}/c$ may be as low as

$$\frac{\theta_{\rm f}}{\theta_0} \simeq \frac{1}{1 + (3L_{\rm x}/c)/t_{\rm cool,T}} . \tag{3.3}$$

Evidently, the simulation with $\gamma_{\rm rad,T} = \sigma_{\rm c,0}/2$ is problematic, with the upstream plasma cooling in time $t_{\rm cool} \sim 1L_x/c$ and potentially falling to $\simeq 26$ per cent of its initial temperature by the end of the run. This simulation also exhibits the worst energy conservation, with the energy error peaking at about 3.6 per cent (all of the other simulations have per cent level or better error). Thus, the results of this most strongly radiative run should not be taken as definitive on their own. Fortunately, the conclusions I draw from the $\gamma_{\rm rad,T}$ scan (Section 3.3.3, Figs 3.19, 3.20,

Table 3.1: Simulation parameters used in this study. Note that I do not scan across all combinations of L_x and $\gamma_{\rm rad,T}$. Instead, I conduct a series of simulations exploring all $\gamma_{\rm rad,T}$ values in the table at a fixed system size $L_x = 320\sigma_{\rm c,0}\rho_0$ and a second series across all L_x values at a fixed radiation-reaction strength $\gamma_{\rm rad,T} = 4\sigma_{\rm c,0}$.

Parameter	Symbol (=definition)	Value
Upstream magnetic field	B_0	
Nominal gyroradius	$\rho_0 = m_e c^2/eB_0$	
Radiation-limited Lorentz factor	$\gamma_{\rm rad,T}$	$(1/2, 1, 2, 4, 6, 8, 16, 32, 64, \infty) \times \sigma_{\rm c,0}$
System size	$L_x = L$	$(80, 120, 160, 240, 320) \times \sigma_{\rm c,0}\rho_0$
Cold magnetization	$\sigma_{\rm c,0} = B_0^2/4\pi n_0 m_e c^2$	10^4
Hot magnetization	$\sigma_{\rm h,0} = B_0^2/16\pi n_0 T_0$	10^2
Background temp.	$\theta_0 = T_0/m_e c^2$	25
Guide field	B_z	$B_0/4$
Peak drift over background density	$\eta = n_{\rm d}/n_0$	5
Harris layer drift velocity	$\beta_{\rm d} c$	$0.3c$
Harris layer (comoving) temp.	$\theta_{\rm d} = T_{\rm d}/mc^2$	1050
Harris layer half-thickness	$\delta = \sigma_{\rm c,0}\rho_0/\beta_{\rm d}\eta$	$0.67\sigma_{\rm c,0}\rho_0$
Cell size	Δx, Δy	$\sigma_{\rm c,0}\rho_0/24$
Time step	Δt	$0.7\Delta x/c$
Macroparticles per cell		80

and 3.21), do not depend on whether I include or exclude this simulation from the quantitative analysis. I have therefore chosen to include it as a tentative endpoint on the data generated by the rest of the $\gamma_{\rm rad,T}$-varying (L_x-constant) simulation series. However, without running larger and more expensive simulations, I am prohibited from runs with $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$ much lower than 1/2 (and hence from exploring, for example, the extremely radiative regime detailed in section 2.2.2).

I note that the benchmark radiative case $\gamma_{\rm rad,T} = \sigma_{\rm c,0}$ is not completely free of the upstream cooling issue. However, according to equation (3.3), the background temperature may decrease by less than a factor of 2 by the end of that simulation. Because such a discrepancy is within the error bars on the main kinetic beaming quantities measured in section 3.3.3, I view the $\gamma_{\rm rad,T} = \sigma_{\rm c,0}$ simulation as marginally acceptable.

Finally, I would like to point out that upstream radiative losses only increase the effective $\sigma_{\rm h,0}$. Because $\sigma_{\rm h,0}$ is quite large to begin with, raising it by order-unity factors preserves (indeed, enhances) the asymptotically large-$\sigma_{\rm h}$ limit. Hence, it is unlikely that acceleration and beaming of high-energy particles are significantly impacted, even in the most strongly cooled ($\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1/2, 1$) simulations.

3.1.2 Evolution of the reconnection layer

Having described the set-up of my simulations, let me now move on to how they evolve in time. In every run, I trigger magnetic reconnection with a small (1 per cent) perturbation to the initial magnetic field. The current sheet then tears into a number of magnetic islands or 'plasmoids' which begin to merge with one another (Fig. 3.1, $t = 0.4L_x/c = 0.4L/c$). Initially, the plasmoids are all about the same size, but eventually – in Fig. 3.1 at about $t = 1.2L/c$ – a single largest plasmoid begins to dominate the reconnection layer. This primary plasmoid proceeds to consume the others that have also accumulated to considerable but smaller sizes, culminating in a spectacular merger between the largest and next-to-largest plasmoids (Fig. 3.1, $t = 2.0L/c$). After this most dramatic merger, additional small plasmoids are continually born from the main X-point and venture across the box to be consumed by the large primary plasmoid (Fig. 3.1, $t = 2.6L/c$).

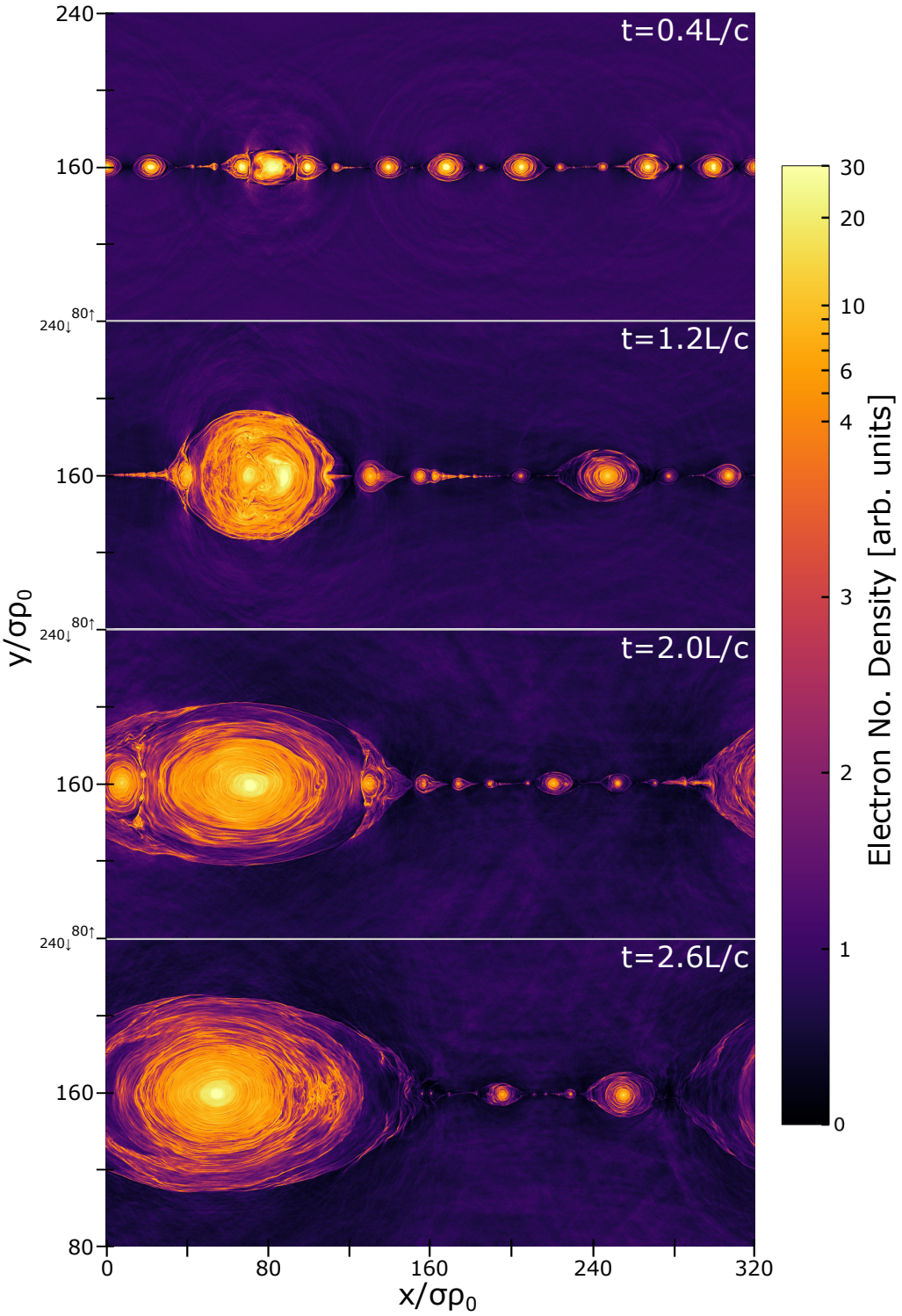

Figure 3.1: Total (drift plus background) electron number density pictured around the lower reconnection layer at key moments during the $\gamma_{\mathrm{rad,T}} = \sigma_{\mathrm{c},0}$ simulation. The full simulation width ($L = L_x = 320\sigma\rho_0$) and a restricted height range ($L_y/4 \pm 80\sigma\rho_0$) are displayed. See text for a description of the various phases of the time evolution.

The simulations exhibit significant NTPA. At late times, the NTPA is bursty: merging plasmoids sporadically punctuate ongoing reconnection from the main X-point with short intense episodes of particle acceleration. In the more radiatively efficient runs, following these episodes, the high-energy particles rapidly cool, leading to a steepening of their non-thermal energy distribution. This effect was first analyzed by Werner et al. (2019) (see also Sironi & Beloborodov, 2020), and I illustrate it in Fig. 3.2, which presents time-dependent particle energy distributions from the lower half of the box in a subset of the simulations. In the limit of weak cooling (e.g., $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 16$ in Fig. 3.2), the distribution develops a shallow power law. However, due to long periods of continuous IC losses interrupted by bursts of plasmoid merger-initiated magnetic reconnection, the particle distributions for the simulations with stronger cooling (e.g. those with $\gamma_{\rm rad,T}/\sigma_{\rm c,0} \leq 4$) all exhibit steeper, more variable power laws at late times. Additionally, for all the displayed simulations, the cut-off particle energy is well approximated by $\gamma_{\rm rad,T}$, indicating that radiative losses control this limit (even when too weak to steepen the non-thermal power-law tail).

3.1.3 A view in angular space

Up until now, I have described the evolution of my magnetic reconnection simulations from spatial and energetic viewpoints. I displayed several snapshots of the electron number density $n(x, y)$ in Fig. 3.1. Then, in Fig. 3.2, I described the electron energy distribution $f(\gamma)$ and how its evolution is impacted by radiative cooling. These pictures represent different ways of viewing the master distribution function in phase space $f(x, y, \gamma, \mathbf{\Omega}; t)$. At a given time t, this master distribution is five-dimensional, containing two spatial and three velocity dimensions, the latter of which I decompose into a Lorentz factor $\gamma = E/mc^2$ and a direction labeled by the solid angle $\mathbf{\Omega}$. In terms of the master distribution, the number density and energy distribution are $n(x, y; t) = \int \mathrm{d}\gamma\, \mathrm{d}\Omega\, f(x, y, \gamma, \mathbf{\Omega}; t)$ and $f(\gamma; t) = \int \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}\Omega\, f(x, y, \gamma, \mathbf{\Omega}; t)$, respectively.

In this work, I am also interested in how particle momenta (and emitted photons) are distributed directionally. As a result, I must keep the angular information in the distribution function $f(x, y, \gamma, \mathbf{\Omega}; t)$, as was first done by Cerutti et al. (2012b). Furthermore, because I am interested

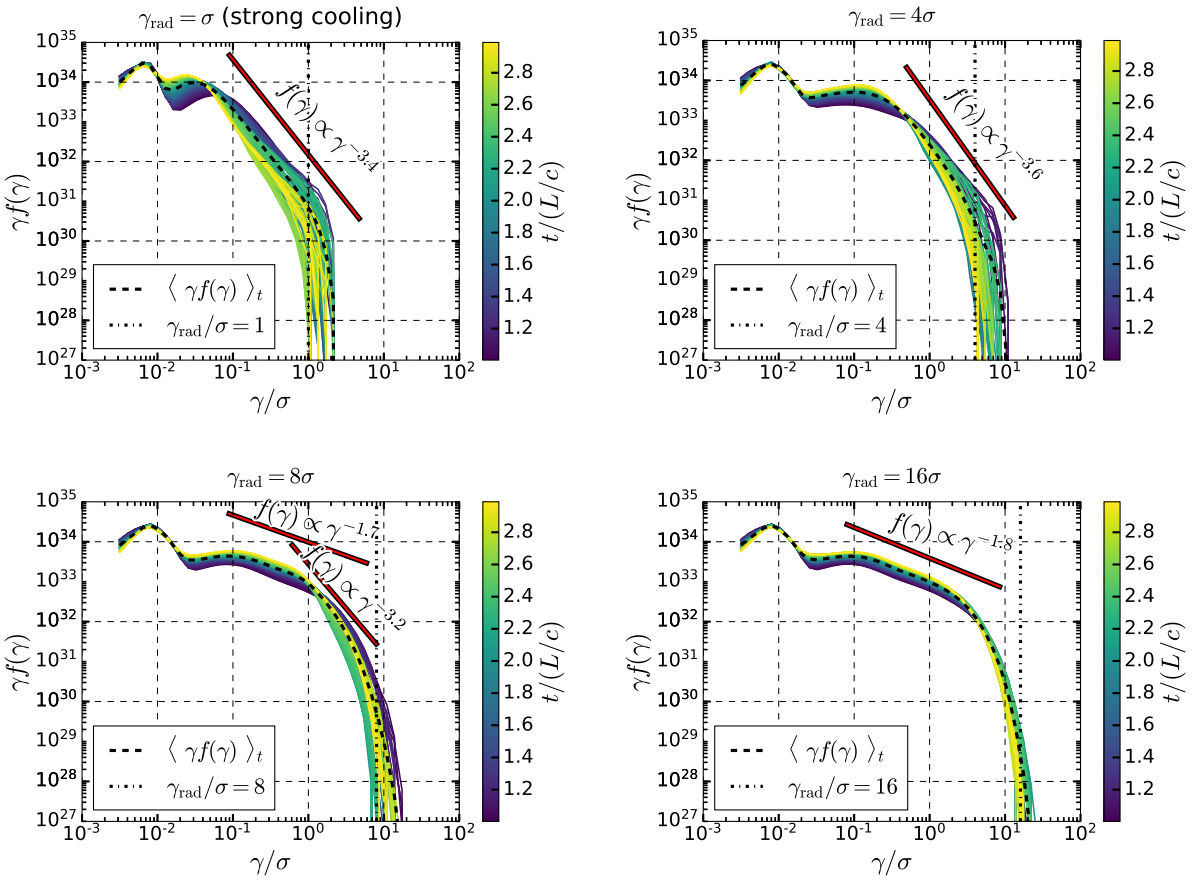

Figure 3.2: Lower layer electron energy distribution as a function of time between L/c and $3L/c$ for a subset of the simulations presented in this work. Colored solid lines denote the distribution at different times. Black dashed lines represent the time-averaged (over $1 \leq ct/L \leq 3$) distribution. Vertical black dot-dashed lines label $\gamma_{\rm rad,T}$. Red bars indicate power-law segments in the time-averaged particle distributions. The qualitative features are consistent with those reported in Werner et al. (2019). Namely, the simulations with strongest cooling ($\gamma_{\rm rad,T} = 1, 4\sigma_{\rm c,0}$) exhibit only a steep power law $f(\gamma) \propto \gamma^{-p}$ with variable index $p \gtrsim 3$; the simulation with intermediate cooling ($\gamma_{\rm rad,T} = 8\sigma_{\rm c,0}$) exhibits both soft variable ($p \gtrsim 3$) and hard steady ($p \lesssim 2$) power-law segments; and the weakly cooled simulation ($\gamma_{\rm rad,T} = 16\sigma_{\rm c,0}$) only contains a hard steady ($p \lesssim 2$) power law.

in *kinetic* beaming – beaming as a function of particle or photon energy – I must preserve correlations between γ and $\mathbf{\Omega}$. To visualize all three velocity dimensions of the distribution function, I separate the energy information from the angular information, viewing the entire angular distribution at a single energy. Examples of this view are the *angular maps* (also *intensity maps* or *heatmaps*) of Figs. 3.3 and 3.4. These display the spatially integrated angular particle distribution $\mathrm{d}N_t/\,\mathrm{d}\gamma\,\mathrm{d}\Omega = \int_{y<L_y/2}\mathrm{d}x\,\mathrm{d}y\,f(x,y,\gamma,\mathbf{\Omega};t)$ at fixed γ and t using the Aitoff projection. A particle contributes to latitude $\varphi \in [-90°, 90°]$ and longitude $\lambda \in [-180°, 180°]$ on a map if its velocity vector parallels the unit vector

$$\hat{\boldsymbol{n}} = \cos(\lambda)\cos(\varphi)\hat{\boldsymbol{z}} + \sin(\lambda)\cos(\varphi)\hat{\boldsymbol{x}} + \sin(\varphi)\hat{\boldsymbol{y}}\,. \tag{3.4}$$

To isolate a single reconnection layer, Fig. 3.3, Fig. 3.4, and all subsequent angular plots are generated using only particles (or, later, photons emitted from particles) located in the lower half of the simulation box.

Let me now describe the physical origins of the basic features in Figs. 3.3 and 3.4. These features can be neatly decomposed into two broad categories: mild horizontal beaming stemming from bulk motion along the primary reconnection current sheet and extreme beaming arising near X-points (not only in the main current layer, but also in secondary ones between merging plasmoids). In the primary current layer near the main X-point, the reconnection electric field points in the $+z$-direction ($\varphi = 0°$, $\lambda = 0°$), resulting in electron acceleration in the $-z$-direction ($\varphi = 0°$, $\lambda = \pm 180°$). As electrons are ejected towards $-z$, they begin to be deflected by lines of reconnected magnetic field, which causes them to disperse towards $\pm x$ ($\varphi = 0°, \lambda = \pm 90°$). This results in the mild concentration of particles along the equator in Fig. 3.4. As these particles radiatively cool, they are simultaneously deposited into plasmoids, and plasmoid bulk motion along the reconnection layer induces a gentle low-energy momentum anisotropy along the $\pm x$-directions as in Fig. 3.3.

I move now to the more extreme beaming. Generally, such pronounced anisotropy occurs only among the higher energy particles, a trend that Fig. 3.4 illustrates nicely. Less universal, but still common, is the fact that the strong beaming patterns in that figure result from plasmoid

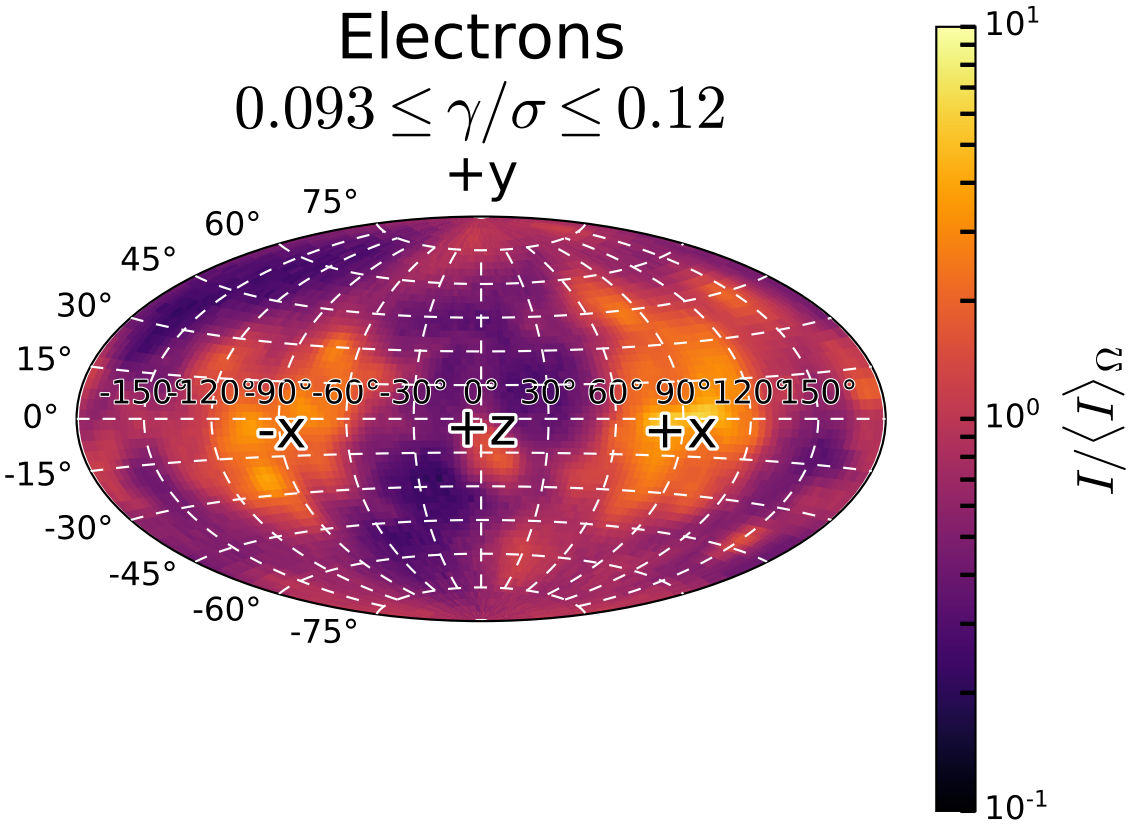

Figure 3.3: An angular map for the $\gamma_{\rm rad,T} = \sigma_{\rm c,0}$ simulation displaying the angular intensity $I = \mathrm{d}N_t/\,\mathrm{d}\gamma\,\mathrm{d}\Omega$ of lower layer electrons at $2.0L/c$. This is a low-energy map – in the sense that the electron Lorentz factors γ are a decade below $\gamma_{\rm rad,T}$ – and exhibits only mild beaming.

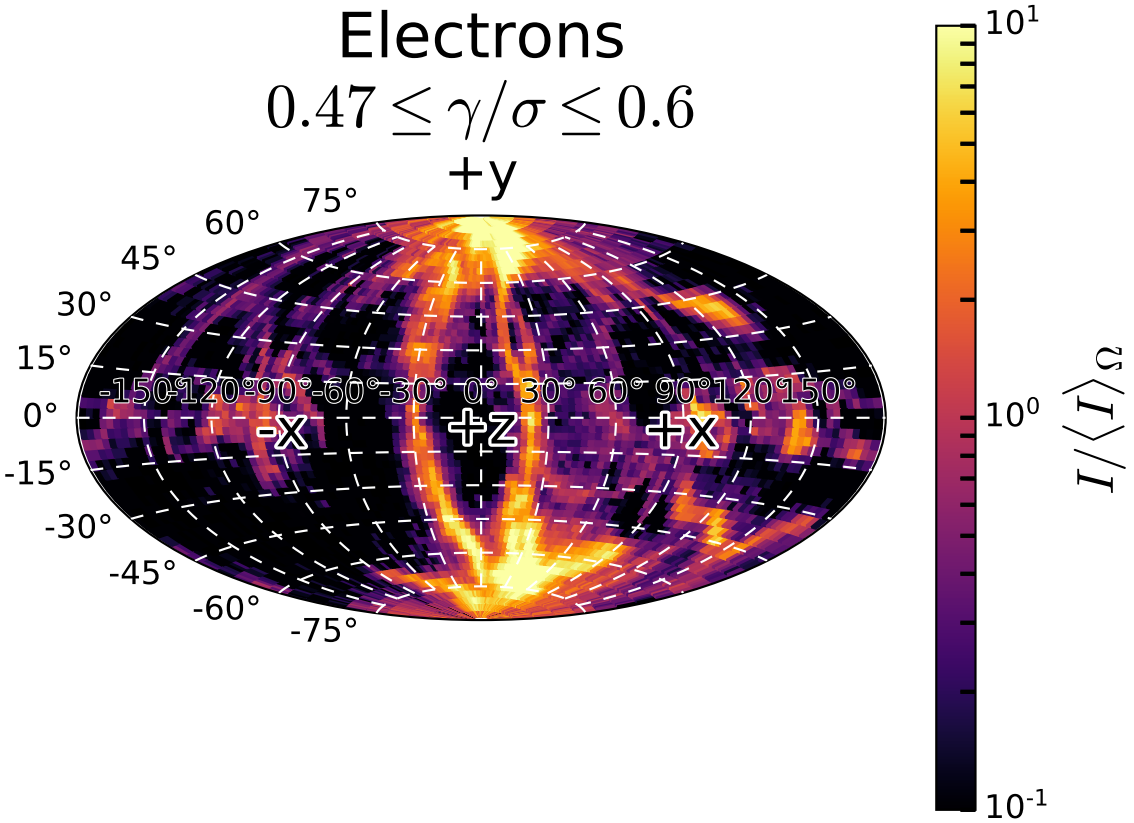

Figure 3.4: An angular map for the $\gamma_{\mathrm{rad,T}} = \sigma_{\mathrm{c},0}$ simulation displaying the angular intensity $I = \mathrm{d}N_t/\,\mathrm{d}\gamma\,\mathrm{d}\Omega$ of lower layer electrons at $2.0L/c$. This is a higher energy map relative to Fig. 3.3 (the electron Lorentz factors are closer to $\gamma_{\mathrm{rad,T}}$) and, as a result, exhibits more intense beaming patterns.

mergers, evidenced by the prominent vertical swaths in the angular particle distribution. This comes about because, between merging plasmoids, a secondary reconnection layer forms approximately parallel to the zy-plane. The reconnection electric field in this secondary layer points along $-z$ and accelerates electrons along $+z$. Owing to the rotated orientation of the reconnecting magnetic field, these electrons begin to fan out towards $\pm y$ – towards the poles – rather than $\pm x$ as in the primary current sheet.

The fan shapes in Fig. 3.4 are not perfectly vertical because plasmoids with unequal sizes and speeds are merging. At this time, there are actually two ongoing mergers – one on either flank of the large primary plasmoid (Fig. 3.6) – both actively accelerating particles. In the merger on the right-hand side, a smaller and faster left-moving plasmoid creates a secondary current sheet that bends and moves to the left, biasing the accelerated particles towards the $-x$-direction. The opposite is true for the merger on the left involving a small/fast right-moving plasmoid. The combined result is that the swaths of high-energy particles shown in Fig. 3.4 do not extend along a single meridional plane running through $\lambda = 0°$, but through two slightly offset planes intersecting $\lambda \simeq \pm 30°$: one for each current sheet created at asymmetric plasmoid mergers.

The features in Figs. 3.3 and 3.4 are nicely mirrored by those in the corresponding positron angular maps. As an example, Fig. 3.5 displays the angular distribution of positrons in the same higher-energy band and at the same time as the electrons in Fig. 3.4. Owing to their opposite response to the reconnection electric field, the positrons yield the same X-point-generated beaming configurations found among the electrons in Fig. 3.4 but reflected about the xy-plane.

To tie the beaming features on angular maps to their spatial origins, Fig. 3.6 shows the spatial electron number density at the time for which Figs 3.3–3.5 were drawn. Also shown are the locations of randomly chosen subsets of electrons with Lorentz factors in the low-energy range of Fig. 3.3 and the high-energy range of Figs 3.4 and 3.5. The positional clustering explicitly demonstrates the dichotomy described above: low-energy particles are confined to plasmoids whose bulk motion governs their momentum anisotropy while high-energy particles exhibit more extreme beaming shaped by reconnection X-points. Fig. 3.6 also shows that X-point acceleration and

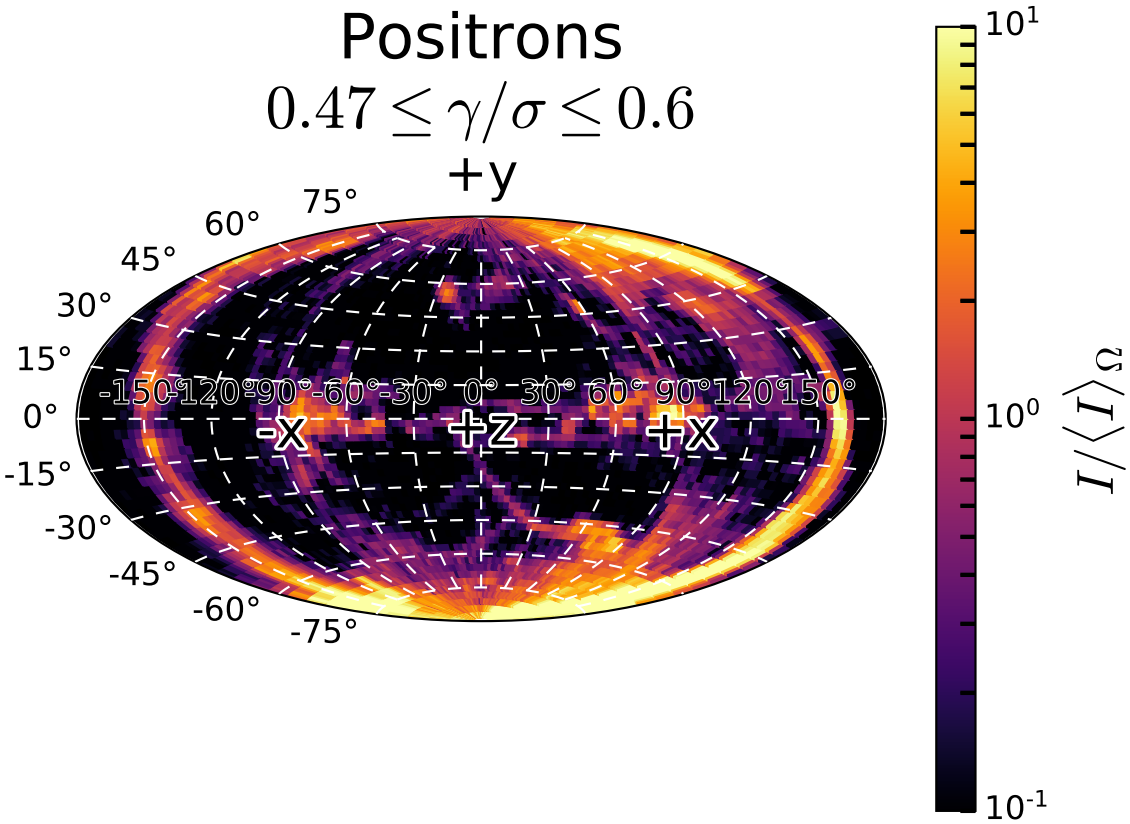

Figure 3.5: A positron intensity map drawn for the same energy band and at the same time as the electron map of Fig. 3.4. Positrons feel the same reconnection electric field as electrons, but experience acceleration in the opposite direction as a result.

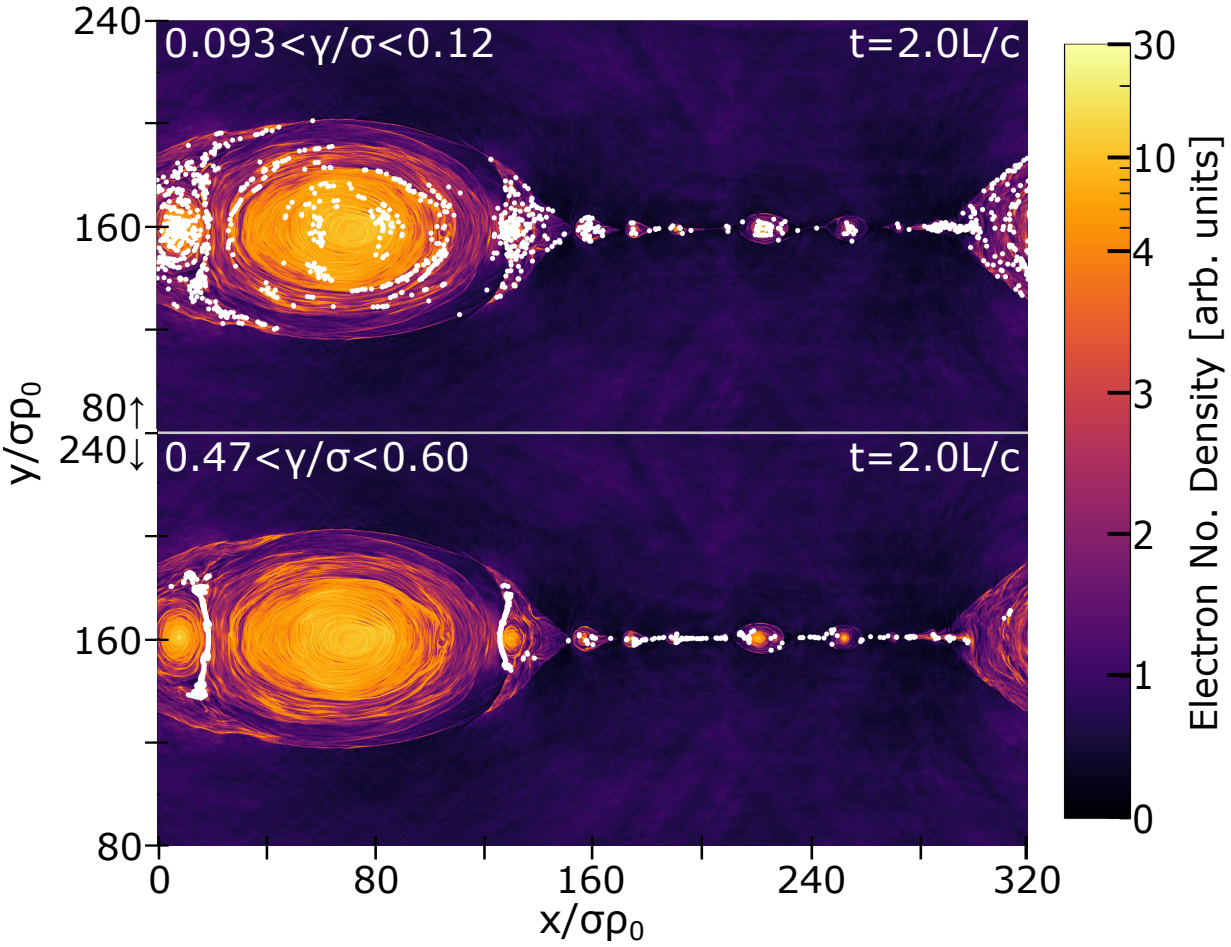

Figure 3.6: Electron density for the same time ($2.0L/c$) used to generate the angular maps of Figs. 3.3-3.5. In the top (bottom) panel, white ovals indicate the locations of a random subset of the electrons contributing to the angular map in Fig. 3.3 (Fig. 3.4) with Lorentz factors in the range $0.093\sigma < \gamma < 0.12\sigma$ ($0.47\sigma < \gamma < 0.60\sigma$). The higher energy particles reside in the hearts of the primary and inter-plasmoid current sheets – near X-points – and the lower energy particles in plasmoids.

collimation is ongoing in the primary reconnection layer even after the plasmoid chain has fully developed. As a result, it is not always the case (as it is in the angular maps shown previously) that the strongest beaming signatures are vertical. In fact, vertical fan shapes tend to be only intermittently prominent: when plasmoids – particularly large ones like in Fig. 3.6 – are actively merging. In between these episodes, strong beaming arising in the primary current sheet can still be significant (e.g., Fig. 3.17).

At this point, I would like to clarify that, of the two beaming origins discussed in this section, it is only the one operating near X-points that was associated with kinetic beaming in the works that originally introduced the concept (Uzdensky et al., 2011; Cerutti et al., 2012a,b). As discussed by those authors, the configuration of electromagnetic fields at these locations is particularly suited to accelerate and collimate high-energy particles: the reconnection electric field delivers energy while the reconnecting magnetic field focuses particles into beams. The particles remaining near X-points longer are consequently more energized and more focused. As demonstrated already (Fig. 3.4), this mechanism is responsible for the most severe beaming at the highest particle energies; as demonstrated later, it also yields the most energy-*dependent* beaming. By comparison, plasmoid motion-generated anisotropy (Fig. 3.3) is milder and tends, because it derives from fluid level motion, to be more achromatic. (Plasmoids, along with their associated bulk motion and Doppler beaming, provide the basis for the 'minijets' model of Giannios et al. 2009, 2010.)

Thus, when I use the term 'kinetic beaming', I am not referring to just any energy-dependence in the degree of particle or photon collimation. I refer specifically to the most extreme and energy-dependent anisotropy at the highest energies generated near X-points. In this sense, the very different signatures of beaming evident in Figs. 3.3 and 3.4, despite occurring at different particle energies, do not illustrate kinetic beaming. Instead, they portray two separate beaming mechanisms that merely dominate at different energy scales. I illustrate kinetic beaming as I and previous authors apply the term – which involves energy-dependent anisotropy sourced only by the X-point mechanism – after I develop a more quantitative description of beaming in the next section.

3.2 Quantifying beaming

In this section, I present two quantitative notions of 'beaming' as manifested in angular maps such as Figs. 3.3–3.5. These facilitate the detailed analysis of observable kinetic beaming, as influenced by varying degrees of radiative efficiency, presented in the next section (3.3). Before proceeding, it is helpful to introduce some additional terminology with which to describe the information on these maps: the angular distribution of particles $\mathrm{d}N/\,\mathrm{d}\Omega$, of the instantaneous radiated power $\mathrm{d}P/\,\mathrm{d}\Omega$, and quantities derived from these.[1] The angular distribution $\mathrm{d}P/\,\mathrm{d}\Omega$ is ordinarily called 'intensity' I, and the power P radiated into a finite solid angle is $P = \int I\,\mathrm{d}\Omega$. In analogy with light, I shall frequently call the angular distribution of particles $\mathrm{d}N/\,\mathrm{d}\Omega$ by the name 'intensity', as well as borrow the symbol I. Furthermore, I use the word 'power' to refer to the total number of particles traveling within a finite angular patch. This language enables me to describe beaming in generic terms. Whether I mean a beam of particles or a beam of photons will be clear from the context.

3.2.1 Two notions of beaming

As demonstrated by the intensity maps of Figs. 3.3 and 3.4, the ZELTRON particle distributions do not necessarily exhibit what one typically imagines as a beam: a spot of high intensity that is nearly symmetric about some axis. Rather, the high intensity regions on angular maps can be quite extended and complicated in shape, particularly at higher energies. Any quantitative definition of beaming one adopts must therefore be sufficiently versatile to handle the diverse set of momentum-space configurations attained by the particles (or photons, but for concreteness this section confines the discussion to the particle distribution).

To meet this challenge, I employ two complementary measures of beaming. The first was originally introduced by Cerutti et al. (2012b), who parameterized beaming by Ω_{50}: the smallest total (possibly non-contiguous) solid angle containing half of the power on an angular map (within

[1] I temporarily omit to explicitly write the dependence on time t, as well as the γ-dependence of $\mathrm{d}N/\,\mathrm{d}\Omega$ and the spectral dependence of $\mathrm{d}P/\,\mathrm{d}\Omega$, while establishing my nomenclature.

the given energy bin). This quantity is illustrated in Figs. 3.7 and 3.8. A smaller value indicates more extreme beaming because a smaller fraction of the sphere contains an order unity fraction of the power.

The second measure of beaming characterizes the angular regions where the intensity exceeds three times the angle-averaged intensity (denoted $\langle I \rangle_\Omega$ in the figures; again within a single energy bin). I define the *beamed fraction*, abbreviated as bf, to be the fraction of the total power contained within these regions. Extraction of the beamed fraction from the angular maps in Figs. 3.3 and 3.4 is demonstrated in Figs. 3.9 and 3.10.

Both measures of beaming – Ω_{50} and bf – have advantages and disadvantages. The beamed fraction does not rely on regions of extreme intensity being confined to small fractions of the sphere. On the other hand, Ω_{50}, when small, is perhaps a more convincing indicator of beaming because it means that the corresponding angular map region contains high power and occupies a small solid angle; large beamed fraction indicates only high power. I use both diagnostics in order to give a more compelling account of kinetic beaming.

One thing that these metrics have in common is that they are insensitive to the shapes and continuities of beams. Although one may conceive of more detailed and observer-centric measures of beaming, perhaps characterizing the morphologies of individual contiguous beams, this would greatly complicate the analysis. Leaving that for a future work, I find that these more coarse-grained measures are sufficient to illustrate a number of intriguing properties of the global system-wide beaming produced by magnetic reconnection.

Having developed two notions of beaming, I am now in a position to analyze *kinetic* beaming, which necessarily involves many maps across the particle energy spectrum. In this effort, the chief utility of the Ω_{50} and bf measures is to enable a reduction of the data contained on any given heatmap to two meaningful numbers, which may then be plotted as a function of particle or photon energy. This procedure is illustrated in Figs. 3.11 and 3.12. In the first figure, I display a collection of electron intensity maps spanning a decade in particle energy at a given instant in the $\gamma_{\rm rad,T} = \sigma_{\rm c,0}$ simulation (strongly radiative). Each map in that figure is distilled to two numbers, its Ω_{50} and its

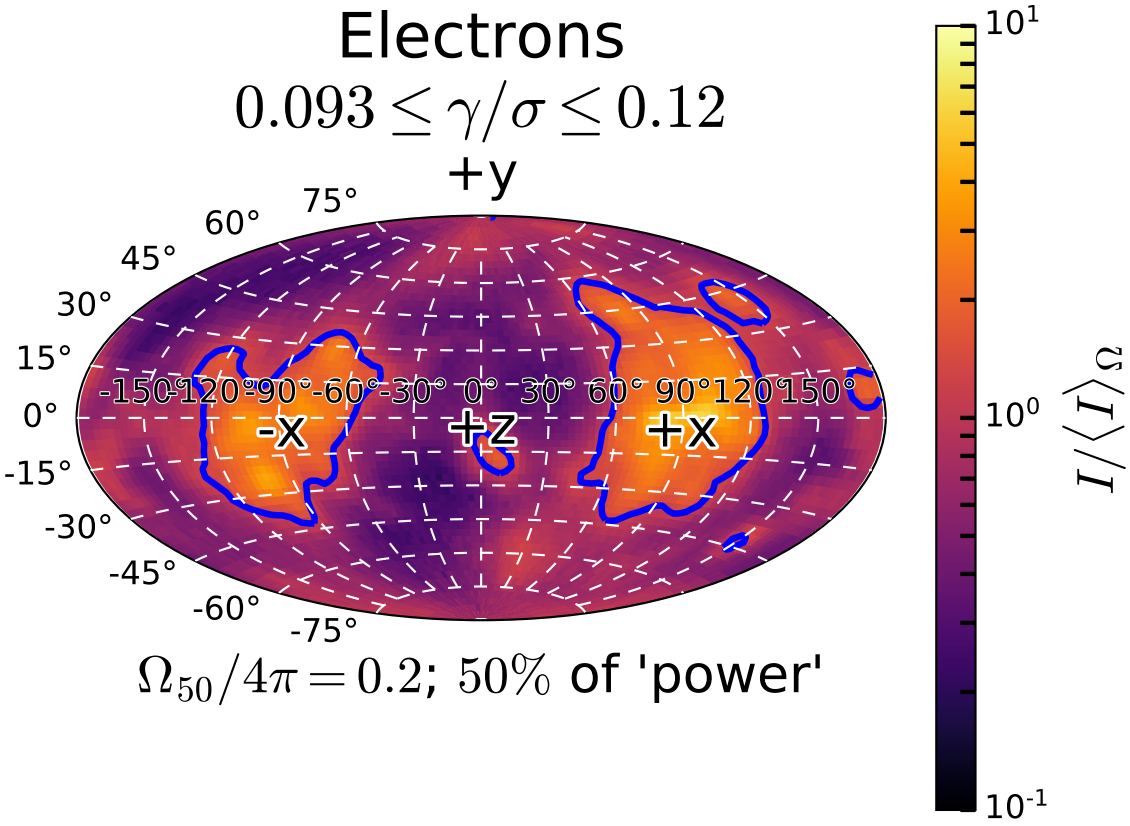

Figure 3.7: The same angular map as Fig. 3.3, but with a blue contour outlining the smallest solid angle that contains half of the total power – i.e., Ω_{50}, which in this case is 20 per cent of 4π.

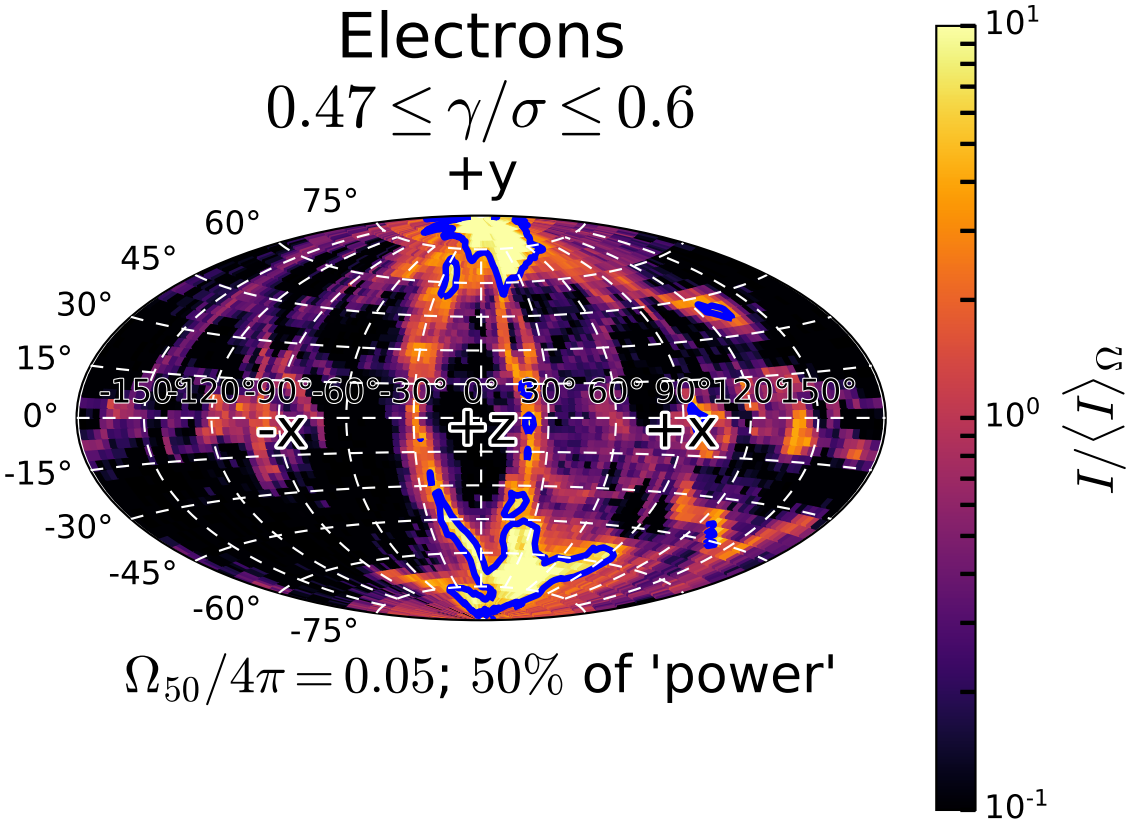

Figure 3.8: The same intensity map as in Fig. 3.4, but with the Ω_{50} contour labeled. In this case, Ω_{50} is 5 per cent of 4π.

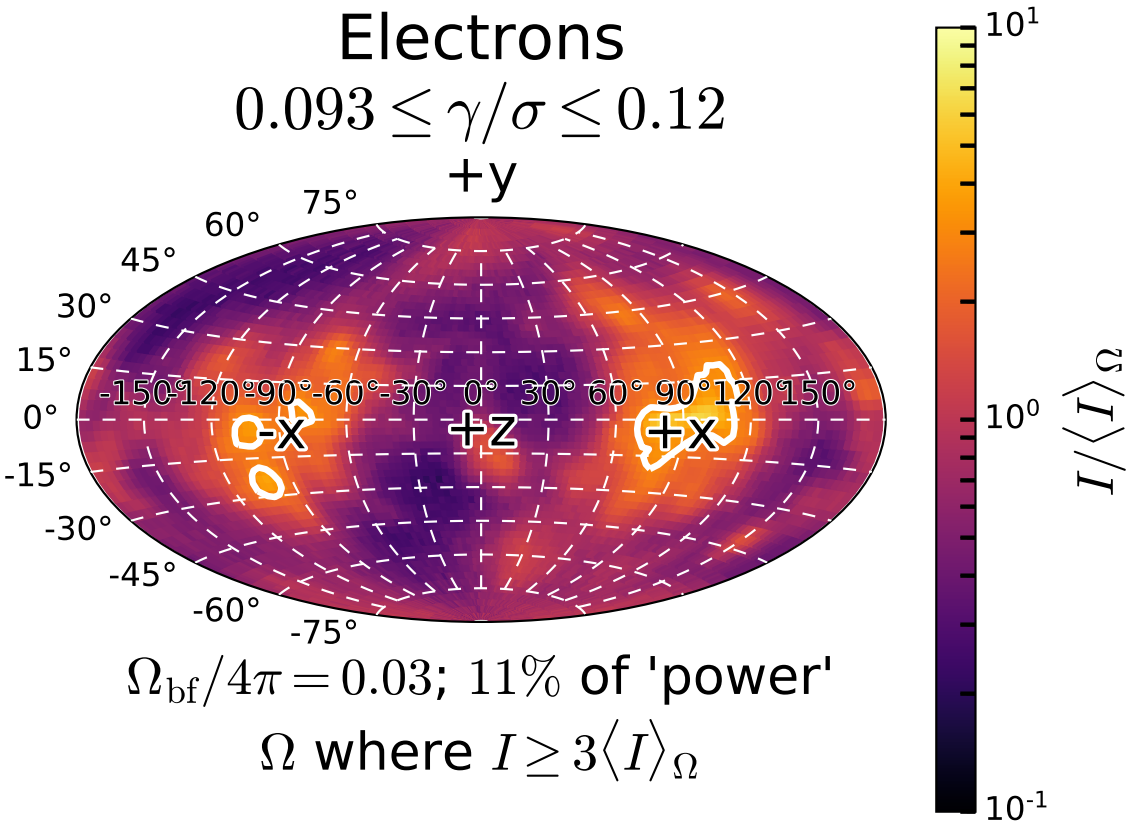

Figure 3.9: The same angular map as in Fig. 3.3, but with a white contour outlining the high intensity region (where the intensity exceeds three times its angle-average). The fraction of the heatmap power contained in this contour is the *beamed fraction* and in this case is equal to 11 per cent. The solid angle footprint of the high intensity region is 3 per cent of the sphere (note that the solid angle Ω_{bf} enclosed by the contour is *not* Ω_{50}). The beamed fraction provides an alternative measure of beaming to Ω_{50}.

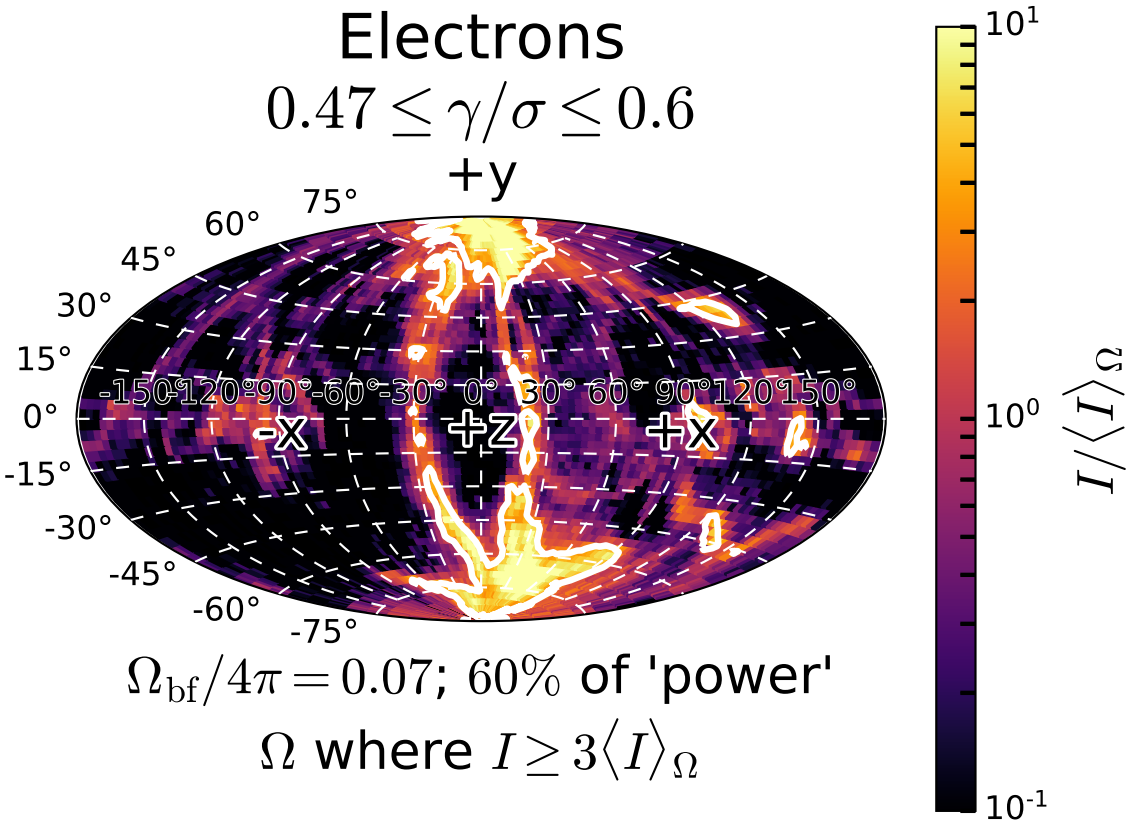

Figure 3.10: The same map as Fig. 3.4, but with a white contour outlining the high intensity region ($I \geq 3\langle I \rangle_\Omega$). Here the beamed fraction is equal to 60 per cent and is contained in 7 per cent of the available solid angle.

beamed fraction, which are then plotted as a function of particle energy in Fig. 3.12. The latter figure depicts the pronounced energy-dependence of beaming more concisely and dramatically, and I make use of many similar plots throughout the remainder of this work.

Before moving on to concentrate more exclusively on the succinct energy-centric view of beaming afforded by plots like Fig. 3.12, I would like to pause and emphasize, once more, the connection between the angular configurations realized in Fig. 3.11 and their underlying physical mechanisms (discussed previously in Section 3.1.3). Namely, Fig. 3.11 demonstrates: (1) mild beaming in the $\pm x$-directions ($\varphi = 0^\circ, \lambda = \pm 90^\circ$) due to bulk plasmoid motion; and (2) dramatic beaming originating near X-points – in this case, X-points between merging plasmoids – and extending from the $+z$-direction ($\varphi = 0^\circ, \lambda = 0^\circ$) towards the poles ($\varphi = \pm 90^\circ$). The former mechanism is most prominent at lower energies but the latter takes precedence at higher energies and gives rise to the steepest energy-dependence in Fig. 3.12. As a reminder, it is this more extreme beaming that I term *kinetic beaming* and to which I devote the majority of the analysis in the next section.

3.3 Kinetic beaming and radiative cooling

In this section, I apply the quantitative measures of beaming described above – Ω_{50} and beamed fraction (bf) – to determine how the observational character of kinetic beaming changes under varying degrees of radiative efficiency ($\gamma_{\rm rad,T}$). First, I consider the question of observable kinetic beaming for two extreme cases: no radiative cooling and strong radiative cooling. After examining these scenarios in detail, I conduct a higher level analysis that makes use of the full parameter scan in $\gamma_{\rm rad,T}$ (Table 3.1) to create a more complete picture of the dependence of kinetic beaming on cooling efficiency.

3.3.1 No cooling: $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = \infty$

For the simulation without IC cooling ($\gamma_{\rm rad,T}/\sigma_{\rm c,0} = \infty$), the time evolution of three quantities as a function of particle Lorentz factor is displayed in Fig. 3.13. From top to bottom, these are the electron energy distribution, electron beamed fraction, and electron Ω_{50}. In the figure, transient

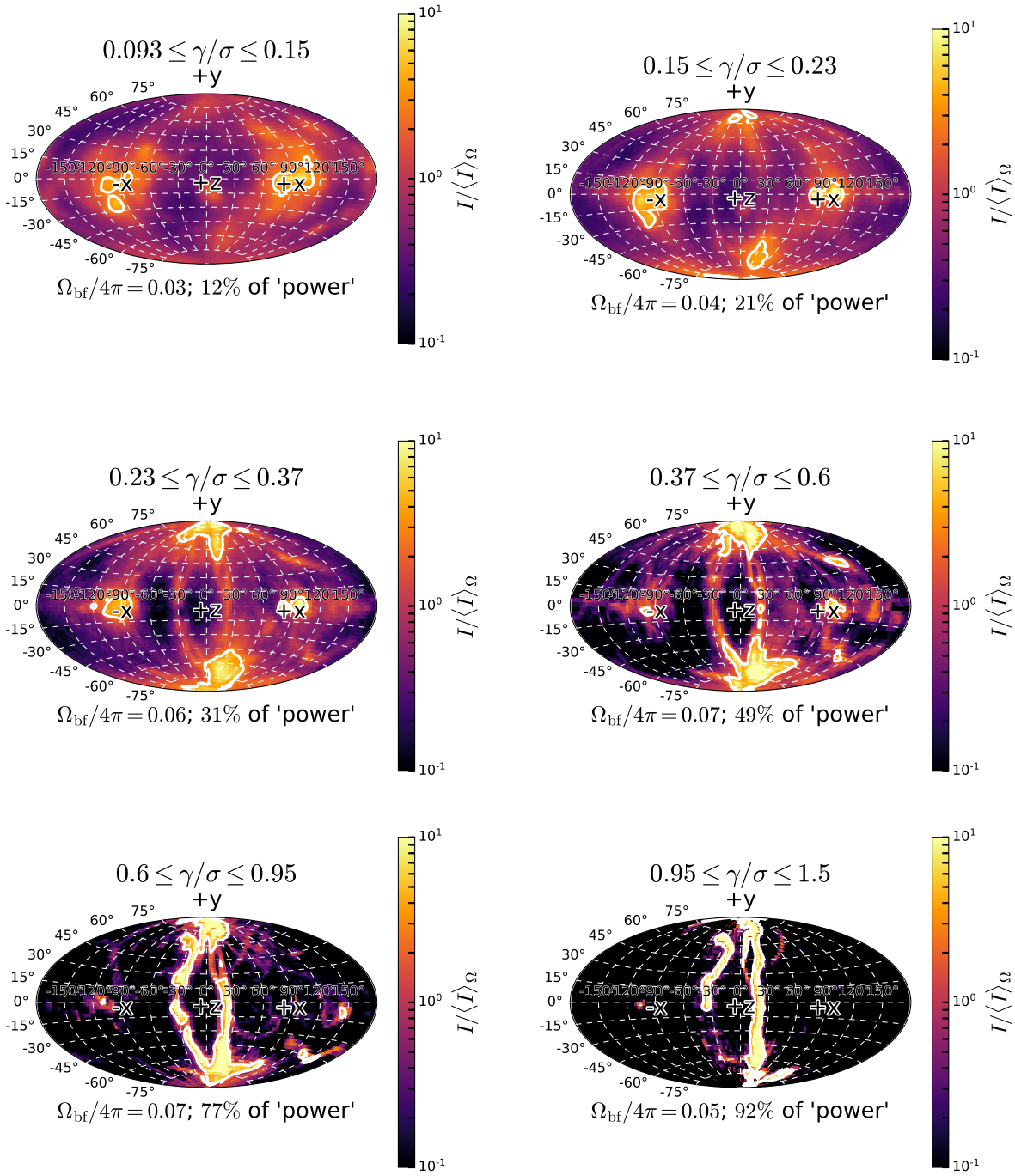

Figure 3.11: Electron intensity maps from the $\gamma_{\rm rad,T} = \sigma_{\rm c,0}$ simulation, with white contours outlining the beamed fraction (where the intensity is more than three times the average), for a series of particle energy bins at a single time $t = 2.0L/c$. Higher energy particles are more strongly beamed and, in this case, originate from reconnection sites between merging plasmoids.

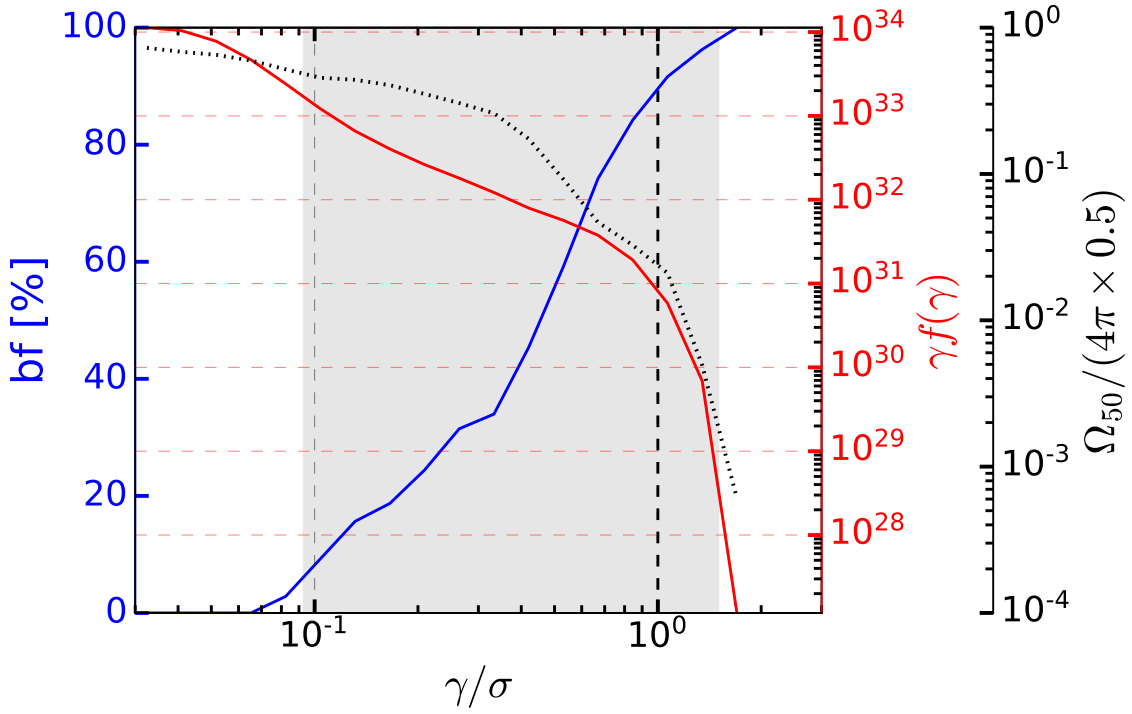

Figure 3.12: Electron beamed fraction (solid blue), Ω_{50} (dotted black), and energy distribution (solid red) as a function of electron energy at $t = 2.0L/c$ in the $\gamma_{\rm rad,T} = \sigma_{\rm c,0}$ simulation. The Lorentz factor $\gamma = \gamma_{\rm rad,T} = \sigma_{\rm c,0}$ is denoted by a vertical dashed line. The Ω_{50} curve is normalized such that a perfectly isotropic angular map produces the value $\Omega_{50}/(4\pi \times 0.5) = 1$. In the same limit (attained at low energies), the beamed fraction tends to zero because the intensity is everywhere less than three times the isotropic intensity. The shaded region indicates the energy range shown in Fig. 3.11. The coincident sharp rise in the beamed fraction and precipitous drop in Ω_{50} demonstrate pronounced kinetic beaming at the highest energies.

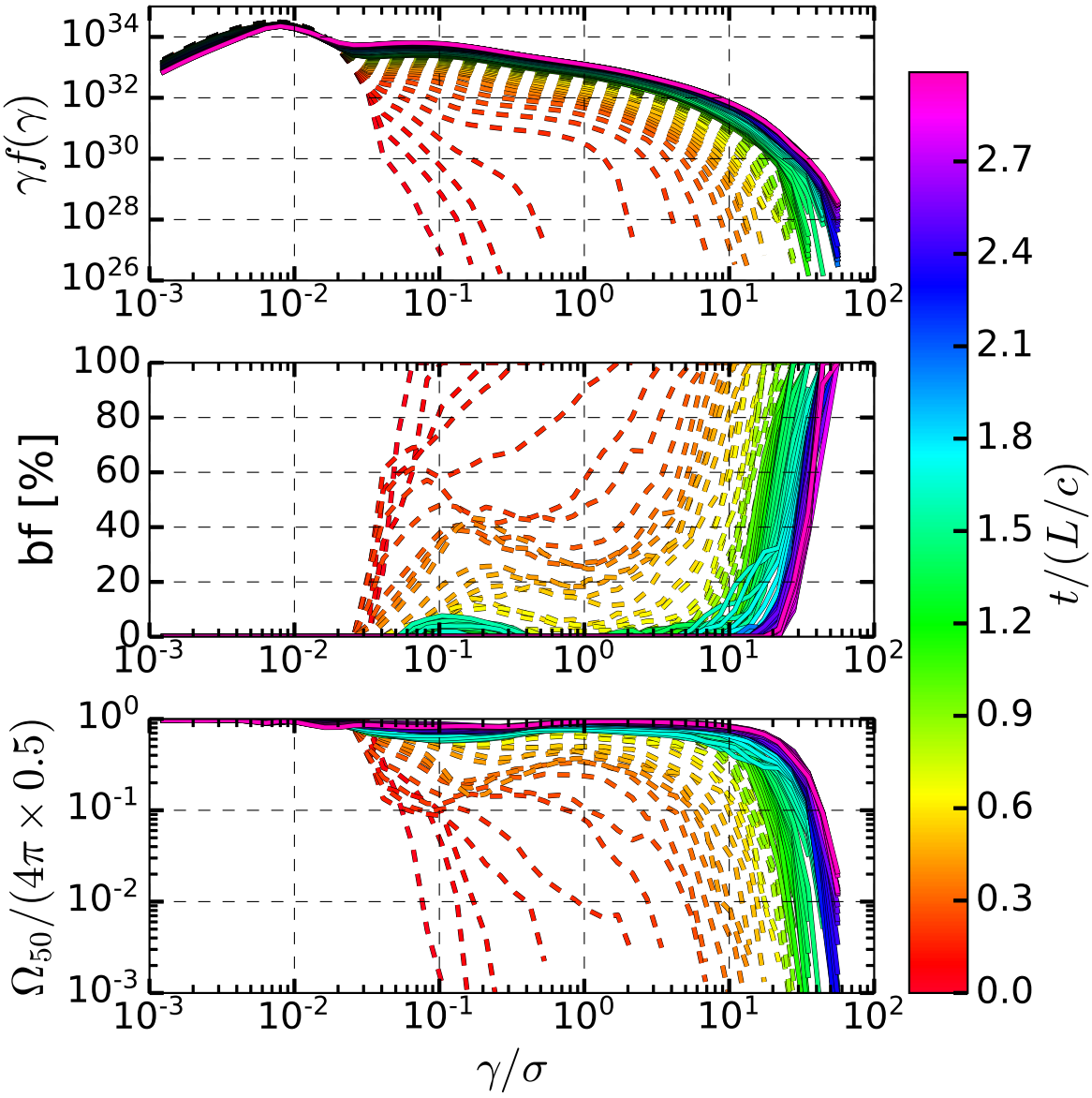

Figure 3.13: The evolution of the electron energy distribution (top panel), beamed fraction (middle panel), and Ω_{50} (bottom panel) for the simulation with $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = \infty$ (no cooling). Dashed lines indicate data prior to the first light crossing time and solid lines data between 1 and 3 light crossings. The solid lines show that, at late times, beaming vanishes at all but the highest (recently populated) energies: the beamed fraction tends to zero and Ω_{50} to $0.5 \times 4\pi$.

behavior in all three quantities persists through about one light crossing time. During this early stage, beaming is both present and energy-dependent, with bf rising sharply and Ω_{50} falling steeply at high Lorentz factors: one observes clear kinetic beaming.

Here, I am restricting the discussion to the highest particle energies, ignoring the non-monotonic behavior in bf and Ω_{50} that takes place at lower energies (and primarily at early times). This behavior stems chiefly from a competition between the two sources of anisotropy in the particle distribution discussed previously in section 3.1.3: plasmoid motion, which induces mild beaming among the low-energy particles, and collimation near reconnection X-points, which has a much more dramatic beaming effect primarily at high particle energies. At intermediate energies, the contributions from both plasmoids and X-points to the global (spatially integrated) distribution of particles can be approximately equal, causing bright regions to cover a larger portion of the angular map and, hence, making it appear more isotropic. By focusing on the highest particle (and, later, photon) energies, where beaming is monotonically increasing [$\mathrm{d}(bf)/\,\mathrm{d}\gamma > 0$ and $\mathrm{d}\,\Omega_{50}/\,\mathrm{d}\gamma < 0$], one isolates the contribution from X-points, the true underlying agents of kinetic beaming as defined here and in previous works (Uzdensky et al., 2011; Cerutti et al., 2012a,b, see also section 3.1.3).

Returning now to Fig. 3.13, one sees that at later times, beaming is quenched. After one light crossing, both the bf and Ω_{50} curves approach their isotropic values – 0 and $0.5 \times 4\pi$, respectively – across nearly all particle energies. As discussed below, this occurs because, after their initial energization, particles quickly isotropize due to gyration about reconnected magnetic field lines. At first glance, it may appear that the highest Lorentz factors – those near the cutoff in the particle distribution – are exceptions to this rule, with dramatic beaming occurring even at late times. This is not really a persistent effect, however, because beaming lasts only temporarily at any fixed Lorentz factor, beginning when the high-energy cutoff crosses (from below to above) that particular energy and ending shortly thereafter. Evidently, high-energy bands retain their beaming only until they may be populated by a significant number of particles. The reason for this is illustrated in Fig. 3.14, and I discuss it here.

High-energy bands in the particle distribution begin the simulation empty. At some point,

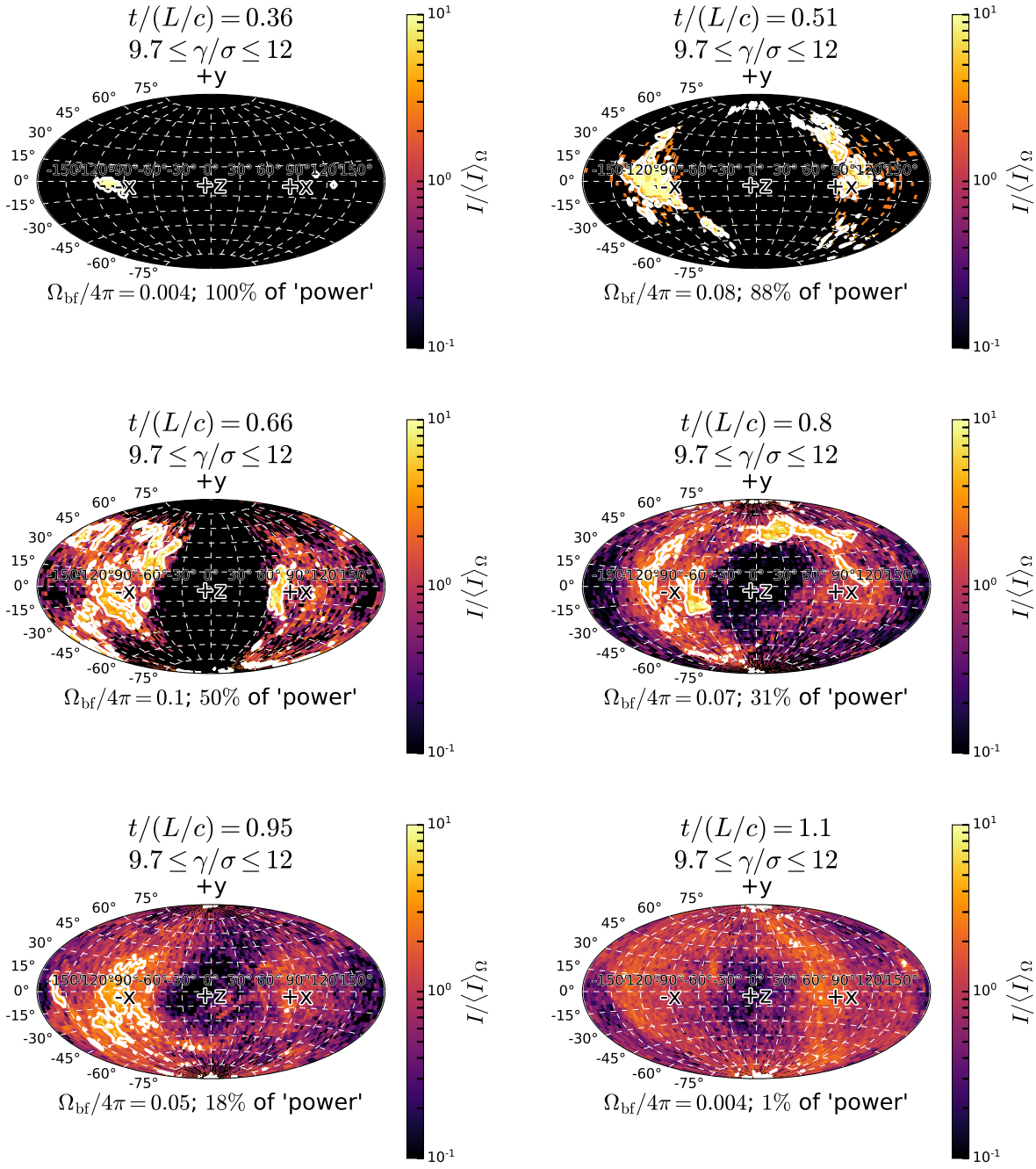

Figure 3.14: Intensity maps for an exemplary high-energy bin across several times for the $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = \infty$ simulation. Beamed fraction contours are in white and beamed fraction values are labeled below each map. Initially, the energy bin contains no particles; only around $0.36L/c$ are particles first accelerated into the bin. Although initially beamed, in the absence of cooling, the particles maintain their energies and isotropize over time. Later, newly accelerated and collimated particles continue to appear in this energy band, but their contribution is drowned out by older, isotropized particles. The generic behavior of the high-energy uncooled particle distribution may be summarized as follows: high-energy bands start empty, they become briefly beamed when they first acquire particles, but they then relax to an isotropic state.

X-point energization begins to populate such a band with particles. These 'young' (recently accelerated) particles are beamed in the same way that particles emerge beamed from the reconnection layer in the early part of the simulation ($t < L/c$, as observed previously by Cerutti et al., 2012b, who confined their analysis to early times). In the absence of radiative cooling, however, particles essentially remain in the energy band to which they are first accelerated. Meanwhile, they settle into plasmoids, where magnetic gyromotion isotropizes their momentum distribution. After a while, a given high-energy band is dominated by 'old' particles whose angular spread has lost the beaming imprint left by X-points. Though a few young particles may still be injected into the energy band, their contribution to the angular intensity is washed out by the large number of older particles that has already accrued there. As a result, the high-energy parts of the particle distribution contain brief, intense beaming when they first acquire particles, but subsequently isotropize as older, increasingly isotropic particles begin to pile up.

I have simplified this explanation by considering particle acceleration to be dominated by the impulsive X-point mechanism. Alternative slower and more isotropic acceleration channels have been studied, for example, by Petropoulou & Sironi (2018), Guo et al. (2019), and Hakobyan et al. (2021). However, these should be suppressed in the strongly radiative regime, unable to keep pace with the rapid cooling of the most energetic particles. Moreover, even when radiative losses are weaker (as in the present case $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = \infty$) and these slower mechanisms are more likely to operate, they can only serve to reduce the amount of beaming I measure, tending to swamp the highly anisotropic angular signatures produced at X-points. Practically speaking, this means that, while an isotropic angular map may be the combined result of dispersing beams *and* intrinsically isotropic energization mechanisms, maps indicating strong kinetic beaming can only be attributed to X-point acceleration.

3.3.2 Strong cooling: $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1$

I now turn to the case of strong radiative cooling $\gamma_{\rm rad,T}/\sigma_{\rm c,0} \lesssim 1$, analyzing the $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1$ simulation in detail. In this regime, the radiative cutoff, $\gamma_{\rm rad,T}$ is not far above the typical

energy ($\langle\gamma\rangle \sim \sigma_{\rm c,0}/4$) – and well below the maximum energy (e.g., $\gamma_{\rm X} \sim 4\sigma_{\rm c,0}$, Werner et al., 2016) – that an energized particle would have in the absence of cooling. Thus, $\gamma_{\rm rad,T}$ chops off the part of the non-thermal power-law tail that could otherwise extend to energies above $\gamma_{\rm rad,T}$ (see Fig. 3.2 and the surrounding discussion).

Paralleling my treatment of the non-radiative case, Fig. 3.15 plots for the simulation with $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1$ what Fig. 3.13 plots for the simulation with $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = \infty$. Here, strong cooling causes the electron distribution, beamed fraction, and Ω_{50} to depart from their previous behavior, where they essentially grew monotonically in horizontal or vertical extent. Because this makes it hard to discern the time evolution in Fig. 3.15, I also supply Fig. 3.16, which presents the post-one-light-crossing time-averaged particle distribution, median beamed fraction, and median Ω_{50}. As evident from Fig. 3.16, beaming persists, when cooling is strong, to late times across almost a decade in particle energy. In contrast, as shown in the previous section, kinetic beaming is only transient when cooling is weak; it is present across a wide range of energies at early times and thereafter relegated to energies near the cutoff of the particle distribution.

As is the case without radiative losses, acceleration from reconnection X-points preferentially collimates the more energetic particles into beams. The crucial difference with strong cooling is that particles radiate away most of their energy before they have had time to isotropize: the most energetic particles are always 'young' (recently accelerated). As a result, the high-energy part of the particle distribution remains beamed at much later times (see Fig. 3.17). Beaming falls off with decreasing particle energy, however, because particles that have been cooling longer have also been isotropizing longer (they are 'older').

In principle, whether kinetic beaming persists at a given particle energy comes down to whether the isotropization time-scale for those particles is longer or shorter than their cooling time-scale. Suppose, for illustration, that the particle isotropization time-scale is the gyration period $t_{\rm iso} \sim \gamma/\omega_0$ where the nominal Larmor frequency is $\omega_0 \equiv c/\rho_0 = eB_0/m_e c$. The cooling time-scale for the same particles is $t_{\rm cool,T} \sim \gamma m_e c^2/P_{\rm T}(\gamma) \sim 10\gamma^2_{\rm rad,T}/\omega_0\gamma$. One expects the smallest Lorentz factor $\gamma_{\rm iso}$ for which kinetic beaming persists to late times to be that for which these time-

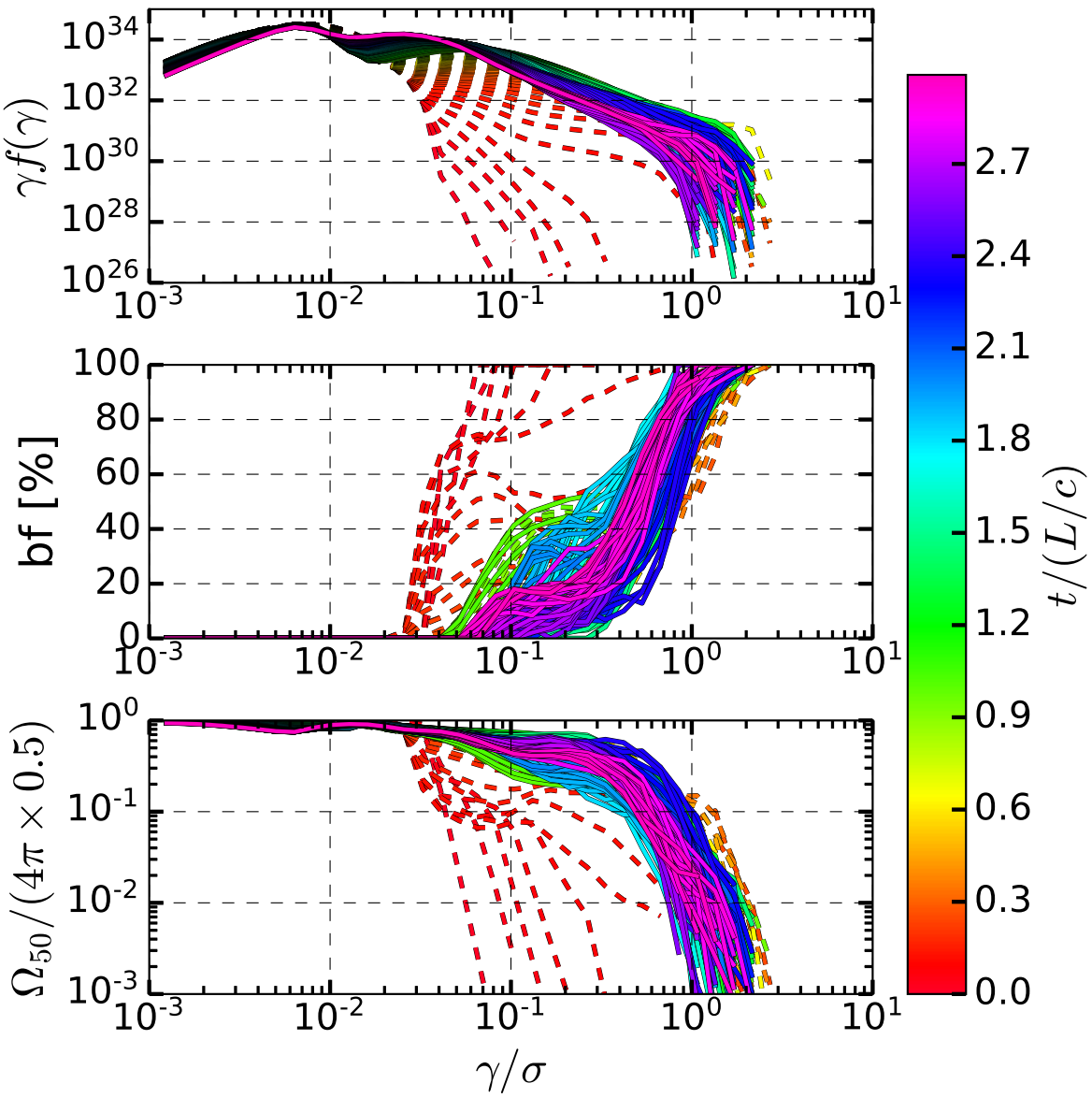

Figure 3.15: The same as Fig. 3.13 but for the simulation with $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1$ (strongly cooled). Unlike the non-radiative case, beaming is present across a moderate range of energies and persists well beyond $t = L/c$. The envelope of the late-time beamed fraction curves indicates at least mild kinetic beaming across a decade in particle energies. The strongest beaming occurs over a somewhat smaller range, where the beamed fraction curves begin to rise steeply and the Ω_{50} curves begin to turn over.

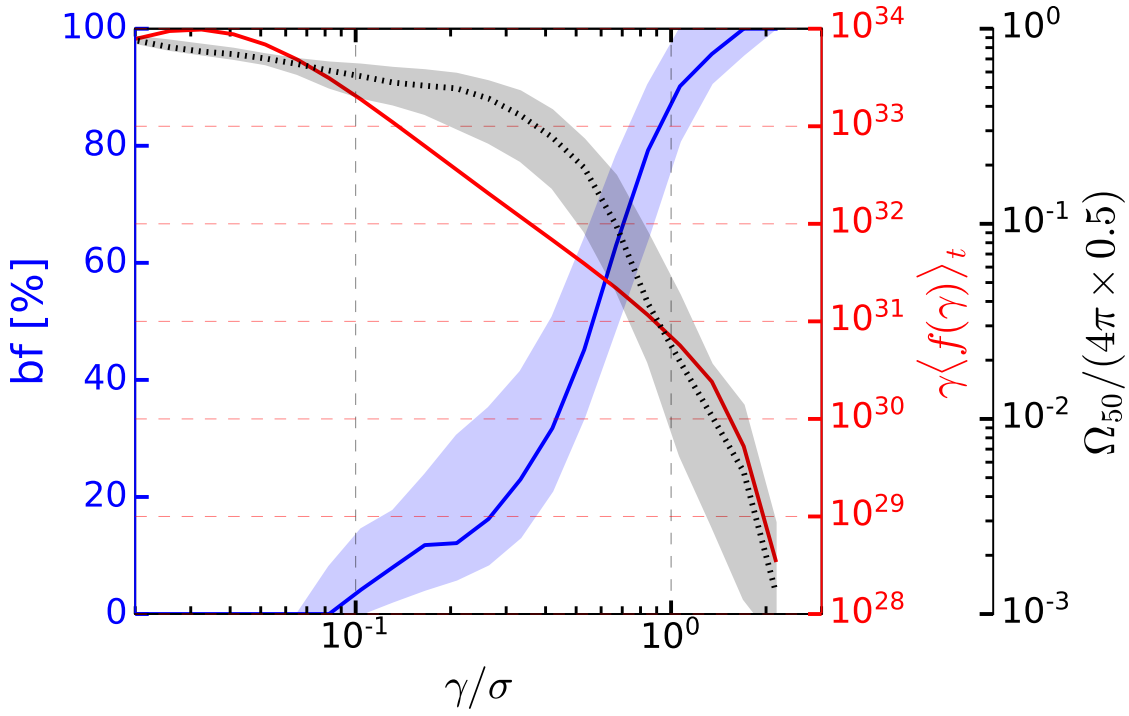

Figure 3.16: The time-averaged electron distribution (red solid line), median electron beamed fraction (blue solid line), and median electron Ω_{50} (black dotted line) as a function of particle energy for times between L/c and $3L/c$ in the $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1$ simulation. The shaded beamed fraction and Ω_{50} envelopes indicate the middle 68 per cent of the time series data at each particle energy. As discussed in Fig. 3.15, but somewhat more obvious here, the beamed fraction indicates at least weak beaming across a decade on the horizontal axis. Both beamed fraction and Ω_{50} indicate strong beaming over a slightly narrower energetic range.

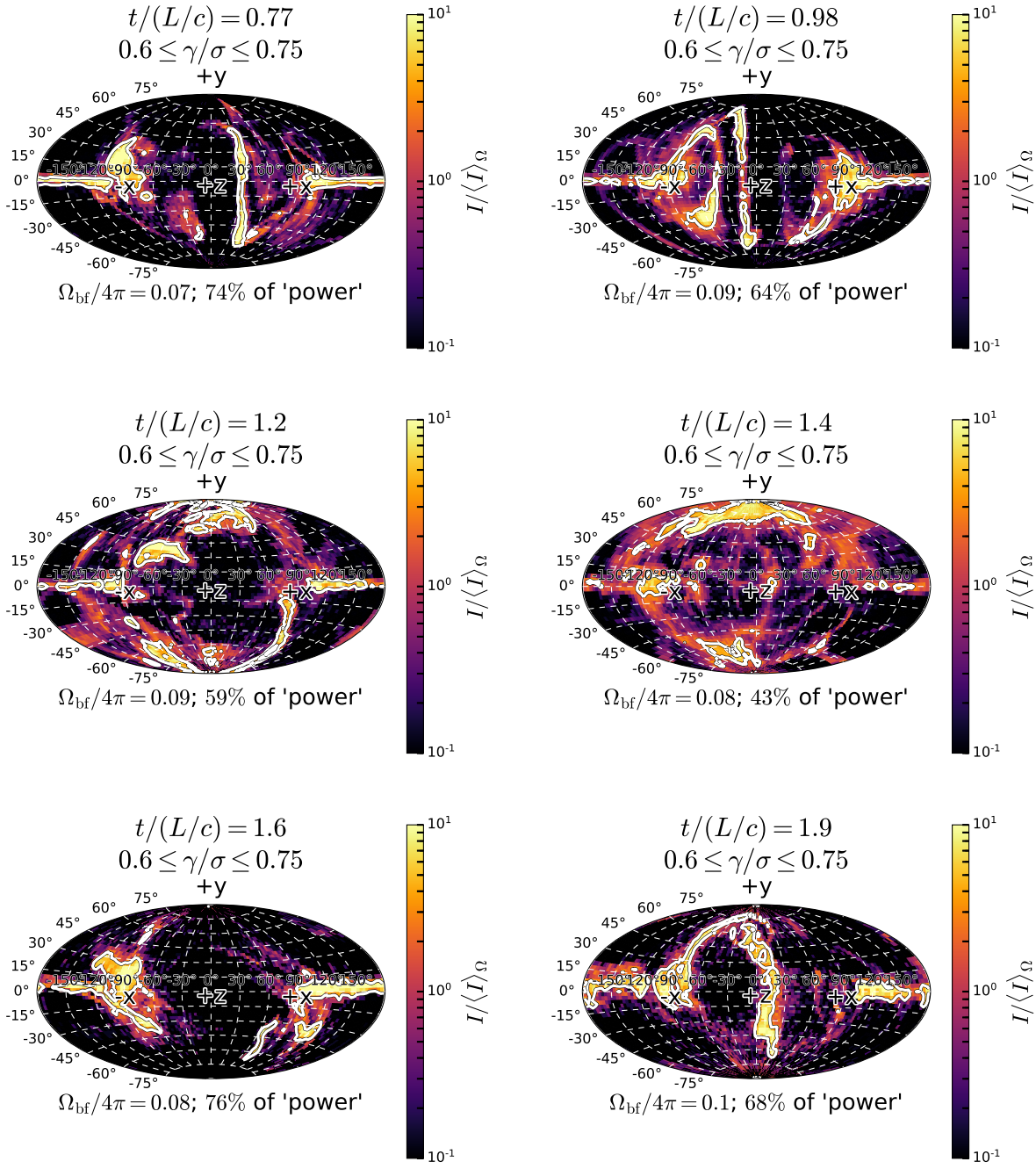

Figure 3.17: The time-dependence of a high-energy heatmap for the $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1$ simulation. In the uncooled simulation, a similar set of figures (Fig. 3.14) demonstrated the transience of beaming in that energy bin. But in this strongly cooled case, spectacular beaming patterns persist to times well past $t = L/c$. Note that sometimes (particularly $t = 1.6L/c$) X-points in the primary reconnection layer dominate the beaming, producing horizontal swaths of high intensity regions; at other times, the prominent vertical swaths indicate beaming produced at X-points between merging plasmoids (cf. section 3.1.3 discussion).

scales are equal: $\gamma_{\rm iso} \sim \sqrt{10}\gamma_{\rm rad,T}$. This picture is oversimplified, for it predicts $\gamma_{\rm iso}$ to exceed (somewhat) the radiation-reaction limit $\gamma_{\rm rad,T}$, and it predicts $\gamma_{\rm rad,T}/\gamma_{\rm iso} \sim$ constant, while I find (see section 3.3.3) a non-trivial scaling of $\gamma_{\rm rad,T}/\gamma_{\rm iso}$ with $\gamma_{\rm rad,T}$. What can be said for certain in the case $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1$ is that particles radiate more quickly than they isotropize over a considerable range of energies.

3.3.3 Kinetic beaming as a function of radiative efficiency

Having demonstrated that strong kinetic beaming persists in the presence of efficient radiative cooling but disappears after about $t = L/c$ when cooling is negligible, I now analyze the transition between these regimes. In particular, I examine how sustained kinetic beaming weakens as the result of decreasing IC radiative efficiency. I also shift my focus from the angular particle distribution $\mathrm{d}N_t/\,\mathrm{d}\gamma\,\mathrm{d}\Omega$ to the IC emission spectrum $\mathrm{d}P_t/\,\mathrm{d}\epsilon\,\mathrm{d}\Omega$ where ϵ and Ω are the energy and direction, respectively, of IC photons.[2] This presents no challenge from an analysis standpoint, since the diagnostics used so far (angular maps, beamed fraction, Ω_{50}) apply as well to photons as to particles. In fact, analyzing the photons themselves rather than the emitting particles enables a more precise measurement of kinetic beaming, a point that I now briefly elaborate.

In this section, I shall be interested in the energetic extent of kinetic beaming: the range of (high) photon or particle energies across which a strong energy-dependent anisotropy is evident. Now, in the Thomson regime, photons Comptonized by a particle of Lorentz factor γ attain increased energies by the factor γ^2 and are emitted along the particle's velocity vector within a cone of half-opening angle $1/\gamma$. So, in the ultrarelativistic limit ($\gamma \gg 1$), not only is the angular distribution of emission nearly identical to that of the radiating particles, but kinetic beaming actually encompasses a wider range of energies in the former than in the latter. Therefore, treating the IC emission rather than the radiating particles directly allows me to measure more precisely the energetic extent of kinetic beaming as it becomes small.

[2] The quantity $\mathrm{d}P_t/\,\mathrm{d}\epsilon\,\mathrm{d}\Omega$ is the instantaneous (at time t), lower layer, volume-integrated IC emission coefficient $j_{\rm IC}$, with $j_{\rm IC}$ as defined by Rybicki & Lightman (1979): $\mathrm{d}P_t/\,\mathrm{d}\epsilon\,\mathrm{d}\Omega = \int_{y<L_y/2} \mathrm{d}x\,\mathrm{d}y\,j_{\rm IC}(x, y, \epsilon, \mathbf{\Omega}; t)$.

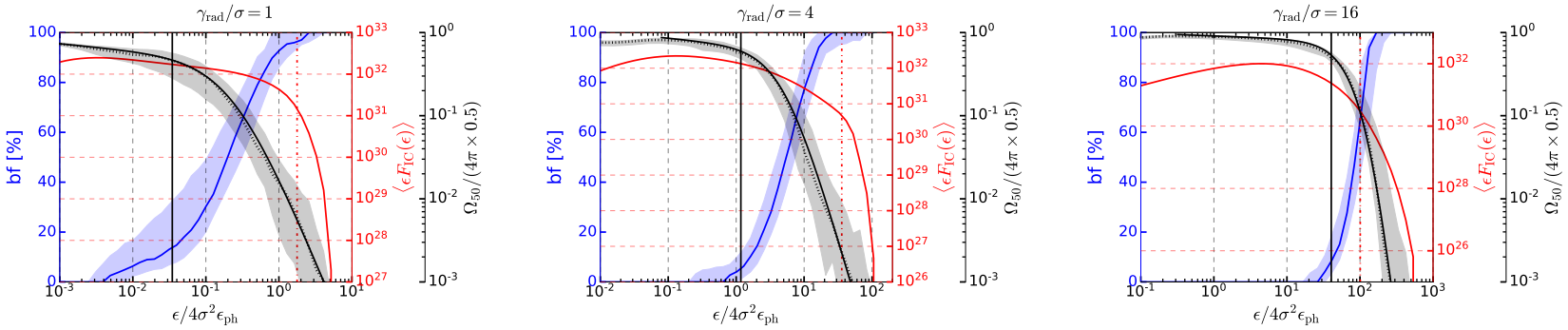

Figure 3.18: The time-averaged (over $1 \leq ct/L \leq 3$) electron IC emission spectrum (solid red), median beamed fraction (bf; solid blue), and median Ω_{50} curve (dotted black) for the simulations with $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1$ (left), 4 (middle), and 16 (right). Shaded regions indicate, for Ω_{50} and bf only, the middle 68 per cent of data. In each panel, the Ω_{50} data (dotted black lines with shaded regions serving as error bars) are fit using a smoothly broken power law parameterized as in equation (3.8). To avoid needing a more complicated fitting formula, only data following the last local maximum in the dotted Ω_{50} curve are fit. The fit is drawn as a solid black line on top of the data that are used. Solid black vertical lines indicate the onset of the best-fitting spectral break ($\epsilon_{\rm iso}$ in the text); dot-dashed red vertical lines show the spectral cutoff ($\epsilon_{\rm c}$ in the text) in the displayed emission spectrum. The horizontal axis is normalized to the maximum energy $4\sigma_{\rm c,0}^2\epsilon_{\rm ph}$ to which a particle of Lorentz factor $\sigma_{\rm c,0}$ can upscatter $\epsilon_{\rm ph}$-energy photons in the Thomson IC regime.

In Fig. 3.18, I illustrate beaming as a function of photon energy for three of the radiative runs ($\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 1, 4,$ and 16). In that figure, one sees that kinetic beaming – marked by rising bf and declining Ω_{50} at the highest energies – persists well beyond the first light crossing time in all simulations. Also, the energy range across which beaming is kinetic widens for the more strongly radiative simulations. Fig. 3.18 quantifies these observations by displaying two characteristic photon energies. The first is the energy $\epsilon_{\rm iso}$ above which beaming acquires pronounced spectral dependence – where the Ω_{50} curves begin to turn downward. The second is $\epsilon_{\rm c}$, the cutoff in the IC emission spectrum. The ratio of the cutoff $\epsilon_{\rm c}$ to the 'isotropic' energy $\epsilon_{\rm iso}$ characterizes the beamed range of photon energies.

These energy scales and their ratio are critical to my eventual quantitative portrait of kinetic beaming as a function of radiative efficiency. Therefore, here I expound upon the techniques I use to measure them and describe the trends in the measurements across the series of simulations. Let me begin with $\epsilon_{\rm c}$. Following the method of equation 37 in Bai et al. (2015) (see also Sironi et al., 2016; Hakobyan et al., 2019), I take

$$\epsilon_{\rm c} = \frac{\int {\rm d}\epsilon\, \epsilon^n F_{\rm IC}(\epsilon)}{\int {\rm d}\epsilon\, \epsilon^{n-1} F_{\rm IC}(\epsilon)}\,, \tag{3.5}$$

where $F_{\rm IC}(\epsilon) = \int {\rm d}\Omega\ {\rm d}P_t/\,{\rm d}\epsilon\,{\rm d}\Omega$ and n is empirically determined. I use $n = 4$ (higher values do not change the power-law scaling of $\epsilon_{\rm c}$ with cooling strength in Fig. 3.19). The result of this calculation is displayed in Fig. 3.18 for three reference simulations, and the functional dependence of $\epsilon_{\rm c}$ on $\gamma_{\rm rad,T}$ is displayed, for my entire series of simulations, in Fig. 3.19. (Note in that figure, as well as in Figs. 3.20 and 3.21, the horizontal axis is $\sigma_{\rm c,0}/\gamma_{\rm rad,T}$, which *increases* with stronger cooling.) I omit $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 64, \infty$ from Figs. 3.19–3.21 because those simulations exhibit secular growth in $\epsilon_{\rm c}$ throughout my analysis interval, $1 \leq ct/L \leq 3$, never reaching a quasi-steady state.

In Fig. 3.19, the large $\gamma_{\rm rad,T}$ (weak radiation-reaction) scaling of $\epsilon_{\rm c}$ with $\gamma_{\rm rad,T}$ significantly deviates from $\epsilon_{\rm c} = 4\gamma_{\rm rad,T}^2\epsilon_{\rm ph}$ (recall here that $\epsilon_{\rm ph}$ is the monochromatic energy of IC seed photons). This scaling is a special case of the more general result, $\epsilon_{\rm c} = 4\gamma_{\rm c}^2\epsilon_{\rm ph}$, which equals the maximum emitted photon energy from a particle at the cutoff Lorentz factor, $\gamma_{\rm c}$, in the particle distribution.

At large $\gamma_{\rm rad,T}$, I measure $\epsilon_{\rm c} < 4\gamma_{\rm rad,T}^2\epsilon_{\rm ph}$. This means that $\gamma_{\rm c} < \gamma_{\rm rad,T}$ and, perhaps, that diminished radiative efficiency allows a slower particle acceleration mechanism to dominate the highest energies.[3]

To see how this might work, suppose that such a mechanism operates and that the associated acceleration time $t_{\rm slow}$ for a particle to double its Lorentz factor $\gamma \to 2\gamma$ scales as $t_{\rm slow} \propto \gamma^{\zeta}$. Generally, one expects $\zeta > 1$ since the linear case $\zeta = 1$ corresponds to the fast time-scale $t_{\rm X} \propto \gamma$ associated with direct X-point acceleration [to be exact, $t_{\rm X}(\gamma) = \gamma m_e c^2/ce\beta_{\rm rec}B_0 \simeq 10\gamma/\omega_0$]. Furthermore, for self-consistency, this picture requires that the fast acceleration scaling $t_{\rm X} \propto \gamma$ does not continue indefinitely to high energies; otherwise there would never be a large-γ regime where $t_{\rm slow} < t_{\rm X}$. I therefore assume that beyond the intrinsic X-point maximum Lorentz factor $\gamma_{\rm X}$, the direct acceleration mechanism ceases. (Formally, $t_{\rm X} \propto \gamma$ for $\gamma < \gamma_{\rm X}$ and $t_{\rm X} = \infty$ otherwise.)

In this scenario, I adopt $\gamma_{\rm X} \sim 4\sigma_{\rm c,0}$. This has already been suggested by Werner et al. (2016) as a natural limit set by the size of elementary current layers in the plasmoid hierarchy. Moreover, $\gamma_{\rm rad,T} = 4\sigma_{\rm c,0}$ is, quite suggestively, where the scaling $\epsilon_{\rm c} \propto \gamma_{\rm rad,T}^2$ appears to break down in Fig. 3.19. Let me therefore tentatively assign $\gamma_{\rm X} = 4\sigma$ (appropriate for my simulations, but, as discussed in section 3.4, not necessarily the case in all astrophysical instances of reconnection).

Assuming radiative losses are weak enough that particles reach Lorentz factors exceeding $\gamma_{\rm X} = 4\sigma_{\rm c,0}$ (i.e., $\gamma_{\rm rad,T} > 4\sigma_{\rm c,0}$), reconnection is in a regime where $t_{\rm slow} < t_{\rm X}$ at the highest energies. Equating, therefore, $t_{\rm slow}$ to the Thomson IC cooling time $t_{\rm T}(\gamma) = 10\gamma_{\rm rad,T}^2/\gamma\omega_0 \propto \gamma_{\rm rad,T}^2/\gamma$ gives an expected cutoff in the particle distribution $\gamma_{\rm c} \propto \gamma_{\rm rad,T}^{2/(\zeta+1)}$ decided not by the competition between radiative losses and X-point acceleration, but by that between radiation and the putative slower acceleration mechanism. Importantly, the corresponding photon energy cutoff $\epsilon_{\rm c} = 4\gamma_{\rm c}^2\epsilon_{\rm ph} \propto \gamma_{\rm rad,T}^{4/(\zeta+1)}$ scales more gently than $\propto \gamma_{\rm rad,T}^2$, which is expected only if $\gamma_{\rm rad,T} \leq \gamma_{\rm X} = 4\sigma_{\rm c,0}$.

The above considerations are more than just an abstract hypothetical exercise. In fact, a slower energization process with $\zeta = 2$ – in which the Lorentz factors of high-energy particles generally follow $\gamma(t) \propto \sqrt{t}$ – has been identified by Petropoulou & Sironi (2018) and recently

[3] I have verified that $\gamma_{\rm c} < \gamma_{\rm rad,T}$, but do not present a corresponding plot.

elaborated by Hakobyan et al. (2021). In contrast to direct acceleration at reconnection X-points, this mechanism operates on particles inside plasmoids. There, particles are accelerated gradually due to conservation of their magnetic moments in the presence of a slowly growing magnetic field. The resultant scaling $\epsilon_{\rm c} \propto \gamma_{\rm rad,T}^{4/3}$ is not far from the apparent weak radiation-reaction scaling in Fig. 3.19.

Let me now use these ideas to construct a theoretical model that explains all of my $\epsilon_{\rm c}$ measurements. In this effort, I regard the cutoffs $\gamma_{\rm c}$ and $\epsilon_{\rm c}$ as dependent functions of the independent variable $\gamma_{\rm rad,T}$. The particular value $\gamma_{\rm rad,T} = \gamma_{\rm X} = 4\sigma_{\rm c,0}$ is special, because I assume that, for $\gamma_{\rm rad,T} > \gamma_{\rm X}$, X-point acceleration is subdominant. Instead, particles are primarily accelerated by a process similar to that of Petropoulou & Sironi (2018) and Hakobyan et al. (2021), which operates on a time-scale $t_{\rm slow} = C\gamma^2$. To fix the proportionality constant C, I require that the slow and fast acceleration time-scales, $t_{\rm slow}$ and $t_{\rm X}$, give equal cutoff Lorentz factors $\gamma_{\rm c}(\gamma_{\rm rad,T}) = \gamma_{\rm rad,T}$ at the transition value $\gamma_{\rm rad,T} = \gamma_{\rm X}$. This can be expressed as the condition $C\gamma_{\rm c}^2 = t_{\rm slow} = t_{\rm X} = t_{\rm cool,T} = 10\gamma_{\rm rad,T}^2/\gamma_{\rm c}\omega_0$, and yields, upon inserting $\gamma_{\rm c} = \gamma_{\rm rad,T} = 4\sigma_{\rm c,0}$, the result $t_{\rm slow} = 5\gamma^2/2\sigma_{\rm c,0}\omega_0$ ($C = 5/2\sigma_{\rm c,0}\omega_0$). For $\gamma_{\rm rad,T} < \gamma_{\rm X}$, the cutoff $\gamma_{\rm c}$ is set by the competition between X-point acceleration and radiative cooling and is given by $t_{\rm X}(\gamma_{\rm c}) = t_{\rm cool,T}(\gamma_{\rm c})$; for larger $\gamma_{\rm rad,T}$, cooling balances the slower acceleration mechanism and $\gamma_{\rm c}$ can be found from the condition $t_{\rm slow}(\gamma_{\rm c}) = t_{\rm cool,T}(\gamma_{\rm c})$. The IC photon cutoff energy $\epsilon_{\rm c} = 4\gamma_{\rm c}^2\epsilon_{\rm ph}$ is then a broken power law in $\gamma_{\rm rad,T}$:

$$\epsilon_{\rm c} = \begin{cases} \epsilon_{\rm c1} = 4\gamma_{\rm rad,T}^2\epsilon_{\rm ph} & \gamma_{\rm rad,T} \leq \gamma_{\rm X} = 4\sigma_{\rm c,0} \\ \epsilon_{\rm c2} = 4\gamma_{\rm rad,T}^2\epsilon_{\rm ph}(4\sigma_{\rm c,0}/\gamma_{\rm rad,T})^{2/3} & \text{otherwise}\,. \end{cases} \tag{3.6}$$

Should one wish to smooth the transition between $\epsilon_{\rm c} = \epsilon_{\rm c1}$ and $\epsilon_{\rm c} = \epsilon_{\rm c2}$, I find that the empirical formula

$$\frac{1}{\epsilon_{\rm c}^2} = \frac{1}{\epsilon_{\rm c1}^2} + \frac{1}{\epsilon_{\rm c2}^2} \tag{3.7}$$

describes the $\epsilon_{\rm c}(\gamma_{\rm rad,T})$ data quite well. Both this smoothed form and $\epsilon_{\rm c1}$ and $\epsilon_{\rm c2}$ individually are displayed in Fig. 3.19. Also shown is a power-law fit $\epsilon_{\rm c} \propto \gamma_{\rm rad,T}^{1.6}$ exhibiting a scaling intermediate

between ϵ_{c1} and ϵ_{c2}. Formally, equations (3.6) and (3.7), and a single power law all acceptably reproduce the $\epsilon_c(\gamma_{rad,T})$ data, but, based on the theoretical considerations discussed here, I suspect that a broken power law more accurately reflects the underlying physics.

To connect back to the discussion in chapter 2, the transition between ϵ_{c1} and ϵ_{c2} occurs at $\gamma_{rad,T} = \gamma_X$, which is precisely the transition between radiative regimes where the high-energy particles marginally attain radiative saturation ($\gamma_c = \gamma_{rad,T}$). When cooling is less efficient ($\gamma_{rad,T} > \gamma_X$), none of the particles energized by reconnection reaches the saturation condition, and, hence, $\gamma_c < \gamma_{rad,T}$. This is consistent with the fact that $\epsilon_{c2} < \epsilon_{c1}$ when $\gamma_{rad,T} > \gamma_X$. At higher cooling strengths, $\gamma_{rad,T} < \gamma_X$, particles energized near reconnection X-points quickly reach the radiative saturation energy $\gamma_{rad,T}$, yielding $\gamma_c = \gamma_{rad,T}$ and, thus, $\epsilon_{c1} = 4\gamma_{rad,T}^2\epsilon_{ph}$.

Let me now move on to describe my other important beaming-related energy scale ϵ_{iso}. As I did with ϵ_c, I first describe how I measure this quantity, interpreting the measurements thereafter. To calculate ϵ_{iso}, the first step is to fit a smoothly broken power law of the form

$$\Omega_{50}(\epsilon) = A\left(\frac{\epsilon}{\epsilon_{br}}\right)^{-p_1}\left\{\frac{1}{2}\left[1+\left(\frac{\epsilon}{\epsilon_{br}}\right)^{1/\Delta}\right]\right\}^{(p_1-p_2)\Delta} \tag{3.8}$$

to the Ω_{50} curve (`astropy.modeling.powerlaws.SmoothlyBrokenPowerLaw1D`, Astropy Collaboration et al., 2018, see Fig. 3.18). The parameters A, p_1, p_2, and ϵ_{br} are the scale, power-law indices, and spectral break of the fit. The parameter Δ controls the width of the break in the sense that equation (3.8) constitutes a pure power law with index p_1 (p_2) at energies below $\epsilon_{br}/10^{\Delta}$ (above $\epsilon_{br}10^{\Delta}$).

Once the parameters in (3.8) are determined, I take $\epsilon_{iso} \equiv \epsilon_{br}/10^{\Delta}$. This definition is empirically motivated from two observations. First, the Ω_{50} fits generally produce p_1 nearly flat and p_2 steep ($|p_1| \leq 0.2$ and $p_2 \geq 1.5$ across all fits). Second, in the (low-energy) p_1 segment of the curve, Ω_{50} hovers near isotropy [$\Omega_{50}(\epsilon) \simeq 0.5 \times 4\pi$]. Hence, at energies below $\epsilon_{br}/10^{\Delta}$, Ω_{50} is both energy-independent and isotropic (i.e., kinetic beaming is absent), but above $\epsilon_{br}/10^{\Delta}$, Ω_{50} begins to turn over, eventually declining precipitously with photon energy. Thus, the choice $\epsilon_{iso} = \epsilon_{br}/10^{\Delta}$

provides a good description for when beaming starts becoming kinetic, as intended.[4]

In order to build confidence in this $\epsilon_{\rm iso}$-extraction method, and to illustrate the utility of having two metrics of beaming (Ω_{50} and bf), I now discuss one subtlety associated with the $\epsilon_{\rm iso}$ measurement procedure. Namely, because 10^{Δ} lies between 2 and 8 across all $\Omega_{50}(\epsilon)$ fits (except for one broad transition $10^{\Delta} = 20$ in the $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 2$ simulation), and because $\epsilon_{\rm br}$ is rather large to begin with, the energy $\epsilon_{\rm br}10^{\Delta}$ signaling the end of the spectral break often falls near the rightmost edge of the $\Omega_{50}(\epsilon)$ data (or, for $\gamma_{\rm rad,T}/\sigma_{\rm c,0} = 2$, well beyond it). This means that p_2 and, to some extent, Δ and $\epsilon_{\rm br}$ are not necessarily well constrained. I deal with this difficulty in two ways. First, I do not rigorously study the p_2 measurements. I only report the lowest value $p_2 = 1.5$ (see above) to generally indicate the pronounced energy-dependence acquired by Ω_{50} beyond $\epsilon_{\rm iso}$. Secondly, for the fitted $\epsilon_{\rm br}$ and Δ values, which are used directly in my definition of $\epsilon_{\rm iso}$, I provide the following sanity check using my second metric of beaming: the beamed fraction. In particular, for each simulation, the location $\epsilon_{\rm iso} = \epsilon_{\rm br}/10^{\Delta}$ – despite being entirely determined from the Ω_{50} data – roughly coincides with photon energies where the beamed fraction slope increases most rapidly [where $\mathrm{d}^2(bf)/\,\mathrm{d}\epsilon^2$ is peaked]. This qualitative agreement between beaming metrics suggests that $\epsilon_{\rm iso}$ flags a real feature in the Ω_{50} curves, and is not merely an artifact of truncated high-energy Ω_{50} information. Fig. 3.20 displays the dependence of $\epsilon_{\rm iso}$ on $\sigma_{\rm c,0}/\gamma_{\rm rad,T}$.

The kinetic beaming range $\epsilon_{\rm c}/\epsilon_{\rm iso}$, measured from actual PIC simulations using the above techniques, allows me to quantify the energetic extent of kinetic beaming versus cooling strength. This is done in Fig. 3.21, where I present the value of $\epsilon_{\rm c}/\epsilon_{\rm iso}$ as a function of $\sigma_{\rm c,0}/\gamma_{\rm rad,T}$ (one data point per simulation). In both Figs. 3.20 and 3.21, I supply power-law fits to the data. This is not meant to indicate a robust theoretical description, but merely to characterize how quickly these quantities change from the weakly radiative ($\gamma_{\rm rad,T} \simeq \gamma_{\rm max} \gg \sigma_{\rm c,0}$) to a more strongly radiative ($\gamma_{\rm rad,T} \lesssim \sigma_{\rm c,0}$) regime.

Fig. 3.21 demonstrates a clear dependence of the kinetic beaming range on IC cooling

[4] Because I wish to flag the onset of kinetic beaming, I choose $\epsilon_{\rm iso} = \epsilon_{\rm br}/10^{\Delta}$ rather than $\epsilon_{\rm iso} = \epsilon_{\rm br}$. This means that, intentionally, the measured $\epsilon_{\rm iso}$ values indicated in Fig. 3.18 are often just before the Ω_{50} curves turn over, rather than in the middle of the spectral break.

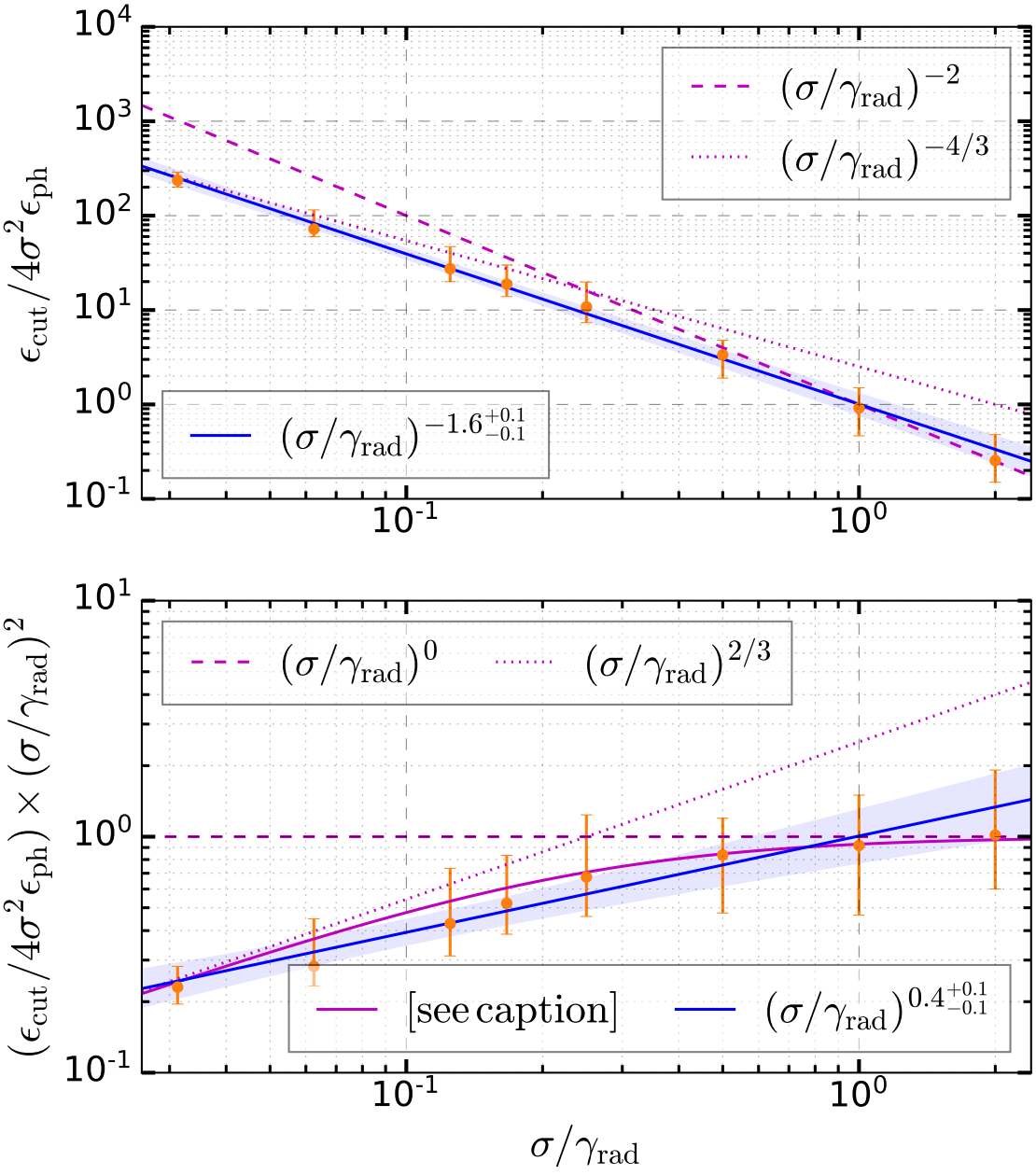

Figure 3.19: The cutoff photon energy displayed as a function of IC cooling strength using two different normalizations. In the upper panel, $\epsilon_{\rm c}$ is normalized to $4\sigma^2_{\rm c,0}\epsilon_{\rm ph}$; in the bottom panel, the normalization is $4\gamma^2_{\rm rad,T}\epsilon_{\rm ph}$. Each data point represents the median cutoff computed as a function of time (over $1 \leq ct/L \leq 3$) for a given simulation. Error bars indicate the middle 68 per cent of data. A power-law fit to the data is presented in blue, with shaded blue envelope indicating the uncertainty in the fit. The expected low-$\gamma_{\rm rad,T}$ scaling $\epsilon_{\rm c1} \sim \gamma^2_{\rm rad,T}$ is displayed in dashed magenta and the high-$\gamma_{\rm rad,T}$ scaling $\epsilon_{\rm c2} \sim \gamma^{4/3}_{\rm rad,T}$ in dotted magenta. Additionally, the lower panel shows the empirical fitting formula (3.7). A χ^2 goodness-of-fit test does not reject equations (3.6) and (3.7), or the single power law.

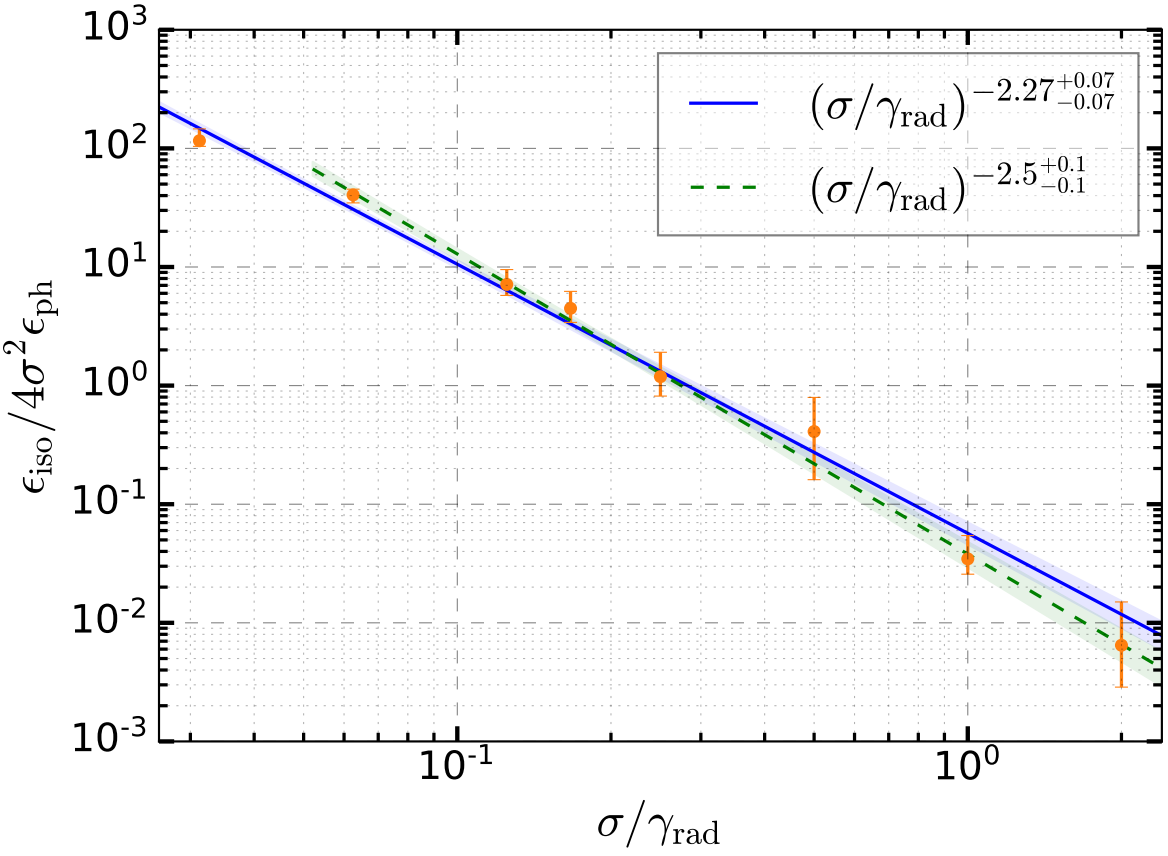

Figure 3.20: The isotropic photon energy $\epsilon_{\rm iso}$, which marks the transition to strong kinetic beaming in the Ω_{50} curve, as a function of IC cooling strength. Error bars indicate 68 per cent confidence intervals given by Markov chain Monte Carlo fits (Foreman-Mackey et al., 2013) to each simulation's $\Omega_{50}(\epsilon)$ curve (see Fig. 3.18) using a smoothly broken power law [as parameterized in equation (3.8)]. In this figure, the $\epsilon_{\rm iso}$ data are fit with unbroken power laws across both the full range of data $1/2 \leq \gamma_{\rm rad,T}/\sigma \leq 32$ (solid blue with shaded error envelope) and a restricted range $1/2 \leq \gamma_{\rm rad,T}/\sigma \leq 16$ (dashed green with shaded error envelope). The restricted fit excludes the non-radiative asymptotic behavior where $\epsilon_{\rm iso}$ is expected to clamp to $\epsilon_{\rm c}$.

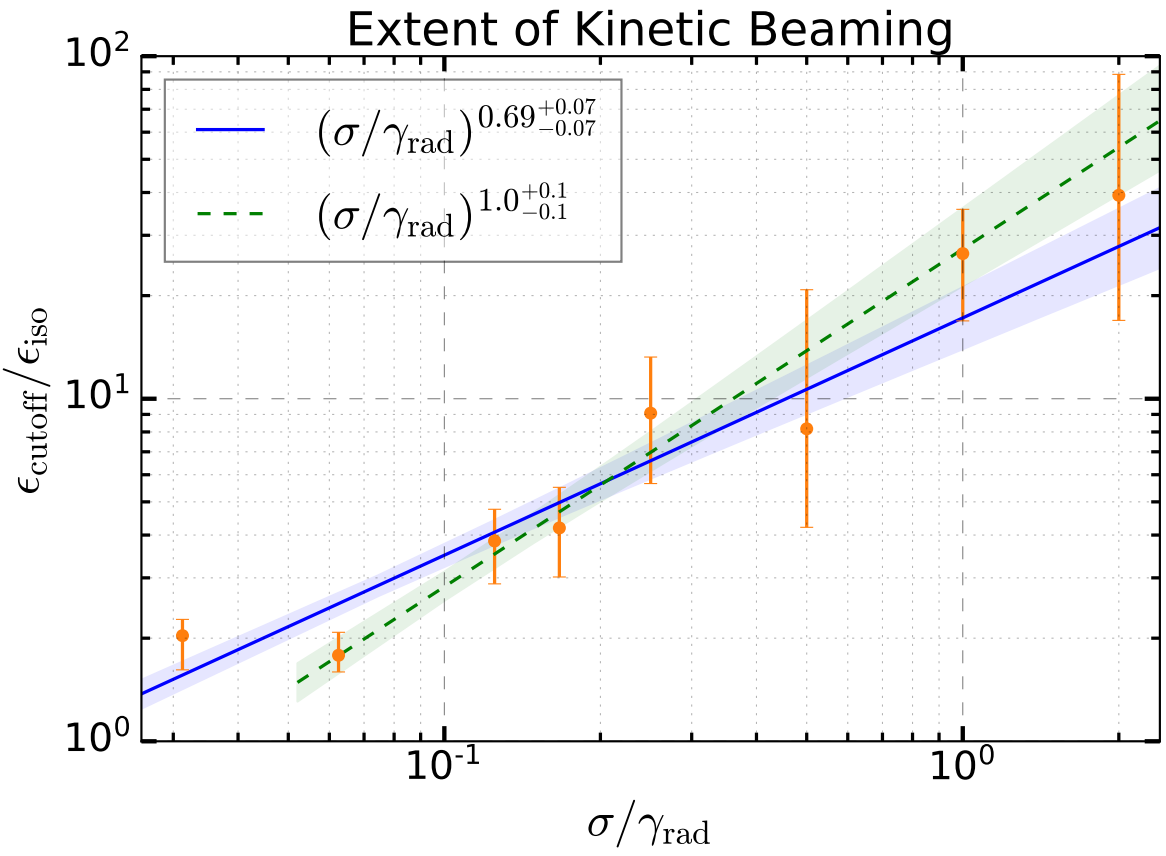

Figure 3.21: The 'range' of kinetic beaming as a function of $\sigma_{\rm c,0}/\gamma_{\rm rad,T}$, expressed as the ratio of cutoff photon energy to photon energy of beaming onset: $\epsilon_{\rm c}/\epsilon_{\rm iso}$. For each simulation, the photon energy $\epsilon_{\rm iso}$ is measured as the break energy in a broken power-law fit to the Ω_{50} curve, as in Fig. 3.18. Error bars presented are 'worst case', given by computing the ratio $\epsilon_{\rm c}/\epsilon_{\rm iso}$ at the extreme ends of the confidence intervals in Fig. 3.19 and Fig. 3.20. The solid blue power-law fit uses the entire range of data $1/2 \leq \gamma_{\rm rad,T}/\sigma_{\rm c,0} \leq 32$ and the dashed green power law uses the restricted range $1/2 \leq \gamma_{\rm rad,T}/\sigma \leq 16$. (As in Figs. 3.19 and 3.20, shaded blue and green envelopes indicate fit uncertainties.) Fits omitting additional data from the weakly radiative end (high $\gamma_{\rm rad,T}$, low $\sigma/\gamma_{\rm rad,T}$) yield power-law scalings consistent with the dashed-green line, suggesting that $\epsilon_{\rm c}/\epsilon_{\rm iso}$ may asymptote to a non-radiative limit of order unity for $\gamma_{\rm rad,T} \gtrsim 16\sigma_{\rm c,0}$.

strength. Moreover, this dependence may be stronger than is suggested by naively fitting a single power law to the entire data set $1/2 \leq \gamma_{\rm rad,T}/\sigma_{\rm c,0} \leq 32$. This is because $\epsilon_{\rm iso}$ can never exceed $\epsilon_{\rm c}$: particles and photons can be neither isotropic nor beamed at energies where none of them exist. Thus, the ratio $\epsilon_{\rm c}/\epsilon_{\rm iso}$, although strongly dependent on $\gamma_{\rm rad,T}$ when $\gamma_{\rm rad,T}$ approaches the strong cooling regime, must ultimately asymptote to unity in the non-radiative limit. According to the data, this occurs closer to $\gamma_{\rm rad,T} = 16\sigma_{\rm c,0}$ than to $\gamma_{\rm rad,T} = 32\sigma_{\rm c,0}$. To demonstrate this, I conduct a series of fits to the data in Fig. 3.21, of which I display only the first two. In each successive iteration, I remove the most weakly radiative simulation (that is, I keep all the data for the first fit, omit $\gamma_{\rm rad,T} = 32\sigma_{\rm c,0}$ for the second, omit $\gamma_{\rm rad,T} = 32\sigma_{\rm c,0}$ and $\gamma_{\rm rad,T} = 16\sigma_{\rm c,0}$ for the third, etc.). The power law becomes insensitive to this procedure once restricted to $\gamma_{\rm rad,T} \leq 16\sigma_{\rm c,0}$ and at that point exhibits nearly linear scaling.

This suggests that my simulation series captures an important transition in the range of beamed photon energies. For the milder radiative cooling ($\gamma_{\rm rad,T} \gtrsim 16\sigma_{\rm c,0}$), kinetic beaming is unobservable, manifesting itself nowhere in the distribution of emitted photons ($\epsilon_{\rm iso} \sim \epsilon_{\rm c}$). However, once the radiative efficiency is increased, kinetic beaming suddenly appears, and persists at late times across a sizable range of photon energies. This range increases throughout the entire set of $\gamma_{\rm rad,T}$ that is numerically accessible to me, and even surpasses one decade when $\gamma_{\rm rad,T} \lesssim 4\sigma_{\rm c,0} \sim \gamma_{\rm X}$.

Finally, I note that the simulation with $\gamma_{\rm rad,T} = \sigma_{\rm c,0}/2$, which (as noted in section 3.1) is problematic from the standpoint of cooling in the upstream plasma, is not essential to any of the findings in this section. In particular, excluding it leaves all best-fitting power-law scalings essentially unchanged. Because this simulation does not modify any of the overall trends, and indeed appears to fall in line with those trends, I have included it in Figs. 3.19–3.21.

3.3.4 System-size dependence

Before placing the numerical results of this chapter in an astrophysical context, let me briefly explore how they depend on the size of the computational box $L_x = L$. To do so, I report on a small

series of simulations with fixed radiation-reaction strength $\gamma_{\rm rad,T} = 4\sigma_{\rm c,0}$ and varying $L/\sigma_{\rm c,0}\rho_0 \in [80, 120, 160, 240, 320]$ (i.e., $\gamma_{\rm max}/\sigma_{\rm c,0} \in [8, 12, 16, 24, 32]$). All other parameters are the same as described in section 3.1 and summarized in Table 3.1.

I have chosen $\gamma_{\rm rad,T} = 4\sigma_{\rm c,0}$ for these simulations because, on the one hand, my results in the previous section indicate that this radiative efficiency is strong enough to yield a substantial range of beamed photon energies. On the other hand, the radiation is weak enough that the upstream plasma does not cool at all during the simulation [$t_{\rm cool,T}/(3L/c) = 22$ in equation (3.2)]. In the $\gamma_{\rm rad,T}$ scan, I tolerated a larger degree of upstream cooling because the expected error imparted to the main measured quantities – $\epsilon_{\rm iso}$, $\epsilon_{\rm c}$, and $\epsilon_{\rm c}/\epsilon_{\rm iso}$ – was both within the measurement error and swamped by the pronounced observed $\gamma_{\rm rad,T}$-dependence. However, here I need to be more strict. My goal is to demonstrate system-size insensitivity of the same beaming quantities, and, hence, I need to eliminate any upstream cooling effects that could selectively come into play at larger L.

For this series of simulations, Fig. 3.22 displays the values of $\epsilon_{\rm c}$, $\epsilon_{\rm iso}$, and the kinetic beaming range $\epsilon_{\rm c}/\epsilon_{\rm iso}$ calculated by the methods described in section 3.3.3. If kinetic beaming were to weaken with system size, presumably because of a diminished importance of kinetic scale phenomena, then one would expect the kinetic beaming range $\epsilon_{\rm c}/\epsilon_{\rm iso}$ to exhibit a downward trend with L. Instead, the data show that this quantity is consistent with being constant with L. This is in fact true not only of the ratio $\epsilon_{\rm c}/\epsilon_{\rm iso}$ but also of $\epsilon_{\rm c}$ and $\epsilon_{\rm iso}$ individually. If there is any non-constant trend at all, although this is not statistically significant, the kinetic beaming range *increases* with L (owing primarily to a decrease in $\epsilon_{\rm iso}$).

In this system-size scan, I have not explored larger L because (1) the existing results establish a compelling trend and (2) the cost of larger boxes is prohibitive: a factor of 2 increase leads to a factor of 2^3 additional computational expense. Pending a future study that more thoroughly elucidates system-size effects – both by going to larger L and by trialing additional values of $\gamma_{\rm rad,T}$ – I move forward assuming that the basic picture of kinetic beaming is captured by the main simulation series (with constant $L = 320\sigma_{\rm c,0}\rho_0$ and varying $\gamma_{\rm rad,T}$). In particular, I assume that those results can be extrapolated to astrophysically large systems, as suggested by the L-invariance

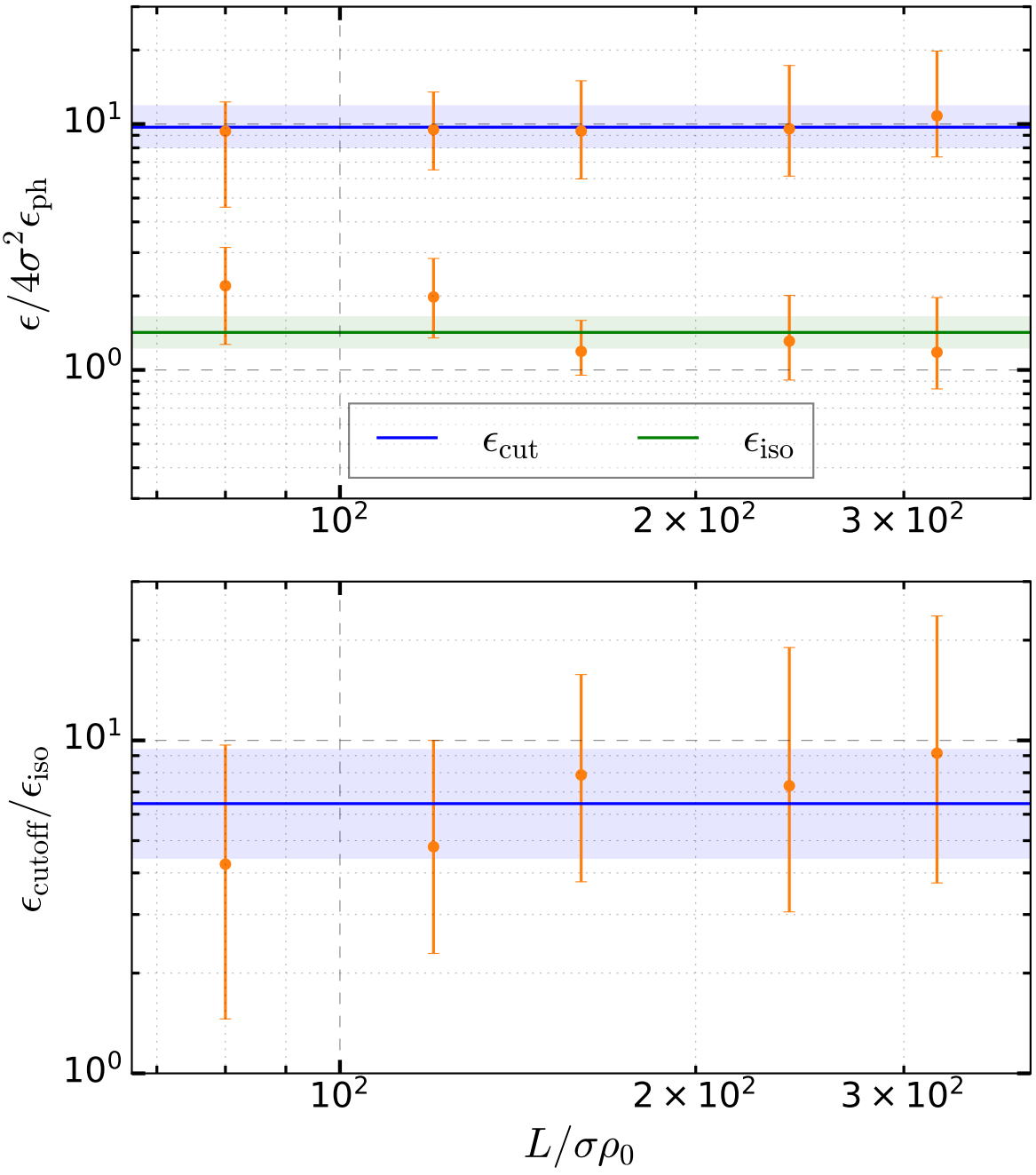

Figure 3.22: Top panel: $\epsilon_{\rm c}$ and $\epsilon_{\rm iso}$ measured using the same techniques as for Figs. 3.19 and 3.20 but for a series of simulations at uniform $\gamma_{\rm rad,T} = 4\sigma_{\rm c,0}$ and varying $L/\sigma_{\rm c,0}\rho_0 = [80, 120, 160, 240, 320]$. Bottom panel: the kinetic beaming range $\epsilon_{\rm c}/\epsilon_{\rm iso}$ for the same series of simulations. Solid lines indicate constant fits to each quantity, with shaded regions displaying 68 per cent confidence intervals. A χ^2 goodness-of-fit test does not reject any of these flat lines, showing that, given the error bars, the data are consistent with L-independence.

apparent in this system-size investigation.

3.4 Summary of kinetic beaming

In this section, I collect and summarize what features of kinetic beaming have been elucidated from the simulations presented above (particularly from section 3.3.3, the fundamental features of which, as suggested in section 3.3.4, may apply even to astrophysically large systems), with an eye towards extrapolating these findings to astrophysical situations. My remarks are mostly at a general level, independent of any particular class of astrophysical objects, but they set the stage for my specific application to TeV blazar flares in chapter 4.

Assuming the connection between beaming and rapid light-curve variability (described in the earlier works of Cerutti et al., 2012b, 2013, 2014a,b), my main result is that kinetic beaming requires a high degree of radiative efficiency in order to leave an observational signature. From my analysis of the kinetic beaming range $\epsilon_c/\epsilon_{iso}$ (Fig. 3.21), I have quantitatively discovered that beaming-induced variability should only exist in the highest energy spectral bands, where the emitting particles are near their radiatively imposed cutoff energy. An implication in the context of gamma-ray (e.g., blazar) flares is that, if high-cadence observations could be made in both bands, TeV flares should not have similar time-scale GeV counterparts (unless multiple radiative processes enable particles near radiative saturation to emit in several bands): kinetic beaming does not appear to be sustained over such a broad energy range.

My findings further suggest that, when radiative cooling is weak, kinetic beaming does not necessarily explain the total duration of rapid flares. While the transient initial phase of beaming seen in my weakly cooled simulations could influence the rising part of a flare, subsequent isotropization would likely prevent, in that case, the fast-rise fast-decay pattern characteristic of a collimated beam crossing the line of sight. On the other hand, kinetic beaming may well shape the entire temporal profile of rapid outbursts when radiative cooling is efficient. Then the energetic particle beams emit corresponding photon beams before diverging. Near the spectral cutoff, beaming is pronounced and kinetic, and the light curves in this band are expected to exhibit increasingly

dramatic variability at higher and higher energies.

Finally, I would like to state a conjecture that may broaden the scope of my results, enhancing the potential variety of astrophysical sources for which kinetic beaming may explain rapid flares. This conjecture concerns the range of $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$ for which kinetic beaming extends across an appreciable span of energies (for which $\epsilon_{\rm c}/\epsilon_{\rm iso}$ is sizable). Fig. 3.21 suggests that this range only exceeds about a decade when $\gamma_{\rm rad,T} \leq 4\sigma_{\rm c,0}$. This, however, does not necessarily mean that sources for which $\gamma_{\rm rad,T} \gg \sigma_{\rm c,0}$ should not exhibit kinetically beamed emission. In particular, I believe that the figure of merit for a kinetic beaming scenario is not whether $\gamma_{\rm rad,T}$ is of order several $\sigma_{\rm c,0}$ or less, but whether the actual energies achieved by particles are close to $\gamma_{\rm rad,T}$. As suggested by my results (Fig. 3.19) and argued in chapter 2 [section 2.2.1 near equation (2.8)], particles are only expected to reach Lorentz factors $\sim \gamma_{\rm rad,T}$ if they are accelerated by the fast X-point mechanism. Slower acceleration channels radiatively saturate at energies less than $\gamma_{\rm rad,T}$. Hence, if particles are somehow able to reach $\gamma_{\rm rad,T}$ even when $\gamma_{\rm rad,T} \gg \sigma_{\rm c,0}$, then they must have been accelerated and, consequently, beamed near an X-point.

Thus, what kinetic beaming really depends on is not whether $\gamma_{\rm rad,T} \lesssim$ (several)$\sigma_{\rm c,0}$, but whether $\gamma_{\rm rad,T}$ is comparable to or less than the maximum Lorentz factor $\gamma_{\rm X}$ achievable due to X-point acceleration. As hypothesized by this and prior studies (Werner et al., 2016; Petropoulou & Sironi, 2018; Hakobyan et al., 2021), $\gamma_{\rm X}$ may be of the order of $4\sigma_{\rm c,0}$ in reconnection set-ups like the one employed by me in this work, meaning that the requirement $\gamma_{\rm rad,T} \leq \gamma_{\rm X}$ in my case simplifies to $\gamma_{\rm rad,T} \leq 4\sigma_{\rm c,0}$. However, in alternative and more astrophysical situations, it may be possible for $\gamma_{\rm X}$ to circumvent this $4\sigma_{\rm c,0}$ limit. For example, many of the sources (e.g., pulsar wind nebulae and active galactic nuclei) for which kinetic beaming nicely explains a number of aspects of observed flares are also highly non-thermal emitters, even in their quiescent states. This suggests that the upstream plasma is itself non-thermal, possessing a long tail of already high-energy particles – very different from the thermal upstream conditions in my (and almost all other) simulations. Injected into the reconnection layer near an X-point, these high-energy particles may not be limited to Lorentz factors $4\sigma_{\rm c,0}$, and could indeed already exceed those Lorentz factors *before*

even experiencing the reconnection electric field.

Such particles could, in principle, reach Lorentz factors all the way up to the Hillas limit $eE_{\rm rec}l/m_ec^2$ where $E_{\rm rec} \simeq 0.1B_0$ is the reconnection electric field, and l is its potentially macroscopic coherence length. Importantly, the effective l should be larger for particles that are more energetic upon entering the reconnection region. Indeed, Werner et al. (2016) found that the characteristic $\gamma_{\rm X} \simeq 4\sigma_{\rm c,0}$ limit arises because cold thermal particles only experience direct linear acceleration in elementary current layers between the smallest-scale plasmoids. Before their energy can grow too large, these particles become magnetized and trapped inside small plasmoids flanking the elementary layer where they were originally accelerated. However, higher energy particles have much larger Larmor radii and therefore sample larger scale fields. They may potentially traverse multiple acceleration regions (spanning several elementary current layers) before finally becoming trapped inside a necessarily large (and therefore rare) plasmoid (cf. Cerutti et al., 2013).

These remarks motivate a future systematic study of the maximum Lorentz factors achievable by X-point acceleration, and of kinetic beaming, in the presence of alternative upstream conditions. However, when it comes to my astrophysical discussion in the next chapter, I simply assume that it is possible to achieve Lorentz factors $\gamma \sim \gamma_{\rm rad,T}$ (i.e., $\gamma_{\rm rad,T} \leq \gamma_{\rm X}$) even if $\gamma_{\rm rad,T} \gg \sigma_{\rm c,0}$, and I do not require $\gamma_{\rm rad,T} \leq 4\sigma_{\rm c,0}$ as a necessary condition for kinetic beaming.

At this point, it is clear that the astrophysical relevance of my findings is predicated on whether they survive under a number of non-trivial generalizations (e.g., to larger systems, to more realistic upstream conditions, and even to 3D), most of which are beyond the scope of this work. With that in mind, one should read my specific astrophysical remarks below not as predictions made by a robust and fully-fledged theory, but as provocative inferences that can be made should the fundamental character of my findings be preserved in real astrophysical systems. It is to those inferences that I now turn. Focusing specifically on TeV blazar flares, I examine whether kinetic beaming – as understood within the simplified framework of this study – can reasonably explain the extreme variability observed in these events in a manner that is consistent with, and possibly constrains, blazar radiative environments.

Chapter 4

Implications of Kinetic Beaming Results for Rapid TeV FSRQ Flares

This chapter is paired with the previous one in presenting material originally published by Mehlhaff et al. (2020). In it, I explore the astrophysical implications of the numerical results, concerning kinetic beaming in Thomson radiative reconnection, detailed in chapter 3. The particular astrophysical application I focus on is rapid TeV blazar (specifically FSRQ) flares. Before getting into the analysis, I provide some additional background concerning blazars – more extensive than that given in the Introduction (section 1.3).

Blazars comprise a class of active galactic nuclei (AGNs) with a relativistic jet pointed towards the Earth. Their observed spectra are generally quite broad, extending from the radio band to the gamma-rays, and characterized by two non-thermal humps. In models where the emission is leptonic, the lower energy (optical/UV/X-ray) hump is thought to stem from synchrotron radiation and the higher energy (gamma-ray) component from the IC process, whereby soft ambient photons are upscattered by relativistic particles (Begelman et al., 2008; Böttcher et al., 2013; Madejski & Sikora, 2016). The photons seeding IC scattering are typically supplied either by synchrotron emission from within the jet itself (synchrotron self-Compton, or SSC, models; e.g., Maraschi et al., 1992; Bloom & Marscher, 1996) or by various external sources (external IC models; e.g. Begelman & Sikora, 1987; Melia & Königl, 1989; Sikora et al., 1994). The most common view is that internal synchrotron emission seeds Compton scattering in BL Lacs – blazars characterized by a lack of strong emission lines – whereas external photons dominate the ambient radiation field in FSRQs, which are more luminous, showing strong emission lines and thermal radiation attributed to an

accretion disk (Tavecchio et al., 2011; Madejski & Sikora, 2016).

Here, I follow in the footsteps of many prior studies (e.g., Giannios et al., 2009; Nalewajko et al., 2011; Giannios, 2013; Sironi et al., 2015; Petropoulou et al., 2016; Werner et al., 2018; Christie et al., 2019, 2020; Ortuño-Macías & Nalewajko, 2020), positing relativistic magnetic reconnection as the driving mechanism behind blazar flares. Because it is difficult for a reconnection layer to sustain internal radiation energy density larger than the upstream (unreconnected) magnetic energy density (Beloborodov, 2017; Christie et al., 2019), the most natural emission model for reconnection-powered Compton-dominated flares – for which the IC spectral component dominates the synchrotron emission – is external IC (e.g., Christie et al., 2020). Conveniently, the objects with the most extreme observed Compton dominance, FSRQs, are also those objects that come pre-equipped with rich external radiation environments.[1] Not only is an externally illuminated reconnection region precisely the set-up addressed in the simulations of chapter 3, but in the case of the first sub-hour TeV FSRQ flare ever observed – that from PKS 1222+21 on 2010 June 17 (Aleksić et al., 2011; Tanaka et al., 2011) – a strong case has already been made by Nalewajko et al. (2012) that kinetic beaming was at play. For these reasons, I devote the main part of my analysis to understanding rapid TeV FSRQ flares, concentrating on the prototypical PKS 1222+21 outburst.

Two prominent sources of external background radiation in FSRQs like PKS 1222+21 are the broad-line region (BLR) and dusty torus (also called the hot dust region; HDR). The BLR contains gas subject to ionizing radiation from the AGN accretion disc, and it reprocesses this light into UV line emission (most prominently Ly α; Tavecchio & Ghisellini, 2008). The HDR is made of dust clouds radiantly heated by the AGN and producing thermal emission predominantly in the IR (Nenkova et al., 2008a,b).

As discussed by Nalewajko et al. (2012) (see also Aleksić et al., 2011; Tavecchio et al., 2011), the very high-energy (VHE; $\gtrsim 0.1\,\mathrm{TeV}$) radiation detected from PKS 1222+21 must have been

[1] A notable exception is the BL Lac PKS 2155-304, which produced a Compton-dominated flare in 2006 (Abramowski et al., 2012).

produced beyond the BLR, at least $\sim 0.5\,\mathrm{pc}$ from the AGN. Otherwise, it would have been absorbed while traversing the intense BLR radiation fields. At that distance, the extremely rapid variability time-scale requires the VHE flare to be fed by an unrealistically high energy density packed into a small fraction of the jet's cross section. However, via kinetic beaming, magnetic reconnection can achieve the same variability time-scale in a much larger space (Cerutti et al., 2013, 2014a,b). This, combined with the highly collimated emitting particles (this work and, originally, Cerutti et al., 2012b), relaxes the necessary energy density, enabling the flare to be fueled at the parsec-scale on a reasonable energy budget (the full details of this argument are presented by Nalewajko et al., 2012).

The numerical results of chapter 3 may be used to further constrain this general picture of kinetic beaming in reconnection-powered VHE FSRQ flares. Namely, one may stipulate that the putative beaming operates with the emitting particles near radiative saturation, and examine what new astrophysical insight may be derived from this requirement. Let me make this idea more quantitative. In my simulations, kinetic beaming was apparent in the late-time distribution of particles only when the highest-energy particles could achieve radiative saturation ($\gamma_{\mathrm{rad,T}} \lesssim 4\sigma_{\mathrm{c},0} \sim \langle\gamma_{\mathrm{X}}\rangle$) and for emitting particle Lorentz factors γ_{emit} above the isotropization threshold γ_{iso}. In the case of particularly efficient radiation ($\gamma_{\mathrm{rad,T}} = \sigma_{\mathrm{c},0}$), γ_{iso} was nearly an order of magnitude smaller than the radiative cutoff, $\gamma_{\mathrm{c}} \simeq \gamma_{\mathrm{rad,T}}$ (see Figs. 3.15 and 3.16 and their surrounding discussion). Thus, a necessary condition for kinetic beaming to yield an observable signature is that the emitting particles bear energies within a fairly narrow band given by

$$\gamma_{\mathrm{rad,T}} \geq \gamma_{\mathrm{emit}} \geq \gamma_{\mathrm{iso}} \sim \frac{\gamma_{\mathrm{rad,T}}}{10}\,. \tag{4.1}$$

As already discussed (section 3.4), despite that it was only in my simulations with $\gamma_{\mathrm{rad,T}} \lesssim 4\sigma_{\mathrm{c},0}$ that I measured an appreciable range of beamed particle and photon energies, I refrain from employing $\gamma_{\mathrm{rad,T}} \lesssim 4\sigma_{\mathrm{c},0}$ as a requirement for kinetic beaming in addition to (4.1). Rather, I assume that it is possible for X-points to accelerate particles up to the radiative limit $\gamma_{\mathrm{rad,T}}$ even if $\gamma_{\mathrm{rad,T}} \gg \sigma_{\mathrm{c},0}$, provided some particles in the reconnection inflow already possess relatively high Lorentz factors.

This could be the case, for example, if the upstream plasma is highly non-thermal, as may reasonably be expected from non-thermal quiescent blazar spectra.

I now present simple estimates to check whether equation (4.1) is satisfied by an external IC model for the PKS 1222+21 VHE flare. Here, unprimed quantities are evaluated in the observer's frame and primed quantities in the frame of the VHE emitting region, which is the assumed rest frame of the reconnection layer. (However, I leave particle Lorentz factors γ unprimed, though they are always evaluated in the reconnection frame.) These frames are connected by the emitting region bulk Lorentz factor Γ. For simplicity, I ignore the source redshift $z \simeq 0.4$, and assume the angle $\theta_{\rm obs}$ between the emitting region bulk velocity and the line of sight to be such that the Doppler factor $\delta = \{\Gamma[1 - (1 - 1/\Gamma^2)^{1/2} \cos\theta_{\rm obs}]\}^{-1}$ is approximately equal to Γ. I follow Nalewajko et al. (2012), adopting $\Gamma = 40$ (sufficient to render external IC radiation more efficient than SSC) and a fiducial comoving (unreconnected) magnetic field strength $B'_0 = 0.1\,{\rm G}$ typical at the parsec-scale.

Beyond the BLR, the likely dominant source of external photons illuminating the jet is the dusty torus. In the observer's frame, the torus radiation is approximately uniform and isotropic, with energy density $U_{\rm HDR} \simeq 9 \times 10^{-5}\,{\rm erg\,cm^{-3}}$ and typical photon energy $\epsilon_{\rm HDR} \simeq 0.3\,{\rm eV}$ (Nenkova et al., 2008a,b; Sikora et al., 2009; Malmrose et al., 2011; Tavecchio et al., 2011; Nalewajko et al., 2012). Particles IC upscattering these photons to the characteristic observed energy $\epsilon_{\rm obs} = 100\,{\rm GeV}$ (Aleksić et al., 2011) have approximate Lorentz factors

$$\gamma_{\rm emit,HDR} \sim \sqrt{\frac{\epsilon'_{\rm obs}}{\epsilon'_{\rm HDR}}} \sim \sqrt{\frac{\epsilon_{\rm obs}/\Gamma}{\epsilon_{\rm HDR}\Gamma}} \sim 1 \times 10^4 \left(\frac{\Gamma}{40}\right)^{-1} \left(\frac{\epsilon_{\rm obs}}{100\,{\rm GeV}}\right)^{1/2} \left(\frac{\epsilon_{\rm HDR}}{0.3\,{\rm eV}}\right)^{-1/2} . \tag{4.2}$$

I note that the Comptonization occurs in the marginal Klein-Nishina regime, since, in the rest frames of the scattering particles, the seed photon energies are close to the electron rest mass [cf. equation (3.1)]:

$$\frac{\gamma_{\rm emit,HDR}\epsilon'_{\rm HDR}}{m_{\rm e}c^2} \sim \sqrt{\frac{\epsilon_{\rm obs}\epsilon_{\rm HDR}}{(m_{\rm e}c^2)^2}} \sim 0.3 \left(\frac{\epsilon_{\rm obs}}{100\,{\rm GeV}}\right)^{1/2} \left(\frac{\epsilon_{\rm HDR}}{0.3\,{\rm eV}}\right)^{1/2} . \tag{4.3}$$

Klein-Nishina effects are even more important for hypothetical particles at the much-higher upper-limit Lorentz factor imposed by the HDR, which can be estimated via equation (2.8) as

$$\gamma_{\rm rad,HDR} = \sqrt{\frac{0.3eB_0'}{4\sigma_{\rm T}U_{\rm HDR}'}} \sim \sqrt{\frac{0.3eB_0'}{4\sigma_{\rm T}U_{\rm HDR}\Gamma^2}} \\ \sim 6\times10^6\left(\frac{\Gamma}{40}\right)^{-1}\left(\frac{B_0'}{0.1\,{\rm G}}\right)^{1/2}\left(\frac{U_{\rm HDR}}{9\times10^{-5}\,{\rm erg\,cm^{-3}}}\right)^{-1/2} . \tag{4.4}$$

Even though the numerical study presented earlier is confined to Thomson IC radiation, with the definition of $\gamma_{\rm rad,T}$ even relying on that fact (however, see the generalization of $\gamma_{\rm rad,T}$ to include Klein-Nishina effects in chapter 6), let me momentarily maintain equation (2.8) as a definition and suppose that the result (4.1) also holds in the deep Klein-Nishina regime. Then, because $\gamma_{\rm rad,HDR}$ and $\gamma_{\rm emit,HDR}$ are widely separated, the emitting particles are far below the expected isotropization threshold:

$$\gamma_{\rm emit} \sim \frac{\gamma_{\rm rad,HDR}}{400} \ll \frac{\gamma_{\rm rad,HDR}}{10} \sim \gamma_{\rm iso,HDR}\,. \tag{4.5}$$

Equation (4.5) suggests that the IC(HDR) process does not impose sufficient radiative losses for kinetic beaming to imprint itself upon the emitted photons. Rather, the radiating particles are expected to emit isotropically. Admittedly, the fact that the VHE photons are produced in the marginal Klein-Nishina regime challenges the applicability of chapter 3's numerical results. In part to more fully explore kinetic beaming in an IC(HDR) scenario, I devote subsequent chapters to analytically and numerically studying reconnection using a fully Klein-Nishina Compton cross section. For now, however, I disfavor a purely Thomson reconnection-based IC(HDR) model of rapid TeV FSRQ flares. Instead, I explore alternative sources of seed photons that, in the Thomson IC limit, better accommodate constraint (4.1).

To that end, I now conduct my analysis in the opposite direction. Rather than modeling the flare's radiative environment, testing afterward whether it is consistent with strongly cooled kinetic beaming, I start by assuming that beaming and efficient Thomson IC radiation operate together and see what this implies about the background photon population. In that spirit, I consider the properties of a hypothetical radiation field, characterized by its (assumed narrowly distributed)

photon energy $\epsilon_{\rm ph}$ and energy density $U_{\rm ph}$, that satisfies the main requirement $\gamma_{\rm emit,ph} \geq \gamma_{\rm iso,ph} \sim \gamma_{\rm rad,ph}/10$ in equation (4.1). This requirement can be recast, using equations (2.8) and (4.2), as the following inequality involving $\epsilon'_{\rm ph}$, $U'_{\rm ph}$, B'_0, and $\epsilon'_{\rm obs}$:

$$10 \geq \frac{\gamma_{\rm rad,ph}}{\gamma_{\rm emit,ph}} \simeq \sqrt{\frac{9}{40}\frac{B'_0}{B_{\rm c}}\frac{U_{\rm c}}{U'_{\rm ph}}\frac{\epsilon'_{\rm ph}}{\epsilon'_{\rm obs}}}\,, \tag{4.6}$$

where $B_{\rm c} = 8\pi e/3\sigma_T = 6.0 \times 10^{15}\,{\rm G}$ is the classical critical field and $U_{\rm c} = B_{\rm c}^2/8\pi$. For a fixed B'_0 and $\epsilon'_{\rm obs}$, saturation of this inequality defines a 1D space of radiation fields $U'_{\rm ph}(\epsilon'_{\rm ph}) \propto \epsilon'_{\rm ph}$ for which $\gamma_{\rm emit,HDR}$ is at the expected isotropization threshold. To pinpoint one candidate combination of $\epsilon'_{\rm ph}$ and $U'_{\rm ph}$, I require that $U'_{\rm ph} \geq U'_{\rm HDR}$, necessary for Comptonization of $U_{\rm ph}$-photons to dominate IC(HDR) emission, and implying $\gamma_{\rm rad,ph} \leq \gamma_{\rm rad,HDR}$. In turn, this yields a smallest permissible emitting particle Lorentz factor [via (4.4) and (4.6)] of $\gamma_{\rm emit,min} \sim \gamma_{\rm rad,ph}/10 \leq \gamma_{\rm rad,HDR}/10 \sim 6\times10^5$. The corresponding Compton seed photons have characteristic energies $\epsilon'_{\rm ph} \sim \epsilon'_{\rm obs}/\gamma^2_{\rm emit,min} \geq 7 \times 10^{-3}\,{\rm eV}$, which are small enough that the IC emission takes place safely in the Thomson regime: $\gamma_{\rm emit,min}\epsilon'_{\rm ph}/m_{\rm e}c^2 \sim 8 \times 10^{-3}$.

Now that I know what kind of seed photon population (i.e., combination of $\epsilon_{\rm ph}$ and $U_{\rm ph}$) is required for efficiently cooled kinetic beaming, I ask whether such a population can be realized in nature. As an affirmative plausibility argument, I briefly consider the possibility of a structured jet. However, since a detailed global flare model is beyond the scope of this study, I discuss only a subset of the possible parameters.

In particular, I consider a spine-sheath configuration, where the transverse jet structure consists of two regions: a central, fast-moving spine surrounded by a slower-moving sheath (Ghisellini et al., 2005; Tavecchio & Ghisellini, 2016; Tavecchio, 2017; Sikora et al., 2016). Photons produced in the sheath are blueshifted to the frame of the spine where they seed reconnection-powered Compton radiation.[2] I suggest that the sheath emission mechanism is synchrotron – perhaps due to a simultaneous but less luminous reconnection event – but, to maintain a simple and general discussion, avoid explicitly invoking this fact. I merely suppose that the spine, which contains the

[2] Strictly speaking, I could equally well consider an emitting blob plowing through an otherwise unstructured jet; the important part is the relative motion.

VHE-producing magnetic reconnection site, inherits the (fast) bulk Lorentz factor $\Gamma_> = \Gamma = 40$, while the sheath moves at more typical (slower) speeds: $\Gamma_< = 10$. The relative Lorentz factor between the two regions is $\Gamma_\mathrm{r} \simeq \Gamma_>/2\Gamma_< = 2$. To generalize my prior convention, primed quantities continue to refer to the reconnection (i.e., spine) rest frame and unprimed quantities (save particle Lorentz factors γ) to the observer's frame; I do not write anything down in the sheath frame itself.

In this set-up, the photon energy $\epsilon'_\mathrm{ph} \sim 7 \times 10^{-3}\,\mathrm{eV}$ corresponds to an observed seed photon energy $\epsilon_\mathrm{ph} \sim \epsilon'_\mathrm{ph}\Gamma_</\Gamma_\mathrm{r} \sim 0.03\,\mathrm{eV}$. This lies on the part of the broad-band spectral energy distribution presented by Tavecchio et al. (2011) attributed to the dusty torus (Malmrose et al., 2011; Tavecchio et al., 2011), and hence is consistent with observed spectral features. Let me see whether the inferred energy density U'_ph – which, again, is expected to be larger than U'_HDR in order for its Comptonization to dominate the flare – is also consistent with observations. To that end, I suppose the sheath luminosity peaks at $L_\mathrm{ph} = 10^{46}\,\mathrm{erg\,s^{-1}}$, similar to that observed in the broad-band spectrum near ϵ_ph (Tavecchio et al. 2011; however, these data are not simultaneous with the VHE flare). The spine-frame seed photon energy density is then $U'_\mathrm{ph} \sim \Gamma_\mathrm{r}^2 L_\mathrm{ph}/4\pi c\Gamma_<^4 R_\mathrm{sh}^2 \sim 0.2\,\mathrm{erg\,cm^{-3}}$, and, importantly, exceeds $U'_\mathrm{HDR} \sim \Gamma_>^2 U_\mathrm{HDR} \sim 0.1\,\mathrm{erg\,cm^{-3}}$, as required. In this estimate, I have assumed that the transverse size R_sh of the sheath photon source is comparable to that of the VHE-emitting region, $R_\mathrm{sh} \simeq R' \sim 10ct_\mathrm{var}\Gamma_> \simeq 2 \times 10^{-3}\,\mathrm{pc}$, implied by the TeV variability timescale $t_\mathrm{var} = 10\,\mathrm{min}$ (Aleksić et al., 2011) and enlarged by a factor of 10 due to kinetic beaming (cf. Cerutti et al., 2012b; Nalewajko et al., 2012; Ackermann et al., 2016). I have checked that the sheath thickness R_sh can be relaxed without substantial change to the model (reducing U'_ph much less severely than the naive expectation $U'_\mathrm{ph} \propto R_\mathrm{sh}^{-2}$).

Kinetic beaming is more viable in this spine-sheath model not so much because $U'_\mathrm{ph} > U'_\mathrm{HDR}$, but because $\epsilon'_\mathrm{ph} \ll \epsilon'_\mathrm{HDR}$. That is, the sheath photons appear much softer than the HDR photons in the spine frame, requiring higher energy particles for Comptonization to the VHE band – particles that are then quite strongly cooled. This is illustrated in Fig. 4.1, which presents the main results of this section through a radiative 'phase diagram'. The blue band in the figure, with lower border given by saturating inequality (4.6), designates seed photon populations conducive to efficiently ra-

diative kinetic beaming. The sheath photons occupy this band, and the HDR photons do not, mostly because of the large energetic disparity between the two populations. This is accentuated by the sheath motion, which enlarges the energy gap to $\epsilon'_{\rm HDR}/\epsilon'_{\rm ph} = (\Gamma_{>}\Gamma_{<}/\Gamma_{\rm r})\epsilon_{\rm HDR}/\epsilon_{\rm ph} = 200\epsilon_{\rm HDR}/\epsilon_{\rm ph}$ in the spine frame.

Hence, purely by invoking relative motion between a VHE-emitting spine and a seed-photon-emitting sheath, one may reconcile the seed photon population required by Thomson IC-mediated kinetic beaming with one that may plausibly be realized during an actual flare. Although I leave a detailed model to future work, I view the above remarks as illustrating the potential utility of a kinetic beaming framework that includes the new ingredient of strong radiative cooling. Whereas kinetic beaming on its own has previously been used to balance the energy budget in VHE FSRQ flares (Nalewajko et al., 2012), one now sees that the added radiative requirement may constrain possible emission mechanisms. Surprisingly, the most appropriate (Thomson IC) radiative model for rapid FSRQ flares may be one that does not rely on external structures at all, and, therefore, presents a potentially universal mechanism for the most rapid TeV flares in all blazars, even BL Lacs.

This concludes the part of this dissertation devoted to technically investigating Thomson IC radiative reconnection and its astrophysical applications – work originally presented by Mehlhaff et al. (2020). In the following chapter, I take a step back, considering the results presented thus far in light of the major, large-scale, guiding themes presented in the Introduction. Simultaneously, I motivate, and transition to, the study of Klein-Nishina radiative reconnection and its astrophysical applications – the second main component of this dissertation.

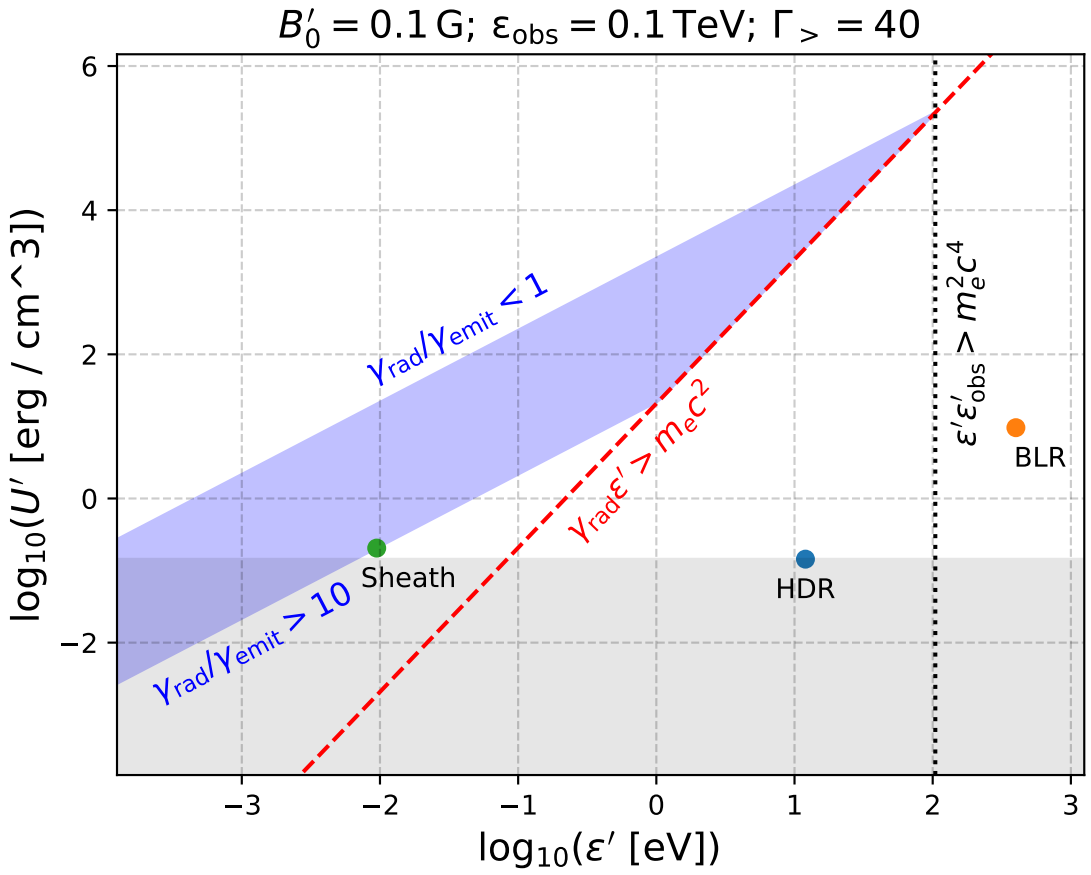

Figure 4.1: Radiative phase diagram for the PKS 1222+21 VHE flare. I project an otherwise high-dimensional parameter space onto the $U' - \epsilon'$ plane by fixing the parameters $\Gamma = \Gamma_> = 40$, $\Gamma_< = 10$, $B'_0 = 0.1\,\mathrm{G}$, and $\epsilon_{\mathrm{obs}} = 100\,\mathrm{GeV}$. Individual radiation fields are assumed to be monochromatic, occupying single points. To the right of the dotted vertical line, VHE photons pair-produce with their seed population. This excludes broad-line region illumination (see text), which I illustrate by adopting a characteristic BLR radiation field: $U_{\mathrm{BLR}} = 6 \times 10^{-3}\,\mathrm{erg\,cm^{-3}}$ and $\epsilon_{\mathrm{BLR}} = 10\,\mathrm{eV}$ (Sikora et al., 2009; Tavecchio et al., 2011; Nalewajko et al., 2012). Right of the dashed red line, particles scatter photons in the Klein-Nishina ($\epsilon'\gamma_{\mathrm{rad,T}} \geq m_e c^2$) regime. In contrast, the blue band indicates efficient but not unphysical ($\gamma_{\mathrm{rad,T}} \geq \gamma_{\mathrm{emit}} \geq \gamma_{\mathrm{iso}} \sim \gamma_{\mathrm{rad,T}}/10$) Thomson ($\epsilon'\gamma_{\mathrm{rad,T}} < m_e c^2$) IC cooling, and its lower border is given by saturating inequality (4.6). Lying outside this band, HDR illumination, at least ignoring Klein-Nishina effects, is probably unable to mediate observable kinetic beaming. A spine-sheath radiation field is more viable, and is above the shaded grey zone, where fields more tenuous than U'_{HDR} yield potentially unobservable IC output.

Chapter 5

Interlude

This chapter serves two purposes. First (section 5.1), I summarize my main results regarding Thomson radiative reconnection, including my numerical investigation of kinetic beaming and subsequent application to rapid TeV FSRQ flares (chapters 2 through 4). Second (section 5.2), I motivate – especially in light of the results obtained so far – studies of radiative reconnection facilitated by more general Klein-Nishina IC losses. These occupy the main focus of the rest of this dissertation (chapters 6-10). Throughout the discussion, I emphasize connections to major overarching themes of this work outlined in the Introduction and summarized in section 1.5.

5.1 Flaring due to kinetic beaming in Thomson radiative reconnection

So far, I have presented results from the first systematic investigation (Mehlhaff et al., 2020) of the role radiative cooling plays in the kinetic (energy-dependent) beaming of particles and their emission in collisionless relativistic magnetic reconnection. The simulations (chapter 3) undergirding this study, by scanning across the radiative saturation Lorentz factor, $\gamma_{\rm rad,T}$, cover a large portion of the Thomson radiative reconnection parameter space: they probe most of the possible regimes identified in chapter 2.

In agreement with prior studies (Cerutti et al., 2012b; Kagan et al., 2016), I measure definite and pronounced kinetic beaming during the early stages of all simulations, independent of cooling strength (see Figs. 3.13 and 3.15). When radiation is inefficient, kinetic beaming quickly gives way to a nearly isotropic distribution of particles and emission, as anticipated (Kagan et al., 2016;

Sironi et al., 2016; Yuan et al., 2016). However, as I demonstrate explicitly, kinetic beaming remains persistently observable when radiative cooling is efficient (see again Figs. 3.13 and 3.15), and may then extend across more than an order of magnitude in photon energies. Moreover, enhanced radiative losses increase the beamed range of photon energies (Figs. 3.18 and 3.21). In every case, kinetic beaming only remains apparent when the emitting particles have energies that are moderately close to (within an order of magnitude of) the radiatively imposed cutoff $\gamma_{\rm rad,T}$.

The underlying picture is a competition of time-scales: that over which the radiating particles cool and that over which they isotropize. As a generic side effect of impulsive X-point acceleration, particles are always initially beamed (Uzdensky et al., 2011; Cerutti et al., 2012a,b). However, in the limit of weak radiative losses, they produce most of their radiation after they have isotropized, and their initial collimation leaves no observable remnant. When cooling is strong, the opposite situation occurs: high-energy radiation comes only from beamed, recently accelerated particles. Reconnection focuses particles in both circumstances, but only in one is this focusing imprinted on the high-energy emission, manifesting itself as rapid variability along the observer's line of sight.

These findings have important consequences for rapid high-energy (HE; $\sim$ GeV) and very high-energy (VHE; $\sim$ TeV) astrophysical flares. At a very general level, they predict that rapid flares observed in a given spectral band (e.g., VHE) should not exhibit similar variability at energies that are lower by more than a couple orders of magnitude (e.g., HE), barring counterparts produced by the same particles simultaneously shining via multiple radiative processes. At a more detailed level, a kinetic beaming framework may constrain emission models in specific flaring systems, and I examine TeV FSRQ flares as an example. Analyzing the 2010 June 17 flare of PKS 1222+21, I find that a kinetic beaming origin of the rapid variability seems at odds with a picture of Thomson IC-scattered dusty torus photons dominating the TeV outburst (I do not consider models invoking Comptonization from inside the broad-line region, which are precluded by pair-production considerations; e.g., Aleksić et al., 2011). I postulate that an alternative spine-sheath model, wherein reconnection-energized particles upscatter photons originating in the outer fringes of the jet, may be more viable in a Thomson IC context. Importantly, this could potentially operate in both FS-

RQs and BL Lacs. Thus, the results of chapters 3 and 4 hint that a single radiative configuration may underlie rapid TeV flares from all blazars, regardless of their class.

More broadly, these inferences concretely illustrate an important overarching theme in this work (e.g., section 1.5). That is, solely by studying reconnection on its own as a fundamental plasma physical process – but from first principles and incorporating relevant extreme (e.g., relativistic and radiative) physical effects – one can derive constraints on the astrophysical systems that may host it. In this case, one sees that simulations of Thomson radiative reconnection may point to an alternative source of seed photons in the context of rapid TeV FSRQ flares than a dusty torus. Invoking a radiation bath stemming from within the jet itself is more compatible with kinetic beaming, which compellingly explains timing and energetics issues associated with these flares (Nalewajko et al., 2012).

However, there is an important caveat here that needs to be addressed in order to more confidently rule out, or perhaps restore viability to, irradiation of the reconnection region by a dusty torus. That issue is Klein-Nishina physics.

5.2 Klein-Nishina reconnection: the missing piece in the FSRQ puzzle

As demonstrated through equation (4.3), in an IC(HDR) scenario, the very high-energy photons observed from FSRQs are almost certainly produced in the marginal Klein-Nishina regime. Thus, in order to investigate Comptonization of dusty torus photons more carefully, one should study IC radiative reconnection not only in the Thomson limit of the IC process, but also in the more general Klein-Nishina regime.

At higher particle energies, Klein-Nishina effects suppress the IC cross section, reducing the radiative efficiency. Although one might expect this to inhibit kinetic beaming, increasing the cooling time-scale relative to the isotropization time-scale of beams, there are several subtle points that suggest the opposite – that Klein-Nishina physics actually promotes observable kinetic beaming. One such consideration (see section 8.1.2 for a more complete discussion) is that Klein-Nishina effects may lower the radiative cutoff energies associated with secondary acceleration mechanisms

relative to direct acceleration near reconnection X-points (despite the fact that the cutoffs for each acceleration mechanism are increased in an absolute sense). This may enable (e.g., Fig. 3.19) X-point-accelerated – and, hence, X-point-collimated – particles to shine across a broader span of high energies that are inaccessible to other energization channels, widening the range of beamed photon energies.

However, one should exercise some caution in attempting to generalize how beaming works in the Thomson IC limit to the Klein-Nishina regime. Klein-Nishina IC physics brings in an entirely new radiative process – pair production – that, as I show in the coming chapters, may significantly impact the qualitative aspects of reconnection. In particular, there is a wide range of reconnection parameters (chapter 6) such that photons Comptonized deeply in the Klein-Nishina IC regime are virtually guaranteed to collide with photons from the ambient seed population to produce high-energy electron-positron pairs inside the reconnection system. These pairs then go on to self-consistently regulate the reconnection dynamics (chapter 7), including, perhaps (section 8.1.2), kinetic beaming.

I would like to point out that, here again, one encounters a concrete example of a more general feature of extreme plasma astrophysics (section 1.5). That is, in this field, plasma phenomena often do not generalize straightforwardly from more traditional (in the sense of section 1.2; e.g., non-relativistic, non-radiative) regimes to the extreme regimes common to astrophysics. Instead, the latter regimes bring in fundamentally new physical processes that qualitatively modify the conventional picture. I have already shown in the first few chapters how kinetic beaming is profoundly affected by Thomson radiative physics, which enable it to remain persistently observable – something that does not happen in the non-radiative regime. So too, in the case of Klein-Nishina reconnection, it is imperative that one self-consistently investigates all of the physical processes at play – not just the modifications made to the radiative cooling of particles, but also how the produced pairs influence and change how reconnection behaves – in order to properly diagnose how kinetic beaming operates.

Besides its application to FSRQ jets, Klein-Nishina radiative reconnection also connects to

a broader set of astrophysical systems. In particular, in the high/soft states of black hole X-ray binaries, observations (Zdziarski et al., 2017) indicate that gamma-ray photons, if Comptonized from photons emitted at soft X-ray energies by the black hole accretion disk, are produced in the marginal Klein-Nishina regime. In fact, this can be shown very simply – similarly to equation (4.3) for an IC(HDR) scenario in FSRQ jets – and I do so in chapter 8. More detailed estimates (also in section 8), indicate that, in a scenario where the observed hard X-ray and gamma-ray emission is powered by reconnection in an accretion disk corona, reconnection is likely to be significantly shaped by Klein-Nishina effects (including pair production).

Hence, while FSRQs strongly motivate studying Klein-Nishina radiative reconnection on their own, black hole X-ray binaries and their associated accretion disk coronae provide another compelling astrophysical application. Therefore, for the remainder of this dissertation, I focus on elucidating reconnection in the Klein-Nishina radiative regime. However, unlike in the case of Thomson IC losses, very little work has been done on this topic. Therefore, instead of jumping straight into simulations, I dedicate most of the following chapters to describing basic features of Klein-Nishina reconnection analytically. In this way, I aim to lay a foundation for more detailed future study. After my analytic discussion in chapters 6-8, I describe preliminary results from a simulation of Klein-Nishina reconnection (chapter 10), including the substantial modifications to my group's PIC code, ZELTRON, that such simulations have required me to make.

Chapter 6

Klein-Nishina Radiative Reconnection Basics

Like chapter 2 for Thomson radiative reconnection, this chapter introduces the basics of Klein-Nishina radiative reconnection, and it is structured very similarly to chapter 2. First, I introduce the basic radiative cooling process, Klein-Nishina IC scattering, at the level of individual particles (section 6.1, the counterpart to section 2.1). Then (section 6.2, the counterpart to section 2.2), I discuss the collective effects that Klein-Nishina physics introduces in the context of reconnection. This includes ordering the fundamentally new energy scale (Lorentz factor) brought in by the radiative physics, $\gamma_{\rm KN}$, with the other energy scales (section 2.2) in the problem. The discussion of collective effects does not exactly mirror the corresponding discussion from chapter 2 (section 2.2) because the collective behavior in this case involves pair production between Comptonized photons and their parent seed photon population – and the consequent feedback this has on reconnection – but this process cannot happen, and, hence, has no parallel in, the Thomson limit. This chapter, along with the next two, make up the portion of this dissertation dedicated to work first presented by Mehlhaff et al. (2021).

6.1 Single-particle Klein-Nishina IC cooling

I begin the discussion of Klein-Nishina IC cooling with the following guiding intuition. One can readily infer that the quadratic Thomson scaling of the IC-scattered photon energy, $\epsilon_{\rm f} \sim \gamma^2 \epsilon$, where γ is the scattering particle's Lorentz factor and ϵ is the unscattered photon energy (see section 2.1), must break down at some point: a particle cannot emit a photon of greater energy

than its own $\gamma m_e c^2$ (ignoring the small initial energy ϵ). A new physical regime must take over when $\gamma^2\epsilon$ becomes of order $\gamma m_e c^2$. At that point, the particle can no longer radiate continuously; it loses an order-unity fraction of its energy in a single scattering event. Moreover, for even higher Lorentz factors, the photon energy can scale, at most, linearly with $\gamma m_e c^2$. The following analysis shows how these basic observations are borne out quantitatively.

At high energies, when the Thomson limit begins to break down, the seed photon energy ϵ becomes a dynamically important variable, influencing not just the spectrum of Comptonized photons, but also the power radiated by a particle $P_{\rm IC}(\gamma)$. To simplify my treatment in the presence of this complication, I specialize to a monochromatic distribution of background radiation

$$u_{\rm ph}(\epsilon) = U_{\rm ph}\delta(\epsilon - \epsilon_{\rm ph})\,. \tag{6.1}$$

To quantify the IC cooling domain, it is useful to define a critical Lorentz factor

$$\gamma_{\rm KN} \equiv \frac{m_e c^2}{4\epsilon_{\rm ph}} \tag{6.2}$$

and a Klein-Nishina parameter

$$q \equiv \frac{4\gamma\epsilon_{\rm ph}}{m_e c^2} = \frac{\gamma}{\gamma_{\rm KN}}\,. \tag{6.3}$$

Scattering particles suffer little recoil from individual photons when $q \ll 1$ ($\gamma \ll \gamma_{\rm KN}$): IC radiation proceeds in the Thomson regime. The opposite, deep Klein-Nishina limit is when $q \gg 1$ ($\gamma \gg \gamma_{\rm KN}$). The crossover point $q = 1$ ($\gamma = \gamma_{\rm KN}$) corresponds to setting the maximum Thomson emission energy, $4\gamma^2\epsilon_{\rm ph}$, equal to the Comptonizing particle energy, $\gamma m_e c^2$.

The Lorentz factor $\gamma_{\rm KN}$, like $\gamma_{\rm rad,T}$ for $U_{\rm ph}$ in the Thomson limit, is the fundamentally new energy scale introduced by Klein-Nishina physics. It serves as a proxy for the underlying physical parameter $\epsilon_{\rm ph}$. By ordering $\gamma_{\rm KN}$ with respect to the other fundamental energy scales discussed in section 2.2, $\gamma_{\rm rad,T}$, $\sigma_{\rm c,0}$, and $\gamma_{\rm max}$, one can determine what new radiative regimes of reconnection are accessible once Klein-Nishina effects have been added to the physical framework.

Before embarking on that task, however, I focus purely on the radiative physics (ignoring collective plasma effects), to build intuition for how individual particles experience IC losses. In the

presence of the seed photon distribution (6.1), the IC power radiated by a single particle becomes

$$P_{\rm IC}(\gamma) = P_{\rm T}(\gamma)\, f_{\rm KN}(\gamma/\gamma_{\rm KN})\,, \tag{6.4}$$

This is the same as the Thomson expression (2.1) but modified by the dimensionless function of q (cf. Jones 1968; Nalewajko et al. 2018; Appendix A)

$$f_{\rm KN}(q) = \frac{9}{q^3}\left[\left(\frac{q}{2}+6+\frac{6}{q}\right)\ln(1+q) - \frac{1}{(1+q)^2}\left(\frac{11}{12}q^3+6q^2+9q+4\right) - 2 + 2\,{\rm Li}_2(-q)\right]\,, \tag{6.5}$$

where $\rm Li_2$ is the dilogarithm. Figure 6.1 displays $f_{\rm KN}(q)$ together with its asymptotic large-argument limit

$$f_{\rm KN}(q \gg 1) \simeq (9/2q^2)\left[\ln(q) - 11/6\right] \tag{6.6}$$

and its approximate form (e.g., Moderski et al., 2005)

$$f_{\rm KN}(q) \simeq \frac{1}{(1+q)^{1.5}}\,, \tag{6.7}$$

which is roughly correct up to $q \sim 10^4$ (at $q \simeq 10^4$, the error reaches a factor of 3 and begins increasing rapidly). The function $f_{\rm KN}(q)$ tends to unity as q becomes small, as required in the Thomson-limit $P_{\rm IC}(\gamma) \to P_{\rm T}(\gamma)$. For large arguments, $f_{\rm KN}(q)$ falls off quadratically with a logarithmic correction: the scattering cross section is suppressed in the deep Klein-Nishina limit.

Let me now examine the 'discreteness' of radiative losses as q increases through unity. To do so, I note that, similar to the radiated power $P_{\rm IC}(\gamma)$, the rate $R_{\rm IC}(\gamma)$ at which a single electron or positron scatters photons distributed according to (6.1) can be written as the corresponding Thomson rate, $R_{\rm T}(\gamma) = R_{\rm T} = c\sigma_{\rm T}U_{\rm ph}/\epsilon_{\rm ph}$, times a dimensionless function:

$$R_{\rm IC}(\gamma) = R_{\rm T}\, g_{\rm KN}(\gamma/\gamma_{\rm KN})\,, \tag{6.8}$$

where $g_{\rm KN}(q)$ reads (Appendix A)

$$g_{\rm KN}(q) = \frac{3}{2q^2}\left[\left(q+9+\frac{8}{q}\right)\ln(1+q) - \frac{1}{1+q}\left(\frac{q^2}{2}+9q+8\right) + 4\,{\rm Li}_2(-q)\right] \tag{6.9}$$

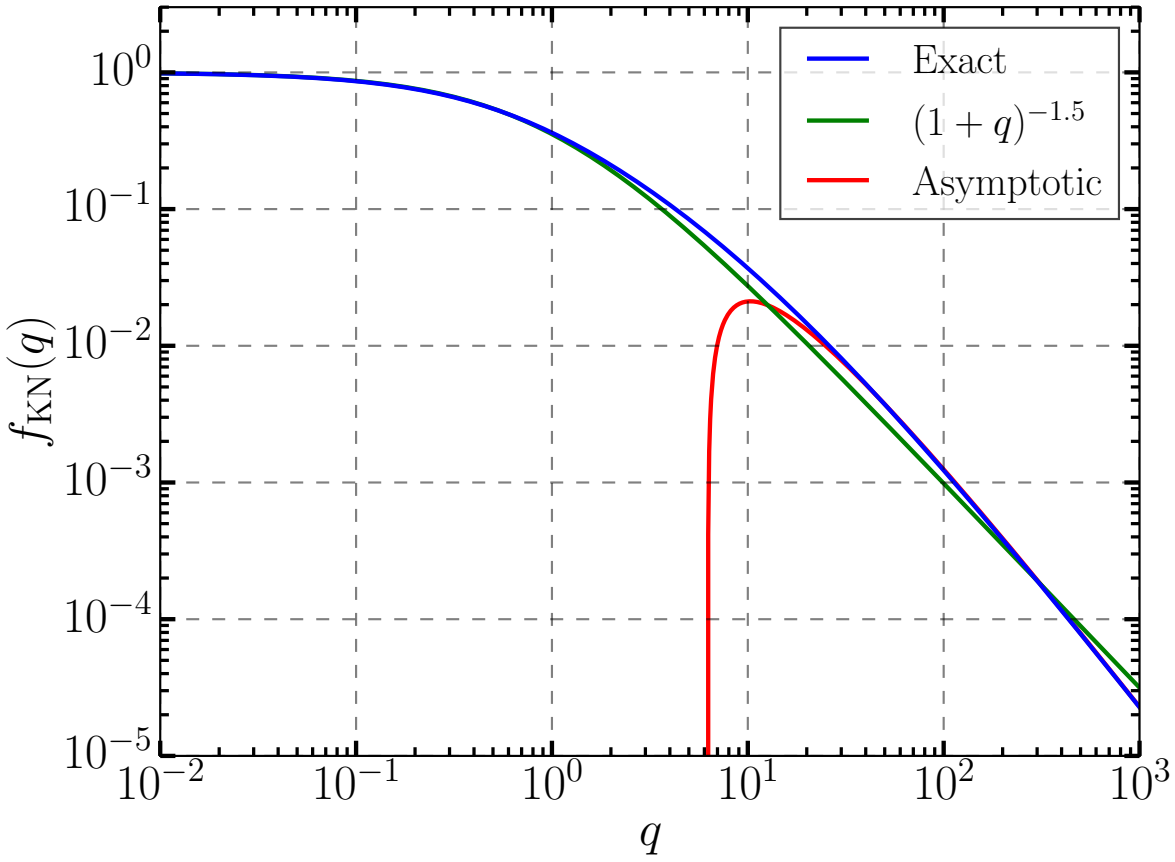

Figure 6.1: The function $f_{\rm KN}(q)$ as defined in equation (6.5) in blue, its asymptotic form (6.6) in red, and its approximation (6.7) in green.

and has asymptotic form

$$g_{\rm KN}(q \gg 1) \simeq \frac{3}{4q}\left[2\ln(q) - 1\right] . \tag{6.10}$$

Using (6.8), one can write down the average photon energy $\langle\epsilon\rangle(\gamma)$ emitted by a particle with Lorentz factor γ:

$$\langle\epsilon\rangle(\gamma) \equiv \frac{P_{\rm IC}(\gamma)}{R_{\rm IC}(\gamma)} = \left[\frac{\gamma}{3\gamma_{\rm KN}}\frac{f_{\rm KN}(\gamma/\gamma_{\rm KN})}{g_{\rm KN}(\gamma/\gamma_{\rm KN})}\right]\gamma m_e c^2 . \tag{6.11}$$

The quantity in square brackets here is the *inelasticity* $\langle\epsilon\rangle(\gamma)/\gamma m_e c^2$, the typical fraction of a particle's energy lost in a single scattering encounter (Moderski et al., 2005). Using the small-argument limits $f_{\rm KN}(q \ll 1) \to 1$ and $g_{\rm KN}(q \ll 1) \to 1$, one may read off the Thomson inelasticity

$$\lim_{\gamma \ll \gamma_{\rm KN}} \frac{\langle\epsilon\rangle(\gamma)}{\gamma m_e c^2} = \frac{\gamma}{3\gamma_{\rm KN}} \ll 1 , \tag{6.12}$$

or, more commonly,

$$\lim_{\gamma \ll \gamma_{\rm KN}} \langle\epsilon\rangle(\gamma) = \frac{\gamma^2 m_e c^2}{3\gamma_{\rm KN}} = \frac{4}{3}\gamma^2 \epsilon_{\rm ph} . \tag{6.13}$$

Similarly, plugging in the asymptotic forms for $f_{\rm KN}$ and $g_{\rm KN}$ verifies that the inelasticity approaches unity as q is taken to infinity:

$$\lim_{\gamma \to \infty} \frac{\langle\epsilon\rangle(\gamma)}{\gamma m_e c^2} = \lim_{\gamma \to \infty} \frac{\ln(\gamma/\gamma_{\rm KN}) - 11/6}{\ln(\gamma/\gamma_{\rm KN}) - 1/2} = 1 . \tag{6.14}$$

However, the limiting value is approached quite slowly, for the ratio in (6.14) is only $\simeq 1$ when $\ln(\gamma/\gamma_{\rm KN}) \gg 11/6$. In fact, $\langle\epsilon\rangle(\gamma)$ does not surpass $3\gamma m_e c^2/4$ until $\gamma \gtrsim 300\gamma_{\rm KN}$. Nevertheless, the inelasticity does obtain a value *of order* unity for much more modest q. For example, $\langle\epsilon\rangle(\gamma) \geq \gamma m_e c^2/4$ when $\gamma \gtrsim 2\gamma_{\rm KN}$.

Both of these effects – the inelasticity slowly approaching, but rapidly rising to the vicinity of, unity for $q \gtrsim 1$ – as well as the function $g_{\rm KN}(q)$ and its asymptotic form (6.10), are plotted in Fig. 6.2. In the figure, one sees that $\langle\epsilon\rangle(\gamma) \simeq \gamma m_e c^2/2$ for a wide range of q. Therefore, when I later (in chapter 7) need $\langle\epsilon\rangle$ for estimates, I adopt $\langle\epsilon\rangle(\gamma) \simeq \gamma m_e c^2/2$ rather than $\langle\epsilon\rangle \simeq \gamma m_e c^2$ whenever $q \gtrsim 1$. The latter becomes a more accurate approximation than the former for $q \gtrsim 300$,

but it is unclear that the astrophysical systems I attempt to model contain particles at such high energies. (However, even if $q \gg 300$ in some systems, the distinction here between factors of order unity is well within the uncertainty of all of the estimates in this work.) Furthermore, my model (chapter 7) is mostly concerned with $q < 280$ [equation (7.24)].

I have now verified the qualitative expectations for the deep Klein-Nishina regime discussed at the beginning of this section: radiative losses become discrete when $q \gtrsim 1$, with particles losing an appreciable fraction of their energy to single photons. Moreover, the scattered photon energy scales approximately linearly with the pre-collision energy of the particle (rather than quadratically, as in the Thomson limit).

Next, because I am concerned primarily with IC radiation in this work, I consider the circumstances in which synchrotron losses may be neglected. For an isotropic particle pitch-angle distribution, the average synchrotron power radiated per particle is

$$P_{\rm s} = \frac{4}{3}\sigma_{\rm T} c \beta^2 \gamma^2 U_{\rm B} \,. \tag{6.15}$$

Note that this is the same as the Thomson IC power (2.1) but with the ambient radiation energy density $U_{\rm ph}$ replaced by the magnetic field energy density $U_{\rm B} \sim B_0^2/8\pi$ (I approximate the magnetic field strength throughout the reconnection system by its upstream value B_0). Equation (6.15) gives a total (IC + synchrotron) radiated power per particle

$$\begin{aligned} P_{\rm tot}(\gamma) &= P_{\rm IC}(\gamma) + P_{\rm s}(\gamma) \\ &= P_{\rm IC}(\gamma) \left(1 + \frac{U_{\rm B}}{U_{\rm ph} f_{\rm KN}(\gamma/\gamma_{\rm KN})} \right) . \end{aligned} \tag{6.16}$$

Clearly, IC losses dominate if

$$f_{\rm KN}(q) > \frac{U_{\rm B}}{U_{\rm ph}} \,. \tag{6.17}$$

Because $f_{\rm KN}(q) \leq 1$, this criterion can only be met for systems whose ambient radiation energy density exceeds the magnetic field energy density. And, importantly, because Klein-Nishina effects begin to suppress IC cooling for $\gamma > \gamma_{\rm KN}$, even when $U_{\rm ph} \gg U_{\rm B}$, there is always a high-energy

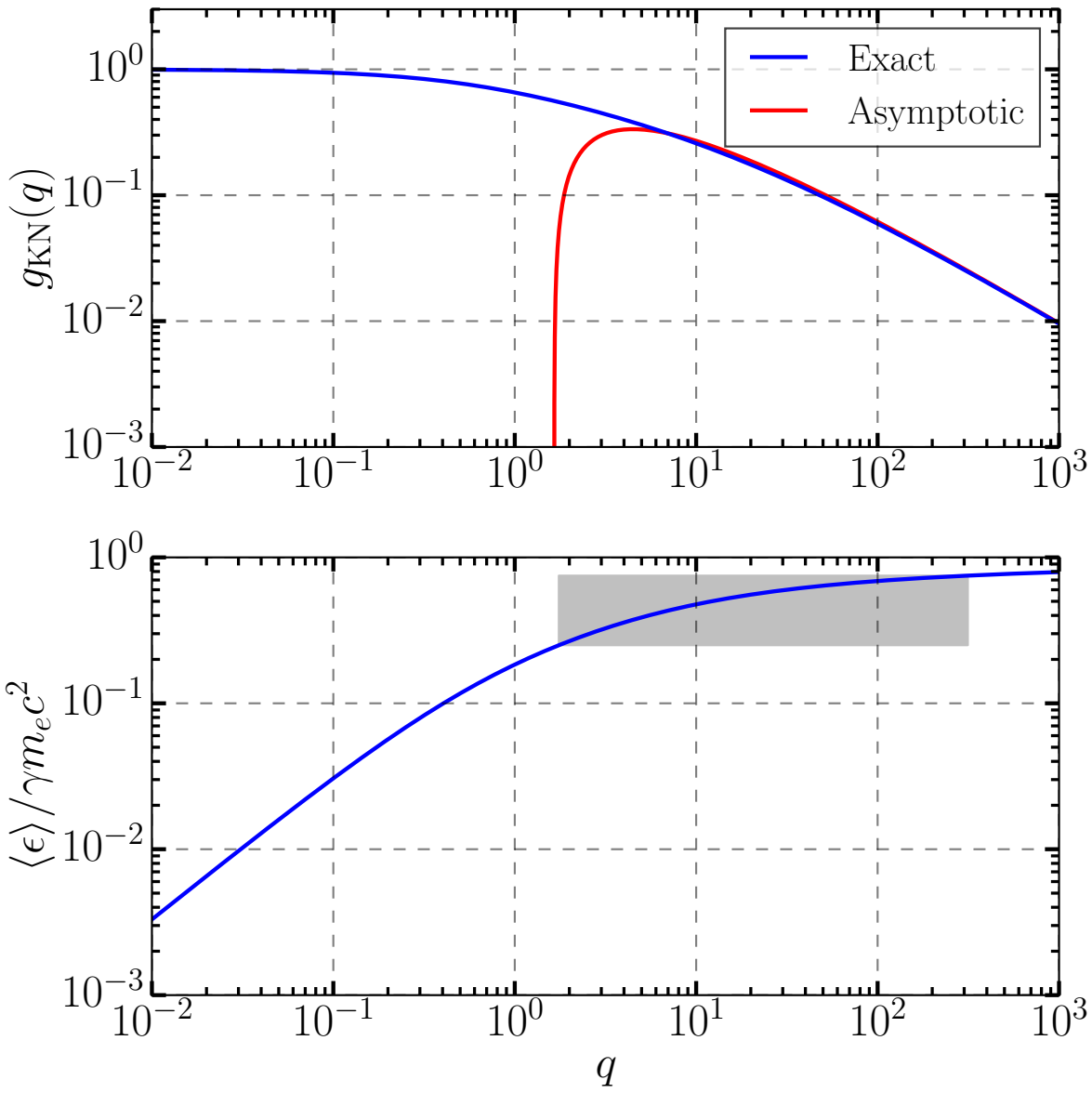

Figure 6.2: Top panel: the function $g_{\rm KN}(q)$ [equation (6.9)] in blue and its asymptotic form (6.10) in red. Bottom panel: the inelasticity $\langle\epsilon\rangle(\gamma)/\gamma m_e c^2 = q f_{\rm KN}(q)/3 g_{\rm KN}(q)$ [equation (6.11)]. The shaded region shows the appreciable range $q \in [1.75, 315]$ over which $\langle\epsilon\rangle$ is between $\gamma m_e c^2/4$ and $3\gamma m_e c^2/4$ (i.e., where it is close to $\gamma m_e c^2/2$). This indicates that, although $\langle\epsilon\rangle(\gamma)$ eventually approaches its asymptotic limit of $\gamma m_e c^2$, it does so rather slowly.

Lorentz factor $\gamma_{\rm s}$ above which synchrotron losses dominate. Using the approximate form $f_{\rm KN}(q) \simeq (1+q)^{-1.5}$ [equation (6.7)], one has (cf Moderski et al., 2005)

$$\gamma_{\rm s} \simeq \gamma_{\rm KN} \left[\left(\frac{U_{\rm ph}}{U_{\rm B}} \right)^{2/3} - 1 \right] . \tag{6.18}$$

Thus, neglecting synchrotron losses is justified when $\gamma_{\rm s}$ exceeds the highest Lorentz factors reached in the system. I assume that this is indeed the case for this chapter and the next. I return to discuss the effects of finite $\gamma_{\rm s}$ as an effective limitation of my analysis in chapter 8.

6.2 Klein-Nishina IC effects on collective plasma behavior

I now examine how the collective reconnection dynamics are influenced by Klein-Nishina radiation-reaction. I focus especially on differences from the case of purely Thomson radiative cooling (previously detailed in chapter 2).

Previously (in the Thomson regime), the highest Lorentz factor to which a particle could be accelerated was $\gamma_{\rm rad,T}$. However, from equation (6.4), radiative losses are suppressed once q exceeds unity. This enables acceleration beyond $\gamma_{\rm rad,T}$, and the definition of the radiative cutoff Lorentz factor can be generalized to include this effect. By equating the force from the reconnection electric field $0.1eB_0$ to the (Klein-Nishina) Compton radiation reaction force $P_{\rm IC}(\gamma)/c$, one may define a generalized cutoff $\gamma_{\rm rad,IC}$ through

$$1 \equiv \frac{\gamma_{\rm rad,IC}^2}{\gamma_{\rm rad,T}^2} f_{\rm KN} \left(\frac{\gamma_{\rm rad,IC}}{\gamma_{\rm KN}} \right) = \frac{\gamma_{\rm rad,IC}^2}{\gamma_{\rm rad,T}^2} f_{\rm KN} \left(\frac{\gamma_{\rm rad,IC}}{\gamma_{\rm rad,T}} \frac{\gamma_{\rm rad,T}}{\gamma_{\rm KN}} \right) . \tag{6.19}$$

The second equality explicitly shows that the ratio $\gamma_{\rm rad,IC}/\gamma_{\rm rad,T}$ is determined solely by the ratio $\gamma_{\rm rad,T}/\gamma_{\rm KN}$. Thus, $\gamma_{\rm rad,IC}$ is a derived scale; it is fixed by the other radiative Lorentz factors $\gamma_{\rm KN}$ and $\gamma_{\rm rad,T}$ (or, equivalently, through the physical parameters $\epsilon_{\rm ph}$ and $U_{\rm ph}$).

Equation (6.19) can always be satisfied for a finite $\gamma_{\rm rad,IC}$, and a numerical solution to the equation is displayed in Fig. 6.3. When $\gamma_{\rm rad,T} \ll \gamma_{\rm KN}$, the equation is satisfied by $\gamma_{\rm rad,IC} \simeq \gamma_{\rm rad,T}$ because, in that case, $f_{\rm KN}(\gamma_{\rm rad,T}/\gamma_{\rm KN} \ll 1) \simeq 1$. However, when $\gamma_{\rm rad,T}$ becomes greater than $\gamma_{\rm KN}$, the cutoff $\gamma_{\rm rad,IC}$ becomes a rapidly increasing function of $\gamma_{\rm rad,T}/\gamma_{\rm KN}$. In fact, in the limit $\gamma_{\rm rad,T} \gg$

$\gamma_{\rm KN}$, for which the large-argument approximation to $f_{\rm KN}$ [equation (6.6)] applies, $\gamma_{\rm rad,IC}$ grows super-exponentially:

$$\gamma_{\rm rad,IC} \simeq \gamma_{\rm KN} \exp\left[\frac{2}{9}\left(\frac{\gamma_{\rm rad,T}}{\gamma_{\rm KN}}\right)^2 + \frac{11}{6}\right] . \tag{6.20}$$

As shown in Fig. 6.3, this limiting form gives a good approximation to $\gamma_{\rm rad,IC}$, even when $\gamma_{\rm rad,T}$ only slightly exceeds $\gamma_{\rm KN}$. I note that the rapid transition between the $\gamma_{\rm KN} \gg \gamma_{\rm rad,T}$ and $\gamma_{\rm KN} \ll \gamma_{\rm rad,T}$ limits is smoothed out in the presence of an extended (e.g., power-law) distribution of ambient radiation. Then, $\gamma_{\rm rad,IC}$ does not grow super-exponentially until $m_e c^2/4\epsilon_{\rm low} \ll \gamma_{\rm rad,T}$, where $\epsilon_{\rm low}$ is the energy of the softest ambient photons.

Hence, when $\gamma_{\rm KN}$ becomes less than $\gamma_{\rm rad,T}$, Klein-Nishina physics inhibits radiative cooling from competing with rapid acceleration near reconnection X-points. It is then highly likely that the cutoff energy for X-point acceleration is set intrinsically rather than by radiative cooling (one expects $\gamma_{\rm X} \ll \gamma_{\rm rad,IC}$).

I now frame the super-exponential divergence in $\gamma_{\rm rad,IC}$ from a different perspective: that of competing acceleration and radiative cooling time-scales. I then predict whether divergences occur in the cutoff energies of secondary acceleration channels by similarly comparing their time-scales against the cooling time. Including Klein-Nishina effects, the IC cooling time-scale is[1]

$$\begin{aligned} t_{\rm cool,IC}(\gamma) &\equiv \frac{\gamma m_e c^2}{P_{\rm IC}(\gamma)} = \frac{\gamma_{\rm cool}}{\gamma f_{\rm KN}(\gamma/\gamma_{\rm KN})}\frac{L}{c} \\ &= \frac{\gamma_{\rm cool}}{\gamma_{\rm KN}}\left[(\gamma/\gamma_{\rm KN}) f_{\rm KN}(\gamma/\gamma_{\rm KN})\right]^{-1}\frac{L}{c} , \end{aligned} \tag{6.21}$$

where $\gamma_{\rm cool}$ [equation (2.9)] is the Lorentz factor of a particle with Thomson cooling time $t_{\rm cool,T}(\gamma) = \gamma m_e c^2/P_{\rm T}(\gamma)$ equal to L/c. A plot of $t_{\rm cool,IC}(\gamma)$ is presented in Fig. 6.4.

The relationship (6.21) encodes a wealth of information. In the Thomson regime $\gamma \ll \gamma_{\rm KN}$, $f_{\rm KN} \to 1$, and $t_{\rm cool,IC} \to t_{\rm cool,T} = \gamma_{\rm cool} L/\gamma c$, which decreases inversely with γ. However, for $\gamma \gg \gamma_{\rm KN}$, $f_{\rm KN}$ assumes its asymptotic form (6.6), inducing the scaling

$$t_{\rm cool,IC}(\gamma \gg \gamma_{\rm KN}) \simeq \frac{2L\gamma_{\rm cool}}{9c\gamma_{\rm KN}}\frac{\gamma/\gamma_{\rm KN}}{\ln(\gamma/\gamma_{\rm KN}) - 11/6} . \tag{6.22}$$

[1] As shown later [equation (6.26)], the factor $\gamma_{\rm cool}/\gamma_{\rm KN}$ in (6.21) equals $3/5\tau_{\gamma\gamma}$ where $\tau_{\gamma\gamma}$ is the characteristic pair-production optical depth of the system.

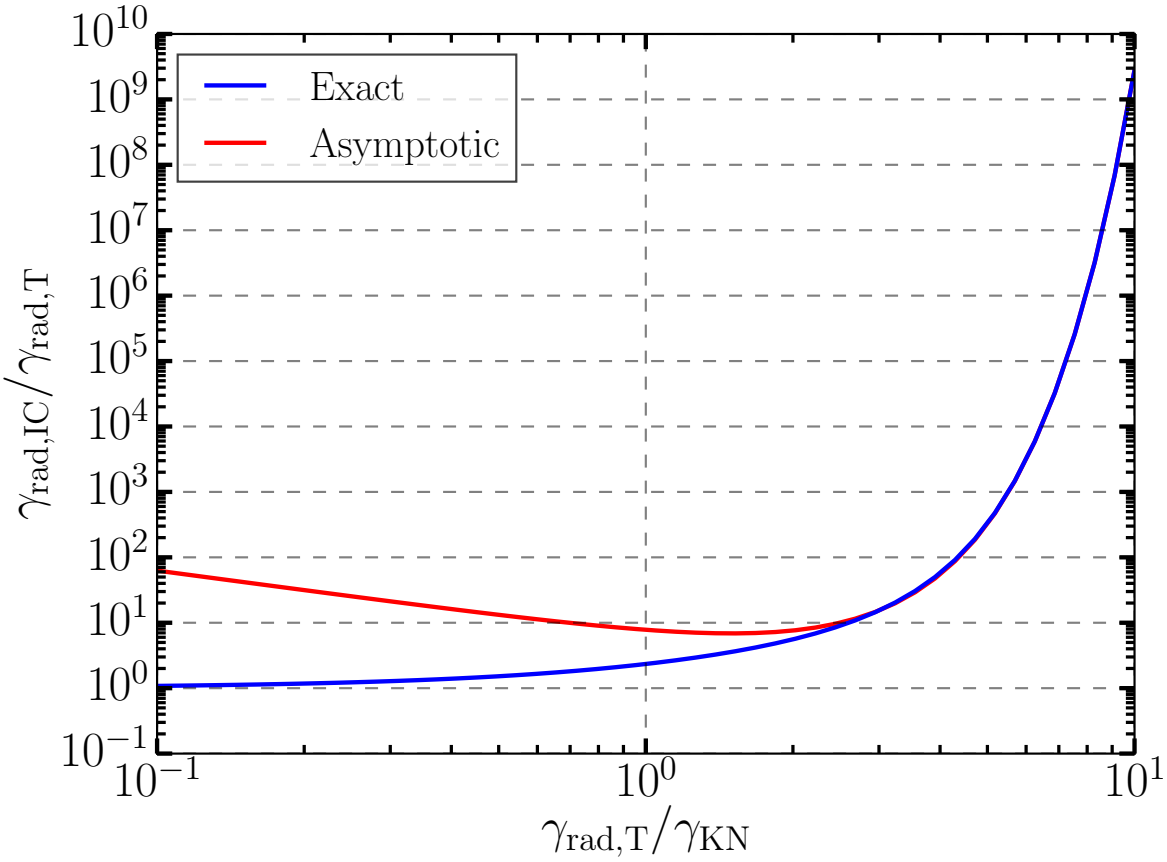

Figure 6.3: The Klein-Nishina cutoff Lorentz factor $\gamma_{\rm rad,IC}$ [equation (6.19)] and its asymptotic form (6.20), both normalized by the Thomson equivalent $\gamma_{\rm rad,T}$. The asymptotic form is already accurate when $\gamma_{\rm rad,T} \simeq 2.5\gamma_{\rm KN}$.

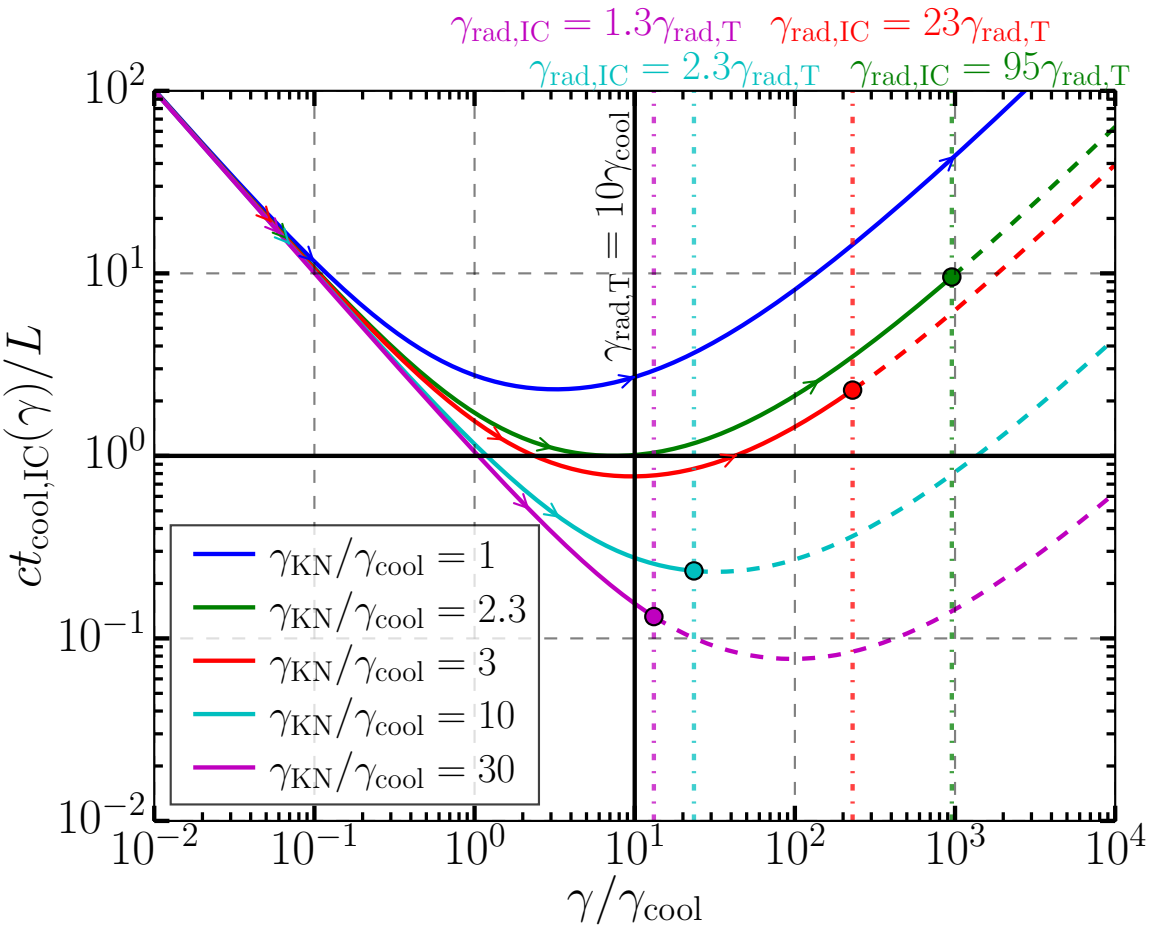

Figure 6.4: The IC cooling time $ct_{\rm cool,IC}(\gamma)/L$ [equation (6.21)]. In the plot, $\gamma_{\rm cool}$ is fixed, $\gamma_{\rm max} = 10\gamma_{\rm rad,T} = 100\gamma_{\rm cool}$, and $\gamma_{\rm KN}$ varies between curves. Arrows indicate the progression of particle cooling times during acceleration near reconnection X-points, which terminates at $\gamma = \gamma_{\rm rad,IC}$ (denoted by filled circles). For reference, each curve continues past $\gamma_{\rm rad,IC}$ as a dashed line. The cooling time is non-monotonic. Depending on $\gamma_{\rm KN}$, some (or even all) of the particles may attain high enough energies not to radiate [possessing cooling times $t_{\rm cool,IC}(\gamma) > L/c$]. If $\gamma_{\rm KN} < 2.3\gamma_{\rm cool}$, then all particles are non-radiative. If $\gamma_{\rm KN} > 2.3\gamma_{\rm cool}$, then only the low-energy ($\gamma < \gamma_{\rm cool,1} \sim \gamma_{\rm cool}$) and high-energy ($\gamma > \gamma_{\rm cool,2}$) particles are non-radiative. For particles to exceed $\gamma_{\rm cool,2}$, the condition $\gamma_{\rm rad,IC} \geq \gamma_{\rm cool,2}$ is necessary, but not sufficient (see text).

Thus, in the deep Klein-Nishina limit, $t_{\rm cool,IC}(\gamma)$ *increases* linearly in γ with a logarithmic correction. This is the source of the super-exponential divergence of $\gamma_{\rm rad,IC}$ with $\gamma_{\rm rad,T}$ in equation (6.20). Equation (2.4) shows that the time-scale, $t_{\rm X} \simeq 10\gamma\rho_0/c$, on which X-point acceleration occurs is also linear in γ. Upon equating $t_{\rm cool,IC}$ and $t_{\rm X}$, which defines the radiative cutoff $\gamma_{\rm rad,IC}$, γ cancels. Thus, $t_{\rm X}$ can only surpass $t_{\rm cool,IC}$ if $\ln(\gamma/\gamma_{\rm KN})$ grows large enough, inducing the super-exponential scaling in equation (6.20).

Importantly, a similar situation does not arise for secondary acceleration channels, which are slower, possessing time-scales super-linear in γ (see, e.g., the discussion in section 3.3.3 and Mehlhaff et al. 2020). As an example, consider the secondary process described in chapter 2 (section 2.2.2), where particles inside contracting plasmoids are gradually energized. The Lorentz factors of such particles grow as $\gamma(t) \propto \sqrt{t}$ (Petropoulou & Sironi, 2018; Hakobyan et al., 2021), yielding the acceleration time $t_{\rm sec}(t) = \gamma(t)/\dot{\gamma}(t) \propto \gamma^2(t)$. Thus, it is much easier for this secondary mechanism – and any other for which $t_{\rm sec} \propto \gamma^\zeta$ with $\zeta > 1$ – to radiatively saturate, as I now show.

For the sake of generality, suppose that $t_{\rm sec} = C\gamma^\zeta$ with C a constant independent of $\epsilon_{\rm ph}$ and $U_{\rm ph}$ (and, hence, of $\gamma_{\rm cool}$, $\gamma_{\rm rad,T}$, $\gamma_{\rm KN}$, and $\gamma_{\rm rad,IC}$; this parallels the argument in section 3.3.3 and Mehlhaff et al. 2020). Then, when IC cooling proceeds deep into the Klein-Nishina regime, the equality $t_{\rm cool,IC} = t_{\rm sec}$ reduces to $\gamma_{\rm sec}^{\zeta-1} \propto \gamma_{\rm cool}/\gamma_{\rm KN}^2$. The key difference from the direct X-point acceleration channel is that here one can ignore the $\ln(\gamma)$ correction – its dependence on γ is much weaker than $\gamma^{\zeta-1}$. As a result, the cutoff $\gamma_{\rm sec}$ scales merely polynomially in $\gamma_{\rm cool}$ (i.e., in $\gamma_{\rm rad,T}$) and $\gamma_{\rm KN}$: $\gamma_{\rm sec} \propto (\gamma_{\rm cool}/\gamma_{\rm KN}^2)^{1/(\zeta-1)}$. Plugging in $\zeta = 2$ for the adiabatic plasmoid compression process gives $\gamma_{\rm sec} \propto \gamma_{\rm cool}/\gamma_{\rm KN}^2$ (different from $\gamma_{\rm sec}$ reported in section 2.2.2 because I am now considering deep Klein-Nishina cooling). Thus, Klein-Nishina physics may effectively remove the high-energy radiation-reaction cap on impulsive X-point acceleration, but *not* on other processes, potentially increasing the relative importance of the primary direct energization channel.

As discussed above, $t_{\rm cool,IC}(\gamma)$ is non-monotonic, decreasing with γ when $\gamma \ll \gamma_{\rm KN}$ and

increasing when $\gamma \gg \gamma_{\rm KN}$. It reaches the minimum

$$\min_{\gamma}\left(\frac{ct_{\rm cool,IC}(\gamma)}{L}\right) \simeq 2.32\frac{\gamma_{\rm cool}}{\gamma_{\rm KN}} \tag{6.23}$$

at a critical *fastest-cooling Lorentz factor*

$$\arg\min_{\gamma}\left(\frac{ct_{\rm cool,IC}(\gamma)}{L}\right) \simeq 3.20\gamma_{\rm KN}\,, \tag{6.24}$$

[i.e., $\min_{\gamma}(t_{\rm cool,IC}(\gamma)) = t_{\rm cool,IC}(3.20\gamma_{\rm KN})$]. The minimum cooling time (6.23) implies that, when $\gamma_{\rm KN} < 2.32\gamma_{\rm cool}$, *all* of the particles in the system radiate weakly (they have cooling times exceeding L/c). Even if $\gamma_{\rm KN}$ falls above this threshold and, hence, $\min(t_{\rm cool,IC}) < L/c$, some high-energy particles may radiate weakly. Namely, if a particle surpasses the fastest-cooling Lorentz factor $3.20\gamma_{\rm KN}$ by a sufficient amount, it reaches a high-energy domain with $t_{\rm cool,IC}(\gamma) > L/c$. This effect does not occur in the Thomson regime.

These remarks are illustrated in Fig. 6.4. The figure displays $ct_{\rm cool,IC}(\gamma)/L$ for fixed $\gamma_{\rm cool} = \gamma_{\rm rad,T}/10 = \gamma_{\rm max}/100$ and several $\gamma_{\rm KN}$. On each curve for which $\gamma_{\rm KN} > 2.32\gamma_{\rm cool}$, the line $t_{\rm cool,IC} = L/c$ is crossed twice, once at a low Lorentz factor $\gamma_{\rm cool,1}$ and once at a high Lorentz factor $\gamma_{\rm cool,2}$. I analyze $\gamma_{\rm cool,1}$ and $\gamma_{\rm cool,2}$ shortly, but I point out some basic features of Fig. 6.4 beforehand. First, when $\gamma_{\rm rad,IC} \simeq \gamma_{\rm rad,T} \leq \gamma_{\rm KN}$ (implying $\gamma_{\rm KN} > 2.32\gamma_{\rm cool}$ because $\gamma_{\rm KN} > \gamma_{\rm rad,T} = 10\gamma_{\rm cool} > 2.32\gamma_{\rm cool}$), particles may access only the Thomson portion of a cooling curve where $t_{\rm cool,IC} \propto \gamma^{-1}$. The $\gamma_{\rm KN} = 30\gamma_{\rm cool} = 3\gamma_{\rm rad,T}$ case illustrates this. Next, in the opposite limit, when $\gamma_{\rm KN}$ becomes smaller than $\gamma_{\rm rad,T}$, the radiative cutoff $\gamma_{\rm rad,IC}$ begins to grow rapidly, opening up the portion of a curve that bends upward. Eventually, at the critical Lorentz factor $\gamma_{\rm cool,2}$, the cooling time $t_{\rm cool,IC}$ once again equals L/c. Thus, if $\gamma_{\rm rad,IC} > \gamma_{\rm cool,2}$, particles accelerated near X-points could break into the high-energy weakly radiative regime. However, in reality, whether particles will actually cross this boundary does not depend solely on whether $\gamma_{\rm rad,IC}$ surpasses $\gamma_{\rm cool,2}$. That is just a necessary condition. In addition, the intrinsic X-point acceleration Lorentz factor $\gamma_{\rm X}$ must exceed $\gamma_{\rm cool,2}$, or – if it does not – secondary acceleration channels must be able to energize particles against radiative cooling past $\gamma_{\rm cool,2}$.

Both $\gamma_{\rm cool,1}$ and $\gamma_{\rm cool,2}$ are illustrated in Fig. 6.5 as functions of $\gamma_{\rm KN}/\gamma_{\rm cool}$ (they only depend on $\gamma_{\rm rad,T}$ and $\gamma_{\rm max}$ through $\gamma_{\rm cool}$). In general, $\gamma_{\rm cool,1}$ is close to $\gamma_{\rm cool}$, because, as illustrated in Fig. 6.4, particles are almost completely in the Thomson limit when $t_{\rm cool,IC}$ crosses L/c from above. In contrast, $\gamma_{\rm cool,2}$, where $t_{\rm cool,IC}$ crosses L/c in the opposite direction, depends rather strongly on $\gamma_{\rm KN}$. Because $\gamma_{\rm cool,2}$ occurs fairly deep into the Klein-Nishina regime, one may employ expression (6.22) to see that $\gamma_{\rm cool,2}$ satisfies

$$\frac{9\gamma_{\rm KN}}{2\gamma_{\rm cool}} \simeq \frac{\gamma_{\rm cool,2}/\gamma_{\rm KN}}{\ln(\gamma_{\rm cool,2}/\gamma_{\rm KN}) - 11/6}\,, \tag{6.25}$$

implying that $\gamma_{\rm cool,2}/\gamma_{\rm cool} \propto (\gamma_{\rm KN}/\gamma_{\rm cool})^2 \times \mathcal{O}(\ln(\gamma_{\rm cool,2}/\gamma_{\rm KN}))$.

Thus, $\gamma_{\rm cool,2}$ is approximately quadratic in $\gamma_{\rm KN}$. This differs from $\gamma_{\rm rad,IC}$, which grows as $\gamma_{\rm KN}$ is *reduced*. For illustration, imagine a progression in $\gamma_{\rm KN}$ with $\gamma_{\rm KN}$ starting high enough that $\gamma_{\rm cool,2} \gg \gamma_{\rm max} \gg \gamma_{\rm rad,IC}$ (for example, as on the $\gamma_{\rm KN} = 30\gamma_{\rm cool}$ curve in Fig. 6.4). Then, dialing down $\gamma_{\rm KN}$, eventually $\gamma_{\rm cool,2}$ crosses $\gamma_{\rm rad,IC}$ and $\gamma_{\rm max}$ from above. It turns out that, by definition, these crossings occur simultaneously. For, if $\gamma_{\rm cool,2} = \gamma_{\rm max}$, then $t_{\rm X}(\gamma_{\rm max}) \equiv L/c$ [equation (2.5)] and $t_{\rm cool,IC}(\gamma_{\rm cool,2}) \equiv L/c$ [equation (6.21)]. Consequently, $t_{\rm cool,IC}(\gamma_{\rm cool,2}) = t_{\rm X}(\gamma_{\rm cool,2})$, which implies $\gamma_{\rm max} = \gamma_{\rm cool,2} = \gamma_{\rm rad,IC}$. By similar reasoning, one can show that, if $\gamma_{\rm cool,2} = \gamma_{\rm rad,IC}$, then $\gamma_{\rm cool,2} = \gamma_{\rm max}$. Continuing to reduce $\gamma_{\rm KN}$ beyond this 'triple point' yields the scale ordering $\gamma_{\rm rad,IC} \gg \gamma_{\rm max} \gg \gamma_{\rm cool,2}$ (typified in the $\gamma_{\rm KN} = 3\gamma_{\rm cool}$ curve of Fig. 6.4, although the separation of scales is rather small in that example). The first of these inequalities, $\gamma_{\rm rad,IC} \gg \gamma_{\rm max}$, means that no particles achieve radiative saturation (in the sense of chapter 2); all have Lorentz factors $\gamma < \gamma_{\rm rad,IC}$. The second inequality $\gamma_{\rm cool,2} \ll \gamma_{\rm max}$, potentially allows various acceleration channels (e.g., X-point acceleration if $\gamma_{\rm X} > \gamma_{\rm cool,2}$) to break a population of high-energy particles through the weakly radiative $\gamma = \gamma_{\rm cool,2}$ barrier.

To summarize up to this point, there are several stark departures from the Thomson picture of radiative cooling in reconnection induced by finite $\gamma_{\rm KN}$. Not only do particles with Lorentz factors $\gamma > \gamma_{\rm KN}$ radiate their energy in discrete chunks, but their radiative cooling times can actually be quite long. When $\gamma_{\rm KN} \lesssim \gamma_{\rm rad,T}$, the effective cutoff Lorentz factor $\gamma_{\rm rad,IC}$ begins to

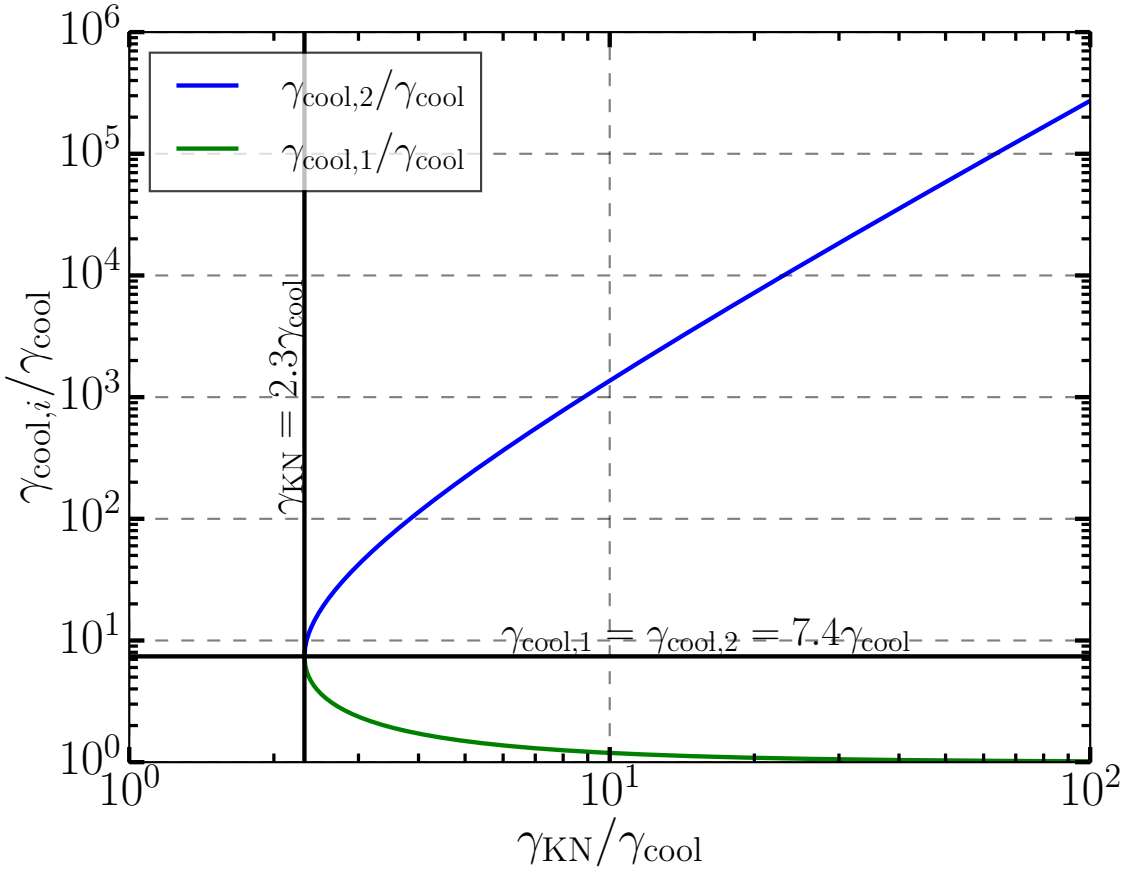

Figure 6.5: The Lorentz factors $\gamma_{\rm cool,1}$ and $\gamma_{\rm cool,2}$ where $t_{\rm cool,IC}(\gamma)$ dips below and above one dynamical time L/c, respectively. Since $\gamma_{\rm cool,1}$ occurs almost entirely in the Thomson regime, it is almost always quite close to $\gamma_{\rm cool}$, differing from $\gamma_{\rm cool}$ by at most the factor $7.4 = 3.20 \times 2.32$ [equations (6.23) and (6.24)] when $\gamma_{\rm KN}$ approaches the non-radiative value $2.32\gamma_{\rm cool}$. In contrast $\gamma_{\rm cool,2}$ grows approximately quadratically in $\gamma_{\rm KN}/\gamma_{\rm cool}$ (with a logarithmic correction).

grow rapidly, and comes, with just a small change in $\gamma_{\rm KN}$, to exceed $\gamma_{\rm max}$, the analog of the Hillas Lorentz factor for relativistic magnetic reconnection. When $\gamma_{\rm KN}$ is decreased even more, eventually it falls below $\gamma_{\rm cool}$, and no particles in the system radiate efficiently. Thus, even in a nominally strongly radiative Thomson scale ordering, by making $\gamma_{\rm KN}$ small enough, a non-radiative regime can be reached. At intermediate $\gamma_{\rm KN}$, a variety of intriguing and exotic physical effects can occur, which I elucidate later in this study. First, I cover one additional piece of physics that is entirely new to the Klein-Nishina realm: pair production.

6.2.1 Pair-production in Klein-Nishina reconnection

A Comptonized photon of energy ϵ may collide with a background photon (energy $\epsilon_{\rm ph}$) to produce an electron-positron pair if the threshold criterion, $\epsilon\epsilon_{\rm ph} \geq (m_e c^2)^2$, is met. In this work, I assume that $\gamma_{\rm KN} = m_e c^2/4\epsilon_{\rm ph} \gg 1$ ($\epsilon_{\rm ph} \ll 100\,{\rm keV}$), implying that $\epsilon \gg m_e c^2 \gg \epsilon_{\rm ph}$ is required to reach pair-production threshold. Such a high-energy photon can only be emitted in the Klein-Nishina IC regime. (I would here like to acknowledge Benoît Cerutti, who originally pointed this out to me.) If one assumes that the IC scattering occurs in the Thomson limit, then a contradiction arises because $1 \leq \epsilon\epsilon_{\rm ph}/(m_e c^2)^2 \sim (\gamma\epsilon_{\rm ph}/m_e c^2)^2 \sim q^2$. In contrast, assuming a Klein-Nishina scaling $\epsilon \sim \gamma m_e c^2/2$ yields the self-consistent result, $q \equiv 4\gamma\epsilon_{\rm ph}/m_e c^2 \sim 8\epsilon\epsilon_{\rm ph}/(m_e c^2)^2 \geq 8$. I adopt $\gamma_{\rm pp} \equiv 8\gamma_{\rm KN}$ (used mainly in the next chapter) as the characteristic minimum particle Lorentz factor to emit above-threshold photons.

Now, the pair-production cross section $\sigma_{\gamma\gamma}$ is zero precisely at threshold, $\epsilon\epsilon_{\rm ph} = (m_e c^2)^2$, but, for the (isotropic, monochromatic) background distribution (6.1), it soon peaks at $\sigma_{\gamma\gamma} \simeq \sigma_{\rm T}/5$ when $\epsilon\epsilon_{\rm ph} \simeq 3.6(m_e c^2)^2$ (Gould & Schréder, 1967). For such photons $q \sim 3.6 \times 8 \simeq 30$. Hence, although the energy scales at which Klein-Nishina IC cooling and pair production occur are both set, fundamentally, by $\gamma_{\rm KN}$, they are offset from one another by a factor of about $10 - 30$. The former kicks in when $q \sim 1$ and the latter when $q \sim 8 - 30$. In this sense, the energy scale $\gamma_{\rm KN}$ 'splits', similarly to $\sigma_{\rm c,0}$ (section 2.2.1), into two that are offset by a fixed ratio.

However, it is not clear that astrophysical reconnection accelerates particles to energies that

are high enough to stray from the Thomson limit but not to emit pair-producing photons. If, contrary to my simplified monochromatic assumption, the seed photons have any spread in energy, then photons Comptonized from the high-energy end of the background will more easily pair-produce with the lower-energy component. Even for a thermal radiation bath, the two frequencies where the Planck spectrum attains half its maximum value are offset from each other by about a factor of 5, reducing the effective splitting of $\gamma_{\rm KN}$ from a factor of $10-30$ to $2-6$. The cutoff in the reconnection-energized particle distribution would then have to fall precisely in a narrow range for Klein-Nishina effects to kick in but for pair-production to remain impossible. And, even in this case, only the very highest-energy sliver of particles would experience Klein-Nishina IC losses; most of the particles would still be cooled in the Thomson regime. Thus, from here on, I assume that Klein-Nishina IC scattering coincides with the emission of above-threshold photons in reconnection.

However, just because a high-energy photon is above threshold does not mean that it gets absorbed inside the reconnection system. One must also consider the optical depth, $\tau_{\gamma\gamma} = U_{\rm ph}\sigma_{\gamma\gamma}L/\epsilon_{\rm ph}$, to pair-production. For simplicity, I evaluate $\tau_{\gamma\gamma}$ at the peak cross section $\sigma_{\gamma\gamma} \simeq \sigma_{\rm T}/5$, which is attained when $\epsilon\epsilon_{\rm ph} \simeq 3.6 m_e c^2$. Thus

$$\tau_{\gamma\gamma} \equiv \frac{U_{\rm ph}\sigma_{\rm T}L}{5\epsilon_{\rm ph}} = \frac{3}{5}\frac{\gamma_{\rm max}\gamma_{\rm KN}}{\gamma_{\rm rad,T}^2} = \frac{3}{5}\frac{\gamma_{\rm KN}}{\gamma_{\rm cool}}\,. \tag{6.26}$$

Actually, $\tau_{\gamma\gamma}$ has been encountered already. It is the (inverse of the) prefactor in the expression for the cooling time $t_{\rm cool,IC}(\gamma)$ in equation (6.21). Thus, the condition $\gamma_{\rm KN} > 2.32\gamma_{\rm cool}$ [equation (6.23)], which ensures that at least some particles cool in times shorter than L/c, is the same as the optically thick condition $\tau_{\gamma\gamma} > (3/5)\times 2.32 \simeq 1$.

This means that there is an appreciable range of parameters where one expects both dynamically-important Klein-Nishina radiative cooling and pair-production. Both mechanisms may actively feed back on the reconnection process when $\gamma_{\rm cool} \ll \gamma_{\rm KN} \ll \gamma_{\rm rad,T}$. The first relationship, $\gamma_{\rm cool} \ll \gamma_{\rm KN}$, is necessary both for $\tau_{\gamma\gamma} \gg 1$ and for $\min[t_{\rm cool,IC}(\gamma)] \ll L/c$. The second criterion, $\gamma_{\rm KN} \ll \gamma_{\rm rad,T}$, is required for at least some particles to enter the regime where Klein-Nishina effects begin to impact their radiative cooling, also enabling them to emit photons above pair threshold.

This gives a simple rule for deciding when Klein-Nishina and pair-production physics become important in reconnection. One must just assemble all of the energy scales – $\gamma_{\rm max}$, $\sigma_{\rm c,0}$ (i.e., $\gamma_{\rm X}$ and $\langle\gamma\rangle$), $\gamma_{\rm rad,T}$, and $\gamma_{\rm cool}$ – arrange these into a familiar Thomson hierarchy (as in section 2.2.2), and insert $\gamma_{\rm KN}$ into a relevant location. If $\gamma_{\rm KN}$ is larger than $\gamma_{\rm rad,T}$, Klein-Nishina effects are absent because $\gamma_{\rm rad,T}$ imposes a hard upper bound on particle acceleration, and, consequently, no particles ever reach $\gamma_{\rm KN}$. If, on the other hand, $\gamma_{\rm KN} < \gamma_{\rm cool}$, then Klein-Nishina effects suppress cooling so much that the whole system becomes non-radiative. Only if $\gamma_{\rm cool} \ll \gamma_{\rm KN} \ll \gamma_{\rm rad,T}$ are Klein-Nishina IC cooling and pair-production both important. And, in that case, it is also necessary to consider how $\gamma_{\rm rad,T}$, $\gamma_{\rm KN}$, and $\gamma_{\rm cool}$ are ordered with respect to the other scales in the problem. These remarks are illustrated in Fig. 6.6 and elaborated in the next subsection.

6.2.2 Regimes of Klein-Nishina radiative reconnection

I now systematically explore, as I did for Thomson IC cooling in section 2.2.2, how to classify regimes of Klein-Nishina radiative reconnection. As an example, consider the Thomson ordering $\gamma_{\rm max} \geq \gamma_{\rm rad,T} \geq \gamma_{\rm X} \gg \langle\gamma\rangle \gg \gamma_{\rm cool}$ (4th row in Table 2.1). If one inserts $\gamma_{\rm KN}$ between $\gamma_{\rm rad,T}$ and $\gamma_{\rm X}$, then Klein-Nishina effects do not affect primary X-point acceleration. They only come into play if secondary energization channels can push particles up to $\gamma \sim \gamma_{\rm KN}$. If $\gamma_{\rm KN}$ is instead placed between $\gamma_{\rm X}$ and $\langle\gamma\rangle$, Klein-Nishina radiative cooling definitely impacts high-energy particles accelerated near X-points, and these particles are also likely to emit pair-producing photons. Klein-Nishina and pair-production physics become even more important if $\gamma_{\rm KN}$ is made smaller than $\langle\gamma\rangle \sim \sigma_{\rm c,0}/4$. Then, the bulk of the accelerated particles – not just the high-energy tail – emit in the Klein-Nishina regime and, likely, many pairs are produced.

Because $\gamma_{\rm KN}$ is not the only new scale, but also introduces a few derived scales (e.g., $\gamma_{\rm cool,2}$ and $\gamma_{\rm rad,IC}$), exhaustively discussing all possible regimes like in section 2.2.2 is prohibitively tedious. Even in the preceding paragraph, I did not consider subtleties such as whether $\gamma_{\rm X} > \gamma_{\rm cool,2}$, in which case some high-energy particles radiate inefficiently. In lieu of an exhaustive discussion, I supply Fig. 6.7, a 'phase diagram', in the $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$–$\gamma_{\rm KN}/\sigma_{\rm c,0}$ plane, of the complex radiative parameter

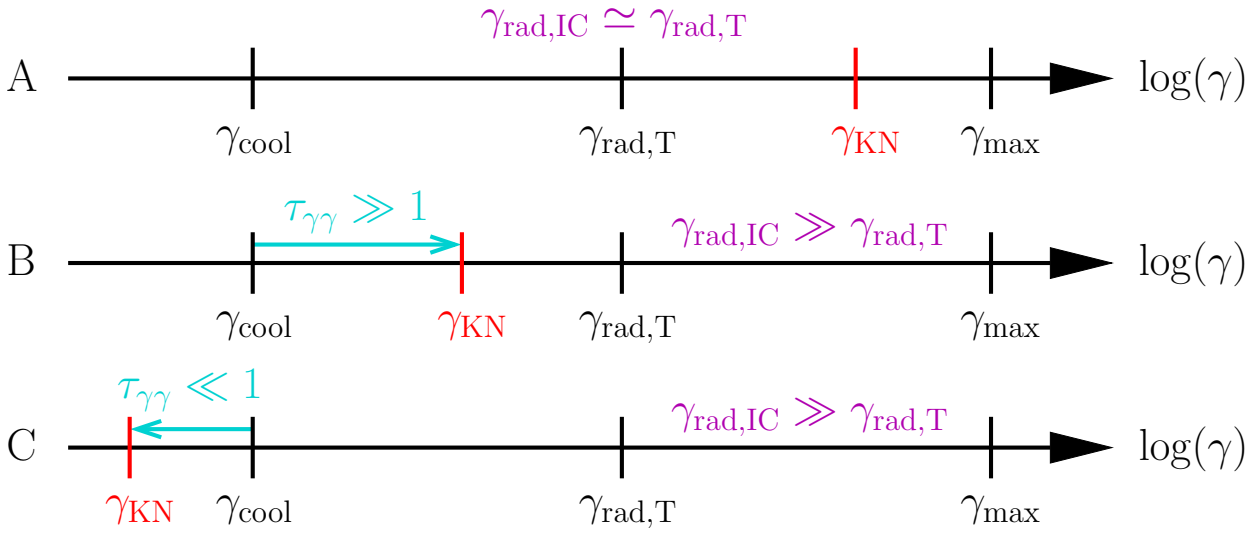

Figure 6.6: Basic orderings of $\gamma_{\rm KN}$ with respect to $\gamma_{\rm cool}$, $\gamma_{\rm rad,T}$, and $\gamma_{\rm max}$. Note that $\gamma_{\rm rad,T} \equiv (\gamma_{\rm cool}\gamma_{\rm max})^{1/2}$ [equation (2.9)]. The case $\gamma_{\rm max} < \gamma_{\rm rad,T} < \gamma_{\rm cool}$, which is non-radiative irrespective of $\gamma_{\rm KN}$, is not considered. The pair-production optical depth $\tau_{\gamma\gamma} \sim \gamma_{\rm KN}/\gamma_{\rm cool}$ is roughly the separation between $\gamma_{\rm KN}$ and $\gamma_{\rm cool}$. The cutoff $\gamma_{\rm rad,IC}$ is virtually identical to $\gamma_{\rm rad,T}$ when $\gamma_{\rm rad,T} \ll \gamma_{\rm KN}$; otherwise, it is much larger than $\gamma_{\rm rad,T}$. Case A: A Thomson radiative reconnection ordering (section 2.2.2); $\gamma_{\rm rad,T} \ll \gamma_{\rm KN}$ prevents particles from accessing the Klein-Nishina regime and from emitting above-threshold photons. Case B: A scale hierarchy where Klein-Nishina IC losses (because $\gamma_{\rm KN} \ll \gamma_{\rm rad,T}$) and pair production (because $\gamma_{\rm cool} \ll \gamma_{\rm KN}$ implies $\tau_{\gamma\gamma} \gg 1$) are both likely. Case C: All orderings with $\gamma_{\rm KN} \ll \gamma_{\rm cool}$ are non/weakly-radiative because $t_{\rm cool,IC} > L/c$. Even the few photons radiated above pair threshold are not absorbed because $\tau_{\gamma\gamma} \ll 1$.

space for Klein-Nishina reconnection. The parameter space is, in reality, 3-dimensional, depending also on $\gamma_{\rm max}/\sigma_{\rm c,0}$. To display it in 2D, I set an exemplary value of $\gamma_{\rm max} = 10^3\sigma_{\rm c,0}$ in Fig. 6.7.

In the figure, contours highlight important values of energy scales and often distinguish different-colored regimes of interest. I caution that the color scheme is somewhat arbitrary. Almost every sliver of parameter space enclosed within a set of contours is its own physical regime, and only a subset of the relevant contours are shown. Without analyzing every possible contour-enclosed region, the best one can do is group regions based on similar expected qualitative behavior, a heuristic that guides the color-coding in Fig. 6.7. However, this exercise is ultimately subjective. A given grouping is useful for conceptualizing some physical similarities, but may need to be reevaluated if the physics of main interest changes.

I now describe the (color-coded) grouping of regions adopted in Fig. 6.7. I begin with the fundamentally new domains corresponding to case B in Fig. 6.6, $\gamma_{\rm cool} < \gamma_{\rm KN} < \gamma_{\rm rad,T}$, where Klein-Nishina effects feature prominently. One of these is the blue area, in which $\gamma_{\rm KN} < \gamma_{\rm rad,IC}$ and $\gamma_{\rm rad,IC} < \gamma_{\rm max} < \gamma_{\rm cool,2}$. In this area, the overall radiative cutoff Lorentz factor $\gamma_{\rm rad,IC}$ is finite (less than $\gamma_{\rm max}$) while $\gamma_{\rm cool,2}$ is not ($\gamma_{\rm cool,2} > \gamma_{\rm max}$). This means that, despite Klein-Nishina suppression of the IC cross section, radiative losses may still regulate the highest achievable energies. To illustrate this, the contour $\gamma_{\rm rad,IC} = 4\sigma_{\rm c,0} = \gamma_{\rm X}$ is drawn (in this discussion, I assume that $\gamma_{\rm X} = 4\sigma_{\rm c,0}$ for definiteness), below which IC radiation limits direct acceleration near reconnection X-points. Straying from the blue region across the line $\gamma_{\rm max} = \gamma_{\rm rad,IC} = \gamma_{\rm cool,2}$ lands one in the red area. Here, the ordering of $\gamma_{\rm cool,2}$ and $\gamma_{\rm rad,IC}$ about $\gamma_{\rm max}$ is flipped: $\gamma_{\rm cool,2} < \gamma_{\rm max} < \gamma_{\rm rad,IC}$. Thus the radiative cutoff energy $\gamma_{\rm rad,IC}$ is no longer finite, but the energy $\gamma_{\rm cool,2}$, beyond which particles are weakly radiative (cooling on time-scales longer than L/c), is now accessible. Here, it may be possible for high-energy particles to surpass $\gamma_{\rm cool,2}$ and enter into a weakly radiative regime. As one moves to the northwest through the red region, $\gamma_{\rm cool,2}$ becomes lower. Eventually, when the contour $\gamma_{\rm cool,2} = 4\sigma_{\rm c,0}$ is crossed, $\gamma_{\rm cool,2}$ falls below $\gamma_{\rm X}\,(= 4\sigma_{\rm c,0})$, guaranteeing that some particles venture into the high-energy Klein-Nishina weakly radiative limit. Continuing even farther upward in the diagram, one eventually crosses the $\gamma_{\rm KN} = 2.32\gamma_{\rm cool}$ line, where the whole system becomes

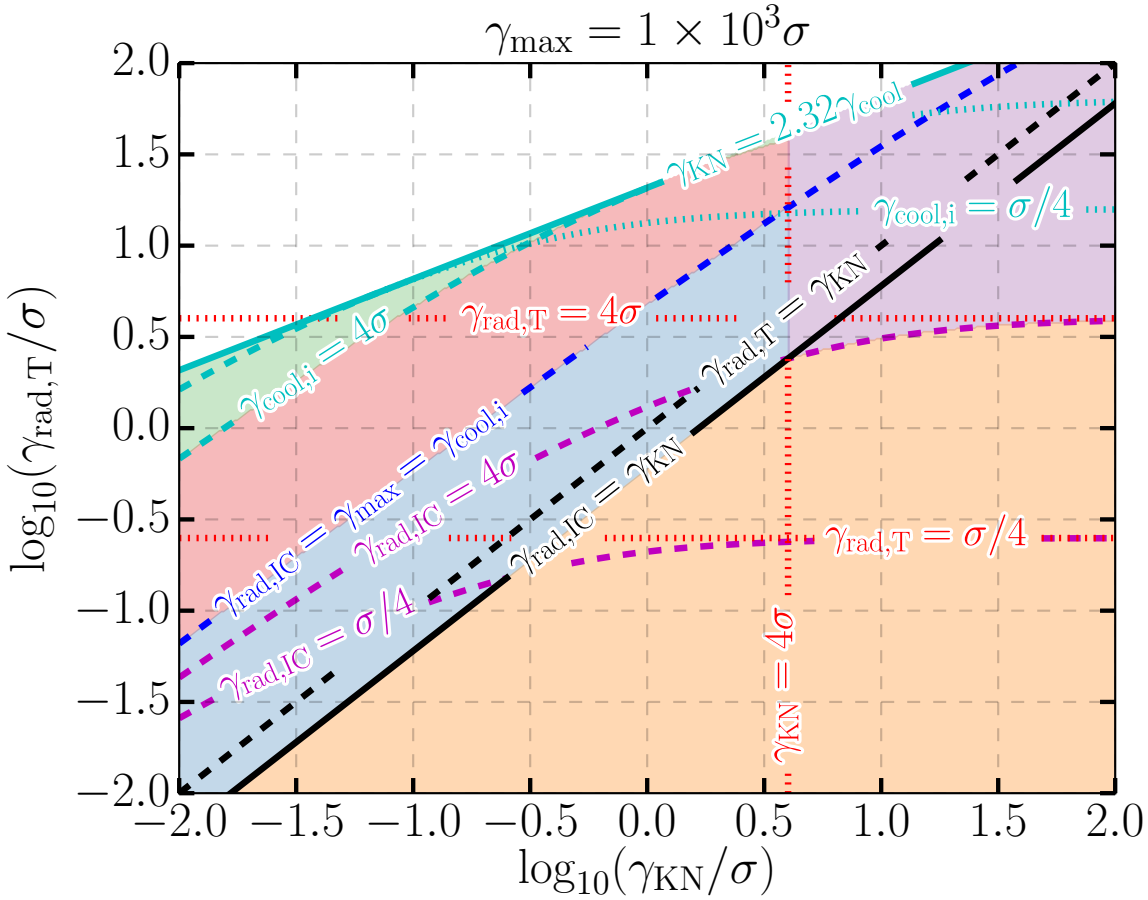

Figure 6.7: Radiative phase diagram for Klein-Nishina relativistic magnetic reconnection. The diagram is inherently 3-dimensional – it depends on $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$, $\gamma_{\rm KN}/\sigma_{\rm c,0}$, and $\gamma_{\rm max}/\sigma_{\rm c,0}$ – but I display a 2D slice where $\gamma_{\rm max} = 10^3\sigma_{\rm c,0}$. In the plot, $\sigma_{\rm c,0}$ is abbreviated as σ. Contours show important equalities between energy scales. The $\gamma_{{\rm cool},i} = $ const. contours transition, when they strike the curve $\gamma_{\rm KN} = 2.32\gamma_{\rm cool}$, from a dashed portion along which $\gamma_{\rm cool,2}$ is fixed, to a dotted portion where $\gamma_{\rm cool,1}$ is fixed to the same value. Regions with a common color exhibit similar qualitative behavior. The colors are inspired by assuming that $\gamma_{\rm X} = 4\sigma_{\rm c,0}$ and, fictitiously, that X-point acceleration is the only operative energization mechanism. Thus, the highest Lorentz factor attained is $\min(\gamma_{\rm rad,IC}, 4\sigma_{\rm c,0})$. In the purple region, $4\sigma_{\rm c,0} < \gamma_{\rm KN}, \gamma_{\rm rad,IC}$. Hence, radiative cooling does not furnish an upper bound on particle acceleration, and no particles suffer Klein-Nishina losses. In the orange region, $\gamma_{\rm rad,IC} < \gamma_{\rm KN}, 4\sigma_{\rm c,0}$. Here, Thomson IC losses firmly cap the maximum particle energy to $\gamma_{\rm rad,IC} \simeq \gamma_{\rm rad,T} < \gamma_{\rm KN}$, and, consequently, the highest energy particles are radiatively saturated. In the blue region, $\gamma_{\rm KN} < \gamma_{\rm X}, \gamma_{\rm rad,IC}$ and $\gamma_{\rm rad,IC} < \gamma_{\rm max} < \gamma_{\rm cool,2}$. Because $\gamma_{\rm KN} < \gamma_{\rm X}$ some particles reach $q \geq 1$. In addition, in the subset of the region where $\gamma_{\rm rad,IC} < \gamma_{\rm X} = 4\sigma_{\rm c,0}$, the highest energy particles are radiatively saturated. The red region is defined by $\gamma_{\rm rad,IC} > \gamma_{\rm max} > \gamma_{\rm cool,2} > 4\sigma_{\rm c,0}$. Thus, radiative losses do not bound particle acceleration, and, as a result, there is likely a broad spectrum of particles above the threshold energy $\gamma \simeq 8\gamma_{\rm KN}$ to emit pair-producing radiation. In the green $\gamma_{\rm cool,2} < 4\sigma_{\rm c,0}$ region, Klein-Nishina effects suppress cooling so that some high-energy particles are weakly radiative, having $\gamma > \gamma_{\rm cool,2}$ and, hence, $t_{\rm cool,IC} > L/c$. In the white $\gamma_{\rm KN} < 2.32\gamma_{\rm cool}$ region, all particles have $t_{\rm cool,IC} > L/c$.

virtually non-radiative (case C in Fig. 6.6).

Lastly, I discuss the Fig.-6.7 regions corresponding to case A of Fig. 6.6 ($\gamma_{\rm cool} < \gamma_{\rm rad,T} < \gamma_{\rm KN}, \gamma_{\rm max}$). The first of these is the orange area, where $\gamma_{\rm rad,IC} \simeq \gamma_{\rm rad,T} < \gamma_{\rm KN}, \gamma_{\rm X}$. Here, IC radiation limits X-point energization to below $\gamma_{\rm KN}$. In the final remaining region, the purple area, one has $4\sigma_{\rm c,0} = \gamma_{\rm X} < \gamma_{\rm rad,IC}, \gamma_{\rm KN}$. In this regime, radiative losses do not inhibit X-point acceleration, and X-point acceleration also cannot promote particles to high enough energies to stray outside the Thomson IC cooling limit.

Having overviewed the rich radiative parameter space available to IC-cooled relativistic reconnection, I now specialize to one as-yet relatively unexplored regime where Klein-Nishina physics profoundly impacts the overall dynamics. Here, pair production and Klein-Nishina radiative cooling can conspire together to form an important self-regulation mechanism. I devote the following chapter to a theoretical exploration of this pair-regulated Klein-Nishina radiative reconnection. Afterwards, in chapter 8, I discuss applications to reconnection-driven emission from ADCe and FSRQ jets.

Chapter 7

A Model of Pair-Regulated Klein-Nishina Reconnection

In this chapter, I explore technical aspects of reconnection with Klein-Nishina radiative cooling and pair production. I begin (section 7.1) with a general, qualitative description of the pair-regulation mechanism and its impact on the reconnection dynamics. I then (section 7.2) spend the majority of this chapter on the technical description and model of pair-regulated Klein-Nishina reconnection that undergirds the qualitative features described in section 7.1.

7.1 Pair-regulated Klein-Nishina reconnection: qualitative description

The self-regulation mechanism through which pair production feeds back on reconnection (a phenomenon I call *pair regulation* or *pair feedback*) is diagrammed in Fig. 7.1, and I describe how it works below. In that description, I adopt a convention that I maintain throughout the remainder of this work. Namely, I use the term *reconnection layer* (sometimes just *layer*) to refer specifically to the region of the reconnection system permeated by reconnected magnetic flux; the term *upstream* (sometimes *inflow*) refers to the region filled with unreconnected flux.

The pair feedback mechanism works as follows. First (step 1 in Fig. 7.1), particles accelerated in the reconnection layer Comptonize ambient seed photons to gamma-ray energies. Second (step 2 in Fig. 7.1), IC-produced gamma-rays penetrate into the upstream region about one pair-production mean free path from the layer. While propagating, these photons are immune to secondary IC scattering because the Thomson optical depth $\tau_{\rm T}$ is very small (even though the pair-production optical depth $\tau_{\gamma\gamma}$ exceeds unity). In step 3, high-energy photons are absorbed by the background

radiation, producing pairs in the upstream plasma. Newborn pairs are then advected toward the layer. While en route, they radiatively cool, and thus some of their initial energy never returns to the layer. Nevertheless, the created pairs remain hot enough that their energy density dominates that of the originally present colder upstream particles. Thus, the plasma feeding the layer in step 4 possesses a reduced hot magnetization $\sigma_{\rm h}$ [equation (2.7)] – the ratio of magnetic energy density to total (original + hot pairs) matter enthalpy density. This inhibits particle acceleration and subsequent photon emission in the reconnection layer, closing the negative feedback loop.

The model of pair-regulated Klein-Nishina reconnection developed in this chapter predicts two types of dynamical behaviors. If only a small fraction of the energy radiated away from the layer is recaptured as hot pairs, reconnection enters a steady state characterized by a *universal* (independent of the initial value) pair-regulated hot magnetization. However, if the nearly all of the radiated energy gets swept back into the layer as pairs, the steady state is never realized. Instead, the system overshoots its theoretical fixed point solution, getting caught between two extreme magnetization states. In these 'swing cycles', the high magnetization state yields efficient above-threshold photon emission from the layer and subsequent injection of hot pairs into the upstream region. This initiates a very low/pair-loaded magnetization. Here, pair production is quenched until the created pairs vacate the inflow plasma by entering the layer, restoring the high magnetization.

In both a steady state and a swing cycle, the created particles, when present, dominate the upstream pressure, but a prolific pair cascade is not expected. The (power-law) distribution of pairs injected into the inflow region, though potentially quite broad, is too steep for later pair generations – born from photons emitted by earlier upstream generations – to outnumber the first generation. Furthermore, for a wide range of parameters, the newborn pairs are also few in number relative to the original upstream plasma particles.

This last aspect of Klein-Nishina reconnection qualitatively departs from earlier treatments of radiative reconnection with pair-production (e.g., Lyubarskii, 1996; Uzdensky, 2011; Hakobyan et al., 2019). Rather than being dressed in a coat of pairs that dominates both the upstream matter

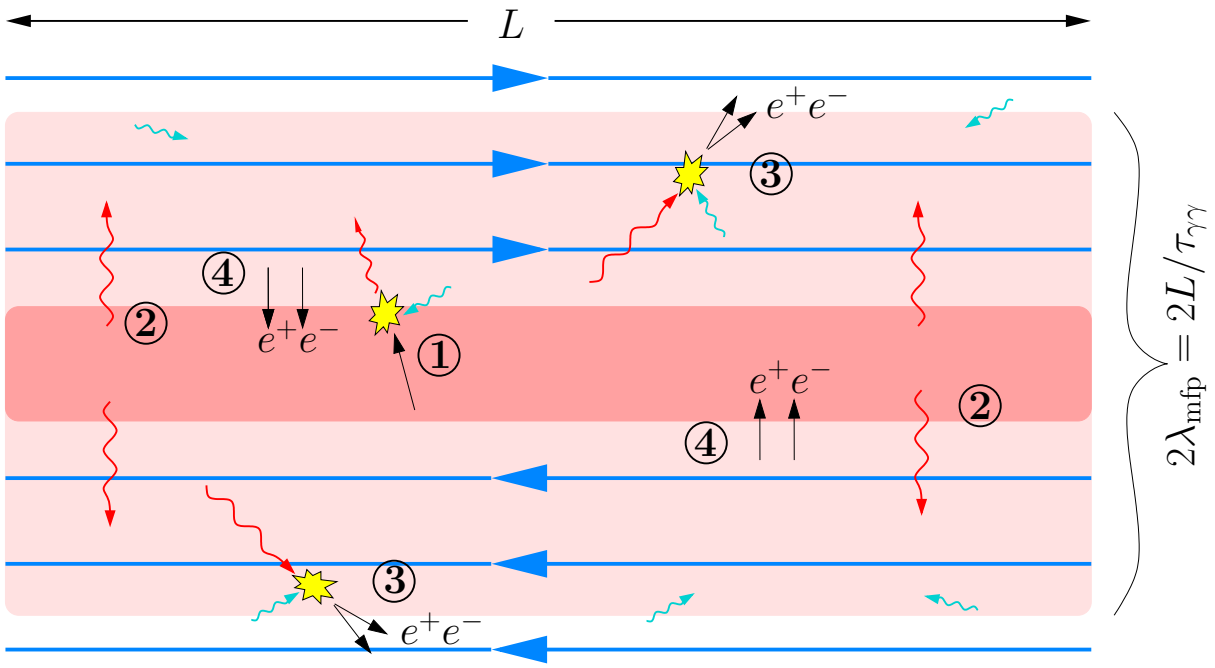

Figure 7.1: Schematic of Klein-Nishina radiative magnetic reconnection when the system is optically thick $\tau_{\gamma\gamma} \gg 1$ to pair-production. In this regime, the mean free path $\lambda_{\rm mfp}$ of a high-energy photon is less than the system size L: $\lambda_{\rm mfp} = L/\tau_{\gamma\gamma} \ll L$. Thus, every high-energy photon (red wiggled arrows) annihilates against an ambient seed photon (cyan wiggled arrows) before escaping the system. Newborn pairs are denoted by black arrows and magnetic field lines by blue arrows. The reconnection layer is opaque pink; the region penetrated by high-energy photons (one mean free path away from the layer) is transparent pink. See text for a description of stages $1-4$.

energy and number densities, the reconnection layer in this regime is self-consistently fed by a few high-energy newborn particles that control only the energy of inflowing material. Hence, pair-regulation in Klein-Nishina reconnection modifies the effective *hot* magnetization, $\sigma_{\rm h}$, of material feeding the layer, but not the *cold* magnetization, $\sigma_{\rm c}$.

7.2 Pair-regulated Klein-Nishina reconnection: technical aspects

I devote the remainder of this chapter to describing the technical analysis that underlies the qualitative picture described above. Before diving in, I state the following basic assumptions to clarify the region of the radiative phase diagram (in the sense of Fig. 6.7) to which the ensuing analysis pertains:

(1) Radiation takes place in the Klein-Nishina regime, where $\gamma_{\rm KN} < \gamma_{\rm rad,IC}$ and $\gamma_{\rm KN} < \gamma_{\rm X}$.

(2) The reconnection region is radiatively efficient, with all particles accelerated above $\gamma_{\rm cool,1} \simeq \gamma_{\rm cool}$ cooling in less than a dynamical time L/c and most particles reaching these energies: $\langle\gamma\rangle \gg \gamma_{\rm cool}$.

(3) The pair-production mean free paths $\lambda_{\rm mfp}$ of all gamma-rays above pair threshold are

   (a) independent of photon energy, and

   (b) between the full thickness Δ of *radiation zones* – the parts of the reconnection layer where above-threshold photons are produced – and the layer's full length L, i.e., $\Delta \ll \lambda_{\rm mfp} \ll L$.

As a reminder, I refer to as the 'layer' the region of the system threaded with reconnected magnetic flux.

Assumption 1 restricts to the Klein-Nishina – i.e., blue or red – region of the radiative phase diagram (Fig. 6.7). Assumption 2 excludes the white and green regions, implying that all of the accelerated particles – from the average energy $\langle\gamma\rangle$ to the cutoff energy – are between $\gamma_{\rm cool,1}$ and $\gamma_{\rm cool,2}$, and hence are strongly cooled. Statement 3a is not strictly true, but the pair-production

cross section $\sigma_{\gamma\gamma}(\epsilon)$ varies relatively weakly with energy beyond its peak $\sigma_{\gamma\gamma}(\epsilon) \simeq \sigma_{\rm T}/5$ when $\epsilon \simeq 3.6(m_e c^2)^2/\epsilon_{\rm ph}$. For example, $\sigma_{\gamma\gamma}(100 m_e^2 c^4/\epsilon_{\rm ph}) \simeq 0.08\sigma_{\rm T}$. Finally, the inequality $\lambda_{\rm mfp} \ll L$ in assumption 3b means that almost all above-threshold photons produced in the system are also absorbed in the system, and $\Delta \ll \lambda_{\rm mfp}$ further means that absorption predominantly occurs in the inflow (upstream) plasma. Note that I distinguish between the effective full thickness, Δ, of the reconnection layer itself, which could be taken as the width of the largest plasmoids $\simeq 0.1L$ (Uzdensky et al., 2010), and the thickness of the radiation zone, which (as discussed below in section 7.2.1) could be much thinner, even approaching the thicknesses of interplasmoid current layers. I illustrate the difference between Δ and $0.1L$ in Fig. 7.2.

In addition to all of these assumptions, I ignore effects due to synchrotron radiation. These enter at energy scales $\gamma > \gamma_{\rm s}$ [see equation (6.18) and its surrounding discussion]. I estimate $\gamma_{\rm s}$ for certain astrophysical systems, and comment on the consequent limitations on the applicability of my model, in chapter 8.

My technical investigation of the basic features of Klein-Nishina reconnection proceeds as follows. I begin (section 7.2.1) with some relatively simple energy-budget arguments. Based on energy considerations alone, I show that a self-regulated steady state or limit cycle should emerge – irrespective, even, of whether a pair cascade develops in the upstream region. I then decorate this basic picture by analyzing the number of produced pairs. This shows (section 7.2.2) that an exponential pair cascade, with each generation containing a constant factor > 1 more particles than the previous one, is not expected except for (almost unrealistically) efficient particle acceleration in the reconnection layer. I further apply detailed information on the distribution of newborn pairs, showing that (section 7.2.3), for $\sigma_{\rm c,0} \ll \gamma_{\rm KN}$, these should be fewer than those originally present in the upstream region.

7.2.1 The large energy density of newborn upstream pairs

Because I assume a radiatively efficient reconnection layer (assumption 2), a sizable fraction (e.g., one half; Werner et al., 2019) of the inflowing Poynting flux is promptly emitted. A fraction $\mathcal{F}$

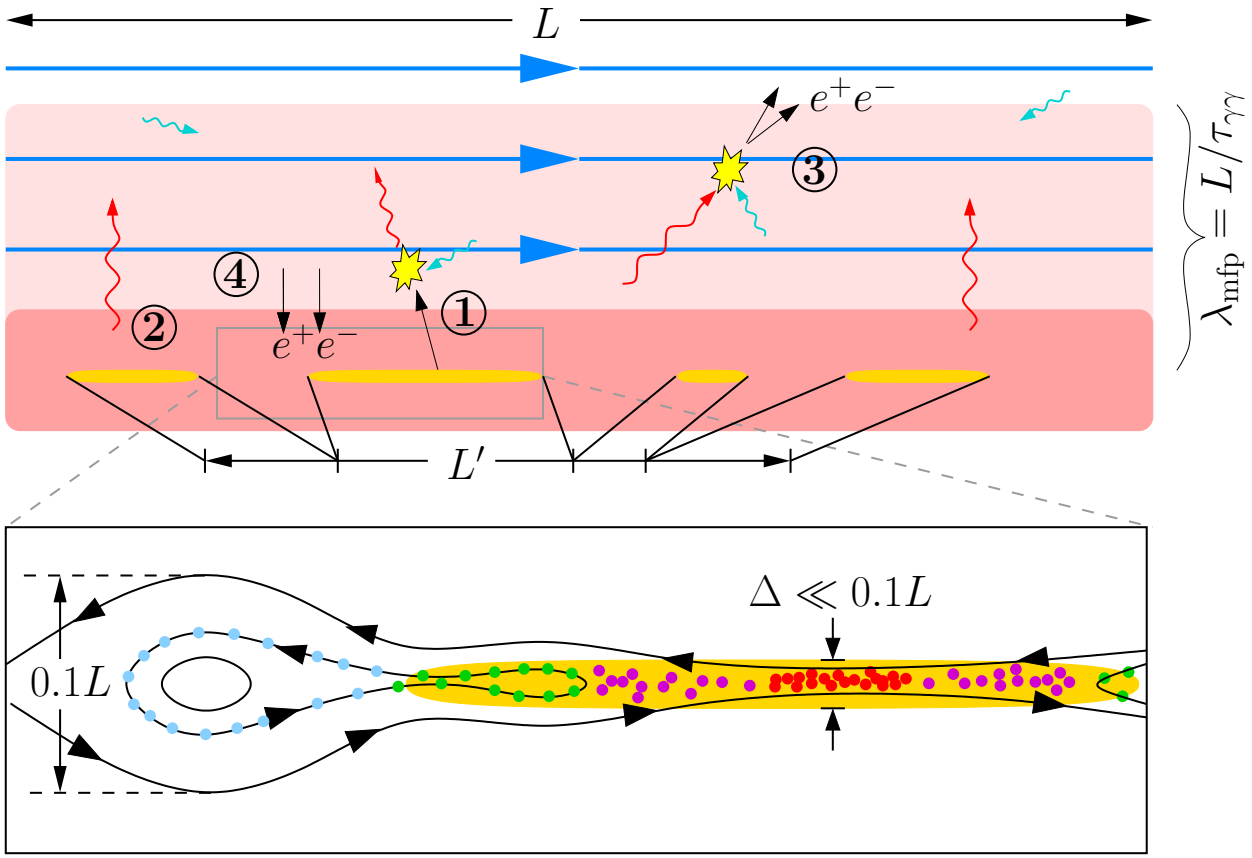

Figure 7.2: A more detailed view of the Klein-Nishina radiative reconnection system. The layer (region threaded with reconnected flux) contains subregions – *radiation zones* – where particles are radiating photons above pair-production threshold (gold). These occupy a combined fraction $L'/L < 1$ of the layer's full length. They are also thin, having full transverse width, $\Delta \ll 0.1L$, much smaller than the expected size of the largest plasmoids. A detailed view of a hypothetical radiation zone is displayed on the lower half of the plot. The thinness of the region ultimately stems from the kinetic-scale current sheets where particles (red dots) are accelerated near reconnection X-points. Following their impulsive acceleration in these locations, the particles (magenta dots) are magnetized by the reconnected magnetic field, which carries them away from the vicinity of an X-point. I concentrate on a subset of these particles (green dots) confined to one reconnected field line. Before these particles are able to fill an entire ring in a large plasmoid where they eventually end up, they cool down below the minimum energy $\gamma_{\rm pp}$ (light blue dots) to radiate photons above pair-production threshold. The transition point where particles cool below $\gamma_{\rm pp}$ – where the coloring changes from green to light blue – determines both the length of the radiation zone (distance from the X-point) and its width, i.e., the spread in red/magenta/green particles (those with $\gamma > \gamma_{\rm pp}$) about the midplane.

of the radiated energy lies above pair threshold with the ambient photon bath. This fraction penetrates a distance $\lambda_{\rm mfp}$ [assumption 3b] into the upstream plasma on both sides of the layer.[1] There, it is recaptured as newborn hot pairs and, ultimately, readvected into the layer. If the energy density of fresh pairs is high enough, the overall enthalpy density $w = w_0 + w_{\gamma\gamma}$ of inflowing material substantially increases. This reduces the effective hot magnetization

$$\sigma_{\rm h} \equiv \frac{B_0^2}{4\pi w} \tag{7.1}$$

below $\sigma_{\rm h,0} = B_0^2/4\pi w_0$ [equation (2.7)], which characterizes the *far upstream* region (beyond $\lambda_{\rm mfp}$ from the layer). In my convention, subscript '0' denotes far upstream quantities and subscript '$\gamma\gamma$' quantities sourced by pair creation within $\lambda_{\rm mfp}$ of the layer. Corresponding naked symbols (e.g., $\sigma_{\rm h}$ or w) are decided by a combination of pair-creation-sourced and far upstream values.

A reduced effective $\sigma_{\rm h}$ may strongly suppress the efficiency of non-thermal particle acceleration (NTPA) in the layer (e.g., Sironi & Spitkovsky, 2014; Sironi et al., 2016; Guo et al., 2014, 2015; Werner et al., 2016; Werner & Uzdensky, 2017; Werner et al., 2018; Ball et al., 2018). This enables a negative feedback loop, in which a layer fed initially by highly magnetized ($\sigma_{\rm h,0} \gg 1$) plasma efficiently accelerates particles to gamma-ray emitting energies. The gamma-rays, in turn, produce pairs in the inflow region, reducing its effective magnetization and, hence, suppressing subsequent NTPA (cf. Hakobyan et al. 2019; see Fig. 7.1). In this section, I calculate the fixed point $\sigma_{\rm h}$ for this feedback loop. Additionally, I determine the conditions governing whether the system asymptotically approaches its fixed point in a late-time steady state. I further show that, if the fixed point is not reached, the system exhibits undamped, large-amplitude cycles of copious pair creation followed by shutdown of NTPA.

The Poynting flux delivered to the reconnection layer (per unit length in the out-of-plane direction) is

$$P_{\rm Poynt} \sim 2L\beta_{\rm rec} c \frac{B_0^2}{4\pi} \,. \tag{7.2}$$

[1] For simplicity, I ignore kinetic beaming (Uzdensky et al., 2011; Cerutti et al., 2012b; Mehlhaff et al., 2020), which produces potentially important anisotropy in the distributions of high-energy particles and their emitted photons. I comment on expected consequences of this beaming in Appendix C but ultimately defer its full treatment to a future simulation study.

The leading factor of 2 results from Poynting flux entering the reconnection region from two directions. If half of this power is given to particles that quickly [within L/c; assumption 2] radiate it away through the IC process, the volume-averaged IC emissivity $j_{\rm IC}$ (power radiated per unit volume) in the reconnection layer satisfies

$$j_{\rm IC}\Delta L' \sim \frac{1}{2}P_{\rm Poynt} \sim L\beta_{\rm rec}c\frac{B_0^2}{4\pi}\,, \tag{7.3}$$

where L' is the combined length of all radiation zones in the reconnection layer. One can ignore all plasmoid/current-sheet substructure, taking the entire layer to be one large radiation zone, by setting $L' = L$. However, given my assumption 2 of a radiatively efficient reconnection system, L' may actually be shorter than L. This is because particles may cool to below the minimum energy, $\gamma_{\rm pp} \equiv 8\gamma_{\rm KN}$ (section 6.2.1), to emit pair-producing photons before traveling far from their primary X-point acceleration sites. Moreover, as particles travel away from an X-point, they also spread out about the reconnection midplane. Thus, a cooling limit on the combined length of radiation zones (such that $L' < L$) also limits their effective thickness, Δ, potentially keeping them much thinner than the characteristic large-plasmoid width (e.g., $\Delta < 0.1L$; Fig. 7.2).

To determine the total enthalpy density w and, from it, the effective magnetization $\sigma_{\rm h}$ [equation (7.1)], one needs to know the fraction $\mathcal{F}$ of power radiated away from the reconnection layer above pair-production threshold (and, hence, captured in the upstream region as electron-positron pairs). Using $\gamma_{\rm pp}$ along with the distribution function of radiating layer particles $\mathrm{d}N/\,\mathrm{d}\gamma$, $\mathcal{F}$ reads

$$\mathcal{F} = \frac{\int_{\gamma_{\rm pp}}^{\infty}\mathrm{d}\gamma\;\mathrm{d}N/\,\mathrm{d}\gamma\,P_{\rm IC}(\gamma)}{\int_{1}^{\infty}\mathrm{d}\gamma\;\mathrm{d}N/\,\mathrm{d}\gamma\,P_{\rm IC}(\gamma)}\,. \tag{7.4}$$

To evaluate $\mathcal{F}$, I insert a power-law reconnection-energized pair-plasma distribution:

$$\frac{\mathrm{d}N}{\mathrm{d}\gamma} = A\begin{cases}\gamma^{-p} & \gamma_1 \le \gamma \le \gamma_2\\ \\ 0 & \text{otherwise}\end{cases}\,, \tag{7.5}$$

where A is a normalization factor and $\gamma_1 \ll \gamma_{\rm KN}$ is assumed. If $p < 3$, the γ^2-dependence of $P_{\rm IC}(\gamma)$ when $\gamma \ll \gamma_{\rm KN}$ suppresses the dependence of $\mathcal{F}$ on the onset energy $\gamma_1 \ll \gamma_{\rm KN}$ of the power law, and γ_1 can thus be taken to unity. If, instead, $p \ge 3$, the onset energy, γ_1, can also be ignored

– the same dependence, $P_{\rm IC}(\gamma \ll \gamma_{\rm KN}) \propto \gamma^2$, pushes $\mathcal{F}$ to zero independently of γ_1. Thus, my assumption $\gamma_1 \ll \gamma_{\rm KN}$ is equivalent to setting $\gamma_1 = 1$.

Substituting, now, (6.4) and (7.5) into (7.4), as well as putting $\gamma_{\rm pp} = 8\gamma_{\rm KN}$ and $\gamma_1/\gamma_{\rm KN} = 0$, gives

$$\mathcal{F}(p,z) = \frac{\int_8^z \mathrm{d}x\, x^{-p+2} f_{\rm KN}(x)}{\int_0^z \mathrm{d}x\, x^{-p+2} f_{\rm KN}(x)}\,, \tag{7.6}$$

where $z \equiv \gamma_2/\gamma_{\rm KN}$. Fig. 7.3 displays $\mathcal{F}(p,z)$ computed according to (7.6). The graphs confirm the above argument that $\lim_{p\to 3}\mathcal{F}(p,z) = 0$ for all z. Furthermore, because $f_{\rm KN}(x) \sim \ln(x)/x^2$ as $x \to \infty$, when $p \le 1$, the integrals in (7.6) diverge with z, but in such a way that $\mathcal{F} = 1$. This signals that virtually all radiation from the layer is emitted above pair threshold. Fig. 7.3 also shows that, modulo a strong z-dependence near pair threshold $z = 8$, $\mathcal{F}(p,z)$ becomes nearly z-independent once $z \gtrsim 12$. Essentially, $\mathcal{F}(p, z \gtrsim 12) \simeq \mathcal{F}(p,\infty)$.

Next, I explicitly connect the fraction $\mathcal{F}$ to the effective hot magnetization $\sigma_{\rm h}$. The power $\mathcal{F} j_{\rm IC}\Delta L'$ shining out of the reconnection layer's radiation zones penetrates a distance $\lambda_{\rm mfp}$ back into the upstream area before being deposited as hot pairs. Assuming this deposition is approximately uniform in space up to a distance $\lambda_{\rm mfp}$ above and below the reconnection layer, hot pairs add to the upstream plasma energy density at a rate $\mathrm{d}u_{\gamma\gamma}/\,\mathrm{d}t$ satisfying

$$\begin{aligned} &\frac{\mathrm{d}u_{\gamma\gamma}}{\mathrm{d}t}(2\lambda_{\rm mfp})L \sim \mathcal{F} j_{\rm IC}\Delta L' \\ &\qquad \Rightarrow \frac{\mathrm{d}u_{\gamma\gamma}}{\mathrm{d}t} \sim \frac{\mathcal{F} j_{\rm IC}}{2}\frac{\Delta}{\lambda_{\rm mfp}}\frac{L'}{L} \sim \mathcal{F}\frac{\beta_{\rm rec}c}{\lambda_{\rm mfp}}\frac{B_0^2}{8\pi}\,. \end{aligned} \tag{7.7}$$

In the first line, I assume that, as the radiation propagates away from the layer, it also fills in the gaps between radiation zones so that the upstream region receives pairs approximately uniformly across its length L. The second line in (7.7) is obtained from the first by plugging in equation (7.3). The factor of 2 accounts for radiated energy being absorbed both below and above the layer.

Consider a plasma parcel with initial energy density u_0 that starts far upstream, $|y| \gg \lambda_{\rm mfp}$, of the layer. The parcel is advected inward at transverse velocity $v_y = -\,\mathrm{sign}(y)\beta_{\rm rec}c \simeq -\,\mathrm{sign}(y)0.1c$, and, upon reaching the pair-creation zone, $|y| \sim \lambda_{\rm mfp}$, begins accruing additional energy at the

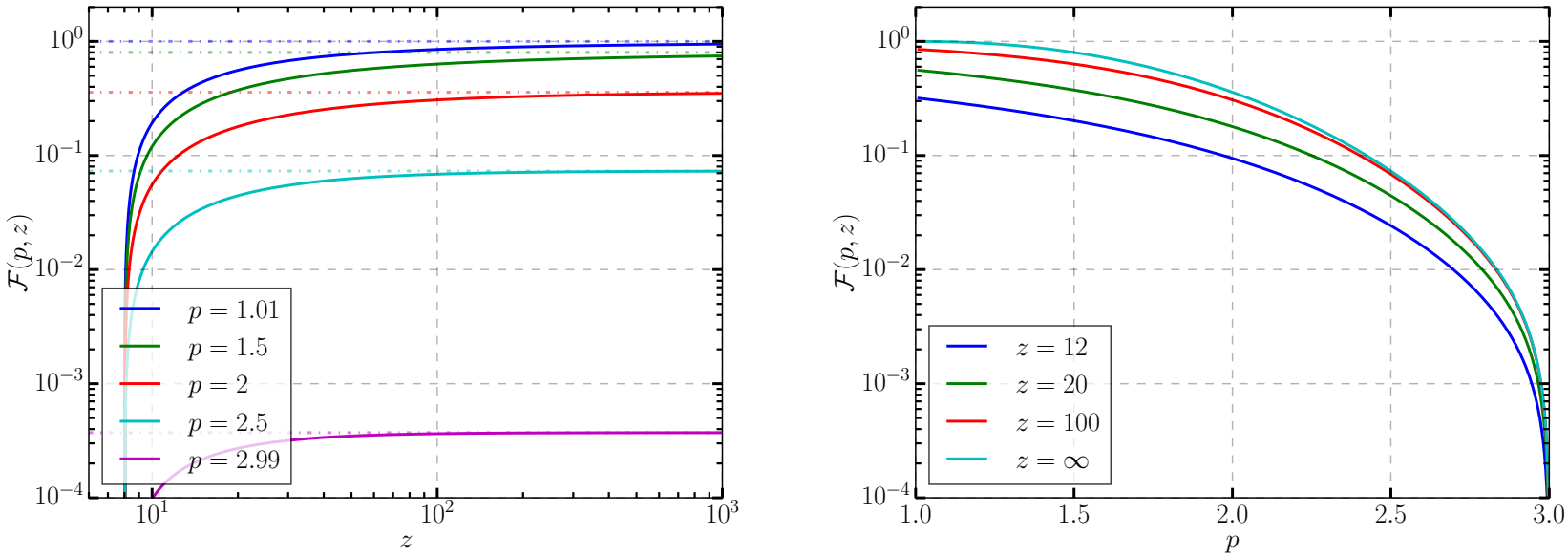

Figure 7.3: Top: A plot of $\mathcal{F}(p, z)$ as a function of $z \equiv \gamma_2/\gamma_{\rm KN}$ for several values of p. Dot-dashed lines indicate $\lim_{z\to\infty} \mathcal{F}(p, z)$. Bottom: A plot of $\mathcal{F}(p, z)$ as a function of p for several values of z. The fraction $\mathcal{F}$ is strongly dependent on p but not on z (at least after the threshold value $z = 8$ is crossed).

rate $\mathrm{d}u_{\gamma\gamma}/\,\mathrm{d}t$. The extra energy acquired in transit from $|y| \sim \lambda_{\rm mfp}$ to the layer ($|y| \sim \Delta$) is simply

$$u^{+}_{\gamma\gamma} \sim \frac{1}{\beta_{\rm rec}c}(\lambda_{\rm mfp} - \Delta)\frac{\mathrm{d}u_{\gamma\gamma}}{\mathrm{d}t} \simeq \frac{\lambda_{\rm mfp}}{\beta_{\rm rec}c}\frac{\mathrm{d}u_{\gamma\gamma}}{\mathrm{d}t} \sim \mathcal{F}\frac{B_0^2}{8\pi}\,. \tag{7.8}$$

The accumulated internal energy density $u^{+}_{\gamma\gamma}$ is less than $B_0^2/8\pi$ because $\mathcal{F} \leq 1$, but it can still far exceed u_0 given sufficient magnetization $\sigma_{\rm h,0} \gg 1$. The superscript '+' denotes that this is only energy *added* to the plasma; I have not yet considered that some energy may be lost en route to the layer – either through radiation or because particles physically escape the system.

Importantly, the 'readvection time' $t_{\rm ra} \equiv \lambda_{\rm mfp}/\beta_{\rm rec}c$ cancels in (7.8). Thus, whether the pair-creation zone is truly confined to transverse distances $|y| \sim \lambda_{\rm mfp}$ or occupies a much larger region (for example, for an N-generation pair cascade, one expects $|y| \sim \sqrt{N}\lambda_{\rm mfp} \gg \lambda_{\rm mfp}$ – a possibility that I entertain in section 7.2.2), $u^{+}_{\gamma\gamma}$ remains approximately the same. For reference, the readvection time is related to the global dynamical time L/c through

$$t_{\rm ra} \equiv \frac{\lambda_{\rm mfp}}{\beta_{\rm rec}c} \simeq \frac{10}{\tau_{\gamma\gamma}}\frac{L}{c}\,, \tag{7.9}$$

where I used $\lambda_{\rm mfp} = L/\tau_{\gamma\gamma}$ and $\beta_{\rm rec} \simeq 0.1$. Note that the prefix 're' in 'readvection' applies only to the energy, which is captured *again* by the reconnection layer. The pairs that carry this energy, by contrast, are advected into the layer for their first time.

I now estimate $u_{\gamma\gamma}$, the energy density retained by the fresh plasma swept into the reconnection layer. This yields the enthalpy density $w_{\gamma\gamma}$ and, through equation (7.1), the effective hot magnetization $\sigma_{\rm h}$. Now, $u_{\gamma\gamma}$ is less than the deposited energy density $u^{+}_{\gamma\gamma}$ because, while traveling to the layer, newborn pairs may both radiatively cool and escape the system. To account for this, I define the *energy recapture efficiency*, $\xi \equiv u_{\gamma\gamma}/u^{+}_{\gamma\gamma} \leq 1$, and write

$$\xi = f_{\rm nocool}f_{\rm noesc}\,. \tag{7.10}$$

Here, $f_{\rm nocool}$ and $f_{\rm noesc}$ are, respectively, the fraction of the accumulated energy that is not radiated away ($f_{\rm nocool}$) and that is not lost through escaping particles ($f_{\rm noesc}$).

I calculate the cooling factor $f_{\rm nocool}$ in detail in Appendix B. There, I identify a physically allowed range $f_{\rm nocool} \in [3/400, 1]$ and show how, within this interval, $f_{\rm nocool}$ depends on the other

parameters in the problem (on the effective magnetization $\sigma_{\rm h}$ and on the cutoff z). While that calculation allows me to compute $\sigma_{\rm h}$ self-consistently (since, in reality, $f_{\rm nocool}$ depends on $\sigma_{\rm h}$), it is mathematically complicated. Furthermore, I find that the main qualitative features of self-regulated Klein-Nishina reconnection are captured by treating $f_{\rm nocool}$ as an independent parameter and scanning it across the allowed interval $[3/400, 1]$. That is the approach I adopt in this section.

In addition to this simplified prescription for $f_{\rm nocool}$, I set the escape factor $f_{\rm noesc}$ to unity, effectively putting $\xi = f_{\rm nocool}$. This is what one expects if the time L/c for a relativistic particle to stream out of the system is longer than the readvection time (7.9), which is true for $\tau_{\gamma\gamma} \gtrsim 10$ (and hence for a broad range of radiative parameters). I comment more thoroughly on the many additional kinetic effects that may influence $f_{\rm noesc}$ in Appendix C. However, because most of these effects tend to push $f_{\rm noesc}$ toward unity, I simply leave $f_{\rm noesc} \simeq 1$ from here onward.

Using ξ, the energy density of fresh pairs entering the reconnection layer is

$$u_{\gamma\gamma} = \xi u^{+}_{\gamma\gamma} \sim \xi \mathcal{F} \frac{B_0^2}{8\pi} \,. \tag{7.11}$$

If these pairs are relativistically hot, then $p_{\gamma\gamma} = u_{\gamma\gamma}/3$ and $w_{\gamma\gamma} = p_{\gamma\gamma} + u_{\gamma\gamma} \simeq (4/3)u_{\gamma\gamma}$; otherwise $w_{\gamma\gamma} = u_{\gamma\gamma}$. I take $w_{\gamma\gamma} = (4/3)u_{\gamma\gamma}$ – still a good approximation in the non-relativistic limit.

The effective inflowing plasma magnetization $\sigma_{\rm h}$ is then

$$\sigma_{\rm h} \equiv \frac{B_0^2}{4\pi(w_0 + w_{\gamma\gamma})} = \frac{B_0^2/4\pi w_0}{1 + w_{\gamma\gamma}/w_0} \sim \frac{\sigma_{\rm h,0}}{1 + 2\xi\mathcal{F}\sigma_{\rm h,0}/3} \,. \tag{7.12}$$

Equation (7.12) encodes two main possible fixed points for $\sigma_{\rm h}$. The first is when $\xi\mathcal{F}\sigma_{\rm h,0} \ll 1$. Then, pair-production is too inefficient to load the upstream plasma substantially and the solution to (7.12) is simply $\sigma_{\rm h} \simeq \sigma_{\rm h,0}$. The other regime is when $\xi\mathcal{F}\sigma_{\rm h,0} \gg 1$. In this situation, hot pairs suppress $\sigma_{\rm h}$ to a universal value

$$\sigma_{\rm h} \sim \frac{3}{2\xi\mathcal{F}} \,, \tag{7.13}$$

which is entirely independent of $\sigma_{\rm h,0}$. Not only is (7.13) universal, but, in principle, it can be solved to yield self-consistent values of $\sigma_{\rm h}$ and p. This is because the effective magnetization governs

the efficiency of NTPA (cf. Werner et al., 2016; Werner & Uzdensky, 2017; Werner et al., 2018; Ball et al., 2018) and ultimately specifies the power-law index p. One only needs to know the reconnection NTPA 'equation of state', $p(\sigma_{\rm h})$.

Let me assume that a suitable $p(\sigma_{\rm h})$ can be borrowed from non-radiative reconnection studies. I take

$$p(\sigma_{\rm h}) = 1 + 2/\sqrt{\sigma_{\rm h}}\,, \tag{7.14}$$

which can be obtained from fitting the data in fig. 3 of Werner et al. (2016) to the general form $p = C_0 + C_1/\sqrt{\sigma_{\rm h}}$ used by Werner & Uzdensky (2017) (see also Werner et al., 2018; Ball et al., 2018). I acknowledge that the distribution $\mathrm{d}N/\,\mathrm{d}\gamma$ in equation (7.5) is the instantaneous distribution of radiating particles in the reconnection layer, which – in this radiative context – may differ from the injected (non-radiative) power-law distribution characterized by $p(\sigma_{\rm h})$. Later on, I account for approximate radiative modifications to the distribution of emitting particles. For mathematical transparency, however, in this first calculation, I plug (7.14) directly into the expression for $\mathcal{F}$.

To simplify further, I take $z = \gamma_2/\gamma_{\rm KN} \to \infty$ when evaluating $\mathcal{F}$ even though calculating $\sigma_{\rm h}$ and p runs the same for any z. As previously remarked, the fraction $\mathcal{F}(p, z)$ is relatively z-independent as long as $z \gtrsim 12$, so taking $z \to \infty$ gives a solution representing a wide range of likely values (i.e., almost all values beyond those very close to the threshold $z = 8$ for pair production to turn on).

I now solve (7.13) and (7.14) for a variety of ξ values and graphically present the solutions in Fig. 7.4. A lower ξ (lower $f_{\rm nocool}$) increases radiative cooling of newborn pairs as they travel toward the reconnection layer. This diminishes their enthalpy density, $w_{\gamma\gamma}$ (which, nevertheless, still dominates over the initial plasma because $w = w_0 + w_{\gamma\gamma} \simeq w_{\gamma\gamma}$), enhancing the effective magnetization $\sigma_{\rm h}$, and, through $p(\sigma_{\rm h})$, hardening the resulting distribution of reconnection-energized particles.

In Fig. 7.4, I solve equations (7.13) and (7.14) rather than the more general form (7.12). This presupposes that the solution $\sigma_{\rm h}$ is much smaller than the original (far upstream) hot mag-

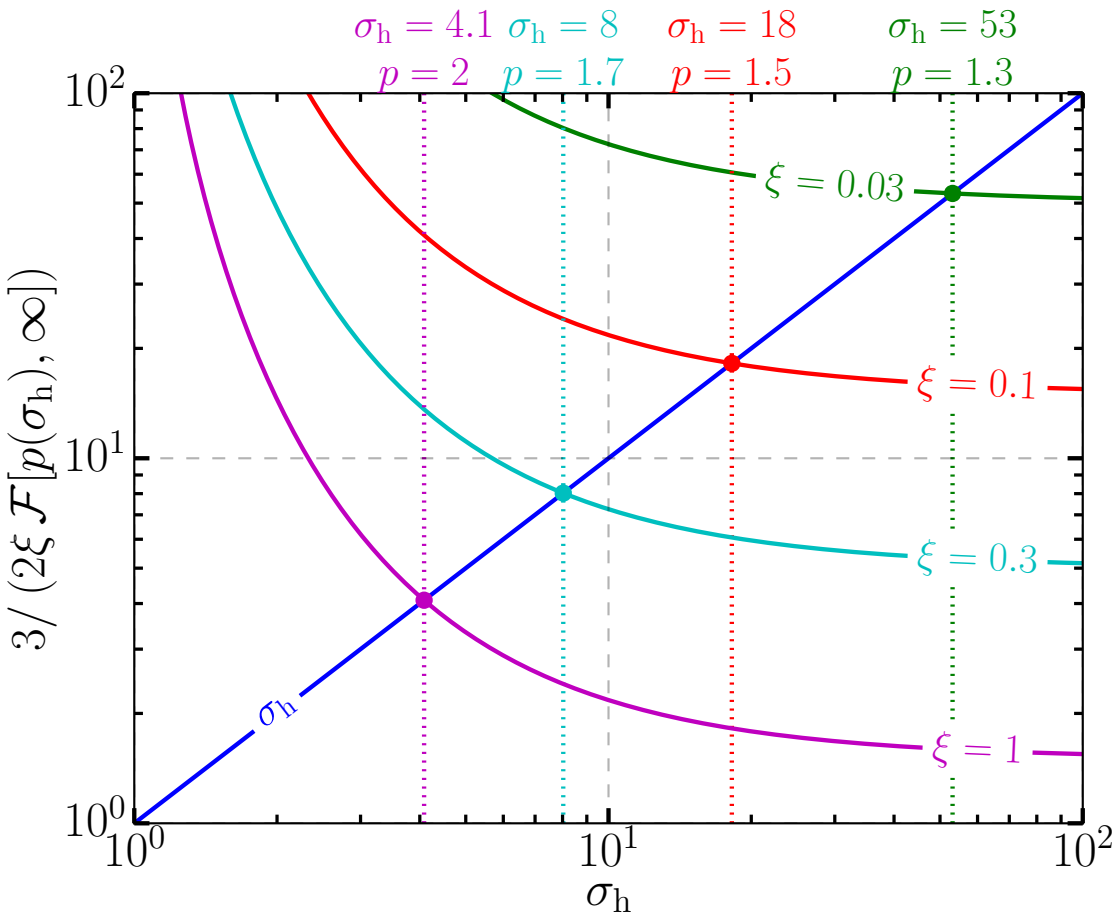

Figure 7.4: Solutions to equations (7.13) and (7.14) for several ξ values. Lower ξ causes the pairs born into the upstream region to cool more. This somewhat inhibits the feedback mechanism by reducing the overall enthalpy density $w = w_0 + w_{\gamma\gamma} \simeq w_{\gamma\gamma}$ of plasma arriving at the layer.

netization $\sigma_{\rm h,0}$. To illustrate the effect of a finite $\sigma_{\rm h,0}$, I also display solutions to equation (7.12) for $\sigma_{\rm h,0} = 100$ in Fig. 7.5. As expected, a finite $\sigma_{\rm h,0}$ has relatively little impact on the value of $\sigma_{\rm h}$ when $\sigma_{\rm h} \ll \sigma_{\rm h,0}$ – the universal regime in which the solution is insensitive to the far upstream magnetization. However, as the resulting solution $\sigma_{\rm h}$ gets closer to $\sigma_{\rm h,0}$, the approximate solution obtained from (7.13) becomes less accurate. This occurs roughly when $\xi \sim 1/\sigma_{\rm h,0}$ (compare, for example, the solutions obtained for $\xi = 0.03 = 3/\sigma_{\rm h,0}$ in Figs. 7.4 and 7.5). In addition, Fig. 7.5 illustrates the stability of the fixed point $\sigma_{\rm h}$, which is the topic of the next section.

7.2.1.1 Stability of the pair-loaded $\sigma_{\rm h}$

Now that it has been seen how to calculate the fixed point $\sigma_{\rm h}$, one can begin to ask whether a system that starts from the initial magnetization $\sigma_{\rm h,0}$ actually approaches $\sigma_{\rm h}$ at some late time. I call the fixed point *stable* if the plasma feeding the reconnection layer approaches a quasi-steady magnetization $\sigma_{\rm h}$, and *unstable* otherwise. Where necessary, I further distinguish *global stability*, which refers to the notion of stability just described, from *local stability*, which is simpler, and only determines whether a system that starts at a magnetization some infinitesimal distance $\delta\sigma_{\rm h}$ away from $\sigma_{\rm h}$ approaches $\sigma_{\rm h}$.

To discuss stability quantitatively, I abbreviate the right-hand-side of (7.12) as

$$h(x) \equiv \frac{\sigma_{\rm h,0}}{1 + 2\xi\, \mathcal{F}[p(x), \infty]\, \sigma_{\rm h,0}/3}\,. \tag{7.15}$$

A system that begins with initial magnetization $\sigma_{\rm h,0}$ (i.e., before any pairs have been produced) accelerates a distribution of particles in the reconnection layer with a power-law index $p(\sigma_{\rm h,0})$ given by equation (7.14). These particles then radiate photons, some of which are above pair threshold with the background radiation and consequently – after a time $\sim \lambda_{\rm mfp}/c$ – pair-produce somewhere in the zone $|y| \lesssim \lambda_{\rm mfp}$. The newborn pairs are then advected toward the layer, reaching it after a readvection time $t_{\rm ra} = \lambda_{\rm mfp}/\beta_{\rm rec}c$. At that point the layer witnesses a new effective magnetization $\sigma_{\rm h,1}$ that is determined by $\sigma_{\rm h,0}$ via $\sigma_{\rm h,1} = h(\sigma_{\rm h,0})$. Assuming that the time for the layer to respond to a new magnetization is smaller than $\lambda_{\rm mfp}/\beta_{\rm rec}c$, the lag time between

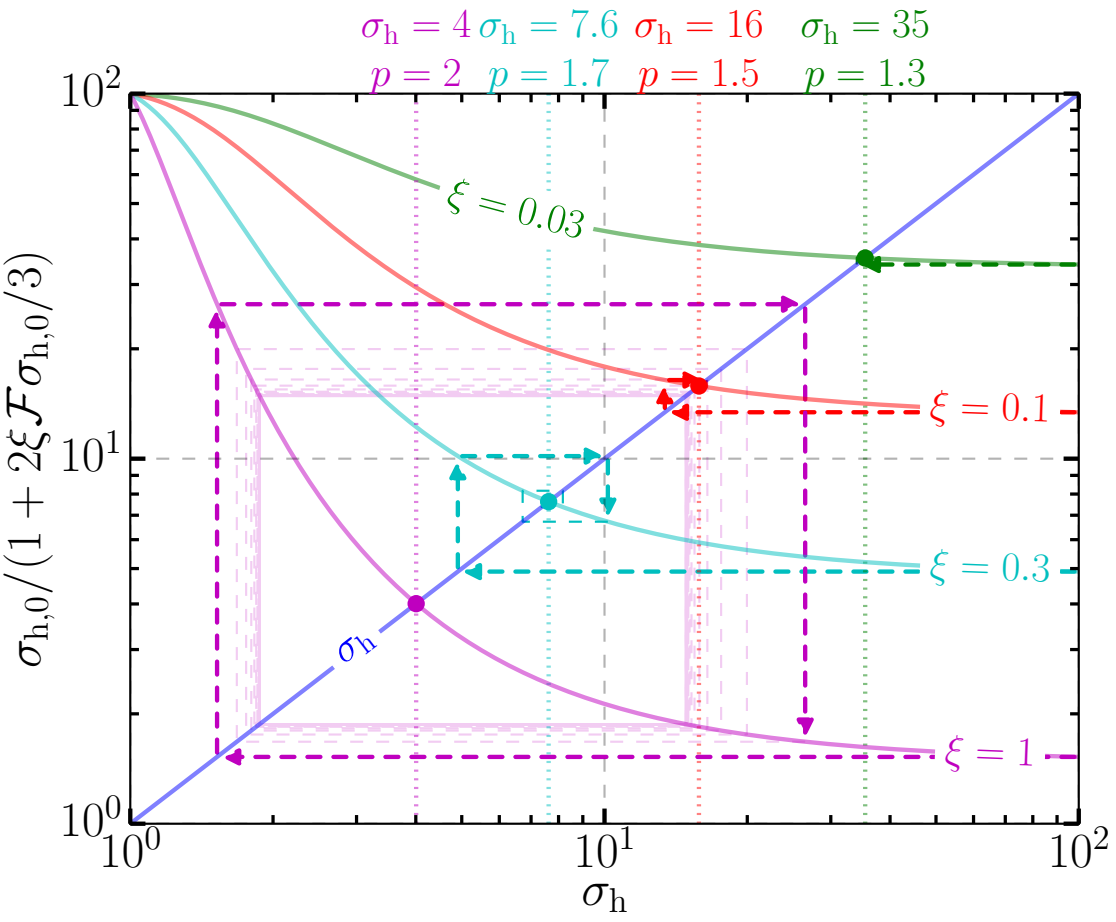

Figure 7.5: Solutions to the more general equation (7.12) and equation (7.14) using $\sigma_{h,0} = 100$. Unlike equation (7.13), equation (7.12) does not assume *a priori* that $\sigma_h \ll \sigma_{h,0}$. The figure also depicts the stability of each solution. Systems starting with magnetization $\sigma_{h,0}$ transition to a new magnetization $\sigma_{h,1} = h(\sigma_{h,0})$ after approximately one readvection time $\lambda_{mfp}/\beta_{rec}c$. This is illustrated for each ξ by a horizontal line running left from $h(\sigma_{h,0})$ to the corresponding new value of $\sigma_{h,1}$ on the blue diagonal [which represents the curve $x = h(x)$]. Then, after another readvection time, the second modified magnetization $\sigma_{h,2} = h(\sigma_{h,1})$ is reached. This is represented by both an upward-running dashed line from $\sigma_{h,1}$ to $h(\sigma_{h,1})$ and a horizontal line, from left to right, terminating on the corresponding value $\sigma_{h,2}$ on the diagonal. Further transitions are denoted by one vertical and one horizontal dashed line: either down-then-left or up-then-right. For low ξ, the system approaches the fixed point σ_h after just a few transitions. For high ξ, pair feedback is so efficient that the system gets stuck in a two-state swing cycle.

when the layer starts processing a $\sigma_{\rm h,0}$-plasma and when it starts to see a $\sigma_{\rm h,1}$-plasma is $\lambda_{\rm mfp}/c + \lambda_{\rm mfp}/\beta_{\rm rec}c \simeq \lambda_{\rm mfp}/\beta_{\rm rec}c$. Thus, the readvection time characterizes the transition period from the initial magnetization $\sigma_{\rm h,0}$ to the first modified magnetization $\sigma_{\rm h,1}$.

By similar reasoning, after about $n\lambda_{\rm mfp}/\beta_{\rm rec}c$, the layer witnesses effective magnetization

$$\sigma_{{\rm h},n} = h(\sigma_{{\rm h},n-1})\,. \tag{7.16}$$

Thus, the system approaches the fixed point $\sigma_{\rm h}$ as an asymptotic steady state if $\lim_{n\to\infty}\sigma_{{\rm h},n} = \sigma_{\rm h}$. Of course, if the convergence is slow, then the system may not actually reach the fixed point $\sigma_{\rm h}$ (even though it is stable) before reconnection finishes. All that is certain in that case is that oscillatory swings about $\sigma_{\rm h}$ are damped: with each successive readvection time, the system reaches a magnetization that is somewhat closer to the asymptotic steady state. However, if the long-time limit of the iterated map $\sigma_{{\rm h},n} = h(\sigma_{{\rm h},n-1})$ does not approach the fixed point $\sigma_{\rm h}$, then some other behavior occurs. That late-time outcome is large-amplitude, undamped oscillations about $\sigma_{\rm h}$ between two states, one with a low magnetization $\sigma_{\rm h,<} < \sigma_{\rm h}$ and one with a high value $\sigma_{\rm h,>} > \sigma_{\rm h}$. The high-$\sigma_{\rm h,>}$ state drives efficient NTPA in the reconnection layer and, consequently, pair production in the upstream region. When the created hot pairs reach the layer, they initiate the low-$\sigma_{\rm h,<}$ state, where NTPA is shut down, pair creation ceases, and, eventually (after another $\lambda_{\rm mfp}/\beta_{\rm rec}c$), the high-$\sigma_{\rm h,>}$ state is restored. The cycle repeats from there.

Both behaviors – an asymptotic steady state and a late-time two-state cycle – are illustrated in Fig. 7.5. For low energy recapture efficiency ξ, the system tends toward a steady state; when ξ is increased beyond a critical threshold, the system bifurcates, asymptotically favoring large amplitude swings. This can be understood as follows. For efficient feedback, so much of the energy radiated from the layer is caught in, and then recaptured from, the upstream region that the system overshoots the fixed point $\sigma_{\rm h}$ by a wide margin. This chokes subsequent pair-producing emission from the layer effectively enough that the system violently ricochets back to a high magnetization, getting caught between two extreme magnetizations $\sigma_{\rm h,<}$ and $\sigma_{\rm h,>}$ as described above.

I now show how this physical picture is encoded in the mathematical machinery. As is well-

known from the theory of iterated maps, the fixed point $\sigma_{\rm h} = h(\sigma_{\rm h})$ is locally stable if $\left|h'(\sigma_{\rm h})\right| \leq 1$ (meaning, as discussed before, that magnetizations $\sigma_{\rm h} + \delta\sigma_{\rm h}$ move toward $\sigma_{\rm h}$ with each successive iteration). Higher ξ lowers the curve $h(x)$, which increases the magnitude of the slope $h'(\sigma_{\rm h})$ (i.e., makes it more negative) at the location of the fixed point (see Fig. 7.5), pushing the system toward the instability threshold $\left|h'(\sigma_{\rm h})\right| = 1$ for swing cycles. Even though this is only a local stability condition, I argue in Appendix D that, for the restricted class of monotonically decreasing $h(x)$ functions, local instability is sufficient for global instability: $\sigma_{\rm h}$ is never approached by systems that start at $\sigma_{\rm h,0}$ when $\left|h'(\sigma_{\rm h})\right| > 1$. In that case, the late-time dynamics consist of a two-state swing cycle. In contrast, if $\left|h'(\sigma_{\rm h})\right| < 1$, the system may or may not converge toward $\sigma_{\rm h}$ depending on the details of the map function $h(x)$. Despite this uncertainty (and as also argued in Appendix D), at late times, all monotonically decreasing $h(x)$ functions result in either a steady state or a swing cycle. For such functions, *no other late-time dynamics are allowed.*

To illustrate these remarks, I supply Fig. 7.6. That figure displays the fixed point $\sigma_{\rm h}$ as a function of ξ. The system bifurcates at a critical value of $\xi = \xi_{\rm c} \simeq 0.84$: the fixed point $\sigma_{\rm h}$ becomes unstable [$\left|h'(\sigma_{\rm h})\right|$ goes above 1] and a new attractor for the late-time dynamics appears – a two-state cycle characterized by a low magnetization $\sigma_{\rm h,<} < \sigma_{\rm h}$ and a high magnetization $\sigma_{\rm h,>} > \sigma_{\rm h}$. Importantly, the condition $\left|h'\right| > 1$ implies that the system traps into a limit cycle at late times. While, in general, it is *not* the case that $\left|h'\right| < 1$ implies convergence to the fixed point $\sigma_{\rm h}$, for this particular system that happens to be true. (I discuss a case where this does not hold in Appendix B.)

7.2.1.2 Calculating $\sigma_{\rm h}$ with more realistic radiative feedback

Having established the mathematical techniques for computing the self-regulated magnetization $\sigma_{\rm h}$ and determining its stability, I now repeat my calculation with more physical realism. In particular, I modify my assumed layer particle distribution $\mathrm{d}N/\,\mathrm{d}\gamma$ from a single power law, as

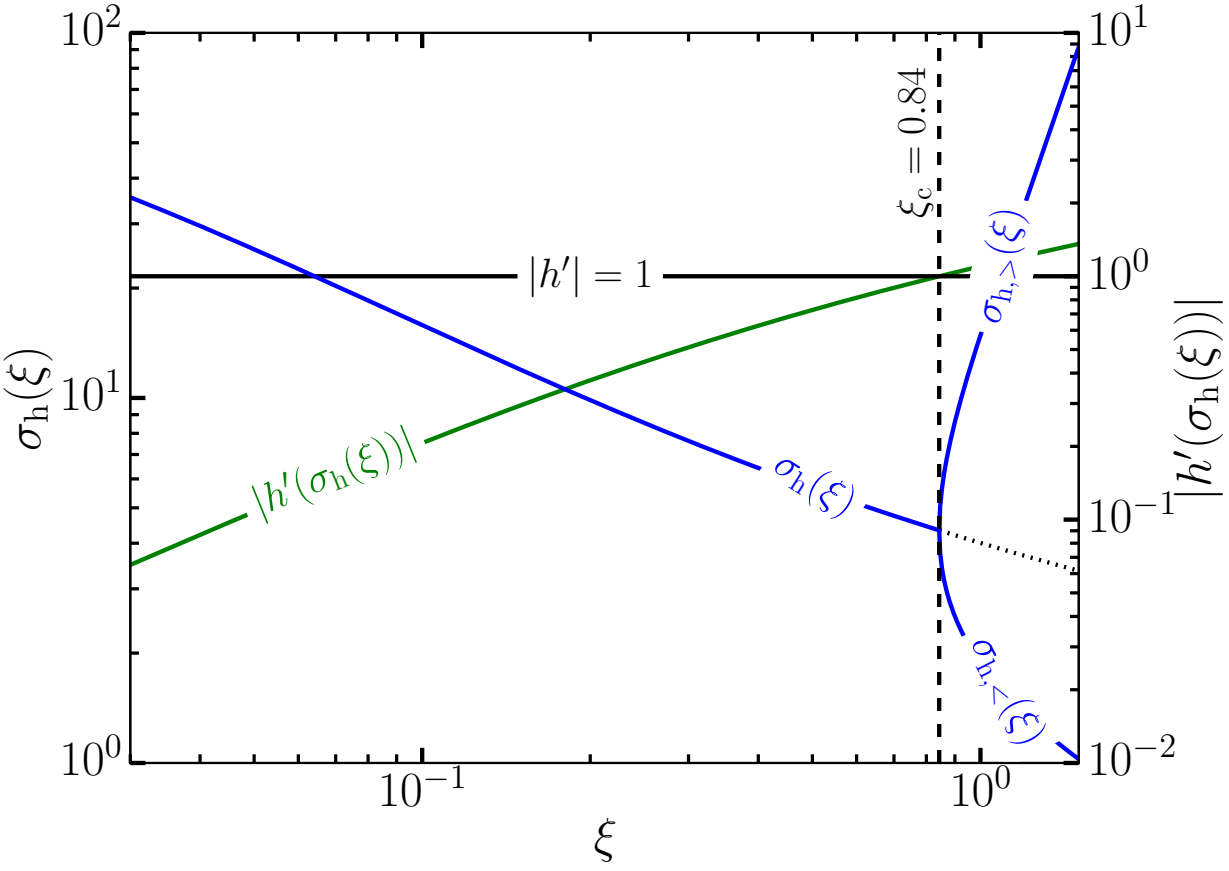

Figure 7.6: The fixed point σ_h of the reconnection layer [i.e., of the iterated map $\sigma_{h,n} = h(\sigma_{h,n-1})$ with $h(x)$ as defined in equation (7.15)] plotted as a function of ξ. Also displayed are the magnitude of the slope of $h(x)$ evaluated at the fixed point and the low $\sigma_{h,<}$ and high $\sigma_{h,>}$ magnetizations associated with two-state cycles. At the critical value $\xi = \xi_c \simeq 0.84$, the system undergoes a bifurcation: the fixed point σ_h goes unstable, as signaled by $\left|h'(\sigma_h(\xi))\right|$ crossing above 1. The fixed point still exists when $\xi > \xi_c$ (and is indicated by a dotted line in that case), but does not control the late-time dynamics. Instead, the system comes to hop between two magnetizations $\sigma_{h,<} < \sigma_h$ and $\sigma_{h,>} > \sigma_h$.

in (7.5), to a broken power law:

$$\frac{\mathrm{d}N}{\mathrm{d}\gamma} = A \begin{cases} (\gamma/\gamma_{\rm KN})^{-p_{\rm T}} & \gamma \leq \gamma_{\rm KN} \\ (\gamma/\gamma_{\rm KN})^{-p_{\rm KN}} & \gamma > \gamma_{\rm KN} \end{cases} . \tag{7.17}$$

This allows me to account for the different expected modifications to the power-law scaling of the distribution function in both the Thomson ($\gamma \leq \gamma_{\rm KN}$) and deep Klein-Nishina ($\gamma > \gamma_{\rm KN}$) regimes.

Given a steady injected particle distribution with power-law index p, the realized distribution, if cooled in the Thomson limit, is steepened to index $p_{\rm T} = p + 1$. In contrast, if one adopts the approximate form for $f_{\rm KN}(q)$ from equation (6.7), then, for $\gamma > \gamma_{\rm KN}$, the actual distribution of radiating particles, in fact, hardens, attaining a power-law scaling $p_{\rm KN} \simeq p - 0.5$ (Moderski et al., 2005). I therefore take $p_{\rm T} = p(\sigma_{\rm h}) + 1$ and $p_{\rm KN} = p(\sigma_{\rm h}) - 0.5$ where $p(\sigma_{\rm h}) = 1 + 2/\sqrt{\sigma_{\rm h}}$ as in equation (7.14). Admittedly, this is still very crude. It ignores the finite time-scale of particle energization in the reconnection layer and the finite time it takes for the full energy range of the distribution function to respond to radiative losses. It ignores, moreover, the bursty nature of magnetic reconnection at the highest energies (e.g., Werner et al., 2019; Mehlhaff et al., 2020). However, these relationships for $p_{\rm T}$ and $p_{\rm KN}$ are at least a first step toward capturing some of the qualitative effects that radiative cooling may have on the pair feedback mechanism.

Using (7.17), the power fraction becomes [cf. equation (7.6)]

$$\tilde{\mathcal{F}}(p_{\rm T}, p_{\rm KN}, z) \equiv \frac{\int_8^z \mathrm{d}x\, x^{-p_{\rm KN}+2} f_{\rm KN}(x)}{\int_0^1 \mathrm{d}x\, x^{-p_{\rm T}+2} f_{\rm KN}(x) + \int_1^z \mathrm{d}x\, x^{-p_{\rm KN}+2} f_{\rm KN}(x)} . \tag{7.18}$$

Note that $\tilde{\mathcal{F}}(p_{\rm T}, p_{\rm KN}, z \leq 8) = 0$, and so it is not necessary to consider the case $z < 1$ in the denominator. The fraction $\tilde{\mathcal{F}}$, with $p_{\rm T} = p + 1$ and $p_{\rm KN} = p - 0.5$ is displayed as a function of z and p in Fig. 7.7. The opposite impacts of radiative losses in the Thomson and deep Klein-Nishina limits – where, respectively, the particle energy distribution tends to steepen and become shallower – have a pronounced effect on the p-dependence of $\tilde{\mathcal{F}}$. Due to the steepening of the distribution at lower energies, the denominator diverges when $p > 2$ ($p_{\rm T} > 3$). Meanwhile, due to the hardening in the Klein-Nishina regime, the fraction $\tilde{\mathcal{F}}$ goes to 1 as $z \to \infty$ when $p < 1.5$ ($p_{\rm KN} < 1$). This

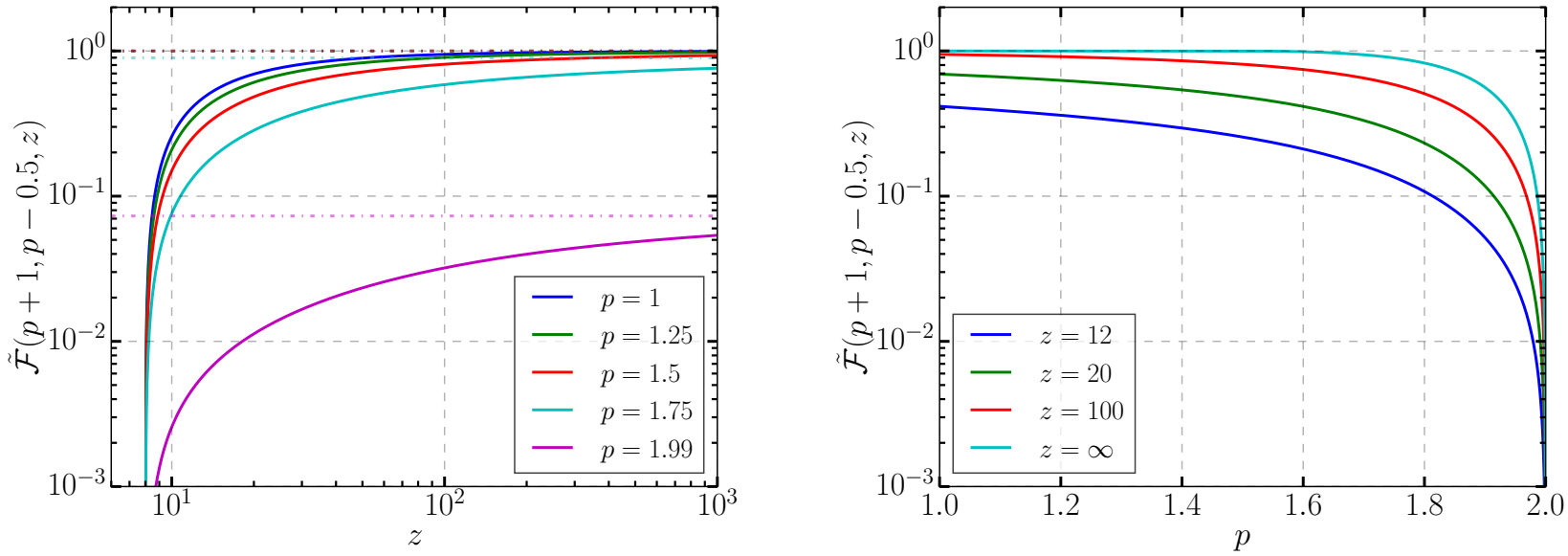

Figure 7.7: Top: A plot of $\tilde{\mathcal{F}}(p{+}1, p{-}0.5, z)$ as a function of $z \equiv \gamma_2/\gamma_{\rm KN}$ for several p. Dot-dashed lines indicate $\lim_{z\to\infty} \tilde{\mathcal{F}}(p{+}1, p{-}0.5, z)$. Bottom: A plot of $\tilde{\mathcal{F}}(p, z)$ as a function of p for several z. Due to the opposite influences of radiative cooling in the Thomson and deep Klein-Nishina regimes, where the particle energy distribution tends to be softened and hardened, respectively, the range in p over which $\tilde{\mathcal{F}}$ significantly varies compresses relative to that of $\mathcal{F}$ ($1.5 < p < 2$ for the former instead of $1.5 < p < 3$ for the latter).

compresses the effective range in p over which $\tilde{\mathcal{F}}$ varies. Whereas my original power fraction $\mathcal{F}(p,\infty)$ in equation (7.6) starts to decline from 1 only once $p \gtrsim 1.5$, not reaching zero until $p = 3$, the new fraction $\tilde{\mathcal{F}}(p+1, p-0.5, \infty)$ departs from 1 when $p > 1.5$ but hits zero already by the time $p = 2$.

This dramatically impacts the fixed points $\sigma_{\rm h}(\xi)$ and their stabilities. In analogy to equation (7.15), I define

$$\tilde{h}(x) = \frac{\sigma_{\rm h,0}}{1 + 2\xi\,\tilde{\mathcal{F}}[p(x)+1, p(x)-0.5, \infty]\sigma_{\rm h,0}/3}\,. \tag{7.19}$$

Fig. 7.8 displays a few $\tilde{h}(x)$ curves, each one with different energy recapture efficiency ξ. As promised, $\tilde{h}(x)$ is flat when $p(x) = 1+2/\sqrt{x} > 1.5$ (i.e., when $x > 16$) because it is here that $\tilde{\mathcal{F}} = 1$. At the same time, $\tilde{h}(x < 4) = \sigma_{\rm h,0}$ since $p(x < 4) > 2$, and, hence, $\tilde{\mathcal{F}} = 0$. The rapid transition in $\tilde{h}(x)$ from nearly vertical to nearly flat between $x = 4$ and $x = 16$ induces a sharp transition in the stability of the fixed point $\sigma_{\rm h}$. The fixed point becomes unstable at $\xi = \xi_{\rm c} \simeq 0.30$, and further increasing ξ beyond this point yields much more dramatic swing cycles than did incremental changes in ξ beyond the threshold $\xi_{\rm c} \simeq 0.84$ of h in equation (7.15). By the time $\xi \gtrsim 0.5$, the reconnection layer *starts* in a swing cycle, hopping between its *initial* magnetization $\sigma_{\rm h,0} = \sigma_{\rm h,>}$ and a much lower magnetization $\sigma_{\rm h,<} < 4$. In the low-magnetization state, pair-production completely ceases ($\tilde{\mathcal{F}} = 0$) and, after one readvection time, the plasma flowing into the layer contains no newborn pair component and once again possesses the initial magnetization $\sigma_{\rm h,0}$. Although this is still a long way from a detailed model, it illustrates what appears to be a robust mechanism for pair feedback in Klein-Nishina relativistic reconnection. An initially highly-magnetized system $\sigma_{\rm h,0} \gg 1$ can efficiently accelerate gamma-ray-radiating leptons in the reconnection layer. These gamma-rays collide with ambient background photons to produce a hot newborn pair component in the upstream plasma. Subsequently, the new pairs are advected into the layer where they suppress NTPA and, hence, the production of additional pairs. This mechanism operates even when radiative cooling of the layer particles is taken into account. If the energy recapture efficiency ξ approaches order unity, then the system undergoes undamped, large-amplitude oscillations between a high magnetization (even as high as the initial value $\sigma_{\rm h,0}$) – sourcing copious upstream pairs – and a low magnetization

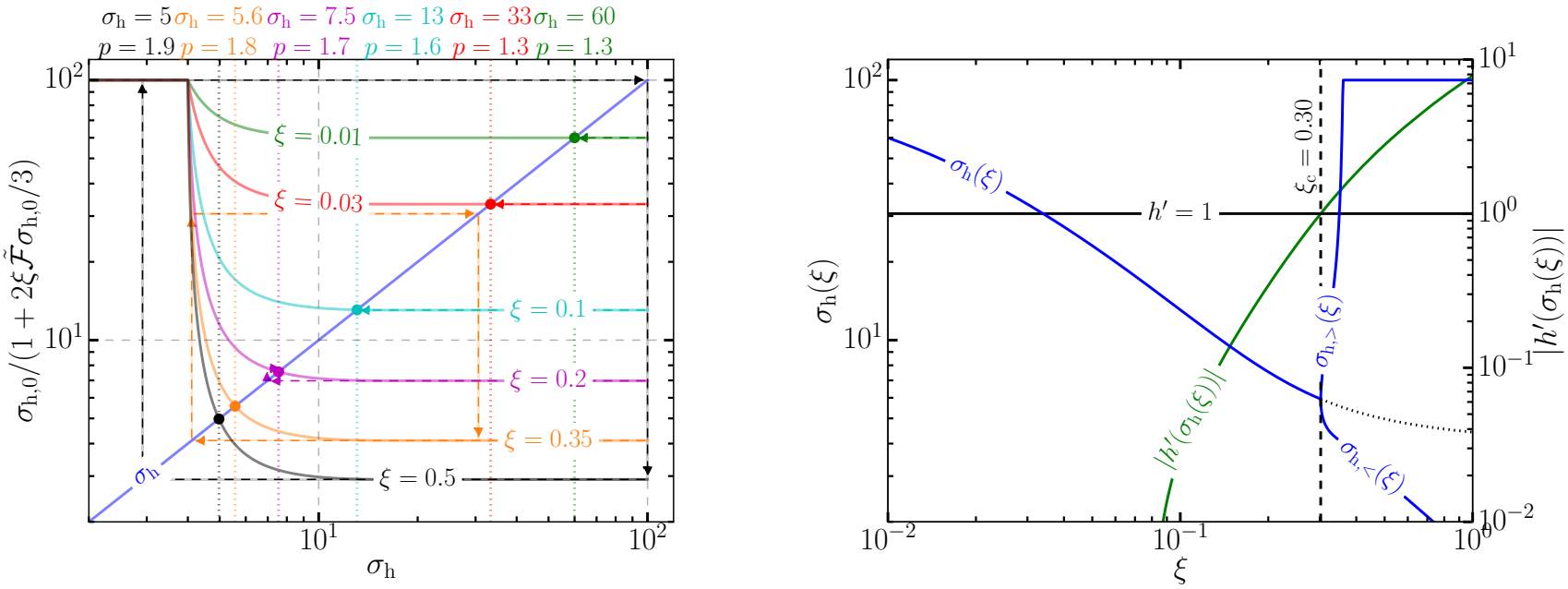

Figure 7.8: Top: The analogue of Fig. 7.5 for $\tilde{\mathcal{F}}$ [as defined in equation (7.18)] instead of $\mathcal{F}$ [as defined in equation (7.6)]. Bottom: The analogue of Fig. 7.6 for $\tilde{\mathcal{F}}$. The tendency for radiative losses to steepen the particle distribution in the Thomson regime and, in contrast, to harden it in the deep Klein-Nishina limit creates a sharp and dramatic bifurcation. The critical value of the energy recapture efficiency $\xi_c \simeq 0.30$ is smaller than $\xi_c = 0.84$ obtained in Fig. 7.6, indicating that radiation back reaction makes the system more 'touchy': more susceptible to late-time limit cycles. Once ξ exceeds this threshold, the system rapidly transitions to a two-state cycle featuring very large-amplitude swings between the initial magnetization $\sigma_{h,>} = \sigma_{h,0}$ and a small magnetization $\sigma_{h,<} \sim 1$.

where pair-production is completely shut down. For $\xi \ll 1$ however, the system approaches a steady state – characterized by a balance between upstream pair-loading and particle energization in the layer – after just one or a few readvection times $t_{\rm ra} = \lambda_{\rm mfp}/\beta_{\rm rec}c$.

7.2.2 Upstream pair cascades are not generally expected

So far, I have described a fixed point solution in which the reconnection layer regulates itself, maintaining a universal effective hot magnetization $\sigma_{\rm h}$, or, potentially, exhibiting large-amplitude two-state limit cycles. The arguments that allowed me to characterize this behavior were based solely upon tracing the flow of energy through the system. I now undertake an analogous program, tracing the plasma particles instead of the energy.

Here, I must deal with an additional complication that was not present in the preceding energy-based arguments. Namely, because of pair production, particle number is not conserved, and one cannot simply equate the number of pairs born into the inflow plasma with (twice) the number of above-threshold photons radiated away from the layer. In principle, additional pair creation in the upstream region can occur: newborn pairs can radiate additional photons that are themselves above pair-production threshold and capable of producing secondary pair generations. I address this issue in this section. As an immediate byproduct of the analysis, I show that a pair cascade is not generally expected – the number of particles only grows exponentially in each subsequent pair generation under certain optimal conditions.

An important quantity here is the distribution of pairs injected into the inflow plasma as the result of high-energy photons, originally emitted from the reconnection layer, getting absorbed in the upstream region. I denote this as

$$Q^{(1)}_{\gamma\gamma}(\gamma) = B_1 \begin{cases} \gamma^{-\Gamma_1} & \gamma^{(1)}_1 \le \gamma < \gamma^{(1)}_2 \\ 0 & \text{otherwise} \end{cases} . \tag{7.20}$$

Here, the superscript '(1)' (subscript '1' on Γ_1 and B_1) indicates the first generation. In this section,

I calculate

$$Q_{\gamma\gamma}^{(n)}(\gamma) = B_n \begin{cases} \gamma^{-\Gamma_n} & \gamma_1^{(n)} \leq \gamma < \gamma_2^{(n)} \\ 0 & \text{otherwise} \end{cases} \tag{7.21}$$

the injected distribution of nth-generation pairs, in terms of $Q_{\gamma\gamma}^{(1)}(\gamma)$. The dimensions of $Q_{\gamma\gamma}^{(n)}(\gamma)$ are that of a rate: particles per unit time per unit energy. I have not normalized by spatial volume. Thus, $Q_{\gamma\gamma}^{(n)}(\gamma)$ is averaged over the spatial region where nth-generation pair-production is active. For now, I assume that $Q_{\gamma\gamma}^{(n)}(\gamma)$ is a power law and check this *a posteriori*.

7.2.2.1 Newborn pair generations: basic observations

Before calculating the $Q_{\gamma\gamma}^{(n)}(\gamma)$ distributions in detail, I offer a few guiding remarks. These primarily pertain to the range of energies present in each generation – i.e., to $\gamma_1^{(n)}$ and $\gamma_2^{(n)}$ – and to the number of generations that the upstream region can support.

As discussed in section 6.1, a particle with Lorentz factor $\gamma \geq \gamma_{\rm pp} = 8\gamma_{\rm KN}$ emits photons of characteristic energy $E_\gamma \sim \gamma m_e c^2/2$. Upon absorption by the soft radiation background, these photons each create two particles with approximate Lorentz factors $E_\gamma/2m_e c^2 \sim \gamma/4$. Thus, the power law (7.20) begins and ends at energies about a factor of 4 less than the corresponding energies of layer particles: $\gamma_1^{(1)} \sim \gamma_{\rm pp}/4 = 2\gamma_{\rm KN}$ and $\gamma_2^{(1)} \sim \gamma_2/4$, [$\gamma_2$ is the cutoff layer particle energy; see equation (7.5)].

In general, the low and high injected energies in each successive pair generation follow by identical reasoning to those of the first generation. One has $\gamma_1^{(n)} \sim \gamma_1^{(1)} \sim \gamma_{\rm pp}/4 = 2\gamma_{\rm KN}$ and $\gamma_2^{(n)} \sim \gamma_2/4^n$. The last generation, N, possesses cutoff at or below the threshold to emit pair-producing radiation: $\gamma_{\rm pp} = 8\gamma_{\rm KN} \geq \gamma_2^{(N)} = \gamma_2/4^N$. This caps the number of generations to $N = \lceil \log_4(\gamma_2/8\gamma_{\rm KN}) \rceil$, where $\lceil x \rceil$ rounds x up to the nearest integer. Note that, if $\log_4(\gamma_2/8\gamma_{\rm KN})$ is not an integer, then a partial additional generation is produced from the subset of pairs in the preceding generation with Lorentz factors above $\gamma_{\rm pp}$. For example, if $\log_4(\gamma_2/8\gamma_{\rm KN})$ is 2.4 – and, hence, $N = 3$ – then two full generations are created in addition to a partial third generation.

One may also place a potentially firmer limit on the number of generations. Note that newborn upstream particles with $t_{\rm cool,IC}(\gamma) > t_{\rm ra}$ do not radiate before entering the layer, and, thus, do not yield additional upstream pairs. This condition is marginally satisfied if $\gamma = \tilde{\gamma}_{\rm cool}$ such that

$$t_{\rm cool,IC}(\tilde{\gamma}_{\rm cool}) \equiv \frac{10}{\tau_{\gamma\gamma}}\frac{L}{c} \simeq \frac{\lambda_{\rm mfp}}{\beta_{\rm rec}c}\,, \tag{7.22}$$

where, for convenience, I use the approximate form $t_{\rm ra} \simeq 10L/\tau_{\gamma\gamma}c$ [equation (7.9)]. Note that the cooling time $t_{\rm cool,IC}$ is the product of the readvection time with a dimensionless function of $q = \gamma/\gamma_{\rm KN}$: namely, $t_{\rm cool,IC}(\gamma) = (3/50)(10L/\tau_{\gamma\gamma}c)[qf_{\rm KN}(q)]^{-1}$, which follows from equations (6.21) and (7.9). This implies that the solution to (7.22) depends only on $\gamma_{\rm KN}$ – it is independent of all other system parameters (e.g., $\gamma_{\rm max}$, $\gamma_{\rm rad,T}$, and $\sigma_{\rm c,0}$). Moreover, because $t_{\rm cool,IC}(\gamma)$ is non-monotonic [see equation (6.21), Fig. 6.4, and the surrounding discussion], there are actually two solutions to (7.22). These can be found numerically and are

$$\tilde{\gamma}_{\rm cool,1} \simeq 0.066\gamma_{\rm KN} \simeq (3/50)\gamma_{\rm KN}\,, \tag{7.23}$$

which is in the Thomson limit, and

$$\tilde{\gamma}_{\rm cool,2} \simeq 280\gamma_{\rm KN}\,, \tag{7.24}$$

which is in the deep Klein-Nishina regime.

All pairs born into the upstream region with $\gamma > \tilde{\gamma}_{\rm cool,2}$ have $t_{\rm cool,IC}(\gamma) > t_{\rm ra}$ and thus do not radiate before being swept into the reconnection layer. However, pairs born with $\gamma < \tilde{\gamma}_{\rm cool,2}$ (which also have $\gamma > \tilde{\gamma}_{\rm cool,1}$ because $\gamma \geq \gamma_1^{(n)} \sim 2\gamma_{\rm KN} > \tilde{\gamma}_{\rm cool,1}$) have $t_{\rm cool,IC}(\gamma) < t_{\rm ra}$. They maintain this condition as they radiatively cool all the way until they reach $\gamma = \tilde{\gamma}_{\rm cool,1}$, at which point $t_{\rm cool,IC}(\gamma) = t_{\rm ra}$. Thus, pairs born between the energies $\tilde{\gamma}_{\rm cool,1}$ and $\tilde{\gamma}_{\rm cool,2}$ may cool all the way down to $\tilde{\gamma}_{\rm cool,1}$ before entering the layer. (They may not cool quite this far if not born a full mean-free-path $\lambda_{\rm mfp}$ upstream of the layer.)

Because first-generation pairs whose parent layer particles had Lorentz factors $\geq 4\tilde{\gamma}_{\rm cool,2}$ do not spawn additional pairs, the number of generations is limited to $N = \lceil \log_4(4\tilde{\gamma}_{\rm cool,2}/8\gamma_{\rm KN}) \rceil \simeq$

$\lceil \log_4(282/2) \rceil = 4$. Combining this with the previous limit yields

$$N = \min\left[\left\lceil \log_4\left(\frac{\gamma_2}{8\gamma_{\rm KN}}\right)\right\rceil, 4\right], \tag{7.25}$$

which slightly modifies the cutoff energies from $\gamma_2^{(n)} \sim \gamma_2/4^n$ to

$$\gamma_2^{(n)} \sim \begin{cases} \gamma_2/4 & n = 1 \\ \min(\gamma_2, 4\tilde{\gamma}_{\rm cool,2})/4^n & n > 1 \end{cases}. \tag{7.26}$$

While the first generation's cutoff is determined entirely in terms of the layer cutoff γ_2, only particles in the first generation with $\gamma < \tilde{\gamma}_{\rm cool,2}$ can give rise to further generations before entering the layer.

The maximum number of generations N places an important constraint on the self-regulated reconnection system. The mean excursion of nth-generation photons (those that create nth-generation pairs) from the reconnection layer is $\left|\bar{y}^{(n)}\right| \sim \sqrt{n}\lambda_{\rm mfp}$, which scales with n as a 1D random walk. Thus, photons would begin to escape the system if $N \geq (L/\lambda_{\rm mfp})^2 \gg 1$. But, because N cannot exceed 4, this is not expected.

7.2.2.2 Newborn pair generations: detailed calculation

I now calculate, in detail, the $Q_{\gamma\gamma}^{(n)}(\gamma)$ distributions in terms of $Q_{\gamma\gamma}^{(1)}(\gamma)$. To do this, I set up a system of coupled differential equations for $N_{\gamma\gamma}^{(n)}(\gamma)$, the (volume-integrated) distributions of pairs born into the inflow plasma, and for $N_{\rm ph}^{(n)}(\epsilon)$, the (volume-integrated) distributions of photons residing in the upstream region. To formulate these equations, I temporarily add time-dependence to the distributions, though I ultimately specialize to the steady state. The time-derivatives of the $N_{\gamma\gamma}^{(n)}$'s can be expressed as

$$\begin{aligned} \frac{\partial}{\partial t} N_{\gamma\gamma}^{(n)}(\gamma, t) &= (\text{injection}) - (\text{cooling}) \\ &= Q_{\gamma\gamma}^{(n)}(\gamma, t) - \frac{\partial}{\partial \gamma}\left(\dot{\gamma} N_{\gamma\gamma}^{(n)}(\gamma, t)\right). \end{aligned} \tag{7.27}$$

Here, I have fictitiously assumed that the discrete Klein-Nishina-limit cooling of particles can be represented continuously. However, this approximation has proven to be quite accurate for non-

mono-energetic pair distributions $N_{\gamma\gamma}^{(n)}$ (Zdziarski, 1989; Moderski et al., 2005). The signed cooling rate is given by $-\dot{\gamma} = |\dot{\gamma}| = \gamma/t_{\rm cool,IC}(\gamma)$.

Simultaneously, the time-derivatives of the $N_{\rm ph}^{(n)}$'s are

$$\begin{aligned}\frac{\partial}{\partial t}N_{\rm ph}^{(n)}(\epsilon,t) &= (\text{emission}) - (\text{annihilation}) \\ &= 2\frac{N_{\gamma\gamma}^{(n-1)}(2\epsilon,t)}{t_{\rm cool,IC}(2\epsilon)} - \frac{N_{\rm ph}^{(n)}(\epsilon,t)}{\lambda_{\rm mfp}/c}\,,\end{aligned} \tag{7.28}$$

where, in this section only, I write photon energies ϵ in units of $m_e c^2$. The leading factor 2 on the first term comes from the Jacobian $\mathrm{d}\gamma/\,\mathrm{d}\epsilon = 2$ corresponding to $\gamma = 2\epsilon$. Equation (7.28) is written explicitly in the Klein-Nishina regime, where each upstream pair cools in time $t_{\rm cool,IC}(\gamma)$, and, when it does, emits a photon of energy $\gamma m_e c^2/2$. Then, each photon travels a distance $\lambda_{\rm mfp}$ in time $\lambda_{\rm mfp}/c$ before annihilating against the background to produce a new pair. Because 2 particles are injected into the nth generation upon each annihilation of an nth-generation photon, one has

$$Q_{\gamma\gamma}^{(n)}(\gamma,t) = 2\times 2\frac{N_{\rm ph}^{(n)}(2\gamma,t)}{\lambda_{\rm mfp}/c}\,. \tag{7.29}$$

Here again, one factor of 2 comes from the Jacobian from particle to photon energies.

Equations (7.27)-(7.29) achieve a steady state if the readvection time, over which the $n = 1$ injection term is constant, is longer than the cooling times of newborn pairs and the photon streaming time $\lambda_{\rm mfp}/c$. The first condition is satisfied because all pairs (at least for $n > 1$) are born with Lorentz factors $\gamma < \tilde{\gamma}_{\rm cool,2}$, while the latter is satisfied because the nth-generation readvection time is $\sqrt{n}\lambda_{\rm mfp}/\beta_{\rm rec}c \sim 10\sqrt{n}\lambda_{\rm mfp}/c \gg \lambda_{\rm mfp}/c$. (The particles in the $n = 1$ generation with $\gamma > \tilde{\gamma}_{\rm cool,2}$ do not produce additional pairs and so can be excluded from these equations.) Let me therefore specialize to the steady state of equations (7.27) and (7.28). This gives, upon plugging (7.29) into (7.28),

$$Q_{\gamma\gamma}^{(n)}(\gamma) = 8\frac{N_{\gamma\gamma}^{(n-1)}(4\gamma)}{t_{\rm cool,IC}(4\gamma)}\,, \tag{7.30}$$

which, in turn, allows one to write

$$\frac{\partial}{\partial\gamma}\left(\dot{\gamma}N_{\gamma\gamma}^{(n)}(\gamma)\right) = -\frac{1}{8}\frac{\partial}{\partial\gamma}\left(\gamma Q_{\gamma\gamma}^{(n+1)}(\gamma/4)\right)\,, \tag{7.31}$$

where I have used $-\dot{\gamma} = \gamma/t_{\rm cool,IC}(\gamma)$.

Then, inserting (7.31) into the steady state of (7.27) gives a recursive formula for the $Q_{\gamma\gamma}^{(n)}(\gamma)$'s:

$$-\frac{1}{8}\frac{\partial}{\partial\gamma}\left(\gamma Q_{\gamma\gamma}^{(n+1)}(\gamma/4)\right) = Q_{\gamma\gamma}^{(n)}(\gamma)\,. \tag{7.32}$$

Let me integrate this equation from $\gamma \geq \gamma_{\rm pp}$ to $\gamma = \infty$. The upper bound gives zero on the left-hand-side of (7.32) since $Q_{\gamma\gamma}^{(n)}(\infty) = 0$. The lower bound does *not* give zero on the left because $\gamma/4$ exceeds the lowest energy $\gamma_1^{(n)} = 2\gamma_{\rm KN}$ in the $(n+1)$st injected distribution. Hence,

$$Q_{\gamma\gamma}^{(n+1)}(\gamma) = \frac{2}{\gamma}\int_{4\gamma}^{\infty} Q_{\gamma\gamma}^{(n)}(\gamma')\,{\rm d}\gamma'\,. \tag{7.33}$$

Let me now use (7.33) to determine $Q_{\gamma\gamma}^{(n+1)}(\gamma)$ explicitly, assuming $Q_{\gamma\gamma}^{(n)}(\gamma)$ is given by a power law as in equation (7.21). Restricting to $2\gamma_{\rm KN} = \gamma_1^{(n+1)} \leq \gamma < \gamma_2^{(n+1)} = \gamma_2^{(n)}/4$, one has

$$\begin{aligned} Q_{\gamma\gamma}^{(n+1)}(\gamma) &= \frac{2B_n}{\gamma}\int_{4\gamma}^{\gamma_2^{(n)}} (\gamma')^{-\Gamma_n}\,{\rm d}\gamma' \\ &= \frac{2B_n}{\gamma(\Gamma_n - 1)}(4\gamma)^{-\Gamma_n+1}\left[1 - \left(\frac{4\gamma}{\gamma_2^{(n)}}\right)^{\Gamma_n - 1}\right]\,. \end{aligned} \tag{7.34}$$

For γ not between $2\gamma_{\rm KN}$ and $\gamma_2^{(n)}/4$, the distribution $Q_{\gamma\gamma}^{(n)}(\gamma)$ is zero. Strictly speaking, for $n = 1$, $\gamma_2^{(1)} = \gamma_2/4$ needs to be modified to $\min(\gamma_2^{(1)}, \tilde{\gamma}_{\rm cool,2})/4$, since only first-generation particles with energies less than $\tilde{\gamma}_{\rm cool,2}$ can spawn additional pairs.

Noting that the term in square brackets in (7.34) is roughly unity except when $\gamma \sim \gamma_2^{(n)}/4$, one sees that $Q_{\gamma\gamma}^{(n+1)}(\gamma)$ is just a power law with a cutoff at $\gamma_2^{(n+1)} = \gamma_2^{(n)}/4$, in agreement with equation (7.26). This verifies the assumed power-law form (7.21) provided the initial injected distribution $Q_{\gamma\gamma}^{(1)}(\gamma)$ is also a power-law. In that case,

$$\Gamma_{n+1} = \Gamma_n = \Gamma_1 \equiv \Gamma \tag{7.35}$$

and

$$B_{n+1} = \mathcal{A}(\Gamma)B_n = \left[\mathcal{A}(\Gamma)\right]^n B_1\,, \tag{7.36}$$

where I have defined

$$\mathcal{A}(\Gamma) \equiv \frac{8}{4^{\Gamma}(\Gamma-1)}\,. \tag{7.37}$$

These recurrence formulae imply that the total pair injection rate in the upstream region is

$$\begin{aligned}\sum_{n=1}^{N}\int Q_{\gamma\gamma}^{(n)}(\gamma)\,\mathrm{d}\gamma &= B_1\sum_{n=1}^{N}\left[\mathcal{A}(\Gamma)\right]^{n-1}\frac{\gamma_{\rm min}^{-\Gamma+1}}{\Gamma-1}\left[1-\left(\frac{\gamma_{\rm min}}{\gamma_2^{(n)}}\right)^{\Gamma-1}\right]\\ &\simeq \frac{B_1\gamma_{\rm min}^{-\Gamma+1}}{\Gamma-1}\sum_{n=0}^{N-1}\left[\mathcal{A}(\Gamma)\right]^{n}\\ &\simeq \left[\int Q_{\gamma\gamma}^{(1)}(\gamma)\,\mathrm{d}\gamma\right]\sum_{n=0}^{N-1}\left[\mathcal{A}(\Gamma)\right]^{n}\\ &= \frac{1-\left[\mathcal{A}(\Gamma)\right]^{N}}{1-\mathcal{A}(\Gamma)}\int Q_{\gamma\gamma}^{(1)}(\gamma)\,\mathrm{d}\gamma\,,\end{aligned} \tag{7.38}$$

where, for convenience, I define $\gamma_{\rm min} \equiv \gamma_1^{(n)} = 2\gamma_{\rm KN}$. In the second and third lines, I assume that $\Gamma > 1$ and, hence, that the term $(\gamma_{\rm min}/\gamma_2^{(n)})^{\Gamma-1}$ can be neglected. The final line is the crux of this section. It determines under what conditions a true pair cascade develops – when the total number of injected pairs in the upstream region is exponential in the number of generations N. One has,

$$\begin{aligned}&\mathcal{A}(\Gamma) > 1 \Longleftrightarrow \text{pair cascade}\\ &\mathcal{A}(\Gamma) < 1 \Longleftrightarrow \text{no pair cascade}\,.\end{aligned} \tag{7.39}$$

The multiplication factor $\mathcal{A}(\Gamma)$ is less than unity for $\Gamma > 1.73$. Whether a pair cascade develops comes down to the expected value of Γ in the first-generation injected distribution.

One expects that $Q_{\gamma\gamma}^{(1)}(\gamma)$ inherits its power-law scaling from the distribution of photons emitted into the upstream region from the layer. For a power-law distribution (7.17) of radiating layer particles, the emitted photon spectrum is, approximately, a power law with index $\Gamma = p_{\rm KN}+1$ (plus a logarithmic correction; see Blumenthal & Gould 1970 and Aharonian & Atoyan 1981). Since one expects the scaling of the layer particle distribution in the Klein-Nishina regime to be $p_{\rm KN} = p - 0.5$, which corresponds to an intrinsic particle acceleration index p hardened due to

Klein-Nishina IC losses, one has $\min(\Gamma) = \min(p + 0.5)$. Now, according to (7.14), $\min(p) = 1$, and thus the distribution of first-generation injected pairs should be no harder than $\min(\Gamma) = 1.5$. This gives a maximum multiplication factor of $\max(\mathcal{A}(\Gamma)) = \mathcal{A}(\min(\Gamma)) = \mathcal{A}(1.5) = 2$, for which the number of injected pairs doubles in each successive generation.

A shallow enough scaling Γ to bring $\mathcal{A}(\Gamma)$ above unity is only achieved for $p < 1.7 - 0.5 = 1.2$. Using equation (7.14), this implies a magnetization $\sigma_{\rm h} > 80$. Also, the theory I have presented thus far ignores the possibility of a guide field, which tends to suppress reconnection-powered NTPA (e.g., Werner & Uzdensky, 2017). Thus, a pair cascade is possible, but requires quite optimal combinations of parameters (high $\sigma_{\rm h}$ and small guide field, for example).

7.2.3 The small number density of upstream pairs

In this section, I employ my tally of the upstream pair creation rate – given by the $Q^{(n)}_{\gamma\gamma}(\gamma)$ distributions – to answer another important question associated with Klein-Nishina radiative reconnection. Namely, I calculate the pair-production *multiplicity*

$$\eta \equiv \frac{n_{\gamma\gamma}}{n_0}\,, \tag{7.40}$$

the ratio of the number density $n_{\gamma\gamma}$ of newborn pairs entering the reconnection layer to that n_0 of pairs in the far upstream region.

Let me define the volumetric (per unit time per unit volume) rate of pair production $\mathrm{d}n^{(n)}_{\gamma\gamma}/\,\mathrm{d}t$ into the nth generation. Assuming that nth-generation pairs are deposited uniformly up to a transverse distance $\bar{y}^{(n)}$ (e.g., $\bar{y}^{(n)} \sim \sqrt{n}\lambda_{\rm mfp}$) away from the layer across the full width L of the system yields

$$\frac{\mathrm{d}n^{(n)}_{\gamma\gamma}}{\mathrm{d}t} \sim \frac{1}{2\bar{y}^{(n)}L}\int \mathrm{d}\gamma\, Q^{(n)}_{\gamma\gamma}(\gamma)\,. \tag{7.41}$$

Noting that a parcel of plasma travels from $|y| \sim \bar{y}^{(n)}$ to the layer at $|y| \sim \Delta \ll \bar{y}^{(n)}$ over time $\bar{y}^{(n)}/\beta_{\rm rec}c$, during which it accrues nth-generation pairs at the rate $\mathrm{d}n^{(n)}_{\gamma\gamma}/\,\mathrm{d}t$, gives

$$n_{\gamma\gamma} \sim f_{\rm noesc}\sum_{n=1}^{N}\frac{\bar{y}^{(n)}}{\beta_{\rm rec}c}\frac{\mathrm{d}n^{(n)}_{\gamma\gamma}}{\mathrm{d}t} \sim \frac{1}{2\beta_{\rm rec}cL}\int \mathrm{d}\gamma\sum_{n=1}^{N} Q^{(n)}_{\gamma\gamma}(\gamma)\,. \tag{7.42}$$

In the second step, I assume $f_{\rm noesc} = 1$. Similar to the cancellation of the readvection time in (7.8), $\bar{y}^{(n)}$ cancels in (7.42).

To evaluate (7.42), I normalize the $Q_{\gamma\gamma}^{(n)}(\gamma)$'s by balancing the pair creation rate in the *first* generation with twice the number of above-threshold photons emitted from the reconnection layer:

$$\int \mathrm{d}\gamma\, Q_{\gamma\gamma}^{(1)}(\gamma) = 2\Delta L' \int_{\gamma_{\rm pp}} \mathrm{d}\gamma\, R_{\rm IC}(\gamma)\, \mathrm{d}N/\,\mathrm{d}\gamma\,, \tag{7.43}$$

where $\mathrm{d}N/\,\mathrm{d}\gamma$ is the distribution of radiating particles as in equations (7.5) and (7.17). In turn, one may normalize the distribution $\mathrm{d}N/\,\mathrm{d}\gamma$ [to find A in equation (7.5) or (7.17)] by setting the power radiated from the layer to $P_{\rm Poynt}/2$ [as in equations (7.2) and (7.3)]:

$$\frac{1}{2}P_{\rm Poynt} = L\beta_{\rm rec}c\frac{B_0^2}{4\pi} \sim j_{\rm IC}\Delta L' = \Delta L' \int \mathrm{d}\gamma P_{\rm IC}(\gamma)\, \mathrm{d}N/\,\mathrm{d}\gamma\,. \tag{7.44}$$

In keeping with the earlier parts of this study, I consider two cases: one in which the radiation reaction force on layer particles is ignored and one in which it is approximately included. In the former case, I assume a single power-law form for $\mathrm{d}N/\,\mathrm{d}\gamma$, as in equation (7.5), with index $p(\sigma_{\rm h}) = 1 + 2/\sqrt{\sigma_{\rm h}}$ [equation (7.14)]. In the latter, I adopt a broken power law, as in equation (7.17), with separate indices $p_{\rm T} = p(\sigma_{\rm h}) + 1$ and $p_{\rm KN} = p(\sigma_{\rm h}) - 0.5$ in the Thomson ($\gamma < \gamma_{\rm KN}$) and Klein-Nishina ($\gamma > \gamma_{\rm KN}$) regimes, respectively. To save space, I simultaneously conduct both analyses by plugging in equation (7.17) for $\mathrm{d}N/\,\mathrm{d}\gamma$ and leaving $p_{\rm T}$ and $p_{\rm KN}$ unspecified until the end of the calculation.

Proceeding in this manner, I find that

$$n_{\gamma\gamma} \sim n_0 \frac{\sigma_{\rm c,0}}{\gamma_{\rm KN}} \tilde{\mathcal{M}}(p_{\rm T}, p_{\rm KN}, z)\,, \tag{7.45}$$

where $z = \gamma_2/\gamma_{\rm KN}$. The right-hand-side does not depend on n_0: the n_0 in the numerator cancels against that in the definition of $\sigma_{\rm c,0}$. The *multiplicity function* $\tilde{\mathcal{M}}(p_{\rm T}, p_{\rm KN}, z)$ is derived, as discussed, by evaluating (7.42) whilst enforcing (7.43) and (7.44). It reads

$$\tilde{\mathcal{M}}(p_{\rm T}, p_{\rm KN}, z) \equiv \frac{3\int_8^z \mathrm{d}x\, g_{\rm KN}(x) x^{-p_{\rm KN}}}{\int_0^1 \mathrm{d}x\, f_{\rm KN}(x) x^{-p_{\rm T}+2} + \int_1^z \mathrm{d}x\, f_{\rm KN}(x) x^{-p_{\rm KN}+2}} \times \frac{\int \mathrm{d}\gamma \sum_{n=1}^{N} Q_{\gamma\gamma}^{(n)}(\gamma)}{\int \mathrm{d}\gamma\, Q_{\gamma\gamma}^{(1)}(\gamma)}\,. \tag{7.46}$$

The factor on the far right encodes the possibility of a pair cascade and is roughly equal to $[1 - \mathcal{A}(\Gamma)^N]/[1 - \mathcal{A}(\Gamma)]$, as in equation (7.38). However, here I evaluate this factor by maintaining finite z-dependent cutoffs in each term, writing

$$\int \mathrm{d}\gamma \, Q^{(n)}_{\gamma\gamma}(\gamma) \simeq \frac{[\mathcal{A}(\Gamma)]^{n-1} B_1 \gamma_{\mathrm{min}}^{-\Gamma+1}}{\Gamma - 1} \left[1 - \left(\frac{\gamma_{\mathrm{min}}}{\gamma_2^{(n)}} \right)^{\Gamma-1} \right] . \tag{7.47}$$

I do not simplify the right-hand-side to $\mathcal{A}(\Gamma)^{n-1} B_1 \gamma_{\mathrm{min}}^{-\Gamma+1}/(\Gamma - 1)$, as done in section 7.2.2 [equation (7.38)]. This ensures that contributions from each generation turn on gradually (as they do in reality), thereby keeping the multiplicity function continuous.

To denote the case where I ignore radiative feedback on the layer particles, in which I put $p_{\mathrm{T}} = p_{\mathrm{KN}} = p$, I write

$$\mathcal{M}(p, z) \equiv \tilde{\mathcal{M}}(p, p, z) \, . \tag{7.48}$$

I display the functions $\mathcal{M}(p, z)$ and $\tilde{\mathcal{M}}(p + 1, p - 0.5, z)$ in Fig. 7.9. Similar to the case of $\mathcal{F}$ and $\tilde{\mathcal{F}}$, the separate impacts of radiative cooling in the Thomson and deep Klein-Nishina limits contract the range in p across which the multiplicity function varies. Just as for the power fraction $\mathcal{F}$, the steepening of the layer particle distribution $p_{\mathrm{T}} = p + 1$ in the Thomson regime completely shuts down pair production whenever $p_{\mathrm{T}} > 3$ (i.e., $p > 2$). However, unlike the power fraction $\mathcal{F}$, the multiplicity function exhibits a much more complicated non-monotonic dependence on the underlying parameters p and z. In addition to this, while the power fraction attains order unity for a wide range of p and z, the multiplicity function never does – it is never larger than $\simeq 0.1$ for any parameter combination, and is very often much smaller than this.

This last fact means that, even when a pair cascade truly does develop in the upstream region – and the total newborn pair count is exponential in the number of generations – the overall multiplicity may still be *small*. In fact, because the multiplicity function obtains a global maximum of order 10^{-1}, one can map out precisely the parameters for which η is guaranteed to be small: $\gamma_{\mathrm{KN}} > \sigma_{\mathrm{c},0}/10$. As I show in the following chapter, this condition is roughly satisfied for reconnection in both FSRQ jets and black hole accretion disc coronae.

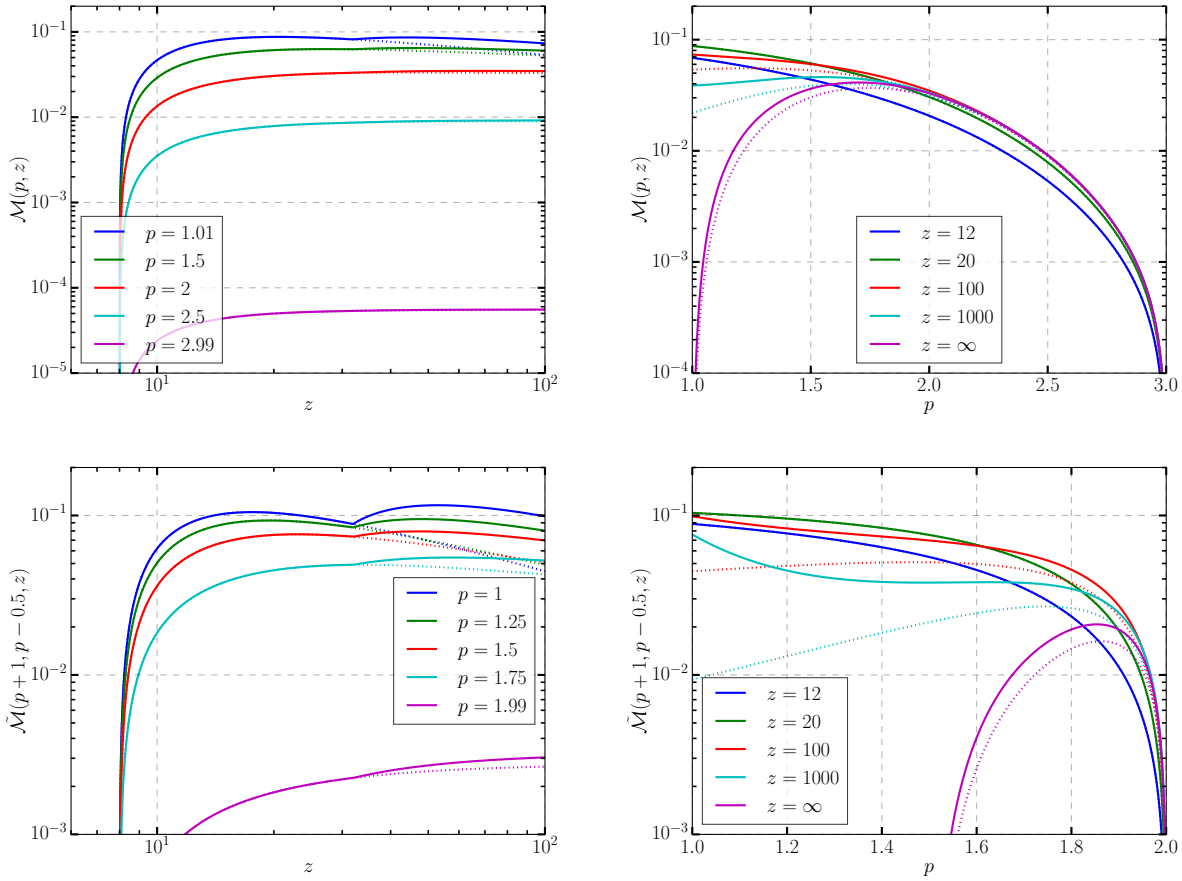

Figure 7.9: The analogue of Figs. 7.3 and 7.7, which respectively display $\mathcal{F}$ and $\tilde{\mathcal{F}}$, for $\mathcal{M}$ and $\tilde{\mathcal{M}}$. In all plots, dotted lines show the value of the multiplicity function divided by the pair cascade factor on the far right of equation (7.46), demonstrating what the multiplicity function would be if only accounting for pairs injected into the first upstream generation (i.e., ignoring a possible pair cascade). Top left: $\mathcal{M}(p,z) \equiv \tilde{\mathcal{M}}(p,p,z)$ displayed as a function of z for several p. Top right: $\mathcal{M}(p,z)$ displayed as a function of p for several z. Bottom left: $\tilde{\mathcal{M}}(p+1,p-0.5,z)$ displayed as a function of z for several p. Bottom right: $\tilde{\mathcal{M}}(p+1,p-0.5,z)$ displayed as a function of p for several z. While the dependence of $\mathcal{M}$ and $\tilde{\mathcal{M}}$ on p and z is more complicated (in particular, non-monotonic) than that of $\mathcal{F}$ and $\tilde{\mathcal{F}}$, the former are always small – no larger than 0.1 – and therefore the multiplicity η is small whenever $\gamma_{\rm KN} > \sigma_{\rm c,0}/10$. Pair cascades significantly influence the multiplicity for large z and small $p_{\rm KN}$, and, hence, have a more pronounced impact on $\tilde{\mathcal{M}}$, where $p_{\rm KN} = p - 0.5$, than on $\mathcal{M}$, where $p_{\rm KN} = p$.

This concludes my detailed discussion of pair-regulated Klein-Nishina magnetic reconnection. In the next chapter, I examine the potential astrophysical ramifications of reconnection in this regime.

Chapter 8

Astrophysical Applications of Klein-Nishina Radiative Reconnection

In this section, I discuss observational aspects of pair-regulated Klein-Nishina magnetic reconnection. I approach this in two stages. First (section 8.1.1), I elaborate observable consequences of my model from the previous chapter that are generic, not requiring an explicit global astrophysical context. This grounds my subsequent discussion, where I estimate parameters for concrete astrophysical systems – FSRQs (sections 8.2.1-8.2.4) and black hole ADCe (sections 8.2.5-8.2.6) – and comment on observations that the model may help to explain. This chapter concludes the component of this dissertation concerned with material first published by Mehlhaff et al. (2021).

8.1 General observable features

I begin by discussing the generic appearance of Klein-Nishina radiative reconnection as viewed through a telescope. I assume that the reconnection system is not spatially resolved.

8.1.1 Observed radiation comes mostly from the layer

In radiative reconnection without pair-production feedback, the radiative output is dominated by high-energy particles in the layer (i.e., the downstream region permeated by reconnected magnetic flux). However, with pair feedback, the upstream and downstream plasmas are radiatively coupled: a substantial portion of the energy emitted from the layer may be intercepted upstream and reprocessed into high-energy newborn pairs. While en route to the layer, these pairs emit potentially observable light (if $f_{\rm nocool} < 1$) that could, in principle, outshine the escaping

(below-threshold) radiation from the layer.

Therefore, I wish to determine whether the light that an observer sees comes predominantly from the upstream or downstream (layer) plasmas. I denote the respective *observable* luminosities (i.e., only of below-threshold and, hence, escaping radiation) of these two regions per unit length in the out-of-plane direction by $L_{\rm upstream}$ and $L_{\rm layer}$. The first luminosity is the fraction of the layer's above-threshold radiated power that is captured by the upstream region and reemitted below threshold before flowing back into the layer:

$$\begin{aligned} L_{\rm upstream} &\sim (1-f_{\rm nocool})(2L\lambda_{\rm mfp})\frac{{\rm d}u_{\gamma\gamma}}{{\rm d}t} \\ &\sim (1-f_{\rm nocool})\mathcal{F}\beta_{\rm rec}cL\frac{B_0^2}{4\pi} \end{aligned} \tag{8.1}$$

[cf. equation (7.7)]. Meanwhile, the fraction of the layer's power emitted below pair threshold is [cf. equation (7.3)]

$$L_{\rm layer} \sim (1-\mathcal{F})\frac{1}{2}P_{\rm Poynt} \sim (1-\mathcal{F})\beta_{\rm rec}cL\frac{B_0^2}{4\pi}\,. \tag{8.2}$$

Note that one may substitute $\tilde{\mathcal{F}}$ for $\mathcal{F}$ in these expressions provided the same symbol is used in both. The luminosity ratio is

$$\frac{L_{\rm upstream}}{L_{\rm layer}} \sim (1-f_{\rm nocool})\frac{\mathcal{F}}{1-\mathcal{F}}\,. \tag{8.3}$$

For the upstream region to outshine the layer, both significant above-threshold radiation ($\mathcal{F} > 1/2$) and relatively low energy recapture efficiency ($f_{\rm nocool} \ll 1$) are required.

However, it is also possible that the system enters a two-state cycle, swinging between a low and a high effective upstream magnetization (see section 7.2.1). Then, the upstream domain may appear brightest in one magnetization state while the downstream region shines the most in the other state. In this case, the relevant luminosities to compare are probably the brightest luminosities achieved by each region (even if in opposite states). This is certainly appropriate if the readvection time $t_{\rm ra}$ is not resolved by the observations. However, even if the readvection time is resolved, different locations along the layer's surface (e.g., different areas in the horizontal

direction of Fig. 7.2) may be decorrelated from each other. One radiating zone may be in a high-magnetization phase while a neighboring zone is in the corresponding low-magnetization state. Then, the observed luminosities from each zone are averaged and (assuming, for simplicity, a 50 per cent duty cycle) dominated by the brighter state of the more luminous region (upstream or downstream) in that zone.

Thus, the luminosity ratio in equation (8.3) may not accurately describe a reconnection layer prone to swing cycles. In fact, according to equation (8.2), the layer is observationally brightest in its *low* magnetization state – this is when $\mathcal{F}$ is smallest and, hence, when most of the incident Poynting flux is radiated at low enough energies (below pair threshold) to escape the system. In contrast, the upstream region becomes brightest when it receives an abundant supply of above-threshold photons – when the layer is in its *high* magnetization phase. These are processed into an energetic fresh pair plasma component that subsequently radiates below threshold [unless $f_{\rm nocool}(\sigma_{\rm h,>}) \simeq 1$]. Thus one expects that, in a two-state swing cycle, the appropriate luminosity ratio is

$$\frac{L_{\rm upstream}(\sigma_{\rm h,>})}{L_{\rm layer}(\sigma_{\rm h,<})} \sim [1 - f_{\rm nocool}(\sigma_{\rm h,>})]\frac{\mathcal{F}(\sigma_{\rm h,>})}{1-\mathcal{F}(\sigma_{\rm h,<})} \sim [1 - f_{\rm nocool}(\sigma_{\rm h,>})] \leq 1\,, \tag{8.4}$$

where, in the last step, I loosely approximated $\mathcal{F}(\sigma_{\rm h,>}) \sim (1-\mathcal{F}(\sigma_{\rm h,<})) \sim 1$, which is often roughly correct in swing cycles (see section 7.2.1). In the above, I also consider $f_{\rm nocool}$ to depend on the magnetization. The detailed dependence (discussed in Appendix B) is not important here beyond that, generally, $f_{\rm nocool}(\sigma_{\rm h,<}) \ll 1$ and $f_{\rm nocool}(\sigma_{\rm h,>}) \sim 1$. Thus, in a swing cycle, the layer probably dominates the observed radiation.

An exceptional case occurs if the swing cycle high-state has energy recapture efficiency $\xi(\sigma_{\rm h,>}) \simeq f_{\rm nocool}(\sigma_{\rm h,>})$ close enough to unity to render the upstream region brightest in the low-magnetization state {i.e., if $[1 - f_{\rm nocool}(\sigma_{\rm h,>})]\mathcal{F}(\sigma_{\rm h,>}) < [1 - f_{\rm nocool}(\sigma_{\rm h,<})]\mathcal{F}(\sigma_{\rm h,<})$}. Then the relevant luminosity ratio, instead of equation (8.4), is $L_{\rm upstream}(\sigma_{\rm h,<})/L_{\rm layer}(\sigma_{\rm h,<})$, which evaluates to $[1-f_{\rm nocool}(\sigma_{\rm h,<})]\mathcal{F}(\sigma_{\rm h,<})/[1-\mathcal{F}(\sigma_{\rm h,<})] \sim \mathcal{F}(\sigma_{\rm h,<}) \ll 1$, and is quite dominated by the layer.

Thus, the layer generally produces most of the observable radiation from pair-regulated Klein-

Nishina reconnection. The upstream luminosity only dominates in a steady state if $\mathcal{F} > 1/2$ and $f_{\rm nocool} \ll 1$, and it may at most be comparable to the layer luminosity in an asymptotic swing cycle. This seems to be corroborated by observations because the expected spectrum produced from the upstream region has spectral index $\alpha = 1/2$ (calculated in Appendices B and E), but the objects I discuss below (in section 8.2) have steeper scalings (as the layer might produce in a low magnetization state).

8.1.2 Klein-Nishina physics may promote rapid variability through kinetic beaming

In the first part of this dissertation (chapters 3 and 4; cf. Mehlhaff et al. 2020), I performed a detailed study of the interplay between radiative physics (Thomson IC radiation reaction) in reconnection and the kinetic beaming phenomenon first discovered by Cerutti et al. (2012b). As a reminder, in collisionless relativistic reconnection, the electromagnetic fields near reconnection X-points tend to simultaneously accelerate *and* collimate particles. The higher energy particles are focused more tightly than the lower energy particles, making this beaming inherently kinetic. A collimated beam of high-energy particles may then sweep across an observer's line of sight, and the synchrotron or IC emission of the bunch – also emitted as a beam – may then create a dramatic blip in the measured lightcurve: a 'lighthouse effect'.

In chapter 3, I showed that efficient radiative losses play a critical role in enabling the kinetic beaming mechanism to impact observations. Without strong radiative cooling, collimated beams of particles isotropize before dumping their reconnection-acquired energy into energetic photons, and thus most of their radiation is emitted quasi-isotropically. Only efficiently cooled beamed particle bunches can radiate their energy before dispersing, leading to sweeping beams of light that may manifest as rapid flares.

I here recall two specific criteria from the earlier chapters that may be necessary for, or at least promote, observable signatures of kinetic beaming:

(1) Particles may need to have cooling times within a certain multiple of their gyroperiods in the reconnecting magnetic field B_0. (If this multiple is 1, then this is the *saturated cooling*

condition from section 2.2.2 – particles are at their radiation-limited energy $\gamma_{\rm rad,IC}$.) The number of gyroperiods before a collimated bunch of particles isotropizes may be large but still finite. In fact, one main result from Mehlhaff et al. (2020) is that kinetic beaming only manifests in the emission from particles with energies within about a decade of the radiative saturation Lorentz factor $\gamma_{\rm rad,T}$, corresponding to $t_{\rm cool,T} \lesssim 100$ gyroperiods.

(2) Direct acceleration (by the reconnection electric field) near reconnection X-points must deliver particles to higher energies than secondary acceleration channels (in the language of chapters 2 and 6, $\gamma_{\rm sec} \ll \gamma_{\rm rad,IC}, \gamma_{\rm X}$ – secondary channels must radiatively stall before both the X-point radiative and intrinsic acceleration limits). Otherwise, secondary – and presumably more isotropic – energization processes may wash out signatures of kinetic beaming in the reconnection-energized distribution of particles.

There are at least 3 reasons why pair-regulated Klein-Nishina reconnection could meet these criteria, perhaps leading to even more pronounced kinetic beaming than when particles are subject to purely Thomson radiative cooling.

(1) The first reason pertains to point 1 above. Namely, Klein-Nishina effects render a particle's *cooling ratio* – the ratio of its cooling length to its Larmor radius – very insensitive, when $\gamma \gg \gamma_{\rm KN}$, to its Lorentz factor. To wit,

$$\begin{aligned}\lim_{\gamma\gg\gamma_{\rm KN}} \frac{ct_{\rm cool,IC}(\gamma)}{2\pi\gamma\rho_0} &= \lim_{\gamma\gg\gamma_{\rm KN}} \left(\frac{\gamma_{\rm rad,T}}{\gamma}\right)^2 \frac{1}{f_{\rm KN}(\gamma/\gamma_{\rm KN})} \\ &= \frac{10}{2\pi}\frac{2}{9}\frac{(\gamma_{\rm rad,T}/\gamma_{\rm KN})^2}{\ln\left(\gamma/\gamma_{\rm KN}\right) - 11/6} \simeq 0.4\frac{(\gamma_{\rm rad,T}/\gamma_{\rm KN})^2}{\ln\left(\gamma/\gamma_{\rm KN}\right) - 11/6}\,,\end{aligned} \tag{8.5}$$

where I used approximation (6.6). Even when Lorentz factors as high as $\gamma_{\rm rad,IC}$, for which $ct_{\rm cool,IC}/2\pi\gamma\rho_0 = 10/2\pi \sim 1$, are not accessible, particles may still reach a high-energy ($\gamma \gg \gamma_{\rm KN}$) regime where their cooling times come within some moderately large, beaming-favoring multiple of their gyroperiods, and where their cooling ratios become essentially γ-independent. In such a scenario, a broad range (e.g., from $\gamma_{\rm KN}$ to $\gamma_{\rm X}$) of energetic particles may radiate efficiently enough to be kinetically beamed. This range could possibly be broader than in the Thomson regime, for

which $t_{\rm cool,T}(\gamma)$ declines as γ^{-1}, making it increasingly difficult to accelerate particles up to higher and higher energies. However, one should invoke this argument with some caution because the smallest accessible cooling ratio in the Klein-Nishina regime may be quite large, scaling, according to (8.5), as $(\gamma_{\rm rad,T}/\gamma_{\rm KN})^2$.

(2) A second way Klein-Nishina effects may promote kinetic beaming pertains to item 2 above. As argued in section 6.2, Klein-Nishina radiative cooling selectively suppresses secondary, slower acceleration channels relative to rapid, impulsive acceleration near reconnection X-points. In the Klein-Nishina regime, the IC cooling time $t_{\rm cool,IC}(\gamma)$ grows as $\gamma/\ln(\gamma)$ [equation (6.22)], which effectively removes the radiative cap on X-point acceleration [because $t_{\rm X}$ is also proportional to γ; see equation (2.4)]. In contrast, secondary energization mechanisms generally operate on time-scales that grow more quickly with γ (e.g., as γ^2), and so maintain a finite cutoff, even for deeply Klein-Nishina radiative cooling.

(3) Finally, and also relevant to point 2 above, the energy distribution of produced pairs is very broad and non-thermal (section 7.2.2). These particles serve precisely as the pre-accelerated upstream population that, as argued in section 3.4, may help overcome the conventional $\gamma_{\rm X} \simeq 4\sigma_{\rm c}$ limit. Upon entering the reconnection layer, newborn particles possess Larmor radii ($\geq \tilde{\gamma}_{\rm cool,1}\rho_0$) that may exceed those, $\sim \sigma_{\rm c}\rho_0$, of typical accelerated particles – most of which come from the much colder and more numerous particle population (section 7.2.3) that was already present in the far upstream region. If so, then newborn pairs sample larger field structures than the elementary current layers and plasmoids, with size scale $\sigma_{\rm c}\rho_0$, at the bottom of the plasmoid hierarchy (Werner et al., 2016; Uzdensky, 2020). Unlike the vast majority of initially present cold upstream particles, these pairs may surf across many elementary layers – that together comprise a much larger acceleration region – becoming energized well beyond $4\sigma_{\rm c}$ before finally becoming magnetized. Thus, $\gamma_{\rm X}$ is effectively raised. This is important because higher $\gamma_{\rm X}$ enables direct acceleration to energize and collimate higher-energy particles, potentially increasing the energy range of kinetically beamed particles and photons.

Detailed predictions for kinetic beaming are outside the scope of this paper; my analysis

only traces the largest-scale bulk flow of energy and particles through the reconnection system. However, these simple observations provide a target for future dedicated simulations, which may study, in detail, the effects of Klein-Nishina and pair-production physics on kinetic beaming.

8.1.3 Caution against detailed spectral predictions

While I believe that the basic qualitative features of the model presented in chapter 7 are fairly robust, the model is not presently quantitatively accurate enough to warrant making specific spectral predictions for various astrophysical sources. The specific power-law scalings of the particle and photon spectra depend on a number of uncertain details.

For one thing, the NTPA 'equation of state' $p(\sigma_{\rm h}) = 1 + 2/\sqrt{\sigma_{\rm h}}$ from equation (7.14) is quite crude and only intended for illustration. The steady-state power-law scalings $p_{\rm T} = p + 1$ and (especially) $p_{\rm KN} = p - 0.5$ are also just rough estimates. For example, the bursty nature of reconnection is already known to modify $p_{\rm T}$ when Thomson IC cooling is quite strong (Werner et al., 2019).

As a related issue, the power-law index of the photon spectrum radiated by particles above $\gamma_{\rm KN}$ (in the $p_{\rm KN}$ portion of the particle distribution) is only very loosely given by $\Gamma = p_{\rm KN} + 1 = p + 0.5$. In fact, Moderski et al. (2005) estimate a harder power-law $\Gamma = p$ and predict little change between the power-law scaling of the emitted photon spectrum from the Thomson to Klein-Nishina regimes. Thus, given the level of quantitative uncertainty in my model, one should avoid inferring numerical values of plasma parameters (like $\sigma_{\rm h}$ and $z = \gamma_2/\gamma_{\rm KN}$) from observed photon spectra, especially at the highest energies (above $\gamma_{\rm KN} m_e c^2$). It is also probably not warranted to test the model by searching for a prominent Klein-Nishina spectral break.

Having discussed some generic observational features of the model of Klein-Nishina reconnection from chapter 7, as well as some of its quantitative limitations (with respect to interpreting power-law scalings of photon spectra), I now turn to exploring its potential applicability to some concrete classes of astrophysical systems.

8.2 Consequences for specific astrophysical systems

In this section, I explicitly estimate reconnection energy scales ($\gamma_{\rm max}$, $\sigma_{\rm c,0}$, $\gamma_{\rm rad,T}$, $\gamma_{\rm KN}$, etc.) for two types of astrophysical systems – flat-spectrum radio quasar (FSRQ) jets and black hole accretion disc coronae (ADCe). I also comment on the prospects for a solid understanding of Klein-Nishina radiative reconnection either to explain certain observational phenomena or to help constrain astrophysical details that are difficult to pin down from observations alone.

8.2.1 The radiative environments of FSRQs

As prerequisite astrophysical material for this discussion, the reader is referred to the Introduction (namely section 1.3) and chapter 4 for an introduction to blazars in general and to FSRQs in particular.

I focus here on FSRQs. Particularly important in the present context are the sources of external radiation in FSRQs that can provide intense illumination to the jet as it propagates relatively far away (up to several parsecs) from the supermassive black hole central engine. As described in chapter 4, two prominent external radiation sources for FSRQs are the broad emission line region (also 'broad-line region'; BLR) and dusty torus (also 'hot dust region'; HDR), which intercept and reprocess light originally emitted from the black hole accretion disc, redirecting some of the disc's radiated energy back onto the jet (e.g., Nalewajko et al., 2014; Madejski & Sikora, 2016). I now supply the details about these circumnuclear regions that are most relevant to my goal: estimating reconnection energy scales.

The smaller of the two regions is the BLR. It is made of gas that is partially ionized by the accretion disc light and hence shines UV line emission onto the jet (most prominently Ly α; Tavecchio & Ghisellini, 2008). Thus, the characteristic BLR photon energy is

$$\epsilon_{\rm BLR} \sim 10\,{\rm eV}\,. \tag{8.6}$$

The BLR extends out to a radius

$$r_{\rm BLR} \sim 0.1 L_{\rm d,46}^{1/2}\,{\rm pc} \tag{8.7}$$

where $L_{\rm d,46}$ is the luminosity of the accretion disc, $L_{\rm d}$, in units of $10^{46}\,{\rm erg\,s^{-1}}$ (Tavecchio & Ghisellini, 2008; Sikora et al., 2009; Nalewajko et al., 2012). At jet propagation distances $r < r_{\rm BLR}$ from the central engine, the jet traverses roughly isotropic ambient radiation, sourced by the BLR, of galaxy-frame energy density

$$U_{\rm BLR} \sim \frac{L_{\rm BLR}}{4\pi r_{\rm BLR}^2 c} \sim 6\times10^{-3}\,{\rm erg\,cm^{-3}} \tag{8.8}$$

(Tavecchio & Ghisellini, 2008; Sikora et al., 2009; Nalewajko et al., 2012). Here I assume that the broad-line region intercepts and reprocesses a certain fraction (2 per cent; Tavecchio et al., 2011) of the accretion disc light, and, therefore, that $L_{\rm BLR} \propto L_{\rm d}$. Thus, $U_{\rm BLR}$ is insensitive to $L_{\rm d}$.

Farther removed from the nucleus than the BLR is the hot dust region, which, radiatively heated by the accretion disc, shines a quasi-thermal spectrum of temperature $T_{\rm HDR} \simeq 1200\,{\rm K}$ onto the jet. Thus (see Nenkova et al. 2008a,b; Nalewajko et al. 2012; chapter 4),

$$\epsilon_{\rm HDR} \sim 3k_{\rm B}T_{\rm HDR} = 0.3\,{\rm eV}\,. \tag{8.9}$$

The HDR extends out to a distance from the central engine of (Nenkova et al., 2008a,b; Sikora et al., 2009)

$$r_{\rm HDR} \sim 4\,L_{\rm d,46}^{1/2} T_{\rm HDR,3}^{-2.6}\,{\rm pc}\,, \tag{8.10}$$

where $T_{\rm HDR,3} \equiv T_{\rm HDR}/1000\,{\rm K} = 1.2$. Thus, when $r < r_{\rm HDR}$, the HDR radiation energy density traversed by the jet is roughly isotropic and approximately

$$U_{\rm HDR} \sim \frac{L_{\rm HDR}}{4\pi r_{\rm HDR}^2 c} \sim 9\times10^{-5}\,{\rm erg\,cm^{-3}}\,. \tag{8.11}$$

Here, I again take a fixed fraction (in this case 10 per cent; Malmrose et al., 2011) of the disc radiation to be reprocessed by the circumnuclear structure. Thus, $U_{\rm HDR}$, like $U_{\rm BLR}$, lacks $L_{\rm d}$-dependence.

I posit that at least some quiescent and flaring blazar gamma-ray observations are powered by magnetic reconnection (Giannios et al., 2009; Nalewajko et al., 2011; Giannios, 2013; Sironi et al., 2015; Petropoulou et al., 2016; Werner et al., 2018; Nalewajko et al., 2018; Christie et al., 2019; Ortuño-Macías & Nalewajko, 2020; Sobacchi et al., 2021), either induced by macroscopic field polarity reversals (e.g., Giannios & Uzdensky, 2019; Sironi et al., 2021) or taking place at the small-scale terminus of a turbulent cascade (e.g., Zhdankin et al., 2013, 2020; Comisso & Sironi, 2018, 2019; Loureiro & Boldyrev, 2020; Boldyrev & Loureiro, 2020; Nättilä & Beloborodov, 2020; Sobacchi et al., 2021).

The blazar emission zone – the distance r where most of the emission is produced – is an important but difficult-to-constrain quantity in blazar research: by nature of being collimated along the observing line of sight, blazar jets appear point-like on the sky, and so it is not possible to deduce r directly from observations. I therefore allow r to vary over an appreciable range. If $r < r_{\rm BLR}$, the BLR dominates the ambient radiation bathing the jet and the emission-powering reconnection zone; if $r_{\rm BLR} < r < r_{\rm HDR}$, the HDR dominates.[1] I consider only blazar zones far enough from the central engine ($r \gtrsim 0.01\,{\rm pc}$; Dermer & Schlickeiser, 2002; Sikora et al., 2009; Nalewajko et al., 2014) that the direct accretion disc light, which illuminates the jet from behind, is redshifted in the jet rest-frame to a lower energy density than the BLR and HDR radiation fields. (Unlike the accretion disc light, the BLR and HDR photons impinge quasi-isotropically on the jet in the galaxy frame and are thus blueshifted when boosted to the jet-frame). I do not consider $r \gg r_{\rm HDR}$.

The BLR and HDR choke gamma-rays above the energies

$$\epsilon_{\rm c,BLR} = \frac{(m_e c^2)^2}{\epsilon_{\rm BLR}} \sim 30\,{\rm GeV} \tag{8.12}$$

and

$$\epsilon_{\rm c,HDR} = \frac{(m_e c^2)^2}{\epsilon_{\rm HDR}} \sim 0.9\,{\rm TeV}\,, \tag{8.13}$$

[1] In reality, the BLR intensity falls off smoothly with r and thus dominates to distances slightly exceeding $r_{\rm BLR}$ (but still much less than $r_{\rm HDR}$; e.g., Nalewajko et al. 2014).

respectively, since the pair-production optical depths suffered by a photon traversing these regions are

$$\tau_{\rm BLR} = \frac{U_{\rm BLR}\sigma_{\rm T}r_{\rm BLR}}{5\epsilon_{\rm BLR}} \sim 15 \tag{8.14}$$

and

$$\tau_{\rm HDR} = \frac{U_{\rm HDR}\sigma_{\rm T}r_{\rm HDR}}{5\epsilon_{\rm HDR}} \sim 300\,. \tag{8.15}$$

Note that these optical depths are *not* equal to those, $\tau_{\gamma\gamma,\rm BLR}$ and $\tau_{\gamma\gamma,\rm HDR}$, of the corresponding reconnection sites where the emission is sourced. Those optical depths are evaluated later using the smaller size of the emission region and with the seed photon number density U_i/ϵ_i ($i = \rm BLR\,or\,HDR$) boosted to the jet-frame. Since quiescent FSRQs are generally not observed at very high energies ($\gtrsim$ 0.1 TeV), the most relevant model for them may be irradiation by the BLR. However, FSRQ flares are sometimes observed at up to several hundred GeV (MAGIC Collaboration et al., 2008; Aleksić et al., 2011; H. E. S. S. Collaboration et al., 2013; Sitarek et al., 2015; Abeysekara et al., 2015; Ahnen et al., 2015) – still below $\epsilon_{\rm c,HDR}$. In these cases, emission between $r_{\rm BLR}$ and $r_{\rm HDR}$, where the external photons come from the HDR, is most likely.

8.2.2 FSRQ jet parameters

In order to estimate reconnection parameters, I must now make a number of further assumptions about the nature of the jet: including its speed, shape, and magnetic field strength.

I assume that reconnection occurs in the rest-frame of the jet and that the jet travels with relativistic bulk Lorentz factor $\Gamma_{\rm j} \gg 1$ with respect to the host galaxy. Like in chapter 4, I denote jet-frame quantities with primes. However, I exclude individual particle Lorentz factors (including cold magnetizations $\sigma_{\rm c,0}$) from this convention, writing them exclusively in the jet-frame and without primes. I adopt a relatively high jet Lorentz factor $\Gamma_{\rm j} = 40$. This is the same as in my earlier model (chapter 4; Mehlhaff et al., 2020) of the rapidly variable very high-energy ($\gtrsim 0.1$ TeV) flare from PKS 1222+21 (Aleksić et al., 2011). A large $\Gamma_{\rm j}$ may be reconciled with more typical

Lorentz factors (e.g., $\Gamma_{\rm j} \simeq 10$) by invoking a structured jet (e.g., chapter 4; Ghisellini et al., 2005; Begelman et al., 2008; Sikora et al., 2016; Tavecchio & Ghisellini, 2016; Tavecchio, 2017; Sironi et al., 2021) in which an inner spine region, moving quickly ($\Gamma_{\rm j} \simeq 40$) and carrying the reconnection current sheet, is surrounded by a slower-moving sheath. This explicit structure is not necessary to my ensuing discussion, however.[2]

I imagine that the jet is conical with opening angle $\theta_{\rm j} \sim 1/5\Gamma_{\rm j}$ (Pushkarev et al., 2009) and model the comoving magnetic field strength by assuming (Nalewajko et al. 2012; Mehlhaff et al. 2020; chapter 4)

$$B'_0(r) \sim 0.1 \left(\frac{r}{1\,{\rm pc}}\right)^{-1} {\rm G}\,, \tag{8.16}$$

corresponding to a total jet Poynting flux luminosity of $(\theta_{\rm j} r)^2(\Gamma_{\rm j}^2 B_0'^2)c \sim 1 \times 10^{43} {\rm erg\,s^{-1}}$ that is conserved in r and independent of $\Gamma_{\rm j}$. If the Poynting flux is instead dissipated (as indeed a reconnection scenario suggests), then the scaling of B'_0 with r should be steeper, but perhaps not by much. For example, the fiducial striped jet model of Giannios & Uzdensky (2019), which includes magnetic dissipation through reconnection, gives a power law close to $B'_0 \propto r^{-5/4}$.

I note that typical isotropic FSRQ luminosities $L_{\rm iso} \sim 10^{48}\,{\rm erg\,s^{-1}}$ require minimum intrinsic jet power, presumably carried by magnetic fields in a reconnection scenario, of $L_{\rm iso}/\Gamma_{\rm j}^2 \sim 6 \times 10^{44} {\rm erg\,s^{-1}}$. I have checked that increasing my fiducial magnetic field strength [e.g., $B'_0(1\,{\rm pc}) \sim 1\,{\rm G}$] to supply such jet power does not affect my conclusions in this chapter, but I adopt (8.16) to maintain continuity the earlier model in chapter 4.

8.2.3 Reconnection parameters in FSRQs

I now explicitly estimate parameters for reconnection in an FSRQ jet illuminated by either the broad-line region or the dusty torus.

[2] Unlike in my earlier model of chapter 4 (Mehlhaff et al., 2020), I do not entertain the sheath as a potential source of Compton seed photons here.

First, the maximum (system-size-limited) Lorentz factor is

$$\gamma_{\rm max} \equiv \frac{0.1L}{\rho_0} \sim \frac{0.1eB_0'(r)r\theta_{\rm j}}{m_e c^2} \sim 9\times10^{10}\left(\frac{\Gamma_{\rm j}}{40}\right)^{-1} . \tag{8.17}$$

Note that this is independent of r because I take the reconnection layer length L to be comparable to the transverse width $r\theta_{\rm j}$ of the jet, and thus the dependence of L on r cancels against that of the magnetic field strength in equation (8.16).

Next I discuss the jet magnetization. In lieu of estimating a fiducial jet electron number density (which is needed to estimate $\sigma_{\rm c,0}$), I take inspiration from previous studies in which the jet is moderately magnetized at the parsec scale (e.g., Giannios, 2013; Giannios & Uzdensky, 2019). If the jet is mass-dominated by protons, this still allows individual electrons/positrons to attain high Lorentz factors. Assuming an electron-proton jet – and that my main results from earlier sections carry over to electron-proton plasmas containing ultrarelativistic initial magnetic energy per electron – I adopt a cold *ion* magnetization (in the jet-frame) of 5, which corresponds to a fiducial jet-frame cold *electron* magnetization $\sigma_{\rm c,0} \sim 10^4$. To illustrate my high degree of uncertainty in this quantity, in Fig. 8.1, I also present the case $\sigma_{\rm c,0} \sim 10^3$, which corresponds to a cold ion magnetization of 0.5 (or of 5 but with 10 positrons per proton in the presence of mixed composition). In addition, I crudely assume that the cold magnetization does not change in r, which is consistent with Poynting flux conservation if one also has particle flux conservation [$\pi(r\theta_{\rm j})^2 n_0(r) =$ constant].

Now I estimate the key radiative parameters. For reconnection illuminated by the broad-line region ($r < r_{\rm BLR}$), one has, plugging the jet-frame seed photon energy $\epsilon'_{\rm BLR} \sim \Gamma_{\rm j}\epsilon_{\rm BLR}$ and energy density $U'_{\rm BLR} \sim \Gamma_{\rm j}^2 U_{\rm BLR}$ [with $\epsilon_{\rm BLR}$ and $U_{\rm BLR}$ given in equations (8.6) and (8.8)] into definitions (2.8), (2.9), (6.2) and (6.26):

$$\gamma_{\rm rad,T}^{\rm (BLR)} \sim 2\times10^6\left(\frac{r}{0.1\,{\rm pc}}\right)^{-1/2}\left(\frac{\Gamma_{\rm j}}{40}\right)^{-1} , \tag{8.18}$$

$$\gamma_{\rm cool,BLR} \sim 60\left(\frac{r}{0.1\,{\rm pc}}\right)^{-1}\left(\frac{\Gamma_{\rm j}}{40}\right)^{-1} , \tag{8.19}$$

$$\gamma_{\rm KN,BLR} \sim 300\left(\frac{\Gamma_{\rm j}}{40}\right)^{-1} \quad \text{and} \tag{8.20}$$

$$\tau_{\gamma\gamma,\rm BLR} \sim 3\left(\frac{r}{0.1\,{\rm pc}}\right) . \tag{8.21}$$

The observed photon energies emitted by particles with these radiative Lorentz factors are

$$\epsilon_{\rm obs}^{\rm (BLR)}\left(\gamma_{\rm rad,T}^{\rm (BLR)}\right) \sim \frac{1}{2}\Gamma_{\rm j}\gamma_{\rm rad,T}^{\rm (BLR)} m_e c^2 \sim 20\left(\frac{r}{0.1\,{\rm pc}}\right)^{-1}\,{\rm TeV}\,, \tag{8.22}$$

$$\epsilon_{\rm obs}^{\rm (BLR)}\left(\gamma_{\rm cool,BLR}\right) \sim \frac{4}{3}\Gamma_{\rm j}^2\gamma_{\rm cool,BLR}^2\epsilon_{\rm BLR} \sim 80\left(\frac{r}{0.1\,{\rm pc}}\right)^{-2}\,{\rm MeV} \quad \text{and} \tag{8.23}$$

$$\epsilon_{\rm obs}^{\rm (BLR)}\left(\gamma_{\rm KN,BLR}\right) \sim \frac{4}{3}\Gamma_{\rm j}^2\gamma_{\rm KN,BLR}^2\epsilon_{\rm BLR} \sim 2\,{\rm GeV}\,. \tag{8.24}$$

These photon energies are independent of the jet Lorentz factor $\Gamma_{\rm j}$. One of them, $\epsilon_{\rm obs}^{\rm (BLR)}(\gamma_{\rm KN,BLR}) \sim 2\,{\rm GeV}$, depends on no unknowns and implies that virtually all the $> 1\,{\rm GeV}$ FSRQ emission, routinely observed by the *Fermi* space telescope (Ghisellini, 2011), is emitted in the Klein-Nishina regime (if Comptonized from BLR photons). Note that, in evaluating the photon energies emitted by $\gamma = \gamma_{\rm KN}$ particles, one may use either the Thomson scaling $4\Gamma_{\rm j}^2\gamma_{\rm KN}^2\epsilon_{\rm ph}/3$ (as I have done) or the Klein-Nishina scaling $\Gamma_{\rm j}\gamma_{\rm KN}m_e c^2/2$; this choice modifies the estimate by less than a factor of 2.

If, instead, $r_{\rm BLR} < r < r_{\rm HDR}$, the radiation from the dusty torus is strongest, yielding, from the HDR estimates (8.9) and (8.11),

$$\gamma_{\rm rad,T}^{\rm (HDR)} \sim 6\times 10^6\left(\frac{r}{1\,{\rm pc}}\right)^{-1/2}\left(\frac{\Gamma_{\rm j}}{40}\right)^{-1}\,, \tag{8.25}$$

$$\gamma_{\rm cool,HDR} \sim 400\left(\frac{r}{1\,{\rm pc}}\right)^{-1}\left(\frac{\Gamma_{\rm j}}{40}\right)^{-1}\,, \tag{8.26}$$

$$\gamma_{\rm KN,HDR} \sim 1\times 10^4\left(\frac{\Gamma_{\rm j}}{40}\right)^{-1} \quad \text{and} \tag{8.27}$$

$$\tau_{\gamma\gamma,{\rm HDR}} \sim 20\left(\frac{r}{1\,{\rm pc}}\right)\,. \tag{8.28}$$

The corresponding observed photon energies are

$$\epsilon_{\rm obs}^{\rm (HDR)}\left(\gamma_{\rm rad,T}^{\rm (HDR)}\right) \sim \frac{1}{2}\Gamma_{\rm j}\gamma_{\rm rad,T}^{\rm (HDR)} m_e c^2 \sim 60\left(\frac{r}{1\,{\rm pc}}\right)^{-1}\,{\rm TeV}\,, \tag{8.29}$$

$$\epsilon_{\rm obs}^{\rm (HDR)}\left(\gamma_{\rm cool,HDR}\right) \sim \frac{4}{3}\Gamma_{\rm j}^2\gamma_{\rm cool,HDR}^2\epsilon_{\rm HDR} \sim 100\left(\frac{r}{1\,{\rm pc}}\right)^{-2}\,{\rm MeV} \quad \text{and} \tag{8.30}$$

$$\epsilon_{\rm obs}^{\rm (HDR)}\left(\gamma_{\rm KN,HDR}\right) \sim \frac{4}{3}\Gamma_{\rm j}^2\gamma_{\rm KN,HDR}^2\epsilon_{\rm HDR} \sim 70\,{\rm GeV}\,. \tag{8.31}$$

The photon energy $\epsilon_{\rm obs}^{\rm (HDR)}(\gamma_{\rm KN,HDR}) \sim 70\,{\rm GeV}$ [like the corresponding energy $\epsilon_{\rm obs}^{\rm (BLR)}(\gamma_{\rm KN,BLR})$ for Comptonization within the BLR] does not depend on any unknowns and implies that IC(HDR)

TeV emission from FSRQs is produced almost entirely in the Klein-Nishina regime. Irrespective of the dominant radiation field (BLR or HDR), the overall IC radiative cutoff energy, $\gamma_{\rm rad,IC}$, is effectively infinite, scaling according to equation (6.20) as $\gamma_{\rm rad,IC} \propto \gamma_{\rm KN} \exp[(2/9)(\gamma_{\rm rad,T}/\gamma_{\rm KN})^2]$.

Fig. 8.1 displays the BLR [(8.18)-(8.21)] and HDR [(8.25)-(8.28)] energy scales by plotting them as points in a radiative reconnection 'phase diagram' in the style of Fig. 6.7. This clearly illustrates the main result of this section: *reconnection in FSRQ jets proceeds in the regime governed by Klein-Nishina and pair production physics.* This does not depend on whether the dominant seed photons stem from the BLR or the HDR. Broadly speaking (I discuss caveats and technical points below), reconnection proceeds in the red region. Here, X-point acceleration is not inhibited by radiative losses ($\gamma_{\rm rad,IC} \gg \gamma_{\rm max} \gg \gamma_{\rm X}$), but subsequently most, if not all, particles cool strongly – on time-scales shorter than L/c – rendering reconnection efficiently radiative. Furthermore, because $\gamma_{\rm X} > \gamma_{\rm KN}$, many particles become impulsively energized up to energies where they radiate in the Klein-Nishina limit, producing above-threshold radiation that may activate pair feedback (chapter 7).

Before moving to additional observational implications, I discuss a few secondary technical details to flesh out this basic picture. First, though one might worry that the IC(HDR) model, depending on $\sigma_{\rm c,0}$, crosses into the purple region of Fig. 8.1 where the nominal intrinsic X-point acceleration cutoff energy, $\gamma_{\rm X} \simeq 4\sigma_{\rm c,0}$, falls below $\gamma_{\rm KN}$, this is not a huge concern. The observed photon energies $\epsilon_{\rm obs}^{\rm (BLR)}(\gamma_{\rm KN,BLR}) \sim 2\,{\rm GeV}$ and $\epsilon_{\rm obs}^{\rm (HDR)}(\gamma_{\rm KN,HDR}) \sim 70\,{\rm GeV}$ are routinely observed in quiescent FSRQ spectra (for the former; Ghisellini, 2011; Madejski & Sikora, 2016) and in TeV outbursts (for the latter; MAGIC Collaboration et al., 2008; Aleksić et al., 2011; H. E. S. S. Collaboration et al., 2013; Sitarek et al., 2015; Abeysekara et al., 2015; Ahnen et al., 2015). Thus, in a reconnection scenario, observations, in fact, suggest that $\gamma_{\rm X}$, whether through high $\sigma_{\rm c,0}$ or through circumventing the conventional $4\sigma_{\rm c,0}$-limit (as discussed in section 8.1.2), is high enough to accelerate particles beyond $\gamma_{\rm KN}$, pushing reconnection into the Klein-Nishina radiative regime.

Next, one should note that the pair-production optical depth of the BLR-illuminated reconnection layer is not very large: $\tau_{\gamma\gamma,\rm BLR} \sim 3$. According to equation (6.23), this means that even

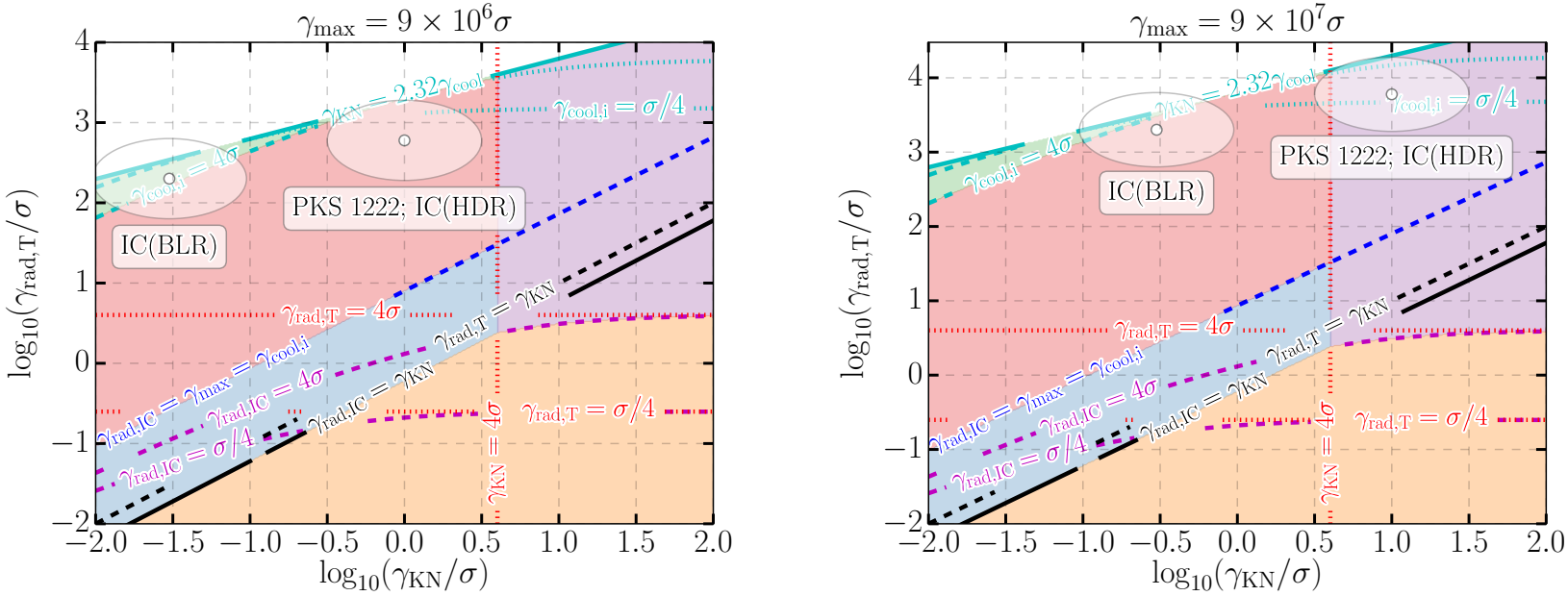

Figure 8.1: Radiative reconnection phase diagrams for the case $\gamma_{\rm max} = 9 \times 10^{10}$ [equation (8.17)] and two different values of $\sigma_{\rm c,0}$ (abbreviated as σ in the plot): $\sigma_{\rm c,0} = 10^4$ (top) and $\sigma_{\rm c,0} = 10^3$ (bottom). Region colorings have the same meaning as in Fig. 6.7. In each diagram, white points indicate the parameters of equations (8.18)-(8.21) and (8.25)-(8.28), where gamma-ray emission comes primarily through Comptonization of BLR photons [IC(BLR)] and HDR photons [IC(HDR)], respectively. Larger shaded white ellipsoids indicate one order-of-magnitude uncertainties in each direction. The IC(HDR) parameters are particularly relevant to the rapid very high-energy flare from PKS 1222+21 (Aleksić et al., 2011) discussed in chapter 4 (Mehlhaff et al., 2020). Because the three parameters $\gamma_{\rm rad,T}$, $\gamma_{\rm KN}$, and $\gamma_{\rm max}$ are all proportional to $\Gamma_{\rm j}^{-1}$, changing $\Gamma_{\rm j}$ is equivalent to changing $\sigma_{\rm c,0}$ by the same factor.

the most strongly radiative particles possess cooling times $t_{\rm cool,IC}(\gamma) \sim L/c\tau_{\gamma\gamma}$ only marginally faster than a dynamical time. This somewhat limits the overall radiative efficiency of reconnection, as signaled by one of the IC(BLR) points in Fig. 8.1 bordering on the green region where Klein-Nishina suppression of the IC cross section begins to shut down radiative cooling for the highest energy ($\gamma > \gamma_{\rm cool,2}$) particles. Plus, since $\tau_{\gamma\gamma} < 10$, some newborn pairs may escape the system before being swept into the layer ($f_{\rm noesc}$ drops below unity; see section 7.2.1), which diminishes the likelihood of swing cycles for the IC(BLR) scenario.

Regarding the multiplicity of produced pairs (section 7.2.3), the estimates of this section yield $\gamma_{\rm KN,BLR}/\sigma_{\rm c,0} \geq 0.03$ and $\gamma_{\rm KN,HDR}/\sigma_{\rm c,0} \geq 1$ for $\sigma_{\rm c,0} \leq 10^4$ and $\Gamma_{\rm j} \leq 40$. Thus, the requirement $\gamma_{\rm KN} > \sigma_{\rm c,0}/10$ to ensure a pair multiplicity (the ratio of number densities of newborn pairs to originally present pairs), η, less than unity is roughly satisfied. The BLR may marginally support $\eta \gtrsim 1$ because $\max(\eta) \simeq \sigma_{\rm c,0}/10\gamma_{\rm KN}$ (section 7.2.3). However this is doubtful for several reasons: (i) $\gamma_{\rm KN,BLR}/\sigma_{\rm c,0} = 0.03$ is a lower bound (increasing for smaller $\sigma_{\rm c,0}$ or $\Gamma_{\rm j}$); (ii) η may never attain its global maximum; and (iii) the result $\max(\eta) \simeq \sigma_{\rm c,0}/10\gamma_{\rm KN}$ ignores particle escape, which may be an issue for reconnection inside the BLR, as mentioned in the previous paragraph. Thus, the estimates of this section support the picture of tenuous newborn pairs that control the energy density of the inflow plasma but not the inflowing particle count.

8.2.4 Relevance to FSRQ observations

Next, I comment on a few observational issues that the model of pair-regulated Klein-Nishina reconnection from chapter 7 may help to address. First, and as discussed in section 8.1.2, Klein-Nishina reconnection may enhance the range of photon energies that are kinetically beamed. This provides an attractive explanation for rapid very high-energy FSRQ flares (of the kind first observed by Aleksić et al., 2011). A kinetic beaming scenario has already been advocated on energy-budget grounds by Nalewajko et al. (2012), and has recently been explored in this work (see also Mehlhaff et al., 2020) in the context of Thomson radiative reconnection. However, Thomson radiative reconnection requires a dense population of seed photons – much denser than those supplied by the

BLR and HDR (chapter 4) – in order to facilitate the high degree of radiative efficiency needed by kinetic beaming. Therefore, in chapter 4, I disfavor a single-zone external IC(HDR) scenario and suggest a two-zone (e.g., spine-sheath) configuration where the seed photons are supplied by the jet itself. However, if Klein-Nishina effects ease the radiative efficiency requirement, an IC(HDR) model may still be compatible with kinetic beaming.

Next, although I would like to refrain from trying to make detailed spectral interpretations (section 8.1.3, I note that the spectra of flaring FSRQs in the very high-energy (VHE; $\gtrsim 0.1\,\mathrm{TeV}$) band are generally quite steep, with intrinsic (deabsorbed) photon number index $\Gamma_{\rm VHE} \gtrsim 2.5$ (MAGIC Collaboration et al., 2008; Aleksić et al., 2011; Ahnen et al., 2015). If one assumes that there is no significant spectral break between the Thomson and deep Klein-Nishina regimes (Moderski et al., 2005), then $\Gamma_{\rm VHE} \simeq (p_{\rm T}+1)/2 \simeq (p+2)/2$. This suggests a steep injected index $p \gtrsim 3$, which is realized in my reconnection 'equation of state' $p(\sigma_{\rm h}) = 1 + 2/\sqrt{\sigma_{\rm h}}$, equation (7.14), if $\sigma_{\rm h,0} \lesssim 1$. This may occur when the ions are moderately magnetized and dominate the initial hot magnetization (e.g., Werner et al., 2018). In this situation, pair feedback, because it cannot reduce $\sigma_{\rm h}$ to below order-unity [equation (7.13)], would probably not strongly modify the spectrum.

However, if $\sigma_{\rm h,0} \gg 1$, pair feedback may produce the observed steep spectral indices during dramatic swing cycles (e.g., Fig. 7.8), wherein the observable emission is dominated by the layer (section 8.1.1) in the low-magnetization $\sigma_{\rm h,<} \sim 1$ part of the cycle. The requirement $\sigma_{\rm h,0} \gg 1$ could be realized if either: (i) the initial hot magnetization governing the distribution of accelerated electrons/positrons decouples from the (order-unity) ion magnetization, or (ii) a highly magnetized pair-dominated jet region furnishes the upstream material for reconnection.

In addition, to support swing cycles, the layer generally requires high energy recapture efficiency $\xi \simeq f_{\rm nocool} \sim 1$ (e.g., Fig. 7.8). In Appendix B, I show that order-unity energy retention factors $f_{\rm nocool}$ require a large cutoff $z = \gamma_2/\gamma_{\rm KN}$ (e.g. $z \gtrsim 1000$ in Fig. B.3). Such high cutoffs might not be realized because the observed Compton dominance in FSRQs is often $U_{\rm ph}/U_B \sim 100$, suggesting that the particle energy, $\gamma_{\rm s}$, beyond which synchrotron losses outcompete IC losses and

may hence limit further particle acceleration, is about $\gamma_{\rm s} \sim 30\gamma_{\rm KN}$ [equation (6.18)].[3]

However, there are a few effects that may serve to promote swing cycles in spite of nonzero synchrotron losses. First, the most energetic particles responsible for emitting pair-producing photons are likely accelerated near reconnection X-points deep inside the current layer. There, the magnetic field is weaker (Uzdensky et al., 2011; Cerutti et al., 2012a), which reduces synchrotron radiation relative to IC losses, potentially allowing the latter to remain dominant to higher energies than $\gamma_{\rm s}$. A second possibility is that swing cycles actually set in at lower values of z than I predict in Appendix B. This is because I conservatively estimate a steep distribution of pair-producing photons penetrating the upstream plasma $\Gamma = p + 0.5$ in section 7.2.2 (rather than, e.g., $\Gamma = p$ as in Moderski et al., 2005), and therefore may artificially underestimate $f_{\rm nocool}$. Finally, and related to the previous point, I find (not presented here) that the threshold on z for swing cycles to kick in is quite sensitive to the precise dependence of $f_{\rm nocool}$ on z. These considerations demonstrate the need for radiative kinetic simulations to examine in detail whether swing cycles are possible, when they occur, and how they impact the spectrum of observed radiation.

As a completely separate prospect, the model of chapter 7 may connect to the matter-antimatter balance in FSRQ jets, which is difficult to constrain from observations alone (Madejski & Sikora, 2016). This ratio may be initially imprinted at the base of the jet – for example, by magnetospheric spark-gap discharges (e.g., Blandford & Znajek, 1977; Beskin et al., 1992; Ford et al., 2018; Chen et al., 2018; Crinquand et al., 2020) or by interaction between the nascent jet and the accretion flow (e.g., Ripperda et al., 2020; Wong et al., 2021) – but it may also be modified *in situ* as the jet propagates. Although the model generally produces relatively few newborn pairs compared to the initial number of upstream particles in each reconnection episode (section 7.2.3), continuous, repeated reconnection occurring as the jet ploughs through ambient radiation fields could still lead to secular growth in the number of pairs present in the plasma. This may be

[3] Generally, $U_{\rm ph}/U_B$ is equated to the observed ratio of IC-to-synchrotron luminosities. This is sound in one-zone emission scenarios where Klein-Nishina effects do not suppress IC emission near the IC spectral peak. The latter requirement is generally satisfied because Compton FSRQ peaks typically fall at lower energies than $\epsilon_{\rm obs}^{\rm (BLR)}(\gamma_{\rm KN,BLR})$ and $\epsilon_{\rm obs}^{\rm (HDR)}(\gamma_{\rm KN,HDR})$.

relevant to FSRQs in their quiescent states, where their spectral cutoffs are broadly consistent with gamma-ray absorption in the broad-line region. Furthermore, because $\tau_{\rm BLR} > 1$, even if the reconnection region itself is optically thin ($\tau_{\gamma\gamma,\rm BLR} < 1$), emitted gamma-rays may still produce pairs in the jet before exiting the BLR (though they do not strongly impact the reconnection dynamics).

In summary, reconnection in FSRQ jets illuminated by either the broad-line region or the dusty torus is expected to occur in the radiative regime governed by pair-production and Klein-Nishina physics. Here, kinetic beaming and pair-feedback-initiated swing cycles may explain the time-scales of rapid TeV flares and typical FSRQ VHE spectral indices, respectively. However, confirming whether these mechanisms are active in Klein-Nishina radiative reconnection depends on several unknowns and remains a key open question for future simulations to address. In addition, *in situ* pair production driven by reconnection could contribute to the pair content of quiescent FSRQ jets.

8.2.5 Reconnection parameters in black hole ADCe

Magnetic reconnection may also power emission from a highly magnetized corona sandwiching a black hole accretion disc (Liang & Price, 1977; Galeev et al., 1979; Di Matteo, 1998; Uzdensky & Goodman, 2008; Goodman & Uzdensky, 2008; Uzdensky, 2016; Beloborodov, 2017; Werner et al., 2019; Sironi & Beloborodov, 2020). Here, I focus on the high/soft states of black hole X-ray binaries, showing that coronal reconnection in this context likely proceeds in a highly radiative Klein-Nishina regime. I term as high/soft any state where the quasi-thermal ($\sim 1\,$keV) spectral component – which is attributed to an optically thick, geometrically thin accretion disc (Shakura & Sunyaev, 1973) – strongly dominates over the non-thermal component, which extends to much higher, hard X- and gamma-ray energies (e.g., Remillard & McClintock, 2006). Because reconnection intrinsically gives rise to a non-thermal distribution of particle energies, a reconnection scenario does not rely on repeated scatterings by an isothermal plasma of moderate optical depth (e.g., Shapiro et al., 1976; Eardley & Lightman, 1976; Rybicki & Lightman, 1979) to explain the observed non-thermal

spectrum.

The bright quasi-thermal disc emission in the high/soft state provides an intense soft photon bath to the coronal reconnection region, and I assume that Comptonization of disc photons is responsible for the observed non-thermal spectrum. I now estimate the energy scales $\gamma_{\rm KN,disc}$ and $\gamma_{\rm rad,T}^{\rm (disc)}$ associated with the disc seed photons, as well as the energy scales $\sigma_{\rm c,0}$ and $\gamma_{\rm max}$. Determining these then allows me to accomplish my main goal for this section: characterizing the radiative regime of reconnection.

Beginning with $\gamma_{\rm KN,disc}$, the characteristic seed photon energy emitted from the disc is

$$\epsilon_{\rm disc} \sim 1\,{\rm keV}\,, \tag{8.32}$$

as both theorized (Shakura & Sunyaev, 1973) and observed (e.g., in Cyg X-1; McConnell et al., 2002; Remillard & McClintock, 2006). Through equation (6.9), the Klein-Nishina Lorentz factor is then

$$\gamma_{\rm KN,disc} \equiv \frac{m_e c^2}{4\epsilon_{\rm disc}} \sim 100\,. \tag{8.33}$$

The corresponding energy to which photons encountering $\gamma_{\rm KN,disc}$-particles are upscattered is

$$\epsilon_{\rm obs}^{\rm (disc)}\left(\gamma_{\rm KN,disc}\right) \sim 20\,{\rm MeV}\,. \tag{8.34}$$

Emission at these energies has been observed by the *Fermi* Large Area Telescope (LAT) from Cyg X-1 (Zdziarski et al., 2017).

I now move on to estimate $\gamma_{\rm rad,T}^{\rm (disc)}$. This requires [equation (2.8)] the energy density of the emission intercepted by the corona at a distance r from the black hole, which is roughly

$$U_{\rm disc} \sim \frac{L_{\rm disc}}{4\pi r^2 c}\,. \tag{8.35}$$

Here, $L_{\rm disc}$ is the luminosity of the disc (emitted mostly near its inner edge) and r is assumed to exceed the innermost disc orbit.

In the high/soft state $L_{\rm disc}$ can be quite high – up to several per cent of the Eddington limit, $L_{\rm Edd} = 4\pi G M_{\rm BH} m_p c/\sigma_{\rm T} = 4\pi r_g m_p c^3/\sigma_{\rm T} \simeq 1\times10^{39}(M_{\rm BH}/10M_\odot)\,{\rm erg\,s^{-1}}$ (e.g., 4 per cent

for Cyg X-1; McConnell et al., 2002). Furthermore, r is at least several gravitational radii, $r_g \equiv GM_{\rm BH}/c^2 \simeq 1\times10^6(M_{\rm BH}/10M_\odot)\,{\rm cm}$. Thus, an upper bound and fiducial scale for the energy density of ambient radiation shining onto the reconnection layer is (cf. Beloborodov, 2017)

$$U_\star \equiv \frac{L_{\rm Edd}}{4\pi r_g^2 c} = \frac{m_p c^2}{r_g \sigma_{\rm T}} \simeq 2\times10^{15}\left(\frac{M_{\rm BH}}{10M_\odot}\right)^{-1}\,{\rm erg\,cm^{-3}}\,. \tag{8.36}$$

Here, the fiducial black hole mass is $10M_\odot$. In the following, I estimate $r\sim r_g$ and $U_{\rm disc}\sim 0.04U_\star$, corresponding to $L_{\rm disc}\simeq 0.04L_{\rm Edd}$ observed from Cyg X-1 (McConnell et al., 2002). The parameter $U_\star$ also provides a fiducial scale for the magnetic field energy density in the disc required to transport angular momentum outward (Shakura & Sunyaev, 1973; Begelman et al., 1984; Goodman & Uzdensky, 2008; Beloborodov, 2017).

Beyond the fiducial scale $U_\star$, there are a number of uncertain geometric factors that control how the radiation and magnetic field energy densities decay as one moves from the vicinity of the black hole in the disc to the reconnection region in the corona. This leads to uncertainty in $\gamma_{\rm rad,T}^{\rm (disc)}$, which depends on both $U_{\rm disc}$ and the coronal magnetic field energy density, $U_{B,\rm cor}$. I parameterize these uncertainties by exhibiting, in the expressions to follow, $U_{\rm disc}$ as a multiple of $U_\star$ and $U_{B,\rm cor}$ as $U_{\rm disc}/C_{\rm d}$, where $C_{\rm d}\equiv U_{\rm disc}/U_{B,\rm cor}$ is the nominal Compton dominance. Without knowing the geometric factors governing how the seed photon and magnetic energy densities decay from the disc to the coronal reconnection region, I assume for simplicity that they fall off in the same way relative to $U_\star$, adopting a fiducial value $C_{\rm d}=1$. I discuss consequences on Klein-Nishina radiative reconnection if $C_{\rm d}$ is truly ~1 at the end of this section.

Following these conventions, one may write the fiducial upstream magnetic field strength as (e.g., Goodman & Uzdensky, 2008; Uzdensky, 2016; Beloborodov, 2017)

$$\begin{aligned} B_0 &\equiv \sqrt{8\pi U_{B,\rm cor}} = \sqrt{8\pi U_{\rm disc}/C_{\rm d}} \\ &\sim 4\times10^7\left(\frac{U_{\rm disc}}{0.04U_\star(M_{\rm BH})}\right)^{1/2}\left(\frac{M_{\rm BH}}{10M_\odot}\right)^{-1/2}\left(\frac{C_{\rm d}}{1}\right)^{-1/2}\,{\rm G} \end{aligned} \tag{8.37}$$

and then use equation (2.8) to write

$$\gamma_{\rm rad,T}^{\rm (disc)} \equiv \sqrt{\frac{0.3eB_{0,\rm cor}}{4\sigma_{\rm T}U_{\rm disc}}} = \left(\frac{0.3e}{4\sigma_{\rm T}}\sqrt{\frac{8\pi}{C_{\rm d}U_{\rm disc}}}\right)^{1/2}$$
$$\sim 6\times10^{3}\left(\frac{U_{\rm disc}}{0.04U_{\star}(M_{\rm BH})}\right)^{-1/4}\left(\frac{C_{\rm d}}{1}\right)^{-1/4}\left(\frac{M_{\rm BH}}{10M_{\odot}}\right)^{1/4}. \quad (8.38)$$

Though the complicated geometry precludes determining $U_{\rm disc}$ and $U_{B,\rm cor}$ precisely, the actual value of $\gamma_{\rm rad,T}^{\rm (disc)}$ depends only weakly on these quantities.

From equations (8.33) and (8.38), one sees that $\gamma_{\rm rad,T}^{\rm (disc)}$ is much higher than $\gamma_{\rm KN,disc}$. This opens up the possibility that radiative reconnection operates in the Klein-Nishina regime, since particles are not radiatively inhibited from becoming accelerated beyond $\gamma_{\rm KN,disc}$. Additionally, the radiative particle acceleration limit is not equal to $\gamma_{\rm rad,T}^{\rm (disc)}$, but is strongly modified by Klein-Nishina effects to $\gamma_{\rm rad,IC}^{\rm (disc)} \gg \gamma_{\rm rad,T}^{\rm (disc)}$ [equation (6.20)]. For reference, particles at $\gamma_{\rm rad,T}^{\rm (disc)}$ Comptonize disc photons to energy

$$\epsilon_{\rm obs}^{\rm (disc)}\left(\gamma_{\rm rad,T}^{\rm (disc)}\right) \sim \frac{1}{2}\gamma_{\rm rad,T}^{\rm (disc)}m_ec^2$$
$$\sim 2\left(\frac{U_{\rm disc}}{0.04U_{\star}(M_{\rm BH})}\right)^{-1/4}\left(\frac{C_{\rm d}}{1}\right)^{-1/4}\left(\frac{M_{\rm BH}}{10M_{\odot}}\right)^{1/4}\,{\rm GeV}\,. \quad (8.39)$$

Next, to fully specify the radiative reconnection regime, I estimate the remaining parameters $\gamma_{\rm max}$ and $\sigma_{\rm c,0}$. I assume that the characteristic current sheet length is

$$r_g \simeq 1\times10^{6}\left(\frac{M_{\rm BH}}{10M_{\odot}}\right)\,{\rm cm}\,. \quad (8.40)$$

This gives an extremely large system-size-limited Lorentz factor

$$\gamma_{\rm max} \equiv \frac{0.1r_geB_0}{m_ec^2} = \frac{0.1r_ge}{m_ec^2}\sqrt{\frac{8\pi U_{\rm disc}}{C_{\rm d}}}$$
$$\simeq 3\times10^{9}\left(\frac{U_{\rm disc}}{0.04U_{\star}(M_{\rm BH})}\right)^{1/2}\left(\frac{C_{\rm d}}{1}\right)^{-1/2}\left(\frac{M_{\rm BH}}{10M_{\odot}}\right)^{1/2}. \quad (8.41)$$

With $\gamma_{\rm KN,disc}$, $\gamma_{\rm rad,T}^{\rm (disc)}$, and $\gamma_{\rm max}$ estimated, $\gamma_{\rm cool,disc}$ [equation (2.9)] and $\tau_{\gamma\gamma,\rm disc}$ [equation (6.26)] follow. Respectively, they are

$$\gamma_{\rm cool,disc} \equiv \frac{3m_ec^2}{4U_{\rm disc}\sigma_{\rm T}r_g} \sim \frac{3}{4}\frac{U_{\star}}{U_{\rm disc}}\frac{m_e}{m_{\rm p}} \simeq 0.01\left(\frac{U_{\rm disc}}{0.04U_{\star}(M_{\rm BH})}\right)^{-1} \quad (8.42)$$

and

$$\tau_{\gamma\gamma,\mathrm{disc}} \equiv \frac{3\gamma_{\mathrm{KN,disc}}}{5\gamma_{\mathrm{cool,disc}}} \sim 8\times10^{3}\left(\frac{U_{\mathrm{disc}}}{0.04U_{\star}(M_{\mathrm{BH}})}\right). \tag{8.43}$$

Neither $\gamma_{\mathrm{cool,disc}}$ nor $\tau_{\gamma\gamma,\mathrm{disc}}$ depend on any unknowns besides the ratio $U_{\mathrm{disc}}/U_{\star}$. Unlike in FSRQ jets, the formal γ_{cool} is now small. Thus, it is not a physical Lorentz factor, but just signals that all particles cool to non-relativistic energies faster than L/c. In other words, the system is highly compact: $\ell_{\mathrm{cor}} \sim 1/\gamma_{\mathrm{cool,cor}} \gg 1$. For particles that have cooled to non-relativistic energies, IC scattering reduces to Thomson scattering and the observed photon energy is just $\epsilon_{\mathrm{obs}}^{(\mathrm{cor})}(1) = \epsilon_{\mathrm{cor}}$.

I move now to my final estimate: the cold electron magnetization. As in the case of FSRQ jets, I am fairly uncertain of this parameter because I do not know the background coronal electron density n_0. To proceed, let me recast n_0 in terms of the Thomson optical depth $\tau_{\mathrm{T}} = n_0\sigma_{\mathrm{T}}r_g$ along the reconnection layer [equation (8.40)]. This coincides with the optical depth of the corona itself assuming the coronal scale height is $\sim r_g$. The cold *electron* magnetization can then be compactly written in terms of τ_{T} and the *magnetic compactness*,

$$\ell_{\mathrm{B}} \equiv \frac{U_{\mathrm{B,cor}}\sigma_{\mathrm{T}}r_g}{m_ec^2} = \frac{U_{\mathrm{disc}}\sigma_{\mathrm{T}}r_g}{C_{\mathrm{d}}m_ec^2} = \frac{U_{\mathrm{disc}}m_{\mathrm{p}}}{C_{\mathrm{d}}U_{\star}m_e} \sim 70\left(\frac{U_{\mathrm{disc}}}{U_{\star}(M_{\mathrm{BH}})}\right)\left(\frac{C_{\mathrm{d}}}{1}\right)^{-1}, \tag{8.44}$$

as

$$\sigma_{\mathrm{c},0} \equiv \frac{2U_{B,\mathrm{cor}}}{n_0m_ec^2} \equiv 2\frac{\ell_{\mathrm{B}}}{\tau_{\mathrm{T}}}. \tag{8.45}$$

Thus, my uncertainty of n_0 is shifted onto τ_{T}. I consider a range of typical values $10^{-2} \leq \tau_{\mathrm{T}} \leq 1$ commonly inferred for coronae of accreting black holes, including in the high/soft state (e.g., Eardley & Lightman, 1976; Poutanen et al., 1997; Gierliński et al., 1999; McConnell et al., 2002; Goodman & Uzdensky, 2008; Beloborodov, 2017). Then, rewriting the right-hand-side of (8.45) as

$$\sigma_{\mathrm{c},0} \equiv 2\frac{\ell_{\mathrm{B}}}{\tau_{\mathrm{T}}} = \frac{2U_{\mathrm{disc}}m_{\mathrm{p}}}{C_{\mathrm{d}}U_{\star}\tau_{\mathrm{T}}m_e} \tag{8.46}$$

and plugging in $10^{-2} \leq \tau_{\mathrm{T}} \leq 1$ gives

$$1\times10^{2} \leq \sigma_{\mathrm{c},0}\left(\frac{C_{\mathrm{d}}}{1}\right)\left(\frac{U_{\mathrm{disc}}}{0.04U_{\star}(M_{\mathrm{BH}})}\right)^{-1} \lesssim 1\times10^{4}. \tag{8.47}$$

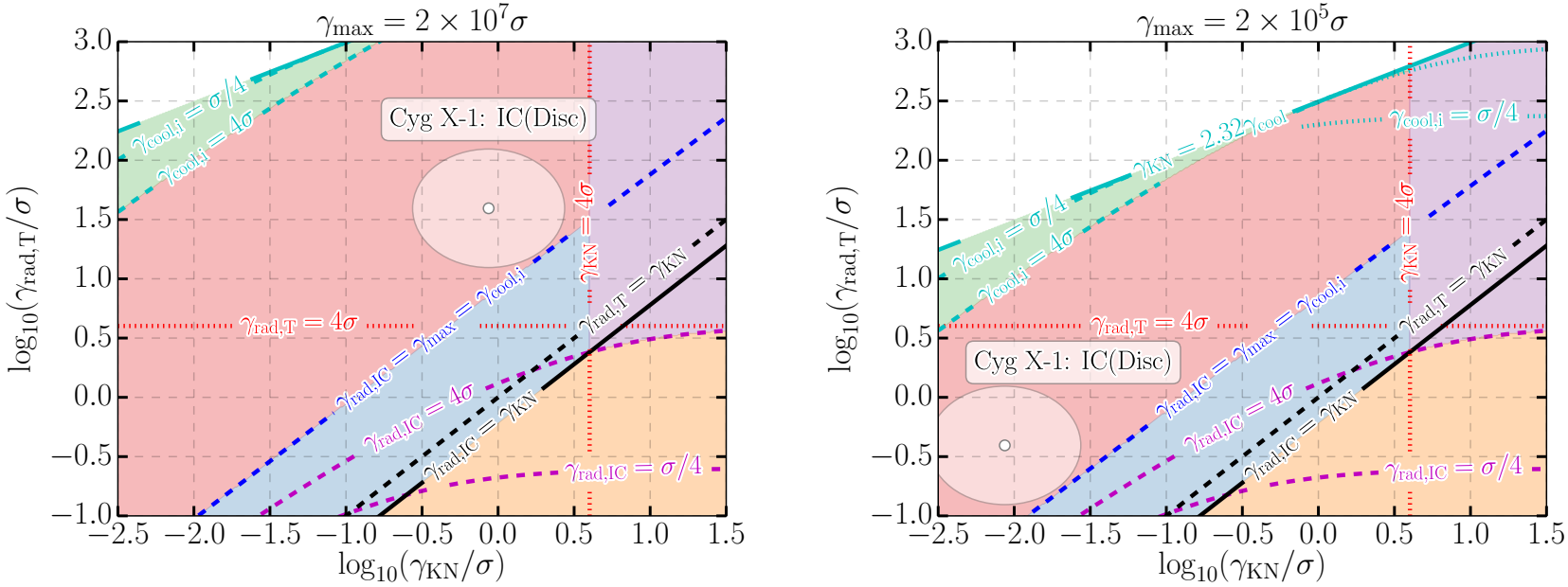

Figure 8.2: Radiative reconnection phase diagrams for parameters relevant to accreting black holes in their high/soft states (particularly Cyg X-1). Color-coding is the same as in Figs. 6.7 and 8.1. As in those figures, $\sigma_{\rm c,0}$ is abbreviated as σ. White points indicate the parameters $\gamma_{\rm rad,T}^{\rm (disc)}$ and $\gamma_{\rm KN,disc}$ estimated in this section; they are surrounded by ellipses that show an arbitrarily chosen uncertainty of one decade in each direction. I display two diagrams corresponding to a range (8.47) of the cold magnetization $\sigma_{\rm c,0}$ (based on typical Thomson optical depths $\tau_{\rm T}$). These are $\sigma_{\rm c,0} = 10^2$ ($\tau_{\rm T} = 1$; top) and $\sigma_{\rm c,0} = 10^4$ ($\tau_{\rm T} = 10^{-2}$; bottom). Throughout the $\sigma_{\rm c,0}$ range, reconnection operates in a domain strongly impacted by Klein-Nishina effects and pair production.

The corresponding range in the number density is $n_0 = \tau_{\rm T}/\sigma_{\rm T} r_g \in [10^{16}\,{\rm cm}^{-3}, 10^{18}\,{\rm cm}^{-3}]$ (cf. Goodman & Uzdensky, 2008).

I display the estimates in this section in a radiative reconnection phase diagram, as done for FSRQs earlier, in Fig. 8.2. This illustrates the main point of all these estimates: *coronal reconnection in the high/soft states of accreting black holes is likely strongly impacted by Klein-Nishina and pair-production physics.* For $\tau_{\rm T} \in [10^{-2}, 1]$, reconnection occurs in the red region of the phase diagram, where $\gamma_{\rm rad,IC} > \gamma_{\rm max} > \gamma_{\rm cool,2} > 4\sigma_{\rm c,0} > \gamma_{\rm KN} > \gamma_{\rm cool,1}$. Thus, radiative losses do not hinder acceleration near reconnection X-points, but virtually all particles are still cooled strongly (on time-scales much shorter than L/c). In addition, a broad distribution of particles develops at energies $\gg \gamma_{\rm KN}$. These particles emit above-threshold photons that then produce pairs which may feed back on reconnection.

I now mention a few subtler details associated with this picture. First, I note that the model of pair-regulated Klein-Nishina reconnection in black hole ADCe is complimentary to the reconnection scenario laid out in the context of the low/hard states of black hole X-ray binaries by Beloborodov (2017). Pair production also features in that earlier work, but, there, it occurs when two layer-Comptonized photons, both carrying energy $< m_e^2 c^4/\epsilon_{\rm ph}$, collide. In contrast, in this study, a pair is produced when a single Comptonized photon with even higher energy, $> m_e^2 c^4/\epsilon_{\rm ph}$, is absorbed by the seed radiation field.

Next, I mention the implications of my estimates for the pair multiplicity, η. Equations (8.33) and (8.47) show that the requirement $\gamma_{\rm KN,cor}/\sigma_{\rm c,0} \geq 0.1$ for η to be less than unity (section 7.2.3) is not always satisfied. Thus, the pairs sourced from gamma-ray absorption by the disc seed photons may sometimes reach a number density comparable to the background upstream plasma. However, a stronger nominal Compton dominance $C_{\rm d} \equiv U_{\rm disc}/U_{B,\rm cor}$ drives down $\sigma_{\rm c,0}$ [equation (8.47)] and pushes reconnection into the regime where the produced pairs are tenuous.

Lastly, I discuss the implications of the assumption $C_{\rm d} \sim 1$. As mentioned earlier, $U_{\rm disc}$ and $U_{B,\rm cor}$ are both set, modulo complicated geometric factors, by the scale $U_\star$ (or, in my specific estimates, by $0.04U_\star$). Thus, it is possible that $C_{\rm d}$ may be close to unity. However, the model

of pair-regulated Klein-Nishina reconnection (chapter 7) ignores synchrotron losses, and so, in the presence of such modest $C_{\rm d}$, synchrotron cooling may need to be suppressed in order for the picture of Compton cooling and pair regulation to remain valid. There are several ways for this to occur. Near reconnection X-points, the magnetic field is weaker, and so synchrotron cooling is reduced there (Uzdensky 2011; Cerutti et al. 2012a; see discussion of FSRQs). In addition, as discussed in Appendix C, the angular distribution of particles radiating above-threshold photons from the layer may cause the birth velocities of newborn upstream pairs to be nearly parallel to the reconnection midplane. Because the upstream magnetic field is also parallel to this plane, the pitch angles of such newborn pairs may be small, allowing their cooling to be IC-dominated. Finally, synchrotron self-absorption may inhibit synchrotron cooling among the lower-energy particles (e.g., with $\gamma \lesssim 10$; Beloborodov, 2017).

However, if $C_{\rm d} > 1$, then my estimates numerically change, but the ordering of scales (region of the radiative phase diagram occupied by reconnection) remains essentially the same. For reference, $\gamma_{\rm KN}/\sigma_{\rm c,0} \propto C_{\rm d}$; $\gamma_{\rm rad,T}/\sigma_{\rm c,0} \propto C_{\rm d}^{3/4}$; and $\gamma_{\rm max}/\sigma_{\rm c,0} \propto C_{\rm d}^{1/2}$. If, for illustration, one scales up $C_{\rm d}$ to 10, the ordering $\gamma_{\rm rad,IC} > \gamma_{\rm max} > \gamma_{\rm cool,2} > \gamma_{\rm rad,T} > 4\sigma_{\rm c,0} > \gamma_{\rm KN}$ is preserved except at the high-$\sigma_{\rm c,0}$ (low-$\tau_{\rm T}$) end of the interval (8.47). There, $\gamma_{\rm KN}$ switches places with $4\sigma_{\rm c,0}$, which brings reconnection into the magenta region (e.g., Fig. 8.2) of the phase diagrams where the nominal impulsive X-point acceleration limit, $4\sigma_{\rm c,0}$, falls below the minimum energy, $\gamma_{\rm KN}$, for Klein-Nishina effects kick in. However, like for FSRQs (section 8.2.3), one may infer from observations [equation (8.34)] that particles are accelerated at least up to $\gamma_{\rm KN}$, suggesting that, if $4\sigma_{\rm c,0} < \gamma_{\rm KN}$, the nominal $4\sigma_{\rm c,0}$ X-point acceleration limit is circumvented and that reconnection may still be strongly influenced by Klein-Nishina radiative physics.

8.2.6 Relevance to observations of black hole X-ray binaries

I now describe the implications of Klein-Nishina reconnection for high/soft states of black hole ADCe in accretion-powered X-ray binaries. I am aware of two such systems for which gamma-ray observations have been reported: Cyg X-1 (detected by the *Fermi* LAT, Zdziarski et al., 2017) and

Cyg X-3 (detected by *AGILE*, Tavani et al. 2009, and by the *Fermi* LAT, Fermi LAT Collaboration et al. 2009; Zdziarski et al. 2018; I adopt the hypothesis that the Cyg X-3 compact object is a black hole).

For Cyg X-1, gamma-ray observations place the high/soft-state photon energy cutoff at $40-80\,\mathrm{MeV}$ (Zdziarski et al., 2017). This is fairly consistent with absorption of gamma-rays propagating through the quasi-thermal radiation field of the disc. Naively, one expects absorption to kick in at around $\sim (m_e c^2)^2/\epsilon_{\rm disc} \sim 300\,\mathrm{MeV}$. However, the quasi-thermal spectrum is not monochromatic, but falls off smoothly at energies beyond its 1-keV peak. In fact, the disc dominates the observed spectrum at up to 3 or 4 keV (e.g., McConnell et al., 2002), which brings the absorption cutoff into the observed range of $40-80\,\mathrm{MeV}$. Note that the pair-production optical depth $\tau_{\gamma\gamma,\rm disc}$ [equation (8.43)] decreases from $\sim 10^4$ at photon energies beyond the spectral peak, but is still much larger than unity at $3-4$ keV. Alternatively, the photon energy cutoff could correspond to the cutoff in the distribution of radiating particles.

Unlike Cyg X-1, Cyg X-3 exhibits gamma-ray emission far above its expected absorption cutoff. However, these high-energy gamma-rays are thought to come from Comptonization of the companion star's radiation by jet electrons at much larger distances (Dubus et al., 2010; Zdziarski et al., 2018). Thus, gamma-ray observations of Cyg X-3 do not provide a strong constraint on the cutoff in the coronal emission, which may be similar to Cyg X-1 and also consistent with pair absorption.

As for their hard X-ray spectra, Cyg X-1 and Cyg X-3 both display high/soft states with power-law hard X-ray photon number indices $\Gamma_{\rm X,HS} \simeq 2.5$ (McConnell et al. 2002; Zdziarski & Gierliński 2004; Szostek et al. 2008; here 'HS' stands for 'high/soft'). In addition, as detailed by Remillard & McClintock (2006), many black hole X-ray binaries are observed with photon indices clustered around this value in their high/soft states (and also in 'steep power-law' states where the observed thermal and non-thermal luminosities are comparable). As discussed for FSRQ jets (section 8.2.4), the relatively steep index $\Gamma_{\rm X,HS} \simeq 2.5$ (corresponding to a reconnection-powered injected electron energy distribution power-law index $p \gtrsim 3$; section 8.2.4) could be associated with

swing cycles in pair-regulated Klein-Nishina reconnection.

Like FSRQs, accreting black holes also exhibit rapid flares. These are observed in the X-rays and have been witnessed on time-scales as short as milliseconds (Gierliński & Zdziarski, 2003) – close to the light crossing time of an orbit at $r = 10r_g$ from a black hole with mass $M_{\rm BH} = 10M_\odot$. Kinetic beaming in the context of Klein-Nishina reconnection may help explain such rapid variability by delivering observed time-scales as short as $0.1L/c$ (Cerutti et al., 2013, 2014b,a). However, this has not been explored in detail as in the case of rapid TeV FSRQ flares and requires further attention.

In summary, pair-regulated Klein-Nishina reconnection in the accretion disc coronae of black hole X-ray binaries may explain the observed steep spectra from these objects, including the gamma-ray cutoff (observed in Cyg X-1, Zdziarski et al. 2017, though this may also result from the cutoff in the distribution of radiating particles). Kinetic beaming facilitated by Klein-Nishina reconnection may additionally explain short flaring time-scales, but this requires much more thorough investigation.

In the next chapter, I collect and summarize the main theoretical and astrophysical features of Klein-Nishina radiative reconnection (chapters 6-8; Mehlhaff et al. 2021). From there, I delve into the details of computationally modeling reconnection in this regime using the PIC technique.

Chapter 9

Summary of Klein-Nishina Radiative Reconnection

In chapters 6-8, I present an analysis of Klein-Nishina radiative reconnection (first detailed by Mehlhaff et al., 2021). This is an important reconnection regime in astrophysical systems – such as FSRQ blazar jets and accreting black holes – where Comptonization of ambient radiation can enter the discrete Klein-Nishina limit. In addition to its astrophysical applications, this type of reconnection is also very interesting on an intrinsic theoretical level (e.g., section 1.5) – not only do Klein-Nishina effects modify the radiative cooling of individual particles, but they also open up an entirely new channel for radiative feedback on reconnection: pair production. In chapter 6, I incorporate Klein-Nishina and pair-production physics into the basic conceptual framework of radiative reconnection. I introduce the new energy scale, $\gamma_{\rm KN}$, brought in by the novel radiative physics and map out the possible regimes of Klein-Nishina reconnection by considering different orderings of this scale with respect to the others already familiar (e.g., chapter 2) from Thomson radiative reconnection.

Specializing to the unexplored hierarchies of reconnection energy scales identified in chapter 6, I find, in chapter 7, a fundamentally new, self-regulated, Klein-Nishina reconnection regime that stems from a negative feedback loop (Fig. 7.1). In this loop, gamma-ray photons Comptonized in the reconnection layer propagate into the upstream plasma where they produce hot pairs by colliding with soft (unscattered) ambient photons, thus lowering the effective upstream magnetization. This inhibits further particle acceleration in the reconnection layer, closing the feedback loop.

The pair-regulation mechanism gives rise to an effective fixed-point upstream magnetiza-

tion $\sigma_{\rm h}$ that is *universal* – independent of the initial value $\sigma_{\rm h,0}$. However, the fixed point is not necessarily reached by the system. If pair feedback is efficient enough, the system overshoots its natural solution, $\sigma_{\rm h}$, by a wide margin and enters a limit (or 'swing') cycle, constantly oscillating between a high- and a low-magnetization state.

Also in chapter 7, I analyze the number of particles produced in the upstream region. A pair cascade, in which the population in each subsequent newborn generation grows exponentially, is not generally expected (except for unrealistically optimal reconnection-driven NTPA). Even in the presence of a cascade, there is a wide parameter range where the total created pair count is dwarfed by that of the pairs originally present. Thus, remarkably, though the newborn pairs are so hot that they can dominate the upstream energy budget, they are often also so tenuous that they contribute negligibly to the upstream lepton density. This feature, in particular, distinguishes Klein-Nishina radiative reconnection from other pair-regulated regimes (Lyubarskii, 1996; Hakobyan et al., 2019).

Finally, in chapter 8, I discuss the observable aspects of pair-regulated Klein-Nishina reconnection. I expect that emission from the layer (i.e., the region threaded with reconnected flux) dominates that from the newborn component of the upstream plasma (section 8.1.1). I also (section 8.1.2) identify several reasons to expect that Klein-Nishina physics may promote kinetic beaming, which is an important mechanism for facilitating rapid flaring variability in reconnection (Cerutti et al., 2012b, 2013, 2014a,b; Nalewajko et al., 2012; Mehlhaff et al., 2020). Furthermore, I explicitly estimate the parameters (i.e., the energy scales introduced in chapters 2 and 6) governing the radiative regime of reconnection in FSRQ blazar jets (sections 8.2.1-8.2.4) and ADCe of black hole X-ray binaries (sections 8.2.5-8.2.6). Reconnection in both types of objects is very likely to be strongly impacted by Klein-Nishina physics, and may even enter the pair-regulated regime modeled in this work. In fact, the generally steep non-thermal spectra produced by FSRQ jets and black hole ADCe appear consistent with strong pair loading in my model, and the observed instances of rapid variability could be caused by the kinetic beaming mechanism. Pair-regulated reconnection may further provide a source for *in situ* pair production in FSRQ jets, where the lepton content is difficult to constrain observationally.

However, many of my observational remarks in chapter 8, as well as the basic physical features of my model from chapter 7, require additional testing through first-principles radiative PIC simulations, which can address some of the following key physical and observational questions:

- How robust is the pair feedback mechanism? In particular,
  - Is the universal magnetization solution ever realized?
  - Can an order-unity energy recapture efficiency ξ be achieved?
  - Is the feedback strong enough to induce late-time limit cycles that jump between high- and low-magnetization states?
- How many pairs are produced? Are they indeed few compared to the number of background particles, and does their number density scale with the reconnection parameters as predicted by my model?
- Are the emitted spectra produced by reconnection in the steady state (or in a limit cycle) consistent with my predictions? Are they consistent with observations?
- Can kinetic beaming operate in Klein-Nishina reconnection as argued here? If so, is beaming more or less prominent than in Thomson radiative reconnection (as diagnosed in chapters 3-4 and by Mehlhaff et al. 2020)?
- To what extent does the physical picture change in 3D, in the presence of ions, with a finite guide field, and (especially) with finite synchrotron losses?

PIC Klein-Nishina reconnection simulations would therefore function essentially as numerical experiments, providing a critical check on the model constructed so far and furnishing new insight upon which to build theoretical refinements. These simulations would also facilitate more straightforward connection to observations, showing in greater detail – and fully self-consistently – the radiative signatures of Klein-Nishina reconnection.

However, such simulations are computationally challenging. Most PIC codes assume, for example, that the number of particles in the system is conserved, a feature of reconnection that is

completely lost in the presence of pair production. Additionally, PIC codes often do not contain support, out-of-the-box, for tracked photons, which are required in order to self-consistently capture the pair-production physics. Thus, in the next chapter, I first detail how I have overcome these computational challenges, implementing within my group's PIC code, ZELTRON (Cerutti et al., 2013; Cerutti & Werner, 2019), all of the necessary radiative physics to model Klein-Nishina reconnection. Then, I briefly present preliminary results from a simulation that leverages these new capabilities, further exploration of which will, in the immediate future, yield answers to some of the above very important questions.

Chapter 10

Simulating Klein-Nishina Radiative Reconnection in PIC

In this chapter, I discuss (section 10.1) the computational aspects of modeling Klein-Nishina reconnection in PIC, including both algorithmic details (section 10.1.1) and resolution challenges (section 10.1.2). Then (section 10.2), I present preliminary results from a novel PIC simulation of Klein-Nishina radiative reconnection.

10.1 Computational issues

I begin by discussing computational challenges associated with modeling Klein-Nishina radiative reconnection. I focus first on algorithmic details, and second on the scales that need to be resolved.

10.1.1 Algorithmic details

Since its inception, the electromagnetic PIC code ZELTRON has supported various radiative physics. This includes both self-consistent synchrotron cooling (e.g., Cerutti et al., 2013, 2014a,b) and Thomson IC cooling (e.g., chapter 3; Werner et al. 2019; Mehlhaff et al. 2020). These are implemented in the low-energy classical regime as continuous drag forces, and they enter into the PIC particle push (i.e., the algorithm used to update particle momenta from the electromagnetic field) following the method of Tamburini et al. (2010).

The first step toward modeling Klein-Nishina reconnection self-consistently involves generalizing the IC radiative cooling of particles from the continuous low-energy Thomson regime to the

discrete high-energy Klein-Nishina limit. When radiative losses are discrete, it no longer makes sense to apply cooling to particles during the particle push, which is meant to numerically approximate a process that is fundamentally continuous. Instead, particles lose a large, order-unity fraction of their energy instantaneously, and this should be reflected in the underlying cooling procedure adopted in the code.

To handle this, I implemented a Monte Carlo IC cooling scheme in ZELTRON. If the product of a given particle's energy with the energy of ambient seed photons, parameterized through the Klein-Nishina parameter $q = \gamma/\gamma_{\rm KN} = 4\gamma\epsilon_{\rm ph}/m_e c^2$ [see equations (6.2), (6.3), and their surrounding discussion], is large enough (signaling departure from the Thomson regime) the continuous cooling of the particle is not applied during its push by the electromagnetic field. Instead, after the particle has been pushed, ZELTRON probabilistically decides whether the particle emits a photon. If so, the photon's momentum is directly subtracted from that of the particle.

For simplicity, the code currently assumes that the simulation is immersed in the static, homogeneous, isotropic, monochromatic ambient radiation field (6.1). Thus, the probability that a particle emits a photon is [cf. equation (6.9)]

$$p_{\rm emit} = \frac{c\sigma_{\rm T} U_{\rm ph}}{\epsilon_{\rm ph}} g_{\rm KN} \left(\frac{\gamma}{\gamma_{\rm KN}}\right) \Delta t \,, \tag{10.1}$$

where Δt is the simulation timestep. If the particle emits a photon, the photon energy ϵ_1 is randomly selected using the Klein-Nishina scattering kernel [cf. Jones 1968; Blumenthal & Gould 1970; equation A.1]

$$\frac{{\rm d}N}{{\rm d}t\,{\rm d}r}\,{\rm d}r \propto \frac{3}{(1+qr)^2}\left[2r\ln r + (1+2r)(1-r) + \frac{1}{2}\frac{(qr)^2}{1+qr}(1-r)\right]{\rm d}r \tag{10.2}$$

as a probability distribution function in r. The parameter r is a proxy for the scattered photon energy ϵ_1, defined as [cf. equation (A.2)]

$$r \equiv \frac{\epsilon_1/\gamma m_e c^2}{q\left(1-\epsilon_1/\gamma m_e c^2\right)} \,. \tag{10.3}$$

For the purposes of a scattering particle's Monte Carlo photon emission, the Klein-Nishina parameter q is a constant, fixed by the particle's pre-scattering energy γ and the (globally constant) seed

photon energy $\epsilon_{\rm ph}$. Thus, equation (10.2) is not a single distribution function on r, but a family of distribution functions – one for each value of q.

Next, because Klein-Nishina reconnection also involves absorption of the Comptonized photons by the static ambient radiation, the photons emitted by each particle need to be tracked[1] by the simulation. This is also a new capability I added to ZELTRON. Photons emitted by particles become tracked as their own chargeless, massless (and, hence, ballistically propagating in straight lines at the speed of light) 'particle' species. Once emitted, a simulated photon has, at each timestep, a certain probability to be absorbed by the seed photon background, producing an electron-positron pair. Thus, pair production is also a Monte Carlo process, with the probability for a photon of energy ϵ to be absorbed in a given timestep set to

$$p_{\rm abs} = \begin{cases} c\sigma_{\gamma\gamma}(s_0)U_{\rm ph}\Delta t/\epsilon_{\rm ph} & s_0 \geq 1 \\ 0 & \text{otherwise} \end{cases} , \tag{10.4}$$

where $\sigma_{\gamma\gamma}(s_0)$ is the pair-production cross section and $s_0 = \epsilon\epsilon_{\rm ph}/(m_e c^2)^2$ is the nominal (maximum possible) center-of-momentum (CM) energy for the interaction (cf. Gould & Schréder, 1967). The pair-production cross section is (cf. Gould & Schréder, 1967)

$$\sigma_{\gamma\gamma}(s_0) = \frac{3}{8}\frac{\sigma_{\rm T}}{s_0^2}\left[\frac{1+\beta_0^2}{1-\beta_0^2}\ln W_0 - \beta_0^2 \ln W_0 - \ln^2 W_0 - \frac{4\beta_0}{1-\beta_0^2} + 2\beta_0 + 4\ln W_0 \ln(1+W_0) + 4\left(\frac{\pi^2}{12} + {\rm Li}_2(-W_0)\right)\right] , \tag{10.5}$$

where β_0 and W_0 are both s_0-dependent and read

$$\beta_0^2(s_0) \equiv 1 - \frac{1}{s_0} \tag{10.6}$$

and

$$W_0(s_0) \equiv \frac{1+\beta_0(s_0)}{1-\beta_0(s_0)} . \tag{10.7}$$

[1] In this chapter, I use the word 'tracked' to denote simulated PIC particles, also sometimes called 'superparticles' or 'macroparticles'. These are usually envisioned as consisting of many physical particles, and the PIC weight of a 'tracked' particle can be thought of as the number of physical particles represented by that tracked particle. Note, by using the word 'tracked', I do *not* mean its other common definition in PIC contexts, where it is used to refer to a subset of the simulated particles for which high-cadence data is dumped to disk.

If a pair is produced, $c\beta_0(s_0)$ is equal to the maximum possible speed of the newborn electron and positron in their CM-frame. The last two terms in (10.5), $4\left(\pi^2/12 + \mathrm{Li}_2(-W_0)\right)$, are equal to the last term, $-4L(W_0)$, in equation 10 of (Gould & Schréder, 1967) but are corrected for the missing factor of 4 in that work. As a reminder, $\mathrm{Li}_2(x)$ is the dilogarithm [cf. equation (6.5) and (6.9)]. Currently, p_{abs} is required by ZELTRON to be less than unity, and this is checked at the beginning of the run using $\max(\sigma_{\gamma\gamma}) \simeq \sigma_{\mathrm{T}}/5$ and the fact that U_{ph}, ϵ_{ph}, and Δt are all global simulation constants.

If a pair is produced, the CM-energy $s = s_0(1-\cos\theta)/2 \in [1, s_0]$ needs to be randomly chosen according to the angle θ between the high-energy Comptonized photon's momentum and that of the (untracked) low-energy seed photon. The cumulative distribution function for s is

$$p_{\mathrm{CM-energy}}(s \leq s_0) = \frac{\sigma_{\gamma\gamma}(s)}{\sigma_{\gamma\gamma}(s_0)}\,. \tag{10.8}$$

In the Monte Carlo pair-production scheme, s plays the role that r does in the photon emission scheme described above – both need to be calculated randomly according to their respective distributions. Meanwhile, s_0 plays the role that q does for photon emission – these are set by the incident photon or particle energy and they decide the distribution functions according to which the random parameters, s and r, respectively, are drawn.

Once s is randomly selected, a new electron and positron are initialized, each with CM-frame speed $c\beta_0(s)$, traveling in randomly-selected (i.e., uniform on a sphere) opposite directions in the CM frame of the pair-creation event. These are then boosted to the simulation frame using the just-absorbed tracked photon's momentum (which very nearly equals the total momentum of the newborn pair because the seed photon momenta are so small) and their current is deposited onto the PIC grid.

I would also like to mention a couple additional features of ZELTRON that I omitted from the above description for simplicity. First, the threshold on q deciding when the numerical scheme switches from continuous (during the PIC particle push) Thomson IC cooling to general Klein-Nishina Monte Carlo cooling is a user-defined input parameter, and can even be set to zero (in

which case the continuous Compton cooling is simulated by the more general Klein-Nishina Monte Carlo algorithm).

Second, it is possible to under- or over-sample the emission of photons by individual particles according to computational needs. For example, in some instances, it may be necessary to undersample the IC-emitted photons in order to conserve memory. In other cases, one may need to overproduce tracked photons in order to more finely sample the distribution function of emitted photons (and, hence, of the pairs that the photons produce). Undersampling and oversampling are accomplished by tuning the PIC weights of the tracked photons with respect to those of the particles. By setting the weight of each tracked photon equal to the weight of its tracked scattering particle, the physical probability (10.1) is equal to the probability with which a tracked particle emits a tracked photon. However, by making the tracked photons 'heavier' with respect to their emitting particles, the probability with which a tracked particle produces a tracked photon can be reduced from $p_{\rm emit}$ by the ratio of the scattering particle weight to that of the emitted photon. In the undersampling case, particles are *always* cooled according to their physical probability $p_{\rm emit}$ to emit photons, but only a *subset* of the emitted photons are promoted to tracked simulation entities. One can also go in the other direction – making the tracked photons 'lighter' than the particles and increasing the probability to emit a tracked photon from $p_{\rm emit}$ by the same ratio. In this oversampling case, ZELTRON splits the tracked emitting particle into two: one copy that did radiate (with new weight equal to the emitted tracked photon weight) and one that did not (with weight equal to the weight of the original particle minus the weight of the new photon). This preserves the discreteness of the emission mechanism – otherwise oversampling the emitted photons would cause the particles to be cooled continuously, which is not how they actually cool in reality. The ratio of each tracked photon's weight to its scattering particle's weight is a user-specified parameter in the ZELTRON input file.

Finally, in ZELTRON, one is free to track photons emitted in the continuous Thomson IC regime, *even if the continuous cooling algorithm is used instead of the Monte Carlo scheme.* Although these photons do not produce pairs, calculating their passive propagation through the

system can provide one with useful observer-centric diagnostic information. At the same time, one is also free to impose a threshold on the minimum tracked photon energy in the ZELTRON input file. Setting this threshold equal to the minimum required energy $m_e^2c^4/\epsilon_{\rm ph}$ to be absorbed by the seed radiation field can allow one to only spend computational resources on tracking those photons that have the potential to feed back on reconnection by creating new pairs.

10.1.2 Resolving the necessary scales

Next, I cover a complementary issue associated with modeling Klein-Nishina reconnection in PIC. That is, even once the Klein-Nishina IC cooling and pair-creation algorithms described above are implemented, one still needs to have sufficient computational resources to order the reconnection energy scales $\gamma_{\rm max}$, $\gamma_{\rm rad,T}$, $\sigma_{\rm c,0}$, and $\gamma_{\rm KN}$ such that Klein-Nishina reconnection is realized.

Recall from chapter 6 that the basic ordering of these scales that brings reconnection into a Klein-Nishina radiative regime is $\gamma_{\rm cool} < \gamma_{\rm KN} < \gamma_{\rm rad,T} < \gamma_{\rm max}$ (Fig. 6.6). Let me briefly review how this hierarchy comes about. First, the pair-production optical depth $\tau_{\gamma\gamma} = 3\gamma_{\rm KN}/5\gamma_{\rm cool}$ [equation (6.26)] is roughly equal to the scale-separation between $\gamma_{\rm KN}$ and $\gamma_{\rm cool}$. Thus, one needs $\gamma_{\rm KN} > \gamma_{\rm cool}$ in order to maximize $\tau_{\gamma\gamma}$, which enhances the radiative efficiency of reconnection [e.g., equation (6.23)] and makes it easier for the photons to produce pairs on time-scales that are short relative to the global time-scale L/c of the reconnection layer (and on length-scales $\lambda_{\rm mfp}$ short compared to the system size L). At the same time, however, one also needs $\gamma_{\rm KN}$ to be somewhat smaller than $\gamma_{\rm rad,T}$ in order to accelerate an appreciable population of particles past the threshold energy $\gamma_{\rm pp} = 8\gamma_{\rm KN}$ to emit pair-producing photons. From these considerations, one sees that, not only must the key radiative Lorentz factors be ordered as $\gamma_{\rm cool} < \gamma_{\rm KN} < \gamma_{\rm rad,T} < \gamma_{\rm max}$, but it is further beneficial (as far as maximizing $\tau_{\gamma\gamma}$ on the one hand and the number of accelerated particles capable of emitting pair-producing photons on the other) if they can all be separated from one another by a wide margin.

However, one is limited in terms of computational resources, because the maximum possible scale-separation, within which all of the other scales must 'fit', is given by the number of cells

across the simulation. Let me define L as the length of the PIC simulation, which corresponds to the length of the reconnection layer in Fig. 7.1. Then, let me define the cell size, Δx, such that the number of cells along the reconnection layer is $N \equiv L/\Delta x$. The cell size in radiative kinetic simulations is set, as in non-radiative kinetic simulations, by the requirement to resolve the Debye length,

$$\lambda_{\rm D} = \sqrt{\frac{T_0}{4\pi n_0 e^2}} \simeq \frac{\sigma_{\rm c,0}\rho_0}{2\sqrt{\sigma_{\rm h,0}}}\,, \tag{10.9}$$

where, as in section 3.1, I define the Debye length in terms of the initial upstream plasma density n_0 and temperature T_0. The approximation on the far right in (10.9) holds when the initial background plasma is relativistically hot: $\theta_0 \equiv T_0/m_e c^2 \gg 1$ [such that $\sigma_{\rm h,0} = \sigma_{\rm c,0}/4\theta_0$; cf. discussion following equation (2.7)]. In terms of the factor $A \equiv \lambda_{\rm D}/\Delta x > 1$ with which the Debye length is numerically resolved, the maximum scale-separation supported in a given reconnection simulation is

$$N = \frac{L}{\Delta x} = \frac{AL}{\lambda_{\rm D}} \simeq 2A\sqrt{\sigma_{\rm h,0}}\frac{L}{\sigma_{\rm c,0}\rho_0}\,. \tag{10.10}$$

From this, it follows that the ratio $\gamma_{\rm max}/\sigma_{\rm c,0}$ is

$$\frac{\gamma_{\rm max}}{\sigma_{\rm c,0}} = \frac{N}{20A\sqrt{\sigma_{\rm h,0}}}\,. \tag{10.11}$$

The upstream Debye length $\lambda_{\rm D}$ also limits how small $\gamma_{\rm cool}$ can be, and, hence, controls the maximum allowed value of the ratios $\gamma_{\rm rad,T}/\gamma_{\rm cool} = \gamma_{\rm max}/\gamma_{\rm rad,T}$ (which one would like to maximize in order to achieve the scale ordering $\gamma_{\rm cool} \ll \gamma_{\rm KN} \ll \gamma_{\rm rad,T} \ll \gamma_{\rm max}$). If $\gamma_{\rm cool} < \theta_0$, then the upstream plasma radiates on time-scales faster than L/c and, hence, faster than the duration of the reconnection event, implying that the plasma feeding the layer is not constant in time. While this alone makes it physically challenging to interpret the simulation results, there is another, more numerical issue associated with efficient upstream cooling. That is, the upstream Debye length can become unresolved, leading to unphysical numerical heating. The smallest that $\gamma_{\rm cool}$ can be to avoid both these issues, such that the upstream plasma only cools marginally over the course of a simulation, is $\gamma_{\rm cool} = 3\theta_0$. In this case, the initial background plasma cooling time is L/c, and so the original upstream material cools by a factor of a few over the course of several L/c, a typical

simulation duration. Assuming $\sigma_{\rm c,0}/4\sigma_{\rm h,0} = \theta_0 \gg 1$, the ratio $\gamma_{\rm cool}/\sigma_{\rm c,0}$ can be written in terms of just the single parameter $\sigma_{\rm h,0}$:

$$\frac{\gamma_{\rm cool}}{\sigma_{\rm c,0}} = \frac{3\theta_0}{\sigma_{\rm c,0}} = \frac{3}{4\sigma_{\rm h,0}}\,. \tag{10.12}$$

This implies, through (10.11), that

$$\frac{\gamma_{\rm max}}{\gamma_{\rm cool}} = \frac{N\sqrt{\sigma_{\rm h,0}}}{15A}\,, \tag{10.13}$$

and, because $\gamma_{\rm rad,T}^2 = \gamma_{\rm max}\gamma_{\rm cool}$, that

$$\frac{\gamma_{\rm rad,T}}{\sigma_{\rm c,0}} = \sqrt{\frac{3N}{80A\sigma_{\rm h,0}^{3/2}}}\,. \tag{10.14}$$

Thus, the scale ratios $\gamma_{\rm max}/\sigma_{\rm c,0}$, $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$, and $\gamma_{\rm cool}/\sigma_{\rm c,0}$ are all set by the two numerical parameters, N and A, along with one physical parameter: the initial upstream hot magnetization $\sigma_{\rm h,0}$. Therefore, N, A, and $\sigma_{\rm h,0}$ specify two ($\gamma_{\rm max}/\sigma_{\rm c,0}$ and $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$) of the three available parameters ($\gamma_{\rm max}/\sigma_{\rm c,0}$, $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$ and $\gamma_{\rm KN}/\sigma_{\rm c,0}$) governing the point in the radiative phase diagram (in the sense of Fig. 6.7) in which reconnection operates. For a given N, A, and $\sigma_{\rm h,0}$, the only parameter that remains to be chosen is $\gamma_{\rm KN}/\sigma_{\rm c,0}$, yielding a line of possible locations in the 3D radiative reconnection parameter space. Fixing $\gamma_{\rm KN}/\sigma_{\rm c,0}$ then amounts to choosing a point on that line.

Now, $\gamma_{\rm KN}$ brings in a variety of secondary, derived scales (e.g., $\gamma_{\rm cool,2}$ and $\gamma_{\rm rad,IC}$; see chapter 6) that cannot be expressed as simple power-law scalings of the other reconnection parameters. Thus, it is not immediately obvious from analytic expressions involving $\gamma_{\rm KN}$, $\gamma_{\rm rad,IC}$, and $\gamma_{\rm cool,2}$ what regimes of reconnection are accessible – even with the other two main radiative scale ratios, $\gamma_{\rm max}/\sigma_{\rm c,0}$ and $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$, fixed by the available computational resources. Therefore, rather than proceeding analytically, the procedure I now follow is to simply choose computationally reasonable values of N and A (and also a theoretically suitable value of $\sigma_{\rm h,0} \gg 1$), and, from these, to construct the radiative reconnection phase diagram. Then, using the phase diagram, I identify appropriate values of $\gamma_{\rm KN}/\sigma_{\rm c,0}$ for realizing Klein-Nishina reconnection by visual inspection.

In 2D, it is feasible for me to run simulations that are a few thousand cells long in one dimension. I therefore set $N = 2560$. I also find that a Debye length resolution of $A = \lambda_{\rm D}/\Delta x = 1.2$ is sufficient. Finally, I set $\sigma_{\rm h,0} = 100$, which is well into the magnetically dominated regime and opens up a large dynamic range of potential pair-loaded effective magnetization values, $\sigma_{\rm h}$, between 1 and $\sigma_{\rm h,0}$ to which the reconnection layer might regulate itself (as in chapter 7). Using these fiducial values, one has

$$\begin{aligned}
\frac{\gamma_{\rm max}}{\gamma_{\rm cool}} &= \frac{N}{15A}\sqrt{\sigma_{\rm h,0}} \simeq 1.4\times10^3 \left(\frac{N}{2560}\right)\left(\frac{A}{1.2}\right)^{-1}\left(\frac{\sigma_{\rm h,0}}{100}\right)^{1/2}, \\
\frac{\gamma_{\rm max}}{\sigma_{\rm c,0}} &= \frac{N}{20A\sqrt{\sigma_{\rm h,0}}} \simeq 11 \left(\frac{N}{2560}\right)\left(\frac{A}{1.2}\right)^{-1}\left(\frac{\sigma_{\rm h,0}}{100}\right)^{-1/2}, \\
\frac{\gamma_{\rm rad,T}}{\sigma_{\rm c,0}} &= \sqrt{\frac{3N}{80A\sigma_{\rm h,0}^{3/2}}} \simeq 0.28 \left(\frac{N}{2560}\right)^{1/2}\left(\frac{A}{1.2}\right)^{-1/2}\left(\frac{\sigma_{\rm h,0}}{100}\right)^{-3/4} \quad \text{and} \\
\frac{\gamma_{\rm cool}}{\sigma_{\rm c,0}} &= \frac{3}{4\sigma_{\rm h,0}} = 7.5\times10^{-3}\left(\frac{\sigma_{\rm h,0}}{100}\right)^{-1}. \qquad (10.15)
\end{aligned}$$

Next, using the ratios $\gamma_{\rm max}/\sigma_{\rm c,0}$ and $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$ as specified above (but leaving $\gamma_{\rm KN}/\sigma_{\rm c,0}$ free for now), I can construct the radiative reconnection phase diagram in Fig. 10.1 after the manner of Fig. 6.7. The figure shows that, in the presence of limited computational resources, the reconnection regimes featuring Klein-Nishina effects occupy smaller slivers of the available parameter space than in the astrophysical systems treated earlier (see section 8.2, Figs. 8.1 and 8.2). I am speaking here of the blue and red areas, where, as a reminder, the nominal intrinsic X-point Lorentz factor, $\gamma_{\rm X} = 4\sigma_{\rm c,0}$, exceeds $\gamma_{\rm KN}$ but also falls below the overall IC radiative saturation Lorentz factor, $\gamma_{\rm rad,IC}$. This allows particles to be accelerated well above $\gamma_{\rm KN}$, where they may emit pair-producing radiation. In addition, both the red and blue regions are efficiently radiative, with $\tau_{\gamma\gamma} \gg 1$ and, hence, with a fastest cooling time of particles $\min(t_{\rm cool,IC}(\gamma)) \sim L/\tau_{\gamma\gamma}c$ [equation 6.23] much shorter than L/c.

The dashed black line in Fig. 10.1 denotes the locations accessible to reconnection by changing the value of $\gamma_{\rm KN}/\sigma_{\rm c,0}$ with both $\gamma_{\rm max}/\sigma_{\rm c,0}$ and $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$ fixed according to (10.15). Note, in particular, that both the blue and red regions are reachable. There is some freedom, even, in $\gamma_{\rm KN}/\sigma_{\rm c,0}$ such that reconnection falls into one of these two areas. For the simulations presented

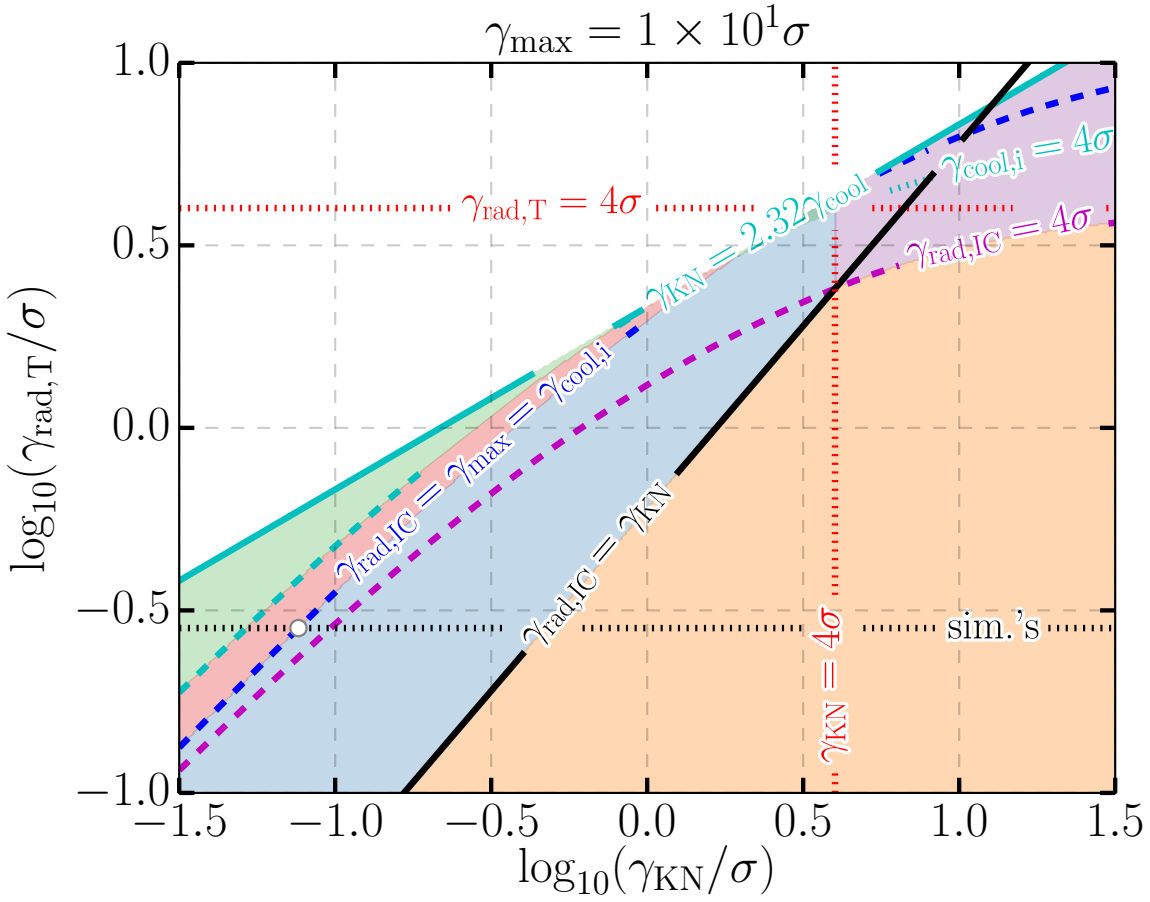

Figure 10.1: Radiative reconnection phase diagram for $N = 2560$, $\sigma_{\rm h,0} = 100$, and $A = 1.2$, which fixes $\gamma_{\rm max}/\sigma_{\rm c,0}$, $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$, and $\gamma_{\rm cool}/\sigma_{\rm c,0}$ as in (10.15). The one-dimensional set of locations in the diagram where reconnection may operate as the result of choosing $\gamma_{\rm KN}/\sigma_{\rm c,0}$ is shown as a black dotted horizontal line (labeled *sim.'s*). The value of $\gamma_{\rm KN}/\sigma_{\rm c,0}$ chosen for the simulation presented in section 10.2 is shown as a white dot. The coloring of the regions in the diagram is the same as in the earlier diagrams: Figs. 6.7, 8.1, and 8.2 (see section 6.2.2 for a complete explanation of the scale hierarchies defining each region). Limited computational resources make the blue and red regions, which feature prominent Klein-Nishina and pair-production physics, much smaller than they are in the presence of astrophysical scale separations (e.g., Figs. 8.1 and 8.2).

below, I choose $\gamma_{\rm KN}/\sigma_{\rm c,0}$ such that reconnection falls on the border between the red and blue regions – along the triple point $\gamma_{\rm max} = \gamma_{\rm rad,IC} = \gamma_{\rm cool,2}$ (discussed in section 6.2). This represents a comfortable compromise: $\gamma_{\rm rad,IC} = \gamma_{\rm max}$ effectively removes the radiative cap on particles accelerated near reconnection X-points; $\gamma_{\rm max} = \gamma_{\rm cool,2}$ ensures that all particles accelerated above $\gamma_{\rm KN}$ are still efficiently radiative, cooling on time-scales $t_{\rm cool,IC}(\gamma < \gamma_{\rm cool,2}) < L/c$; and the pair-production optical depth $\tau_{\gamma\gamma} = 3\gamma_{\rm KN}/5\gamma_{\rm cool} \simeq 5.9$ is fairly large. Straying farther into the blue region would increase the pair-production optical depth at the expense of lowering $\gamma_{\rm rad,IC}$, potentially reducing the number of particles at high-enough energies to emit pair-producing photons. Straying farther into the red region reduces $\tau_{\gamma\gamma}$ and brings $\gamma_{\rm cool,2}$ into the range where some particles at the highest energies may be weakly radiative.

In the next section, I present preliminary results from a simulation run using $N = 2560$, $A = 1.2$, and $\sigma_{\rm h,0} = 100$, which fixes the parameters $\gamma_{\rm max}/\sigma_{\rm c,0}$, $\gamma_{\rm rad,T}/\sigma_{\rm c,0}$, and $\gamma_{\rm cool}/\sigma_{\rm c,0}$ to their values in (10.15). The simulation also uses the value $\gamma_{\rm KN}/\sigma_{\rm c,0} = 0.076$, which is plotted in Fig. 10.1 as a white dot.

10.2 Simulation results

The simulation presented in this section is very preliminary. At the time of this writing, it is only about one week old. With it, I aim to make two basic points. The first is that the radiative physics discussed in section 10.1.1 is implemented correctly and working at scale. The second is that, even this early simulation already shows evidence for a very basic aspect of the pair-regulated reconnection model presented in chapter 7: the created pairs are quite tenuous – everywhere much less dense than the originally present plasma particles – but they carry a significant fraction of the material (rest mass + kinetic) energy, even dominating the matter energy flux across the reconnection separatrices.

The simulation setup is the same double-periodic box geometry as described for the simulations in section 3.1. Many of the parameters are also set to the same values as in those earlier simulations. In table 10.1, I summarize the simulation parameters used in this section, highlighting

in red the values that differ from the simulations presented in chapter 3. I note that, besides the system-size and radiative Lorentz factors $\gamma_{\rm max}$, $\gamma_{\rm rad,T}$, and $\gamma_{\rm KN}$, the guide field strength is also different from the earlier simulations: reduced from $B_z = B_0/4$ to $B_z = B_0/10$. This is to get as close as possible to the zero-guide-field limit, which promotes non-thermal particle acceleration (NTPA; see Werner & Uzdensky, 2017), but to still retain finite B_z, which protects plasmoids from radiatively cooling and compressing until the Debye length in their cores becomes unresolved.

Over the next several pages, I show a variety of figures that display different aspects of the evolution of the reconnection system at the times $ct/L = 0.33, 0.57, 0.82, 1.07, 1.31,$ and 1.72. First, in Fig. 10.2, I show the evolution of three densities: the number density of the original background plasma particles (those present from the beginning of the simulation as described in section 3.1), the energy density of above-threshold photons in the simulation, and the number density of pairs produced during the simulation. Concentrating first on the (original) background particles, one sees that the overall qualitative evolution of the reconnection layer is quite similar to more familiar regimes, such as that treated in chapter 3. The layer first fragments into many small plasmoids that begin to move about and merge with one another (e.g., the first three time snapshots in Fig. 10.2). After some time, the layer becomes dominated by just two larger plasmoids that undergo one large spectacular merger (e.g., $ct/L = 1.72$ in Fig. 10.2).

However, unique to reconnection in this regime, one can now view how the photons and produced pairs play into this picture. Looking at the above-threshold photon radiation energy density in Fig. 10.2, one sees that the reconnection system literally lights up at active reconnection sites. This includes both the main reconnection X-point (especially evident as a source of above-threshold photons at early times; e.g., $ct/L = 0.33$ in Fig. 10.2) and merging-plasmoid reconnection layers (e.g., $ct/L = 0.82$ in Fig. 10.2). In both locations, the reconnection electric field linearly accelerates unmagnetized particles to high, gamma-ray emitting energies.

Once emitted, photons propagate in straight lines at the speed of light. They are insensitive to the electromagnetic field and therefore do not respect the separatrices; they freely pass back into the upstream region. This has important implications on the spatial distribution of the pairs produced

Table 10.1: Simulation parameters used in this section. The parameters are largely the same as in chapter 3, but differences are highlighted in red.

Parameter	Symbol (=definition)	Value
Upstream magnetic field	B_0	
Nominal gyroradius	$\rho_0 = m_e c^2/eB_0$	
Thomson saturation Lorentz factor	$\gamma_{\rm rad,T}$	$0.282843 \times \sigma_{\rm c,0}$
Klein-Nishina Lorentz factor	$\gamma_{\rm KN}$	$0.0762437 \times \sigma_{\rm c,0}$
System size	$L_x = L$	$106.667 \times \sigma_{\rm c,0}\rho_0$
Reconnection Hillas Lorentz factor	$\gamma_{\rm max}$	$10.6667 \times \sigma_{\rm c,0}$
Cold magnetization	$\sigma_{\rm c,0} = B_0^2/4\pi n_0 m_e c^2$	10^4
Hot magnetization	$\sigma_{\rm h,0} = B_0^2/16\pi n_0 T_0$	10^2
Background temp.	$\theta_0 = T_0/m_e c^2$	25
Guide field	B_z	$B_0/10$
Peak drift over background density	$\eta = n_{\rm d}/n_0$	5
Harris layer drift velocity	$\beta_{\rm d} c$	$0.3c$
Harris layer (comoving) temp.	$\theta_{\rm d} = T_{\rm d}/mc^2$	1050
Harris layer half-thickness	$\delta = \sigma_{\rm c,0}\rho_0/\beta_{\rm d}\eta$	$0.67\sigma_{\rm c,0}\rho_0$
Cell size	Δx, Δy	$\sigma_{\rm c,0}\rho_0/24$
Time step	Δt	$0.7\Delta x/c$
Initial macroparticles per cell		80

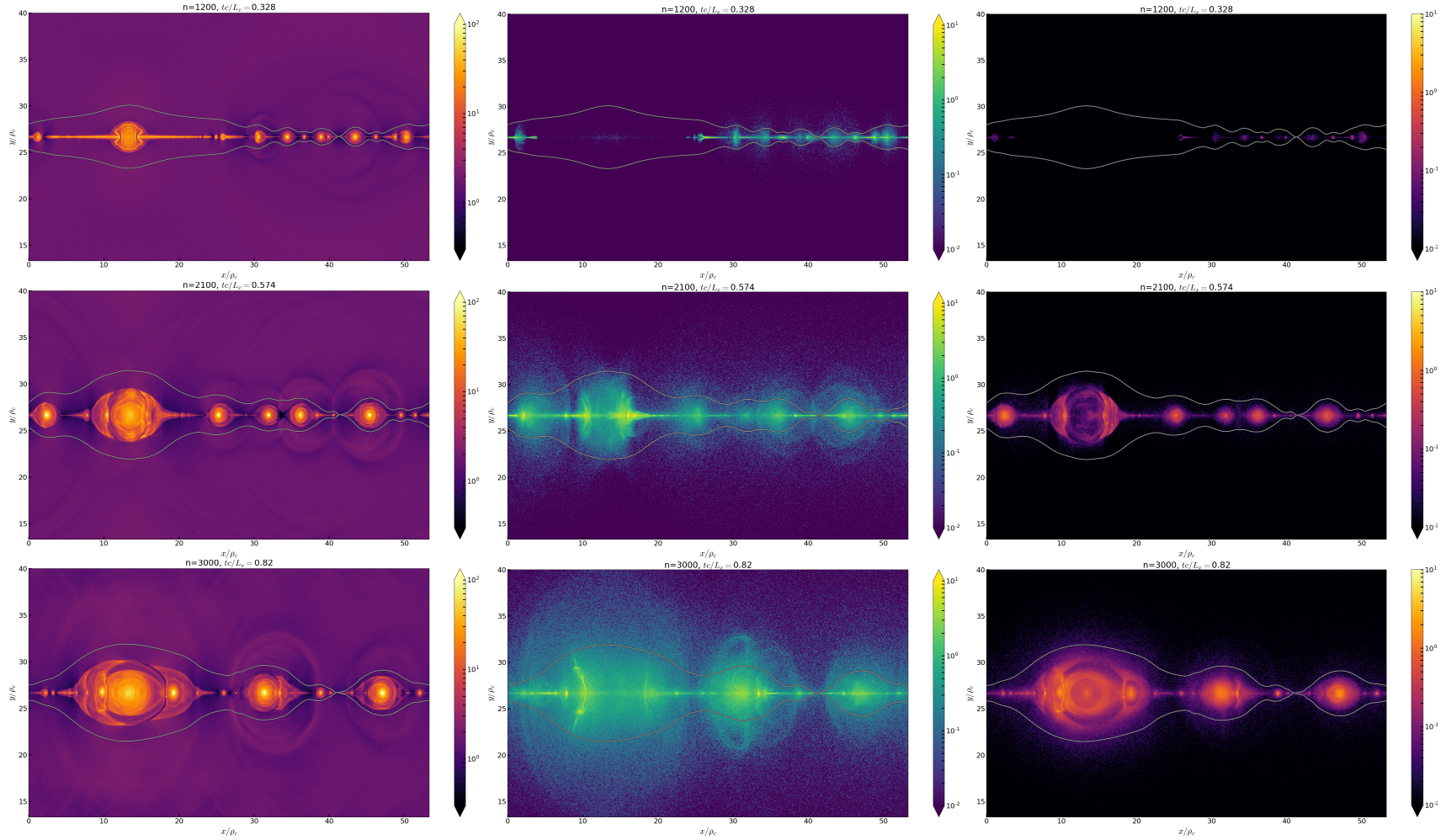

Figure 10.2: Left column: The number density of original background particles (i.e., not created by gamma-ray absorption) on a logarithmic colorscale spanning $[3 \times 10^{-2}, 10^2]$ in units of the background electron number density, $n_0/2$. Middle column: The energy density of above-threshold photons on a logarithmic colorscale spanning $[10^{-2}, 10^1]$ in units of $0.1B_0^2/8\pi$ – one-tenth of the upstream reconnecting magnetic field energy density. Right column: The number density of pairs produced during the simulation on a logarithmic colorscale spanning $[10^{-2}, 10^1]$ in units of the background electron number density, $n_0/2$.

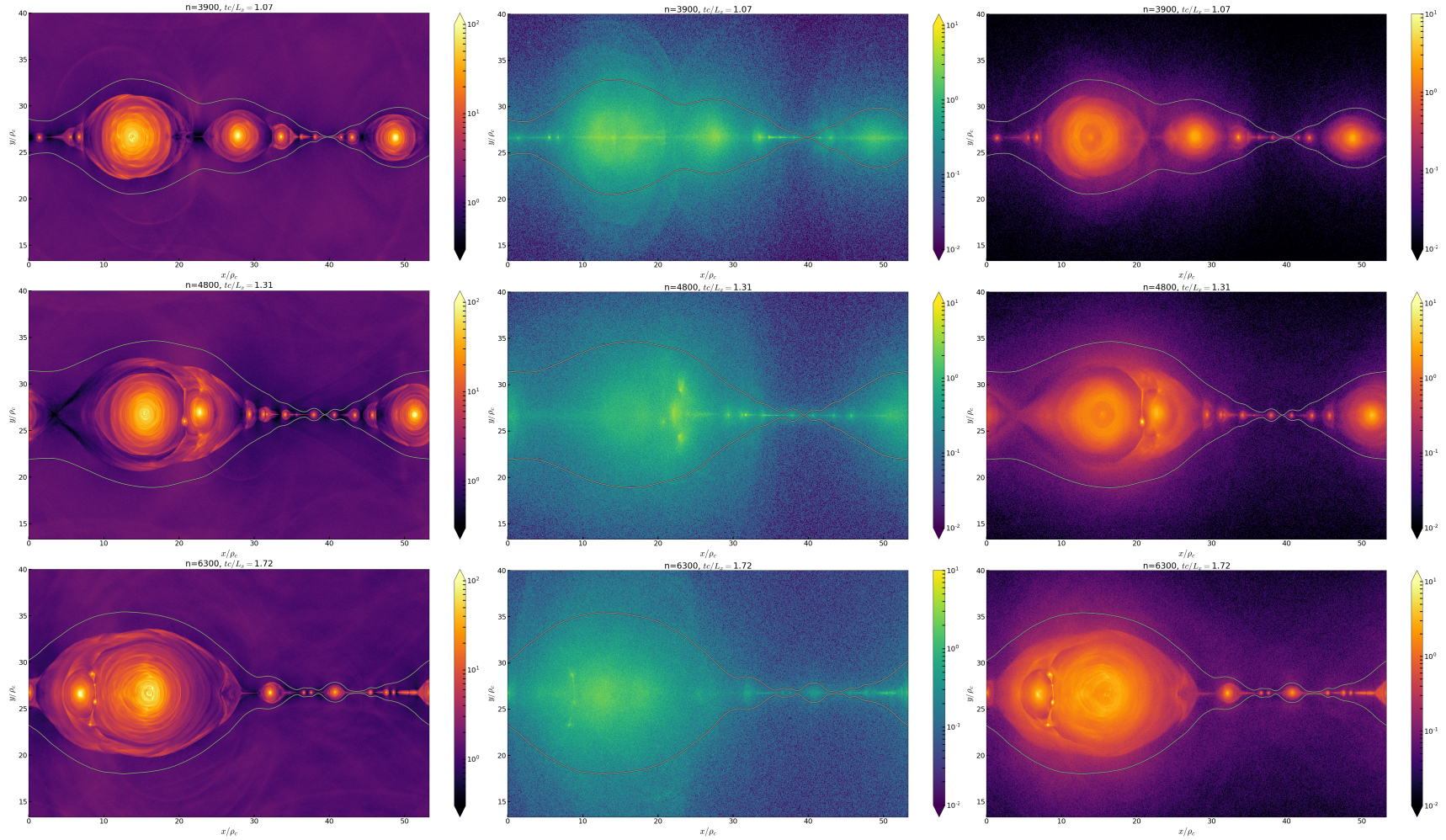

Figure 10.2: (Continued.) Each successive row corresponds to the simulation time $ct/L = 0.33, 0.57, 0.82, 1.07, 1.31$, and 1.72. For reference, in each panel, the separatrices are displayed as a contour (which corresponds to the value of the magnetic flux function A_z at the main reconnection X-point).

during reconnection. Focusing on the created particles in Fig. 10.2, one sees that, initially, pairs are produced inside the region of reconnected flux. However, later on, once a significant number of photons have crossed back into the upstream region, one begins to observe a diffuse population of tenuous pairs cropping up there (e.g., $ct/L = 1.07$ and onward in Fig. 10.2).

Beyond transferring some of the dissipated magnetic energy back into the upstream area as new pairs, the photons also influence the spatial distribution of created particles inside the layer. Comparing the spatial maps of original background particles to produced particles in Fig. 10.2, one can see that the background particles form clean concentric rings in the plasmoids where they are deposited, a morphological feature that can be traced back to how the particles end up in those locations. The background particles can only be brought into a plasmoid by following a reconnected field line – either after passing near an X-point and becoming remagnetized by the reconnected magnetic field, or simply by being carried along a field line that reconnects far away and begins to wrap its plasma around an accreting plasmoid. Thus, in the density maps of background particles, every particle that is bound to a field line when it reconnects remains on that field line. And so, the field lines, like the rings of a tree, tell the story of an aging plasmoid, bearing memory of the plasma conditions when they first reconnected and accreted onto the plasmoid. In contrast, the produced particles can be brought into a plasmoid through a completely separate channel: by being born there. Thus, while the created pairs also form concentric rings around plasmoids, these rings appear more fuzzy and diffuse: pair-production deposits pairs onto field lines inside plasmoids that were not there when that field line reconnected, washing out the clean record the field line would otherwise keep of the plasma it carried onto the plasmoid.

Next, I note that the colorscales on the plasma density maps of Fig. 10.2 are not the same: even the brightest regions on the spatial maps of created particles would appear only warm on the background particle maps. It is therefore somewhat clear from visual inspection of the density maps that the produced particles are tenuous relative to the particles that are initially present in the simulation. However, let me show this more explicitly. To do so, I present Fig. 10.3, which collapses the number densities of background and produced pairs in the simulation down to one

dimension. In that figure, the number densities are plotted as a function of magnetic flux (value of the z-component of the vector potential A_z; contours of A_z coincide, in 2D, with in-plane magnetic field lines). The procedure to compute these plots is very similar to averaging the number densities along slices of constant y (vertical direction in the spatial maps), and indeed reduces to this far from the reconnection layer where the in-plane field lines are evenly spaced and completely horizontal. However, close to the layer, the field lines begin to bend, and so averaging along A_z-contours instead of y-contours helps to isolate variations between field lines, which become very important as one crosses the reconnection separatrices. In particular, this diagnostic allows one to concisely see how pair production influences the properties of the plasma flowing across the separatrices and, hence, feeding reconnection – plasma that is very important to isolate in order to determine whether pair loading occurs as predicted by the model of chapter 7.

Using Fig. 10.3 one sees that, indeed, irrespective of the particular field line (value of A_z) that one considers, the produced pair plasma is always low in density compared to the plasma that is there from the beginning of the simulation. This is consistent with the basic picture of Klein-Nishina reconnection put forward in chapter 7.

However, the other basic qualitative prediction of the reconnection model of 7 is that, even if the produced pairs are tenuous, they may still be able to rival, or even dominate, the energy density of the plasma feeding reconnection. To determine whether this is consistent with the simulation, I supply Fig. 10.4. This is similar to Fig. 10.3 but plots, instead of the plasma number density as a function of A_z, the energy densities of background and produced particles versus A_z.

In Fig. 10.4, one sees that, after a short while – as early as $ct/L = 0.82$ – the energy density of produced pairs along the separatrices (at $A_z = A_{z,X}$ in the figure) begins to rival, and even surpasses, the energy density of background particles feeding reconnection. Thus, the diffuse cloud of pairs produced in the upstream region evident in Fig. 10.2, though low in number density, actually carries a significant amount of the material energy flowing into the reconnection layer: consistent with the model of Klein-Nishina reconnection in chapter 7.

Finally, I also present in Fig. 10.5, at the same time snapshots as the earlier figures in this

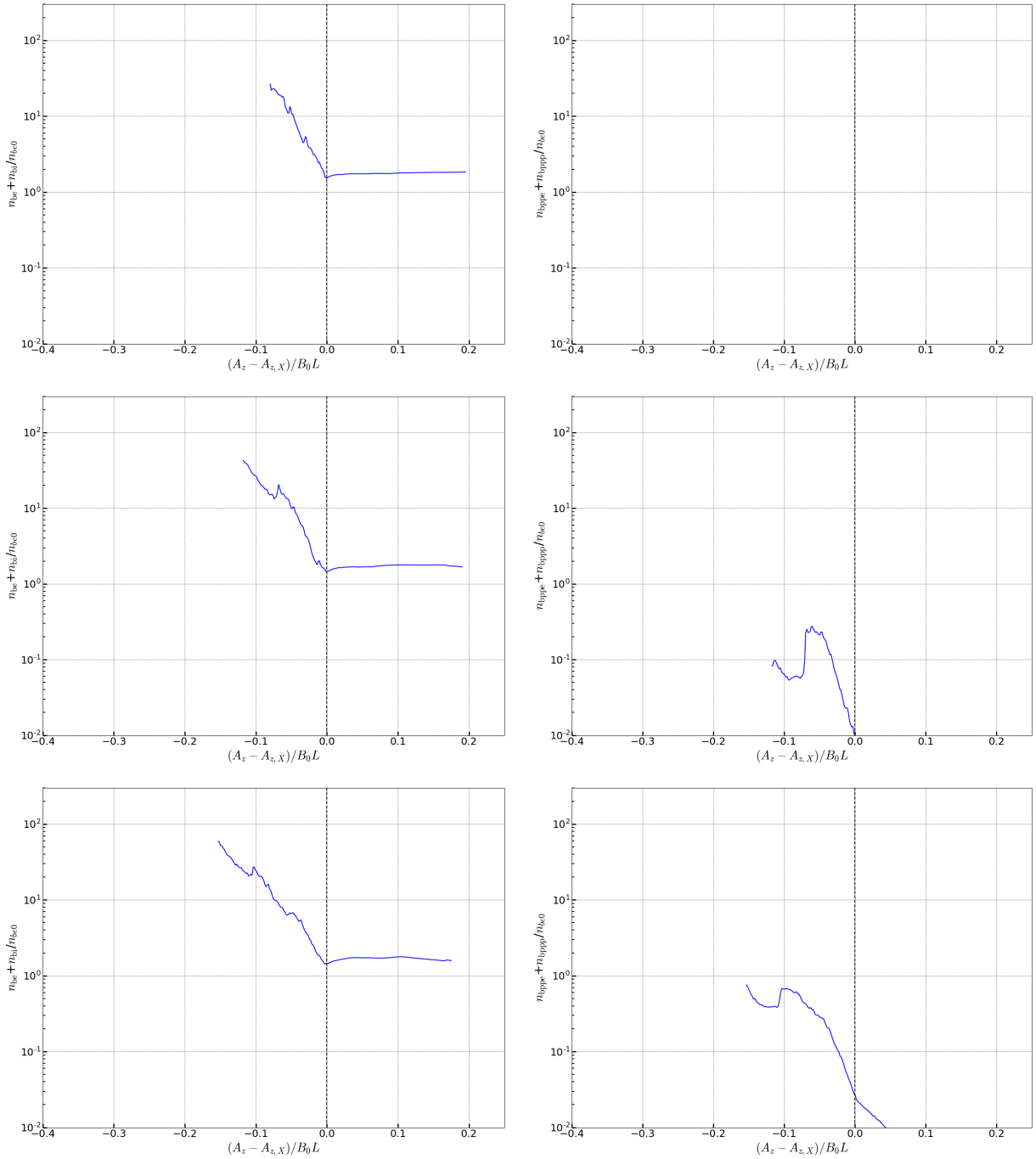

Figure 10.3: Left column: The average background number density (i.e., of originally present particles – those not created by gamma-ray absorption) along each in-plane magnetic field line (value of the magnetic flux function A_z). Right column: Same but for the number density of particles created by gamma-ray absorption. All vertical axes span $[10^{-2}, 3 \times 10^2]$ and are in units of background electron number density $n_0/2$. Each successive row corresponds to the simulation time $ct/L = 0.33, 0.57, 0.82, 1.07, 1.31,$ and 1.72.

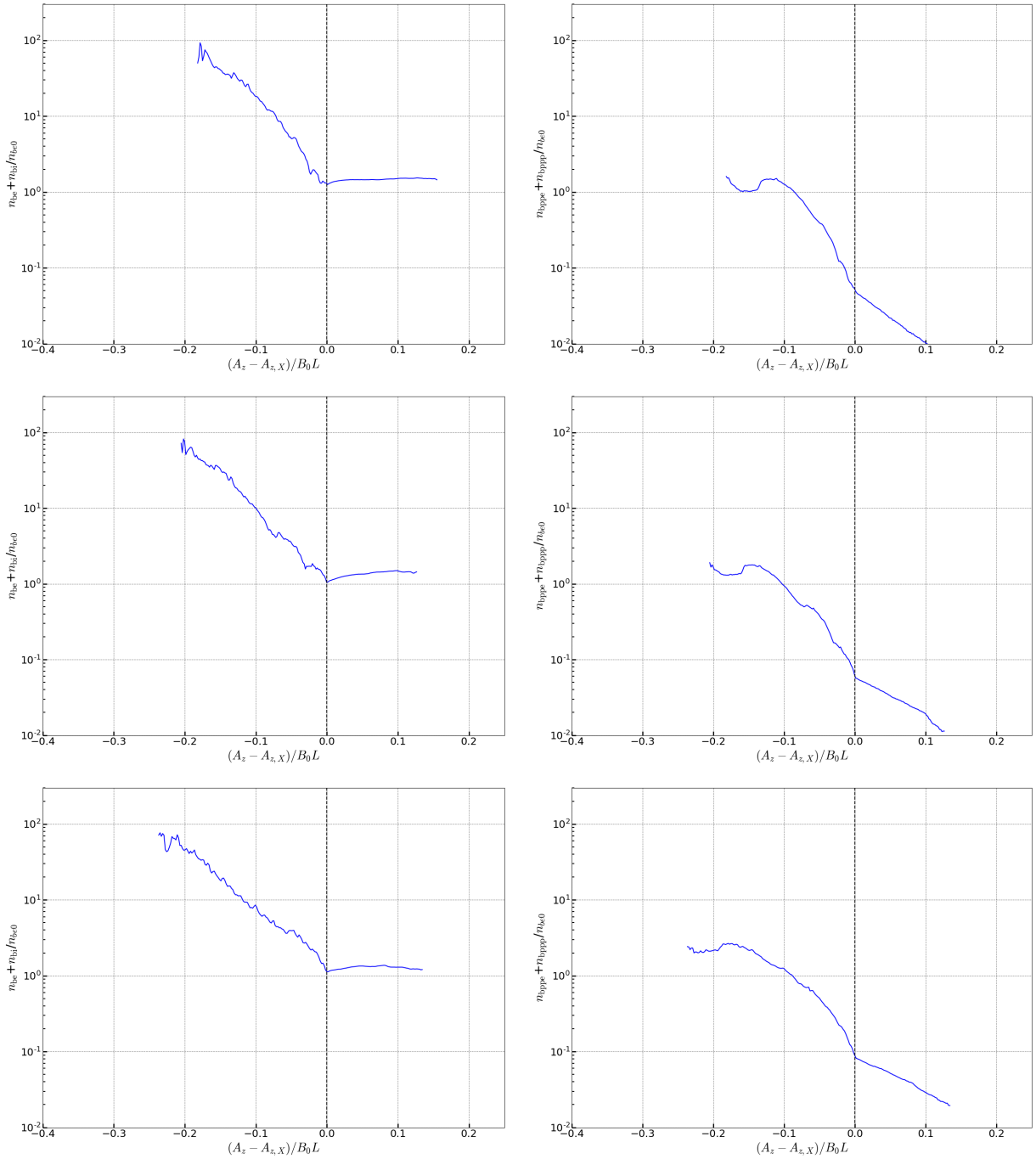

Figure 10.3: (Continued.) The horizontal axis in all plots is $(A_z - A_{z,X})/B_0L$ where $A_{z,X}$ is the value of the flux function at the main X-point and, thus, corresponds to the separatrices indicated in Fig. 10.2. Produced particles are everywhere tenuous compared to background particles along the same field line.

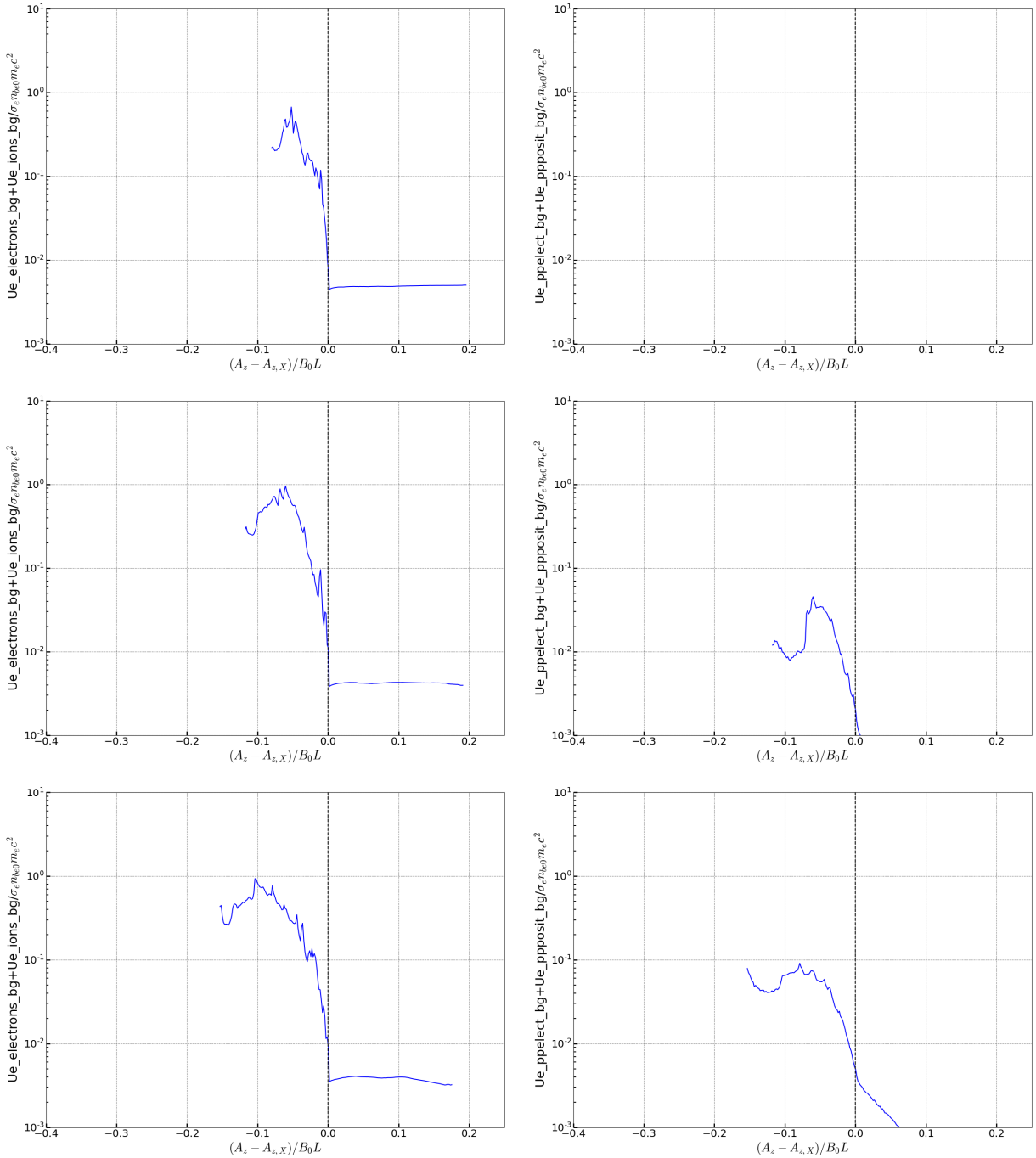

Figure 10.4: Left column: The average background energy density (i.e., of originally present particles – those not created by gamma-ray absorption) along each in-plane magnetic field line (value of the magnetic flux function A_z). Right column: Same but for the energy density of particles created by gamma-ray absorption. All vertical axes span $[10^{-3}, 10^{1}]$ and are in units of $\sigma_{\rm c,0} n_0 m_e c^2$ – the energy density obtained if the number density is n_0 and all particles achieve Lorentz factors $\sigma_{\rm c,0}$.

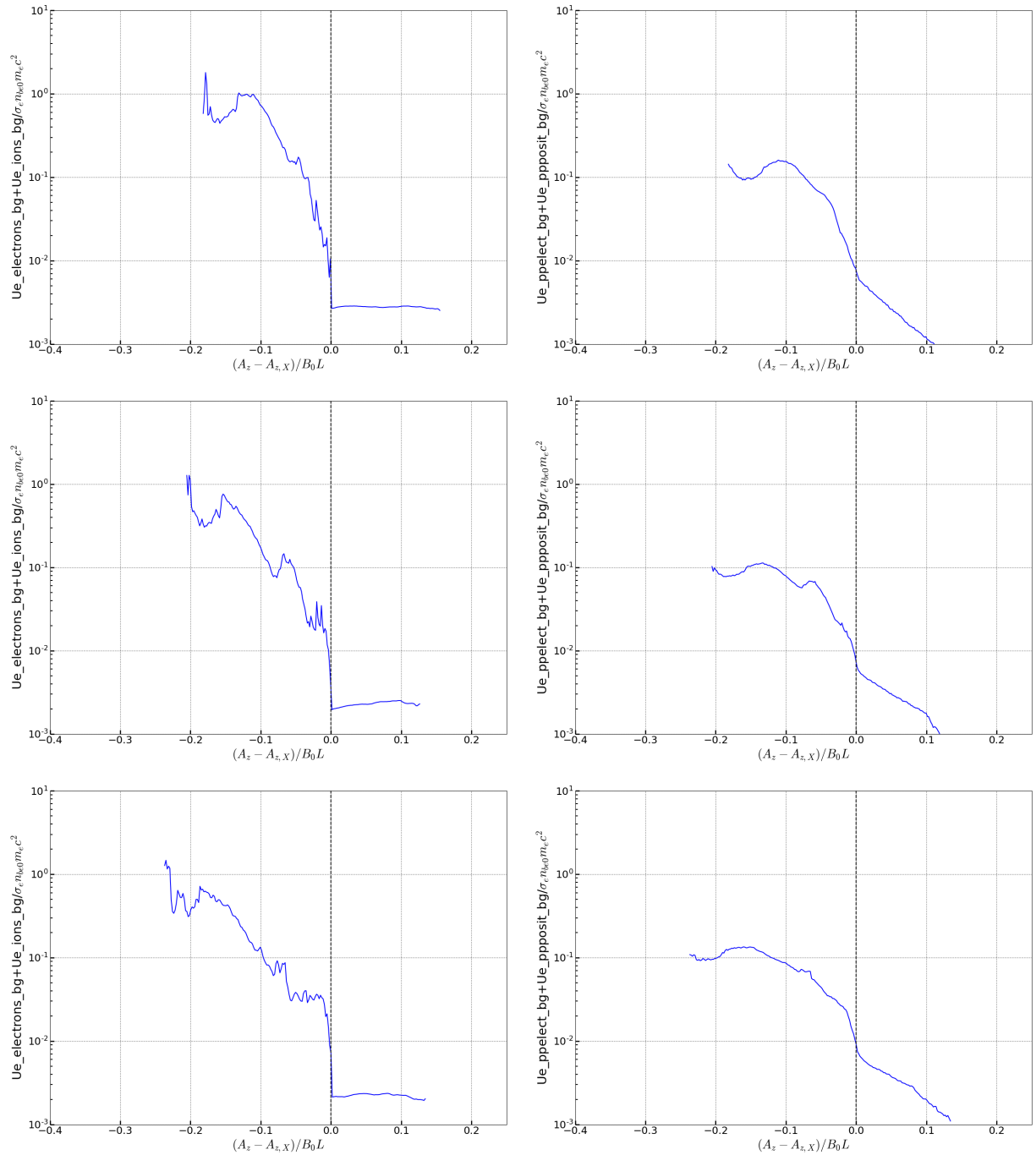

Figure 10.4: (Continued.) Each successive row corresponds to the simulation time $ct/L = 0.33, 0.57, 0.82, 1.07, 1.31,$ and 1.72. The horizontal axis in all plots is $(A_z - A_{z,X})/B_0 L$ – the same as in Fig. 10.3. Eventually (starting at around $ct/L = 0.82$), produced particles begin to dominate the energy density of material right outside the separatrices and give an order-unity contribution to the energy density of material just inside the separatrices [e.g., for $-0.1 < (A_z - A_{z,X})/B_0 L < 0$].

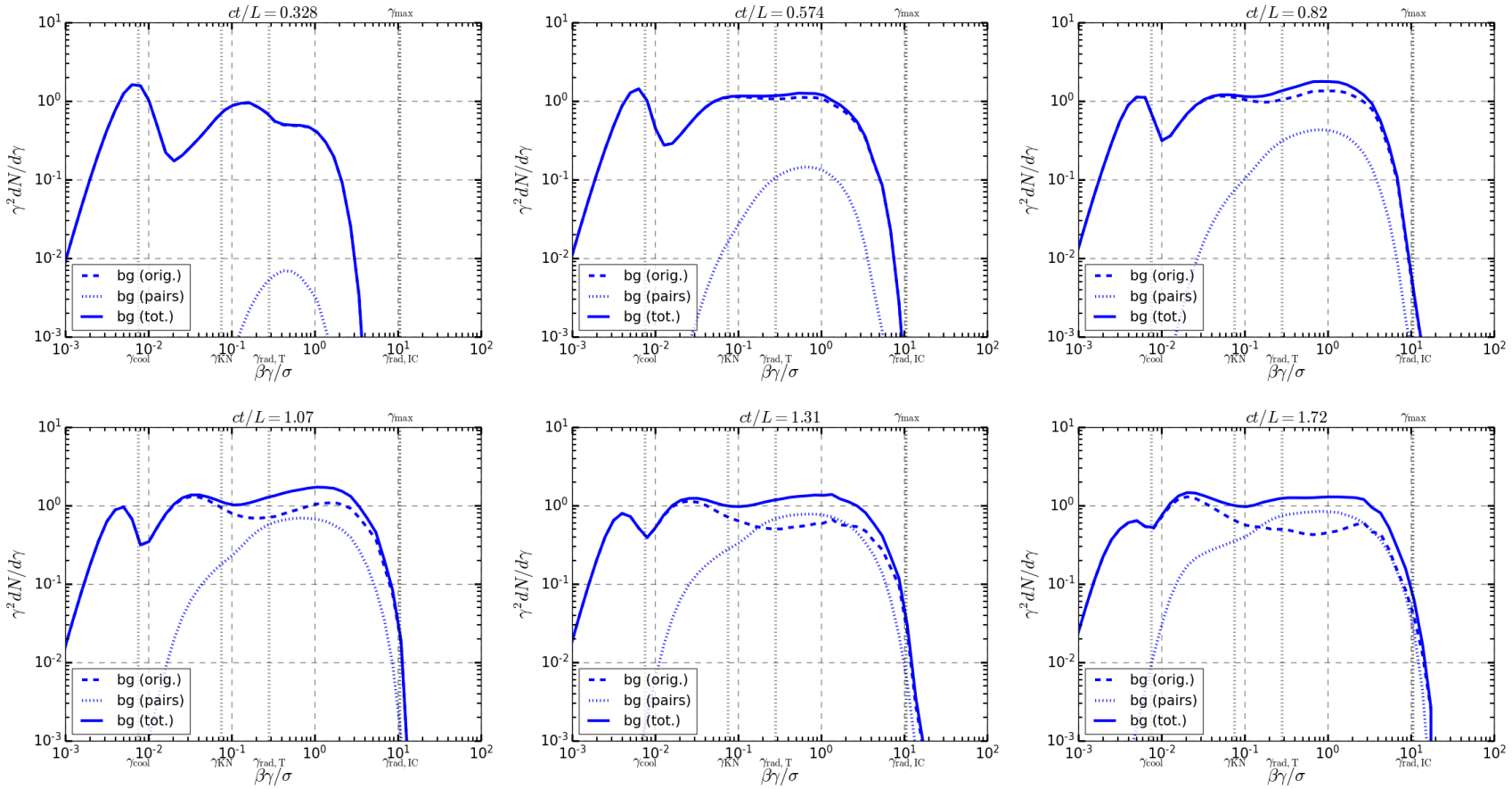

Figure 10.5: The total particle energy distribution for the times, $ct/L = 0.33, 0.57, 0.82, 1.07, 1.31$, and 1.72, shown in Figs. 10.2-10.4. Solid lines indicate the total distribution. Dashed lines show the contribution just from background particles that are initially present in the simulation. Dotted lines show the contribution from all produced pairs stemming from background particles. Thus, any first-generation pair – created from a photon emitted by a background particle – is tallied in the dotted line. Additionally, second-generation pairs – created from photons emitted from the first generation – third-generation pairs – created from photons emitted from the second generation – and all higher-generation pairs are counted in the dotted line. Vertical dotted black lines show values of important radiative energy scales. The scale ordering is as in case B of Fig. 6.6. Pairs contribute significantly to, and sometimes even dominate, the energy contained in the high-energy non-thermal tail of the distribution.

section, the particle energy distribution $\mathrm{d}N/\,\mathrm{d}\gamma$ realized in the simulation. The distribution is decomposed into contributions from the background (originally present) pairs and the produced pairs. It is compensated by a factor of γ^2 and therefore displays the energy carried per unit logarithmic interval in γ. Though I use the same symbol for this distribution $\mathrm{d}N/\,\mathrm{d}\gamma$ as in equations (7.5) and (7.17), I note that the distribution displayed in Fig. 10.5 contains all of the particles in the simulation – not just those confined to within the separatrices as in equations (7.5) and (7.17).

A remarkable feature of Fig. 10.5 is that, at some time snapshots (e.g., $ct/L = 1.31, 1.72$) the amount of energy carried by the high-γ non-thermal tail of the particle distribution is dominated by the produced pairs. At even later times (not shown), the newborn particles go on to dominate the *total matter energy* in the simulation box. This means that a relatively large spatial volume is occupied by the range of A_z values in Fig. 10.4 (typically outside the separatrices: i.e., when $A_z > A_{\mathrm{z,X}}$) where the created particles have higher energy density than the background particles. Evidently, these zones fill a large-enough fraction of the overall system that the created particles can dominate the *total material energy* in the system – not just the energy in small local patches.

Overall, this section is just the tip of the iceberg. Though the simulation presented here is consistent with an important, foundational, qualitative aspect of the Klein-Nishina reconnection model of chapter 7 – that the newborn pairs are fewer than the original background particles, but can dominate the upstream pressure – many more features of the model remain to be tested. It will be interesting, in the immediate future, to continue to analyze even just the one simulation presented here, further exploring reconnection and its observational implications in this fascinating regime.

Chapter 11

Conclusions

In this dissertation, I have explored radiative relativistic magnetic reconnection facilitated by inverse Compton (IC) cooling – an extreme regime with considerable astrophysical relevance, especially to jetted outflows from active galactic nuclei observed as blazars. This exploration has proceeded in two main phases.

First, in chapters 2-4, I studied radiative reconnection facilitated by optically thin Thomson IC cooling. I outlined the main reconnection parameters, phrased as energy scales, that control the regime of reconnection (chapter 2), and I explored a variety of these regimes using first-principles PIC simulations, focusing on the kinetic beaming phenomenon and how its observational signatures are mediated by strong radiative cooling (chapter 3). Then, to conclude the first half of this dissertation, I explored (chapter 4) the consequences of my numerical results for rapid TeV FSRQ blazar flares (e.g., Aleksić et al., 2011). I showed that a typical picture, in which the observed TeV photons are Comptonized from seed radiation shining onto the blazar jet from its circumnuclear environment (in particular, from the dusty torus), is strained if one posits magnetic reconnection as the energization mechanism for the emitting particles and naively extrapolates my Thomson IC results for kinetic beaming from chapter 3. Then, radiative cooling is not efficient enough for kinetically beamed high-energy particle bunches to radiate away their energy before they disperse and isotropize. Nevertheless, a kinetic beaming scenario nicely explains severe timing and energetics constraints imposed by observations (Nalewajko et al., 2012), and I supplied an alternative scenario, in which the emitting particles, energized by reconnection, Comptonize lower energy photons

produced from within the jet. However, one important caveat associated with these astrophysical conclusions is that, if the seed radiation field for the observed TeV emission is indeed a dusty torus, then the Comptonization does not occur in the Thomson IC regime. Thus, judging the feasibility of an external IC scenario requires self-consistently incorporating high-energy Klein-Nishina Compton losses.

Thus, in the second half of this dissertation (chapters 6-10), I considered more general Klein-Nishina IC radiative cooling. This includes the low-energy Thomson limit (studied in the earlier part of this work), in which particles lose their energy continuously, but it also accounts for the higher-energy regime, in which radiative losses become discrete (particles deliver an order-unity fraction of their energy to individual photons) and the scattered photons are above pair-production threshold with the ambient radiation.

After charting out the radiative parameter space governing reconnection with Klein-Nishina IC losses (chapter 6), I specialized to the new and unexplored regions of this space, outlining an analytic model for pair-regulated Klein-Nishina reconnection (chapter 7). The model hinges on a negative feedback loop that self-regulates the effective hot magnetization parameter of the plasma feeding reconnection. I then pointed out (chapter 8) several basic observational implications of the pair-regulated Klein-Nishina reconnection model, and I estimated the fundamental radiative reconnection parameters for two astrophysical systems – FSRQ blazar jets and the accretion disk coronae of black hole X-ray binaries in their high/soft states. These estimates showed that, if reconnection occurs in these systems as is often thought, then it is very likely to operate in a regime where high-energy discrete Compton losses and pair production significantly impact the collective plasma physical behavior.

Finally, in chapter 10, I presented preliminary results from a radiative PIC study of Klein-Nishina reconnection. There, I overviewed the modifications I made to the PIC code ZELTRON in order to model Klein-Nishina reconnection, I demonstrated the feasibility of computationally resolving all of the required scales, and I discussed early findings from a recent simulation I performed. Though preliminary, the results of that simulation are consistent with the picture of

pair-loading in Klein-Nishina reconnection advanced in chapter 7. They show that, as predicted, the number density of the particles created *in situ* during reconnection is small compared to that of the originally present background plasma. Nevertheless, the newborn pairs give a significant, and even dominant, contribution to the plasma material energy density, modifying the upstream hot magnetization from the value it would otherwise have in the presence of only the initial particles.

There is a rich and broad scope of future work that might extend and build upon the results described in this dissertation. In the immediate future, I will continue to run and analyze simulations of Klein-Nishina reconnection, testing from first principles several additional plasma physical and observational consequences of my analytic model described in chapters 7-8 (see chapter 9 for a discussion of specific follow-up questions that these simulations may address). More remotely, kinetic simulations mimicking global jet-like geometries (e.g., Alves et al., 2018; Sironi et al., 2021) and also incorporating Klein-Nishina and pair-production physics would determine whether and how this type of reconnection occurs in an explicitly global context. The same could be said for accretion-like numerical setups (e.g., Crinquand et al., 2021).

More broadly, I would like to frame this work, as I did in the Introduction (e.g., section 1.5), as a specific realization of several aspects of the exciting overarching field in which it is situated. As is commonly done in high-energy plasma astrophysics, I have, in this dissertation, explored a classic plasma process, magnetic reconnection, in the context of several new and extreme physical effects (in this case, special relativity, strong radiative cooling, and pair production) that are absolutely critical for understanding how this process works in astrophysical settings. Far from being passive participants on top of the traditional picture of reconnection gleaned from lower-energy environments, these physical effects fundamentally change how this process works and, in a very literal sense, looks – they decide the observational appearance of reconnection through a telescope. Living, as exemplified here, at the intersection of plasma physics, astrophysics, and high-energy physics, the study of high-energy plasma astrophysics presents both intellectually challenging new phenomena to understand and promises enhanced insight into some of the universe's most powerful and exciting objects. In the specific corner of this fascinating field occupied by work on radiative reconnection,

the novel physics treated in this dissertation (but see also Lyubarskii, 1996; Beloborodov, 2017; Werner et al., 2019; Schoeffler et al., 2019; Hakobyan et al., 2019; Sironi & Beloborodov, 2020) may help to pave the way to theories of increasingly realistic reconnection regimes with tremendous impact on our understanding of, and appreciation for, the high-energy universe.

Bibliography

Abdo A. A., et al., 2011, Science, 331, 739

Abdollahi S., et al., 2020, ApJS, 247, 33

Abeysekara A. U., et al., 2015, ApJ, 815, L22

Abramowski A., et al., 2012, ApJ, 746, 151

Acciari V. A., et al., 2009, Science, 325, 444

Ackermann M., et al., 2016, ApJ, 824, L20

Aharonian F. A., Atoyan A. M., 1981, Ap&SS, 79, 321

Aharonian F. A., Belyanin A. A., Derishev E. V., Kocharovsky V. V., Kocharovsky V. V., 2002, Phys. Rev. D, 66, 023005

Aharonian F., et al., 2007, ApJ, 664, L71

Ahnen M. L., et al., 2015, ApJ, 815, L23

Albert J., et al., 2007, ApJ, 669, 862

Aleksić J., et al., 2011, ApJ, 730, L8

Aleksić J., et al., 2014, Science, 346, 1080

Alves E. P., Zrake J., Fiuza F., 2018, Phys. Rev. Lett., 121, 245101

Astropy Collaboration et al., 2018, AJ, 156, 123

Bai X.-N., Caprioli D., Sironi L., Spitkovsky A., 2015, ApJ, 809, 55

Ball D., Sironi L., Özel F., 2018, ApJ, 862, 80

Begelman M. C., Sikora M., 1987, ApJ, 322, 650

Begelman M. C., Blandford R. D., Rees M. J., 1984, Reviews of Modern Physics, 56, 255

Begelman M. C., Fabian A. C., Rees M. J., 2008, MNRAS, 384, L19

Beloborodov A. M., 2017, ApJ, 850, 141

Beskin V. S., Istomin Y. N., Parev V. I., 1992, Soviet Ast., 36, 642

Bessho N., Bhattacharjee A., 2007, Physics of Plasmas, 14, 056503

Bessho N., Bhattacharjee A., 2012, ApJ, 750, 129

Birdsall C. K., Langdon A. B., of Physics (Great Britain) I., 2005, Plasma physics via computer simulation. Taylor & Francis, New York

Blackman E. G., Field G. B., 1994, Phys. Rev. Lett., 72, 494

Blandford R. D., Königl A., 1979, ApJ, 232, 34

Blandford R. D., Znajek R. L., 1977, MNRAS, 179, 433

Blandford R., Meier D., Readhead A., 2019, ARA&A, 57, 467

Bloom S. D., Marscher A. P., 1996, ApJ, 461, 657

Blumenthal G. R., Gould R. J., 1970, Reviews of Modern Physics, 42, 237

Boldyrev S., Loureiro N. F., 2020, Journal of Geophysical Research (Space Physics), 125, e28185

Böttcher M., Reimer A., Sweeney K., Prakash A., 2013, ApJ, 768, 54

Britto R. J., Bottacini E., Lott B., Razzaque S., Buson S., 2016, ApJ, 830, 162

Buehler R., et al., 2012, ApJ, 749, 26

Cerutti B., Philippov A. A., 2017, A&A, 607, A134

Cerutti B., Werner G., 2019, Zeltron: Explicit 3D relativistic electromagnetic Particle-In-Cell code (ascl:1911.012)

Cerutti B., Uzdensky D. A., Begelman M. C., 2012a, ApJ, 746, 148

Cerutti B., Werner G. R., Uzdensky D. A., Begelman M. C., 2012b, ApJ, 754, L33

Cerutti B., Werner G. R., Uzdensky D. A., Begelman M. C., 2013, ApJ, 770, 147

Cerutti B., Werner G. R., Uzdensky D. A., Begelman M. C., 2014a, Physics of Plasmas, 21, 056501

Cerutti B., Werner G. R., Uzdensky D. A., Begelman M. C., 2014b, ApJ, 782, 104

Cerutti B., Philippov A. A., Spitkovsky A., 2016, MNRAS, 457, 2401

Cerutti B., Philippov A. A., Dubus G., 2020, A&A, 642, A204

Chen A. Y., Yuan Y., Yang H., 2018, ApJ, 863, L31

Christie I. M., Petropoulou M., Sironi L., Giannios D., 2019, MNRAS, 482, 65

Christie I. M., Petropoulou M., Sironi L., Giannios D., 2020, MNRAS, 492, 549

Comisso L., Sironi L., 2018, Phys. Rev. Lett., 121, 255101

Comisso L., Sironi L., 2019, ApJ, 886, 122

Crinquand B., Cerutti B., Philippov A., Parfrey K., Dubus G., 2020, Phys. Rev. Lett., 124, 145101

Crinquand B., Cerutti B., Dubus G., Parfrey K., Philippov A., 2021, A&A, 650, A163

Dahlin J. T., Drake J. F., Swisdak M., 2014, Physics of Plasmas, 21, 092304

Dahlin J. T., Drake J. F., Swisdak M., 2015, Physics of Plasmas, 22, 100704

Dahlin J. T., Drake J. F., Swisdak M., 2017, Physics of Plasmas, 24, 092110

Davis S. W., Tchekhovskoy A., 2020, ARA&A, 58, 407

Dermer C. D., Schlickeiser R., 2002, ApJ, 575, 667

Di Matteo T., 1998, MNRAS, 299, L15

Drake J. F., Swisdak M., Che H., Shay M. A., 2006, Nature, 443, 553

Dubus G., Cerutti B., Henri G., 2010, MNRAS, 404, L55

Eardley D. M., Lightman A. P., 1976, Nature, 262, 196

Fermi LAT Collaboration et al., 2009, Science, 326, 1512

Ford A. L., Keenan B. D., Medvedev M. V., 2018, Phys. Rev. D, 98, 063016

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fossati G., Maraschi L., Celotti A., Comastri A., Ghisellini G., 1998, MNRAS, 299, 433

Galeev A. A., Rosner R., Vaiana G. S., 1979, ApJ, 229, 318

Ghisellini G., 2011, in Aharonian F. A., Hofmann W., Rieger F. M., eds, American Institute of Physics Conference Series Vol. 1381, 25th Texas Symposium on Relativistic AstroPhysics (Texas 2010). pp 180–198 (arXiv:1104.0006), doi:10.1063/1.3635832

Ghisellini G., Tavecchio F., Chiaberge M., 2005, A&A, 432, 401

Giannios D., 2013, MNRAS, 431, 355

Giannios D., Uzdensky D. A., 2019, MNRAS, 484, 1378

Giannios D., Uzdensky D. A., Begelman M. C., 2009, MNRAS, 395, L29

Giannios D., Uzdensky D. A., Begelman M. C., 2010, MNRAS, 402, 1649

Gierliński M., Zdziarski A. A., 2003, MNRAS, 343, L84

Gierliński M., Zdziarski A. A., Poutanen J., Coppi P. S., Ebisawa K., Johnson W. N., 1999, MNRAS, 309, 496

Goodman J., Uzdensky D., 2008, ApJ, 688, 555

Gould R. J., Schréder G. P., 1967, Physical Review, 155, 1404

Guo F., Li H., Daughton W., Liu Y.-H., 2014, Physical Review Letters, 113, 155005

Guo F., Liu Y.-H., Daughton W., Li H., 2015, ApJ, 806, 167

Guo F., et al., 2016, ApJ, 818, L9

Guo F., Li X., Daughton W., Kilian P., Li H., Liu Y.-H., Yan W., Ma D., 2019, ApJ, 879, L23

Guo F., Li X., Daughton W., Li H., Kilian P., Liu Y.-H., Zhang Q., Zhang H., 2020, arXiv e-prints, p. arXiv:2008.02743

Gurnett D. A., Bhattacharjee A., 2017, Introduction to Plasma Physics: With Space, Laboratory and Astrophysical Applications, 2 edn. Cambridge University Press, doi:10.1017/9781139226059

H. E. S. S. Collaboration et al., 2010, A&A, 520, A83

H. E. S. S. Collaboration et al., 2013, A&A, 554, A107

Hakobyan H., Philippov A., Spitkovsky A., 2019, ApJ, 877, 53

Hakobyan H., Petropoulou M., Spitkovsky A., Sironi L., 2021, ApJ, 912, 48

Hillas A. M., 1984, ARA&A, 22, 425

Jaroschek C. H., Hoshino M., 2009, Phys. Rev. Lett., 103, 075002

Jaroschek C. H., Treumann R. A., Lesch H., Scholer M., 2004, Physics of Plasmas, 11, 1151

Ji H., Daughton W., 2011, Physics of Plasmas, 18, 111207

Jones F. C., 1968, Physical Review, 167, 1159

Kagan D., Nakar E., Piran T., 2016, ApJ, 826, 221

Kagan D., Nakar E., Piran T., 2018, MNRAS, 476, 3902

Kirk J. G., Skjæraasen O., 2003, ApJ, 591, 366

Li T.-Y., Yorke J. A., 2004, Period Three Implies Chaos. Springer New York, New York, NY, pp 77–84, doi:10.1007/978-0-387-21830-4˙6, https://doi.org/10.1007/978-0-387-21830-4_6

Liang E. P. T., Price R. H., 1977, ApJ, 218, 247

Loureiro N. F., Boldyrev S., 2020, ApJ, 890, 55

Lyubarskii Y. E., 1996, A&A, 311, 172

Lyubarsky Y. E., 2005, MNRAS, 358, 113

Lyubarsky Y., Liverts M., 2008, ApJ, 682, 1436

Lyutikov M., Uzdensky D., 2003, ApJ, 589, 893

MAGIC Collaboration et al., 2008, Science, 320, 1752

MAGIC Collaboration et al., 2019, A&A, 623, A175

Madejski G. G., Sikora M., 2016, ARA&A, 54, 725

Malmrose M. P., Marscher A. P., Jorstad S. G., Nikutta R., Elitzur M., 2011, ApJ, 732, 116

Maraschi L., Ghisellini G., Celotti A., 1992, ApJ, 397, L5

McConnell M. L., et al., 2002, ApJ, 572, 984

McKinney J. C., Uzdensky D. A., 2012, MNRAS, 419, 573

Mehlhaff J. M., Werner G. R., Uzdensky D. A., Begelman M. C., 2020, MNRAS, 498, 799

Mehlhaff J. M., Werner G. R., Uzdensky D. A., Begelman M. C., 2021, arXiv e-prints, p. arXiv:2107.01297

Melia F., Königl A., 1989, ApJ, 340, 162

Melzani M., Walder R., Folini D., Winisdoerffer C., Favre J. M., 2014a, A&A, 570, A111

Melzani M., Walder R., Folini D., Winisdoerffer C., Favre J. M., 2014b, A&A, 570, A112

Moderski R., Sikora M., Coppi P. S., Aharonian F., 2005, MNRAS, 363, 954

Nalewajko K., 2016, Galaxies, 4, 28

Nalewajko K., Giannios D., Begelman M. C., Uzdensky D. A., Sikora M., 2011, MNRAS, 413, 333

Nalewajko K., Begelman M. C., Cerutti B., Uzdensky D. A., Sikora M., 2012, MNRAS, 425, 2519

Nalewajko K., Begelman M. C., Sikora M., 2014, ApJ, 789, 161

Nalewajko K., Yuan Y., Chruślińska M., 2018, Journal of Plasma Physics, 84, 755840301

National Academies of Sciences, Engineering, and Medicine 2021, Plasma Science: Enabling Technology, Sustainability, Security, and Exploration. The National Academies Press, Washington, DC, doi:10.17226/25802, https://www.nap.edu/catalog/25802/plasma-science-enabling-technology-sustainability-security-and-exploration

Nättilä J., Beloborodov A. M., 2020, arXiv e-prints, p. arXiv:2012.03043

Nenkova M., Sirocky M. M., Ivezić Ž., Elitzur M., 2008a, ApJ, 685, 147

Nenkova M., Sirocky M. M., Nikutta R., Ivezić Ž., Elitzur M., 2008b, ApJ, 685, 160

Ortuño-Macías J., Nalewajko K., 2020, MNRAS, 497, 1365

Parfrey K., Philippov A., Cerutti B., 2019, Phys. Rev. Lett., 122, 035101

Pétri J., 2012, MNRAS, 424, 2023

Petropoulou M., Sironi L., 2018, MNRAS, 481, 5687

Petropoulou M., Giannios D., Sironi L., 2016, MNRAS, 462, 3325

Philippov A. A., Spitkovsky A., 2018, ApJ, 855, 94

Philippov A. A., Spitkovsky A., Cerutti B., 2015, ApJ, 801, L19

Philippov A., Uzdensky D. A., Spitkovsky A., Cerutti B., 2019, ApJ, 876, L6

Phinney E. S., 1982, MNRAS, 198, 1109

Poutanen J., Krolik J. H., Ryde F., 1997, MNRAS, 292, L21

Pozdnyakov L. A., Sobol I. M., Syunyaev R. A., 1983, Astrophys. Space Phys. Res., 2, 189

Pushkarev A. B., Kovalev Y. Y., Lister M. L., Savolainen T., 2009, A&A, 507, L33

Rees M. J., 1966, Nature, 211, 468

Remillard R. A., McClintock J. E., 2006, ARA&A, 44, 49

Ripperda B., Bacchini F., Philippov A. A., 2020, ApJ, 900, 100

Rybicki G. B., Lightman A. P., 1979, Radiative Processes in Astrophysics. Wiley, New York

Schoeffler K. M., Grismayer T., Uzdensky D., Fonseca R. A., Silva L. O., 2019, ApJ, 870, 49

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337

Shapiro S. L., Lightman A. P., Eardley D. M., 1976, ApJ, 204, 187

Sikora M., Begelman M. C., Rees M. J., 1994, ApJ, 421, 153

Sikora M., Stawarz Ł., Moderski R., Nalewajko K., Madejski G. M., 2009, ApJ, 704, 38

Sikora M., Rutkowski M., Begelman M. C., 2016, MNRAS, 457, 1352

Sironi L., Beloborodov A. M., 2020, ApJ, 899, 52

Sironi L., Spitkovsky A., 2009, ApJ, 698, 1523

Sironi L., Spitkovsky A., 2014, ApJ, 783, L21

Sironi L., Petropoulou M., Giannios D., 2015, MNRAS, 450, 183

Sironi L., Giannios D., Petropoulou M., 2016, MNRAS, 462, 48

Sironi L., Rowan M. E., Narayan R., 2021, ApJ, 907, L44

Sitarek J., et al., 2015, in 34th International Cosmic Ray Conference (ICRC2015). p. 825 (`arXiv:1508.04580`)

Sobacchi E., Nättilä J., Sironi L., 2021, MNRAS, 503, 688

Szostek A., Zdziarski A. A., McCollough M. L., 2008, MNRAS, 388, 1001

Tamburini M., Pegoraro F., Di Piazza A., Keitel C. H., Macchi A., 2010, New Journal of Physics, 12, 123005

Tanaka Y. T., et al., 2011, ApJ, 733, 19

Tavani M., et al., 2009, Nature, 462, 620

Tavani M., et al., 2011, Science, 331, 736

Tavecchio F., 2017, AIP Conf. Proc., 1792, 020007

Tavecchio F., Ghisellini G., 2008, MNRAS, 386, 945

Tavecchio F., Ghisellini G., 2016, MNRAS, 456, 2374

Tavecchio F., Becerra-Gonzalez J., Ghisellini G., Stamerra A., Bonnoli G., Foschini L., Maraschi L., 2011, A&A, 534, A86

Uzdensky D. A., 2011, Space Sci. Rev., 160, 45

Uzdensky D. A., 2016, Radiative Magnetic Reconnection in Astrophysics. Springer International Publishing, Cham, p. 473 (arXiv:1510.05397), doi:10.1007/978-3-319-26432-5˙12

Uzdensky D. A., 2020, arXiv e-prints, p. arXiv:2007.09533

Uzdensky D. A., Goodman J., 2008, ApJ, 682, 608

Uzdensky D. A., McKinney J. C., 2011, Physics of Plasmas, 18, 042105

Uzdensky D. A., Spitkovsky A., 2014, ApJ, 780, 3

Uzdensky D. A., Loureiro N. F., Schekochihin A. A., 2010, Physical Review Letters, 105, 235002

Uzdensky D. A., Cerutti B., Begelman M. C., 2011, ApJ, 737, L40

Uzdensky D., et al., 2019, BAAS, 51, 362

Werner G. R., Uzdensky D. A., 2017, ApJ, 843, L27

Werner G. R., Uzdensky D. A., Cerutti B., Nalewajko K., Begelman M. C., 2016, ApJ, 816, L8

Werner G. R., Uzdensky D. A., Begelman M. C., Cerutti B., Nalewajko K., 2018, MNRAS, 473, 4840

Werner G. R., Philippov A. A., Uzdensky D. A., 2019, MNRAS, 482, L60

Wong G. N., Ryan B. R., Gammie C. F., 2021, ApJ, 907, 73

Yuan Y., Nalewajko K., Zrake J., East W. E., Blandford R. D., 2016, ApJ, 828, 92

Zdziarski A. A., 1989, ApJ, 342, 1108

Zdziarski A. A., Gierliński M., 2004, Progress of Theoretical Physics Supplement, 155, 99

Zdziarski A. A., Malyshev D., Chernyakova M., Pooley G. G., 2017, MNRAS, 471, 3657

Zdziarski A. A., et al., 2018, MNRAS, 479, 4399

Zenitani S., Hoshino M., 2001, ApJ, 562, L63

Zenitani S., Hoshino M., 2007, ApJ, 670, 702

Zenitani S., Hoshino M., 2008, ApJ, 677, 530

Zhdankin V., Uzdensky D. A., Perez J. C., Boldyrev S., 2013, ApJ, 771, 124

Zhdankin V., Uzdensky D. A., Werner G. R., Begelman M. C., 2020, MNRAS, 493, 603

Zhdankin V., Uzdensky D. A., Kunz M. W., 2021, ApJ, 908, 71

Zweibel E. G., Yamada M., 2009, ARA&A, 47, 291

Appendix A

The Klein-Nishina Scattering Power and Rate

Here, I sketch the derivations of the functions $f_{\rm KN}(q)$ and $g_{\rm KN}(q)$ as defined in equations (6.5) and (6.9). Jones (1968) and Blumenthal & Gould (1970) report the Klein-Nishina scattering kernel

$$\begin{aligned}\frac{\mathrm{d}N}{\mathrm{d}t\,\mathrm{d}r}\,\mathrm{d}r &= c\sigma_{\rm T}\frac{U_{\rm ph}}{\epsilon_{\rm ph}} \\ &\quad\times\frac{3}{(1+qr)^2}\left[2r\ln r+(1+2r)(1-r)+\frac{1}{2}\frac{(qr)^2}{1+qr}(1-r)\right]\mathrm{d}r \\ &\equiv c\sigma_{\rm T}\frac{U_{\rm ph}}{\epsilon_{\rm ph}}K(r,q)\,\mathrm{d}r\,. \end{aligned} \tag{A.1}$$

This is the number of photons scattered per unit time by a particle with Klein-Nishina parameter $q=\gamma/\gamma_{\rm KN}=4\gamma\epsilon_{\rm ph}/m_ec^2$ to final photon energies between r and $r+\mathrm{d}r$. The parameter r is not actually equal to the scattered photon energy ϵ_1 but is defined in terms of it through

$$r\equiv\frac{\epsilon_1/\gamma m_ec^2}{q\left(1-\epsilon_1/\gamma m_ec^2\right)}\,. \tag{A.2}$$

Equivalently,

$$\epsilon_1=\gamma m_ec^2\frac{qr}{1+qr}\,. \tag{A.3}$$

It is convenient to take integrals over r as a proxy for ϵ_1 because, as ϵ_1 spans its kinematically allowed range $[\epsilon_{\rm ph},\gamma m_ec^2q/(1+q)]$, r runs approximately from 0 to 1 (Jones, 1968; Blumenthal & Gould, 1970).

For completeness, I mention the validity conditions for the kernel (A.1): $\gamma m_ec^2/\epsilon_{\rm ph}=4\gamma_{\rm KN}\gamma\gg 1$ and $\gamma\gg 1$ (Jones, 1968). The first condition is always satisfied for $\gamma_{\rm KN}\gg 1$, which is true of

every astrophysical system to which I apply my results (and also of many others). The relativistic requirement $\gamma \gg 1$ is then taken care of because the general Klein-Nishina expressions are only needed for $\gamma \gg \gamma_{\rm KN} \gg 1$. At non-relativistic energies $\gamma \simeq 1$, the Thomson prefactors in equations (6.4) and (6.8) are all that remain [$f_{\rm KN}(q \ll 1) \to 1$ and $g_{\rm KN}(q \ll 1) \to 1$], and these are non-relativistically correct. Thus, for $\gamma_{\rm KN} \gg 1$, equations (6.4) and (6.8) are correct even in the non-relativistic case.

To obtain the rates at which a scattering particle loses energy [equation (6.4)] and encounters soft seed photons [equation (6.8)] requires integrating over ϵ_1 – or, equivalently, over r – while keeping γ, $\epsilon_{\rm ph}$, and, hence, $q = \gamma/\gamma_{\rm KN}$ fixed. One obtains

$$\begin{aligned}\begin{bmatrix} R_{\rm IC}(\gamma) \\ P_{\rm IC}(\gamma) \end{bmatrix} &= c\sigma_{\rm T}\frac{U_{\rm ph}}{\epsilon_{\rm ph}}\int_0^1 {\rm d}r\, K(r,q) \begin{bmatrix} 1 \\ \epsilon_1(r) \end{bmatrix} \\ &= \begin{bmatrix} c\sigma_{\rm T}U_{\rm ph}/\epsilon_{\rm ph} \\ (4/3)c\sigma_{\rm T}\gamma^2 U_{\rm ph} \end{bmatrix} \int_0^1 {\rm d}r\, K(r,q) \begin{bmatrix} 1 \\ 3r/(1+qr) \end{bmatrix} . \end{aligned} \tag{A.4}$$

Then, using $\beta^2 = 1 - 1/\gamma^2 \simeq 1$ to write $P_{\rm T}(\gamma) \simeq (4/3)c\sigma_{\rm T}\gamma^2 U_{\rm ph}$, one can read off $g_{\rm KN}(q)$ and $f_{\rm KN}(q)$ from equation (A.4):

$$g_{\rm KN}(q) = \int_0^1 {\rm d}r\, K(r,q) \tag{A.5}$$

and

$$f_{\rm KN}(q) = 3\int_0^1 {\rm d}r\, K(r,q)\frac{r}{1+qr}\,. \tag{A.6}$$

The integrals can be evaluated to yield equations (6.5) and (6.9). Should the reader wish to verify by explicit computation, I find that the identity

$${\rm Li}_2\left(\frac{1}{1+q}\right) = {\rm Li}_2(-q) + \frac{\pi^2}{6} + \frac{1}{2}\log^2(q) - \frac{1}{2}\log^2\left(\frac{q}{1+q}\right) \tag{A.7}$$

is useful.

Appendix B

The Energy Retention Factor $f_{\rm nocool}$

Recall that in section 7.2.1, I left ξ as a free parameter. This stemmed chiefly from my uncertainty regarding the fraction of energy $f_{\rm nocool}$ retained by the fresh pair plasma as it travels toward the reconnection layer. In order to demonstrate the effect of this parameter without knowing it precisely, I entertained a class of models where ξ (specifically $f_{\rm nocool}$) was constant: independent of NTPA in the layer [i.e., of $p(\sigma_{\rm h})$ and z]. However, in principle, one can explicitly compute the distributions $N_{\gamma\gamma}^{(n)}(\gamma)$ of newborn pairs (by straightforwardly extending the analysis of section 7.2.2), and, from them, calculate the energy carried by fresh plasma reaching the layer. This can then be compared to the energy *deposited* into the upstream region to yield $f_{\rm nocool}$.

In this section, I adopt this strategy to calculate $\xi = f_{\rm nocool}$. I find that, like $\mathcal{F}$, ξ is a function of $p(\sigma_{\rm h})$ and z: $\xi = \Xi[p(\sigma_{\rm h}), z]$. Furthermore, the basic intuition gleaned in section 7.2.1 remains intact: extremely efficient energy delivery to the layer, $\xi \sim 1$, may still push the system into a 2-state swing cycle, and radiative feedback on NTPA still makes the system more prone to these oscillations. Through the present analysis, one merely gains a more precise notion of the values of physical parameters – particularly the cutoff z in the distribution of layer particles – required to initiate swing cycles versus those that cause rapid progression toward the fixed point $\sigma_{\rm h}$.

My first step is to write down a quantitative expression for $f_{\rm nocool}$. In order to do this, I define several auxiliary quantities: the nth-generation steady-state particle count

$$N_{\gamma\gamma}^{(n)} \equiv \int N_{\gamma\gamma}^{(n)}(\gamma)\,{\rm d}\gamma\,, \tag{B.1}$$

which I distinguish from $N_{\gamma\gamma}^{(n)}(\gamma)$ by omitting the functional argument; the nth-generation particle injection rate

$$Q_{\gamma\gamma}^{(n)} \equiv \int Q_{\gamma\gamma}^{(n)}(\gamma)\,\mathrm{d}\gamma\,, \tag{B.2}$$

also distinguished from $Q_{\gamma\gamma}^{(n)}(\gamma)$ by argument omission; the average particle energy *injected* into the nth generation

$$\bar{\gamma}_Q^{(n)} \equiv \frac{\int \gamma Q_{\gamma\gamma}^{(n)}(\gamma)\,\mathrm{d}\gamma}{Q_{\gamma\gamma}^{(n)}}\,; \tag{B.3}$$

and the average particle energy *carried* by the nth generation in its steady state

$$\bar{\gamma}_N^{(n)} \equiv \frac{\int \gamma N_{\gamma\gamma}^{(n)}(\gamma)\,\mathrm{d}\gamma}{N_{\gamma\gamma}^{(n)}}\,. \tag{B.4}$$

In terms of these quantities, one has

$$\begin{aligned} f_{\mathrm{nocool}} &\equiv \frac{\text{newborn pair energy carried into layer}}{\text{energy injected upstream}} \\ &= \frac{\sum_{n=1}^{N} N_{\gamma\gamma}^{(n)} \bar{\gamma}_N^{(n)}}{\sum_{n=1}^{N} N_{\gamma\gamma}^{(n)} \bar{\gamma}_Q^{(n)}} \\ &= \frac{\sum_{n=1}^{N} N_{\gamma\gamma}^{(n)} \bar{\gamma}_Q^{(n)} f_{\mathrm{nocool}}^{(n)}}{\sum_{n=1}^{N} N_{\gamma\gamma}^{(n)} \bar{\gamma}_Q^{(n)}}\,, \end{aligned} \tag{B.5}$$

where

$$f_{\mathrm{nocool}}^{(n)} \equiv \frac{N_{\gamma\gamma}^{(n)} \bar{\gamma}_N^{(n)}}{N_{\gamma\gamma}^{(n)} \bar{\gamma}_Q^{(n)}} = \frac{\bar{\gamma}_N^{(n)}}{\bar{\gamma}_Q^{(n)}} \tag{B.6}$$

is the average fractional energy retained by the nth generation.

Before I calculate f_{nocool} in detail using equation (B.5), let me briefly argue why, as claimed in section 7.2.1, f_{nocool} should be physically confined to the range $[3/400, 1]$. The upper bound is trivial: f_{nocool} cannot exceed one by definition. Let me now show how the lower bound arises.

In the case where f_{nocool} is as small as possible, all particles born into the upstream region are quite energetic, with $\gamma_{\mathrm{KN}} \ll \bar{\gamma}_Q^{(n)} \leq \tilde{\gamma}_{\mathrm{cool},2}$, but they cool quickly – until their cooling times match their readvection time t_{ra}, and hence until their Lorentz factors equal $\tilde{\gamma}_{\mathrm{cool},1}$ [equation (7.23)]. However, when these particles' energies exceed γ_{pp}, none of their emitted photons escape the

system. Rather, the power $Q^{(1)}_{\gamma\gamma}\bar{\gamma}^{(1)}_{Q}m_ec^2$ injected into the upstream region losslessly converts to secondary pairs, and therefore matches the steady-state energy flux $Q'\gamma_{\rm pp}m_ec^2$ into the particle energy bin $\gamma < \gamma_{\rm pp} < \gamma + {\rm d}\gamma$. Here, Q' is the (generationally summed) particle flux into the same bin. Only once particles cool past $\gamma_{\rm pp}$ does their emission leak out of the system. Thus, if the injected particles have high initial energies but all cool down to $\tilde{\gamma}_{\rm cool,1}$, $f_{\rm nocool}$ reaches the (lowest possible) value

$$\begin{aligned} f_{\rm nocool,min} &\equiv 1 - \frac{\text{power radiated by upstream particles}}{\text{power injected upstream}} \\ &= 1 - \frac{Q'(\gamma_{\rm pp} - \tilde{\gamma}_{\rm cool,1})}{Q^{(1)}_{\gamma\gamma}\bar{\gamma}^{(n)}_{Q}} \\ &= 1 - \frac{\gamma_{\rm pp} - \tilde{\gamma}_{\rm cool,1}}{\gamma_{\rm pp}} \\ &= \frac{\tilde{\gamma}_{\rm cool,1}}{\gamma_{\rm pp}} = \frac{1}{8}\frac{3}{50} = \frac{3}{400}\,. \end{aligned} \tag{B.7}$$

The reason that (as I show quantitatively in this section) $f_{\rm nocool,min} = 3/400$ is an overly pessimistic estimate for $f_{\rm nocool}$ is that, in reality, upstream particles are constantly being replenished at high energies even as they rapidly cool toward $\tilde{\gamma}_{\rm cool,1}$. This results in a pronounced high-energy tail of hot upstream particles entering the reconnection layer, greatly enhancing $f_{\rm nocool}$, even, in some cases, to order unity.

B.1 The steady-state pair distributions $N^{(n)}_{\gamma\gamma}(\gamma)$

Evaluating equation (B.5) for $f_{\rm nocool}$ explicitly requires the steady-state pair distributions $N^{(n)}_{\gamma\gamma}(\gamma)$. I now retrieve these distributions by extending the analysis of section 7.2.2. Rearranging equation (7.30), one may write

$$N^{(n)}_{\gamma\gamma}(\gamma) = \frac{1}{8}t_{\rm cool,IC}(\gamma)Q^{(n+1)}_{\gamma\gamma}(\gamma/4)\,. \tag{B.8}$$

Plugging in $t_{\rm cool,IC}(\gamma) \propto [\gamma f_{\rm KN}(\gamma/\gamma_{\rm KN})]^{-1}$ and using the approximate expression $f_{\rm KN}(q) \simeq (1+q)^{-1.5}$ from equation (6.7), one sees that $N^{(n)}_{\gamma\gamma}(\gamma) \propto \gamma^{-\Gamma+0.5}$ provided $Q^{(n)}_{\gamma\gamma} \propto \gamma^{-\Gamma}$ [i.e., as in (7.21)] and $\gamma \gg \gamma_{\rm KN}$. The latter condition is satisfied because the deep Klein-Nishina regime is assumed in writing equations (7.27) and (7.28), from which (B.8) follows.

However, this is not the full story for $N^{(n)}_{\gamma\gamma}(\gamma)$. Even though $Q^{(n)}_{\gamma\gamma}(\gamma)$ is zero when $\gamma < \gamma_{\rm min} \equiv 2\gamma_{\rm KN}$, which comes from the fact that no pairs are injected at energies below $\gamma_{\rm min}$, $N^{(n)}_{\gamma\gamma}(\gamma)$ is *not* zero at these low energies. Instead, particles in every generation cool continuously – in the Thomson regime – once they reach energies $\sim \gamma_{\rm KN} = \gamma_{\rm min}/2$, populating a low-energy component of each distribution $N^{(n)}_{\gamma\gamma}(\gamma)$.

To model this situation, I assume that, in each generation, a constant flux of particles leaks from above to below $\gamma_{\rm KN}$ in energy space [through the advective term in equation (7.27), which remains nonzero in the Thomson regime even though $Q^{(n)}_{\gamma\gamma}(\gamma)$ vanishes]. Once there, the rate of change of each particle's energy is approximately Thomson: $-\dot{\gamma} \simeq -\dot{\gamma}_{\rm T} \equiv \gamma/t_{\rm cool,T}(\gamma)$. To determine $N^{(n)}_{\gamma\gamma}(\gamma)$ in the Thomson limit $\gamma < \gamma_{\rm KN}$ then requires solving a simplified version of equation (7.27),

$$\frac{\partial}{\partial t}N^{(n)}_{\gamma\gamma}(\gamma,t) + \frac{\partial}{\partial\gamma}\left(\dot{\gamma}_{\rm T}N^{(n)}_{\gamma\gamma}(\gamma,t)\right) = Q^{(n)}_{\gamma\gamma}\delta(\gamma-\gamma_{\rm KN})\,, \tag{B.9}$$

where the source term $Q^{(n)}_{\gamma\gamma}\delta(\gamma-\gamma_{\rm KN}) = \delta(\gamma-\gamma_{\rm KN})\int Q^{(n)}_{\gamma\gamma}(\gamma)\,{\rm d}\gamma$ gives the flux into $\gamma_{\rm KN}$ of particles from higher energies.[1] The solution to (B.9) can be obtained exactly, and the steps are detailed in Appendix E. The result is a low-energy distribution of particles arriving at the layer in each generation of

$$N^{(n)}_{\gamma\gamma}(\gamma \lesssim \gamma_{\rm KN}) = Q^{(n)}_{\gamma\gamma}\gamma_{\rm cool}\frac{L}{c}\begin{cases}\gamma^{-2} & \gamma \geq \tilde{\gamma}_{\rm cool,1}\\ 0 & \gamma < \tilde{\gamma}_{\rm cool,1}\end{cases}. \tag{B.10}$$

Equivalently, one can use $\gamma_{\rm cool} = (10/\tau_{\gamma\gamma})\tilde{\gamma}_{\rm cool,1} = (10/\tau_{\gamma\gamma})(3/50)\gamma_{\rm KN}$ to express this as

$$N^{(n)}_{\gamma\gamma}(\gamma \lesssim \gamma_{\rm KN}) = Q^{(n)}_{\gamma\gamma}\frac{3}{50}\frac{10L}{\tau_{\gamma\gamma}c}\begin{cases}\gamma_{\rm KN}/\gamma^2 & \gamma \geq \tilde{\gamma}_{\rm cool,1}\\ 0 & \gamma < \tilde{\gamma}_{\rm cool,1}\end{cases}, \tag{B.11}$$

[1] Strictly speaking, the first generation of pairs $Q^{(1)}_{\gamma\gamma}(\gamma)$ may have particles at energies between $\tilde{\gamma}_{\rm cool,2}$ and $\gamma^{(1)}_2 = \gamma_2/4$ if $\gamma_2 > 4\tilde{\gamma}_{\rm cool,2}$. In that case, taking $Q^{(n)}_{\gamma\gamma} = \int Q^{(n)}_{\gamma\gamma}(\gamma)\,{\rm d}\gamma$, and not cutting off the integral at $\tilde{\gamma}_{\rm cool,2}$, overestimates the number of particles in the low-energy Thomson regime, artifically reducing the estimate of $\bar{\gamma}^{(n)}_N$ and, hence, of $f^{(n)}_{\rm nocool}$.

which will be a slightly more useful form later. For reference, the average energy of this distribution is

$$\frac{\int_1^{\gamma_{\rm min}} \gamma N_{\gamma\gamma}^{(n)}(\gamma)\,{\rm d}\gamma}{\int_1^{\gamma_{\rm min}} N_{\gamma\gamma}^{(n)}(\gamma)\,{\rm d}\gamma} = \tilde{\gamma}_{\rm cool,1} \ln\left(\frac{2\gamma_{\rm KN}}{\tilde{\gamma}_{\rm cool,1}}\right) \simeq 3.5\tilde{\gamma}_{\rm cool,1}\,, \tag{B.12}$$

where the integrals are taken through $\gamma_{\rm min} = 2\gamma_{\rm KN}$ because, as detailed below, this is where I match the Thomson solution (B.11) to the one I will obtain below in the deep Klein-Nishina regime.

Next, I determine the high-energy ($\gamma > \gamma_{\rm KN}$) part of the steady-state distributions $N_{\gamma\gamma}^{(n)}(\gamma)$. I note that $Q_{\gamma\gamma}^{(n)}$ can be written [cf. equation (7.38)] as

$$Q_{\gamma\gamma}^{(n)} \simeq B_n \frac{\gamma_{\rm min}^{-\Gamma+1}}{\Gamma-1}\left[1-\left(\frac{\gamma_{\rm min}}{\gamma_2^{(n)}}\right)^{\Gamma-1}\right] \simeq B_n \frac{\gamma_{\rm min}^{-\Gamma+1}}{\Gamma-1}\,. \tag{B.13}$$

Plugging this into the right-hand-side of (B.8) gives

$$\begin{aligned} N_{\gamma\gamma}^{(n)}(\gamma \gtrsim \gamma_{\rm KN}) &= \frac{1}{8}\left[t_{\rm cool,IC}(\gamma)\right]\left[Q_{\gamma\gamma}^{(n+1)}(\gamma/4)\right] \\ &= \frac{1}{8}\left[\frac{3}{50}\frac{10L}{\tau_{\gamma\gamma}c}\frac{1}{qf_{\rm KN}(q)}\right]\left[B_{n+1}\left(\frac{\gamma}{4}\right)^{-\Gamma}\right] \\ &= \frac{1}{8}\left[\frac{3}{50}\frac{10L}{\tau_{\gamma\gamma}c}\frac{1}{qf_{\rm KN}(q)}\right]\left[\mathcal{A}(\Gamma)B_n 4^{\Gamma}\gamma^{-\Gamma}\right] \\ &\simeq \frac{1}{8}\left[\frac{3}{50}\frac{10L}{\tau_{\gamma\gamma}c}\frac{1}{qf_{\rm KN}(q)}\right]\left[\mathcal{A}(\Gamma)Q_{\gamma\gamma}^{(n)}4^{\Gamma}(\Gamma-1)\gamma_{\rm min}^{-1}\left(\frac{\gamma}{\gamma_{\rm min}}\right)^{-\Gamma}\right] \\ &= \frac{1}{2\gamma_{\rm KN}}Q_{\gamma\gamma}^{(n)}\left[\frac{3}{50}\frac{10L}{\tau_{\gamma\gamma}c}\frac{1}{qf_{\rm KN}(q)}\right]\left(\frac{\gamma}{\gamma_{\rm min}}\right)^{-\Gamma}\,. \end{aligned} \tag{B.14}$$

The third line follows from (7.36), the fourth line from (B.13), and the last line from (7.37) along with $\gamma_{\rm min} = 2\gamma_{\rm KN}$. Rigorously matching the low-energy Thomson solution (B.11) to the high-energy Klein-Nishina result (B.14) requires a detailed analysis of the radiative physics near $\gamma \sim \gamma_{\rm KN}$. I expect this, at most, to modify (B.11) and (B.14) by order-unity factors near $\gamma \sim \gamma_{\rm KN}$, since the assumptions from which these solutions are derived are fairly robust in their respective limits (far from $\gamma_{\rm KN}$). Moreover, the solutions are already of similar scale at a natural matching point, $\gamma = \gamma_{\rm min} = 2\gamma_{\rm KN}$, with equation (B.14) a factor of about $1/f_{\rm KN}(2) \sim 5$ larger than (B.11) when both

are evaluated at $\gamma_{\rm min}$. This is independent of Γ. Thus, as a rough estimate, I take the overall distribution of nth-generation pairs (including both $\gamma > \gamma_{\rm KN}$ and $\gamma < \gamma_{\rm KN}$) to be

$$N_{\gamma\gamma}^{(n)}(\gamma) \simeq Q_{\gamma\gamma}^{(n)} \frac{3}{50} \frac{10L}{\tau_{\gamma\gamma} c} \begin{cases} \gamma_{\rm min}/2\gamma^2 & \tilde{\gamma}_{\rm cool,1} \leq \gamma < \gamma_{\rm min} \\ [\gamma_{\rm min} q f_{\rm KN}(q)]^{-1} (\gamma/\gamma_{\rm min})^{-\Gamma} & \gamma_{\rm min} \leq \gamma < \gamma_2^{(n)} \\ 0 & \text{otherwise} \end{cases} . \tag{B.15}$$

If one leverages the fact that $q > 2$ when $\gamma > \gamma_{\rm min}$ to approximate $q f_{\rm KN}(q) \simeq q^{-0.5}$, equation (B.15) becomes

$$N_{\gamma\gamma}^{(n)}(\gamma) \simeq Q_{\gamma\gamma}^{(n)} \frac{3}{50} \frac{10L}{\tau_{\gamma\gamma} c} \begin{cases} \gamma_{\rm min}/2\gamma^2 & \tilde{\gamma}_{\rm cool,1} \leq \gamma < \gamma_{\rm min} \\ (\sqrt{2}/\gamma_{\rm min})(\gamma/\gamma_{\rm min})^{-\Gamma+0.5} & \gamma_{\rm min} \leq \gamma < \gamma_2^{(n)} \\ 0 & \text{otherwise} \end{cases} . \tag{B.16}$$

B.2 Evaluating $f_{\rm nocool}$

Armed with the pair distributions $N_{\gamma\gamma}^{(n)}(\gamma)$, and knowing, from section 7.2.2, the injected distributions $Q_{\gamma\gamma}^{(n)}(\gamma)$, I can now explicitly evaluate equation (B.5) to obtain $f_{\rm nocool}$. Assembling all of the ingredients – equations (7.21), (B.1)-(B.4), (B.13), and (B.16) – one has

$$\frac{\bar{\gamma}_Q^{(n)}}{\gamma_{\rm pp}} = \frac{1}{4} \frac{\Gamma - 1}{\Gamma - 2} \left[\frac{1 - \left(z/4^{n+1/2} \right)^{2-\Gamma}}{1 - \left(z/4^{n+1/2} \right)^{1-\Gamma}} \right] , \tag{B.17}$$

$$N_{\gamma\gamma}^{(n)} = Q_{\gamma\gamma}^{(n)} \frac{10L}{\tau_{\gamma\gamma} c} \left\{ 1 + \sqrt{2} \frac{3/50}{\Gamma - 1.5} \left[1 - \left(\frac{z}{4^{n+1/2}} \right)^{1.5-\Gamma} \right] \right\} , \tag{B.18}$$

and

$$\frac{\bar{\gamma}_N^{(n)}}{\gamma_{\rm pp}} = \frac{3}{400} \left\{ 3.5 + \frac{2\sqrt{2}}{2.5 - \Gamma} \left[\left(\frac{z}{4^{n+1/2}} \right)^{2.5-\Gamma} - 1 \right] \right\} \times \left\{ 1 + \sqrt{2} \frac{3/50}{\Gamma - 1.5} \left[1 - \left(\frac{z}{4^{n+1/2}} \right)^{1.5-\Gamma} \right] \right\}^{-1} . \tag{B.19}$$

Some special numbers that appear in these expressions are: $3.5 \simeq \ln(2 \times 3/50) = \ln(\gamma_{\rm min}/\tilde{\gamma}_{\rm cool,1})$ in the numerator of (B.19) [cf. equation (B.12)]; $1/4 \equiv \gamma_{\rm min}/\gamma_{\rm pp}$, the prefactor of (B.17); and $3/400 \equiv \tilde{\gamma}_{\rm cool,1}/\gamma_{\rm pp} \equiv f_{\rm nocool,min}$, the prefactor of (B.19). In writing these expressions, I assume that the layer particle distribution cuts off at $\gamma_2 \leq 4\tilde{\gamma}_{\rm cool,2}$ and, therefore, that it is justified to replace $\gamma_2^{(n)}/\gamma_{\rm min}$ with $\gamma_2/2\gamma_{\rm KN}4^n = z/4^{n+1/2}$, as I have done.

Some important features and limits of these formulae are as follows. First, noting that $\Gamma \simeq p_{\rm KN}+1 \simeq p+0.5$, one sees that $\Gamma \in [1.5, 3.5]$. Thus, the second term in the braces of equation (B.18) is almost always small, and, generally, $N_{\gamma\gamma}^{(n)} \sim Q_{\gamma\gamma}^{(n)} 10L/\tau_{\gamma\gamma c}$: the number of particles in the steady state is dominated by those in the Thomson regime. Next, when Γ is on the softer side ($\Gamma \gtrsim 2.5$), all expressions become virtually z-independent (unless $z \sim 4^{n+1/2}$). This is because the distributions of pairs are so steep that all quantities are dominated by the low energies [near $\gamma_{\rm min}$ for the $Q_{\gamma\gamma}^{(n)}(\gamma)$'s and near $\tilde{\gamma}_{\rm cool,1}$ for the $N_{\gamma\gamma}^{(n)}(\gamma)$'s]. In this limit, the weights $\bar{\gamma}_Q^{(n)} N_{\gamma\gamma}^{(n)} / \sum_k \bar{\gamma}_Q^{(k)} N_{\gamma\gamma}^{(k)}$ used to average the $f_{\rm nocool}^{(n)}$'s in (B.5) exhibit a very simple n-dependence: $\bar{\gamma}_Q^{(n)} N_{\gamma\gamma}^{(n)} \propto Q_{\gamma\gamma}^{(n)} \propto \mathcal{A}(\Gamma)^n \ll 1$. Thus, $f_{\rm nocool}$ is dominated by the first generation, tending to

$$\lim_{\Gamma\to 3.5} f_{\rm nocool} = \lim_{\Gamma\to 3.5} f_{\rm nocool}^{(1)} = \lim_{\Gamma\to 3.5} \frac{\bar{\gamma}_N^{(n)}}{\bar{\gamma}_Q^{(n)}}$$
$$= \frac{3}{400} \frac{3.5 + 2\sqrt{2}}{(3.5-1)/4(3.5-2)} \simeq 0.1\,. \qquad \text{(B.20)}$$

This is much higher than the pessimistic (small $f_{\rm nocool}$) estimate $f_{\rm nocool,min} = 3/400$ from equation (B.7), and owes to the extended nature of the power-law distribution of particles. Though dominated by the low energies, this distribution still carries an average particle energy that is a factor of several higher than $\tilde{\gamma}_{\rm cool,1}$.

Finally, I note that $\bar{\gamma}_N^{(n)}$ becomes highly z-dependent when $\Gamma < 2.5$, and so does $\bar{\gamma}_Q^{(n)}$ when $\Gamma < 2$. As the result of this, the highest values of z that I study – all the way up to $z = 4\tilde{\gamma}_{\rm cool,2}/\gamma_{\rm KN} \simeq 1100$ – produce $f_{\rm nocool}$ of order unity (between 0.4 and 0.5). Thus, in my quantitative framework, $f_{\rm nocool}$ is rather weakly dependent on the physical parameters z and p, only varying between about 0.1 and 0.5 over the broad parameter space, $(p, z) \in [1, 3] \times [8, 1100]$. However, one should bear in mind that I have made many simplifying assumptions in this model,

and so a wider range of $f_{\rm nocool}$ may be possible in reality. Thus I would like to stress the overall qualitative insight – that $f_{\rm nocool}$ is never really too small (always at least several per cent) and may potentially reach order unity – more than the exact quantitative values. So, even though I now repeat my analysis from section 7.2.1 using the functional form for $f_{\rm nocool}$ (via those for $\bar{\gamma}_N^{(n)}$, $\bar{\gamma}_Q^{(n)}$ and $N_{\gamma\gamma}^{(n)}$) obtained in this section, I aim to stress general features, showing how the (now somewhat more self-consistent) dynamics predicted by the pair-regulated Klein-Nishina reconnection model depend in a very qualitative sense on z and p.

B.3 Reconnection dynamics with self-consistent $f_{\rm nocool}$

Putting $f_{\rm noesc} = 1$ as in section 7.2.1 (see also Appendix C), the energy recapture efficiency is just $\xi = f_{\rm noesc} f_{\rm nocool} = f_{\rm nocool}$. I denote the (new in this Appendix) self-consistent functional dependence of ξ, through $f_{\rm nocool}$, on $p_{\rm KN}$ and z by writing

$$\xi = \Xi(p_{\rm KN}, z) = f_{\rm nocool}\,, \tag{B.21}$$

where $f_{\rm nocool}$ is evaluating according to equation (B.5) using equations (B.13) and (B.17)-(B.19). Note that Ξ is a function of p only through $p_{\rm KN}$. This is because the power-law scaling of the upstream particle distributions Γ is inherited strictly from the deep Klein-Nishina-regime layer particles: $\Gamma \simeq p_{\rm KN} + 1$. Equation (B.21) eliminates ξ as a free parameter from the problem – just like $\mathcal{F}$, it is entirely determined in terms of $p_{\rm KN}$ (and, hence, in terms of $\sigma_{\rm h}$) and $z = \gamma_2/\gamma_{\rm KN}$. This allows me to repeat the analysis from section 7.2.1, except using the expression for ξ in (B.21) instead of scanning across it as a free parameter. This I do in Figs. B.1-B.3.

Each of these figures presents results from two cases, one where radiation back reaction on the layer distribution of emitting particles is ignored [i.e., where $p_{\rm T} = p_{\rm KN} = p(\sigma_{\rm h})$] and one where it is crudely taken into account [i.e., where $p_{\rm T} = p(\sigma_{\rm h}) + 1$ and $p_{\rm KN} = p(\sigma_{\rm h}) - 0.5$]. In the former case, equation (7.15) generalizes to

$$\sigma_{{\rm h},n+1} = H(\sigma_{{\rm h},n}) \equiv \frac{\sigma_{\rm h,0}}{1 + 2\,\Xi[p(\sigma_{{\rm h},n}), z]\mathcal{F}[p(\sigma_{{\rm h},n}), z]\sigma_{\rm h,0}/3}\,, \tag{B.22}$$

whereas, in the latter, equation (7.19) generalizes to

$$\sigma_{\mathrm{h},n+1} = \tilde{H}(\sigma_{\mathrm{h},n}) \equiv \frac{\sigma_{\mathrm{h},0}}{1 + 2\,\tilde{\Xi}[p(\sigma_{\mathrm{h},n}), z]\tilde{\mathcal{F}}[p(\sigma_{\mathrm{h},n}), z]\sigma_{\mathrm{h},0}/3}\,. \tag{B.23}$$

Here, I have defined

$$\tilde{\Xi}(p, z) \equiv \Xi(p - 0.5, z)\,. \tag{B.24}$$

I note that, when evaluating Ξ and $\tilde{\Xi}$ through equation (B.5), I replace the symbol $Q_{\gamma\gamma}^{(n)}$ in the expression for $N_{\gamma\gamma}^{(n)}$ [equation (B.18)] with $B_n\gamma_{\mathrm{min}}^{-\Gamma+1}[1 - (4^{n+1/2}/z)^{\Gamma-1}]/(\Gamma - 1)$ rather than adopt the cruder approximation $Q_{\gamma\gamma}^{(n)} \simeq B_n\gamma_{\mathrm{min}}^{-\Gamma+1}/(\Gamma-1)$ [see equation (B.13)]. This maintains continuity of f_{nocool}, ensuring that contributions from each successive generation turn on gradually with z (as they do in reality) rather than discretely. I have checked that this does not introduce significant error into the calculation, even though it is slightly inconsistent with the derivations of equations (B.17)-(B.19) [particularly one of the steps in equation (B.14)], which assume $Q_{\gamma\gamma}^{(n)} \simeq B_n\gamma_{\mathrm{min}}^{-\Gamma+1}/(\Gamma - 1)$.

In Figs. B.1-B.3, I analyze more thoroughly the dependence of the solution σ_{h} on z than in section 7.2.1. There are two reasons for this. First, eliminating ξ as a free parameter renders z the only independent variable in the problem, and so examining z-dependence is now easier. Second, unlike when I treated ξ as a free parameter, Ξ and $\tilde{\Xi}$ both depend on z, so examining z-dependence is now more necessary. I conduct an analysis in the broad range of $z = \gamma_2/\gamma_{\mathrm{KN}}$ spanning from the minimum for pair production to occur in the upstream region, $z = 8$, to the maximum such that all newborn particles are injected with Lorentz factors less than $\tilde{\gamma}_{\mathrm{cool},2}$. In the latter case, z is determined by setting the cutoff in the first generation's injected distribution $\gamma_2^{(1)}$ to $\tilde{\gamma}_{\mathrm{cool},2}$, yielding $z = \gamma_2/\gamma_{\mathrm{KN}} = (4\gamma_2^{(1)})/\gamma_{\mathrm{KN}} = 4\tilde{\gamma}_{\mathrm{cool},2}/\gamma_{\mathrm{KN}} \simeq 1100$ [equation (7.24)].

Fig. B.1 shows how the solutions $\sigma_{\mathrm{h}} = H(\sigma_{\mathrm{h}})$ and $\sigma_{\mathrm{h}} = \tilde{H}(\sigma_{\mathrm{h}})$, where ξ is determined self-consistently, differ from the corresponding solutions $\sigma_{\mathrm{h}} = h(\sigma_{\mathrm{h}})$ and $\sigma_{\mathrm{h}} = \tilde{h}(\sigma_{\mathrm{h}})$, when ξ is taken to be constant. This is done for the extreme case $z = 1100$, which is most suitable for comparing with section 7.2.1, where z is infinite. Additionally, Fig. B.1 also displays the functions Ξ and $\tilde{\Xi}$. One can see that, at small σ_{h}, Ξ and $\tilde{\Xi}$ are small [~ 0.1, as predicted by equation (B.20)]. Furthermore, Ξ

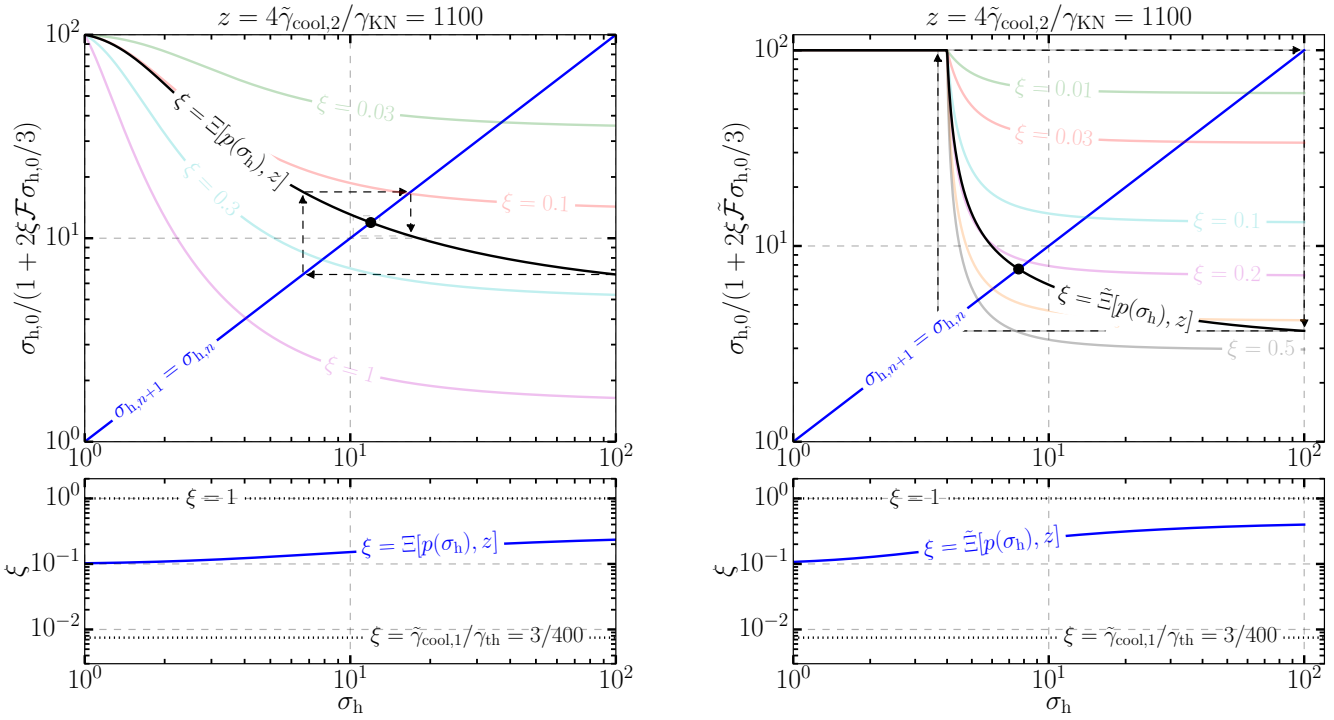

Figure B.1: Top subpanels: On the left (right), the function $H(\sigma_{\rm h})$ $[\tilde{H}(\sigma_{\rm h})]$, which includes a calculation of ξ including (excluding) the effect of radiation-reaction on the distribution of radiating particles in the reconnection layer, for the case $z = \gamma_2/\gamma_{\rm KN} = 1100$. This function is overlaid on several constant-ξ contours of $h(\sigma_{\rm h})$ $[\tilde{h}(\sigma_{\rm h})]$, which are taken from Fig. 7.5 (Fig. 7.8, left panel). When a contour is crossed, it means that the value of $\xi = \Xi$ $(\xi = \tilde{\Xi})$ equals the corresponding contour value. Bottom subpanels: The function Ξ $(\tilde{\Xi})$ plotted explicitly, alongside the minimum and maximum allowed values, $3/400 \sim 0.01$ and 1, respectively. The energy recapture efficiency ξ is always much higher than the minimum, regardless of the assumptions made about radiation reaction, and, if radiation reaction is accounted for, may even approach order unity.

and $\tilde{\Xi}$ both increase with $\sigma_{\rm h}$. This results from the power-law index $p(\sigma_{\rm h})$ of radiating layer particles becoming harder, which causes $f_{\rm nocool}$ to be dominated by the high-energy particles. These particles have their radiative cooling somewhat suppressed by Klein-Nishina effects and, for high enough z, may retain a large portion of the initially injected energy, substantially increasing $f_{\rm nocool}$. Thus, as $\sigma_{\rm h}$ grows, one observes the solution $H(\sigma_{\rm h})$ $[\tilde{H}(\sigma_{\rm h})]$ cross progressively larger constant-ξ contours of $h(\sigma_{\rm h})$ $[\tilde{h}(\sigma_{\rm h})]$. In fact, when radiation reaction on the layer particles is taken into account, $\tilde{\Xi}$ grows large enough at high z to initiate a dramatic 2-state swing cycle – a very stark difference from when ξ is pessimistically estimated as $f_{\rm nocool,min} \sim 0.01$.

In Fig. B.2, I illustrate how the picture changes at smaller z. Here, the monotonic dependence of ξ on $\sigma_{\rm h}$ from Fig. B.1 remains but is weaker. In particular, the largest value that ξ reaches, which occurs at high-$\sigma_{\rm h}$, diminishes as one reduces z. Thus, a high cutoff in the layer particle energy distribution is needed to achieve order-unity efficiency. However, it is *not* the case that a lower cutoff causes ξ to plummet. As long as $z \geq 8$, and consistent with equation (B.20), all models have appreciable efficiency, with $\xi \sim 0.1 \gg f_{\rm nocool,min}$. This implies an important conclusion: even for a radiating layer particle distribution that cuts off barely beyond $z = 8$, establishing a universal steady state is still possible for finite $\sigma_{\rm h,0}$. The initial magnetization must only be greater than $\sim 1/\min(\xi) \sim 10$.

Finally, in Fig. B.3, I present a complete stability analysis in z between 8 and 1100. This includes both how the fixed point solution $\sigma_{\rm h}(z)$ varies with z, and whether a two-state swing cycle appears, with respective high and low magnetizations $\sigma_{\rm h,>}(z)$ and $\sigma_{\rm h,<}(z)$. The model that neglects radiation reaction on the layer particles never develops oscillatory behavior, but always converges toward the fixed points $\sigma_{\rm h}(z)$. (However, this could change if z is made even larger. In that case, layer particles with energies $> 4\tilde{\gamma}_{\rm cool,2}$ spawn upstream pairs that retain nearly *all* of their energy while traveling back toward the layer, increasing ξ, and likely raising $\left|H'(\sigma_{\rm h})\right|$ above 1 in Fig. B.3.) However, in the more radiatively self-consistent model, a swing cycle develops at a critical $z = z_{\rm c}^* \simeq 880$. Intriguingly, this is before the fixed point $\sigma_{\rm h}(z)$ becomes unstable, which does not happen for any $z < 1100$. Instead, an *unstable* 2-cycle appears and intercepts the flow

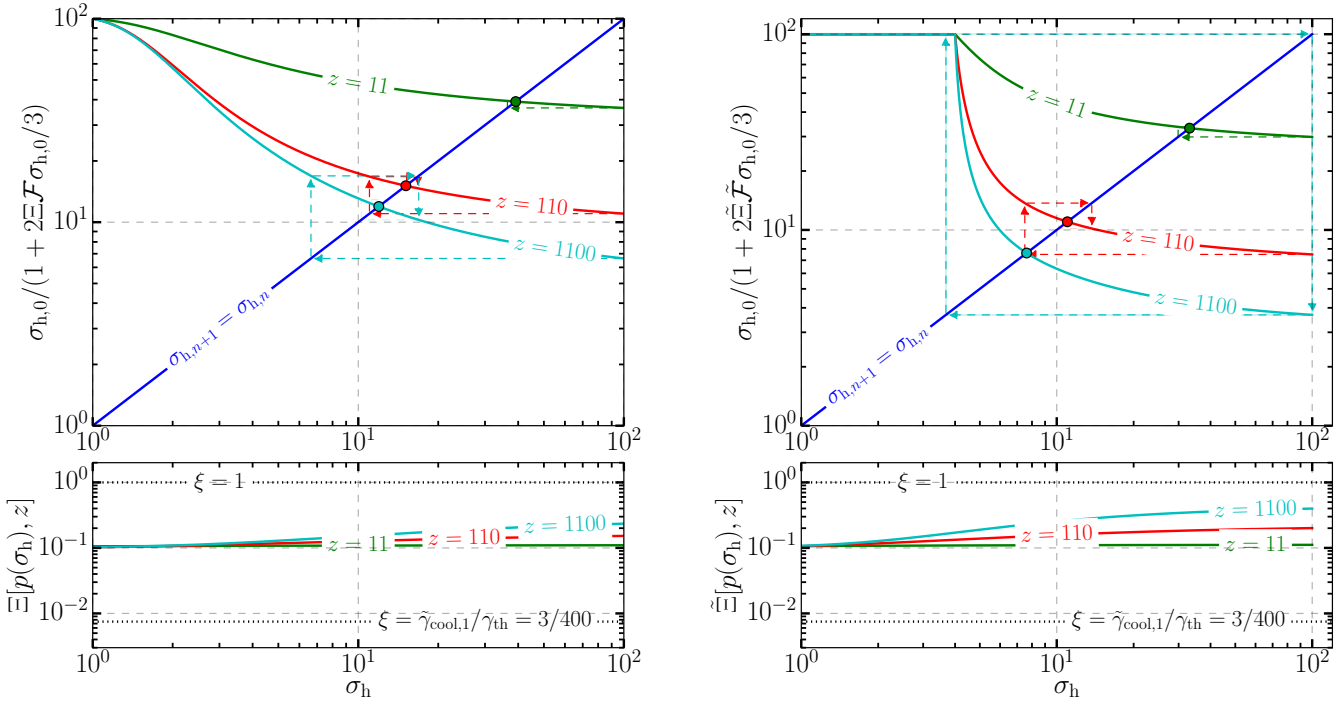

Figure B.2: The same as Fig. B.1, but without constant-ξ contours of the functions h and $\tilde{h}$. Instead, on the left (right), the functions H ($\tilde{H}$) and Ξ ($\tilde{\Xi}$) are plotted for several different values of z. On both sides, the maximum ξ-value occurs at high-$\sigma_{\rm h}$ and is sensitive to z. Both radiation reaction and high-z are required for ξ to reach order unity. In contrast, the low value of ξ, occuring at low-$\sigma_{\rm h}$, is virtually independent of z and is always around $\xi = 0.1 \gg f_{\rm nocool,min}$. Thus, order-unity efficiency requires both high z and high $\sigma_{\rm h}$, but extremely low efficiency ($\xi \sim 10^{-2}$) is averted across the domain of all models.

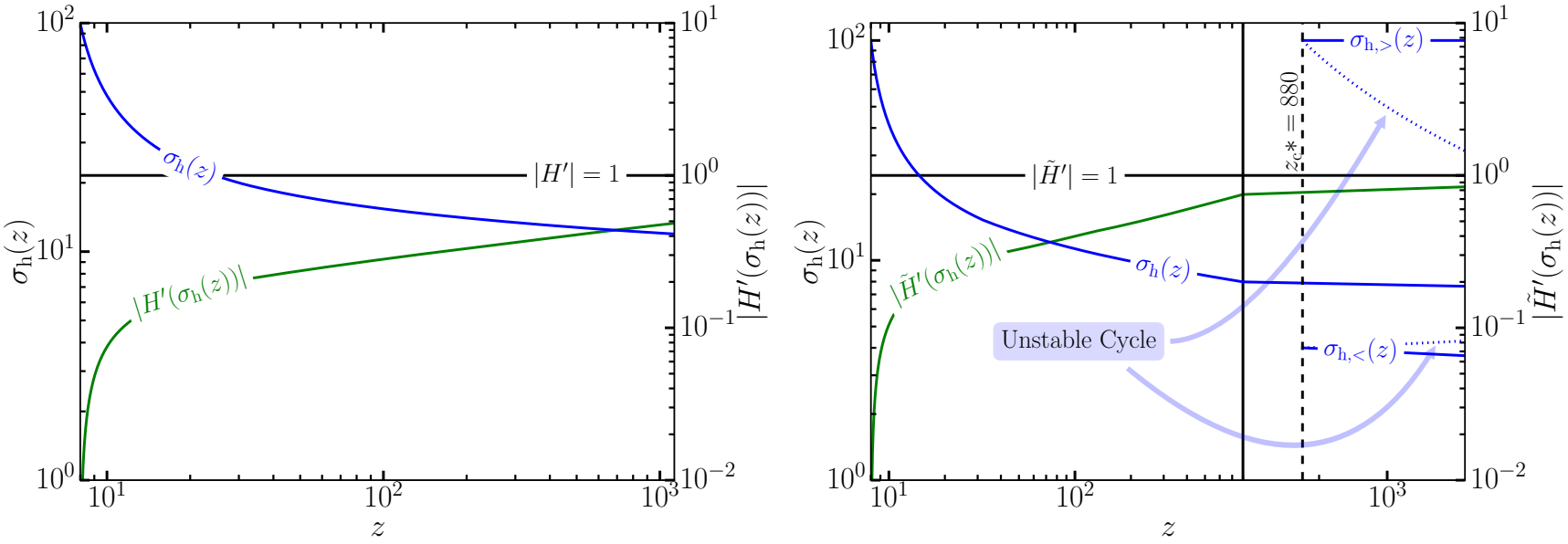

Figure B.3: On the left (right), a stability analysis of the function $H(\sigma_{\rm h})$ $[\tilde{H}(\sigma_{\rm h})]$ for all z between 8 and 1100. For visual clarity on the right plot, the horizontal axis scale beyond $z = 800$ is highly zoomed (and thus the kinks in curves crossing $z = 800$ are not real). The model without radiation back reaction on the layer particles does not possess high enough efficiency for its fixed point $\sigma_{\rm h}(z)$ to go unstable, or to develop a two-state swing cycle. In contrast, the model with radiation reaction displays limit-cycle behavior. There, pair feedback is so efficient that the system immediately starts, from the first readvection time, in a swing cycle whenever $z > z_{\rm c}* = 880$. This is true even though the fixed point $\sigma_{\rm h}(z)$ does not become locally unstable for any $z < 1100$. Instead, an *unstable* cycle appears and intercepts the flow on the iterated map, blocking all states that start from $\sigma_{\rm h,0}$ from ever reaching the fixed point. Thus, the model on the right illustrates the argument in Appendix D that a locally unstable fixed point is a sufficient but not necessary condition for the system to globally asymptote to a limit cycle.

of the interated map $\sigma_{\rm h,n+1} = \tilde{H}(\sigma_{\rm h,n})$, blocking states that start at $\sigma_{\rm h,0}$ from reaching the fixed point.

While swing cycles appear to require quite extended particle distributions in this model ($z_{\rm c}* = 880$ is close to the upper limit $z = 1100$ that I study), I do not think that this necessarily precludes their operation in all but the most extreme systems. Instead, I have found that the particular value of z where swing-cycles set in is highly sensitive to order-unity changes in ξ, and, thus, this value could be either much higher (in which case swing cycles would be unlikely in reality) or much lower (in which case they may be quite common).

B.4 Summary of self-consistent $f_{\rm nocool}$

This discussion shows how, by self-consistently determining $\xi = f_{\rm nocool}$ (through the functions Ξ and $\tilde{\Xi}$), one may achieve significantly higher efficiencies than a simple and pessismistic estimate, $\xi = f_{\rm nocool,min} = 3/400$, would predict. Rather than radiating away most of the energy it receives from the reconnection layer, the newly created upstream plasma may catch and hold onto this energy, delivering an appreciable portion of it back to the layer. The reason is not that the fresh pairs radiate inefficiently. On the contrary, given a full readvection time, all of them would cool to energies $\sim \tilde{\gamma}_{\rm cool,1}$. Rather, the broad distribution of newborn pairs is constantly being replenished, due to injection from photon annihilation, at high energies, and this allows the typical energy of radiating particles, as they enter the layer, to be quite high.

Thus, at some finite large (but not excessively large) initial magnetization $\sigma_{\rm h,0} \leq 1/\min(\xi) \sim 10$, the layer *always* has a universal fixed point solution. Furthermore – and depending on how significantly radiation back reaction modifies the distribution of layer particles – energy recapture by the layer may reach order-unity effiency when a long high-energy tail of layer particles extends beyond $\gamma_{\rm KN}$. In this case, the reconnection layer may overshoot its fixed point steady state, and undergo late-time limit-cycle oscillations about this solution rather than converge toward it.

Appendix C

The Particle Escape Factor $f_{\rm noesc}$

In this Appendix, I discuss a complementary channel, besides radiative cooling, through which newly created upstream matter may lose energy: particle escape. Escape occurs when newborn pairs stream a distance $\sim L$ along the unreconnected magnetic field, vacating the system. Pairs can be born with essentially any pitch angle, so the typical escape time is $\sim L/c$ (except for a few particles with very small pitch angles). Therefore, the escape factor $f_{\rm noesc}$ is close to unity if the readvection time, $t_{\rm ra}$, satisfies $t_{\rm ra} \simeq 10L/\tau_{\gamma\gamma}c < L/c$, which requires $\tau_{\gamma\gamma} > 10$.

Beyond this simple consideration, there are also other, more complicated kinetic effects that may influence the value of $f_{\rm noesc}$. Many of these effects, at the same time, also help to decide whether above-threshold photons radiated from the layer ever reach the upstream region in the first place, as assumed throughout this work [e.g., assumption 3b of section 7.2]. These are, in principle, independent concepts: $f_{\rm noesc}$ pertains to pairs that have already been born into the upstream region; photon escape from the layer to the upstream region concerns the emission and propagation of radiation before pairs are ever produced. However, the same processes dictate both, and I discuss these parallel influences simultaneously.

Consider first the shape of the unreconnected magnetic field lines at a transverse distance $|y| < 0.1L$ from the layer. In this region, the unreconnected field is perturbed by the presence of large plasmoids, the largest of which may extend a distance $\sim 0.1L$ into the upstream region (Uzdensky et al., 2010; Sironi et al., 2016). Because the field is asymptotically uniform as $|y| \to \infty$, it is necessarily stronger (i.e., compressed), when $|y| \lesssim 0.1L$, in regions above plasmoids. This may lead

to a magnetic bottling effect, where particles produced at $|y| < 0.1L$ tend to be mirror-confined between large plasmoids (even when they might otherwise escape), pushing $f_{\rm noesc}$ closer to unity. Unlike the effects discussed below, this confinement mechanism only pertains to particles already born and, therefore, does not affect what fraction of photons escape the layer to the upstream region.

I move now to a separate issue: anisotropy in the distribution of radiating layer particles. The layer tends to drive bulk plasma motion and, through the kinetic beaming mechanism (e.g., Cerutti et al., 2012b, 2013; Mehlhaff et al., 2020), collimated bunches of high-energy particles, into the $\pm x$-directions (either to the left or to the right in Figs. 7.1 and 7.2). This biases the emitted photons (which are relativistically beamed along the directions of their emitters) into these same directions, and may, in turn, increase $f_{\rm noesc}$. Suppose, for example, that a given above-threshold photon travels at an angle θ from the reconnection midplane. This reduces the total distance it propagates into the upstream region from $|y| \sim \lambda_{\rm mfp}$ to $|y| \sim \lambda_{\rm mfp} \sin\theta$. The readvection time for the produced pair is therefore reduced by the same factor: $\sin\theta$. Meanwhile, the escape time is still $\sim L/c$. Thus, the readvection time decreases relative to the escape time. This increases the fraction $f_{\rm noesc}$ of particles captured by the reconnection layer.

On the other hand, if the beaming of plasma motion (both on bulk and kinetic levels) is strong enough, the angle θ discussed above could be so small that most photons do not cross the separatrices into the upstream plasma before annihilating. This would occur if the transverse propagation distance $|y| \sim \lambda_{\rm mfp} \sin\theta$ is smaller than the thickness Δ of the layer radiation zone (discussed in more detail below). In such a scenario, the created pairs would not load the inflow plasma and thus the pair-regulation mechanism would be somewhat suppressed.

I next discuss how the thickness, Δ, of radiation zones, where most pair-producing photons are emitted (discussed in section 7.2.1 and Fig. 7.2), influences the escape of photons to the upstream region and $f_{\rm noesc}$. If most photons are radiated from inside large plasmoids, instead of from thin strip-like radiation zones (of the kind argued for in section 7.2.1), then the condition $\tau_{\gamma\gamma} > 10$, although it maintains $t_{\rm ra} < L/c$ and helps keep $f_{\rm noesc}$ near unity, also confines most of the above-

threshold photons to within the separatrices: inside plasmoids exceeding $\lambda_{\rm mfp}$ in size. This inhibits pair feedback. If, instead, the primary radiation sites are thin, even kinetic-scale, current sheets (e.g., with $\Delta \ll \lambda_{\rm mfp}$, as argued in section 7.2.1) most photons travel to the upstream region and effectively load its plasma when they annihilate (modulo potential extreme beaming effects described above).[1]

Finally, let me consider how reconnection current sheets between merging plasmoids play into this picture. These miniature reconnection layers are oriented perpendicularly with respect to the main reconnection current sheet and therefore bias the motion of particles and emitted photons in the transverse $\pm y$-directions. The produced upstream pairs then have suppressed pitch angles (the $\pm y$ direction is perpendicular to the unreconnected magnetic field), which inhibits particle escape. However, if $\tau_{\gamma\gamma} > 10$, then not all radiation produced at merging-plasmoid current sheets escapes back into the inflow plasma: there will at least be some photon-confining plasmoids larger than $\lambda_{\rm mfp} < L/10$. Mergers between such plasmoids would not source significant self-regulating pair production. This, however, does not preclude merging-plasmoid reconnection from regulating itself (just on a smaller scale) in the same way that the main reconnection layer does as a whole.

In light of this discussion, there are clearly several details, influencing both $f_{\rm noesc}$ and the delivery of photons to the upstream region, that are beyond the scope of the present work to calculate quantitatively. These will require a future computational study in order to properly diagnose. Nevertheless, many of these effects (especially when $\tau_{\gamma\gamma} > 10$) promote order-unity $f_{\rm noesc}$, and – modulo extreme beaming near reconnection X-points – allow photon escape to the upstream region from the primary large-scale current sheet.

[1] Note that, if $\lambda_{\rm mfp} > 0.1L$, it is irrelevant whether most of the pair-producing photons come from thin structures or from large round plasmoids. Either way, the photons escape to the upstream region.

Appendix D

Global Stability of the Iterated Map $x_{n+1} = h(x_n)$

In this Appendix, the phrase 'late-time' or 'late times' refers to the limit $\lim_{n\to\infty} t_n$ where $t_n \equiv n\lambda_{\mathrm{mfp}}/\beta_{\mathrm{rec}}c$. Late times defined in this sense are not necessarily reached by the reconnection system before reconnection terminates.

Here, I argue that the condition deciding the local stability of the iterated magnetization map $\sigma_{\mathrm{h},n+1} = h(\sigma_{\mathrm{h},n})$ from section 7.2.1 also reveals, in some cases, its global stability. In particular, I argue that local instability $\left|h'(\sigma_{\mathrm{h}})\right| > 1$ implies that a system starting at magnetization $\sigma_{\mathrm{h},0}$ asymptotically approaches a two-state swing-cycle. If, on the other hand, the fixed point is stable, $\left|h'(\sigma_{\mathrm{h}})\right| < 1$, then the system may or may not converge toward it. Thus, local instability is sufficient, but not necessary, for global instability.

I focus here only on strictly decreasing functions $h(x)$ [such that $h(y) < h(x)$ if $y > x$] that map into their own domain $[1, \sigma_{\mathrm{h},0}]$ (i.e., $h : [1, \sigma_{\mathrm{h},0}] \to [1, \sigma_{\mathrm{h},0}]$). Here, the upper end of the domain happens to coincide with the starting point of the map $\sigma_{\mathrm{h},0}$. The particular functions $h(x)$ and $H(x)$ specified in equations (7.15) and (B.22) fall into this class. [Note here that I am overloading the symbol $h(x)$, designating with it a general class of functions and not necessarily the particular form in equation (7.15).] The arguments in this section can be generalized to non-increasing functions $h^*(x)$ such that $h^*(y) \leq h^*(x)$ if $y > x$ [into which category fall the functions $\tilde{h}(x)$ and $\tilde{H}(x)$ in equations (7.19) and (B.23)] without difficulty, but require additional edge cases (saturation of the inequalities) to be considered, and so I do not formally treat them here.

I begin with a few basic observations. First, if the map does not start on the unique fixed

point $\sigma_{\rm h}$, then each successive iteration lands on the opposite side of the fixed point from the preceding iteration. That is, if $\sigma_{{\rm h},n} < \sigma_{\rm h}$, then $\sigma_{{\rm h},n+1} > \sigma_{\rm h}$, and if $\sigma_{{\rm h},n} > \sigma_{\rm h}$, then $\sigma_{{\rm h},n+1} < \sigma_{\rm h}$. This is because

$$\sigma_{{\rm h},n+1} = h(\sigma_{{\rm h},n}) < h(\sigma_{\rm h}) = \sigma_{\rm h}\,, \tag{D.1}$$

where the '$<$' follows when $\sigma_{{\rm h},n} > \sigma_{\rm h}$ by the strictly decreasing hypothesis. The same proof, but with a '$>$' sign, follows when $\sigma_{{\rm h},n} < \sigma_{\rm h}$. This precludes all non-fixed-point odd-period orbits. Hence, one cannot automatically infer the presence of chaos in the system using the 3-period theorem (Li & Yorke, 2004). As I now show, the dynamics are even more constrained: chaotic behavior is, in fact, completely precluded.

For my second observation, I note that if, for any starting index m, one discovers that $\sigma_{{\rm h},m+2} < \sigma_{{\rm h},m}$, then for all integers $n \geq 1$, it is necessarily the case that $\sigma_{{\rm h},m+2n} < \sigma_{{\rm h},m+2(n-1)}$. That is, even though each successive iteration bounces to the opposite side of the fixed point, every consecutive even (or odd) iteration moves strictly in one direction. This follows inductively because

$$\begin{aligned} \sigma_{{\rm h},m+3} = h(\sigma_{{\rm h},m+2}) > h(\sigma_{{\rm h},m}) = \sigma_{{\rm h},m+1} \\ \Rightarrow \sigma_{{\rm h},m+4} = h(\sigma_{{\rm h},m+3}) < h(\sigma_{{\rm h},m+1}) = \sigma_{{\rm h},m+2}. \end{aligned} \tag{D.2}$$

The first line follows from the strictly decreasing hypothesis and the starting assumption that $\sigma_{{\rm h},m+2} < \sigma_{{\rm h},m}$. The second line follows from the strictly decreasing hypothesis and the first line.

This is a very powerful constraint because it severely limits the potential late-time dynamics of the system. In effect, the system can only either approach the fixed point $\sigma_{\rm h} = h(\sigma_{\rm h})$ or a two-state swing-cycle. To see this, let me suppose that $\sigma_{{\rm h},2} < \sigma_{{\rm h},0}$. (Note that $\sigma_{{\rm h},2} = \sigma_{{\rm h},0}$ implies that the system begins in a two-cycle, and the condition $\sigma_{{\rm h},2} > \sigma_{{\rm h},0}$ is impossible because $h(x)$ maps onto its own domain.) Then each successive $\sigma_{{\rm h},2n}$ marches resolutely toward smaller values. This continuing reduction in $\sigma_{{\rm h},2n}$ can only be terminated in one way. It must be the case that there exists some $\sigma_{{\rm h},\infty}$ such that $\sigma_{\rm h} \leq \sigma_{{\rm h},\infty} = \lim_{n\to\infty} \sigma_{{\rm h},2n}$. If $\sigma_{{\rm h},\infty} = \sigma_{\rm h}$, then the system approaches

the fixed point $\sigma_{\rm h}$ in a late-time steady state. If $\sigma_{{\rm h},\infty} > \sigma_{\rm h}$, then the system approaches a two-cycle. Note that it is not possible for the system to approach any $2n$-cycle for $n > 1$ that is not also a two-cycle. If that were the case, then one of the $\sigma_{{\rm h},n}$'s on the cycle would be on the same side of the fixed point as another $\sigma_{{\rm h},n}$ of the same parity (even or odd), and the two would not be equal. This would contradict the result that $\sigma_{{\rm h},2n+2} < \sigma_{{\rm h},2n}$.

Thus, there are only two possible asymptotic behaviors of the map $\sigma_{{\rm h},n+1} = h(\sigma_{{\rm h},n})$: either $\sigma_{{\rm h},n}$ approaches the fixed point $\sigma_{\rm h}$ or the map converges to a two-state cycle, and one of the states has magnetization $\sigma_{{\rm h},>} > \sigma_{\rm h}$. It follows that the flow of the iterated map from $\sigma_{{\rm h},0}$ to $\sigma_{\rm h}$ is always intercepted by a two-cycle when the local instability criterion $\left|h'(\sigma_{\rm h})\right| > 1$ is met. Otherwise, there would not be a suitable attractor [i.e., one consistent with both equations (D.1) and (D.2)] to catch the strictly decreasing flow of the map $\sigma_{{\rm h},n+2} = h(h(\sigma_{{\rm h},n}))$. Let me illustrate these remarks with an example.

I display the particular map $\sigma_{{\rm h},n+2} = h(h(\sigma_{{\rm h},n}))$, where h is defined as in equation (7.15), in Fig. D.1. For low ξ (in this case, lower than $\xi_{\rm c} = 0.84$), all of the $h(h(\sigma_{{\rm h},n}))$ curves intersect the diagonal line only once and with a gentle slope $\left|{\rm d}h(h(\sigma_{{\rm h},n}))/\,{\rm d}\sigma_{{\rm h},n}\right| < 1$. These intersection points coincide with the stable fixed points $\sigma_{\rm h}$ illustrated in Fig. 7.5. The fixed point becomes unstable when the slope of the twice-iterated map becomes tangent to the diagonal at a critical value $\xi_{\rm c} \simeq 0.84$. For ξ higher than this value, the map must intercept the diagonal at at least two other locations – the values $\sigma_{{\rm h},<}$ and $\sigma_{{\rm h},>}$ corresponding to a two-cycle. This is demanded in order for $h(h(x))$ to be at or below the diagonal as $\sigma_{{\rm h},n}$ approaches $\sigma_{{\rm h},0}$ and to be at or above the diagonal as $\sigma_{{\rm h},n}$ approaches 1 (which itself is required by the fact that h maps into its own domain).

I stress that this general behavior – where a naked fixed point $\sigma_{\rm h}$ is immediately concealed behind a two-cycle as soon as it goes unstable – is demanded by equations (D.1) and (D.2). However, the inverse – that all stable fixed points occur in isolation from two-cycles – does not hold in general. The topological argument of the preceding paragraph, which was based on the fact that $h(h(\sigma_{{\rm h},0})) \leq \sigma_{{\rm h},0}$ and $h(h(1)) \geq 1$, does not preclude the possibility that a stable fixed point could spontaneously

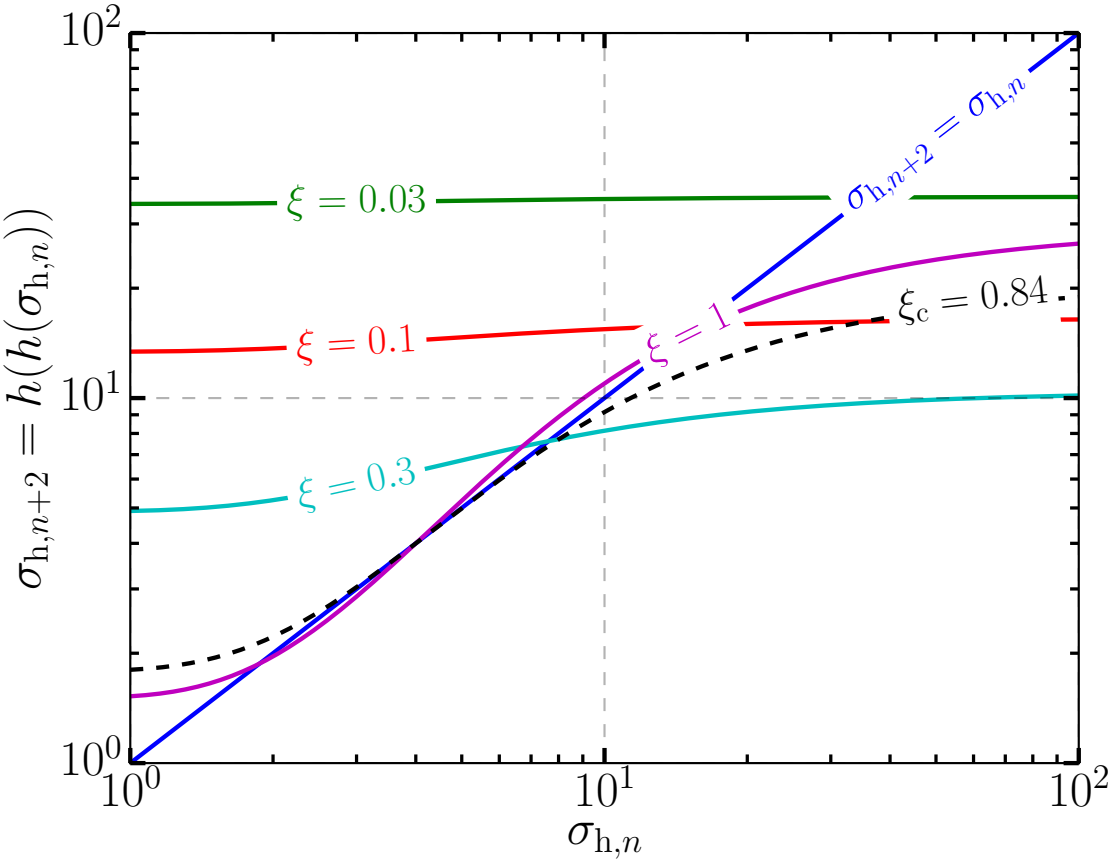

Figure D.1: The twice-iterated map $\sigma_{\mathrm{h},n+2} = h(h(\sigma_{\mathrm{h},n}))$ with $h(x)$ as defined in equation (7.15). The appearance of a stable two-cycle coincides with the critical value $\xi_{\mathrm{c}} \simeq 0.84$ where the fixed point at σ_{h} becomes unstable. This is consistent with the fact that there must be an attractor on the map $\sigma_{\mathrm{h},n+1} = h(\sigma_{\mathrm{h},n})$ to intercept the flow $\sigma_{\mathrm{h},2n+2} < \sigma_{\mathrm{h},2n} \leq \sigma_{\mathrm{h},0}$ when the local instability criterion $\left|h'(\sigma_{\mathrm{h}})\right| > 1$ is met. If locally unstable, the fixed point cannot be naked: it must be blocked from the initial state $\sigma_{\mathrm{h},0}$ by a two-state cycle.

become enshrouded by an even number of two-cycles with alternating stabilities. In this case, the function $h(h(\sigma_{\mathrm{h},n}))$ would simply intersect the diagonal an even number of times on either side of the fixed point. The first of these intersections, counted as one moves outward from the fixed point to the boundaries of the domain (1 and $\sigma_{\mathrm{h},0}$) would be necessarily unstable (since $\left|\mathrm{d}h(h(\sigma_{\mathrm{h}}))/\,\mathrm{d}\sigma_{\mathrm{h}}\right| > 1$), the next would be stable, the one after unstable, and so on until the outermost stable two-cycle. This would furthermore be consistent with equations (D.1) and (D.2), ensuring that a stable two-cycle attracts the late-time dynamics. Although such a state of affairs may seem at first rather unlikely, this is actually precisely what occurs for the map $\tilde{H}$ defined in equation (B.23). As shown in Fig. B.3, the fixed point on that map becomes enclosed inside two 2-cycles without *ever* becoming locally unstable.

All in all, if the fixed point σ_{h} is unstable ($\left|h'(\sigma_{\mathrm{h}})\right| > 1$) then the system necessarily asymptotes to a two-cycle. On the other hand, if σ_{h} is stable, then whether the system approaches it at late-times depends on whether a two-cycle is present to intercept the flow on the iterated map. If a stable two-cycle exists, σ_{h} is never reached; otherwise, the system converges to σ_{h}. Fixed points and two-cycles are the only allowed late-time behaviors for monotonically decreasing maps.

Appendix E

Exact Solution to the Thomson-Limit Energy-Advection Equation

In this section, I solve equation (B.9) exactly subject to the initial condition $N_{\gamma\gamma}^{(n)}(\gamma, 0) = 0$. I evolve the solution for one readvection time $t_{\rm ra} = \lambda_{\rm mfp}/\beta_{\rm rec}c$, which simulates the continual injection of particles as a parcel of plasma moves from $|y| \sim \lambda_{\rm mfp}$ to the layer. The solution follows identically for all generations, so I simplify my notation to

$$\frac{\partial}{\partial t}N(\gamma, t) + \frac{\partial}{\partial \gamma}\left[\dot{\gamma}_{\rm T}N(\gamma, t)\right] = Q\delta(\gamma - \gamma_{\rm KN})\,, \tag{E.1}$$

and write $-\dot{\gamma}_{\rm T} = \gamma/t_{\rm cool,T}(\gamma) = c\gamma^2/L\gamma_{\rm cool} \equiv \gamma^2/\tau$, abbreviating $\tau \equiv L\gamma_{\rm cool}/c$. (Note that τ is a time and not an optical depth.)

I denote the Laplace transform in time with a tilde:

$$\tilde{N}(\gamma, s) \equiv \int_0^\infty N(\gamma, t)e^{-st}\,{\rm d}t\,. \tag{E.2}$$

Laplace-transforming equation (E.1) gives

$$s\tilde{N} - \frac{1}{\tau}\frac{\partial}{\partial\gamma}\left(\gamma^2\tilde{N}\right) = \frac{Q}{s}\delta(\gamma - \gamma_{\rm KN})\,, \tag{E.3}$$

where I used the initial condition $N(\gamma, 0) = 0$. Expanding the γ-derivative, rearranging, and multiplying through by the integrating factor

$$\mu(\gamma) = \gamma^2 e^{-s\tau(1-1/\gamma)} \tag{E.4}$$

gives

$$\frac{\partial}{\partial\gamma}\left[\mu(\gamma)\tilde{N}(\gamma, s)\right] = -\frac{Q\tau}{s\gamma^2}\mu(\gamma)\delta(\gamma - \gamma_{\rm KN})\,. \tag{E.5}$$

Next, I integrate (E.5) from γ to some arbitrary $\gamma_{\rm hi} > \gamma_{\rm KN}, \gamma$. Using $\tilde{N}(\gamma_{\rm hi}, s) = 0$ (radiative cooling only populates energies lower than $\gamma_{\rm KN}$) yields

$$\begin{aligned}\tilde{N}(\gamma, s) &= \frac{Q\tau}{s\gamma_{\rm KN}^2}\frac{\mu(\gamma_{\rm KN})}{\mu(\gamma)}\Theta(\gamma_{\rm KN} - \gamma) \\ &= \frac{Q\tau}{s\gamma_{\rm KN}^2}e^{-s\tau(1/\gamma - 1/\gamma_{\rm KN})}\Theta(\gamma_{\rm KN} - \gamma)\,, \end{aligned} \tag{E.6}$$

where $\Theta(x)$ is the Heaviside step function. Equation (E.6) is the Laplace transform of the sought solution:

$$N(\gamma, t) = \frac{Q\tau}{\gamma^2}\Theta\left[t - \tau\left(\frac{1}{\gamma} - \frac{1}{\gamma_{\rm KN}}\right)\right]\Theta(\gamma_{\rm KN} - \gamma)\,. \tag{E.7}$$

Thus, constant injection at $\gamma = \gamma_{\rm KN}$ develops into a power-law γ^{-2} extending into lower and lower energies with time. Putting $t = \lambda_{\rm mfp}/\beta_{\rm rec}c \simeq 10L/\tau_{\gamma\gamma}c$ allows the power-law to extend from $\gamma_{\rm KN}$ all the way down to

$$\left(\frac{1}{\gamma_{\rm KN}} + \frac{1}{\tilde{\gamma}_{\rm cool,1}}\right)^{-1} \simeq \tilde{\gamma}_{\rm cool,1} \equiv \frac{3}{50}\gamma_{\rm KN}\,. \tag{E.8}$$

Hence [cf. equation (B.10)],

$$N(\gamma, t_{\rm ra}) = Q\gamma_{\rm cool}\frac{L}{c}\begin{cases}\gamma^{-2} & \tilde{\gamma}_{\rm cool,1} \le \gamma < \gamma_{\rm KN} \\ 0 & \text{otherwise}\end{cases}\,. \tag{E.9}$$